



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA MECÂNICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

KAMILA FERNANDA FERREIRA DA CUNHA QUEIROZ

**INTEGRAÇÃO DE FERRAMENTAS DE SIMULAÇÃO COMPUTACIONAL E DE  
CLASSIFICAÇÃO PARA ANÁLISE DE PATOLOGIAS MAMÁRIAS EM IMAGENS  
TERMOGRÁFICAS**

Recife

2022

KAMILA FERNANDA FERREIRA DA CUNHA QUEIROZ

**INTEGRAÇÃO DE FERRAMENTAS DE SIMULAÇÃO COMPUTACIONAL E DE  
CLASSIFICAÇÃO PARA ANÁLISE DE PATOLOGIAS MAMÁRIAS EM IMAGENS  
TERMOGRÁFICAS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutora em Engenharia Mecânica.

Área de concentração: Energia.

Orientadora: Profa. Dra. Rita de Cássia Fernandes de Lima.

Recife

2022

Catálogo na fonte:  
Bibliotecária Sandra Maria Neri Santiago, CRB-4 / 1267

Q3i

Queiroz, Kamila Fernanda Ferreira da Cunha.

Integração de ferramentas de simulação computacional e de classificação para análise de patologias mamárias em imagens termográficas / Kamila Fernanda Ferreira da Cunha Queiroz. – 2022.

227 f.: il., figs., quads., tabs., abev. e siglas.

Orientadora: Profa. Dra. Rita de Cássia Fernandes de Lima.

Tese (Doutorado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Mecânica. Recife, 2022.

Inclui referências e apêndices.

1. Engenharia mecânica. 2. Câncer de mama. 3. Termografia. 4. Ground-truth. 5. Sobreamostragem. 6. Algoritmos evolutivos. 7. Diagnóstico auxiliado por computador. I. Lima, Rita de Cássia Fernandes de (Orientadora). II. Título.

UFPE

621 CDD (22. ed.)

BCTG/2022-374

KAMILA FERNANDA FERREIRA DA CUNHA QUEIROZ

**INTEGRAÇÃO DE FERRAMENTAS DE SIMULAÇÃO COMPUTACIONAL E DE  
CLASSIFICAÇÃO PARA ANÁLISE DE PATOLOGIAS MAMÁRIAS EM IMAGENS  
TERMOGRÁFICAS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para a obtenção do título de doutora em Engenharia Mecânica. Área de concentração: Energia.

Aprovada em: 26/08/2022.

**BANCA EXAMINADORA**

---

Profa. Dra. Rita de Cássia Fernandes de Lima (Orientadora)  
Universidade Federal de Pernambuco

---

Prof. Dr. Paulo Roberto Maciel Lyra (Examinador Interno)  
Universidade Federal de Pernambuco

---

Prof. Dr. Wellington Pinheiro dos Santos (Examinador Externo)  
Universidade Federal de Pernambuco

---

Profa. Dra. Ana Lúcia Bezerra Candeias (Examinadora Externa)  
Universidade Federal de Pernambuco

---

Profa. Dra. Aura Conci (Examinadora Externa)  
Universidade Federal Fluminense

## AGRADECIMENTOS

Agradeço a Deus pelas oportunidades que Ele tem me proporcionado, por me iluminar em todos os momentos e por Ele ter colocado em minha vida pessoas que são sustento, inspiração e motivo de sorrisos. À minha Fatinha (Nossa Senhora de Fátima) pela intercessão de sempre e por estar sempre comigo, confirmando diariamente que minha profissão é minha verdadeira vocação.

Agradeço ao meu pequeno Miguel por ter chegado em nossas vidas no momento certo, sendo o combustível que me faltava para conseguir finalizar esta etapa. Ao meu marido, Rodrigo, obrigada por me apoiar, pelo incentivo e pela preocupação de quem sempre acompanhou meu doutorado. Obrigada, também, pela paciência e pelo amor demonstrados nos meus momentos não tão bons.

Aos meus pais que educaram a mim e ao meu irmão com amor e com respeito às nossas decisões. Obrigada por nunca medirem esforços para nos proporcionar uma formação de qualidade e por nos transmitir bons valores, como a empatia, o comprometimento e a humildade.

Agradeço a Reginaldo, meu irmão favorito, pelas sugestões, pelo incentivo e por toda ajuda durante o doutorado. Me faltam palavras para agradecer tudo que fez por mim nesse longo período, sei o quanto torceu e esperou por esse momento. Eis que no final, o título é meu, mas o Dr. vai no seu nome.

Aos meus avós pelos ensinamentos, preocupação e por toda ajuda que me deram. A Vovô Dé, que certamente teria ficado feliz e emocionado por esse momento. Agradeço por todo amor, toda alegria e todos os acordos que compartilhamos.

Agradeço à Professora Rita, que me acompanha desde a iniciação científica (2011), pela orientação, pelas sugestões, pelo incentivo, disponibilidade e amizade. Aos professores Paulo Lyra e Wellington Pinheiro pelas contribuições no seminário e na organização do trabalho. Ao médico Francisco George, pela participação e apoio fundamentais para o andamento de toda a pesquisa.

Aos amigos do LABTERMO pelos sorrisos e cafés compartilhados, pelo incentivo e por se fazerem sempre presentes. Agradeço às minhas queridas amigas do IFRN, Leila e Camille. Obrigada pela acolhida no Rio Grande do Norte, pelo apoio incondicional em todos os momentos e pelo companheirismo sem limites.

Ao Programa de Pós-Graduação em Engenharia Mecânica pela oportunidade e por fornecer os recursos necessários para o desenvolvimento desta tese.

## RESUMO

A inspeção termográfica é considerada um método potencial para melhorar a eficiência da detecção precoce de doenças mamárias. Nesse contexto, o objetivo deste estudo foi o desenvolvimento da ferramenta AGATA (versões Triagem e Multiclassificação) destinada a detectar automaticamente alterações termográficas que indicam risco de malignidade mamária. Para isso, inicialmente, foi validado o método de segmentação automática comparando com cada imagem de referência correspondente, obtida por segmentação manual. Nesta etapa, foram utilizados 90 termogramas (180 segmentações) e a técnica foi avaliada por medidas de sobreposição (Índice de Dados, Índice de Jaccard, Taxa de Falso Positivo e Taxa de Falso Negativo) e por análise estatística para determinar concordância e correlação (gráfico Bland-Altman e correlação de Pearson). Em seguida, o método de segmentação automática foi aplicado a um banco de dados com 233 termogramas mamários. Cinco classificadores (k-Vizinhos mais próximos, Máquinas de Vetores Suporte - SVM, Árvore de Decisão, Análise Discriminante e Naive Bayes) foram avaliados em combinação com as técnicas de sobreamostragem mais comuns na literatura. No problema multiclasse, os elementos da base de dados foram categorizados (*safe*, *borderline*, *rare* e *outlier*) e foram aplicadas estratégias distintas de sobreamostragem, além das tradicionais. Abordagens de seleção de atributos foram consideradas usando métodos evolucionários (Algoritmo Genético e Evolução Diferencial), mRMR (Redundância mínima e Relevância Máxima) e ganho de informação (WEKA – *Waikato Environment for Knowledge Analysis*). Acurácia, sensibilidade, especificidade, área sob a curva ROC (*Receiver Operating Characteristic*) (AUC) e estatística Kappa foram usados para avaliar o desempenho dos classificadores. De modo complementar aos casos de lesão maligna, foi desenvolvida uma estratégia para estimar a profundidade e o tamanho do tumor através do perfil de distribuição de temperatura das superfícies de modelos simulados e através do uso do SVM para problemas de regressão. Como resultado da validação da segmentação automática, o gráfico de Bland-Altman apresentou um viés médio de  $2458,46 \pm 2641,89$  pixels entre as áreas obtidas pela segmentação automática e o *ground-truth*, com coeficiente de correlação de Pearson de 0,68. O índice de similaridade Dice mostrou concordância substancial de aproximadamente 69%. A metodologia adotada para validação sugere que as técnicas analisadas de segmentação das regiões mamárias foram adequadas e consistentes e podem ser utilizadas no processo CAD (*Computer-Aided Diagnosis*). No sistema AGATA Triagem foi implementada o classificador SVM combinado com a seleção de atributos por algoritmo genético e a técnica de sobreamostragem ASUWO (*Adaptive Semi-Unsupervised Weighted Oversampling*), atingindo 95,23% de acurácia, 93,65% de

sensibilidade e 96,81% de especificidade. Para o AGATA Multiclasses foi implementado o classificador SVM e a técnica de balanceamento por sobreamostragem de elementos *safes*, atingindo 92,20% de acurácia, 92,32% de sensibilidade e 97,40% de especificidade. Quanto a profundidade e o tamanho das lesões malignas, foram obtidos, respectivamente, sistemas com  $R^2$  (Coeficiente de Determinação) médio de 0,53 e REQM de 2,95 e  $R^2$  médio de 1,16 e REQM (Raiz do Erro Quadrático Médio) de 2,26. O desenvolvimento de sistemas de alto desempenho utilizando uma nova modalidade de imagem mamária pode auxiliar positivamente no rastreamento do câncer de mama.

Palavras-chave: câncer de mama; termografia; *ground-truth*; sobreamostragem; algoritmos evolutivos; diagnóstico auxiliado por computador.

## ABSTRACT

Thermographic inspection is considered a potential method to improve the efficiency of early detection of breast diseases. In this context, the aim of this study was to develop the AGATA tool (Screening and Multiclassification versions) to automatically detect thermographic changes that indicate a risk of breast malignancy. For this, initially, the automatic segmentation method was validated by comparing it with its corresponding reference image, obtained by manual segmentation. In this step, 90 thermograms (180 segmentations) were used and the technique was evaluated by measures of overlap (Data Index, Jaccard Index, False Positive Rate and False Negative Rate) and by statistical analysis to determine agreement and correlation (Bland-Altman graph and Pearson correlation). Then, the automatic segmentation method was applied to a database with 233 breast thermograms. Five classifiers (k-Nearest Neighbors, Support Vector Machines - SVM, Decision Tree, Discriminant Analysis and Naive Bayes) were evaluated in combination with the most common oversampling techniques in the literature. In the multiclass problem, the elements of the database were categorized (safe, borderline, rare and outlier) and different oversampling strategies were applied, in addition to the traditional ones. Attribute selection approaches were considered using evolutionary methods (Genetic Algorithm and Differential Evolution), mRMR (Minimum Redundancy and Maximum Relevance) and information gain (WEKA – Waikato Environment for Knowledge Analysis). Accuracy, sensitivity, specificity, area under the ROC curve (Receiver Operating Characteristic) (AUC) and Kappa statistics were used to evaluate the performance of the classifiers. In addition to cases of malignant lesion, a strategy was developed to estimate the depth and size of the tumor through the surface temperature distribution profile of simulated models and using SVM for regression problems. As a result of the automatic segmentation validation, the Bland-Altman plot presented an average bias of  $2458.46 \pm 2641.89$  pixels between the areas obtained by the automatic segmentation and the ground-truth, with a Pearson correlation coefficient of 0.68. The Dice similarity index showed substantial agreement of approximately 69%. The methodology adopted for validation suggests that the analyzed techniques of segmentation of the mammary regions were adequate and consistent and can be used in the CAD (Computer-Aided Diagnosis) process. In the AGATA Triagem system, the SVM classifier was implemented, combined with the selection of attributes by genetic algorithm and the oversampling technique ASUWO (Adaptive Semi-Unsupervised Weighted Oversampling), reaching 95.23% of accuracy, 93.65% of sensitivity and 96, 81% specificity. For AGATA Multiclass, the SVM classifier and the balancing technique by oversampling of safe elements were implemented,

reaching 92.20% of accuracy, 92.32% of sensitivity and 97.40% of specificity. Regarding the depth and size of malignant lesions, systems with mean  $R^2$  (Coefficient of determination) of 0.53 and 2.95 RMSE (Root Mean Squared Error) and mean  $R^2$  of 1.16 and RMSE of 2.26 were obtained, respectively. The development of high-performance systems using a new breast imaging modality can positively assist in breast cancer screening.

Keywords: breast cancer; thermography; ground-truth; oversampling; evolutionary algorithms; computer-assisted diagnosis.

## LISTA DE FIGURAS

Figura 1 –	Simplificação do fluxo assistencial à paciente com patologia mamária na rede de saúde, com proposta geral de sistema de apoio ao diagnóstico .....	28
Figura 2 –	Tela principal (Menu Arquivo) .....	43
Figura 3 –	Tela da interface com resultados .....	44
Figura 4 –	Tela da inicial do sistema PCA Diagnostic .....	45
Figura 5 –	GUI do software ThermoLabAnimal .....	46
Figura 6 –	Visão da GUI desenvolvida por Brzezinski et al. (2020) .....	46
Figura 7 –	Página principal da GUI construída para análise de diabetes tipo II em termogramas faciais .....	47
Figura 8 –	GUI do sistema proposto .....	48
Figura 9 –	Detalhes da interface gráfica criada .....	49
Figura 10 –	Interface gráfica AVMTM (Menu Arquivo) .....	49
Figura 11 –	Tela inicial da interface gráfica desenvolvida por Bayas (2015) .....	50
Figura 12 –	Interface para auxiliar o diagnóstico precoce do câncer de mama desenvolvida por Granadillo et al (2019) .....	51
Figura 13 –	GUI para classificação de imagens termográficas de mamas criada por Ekici e Jawzal (2019) .....	51
Figura 14 –	Anatomia feminina da mama com regiões de maior incidência de câncer destacadas .....	52
Figura 15 –	Gráficos de dispersão (a) Amostra desbalanceada e (b) Amostra balanceada por SMOTE .....	60
Figura 16 –	Gráficos de dispersão (a) Amostra desbalanceada e (b) Amostra balanceada por Borderline-SMOTE .....	61
Figura 17 –	Gráficos de dispersão (a) Amostra desbalanceada e (b) Amostra balanceada por ADASYN .....	62
Figura 18 –	Etapas do algoritmo CBO (a) Amostra desbalanceada, (b) Vetor de distância entre exemplos e centros dos clusters, (c) Clusters atualizados e (d) Amostra balanceada por CBO .....	64
Figura 19 –	Casos correspondentes a taxas de nível de segurança .....	65

Figura 20 –	Abordagem de sobreamostragem entre instâncias selecionadas e seus 4 vizinhos mais próximos, considerando que pertençam ao mesmo cluster .....	66
Figura 21 –	Criação de instâncias sintéticas, geradas mais próximas das instâncias majoritárias .....	67
Figura 22 –	Exemplo ilustrativo do algoritmo k-NN, para $k = 3$ .....	71
Figura 23 –	Representação de uma árvore como um tipo de grafo .....	73
Figura 24 –	Exemplo LDA para três classes .....	77
Figura 25 –	Ilustração das principais características de um espaço ROC .....	81
Figura 26 –	Fluxograma do método Algoritmo Genético .....	84
Figura 27 –	Operadores genéticos .....	85
Figura 28 –	Fluxograma do método Evolução Diferencial .....	86
Figura 29 –	Operador crossover para Evolução Diferencial .....	87
Figura 30 –	Fluxograma da metodologia adotada para validação da segmentação automática .....	90
Figura 31 –	Resultado da segmentação automática .....	91
Figura 32 –	Exemplo de resultados da segmentação manual .....	92
Figura 33 –	Fluxograma da metodologia adotada para desenvolvimento do sistema de classificação binária .....	93
Figura 34 –	Exemplos das posições frontais e laterais realizadas na aquisição das imagens termográficas.....	94
Figura 35 –	Fluxograma da metodologia aplicada à seleção de características, onde a cor verde representa a seleção por otimização da classificação e a cor cinza representa a redução do subconjunto selecionado a um par de atributos .....	98
Figura 36 –	Representação da associação entre o cromossomo binária e os atributos selecionados .....	99
Figura 37 –	Fluxograma da metodologia adotada para desenvolvimento do sistema CAD de multiclassificação, com base de dados balanceada .....	106
Figura 38 –	Fluxograma da metodologia adotada para desenvolvimento do sistema CAD de multiclassificação, com base de dados desbalanceada e normalizada .....	107
Figura 39 –	Representação dos tipos de exemplos em uma amostra desbalanceada .....	108

Figura 40 –	Fluxograma da metodologia adotada para predição do tamanho e da profundidade da lesão maligna .....	111
Figura 41 –	Representação das mamas do grupo de pesquisa. (a) Semiesfera (b) Fantoma (c) Prótese de silicone (d) Prótese de silicone ajustada à paciente (e) Geometria com estruturas internas (f) Geometria de mama constituída por tecidos glandular e adiposo criada a partir de termograma .....	112
Figura 42 –	Ajuste da base torácica com a medida da semirreta MH .....	113
Figura 43 –	Representação da extração manual da curva de perfil de um termograma...	113
Figura 44 –	Visão transversal de modelos de mama com diferentes composições mamárias. (a) ExD (b) HD (c) PF (d) PG .....	115
Figura 45 –	Representação das condições de contorno adotadas para o problema em estudo .....	116
Figura 46 –	(a) Gráfico de dispersão para as variáveis área S e área GT. (b) Gráfico de Bland-Altman para média e diferenças entre as áreas segmentadas. LSC: Limite Superior de Concordância; LIC: Limite Inferior de Concordância...	123
Figura 47 –	Resultados da segmentação para diferentes tipos de mama (a) Mama pequena (b) Mama média e (c) Mama grande .....	123
Figura 48 –	Métricas para avaliar a segmentação automática e sua variação em relação à verdade do terreno (a) Índice de Jaccard. (b) Dados de Índice. (c) Falso positivo. (d) Falso negativo .....	124
Figura 49 –	Resultados da (a) AUC e do (b) Índice Kappa utilizando sobreamostragem por SMOTE, para base de dados binária .....	130
Figura 50 –	Resultados da (a) AUC e do (b) Índice Kappa utilizando sobreamostragem por SMOTE Borderline, para base de dados binária .....	130
Figura 51 –	Resultados da (a) AUC e do (b) Índice Kappa utilizando sobreamostragem por ASUWO, para base de dados binária .....	130
Figura 52 –	Interface gráfica AGATA Triagem .....	133
Figura 53 –	Variação dos índices (a) Kappa e (b) AUC de multiclassificação em função das técnicas de seleção de atributos .....	136
Figura 54 –	Distribuição das múltiplas classes do conjunto usado .....	137
Figura 55 –	Resultados da (a) ACC, da (b) AUC do (c) Índice Kappa utilizando os métodos de sobreamostragem, para base de dados multiclases .....	138

Figura 56 –	Distribuição categórica dos exemplos em função do tamanho da vizinhança .....	139
Figura 57 –	Resultados da (a) ACC, da (b) AUC do (c) Índice Kappa utilizando os métodos de sobreamostragem, para base de dados multiclases .....	142
Figura 58 –	Espaços de características resultantes das técnicas de normalização de dados .....	143
Figura 59 –	Comparação entre valores AUC dos classificadores para base não normalizada e bases normalizadas .....	146
Figura 60 –	Comparação entre valores ACC dos classificadores para base não normalizada e bases normalizadas .....	146
Figura 61 –	Interface gráfica AGATA Versão Multiclases .....	159
Figura 62 –	(a) Sub-interface Estatística (b) Gráfico Casos clínicos x Idade gerado com dados hipotéticos .....	160

## LISTA DE QUADROS

Quadro 1 –	Classificação BI-RADS.....	25
Quadro 2 –	Resumo dos estudos citados de sistemas CAD destinados à análise de imagens termográficas de mamas .....	40
Quadro 3 –	Equações das métricas de análise de segmentação .....	56
Quadro 4 –	Funções kernel mais comuns .....	76
Quadro 5 –	Matriz de confusão para um problema binário .....	79
Quadro 6 –	Demais medidas baseadas na matriz de confusão .....	79
Quadro 7 –	Aspectos importantes sobre os índices de sensibilidade e especificidade ....	80
Quadro 8 –	Valores de referência para avaliação segundo a medida AUC.....	82
Quadro 9 –	Interpretação do índice de Kappa .....	83
Quadro 10 –	Lista de características extraídas dos termogramas de mamas .....	96
Quadro 11 –	Parâmetros adotados na configuração do AG .....	98
Quadro 12 –	Parâmetros adotados na configuração da ED .....	100
Quadro 13 –	Parâmetros adotados na configuração dos classificadores .....	103
Quadro 14 –	Resultado do teste sequencial, a partir dos resultados dos testes individuais X e Y .....	104
Quadro 15 –	Técnicas de normalização testadas.....	109
Quadro 16 –	Composição percentual de tecido .....	114
Quadro 17 –	Parâmetros termofísicos para todos os modelos de mamas .....	117
Quadro 18 –	Classificação da densidade mamária de pacientes de acordo com a idade ...	120
Quadro 19 –	Aplicações das plataformas computacionais utilizadas .....	121
Quadro 20 –	Síntese dos resultados de classificação binária de imagens termográficas mamárias de trabalhos encontrados na literatura .....	132
Quadro 21 –	Classificação ICT .....	134
Quadro 22 –	Síntese dos resultados de classificação multiclases de imagens termográficas mamárias de trabalhos desenvolvidos no grupo de pesquisa e encontrados na literatura .....	150
Quadro 23 –	Resultados obtidos na classificação das pacientes estudadas .....	161

## LISTA DE TABELAS

Tabela 1 –	Amostra desbalanceada utilizada na classificação binária.....	94
Tabela 2 –	Resultados da seleção de características para base de dados binária desbalanceada .....	126
Tabela 3 –	Resultados do balanceamento da base de dados binária .....	127
Tabela 4 –	Resultados da seleção de características para base de dados binária balanceada por SMOTE, SMOTE Boderline e ASUWO .....	129
Tabela 5 –	Resultados da seleção de características para base de dados multiclassés desbalanceada .....	135
Tabela 6 –	Distribuição categórica dos exemplos em função do tamanho da vizinhança .....	140
Tabela 7 –	Resultados da classificação multiclassés da amostra balanceada (aumento proporcional dentro de cada categoria) .....	140
Tabela 8 –	Teste overfitting em amostra desbalanceada e normalizada pelo método Escala (Scale) .....	148
Tabela 9 –	Estudo da diferença estatística entre as características extraídas de mamas saudáveis e de mamas com tumor maligno .....	151
Tabela 10 –	Características dos modelos preditivos para o tamanho do tumor maligno e seus valores de performance .....	153
Tabela 11 –	Resultados das estimativas de tamanho de tumores malignos usando imagens por infravermelho .....	155
Tabela 12 –	Características dos modelos preditivos para o tamanho do tumor maligno e seus valores de performance .....	157
Tabela 13 –	Resultados das estimativas da profundidade de tumores malignos usando imagens infravermelhas .....	158

## LISTA DE ABREVIATURAS E SIGLAS

ADASYN	Amostragem Sintética Adaptativa
AG	Algoritmo Genético
ASUWO	<i>Adaptive Semi-Unsupervised Weighted Oversampling</i>
AUC	<i>Area Under ROC Curve</i>
BHTE	<i>Bioheat Transfer Equation</i>
BI-RADS	<i>Breast Imaging Reporting and Data System</i>
C	Cisto
CAD	<i>Computer-Aided Diagnosis</i>
CART	<i>Classification and Regression Trees</i>
CBO	<i>Cluster Based Oversampling</i>
ED	Evolução Diferencial
ELM	<i>Extreme Learning Machine</i>
EQM	Erro Quadrático Médio
ESP	Especificidade
ExD	Extremamente Densa
FN	Falso Negativo
FP	Falso Positivo
Globocan	<i>Global Cancer Observatory</i>
GUI	<i>Graphical User Interface</i>
GUIDE	<i>Graphical User Interface Development Environment</i>
HC	Hospital das Clínicas
HD	Heterogeneamente Densa
HUAP	Hospital Universitário Antônio Pedro
IC	Intervalo de Confiança
ID	Índice Dice
IJ	Índice Jaccard
kNN	<i>k-Nearest Neighbor</i>
LB	Lesão Benigna
LDA	<i>Linear Discriminant Analysis</i>
LM	Lesão Maligna
MCS	<i>Multiple Classifier Systems</i>
MD	Mama Direita

ME	Mama Esquerda
MLP	<i>Multiplayer Perceptron</i>
mRMR	Mínima Redundância e Máxima Relevância
NB	Naive Bayes
OMS	Organização Mundial da Saúde
PF	Parcialmente Fibroglandular
PG	Predominantemente Gordurosa
REMQ	Raiz do Erro Médio Quadrático
ROC	<i>Receiver Operating Characteristic</i>
ROI	<i>Region Of Interest</i>
SENS	Sensibilidade
SC-AG	Seleção de Características por Algoritmo Genético
SC-ED	Seleção de Características por Evolução Diferencial
SC-WEKA	Seleção de Características pelo WEKA
SD	Saudável
SL	<i>Safe Level</i>
SMOTE	<i>Synthetic Minority Oversampling Technique</i>
SVM	<i>Support Vector Machine</i>
TC	Todas Características
TFN	Taxa de Falsos Negativos
TFP	Taxa de Falsos Positivos
TSL	Taxa <i>Safe Level</i>
UFF	Universidade Federal Fluminense
UFPE	Universidade Federal de Pernambuco
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
VPN	Valor Preditivo Negativo
VPP	Valor Preditivo Positivo
WEKA	<i>Waikato Environment for Knowledge Analysis</i>

## LISTA DE SÍMBOLOS

$\rho$	Massa específica (kg m <sup>-3</sup> )
$c$	Calor específico (J kg <sup>-1</sup> K <sup>-1</sup> )
$\nabla$	Operador nabla
$k$	Condutividade térmica (W m <sup>-1</sup> K <sup>-1</sup> )
$T$	Temperatura (K)
$Q_m$	Taxa volumétrica de geração de calor metabólico (W m <sup>-3</sup> )
$Q_p$	Taxa volumétrica de geração de calor devido à perfusão sanguínea (W m <sup>-3</sup> )
$Q$	Taxa volumétrica de geração de calor devido à fonte externa (W m <sup>-3</sup> )
$\omega$	Taxa de perfusão sanguínea (s <sup>-1</sup> )
$\tau$	Tempo de duplicação do tumor maligno (dias)
$D$	Diâmetro do tumor maligno (cm)
$G$	<i>Ground-truth</i>
$S$	Segmentação automática
$X_n^d$	Matriz base de dados
$x_{ij}$	Elemento da matriz
$B_{treino}$	Base de dados de treinamento
$m$	Número de atributos
$x_i$	Dado atributo
$X$	Matriz de atributos
$y_i$	Dado rótulo de classe
$Y$	Vetor de rótulos
$B_{treino,min}$	Exemplos da classe minoritária
$B_{treino,maj}$	Exemplos da classe majoritária
$k$	Número de vizinhos
$\hat{x}_i$	Vizinho selecionado aleatoriamente
$\delta$	Número aleatório entre 0 e 1
$x_{sintética}$	Instância sintética
$n_{x,sintética}$	Número de instâncias sintéticas para classe minoritária
$\beta$	Nível de equilíbrio após criação dos dados sintéticos ( $\beta \in [0,1]$ )
$\Gamma_i$	Distribuição de densidade

$\Delta_i$	Número de exemplos vizinhos de $x_i$ que pertencem a $B_{treino,maj}$
$Z$	Constante de normalização
$g_i$	Número de instâncias sintéticas para cada elemento de $B_{treino,min}$
$v_{novo,i}$	Instância normalizada
mín	Valor mínimo
max	Valor máximo
$\mu$	Média
$\sigma$	Desvio padrão
$n$	Tamanho da amostra
$x_t$	Objeto teste
$s$	Valor que define variações da métrica de Minkowski ( $1 \leq s < \infty$ )
$K$	Kernel
$\varphi$	Função mapeamento
$p$	Número de preditores
$\Sigma$	Matriz de covariância
$\kappa$	Índice <i>Kappa</i>
$P_o$	Acurácia observada
$P_e$	Acurácia esperada
$v_{novo}$	Novo indivíduo gerado pela mutação na Evolução Diferencial
$v_{base}$	Indivíduo base (Evolução Diferencial)
$F$	Fator de perturbação
$u_i^j$	Vetor experimental
$v^j$	Vetor alvo
$r_i$	Número randômico entre 0 e 1
CR	Taxa de cruzamento
$R^2$	Coefficiente de determinação

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>23</b>
1.1	OBJETIVOS .....	28
<b>1.1.1</b>	<b>Objetivo geral .....</b>	<b>28</b>
<b>1.1.2</b>	<b>Objetivos específicos .....</b>	<b>29</b>
1.2	ORGANIZAÇÃO DO TRABALHO .....	29
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA .....</b>	<b>30</b>
2.1	SISTEMAS DE DIAGNÓSTICO APLICADOS ÀS IMAGENS TERMOGRÁFICAS DE MAMA .....	30
2.2	INTERFACES GRÁFICAS RELACIONADAS À TERMOGRAFIA .....	43
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>52</b>
3.1	O CÂNCER DE MAMA .....	52
3.2	TRANSFERÊNCIA DE CALOR NO CORPO HUMANO .....	53
3.3	MÉTRICAS DE ANÁLISE DE PERFORMANCE DA SEGMENTAÇÃO DE IMAGENS DIGITAIS .....	55
3.4	EXTRAÇÃO DE ATRIBUTOS .....	57
3.5	PRÉ-PROCESSAMENTO DE DADOS .....	58
<b>3.5.1</b>	<b>Dados desbalanceados .....</b>	<b>58</b>
<b>3.5.2</b>	<b>Transformação de atributos numéricos .....</b>	<b>67</b>
<b>3.5.3</b>	<b>Redução de dados .....</b>	<b>68</b>
3.6	MODELOS PREDITIVOS .....	69
<b>3.6.1</b>	<b>Algoritmo do k vizinho mais próximo .....</b>	<b>70</b>
<b>3.6.2</b>	<b>Naive Bayes .....</b>	<b>71</b>
<b>3.6.3</b>	<b>Árvore de decisão .....</b>	<b>73</b>
<b>3.6.4</b>	<b>Máquinas de vetores de suporte .....</b>	<b>75</b>
<b>3.6.5</b>	<b>Análise discriminante linear .....</b>	<b>76</b>
3.7	AVALIAÇÃO DE MODELOS PREDITIVOS .....	78
<b>3.7.1</b>	<b>Matriz de confusão .....</b>	<b>78</b>
<b>3.7.2</b>	<b>Análise ROC .....</b>	<b>81</b>
<b>3.7.3</b>	<b>Teste de concordância Kappa .....</b>	<b>82</b>
3.8	ALGORITMOS DE OTIMIZAÇÃO GLOBAL .....	83
<b>4</b>	<b>METODOLOGIA .....</b>	<b>89</b>

4.1	O PROBLEMA DA VALIDAÇÃO DA SEGMENTAÇÃO AUTOMÁTICA ...	89
4.2	ELABORAÇÃO DO SISTEMA DE CLASSIFICAÇÃO BINÁRIA .....	93
<b>4.2.1</b>	<b>Base de dados .....</b>	<b>93</b>
<b>4.2.2</b>	<b>Extração de características .....</b>	<b>95</b>
4.3	SOBREAMOSTRAGEM POR GERAÇÃO DE DADOS SINTÉTICOS .....	96
4.4	SELEÇÃO DE CARACTERÍSTICAS .....	97
<b>4.4.1</b>	<b>Seleção de características baseada no Algoritmo Genético .....</b>	<b>98</b>
<b>4.4.2</b>	<b>Seleção de características baseada na Evolução Diferencial.....</b>	<b>100</b>
<b>4.4.3</b>	<b>Seleção de um par de características – mRMR e WEKA .....</b>	<b>101</b>
4.5	CLASSIFICAÇÃO DA AMOSTRA .....	102
4.6	CLASSIFICAÇÃO INDIVIDUAL NÃO SUPERVISIONADA .....	103
4.7	ELABORAÇÃO DO SISTEMA DE CLASSIFICAÇÃO MULTICLASSES....	105
<b>4.7.1</b>	<b>Base de dados .....</b>	<b>105</b>
<b>4.7.2</b>	<b>Categorização e balanceamento da amostra .....</b>	<b>107</b>
<b>4.7.3</b>	<b>Normalização .....</b>	<b>109</b>
4.8	ESTRATÉGIA PARA ESTIMATIVA DA LOCALIZAÇÃO DO TUMOR MALIGNO .....	110
<b>4.8.1</b>	<b>Modelagem geométrica da mama .....</b>	<b>111</b>
<b>4.8.2</b>	<b>Modelagem matemática utilizada .....</b>	<b>115</b>
<b>4.8.3</b>	<b>Propriedades termofísicas e simulação numérica .....</b>	<b>116</b>
<b>4.8.4</b>	<b>Modelo de aprendizado de máquina .....</b>	<b>118</b>
<b>4.8.5</b>	<b>AGATA Multiclasses – Módulo localização .....</b>	<b>119</b>
4.9	INFRAESTRUTURA COMPUTACIONAL .....	120
<b>5</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>122</b>
5.1	VALIDAÇÃO DA SEGMENTAÇÃO AUTOMÁTICA .....	122
5.2	ANÁLISE DA CLASSIFICAÇÃO BINÁRIA .....	125
<b>5.2.1</b>	<b>AGATA Triagem - proposta de um sistema CAD de classificação binária...</b>	<b>132</b>
5.3	ANÁLISE DA CLASSIFICAÇÃO MULTICLASSES .....	135
<b>5.3.1</b>	<b>Base de dados desbalanceada .....</b>	<b>135</b>
<b>5.3.2</b>	<b>Base de dados balanceada .....</b>	<b>137</b>
<b>5.3.3</b>	<b>Base de dados normalizada .....</b>	<b>142</b>
5.4	ANÁLISE DA PREVISÃO DO TAMANHO E PROFUNDIDADE DO TUMOR MALIGNO .....	150

5.4.1	<b>Avaliação das características .....</b>	<b>150</b>
5.4.2	<b>Previsão do tamanho do tumor maligno .....</b>	<b>152</b>
5.4.3	<b>Previsão da profundidade do tumor maligno .....</b>	<b>155</b>
5.5	<b>AGATA MULTICLASSES - PROPOSTA DE UM SISTEMA CAD DE MULTICLASSIFICAÇÃO .....</b>	<b>158</b>
5.6	<b>ESTUDO DE CASOS .....</b>	<b>160</b>
5.7	<b>SÍNTESE DAS CONTRIBUIÇÕES DO TRABALHO .....</b>	<b>161</b>
6	<b>CONCLUSÕES .....</b>	<b>163</b>
	<b>REFERÊNCIAS .....</b>	<b>166</b>
	<b>APÊNDICE A – TERMOGRAFIA APLICADA ÀS CIÊNCIAS DA SAÚDE .....</b>	<b>182</b>
	<b>APÊNDICE B - INTERFACES GRÁFICAS NÃO-RELACIONADAS À TERMOGRAFIA .....</b>	<b>193</b>
	<b>APÊNDICE C - APLICAÇÕES DE MÉTODOS DE OTIMIZAÇÃO .....</b>	<b>200</b>
	<b>APÊNDICE D - ÍNDICES DE AVALIAÇÃO TERMOGRÁFICA DE ANOMALIAS MAMÁRIAS .....</b>	<b>203</b>
	<b>APÊNDICE E – DESCRIÇÃO DOS ATRIBUTOS EXTRAÍDOS DOS TERMOGRAMAS DE MAMA .....</b>	<b>207</b>
	<b>APÊNDICE F – GRÁFICOS BOXPLOT DOS ATRIBUTOS EXTRAÍDOS DA BASE DE DADOS BINÁRIA .....</b>	<b>212</b>
	<b>APÊNDICE G – RESULTADOS DA CLASSIFICAÇÃO BINÁRIA PARA BASE DE DADOS DESBALANCEADA .....</b>	<b>216</b>
	<b>APÊNDICE H - RESULTADO DA CLASSIFICAÇÃO BINÁRIA PARA OS DIFERENTES MÉTODOS DE BALANCEAMENTO DA AMOSTRA, UTILIZANDO TODAS AS CARACTERÍSTICAS .....</b>	<b>217</b>
	<b>APÊNDICE I - RESULTADOS DA CLASSIFICAÇÃO BINÁRIA PARA AMOSTRA BALANCEADA POR SMOTE, COM CONJUNTO SELECIONADO DE ATRIBUTOS .....</b>	<b>218</b>
	<b>APÊNDICE J - RESULTADOS DA CLASSIFICAÇÃO BINÁRIA PARA AMOSTRA BALANCEADA POR BORDERLINE-SMOTE, COM CONJUNTO SELECIONADO DE ATRIBUTO .....</b>	<b>219</b>

<b>APÊNDICE K – RESULTADOS DA CLASSIFICAÇÃO BINÁRIA PARA AMOSTRA BALANCEADA POR ASUWO, COM CONJUNTO SELECIONADO DE ATRIBUTOS .....</b>	<b>220</b>
<b>APÊNDICE L – PERFORMANCE DOS CLASSIFICADORES EM BASE MULTICLASSES DESBALANCEADA .....</b>	<b>221</b>
<b>APÊNDICE M – PERFORMANCE DOS CLASSIFICADORES EM BASE MULTICLASSES BALANCEADA .....</b>	<b>222</b>
<b>APÊNDICE N – PERFORMANCE DOS CLASSIFICADORES EM BASE MULTICLASSES BALANCEADA PELA TÉCNICA SAES .....</b>	<b>223</b>
<b>APÊNDICE O – PERFORMANCE DOS CLASSIFICADORES EM BASE MULTICLASSES DESBALANCEADA E NORMALIZADA .....</b>	<b>224</b>
<b>APÊNDICE P – HISTOGRAMAS DAS CARACTERÍSTICAS EXTRAÍDAS DA BASE MULTICLASSES .....</b>	<b>225</b>

## 1 INTRODUÇÃO

O câncer de mama é hoje um relevante problema de saúde pública. É a neoplasia maligna de mais incidência em mulheres na maior parte do mundo. Segundo as últimas estatísticas mundiais do Globocan 2020 (*Global Cancer Observatory*) foram estimados, aproximadamente, 2,3 milhões de novos casos de câncer e 685 mil óbitos pela doença (SUNG et al, 2021).

No Brasil, as estimativas de incidência de câncer de mama para cada ano do triênio 2020 – 2022 são de 66.280 casos novos. Desconsiderando os tumores de pele não melanoma, o câncer de mama feminino ocupa a primeira posição mais frequente em todas as regiões do país (INCA, 2019). O prognóstico da doença pode ser pouco promissor pela possibilidade de desenvolvimento silencioso, o que agrava ainda mais a situação, pois retarda o diagnóstico e dificulta a efetividade do tratamento (BARBOSA, 2012).

Para o câncer de mama, a idade é um dos principais fatores de risco. O processo de envelhecimento vem acompanhado de alterações biológicas e do acúmulo de exposições à radiação ao longo da vida, fatos que aumentam a probabilidade de se desenvolver o referido câncer. As mulheres com idade superior a 50 anos são mais predispostas a desenvolver a doença, aumentando a taxa de mortalidade a partir dessa idade (INCA, 2019).

Os altos índices de casos e de óbitos por câncer de mama na população feminina, somados aos elevados custos da assistência levaram à necessidade de se traçar estratégias que favoreçam o seu controle no futuro. Segundo a Organização Mundial da Saúde (OMS), para o aumento da sobrevivência associado a esse tipo de câncer, é necessário que haja adoção de políticas voltadas para a detecção precoce e para o tratamento adequado em tempo hábil (BRITO et al., 2005). As estratégias para a prévia detecção são o rastreamento, que é a realização de exames de triagem em indivíduos assintomáticos para detectar o câncer ou lesões sugestivas de câncer, e o diagnóstico precoce.

O diagnóstico precoce necessita da conscientização da população e de profissionais de saúde para os sinais e sintomas antecipados do câncer, promovendo a realização de atendimento médico regular para os pacientes avaliados (SILVA & HORTALE, 2012).

Algumas técnicas de imagens têm sido comumente usadas na investigação de patologias mamárias, como o câncer de mama. Entre elas destacam-se: mamografia, ressonância magnética, termografia e ultrassonografia. Esses métodos têm suas próprias abordagens e ferramentas distintas; os resultados esperados desses métodos dependem de diferentes tipos de fatores, e é recomendado o uso de mais de um método para validar os resultados (EKICI & JAWZAL, 2019).

A mamografia é o exame de imagem mais largamente utilizado para o rastreio e a detecção do câncer de mama, mas, como já bem descrito na literatura, apresenta uma série de limitações. Ela pode ser ineficiente para mulheres com mamas fibrocísticas, densas ou cirurgicamente implantadas, e por possuir baixa especificidade, requer a confirmação por biópsias. Além disso, há o perigo do uso da radiação ionizante e o desconforto gerado pela compressão das mamas entre as superfícies planas do mamógrafo para melhorar a qualidade da imagem (NG; KEE; ACHARYA, 2005; RAMÍREZ-TORRES et al., 2017).

Uma segunda técnica com grande utilização na detecção de tumores mamários é a ultrassonografia, que se baseia na emissão de pulsos de onda sonora. Normalmente, a ultrassonografia é usada na investigação de áreas suspeitas da mama encontradas através da mamografia ou do exame clínico. É uma técnica que costuma ser muito eficaz na distinção entre tumores líquido e sólidos. Ainda é utilizada na triagem ou como forma suplementar à mamografia, possibilitando a investigação de mamas densas (KANDLIKAR et al., 2017). Quando utilizada complementarmente à mamografia, é capaz de aumentar a sensibilidade do rastreamento (BERG et al., 2008; KANDLIKAR et al., 2017).

Um estudo efetuado no Japão avaliou a sensibilidade e a especificidade da mamografia associada à ultrassonografia das mamas como método de rastreamento do câncer de mama (OHUCHI et al., 2016). O estudo foi realizado com 72.998 mulheres. As voluntárias foram distribuídas aleatoriamente entre dois grupos: grupo intervenção (submetido aos exames de mamografia e ultrassonografia mamárias) e grupo controle (submetido ao exame de mamografia). A sensibilidade e a especificidade para o grupo intervenção foram de, respectivamente, 91,1% e 87,7%. Enquanto para o grupo controle, a sensibilidade foi de 77,0% e a especificidade de 91,4%. Dessa forma, o estudo concluiu que a ultrassonografia adjuvante é capaz de aumentar as taxas de detecção do câncer de mama.

Os laudos mamográfico e ultrassonográfico adotam as orientações do Colégio Americano de Radiologia que classifica os resultados do exame conforme o sistema *Breast Imaging Reporting and Data System* (BI-RADS®), sintetizados no Quadro 1.

Quadro 1– Classificação BI-RADS

Classificação	Avaliação	Descrição	Recomendação
BI-RADS 0	Incompleto	Inconclusiva	Realizar novos exames
BI-RADS 1	Negativo	Nada a comentar	Rastreamento de rotina
BI-RADS 2	Benigno	Achado benigno	Rastreamento de rotina
BI-RADS 3	Provavelmente benigno	Achado com alta probabilidade de benignidade (>95%)	Acompanhamento em 6 meses
BI-RADS 4	Anormalidade suspeita	Razoável probabilidade de malignidade	Biópsia deve ser considerada
BI-RADS 5	Altamente suspeito de malignidade	Achado com alta probabilidade de malignidade (>95%)	Realizar ações apropriadas
BI-RADS 6	Malignidade comprovada por biópsia	Achado maligno sendo visualizado para tratamento	Tratamento

Fonte: Spak et al. (2017)

Nos últimos anos, a termografia vem ganhando interesse de pesquisadores, especialmente para fins de rastreamento do câncer de mama (KANDLIKAR et al., 2017). Baseia-se no princípio de que as temperaturas em torno do tecido mamário que apresenta algum tipo de anomalia são maiores do que as do tecido normal (ACHARYA et al., 2014). Estudos que combinaram a mamografia e a termografia para a detecção do câncer de mama obtiveram uma sensibilidade de 95%. Houve uma melhora da sensibilidade de 95% para 98% quando o exame clínico também foi incluído (NG & SUDHARSAN, 2004).

A vantagem adicional da termografia, além do fato de não utilizar radiação ionizante, é a facilidade da realização de exames de mama em homens e a detecção de alterações mamárias em mulheres mais jovens que ainda não são submetidas a exames de mamografia pelas questões discutidas anteriormente (ARAÚJO et al., 2021).

No entanto, a termografia ainda não foi reconhecida como procedimento padrão para rastreamento do câncer de mama, sendo considerada um exame complementar ainda em estudo. A literatura mostra que rastreamento termográfico do câncer de mama pode ser incluído à rotina de exames de pacientes sintomáticas e assintomáticas. Para isso, é necessário investir, principalmente, em técnicas de processamento de imagens e de inteligência artificial.

Neste contexto, os sistemas de diagnóstico auxiliado por computador (CAD – *Computer-Aided Diagnosis*) são considerados uma importante ferramenta no auxílio ao diagnóstico médico do câncer de mama. A sua intenção é melhorar a consistência da interpretação das imagens considerando-se a resposta do computador. Esta resposta pode ser útil notadamente quando o diagnóstico de uma doença está relacionado à subjetividade de uma análise visual. Assim, uma dupla leitura por um médico especialista e por um computador pode melhorar significativamente a eficiência do diagnóstico (FURUIE et al., 1999 apud SILVA, 2007).

Um sistema CAD clássico é, geralmente, formado por processos que envolvem a segmentação da região de interesse, a extração de características e a classificação da imagem. Como é o caso de diversos problemas de tomada de decisão na área médica, há uma quantidade muito maior de pacientes considerados normais, com tumor benigno ou com cistos disponíveis para o treinamento de classificadores em comparação com o número de pacientes com tumor maligno. Esse fato leva a um problema de classificação de amostras desbalanceadas.

O desbalanceamento de classes pode gerar um sistema CAD que funciona bem na classe majoritária (por exemplo, a classe de Tumor Benigno), mas que possui um funcionamento não tão adequado para a classe minoritária (por exemplo, a classe de Tumor Maligno). Isso resulta numa baixa sensibilidade, o que torna essa abordagem inadequada. Dessa forma, considerar e resolver o desequilíbrio das classes é essencial para o desenvolvimento de um classificador com boa performance e diagnóstico confiável.

O desequilíbrio existente na amostra torna a etapa de aprendizagem mais complexa e, geralmente, esse desequilíbrio é acompanhado de outros pontos, como: o pequeno tamanho da amostra, pequenas disjunções relacionadas à classe minoritária e à sobreposição de classes. Segundo Batista, Prati e Monard (2004), o desbalanceamento entre as classes, por si só, não aparenta ser um problema difícil de resolver. Mas quando associado a classes altamente sobrepostas, pode diminuir significativamente o número de exemplos das classes minoritárias classificados corretamente.

Além de aplicações relacionadas à classificação de patologias mamárias, as imagens termográficas em conjunto com as simulações numéricas podem se tornar uma ferramenta importante no auxílio a detecção do câncer de mama. Segundo Jiang et al. (2011 apud BEZERRA, 2013), técnicas de modelagem numérica podem quantificar complexas relações entre os comportamentos térmicos e as condições fisiológicas da mama. Os perfis de temperatura obtidos através de simulação numérica de regiões mamárias considerados anormais podem ser usados para prever a localização e o tamanho de um tumor, assim como estimar os parâmetros térmicos da região. Tais perfis também podem ser usados para acompanhar o procedimento de tratamento (HOSSAIN & MOHAMMADI, 2016).

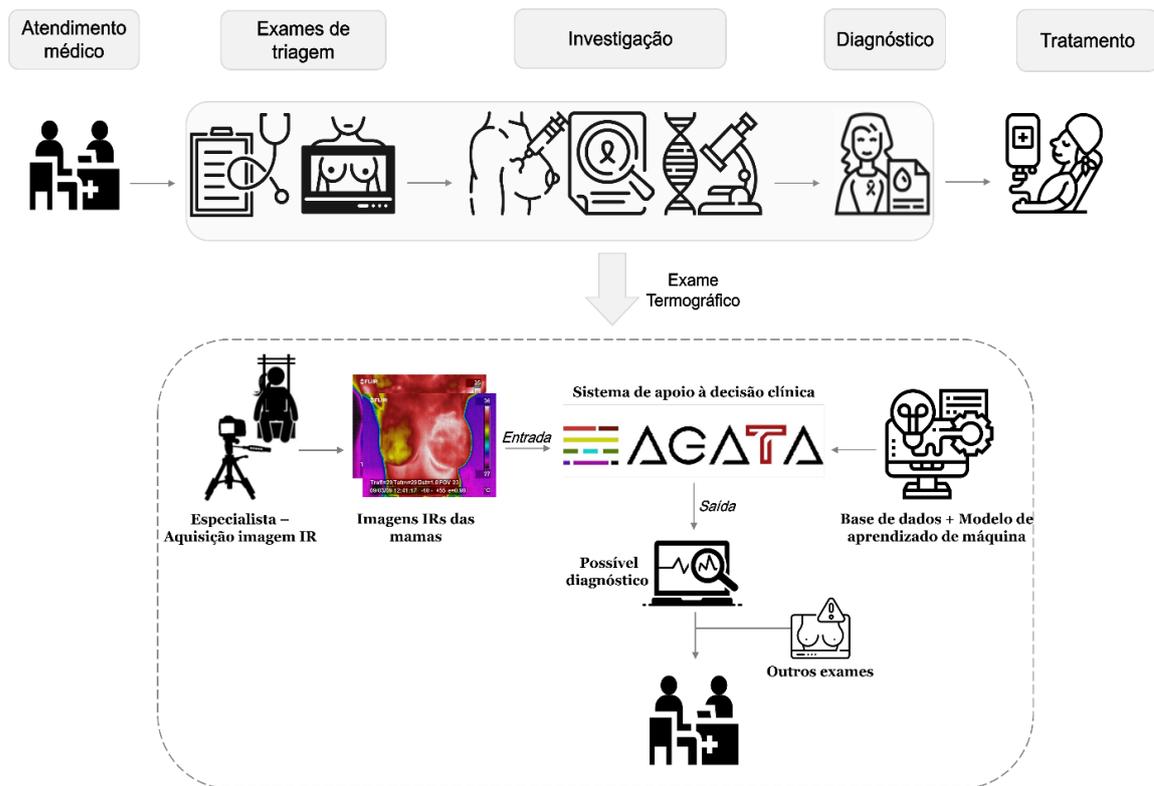
As imagens médicas dos métodos tradicionais de diagnóstico do câncer de mama fornecem informações clínicas relevantes para a tomada de decisão de um especialista. Utilizar ferramentas de modelagem e de simulações numéricas, associadas aos algoritmos de aprendizado de máquina e de otimização, baseadas em imagens termográficas pode proporcionar um aprimoramento na estimativa de parâmetros físicos, da posição do tumor, e na técnica de rastreamento de patologias mamárias.

Além da comprovação da termografia como método complementar no rastreamento do câncer de mama, o presente trabalho possui como contribuição a criação de sistemas de apoio ao diagnóstico validados do início (segmentação automática) ao fim (classificação), a fim de desenvolver uma ferramenta, aqui chamada AGATA<sup>1</sup>, de análise individual para aplicação clínica real. As imagens utilizadas foram obtidas de pacientes do Hospital das Clínicas da Universidade Federal de Pernambuco e inclui pacientes saudáveis, além de pacientes com cisto, lesão benigna e lesão maligna. Como os sistemas de imagem se propõem a não só diagnosticar alterações, mas também a identificar padrões de normalidade, essa é uma característica singular desse trabalho uma vez que a grande parte dos estudos não incluem a classe Saudável em suas análises (ver Capítulo 2). Métodos de sobreamostragem para lidar com bases desbalanceadas, métodos de seleção de atributos e de classificação de amostras e método de avaliação objetiva e rigorosa da ocorrência de sobreajustamento foram estudados, desenvolvidos e implementados em uma ferramenta destinada a contribuir em uma eventual triagem ou um diagnóstico do câncer de mama. A Figura 1 apresenta um esquema geral da solução proposta nesse trabalho, indicando onde a ferramenta AGATA auxiliaria os especialistas na investigação diagnóstica de patologias mamárias de uma paciente.

Figura 1-Simplificação do fluxo assistencial à paciente com patologia mamária na rede de saúde, com proposta geral de sistema de apoio ao diagnóstico

---

<sup>1</sup> Nome inspirado na história de Santa Ágata, considerada protetora das mulheres com câncer de mama.



## 1.1 OBJETIVOS

A seguir, são apresentados os objetivos geral e específicos do presente trabalho.

### 1.1.1 Objetivo geral

O presente trabalho tem como objetivo a elaboração de modelos eficientes de classificação de imagens mamárias obtidas por câmera infravermelho, com o intuito de auxiliar na triagem de pacientes e no diagnóstico precoce do câncer de mama. O modelo computacional tridimensional de mama desenvolvido por Melo et al. (2019) foi utilizado para fins de simulação de diferentes composições de tecido mamário, posições e tamanhos de tumor. O objetivo foi desenvolver sistemas de classificação confiáveis para, a partir do aprendizado baseado nas características extraídas dos perfis de temperatura simulados, indicar a localização e o tamanho de um tumor mamário em imagens termográficas. Todas as ferramentas numéricas e de classificação estão agrupadas em uma interface gráfica, aqui denominada AGATA, viabilizando a utilização ágil e prática das metodologias desenvolvidas.

### 1.1.2 Objetivos específicos

A fim de alcançar o objetivo geral da pesquisa, serão atendidos os seguintes objetivos específicos:

- Validar a técnica de segmentação automática (DOURADO, 2014) comparando com um *ground-truth*, aplicando métricas de avaliação presentes na literatura;
- Testar e analisar novas estratégias para classificação a fim de solucionar o problema de desbalanceamento da amostra e de sobreposição de classes;
- Calcular perfis de temperatura com a nova geometria 3D e ferramentas de CFD (*Computational Fluid Dynamics*) para prever a localização e o tamanho de tumores mamários em imagens termográficas;
- Desenvolver ferramentas computacionais que permitam uma triagem mais rápida e um auxílio ao diagnóstico precoce do câncer de mama.

## 1.2 ORGANIZAÇÃO DO TRABALHO

Este trabalho está dividido em seis principais capítulos. No Capítulo 1 foi apresentada uma introdução que engloba a motivação e os objetivos fundamentais do projeto desenvolvido. No Capítulo 2 são abordados trabalhos correlatos à temática da tese, destacando a aplicação da termografia na área da saúde e aplicações de sistemas CAD, associados a interfaces gráficas e a algoritmos de otimização. No Capítulo 3 são apresentados os conceitos necessários para compreensão do trabalho, tendo como exemplo os fundamentos da transferência de calor no corpo humano, as métricas de validação de segmentação de imagens e os conceitos de pré-processamento de imagens e de classificação. No Capítulo 4 é descrita a metodologia adotada para o desenvolvimento do projeto. Essa metodologia aborda, basicamente, as etapas para validação da segmentação automática; desenvolvimento de sistemas de classificação binária e multiclases e previsão do tamanho e da profundidade de lesão mamária em imagens termográficas. No Capítulo 5 são apresentados e discutidos os resultados obtidos, além da ferramenta computacional AGATA para análise individual de pacientes. Por fim, o Capítulo 6 fornece as conclusões e as sugestões para trabalhos futuros, seguido pelas Referências e pelo Apêndice.

## 2 REVISÃO BIBLIOGRÁFICA

A seguir são abordados trabalhos correlatos ao tema do presente trabalho.

### 2.1 SISTEMAS DE DIAGNÓSTICO APLICADOS ÀS IMAGENS TERMOGRÁFICAS DE MAMA

A detecção precoce do câncer de mama proporciona maiores chances de sobrevivência das pessoas assintomáticas e/ou sintomáticas. Em Ng (2009) foi preconizado que a termografia pode ter o potencial de detectar o câncer mamário dez anos antes do que a mamografia.

Apesar disto, existe ainda certa resistência da comunidade médica na utilização da termografia devido à baixa sensibilidade para detectar tumores pequenos e profundos, à impossibilidade de diferenciar tumores de pontos quentes correspondentes às inflamações locais e à subjetividade utilizada pelos médicos para interpretar o diagnóstico através dos termogramas (BORCHARTT, 2013). Como solução destes problemas, e como forma de proporcionar maior confiabilidade ao diagnóstico do câncer de mama, estão sendo desenvolvidos diversos sistemas CAD.

As técnicas básicas que compõem um sistema de CAD são: o processamento e o estudo da imagem para a detecção de anomalias, a extração de características da imagem processada, a análise de dados para a classificação da imagem (por exemplo, anomalia benigna ou maligna) e a avaliação quantitativa do desempenho do teste de diagnóstico utilizando, por exemplo, a análise ROC (*Receiver Operating Characteristic*) (PINHO, 2015).

Normalmente, os sistemas de CAD, baseados em imagens, consideram a contralateralidade entre as duas mamas. Em todos os tipos de exames utilizados para detectar alterações mamárias, espera-se encontrar determinada simetria entre as mamas. Então, evidências de assimetrias podem ser indicativas de doenças na mama ou de alguma alteração nela. No caso das imagens termográficas, uma elevada diferença nos padrões de distribuição de temperatura entre as mamas sugere um sinal de alerta (QUEIROZ, 2016).

Diversos estudos procuram identificar a possibilidade do uso de reconhecimento de padrões na classificação das imagens termográficas. Arora et al. (2008) analisaram um sistema de CAD térmico com três modelos de operação: triagem, clínico e rede neural artificial. Os autores não relataram sobre as características extraídas e nem como o *software* trabalha, mas expuseram as acurácias obtidas por cada modo: triagem, 66,70%; clínico, 71,40%; e rede neural artificial, 81,80%.

Schaefer, Závisek, Nakashima (2009) utilizaram uma amostra de 146 termogramas (29 casos malignos e 117 casos benignos) e trabalharam com 38 características, das quais 11 foram selecionadas após o uso de um filtro Laplaciano. O universo das características foi formado a partir de medidas estatísticas e de descritores de Fourier (medida de forma). Estas características foram analisadas através da lógica *fuzzy* com uma técnica de validação cruzada e os melhores resultados obtidos nesse estudo foram: 79,53% de acurácia, 79,86% de sensibilidade e 79,49% de especificidade.

No estudo de *softwares* que auxiliam no diagnóstico do câncer de mama, Acharya et al. (2012) utilizaram 50 imagens de infravermelho (25 de mamas saudáveis e 25 de mamas com câncer) para detectar automaticamente anormalidades mamárias. O processo de segmentação ocorreu a partir de um retângulo que envolve toda mama, o que resulta na inclusão de áreas que não pertencem à região de interesse (ROI – *Region Of Interest*) podendo comprometer as demais análises. Foram extraídas 16 características de textura, embora que no decorrer do desenvolvimento do método somente quatro características foram empregadas por terem sido consideradas como clinicamente significantes se comparadas com as demais. O método de classificação utilizado foi o *Support Vector Machine* (SVM) que mostrou um bom desempenho no reconhecimento de padrões de maior parte da amostra. Com esta técnica, os autores obtiveram 88,10% de acurácia, 85,71% de sensibilidade e 90,48% de especificidade.

Borchardt (2013) utilizou uma amostra composta por 69 imagens termográficas (19 de mamas saudáveis e 50 de mamas com patologias), sendo que 51 imagens foram obtidas no Hospital das Clínicas da Universidade Federal de Pernambuco (HC/UFPE) e 18 imagens foram obtidas no Hospital Universitário Antônio Pedro da Universidade Federal Fluminense (HUAP/UFF). É importante ressaltar que as imagens do HUAP/UFF que representam alterações foram classificadas segundo o seu BI-RADS: BI-RADS-1 para as pacientes normais e BI-RADS-2 para as pacientes anormais. Esta última metodologia é questionável, visto que a classe BI-RADS-2 também indica ausência de risco de nódulo maligno. Um pré-processamento foi realizado para corrigir a posição da paciente na imagem com intuito de melhorar os resultados da segmentação. Tal segmentação foi realizada automaticamente com auxílio do *software* desenvolvido por Marques (2012), porém, em alguns casos, ela não foi satisfatória sendo necessária uma segmentação manual.

No trabalho de Borchardt (2013) foram extraídas 40 características obtidas de medidas estatísticas simples, de dimensão fractal da região mamária e de medidas baseadas em geoestatística. A seleção das melhores características foi baseada no algoritmo genético. O classificador SVM foi utilizado para classificar a amostra em duas classes (normal e anormal) e o

processo de validação da classificação foi realizado pela técnica *leave-one-out* e pela *17-fold*. Como resultados foram obtidos os seguintes valores: 88,2% para taxa de acerto, 91,9% para sensibilidade e 78,6% para especificidade. Porém, estes resultados simplesmente apontam para a existência de alguma anormalidade, das quais a maior parte é benigna, sendo o teste pouco informativo sobre o efetivo risco de câncer.

Araújo (2014) propôs um sistema para classificação de anomalias mamárias (tumor maligno, tumor benigno e cisto) usando imagens termográficas. O pré-processamento da imagem foi iniciado pela segmentação da região das mamas, a partir de elementos elipsoidais gerados de forma manual por meio do Matlab<sup>®</sup>. Em seguida, um processo morfológico foi realizado sobre cada imagem obtida da segmentação, que incluiu a definição de um elemento estruturante e de processos de erosão, dilatação e *Top Hat*, para determinar a altura dos máximos locais de temperatura com relação à temperatura da sua vizinhança. A extração de características foi realizada sobre as imagens segmentadas e sobre aquelas obtidas da morfologia. Para tal extração foram propostas duas abordagens: a primeira baseou-se na extração de variáveis contínuas obtidas de medidas de dissimilaridade intervalar e que, em seguida, foram transformadas em um novo espaço de características pelo critério de Fisher. Na segunda abordagem foram extraídas variáveis intervalares de temperatura referentes às quatro matrizes de entrada. Foram usadas 50 pacientes e todos os classificadores usados (classificação por discriminação linear, classificação por distância mínima e por Janela de Parzen) foram validados usando a técnica *leave-one-out*. O resultado obtido na primeira metodologia foi de 84% para taxa de acerto da classificação; 85,7% para sensibilidade e 86,5% para especificidade da Classe Maligno; e na segunda metodologia atingiu 84% de taxa de classificação correta com 93% de sensibilidade para a Classe Maligno.

Dourado (2014) desenvolveu um sistema de CAD baseado na segmentação automática da região das mamas. Para isto, a matriz de temperatura associada a cada imagem termográfica foi tratada em tons de cinza. A primeira etapa da segmentação foi a separação entre a região do corpo e do fundo da imagem, a partir do processo de limiarização. O próximo passo foi a eliminação das regiões do pescoço e axilas, pela definição do ponto mais acima onde se iniciam os braços. A parte inferior foi eliminada através da seleção de bordas que começa na região central, entre as duas mamas, e segue em descida para cada limite inferior das mamas, excluindo cada pixel e toda sua linha abaixo se este possui o valor máximo da linha vertical a qual o pixel pertence e se o pixel vizinho lateral mais central já foi excluído. A última etapa da segmentação foi a retirada de possíveis regiões laterais superiores residuais, com a definição de linhas retas que se iniciam nos pontos entre as mamas e as axilas e que formam um ângulo de 60° com a

horizontal. Foram extraídas 20 características, dentre elas características intervalares e estatísticas, e a escolha das melhores foi realizada pelo WEKA (*Waikato Environment for Knowledge Analysis*). Para as melhores características, o SVM foi o classificador mais adequado com os seguintes resultados: 89,36% de taxa de acerto, 87,50% de sensibilidade e 90,32% de especificidade.

No estudo desenvolvido por Schaefer (2014) é mostrado como um sistema de classificação baseado na otimização de colônias de formigas pode ser aplicado ao problema de análise de termogramas mamários. Para isso, o autor empregou o *cAnt-Miner*, uma variação do classificador clássico *Ant-Miner*. Para todos os termogramas, foram extraídas 38 características. Na implementação do *cAntMiner*, foram usadas 2.000 formigas para a construção de regras, o número mínimo de amostras cobertas por regra foi definido como 3, o número máximo de amostras de treinamento descobertas pela base de regra como 3 e o número de regras usadas para testar a convergência foi definido como 5. O processo de validação da classificação foi realizado pela técnica *cross-validation* com 10-fold. Com essa metodologia, Schaefer (2014) alcançou uma acurácia de 79,52%.

Ali et al. (2015) propõem um novo método de segmentação automática de imagens por infravermelho de mamas. A fim de avaliar a precisão da abordagem de segmentação proposta, diferentes características seletivas foram extraídas das regiões segmentadas e, em seguida, o classificador SVM foi usado para detectar as anormalidades mamárias. O método de segmentação proposto é dependente da distância entre a câmera e o paciente durante a aquisição das imagens térmicas da mama. Uma vez que os termogramas foram obtidos, a primeira etapa foi para isolar a região das mamas. Em seguida, os autores utilizaram a equalização do histograma para realçar as imagens isoladas das mamas. Com as imagens que foram aprimoradas, os autores extraíram um total de 21 características, mas apenas 16 foram usadas para treinar o classificador SVM. Dentre os testes realizados, a acurácia de 80% foi a maior encontrada.

Utilizando lógica *fuzzy*, Calderon-Contreras et al. (2015) avaliaram 68 termogramas, sendo 20 de pacientes com neoplasia maligna e 48 pacientes saudáveis. O modelo proposto usa análise de assimetria com características estatísticas extraídas de termogramas comparando as características da área da mama direita e da área da mama esquerda. Na avaliação do desempenho do sistema *fuzzy* obteve-se uma acurácia de 89,71%, uma sensibilidade de 82,35% e especificidade de 92,15%.

Silva et al. (2015) analisaram as imagens termográficas de mamas para identificar pacientes com risco de câncer. O teste foi realizado com 22 pacientes (11 casos malignos e 11 normais), com a obtenção de 20 imagens para cada uma delas. Após o registro destas imagens,

foi obtida a formação do sinal térmico a partir da divisão da mama em regiões com tamanho de 3x3 pixels, do cálculo da temperatura média sobre cada região e da interpolação por convolução cúbica desses pontos. As características dos sinais térmicos foram calculadas através de medidas de complexidade e agrupadas automaticamente segundo seu grau de semelhança (clusterização). Os autores utilizaram técnicas de aprendizagem de máquina (rede neural, rede bayesiana e árvore de decisão) para classificar os vetores de características e obtiveram uma variação percentual de taxa de acerto de 86,36-91,90%, considerando os resultados de todos os classificadores.

Em estudos mais recentes, Schaefer e Nakashima (2015) utilizaram algumas estratégias para tratar do desequilíbrio de classes (maior número de casos benignos que malignos) no contexto de classificar o câncer de mama em imagens por infravermelho. A extração de características ocorreu de forma similar ao trabalho anterior desses autores (SCHAEFER; ZÁVISEK; NAKASHIMA, 2009). A principal sugestão do trabalho foi o uso de Sistemas com Múltiplos Classificadores (MCS – *Multiple Classifier Systems*) para melhorar o desempenho dos classificadores base. A primeira estratégia utilizada foi a realização da *classificação cost-sensitive* que incorpora diferentes fatores (ou funções) de custo na função de distribuição das classes, visando a obtenção de uma regra de decisão final que dá mais importância à classe maligna. A segunda estratégia foi a formação de um conjunto cujos classificadores base são treinados sobre subespaços equilibrados, definidos a partir de correlações, e analisados pela lógica *fuzzy*. Por fim, a terceira estratégia foi utilizar o classificador *on-class* sobre a amostra desbalanceada e tentar identificar objetos de uma classe específica, aprendendo a partir de um conjunto de treino que contém apenas os objetos dessa classe.

As três estratégias foram avaliadas para um conjunto de 146 termogramas (29 casos malignos e 117 casos benignos) e mostraram um bom desempenho no processo de classificação. Para o classificador *cost-sensitive*, para o classificador com amostra equilibrada e para o classificador *on-class* foram obtidos os seguintes grupos de valores, respectivamente: 81,02% de sensibilidade, 92,35% de especificidade e 90,10% de acurácia; 81,28% de sensibilidade, 90,53% de especificidade e 90,43% de acurácia; e 82,13% de sensibilidade, 88,78% de especificidade e 87,46% de acurácia.

No sistema proposto por Bhowmik et al. (2016), foram avaliados 20 termogramas de mama (10 saudáveis e 10 com alguma anormalidade) com o auxílio de análise estatística para avaliar a existência de assimetria térmica entre as mamas direita e esquerda. Como nos demais estudos citados, o processo de análise de um termograma inicia-se com o pré-processamento e a segmentação. Em seu estudo, os autores converteram a imagem para escala de cinza no intuito

de produzir uma nova imagem apenas com as bordas. Esta última é utilizada na confecção de uma máscara para extração manual da região de interesse da mama. Foram extraídos recursos estatísticos de primeira e segunda ordem que fornecem informações sobre a textura ou distribuição espacial dos valores de cinza do termograma. Os recursos estatísticos de primeira ordem calculam as propriedades do valor de *pixel* individual, enquanto os recursos estatísticos de segunda ordem estimam as propriedades de dois *pixels* de um determinado local. No total, os autores extraíram 12 características de cada mama do termograma analisado. Dessas, apenas sete foram usadas no processo de classificação. A sensibilidade, especificidade e acurácia do sistema de detecção de anormalidade baseado na análise de assimetria são de 90% cada.

O algoritmo proposto por Lashkari, Pak e Firouzmand (2016) consiste em quatro etapas principais. Na primeira etapa, usando operação totalmente automática, a ROI é determinada e a qualidade da imagem é melhorada e, com o auxílio de técnicas de limiarização e detecção de bordas, as mamas direita e esquerda são isoladas. Então, após esse processo, a matriz da imagem é normalizada e segue-se com a etapa de extração de 23 características, incluindo estatísticas, morfológicas, do domínio da frequência e dos recursos baseados em histograma e da matriz de coocorrência. Para obter os melhores recursos a serem utilizados na etapa de classificação, métodos de seleção de recursos, como mínima Redundância e Máxima Relevância (mRMR), Seleção Sequencial para Frente, Seleção Sequencial para Trás, Seleção Sequencial Flutuante para Frente, Seleção Sequencial Flutuante para Trás e Algoritmo Genético (AG) foram usados na terceira etapa. Finalmente, foram avaliados classificadores como AdaBoost, SVM, kNN (*k-Nearest Neighbor*), NB (Naive Bayes) e rede neural de probabilidade para encontrar o mais adequado. Essas etapas são aplicadas em diferentes graus de rotação das imagens por infravermelho. De acordo com os resultados experimentais, AG com AdaBoost foi a melhor combinação de seleção de recursos e de classificação de imagens termográficas de mama, com uma acurácia média de 85,33% e 87,42% para as imagens das mamas esquerda e direita, respectivamente, com 0 grau de rotação.

Silva et al. (2016) propuseram uma metodologia para classificar automaticamente os termogramas de indivíduos obtidos através do método dinâmico. A primeira etapa foi registrar as imagens na série temporal para reduzir o movimento entre os quadros usando o registro baseado em intensidade. O processo de registro é necessário na etapa de subtração para comparar as áreas correspondentes nas imagens. Após a etapa de registro das imagens, foi construída uma série temporal segmentando manualmente a região das mamas. Em seguida, a região das mamas é dividida em seções de 11x11 pixels. Os autores agruparam regiões nas imagens com base na temperatura usando o algoritmo *k-means*. Eles usaram aprendizado de máquina

não-supervisionado para avaliar o *clustering*. Em seguida, usaram o aprendizado de máquina supervisionado para gerar um esquema de classificação. Os autores utilizaram termogramas de 80 indivíduos: 40 com malignidade comprovada e 40 indivíduos sem câncer de mama. Foi testada uma variedade de classificadores, incluindo Redes Bayesianas e Redes Neurais. Os autores relataram 100% de sensibilidade e especificidade usando Rede Bayesiana, 92,59% de sensibilidade e 96,30% de especificidade usando Redes Neurais.

Em seu estudo, Madhavi e Bobby (2017) realizaram uma tentativa de distinguir indivíduos saudáveis e patológicos usando o método de decomposição empírica bidirecional (Método 1) e o padrão binário local girado uniforme (Método 2). Termogramas frontais foram pré-processados usando filtragem de difusão anisotrópica que preserva bordas nítidas, detalhes finos e remove efetivamente o ruído. Em seguida, a técnica de segmentação de conjunto de nível foi aplicada na imagem pré-processada para obter a ROI necessária. O Método 1 foi então aplicado as ROIs, sendo decomposta em funções de modo intrínseco e um resíduo. A análise de textura foi feita utilizando funções de modo intrínseco, usando o Método 2 para obter o vetor de recursos composto por 59 elementos únicos. A dimensionalidade dos recursos é reduzida para 20 usando *Kernel Principal Component Analysis*. Para a classificação, foi empregado o classificador *Least Square Support Vector Machine* com diferentes funções de *kernel*. Os melhores resultados experimentais foram obtidos com o classificador LSSVM usando *kernel* de função de base radial, alcançando 89% de precisão, 86% de sensibilidade e 100% de especificidade.

Vila Nova (2017) apresentou em seu estudo duas propostas de classificação de imagens por infravermelho. A primeira proposta consistiu em utilizar um classificador de quatro classes (Maligno, Benigno, Cisto e Normal), baseado na distância mínima de Mahalanobis para dados intervalares. A segunda proposta utilizou a combinação de classificadores com o objetivo de melhorar os resultados de classificação para anomalias mamárias, dando ênfase à acurácia e à sensibilidade à Classe Maligno. A primeira abordagem de classificação obteve 72,71% de taxa de acerto e 96,30% de sensibilidade à Classe Maligno. A segunda abordagem obteve 71,15% de taxa de acerto e 87,18% de sensibilidade à Classe Maligno, para um classificador multiclases, enquanto para um classificador binário, os resultados obtidos com a combinação de classificadores foram de 94,21% de taxa de acerto do classificador e 95,26% de sensibilidade à Classe Maligno.

Vasconcelos, Santos e Lima (2018) avaliaram a eficiência de diversos classificadores para o diagnóstico de alterações mamárias em termogramas. Foram utilizadas 175 pacientes em uma abordagem multiclases (Maligno, Benigno, Cisto e Normal). As etapas de pré-processamento e segmentação aplicadas foram as mesmas daquelas desenvolvidas por Dourado (2014).

Na extração de características, foram utilizadas abordagens fundamentadas em medidas estatísticas e em medidas intervalares. Dentre os classificadores testados pelos autores, o *Social Media Optimization* foi o que apresentou as melhores métricas de análise de classificação: 63,46% de acurácia, 80,77% de sensibilidade e 86,54% de especificidade.

Santana et al. (2018) extraíram os momentos de Zernike e Haralick com base na geometria e texturas, respectivamente. Termogramas de 100 pacientes foram usados em um total de oito classificadores. O classificador *Extreme Learning Machine* (ELM) usado no estudo compreende uma rede *feed forward* de camada única com geração aleatória de neurônios na camada oculta. Os resultados mostraram que o ELM obteve uma precisão de 65,95% e índice *kappa* de 0,4822 (usando divisão percentual) e ELM com precisão de 71,221% e índice *kappa* de 0,6676 (usando validação cruzada) forneceram os melhores resultados em comparação com outros classificadores. Esses resultados foram obtidos usando apenas características de Haralick. O Perceptron Multicamadas (do inglês, *Multiplayer Perceptron* – MLP) aumentou a precisão para 76,01% com *kappa* de 0,6402 (usando validação cruzada de 10 vezes) com a combinação dos momentos Haralick e Zernike. Os autores concluíram que o ELM provou ser um classificador promissor com treinamento rápido e custo computacional reduzido.

Jeyanathan, Jeyashree e Shenbagavalli (2018) aplicaram a transformada *wavelet*, *curvelet* e contorno para extrair características de termogramas. As características foram avaliadas estatisticamente usando teste t de amostra e teste t independente. A classificação usando rede neural Perceptron Multicamadas, treinada através da retropropagação, mostrou que as características *wavelet* do tipo Gabor classificaram os termogramas em normais e anormais, com valor de erro quadrático médio (EQM) de 0,3.

Gogoi et al. (2019), em seu estudo, investigaram a eficiência da termografia fazendo análise baseada em temperatura, análise baseada em intensidade e correspondência de localização de tumor. Várias características de temperatura e intensidade foram extraídas de cada termograma para caracterizar os termogramas de mama saudáveis, com tumores benignos e malignos. Treze conjuntos diferentes de recursos foram criados a partir das características de temperatura e intensidade extraídos e seus desempenhos de classificação foram avaliados usando o SVM com kernel de função de base radial. Entre todos os conjuntos de recursos, o conjunto formado por aqueles estatisticamente significativos ( $p < 0,05$ ) forneceu a maior acurácia com 83,22%, sensibilidade de 85,56% e especificidade de 73,23%.

Pramanik et al. (2019) em seu estudo realizaram a classificação de termogramas em saudáveis e malignos após o processo de extração das regiões suspeitas em potencial. Para a segmentação automática, um novo método de definição de nível baseado em região

denominado GRL-LSM foi proposto pelos autores. Inicialmente, as regiões potencialmente suspeitas são estimadas pelo método de limiarização adaptativo proposto, denominado GRL. Então, um método de definição de nível baseado em região (LSM) é empregado para segmentar precisamente as regiões potencialmente suspeitas. Após a fase de segmentação, alguns recursos de textura de ordem superior são extraídos da matriz de coocorrência. Uma rede neural artificial foi usada como classificador. A avaliação experimental mostrou que a estrutura proposta pode diferenciar as mamas malignas e saudáveis com uma acurácia de 89,4%, sensibilidade de 86% e especificidade de 90%.

Alfayez, El-Soud e Gaber (2020) propuseram uma abordagem de quatro fases para detecção de câncer de mama baseada em termograma: (1) Pré-processamento de imagem usando filtragem homomórfica, transformação *top-hat* e equalização adaptativa de histograma, (2) segmentação de ROI usando uma máscara binária e agrupamento *K-mean*, (3) extração de recursos geométricos, de textura e usando limite de assinatura (sinal característico para cada borda avaliada), e (4) classificação utilizando o ELM e o MLP. Comparando os classificadores testados, o ELM apresentou uma performance superior, com acurácia de 99,10%, sensibilidade de 98,05% e especificidade de 97,03%.

Arul Edwin Raj, Sundaram e Jaya (2020) apresentaram uma abordagem de equalização de histograma de nível de cinza auto-adaptativa para realçar a cor da imagem de infravermelho para detecção precoce de tumores. As características extraídas foram todas baseadas na matriz de coocorrência. Após a extração dos recursos, estes foram submetidos a uma redução de dimensionalidade por meio da técnica Análise de Componentes Principais (do inglês, *Principle Component Analysis*). A classificação final das imagens tumorais e não tumorais da mama foi obtida com o uso do classificador SVM. No estudo, os autores avaliaram uma base com 40 termogramas, sendo 16 com neoplasia maligna e 24 saudáveis. Com a metodologia aplicada pelos autores, foi possível obter uma acurácia, sensibilidade e especificidade de, respectivamente, 90,0%, 91,6% e 87,5%.

A ferramenta computacional desenvolvida por Ekici e Jawzal (2020) integrou uma rede neural convolucional (do inglês, *Convolution neural network*) otimizada pelo algoritmo de Bayes. Usando o algoritmo proposto pelos autores, foi possível alcançar uma taxa de acurácia de 98,95% para uma classificação binária de imagens térmicas no conjunto de dados para 140 pacientes.

Lennox e Haskins (2020) utilizaram uma base contendo termogramas de mamas de 30 pacientes doentes e 30 saudáveis. As imagens foram então segmentadas usando uma abordagem semiautomática, onde os sulcos inframamários foram detectados automaticamente e removidos

após alguns cortes manuais, de acordo com as orientações dos especialistas. Os recursos estatísticos e de texturas estatisticamente significativos foram extraídos das imagens em escala de cinza de alta resolução para serem usadas na classificação automática. Este estudo utilizou e comparou os resultados de classificadores comuns usados em tarefas de visão computacional (SVM, kNN, NB, MLP, *Random Forest* e AdaBoost). O classificador de melhor desempenho foi o *Random Forest* com uma acurácia de 90%.

No estudo de Mishra e Rath (2020), os termogramas de mama de 56 indivíduos com registros de temperatura disponíveis no *Database Mastology Research* foram considerados. Foram extraídos recursos de textura na matriz de nível de cinza e na matriz de coocorrência dessas imagens. Técnicas de redução de dimensionalidade, como PCA e *Autoencoder*, foram utilizadas para seleção das características. Para classificação binária de pacientes saudáveis e não-saudáveis, os autores testaram diferentes classificadores (SVM, árvore de decisão, *Random Forest*, kNN, regressão linear e lógica *fuzzy*) em combinação com as técnicas de redução de dimensionalidade mencionadas. Entre todos os classificadores, *Random Forest* com PCA apresentou uma acurácia de 95,45%.

Silva et al. (2020), utilizando a termografia infravermelha dinâmica propuseram uma rotina computacional para detecção de pacientes saudáveis ou com alguma anomalia maligna da mama. Após a obtenção das imagens sequenciais de cada paciente, seus arranjos de temperatura foram computados e novas imagens em escala de cinza foram geradas. Em seguida, o processo de segmentação extraiu a ROI e foram calculadas características baseadas em estatísticas, agrupamento, comparação de histogramas, geometria fractal, índices de diversidade e estatísticas espaciais. As séries temporais que são divididas em subconjuntos de cardinalidades diferentes foram geradas a partir de tais recursos. O classificador empregado foi o SVM, o qual atingiu taxas de acurácia, sensibilidade e especificidade de 100%.

Quadro 2 - Resumo dos estudos citados de sistemas CAD destinados à análise de imagens termográficas de mamas

Continua

Autor	Ano	Tipo de análise	Base de dados	Classificador	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
Arora et al.	2008	Binária	60 M 34 B	ANN	–	96,70	26,50
Schaefer, Závisek e Nakashima	2009	Binária	29 M 117 B	Lógica <i>fuzzy</i>	79,53	79,86	79,49
Acharya et al.	2012	Binária	25 M 25 S	SVM	88,10	85,71	90,48
Borchardt	2013	Binária	50 A 19 S	SVM	88,20	91,90	78,60
Araújo	2014	Multiclasses	14 M 19 B 17 C	Distância de <i>Mahala-</i> <i>nobis</i>	84,00	85,70	86,50
Dourado Neto	2014	Multiclasses	43 M 79 B 42 C 70 S	SVM	89,36	87,50	90,32
Schaefer	2014	Binário	29 M 117 B	ACO	79,52	–	–
Ali et al.	2015	Binário	34 M 29 S	SVM	80,00	–	–
Calderon-Con- treras et al.	2015	Binário	20 M 48 S	Lógica <i>fuzzy</i>	89,71	82,35	92,15
Silva et al.	2015	Binário	11 M 11 S	Árvore de decisão ANN Rede <i>Bayesiana</i>	90,91 90,91 86,36	90,91 100,00 100,00	90,91 84,61 78,57
Schaefer e Nakashima	2015	Binário	29 M 117 B	<i>Cost-sensitive</i> <i>On-class</i>	90,10 87,46	81,02 82,13	92,35 88,78

Quadro 2 - Resumo dos estudos citados de sistemas CAD destinados à análise de imagens termográficas de mamas

Continuação

<b>Autor</b>	<b>Ano</b>	<b>Tipo de análise</b>	<b>Base de dados</b>	<b>Classificador</b>	<b>Acurácia (%)</b>	<b>Sensibilidade (%)</b>	<b>Especificidade (%)</b>
Bhowmik et al.	2016	Binário	10 A 10 S	Estatística (assimetria térmica)	90,00	90,00	90,00
Lashkari, Pak e Firouzmand	2016	Binário	–	AdaBoost**	85,33 (ME); 87,42 (MD)	–	–
Silva et al.	2016	Binário	40 M 40 S	Redes Bayesianas ANN	– –	100,00 92,59	100,00 96,30
Madhavi e Bobby	2017	Binário	24 M 43 S	LSSVM	89,00	86,00	100,00
Vila Nova	2017	Multiclasses	27 M 22 B 22 C 16 S	Distância de <i>Mahalanobis</i> Multiclassificadores	72,71 71,15	96,30 87,18	– –
Jeyanathan, Jeyashree e Shenbagavalli	2018	Binário	25 A 25 N	Rede Neural de Retropropagação Multicamadas <i>Perceptron</i> (Agrupamento baseado no teste de Wilcoxon-Mann-Whitney)	–	–	–
Santana et al.	2018	Multiclasses	235 M 371 B 219 C	ELM MLP	71,22% 76,01%	–	–
Vasconcelos, Santos e Lima	2018	Multiclasses	41 M 47 B 46 C 41 N	SMO	63,46	80,77	86,54
Gogoi et al.	2019	Multiclasses	12 M 23 B 25 S	SVM	83,22	85,56	73,23

Quadro 2 - Resumo dos estudos citados de sistemas CAD destinados à análise de imagens termográficas de mamas

Conclusão

Autor	Ano	Tipo de análise	Base de dados	Classificador	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
Pramanik et al.	2019	Binário	25 M 25 S	ANN	89,40	86,00	90,00
Alfayez, El-Soud e Gaber	2020	Multiclasses	440 M 200 B 705 S	MLP ELM	80,04 99,10	84,00 98,05	61,60 97,03
Arul Edwin Raj, Sundaram e Jaya	2020	Binário	16 M 24 S	SVM	90,00	91,60	87,50
Ekici e Jawzal	2020	Binário	42 M 98 B	CNN	98,95	–	–
Lennox e Has-kins	2020	Binário	30 A 30 S	<i>Random Forest</i>	90,00	–	–
Mishra e Rath	2020	Binário	37 A 19 S	<i>Random Forest</i>	95,45	99,17	88,07
Silva et al.	2020	Binário	32 A 32 S	SVM	100,00	100,00	100,00

Fonte: A Autora (2022)

\*A = Anormal (inclui qualquer patologia mamária); B = neoplasia benigna de mama; C = cisto mamário; M = neoplasia maligna de mama; S = saudável; MD = mama direita; ME = Mama esquerda.

\*\* Foi utilizado nesta revisão apenas o melhor resultado do estudo.

## 2.2 INTERFACES GRÁFICAS RELACIONADAS À TERMOGRAFIA

Uma interface gráfica de usuário (GUI - *Graphical User Interface*) é uma interface pictórica com um programa. Uma boa GUI pode, sobretudo, tornar os programas mais fáceis de usar, ao fornecer para eles uma aparência consistente e com controles intuitivos. A GUI deve se comportar de maneira compreensível e previsível, para que o usuário saiba o que esperar ao executar uma ação (CHAPMAN, 2003).

O estudo de Santos et al. (2007), foi motivado pela dificuldade na medição de temperaturas in vivo de pacientes submetidos à radiofrequência para tratamento de tumores de próstata. A principal ferramenta deste trabalho foi o software API (Aquisição de Pontos em Imagem), desenvolvido na plataforma Matlab®, que facilita a manipulação e extração dos pontos e domínios em imagens médicas bidimensionais. A interface desenvolvida para este programa possui quatro menus: Arquivo, Operações, Visualizar e Ajuda. O menu Arquivo (Figura 2) é o local onde é possível abrir e salvar imagens, além de permitir ao usuário o dimensionamento da imagem para definir medidas reais do modelo. No menu Operações encontram-se ferramentas para manipulação da imagem, possibilitando a determinação dos domínios e dos segmentos em que ela será segmentada e a definição da condição de contorno do segmento. O menu Visualizar permite que o usuário verifique o resultado dos arquivos de saída, disponibilizados para serem abertos por qualquer editor de texto comercial. Finalmente, o menu Ajuda fornece ao usuário a opção de esclarecer dúvidas quanto ao funcionamento do programa, bem como a versão do aplicativo e as últimas atualizações.

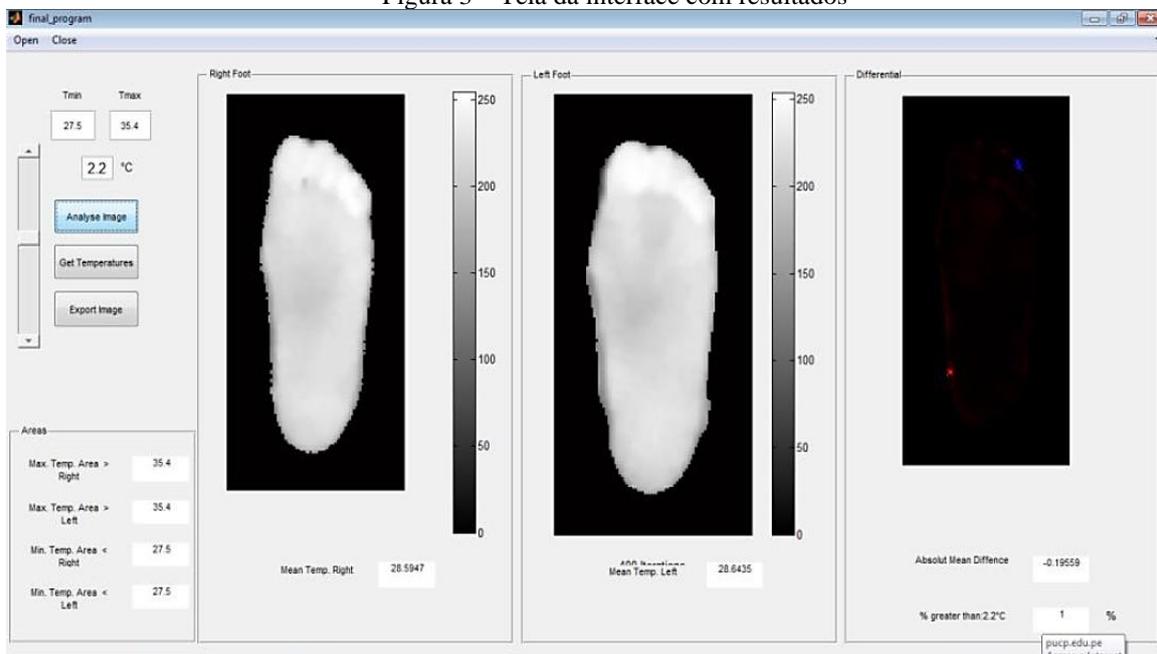
Figura 2 – Tela principal (Menu Arquivo)



Fonte: Santos et al. (2007)

Pomalaza (2016) desenvolveu uma interface gráfica no Matlab<sup>®</sup> que realiza a segmentação e classificação de imagens termográficas de pés de pacientes diabéticos, auxiliando no diagnóstico precoce de patologias associadas a complicações da diabetes. Com a interface (Figura 3), o usuário pode selecionar a imagem a ser analisada clicando em “Open”. Em seguida, deve-se usar o botão “Get Temepratures” para obter-se as temperaturas máximas e mínima da imagem. Após essa etapa, pode-se iniciar o processo de análise com o botão “Analyse image”. Com isso, a imagem será processada, segmentada e classificada. O processo é finalizado com a visualização dos resultados na interface.

Figura 3 – Tela da interface com resultados



Fonte: Pomalaza (2016)

Farias et al. (2019) propuseram um sistema computacional preditivo para a caracterização da Persistência do Canal Arterial (PCA) em imagens termográficas de recém-nascidos, baseado no uso de uma Rede Neural Artificial (RNA). O PCA *Diagnostic* (Figura 4) é um sistema *standalone* que permite ao usuário ter um auxílio automatizado ao diagnóstico da PCA a partir do processamento de imagens termográficas de neonatos com suspeita da doença. Esse sistema foi desenvolvido no *software* Matlab<sup>®</sup>. Informações sobre a qualidade de predição que a RNA adquiriu após o carregamento do programa é exibida no painel de *Status*. Caso o último desempenho de treinamento obtido, baseado na taxa de erro de predição, não tenha sido satisfatório, o usuário tem a opção de realizar um novo treinamento clicando no botão “Treinar RNA”, Para iniciar a predição, o usuário deve abrir uma imagem termográfica clicando em “Abrir Imagem Térmica”. A imagem termográfica selecionada é exibida no painel “Imagem

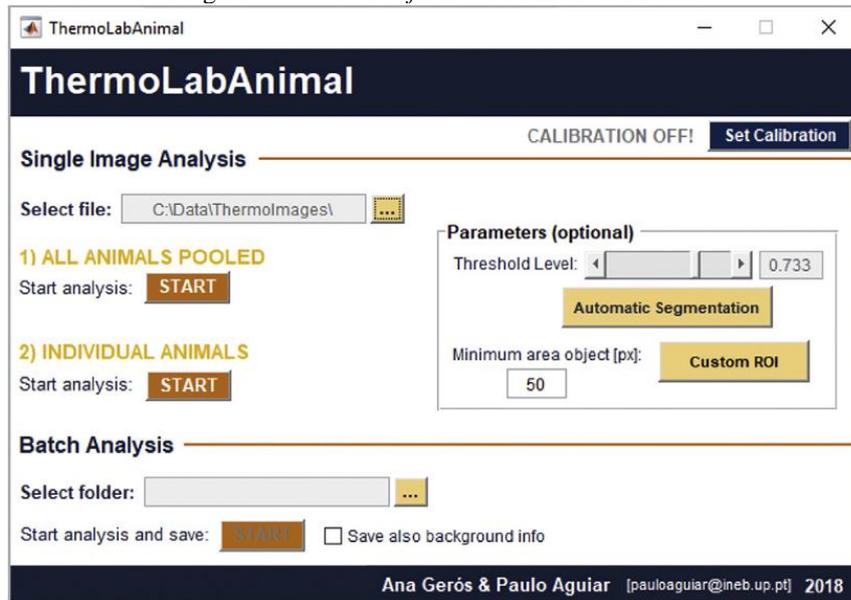
Original” e, após o processo de segmentação automática de tronco, a imagem segmentada é apresentada no painel “Tronco Segmentado” e as características extraídas são exibidas no painel “Console”. Com as características da imagem térmica disponíveis, o sistema realiza a classificação da imagem. O resultado da classificação da rede é apresentado logo em seguida no painel *Status*.

Figura 4 – Tela da inicial do sistema PCA *Diagnostic*



Fonte: Farias (2019)

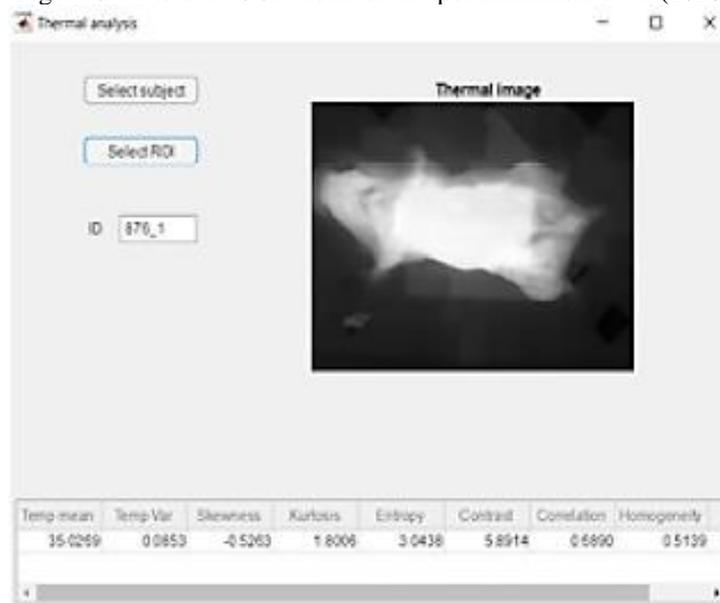
Franco et al. (2019) apresentaram a ferramenta ThermoLabAnimal, uma interface gráfica desenvolvida no Matlab<sup>®</sup> que possui um método para avaliação automática da temperatura média da superfície corporal em ratos utilizados em ambiente laboratorial. A GUI principal (Figura 5) fornece um acesso simples às ferramentas/funcionalidades do *software*, com a possibilidade de se optar por uma análise de imagem única ou pela análise em lote. Depois de carregar uma imagem termográfica na análise de imagem única, o usuário pode inspecionar a imagem usando ferramentas de visualização padrão como zoom ou ler valores de pixel específicos. O usuário pode modificar parâmetros da segmentação, como o valor do nível de separação dos pixels dos corpos dos animais e do fundo da imagem (*Threshold level*) e o valor mínimo da área da máscara de segmentação (*Minimum area object*). Ao fim do processo, são exibidas a imagem segmentada com os animais e um histograma com a distribuição da temperatura da superfície de todos os ratos segmentados.

Figura 5 – GUI do *software* ThermoLabAnimal

Fonte: Franco et al. (2019)

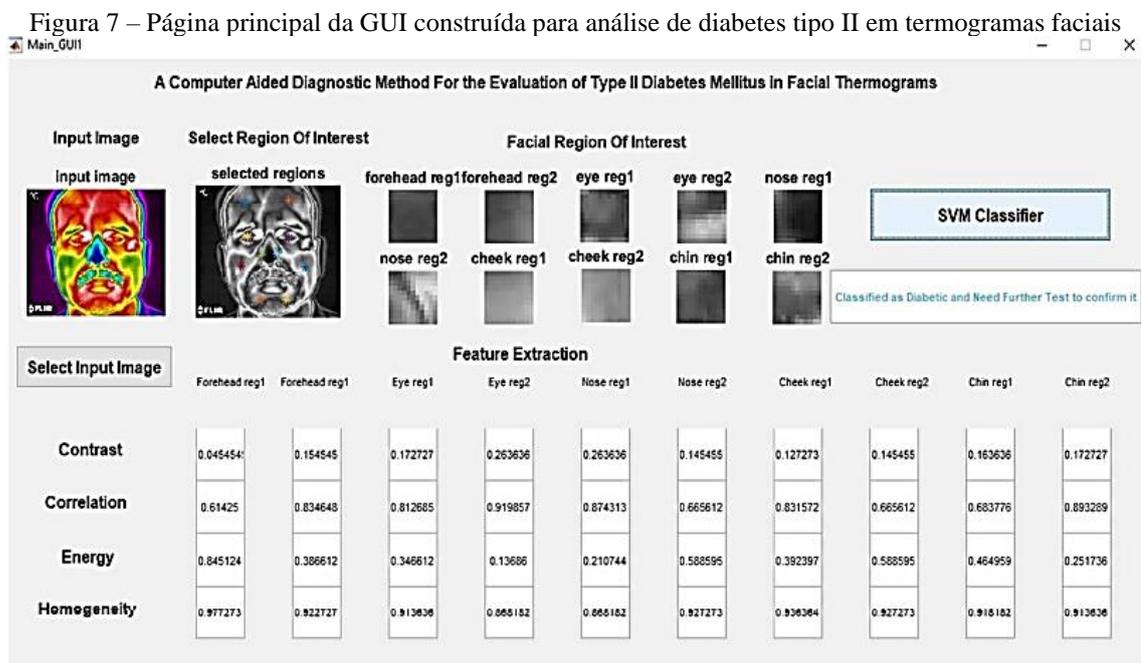
Brzezinski et al. (2020) desenvolveram um novo algoritmo de processamento de imagem para prever automaticamente o estado de doença hepática gordurosa, em um modelo animal. Os autores acoplaram esse novo algoritmo a uma interface gráfica (Figura 6) construída no Matlab®. O usuário, por meio dos botões, pode selecionar uma imagem a ser avaliada, extrair sua ROI e, após esse processo, são exibidas na parte inferior da GUI, as nove características usadas para realizar a classificação.

Figura 6 – Visão da GUI desenvolvida por Brzezinski et al. (2020)



Fonte: Brzezinski et al. (2020)

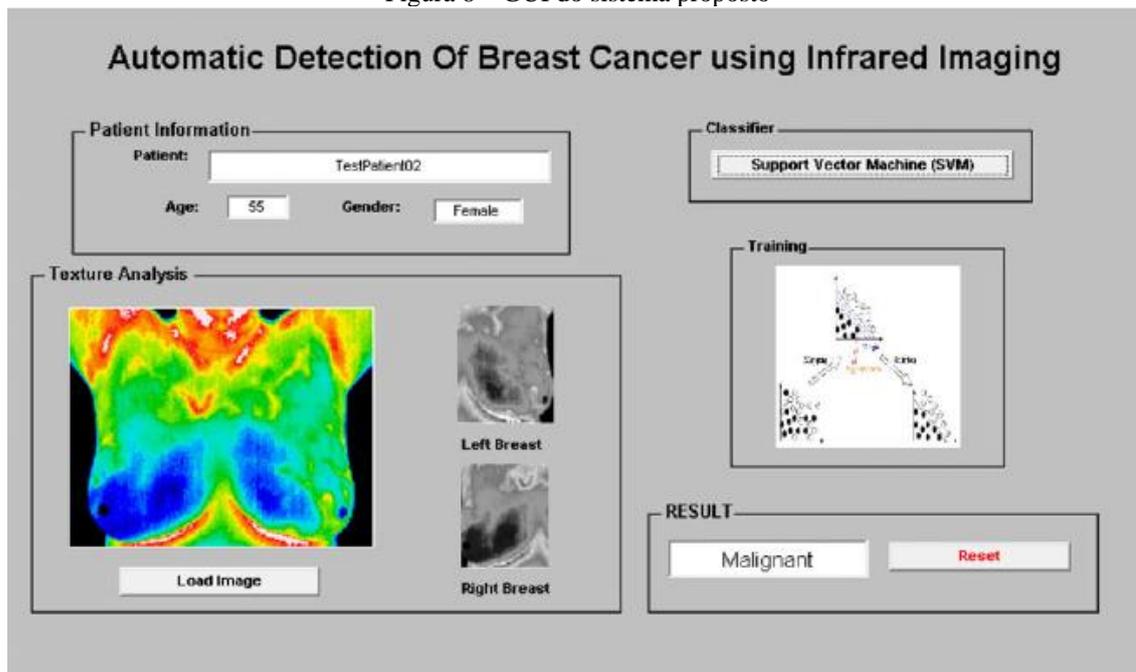
Thirunavukkarasu et al. (2020) avaliaram o diabetes tipo II em termogramas faciais por meio de um sistema desenvolvido no Matlab®. O classificador implementado na GUI (Figura 7) foi o SVM que, dentre os classificadores testados pelos autores, foi o que apresentou melhor acurácia. Na parte inferior da interface, são apresentadas as quatro características mais relevantes estatisticamente, quanto à separação das classes diabético e normal. Após a seleção da imagem, o usuário pode iniciar o processo de análise. Na região superior do *software* são apresentadas as ROIs obtidas da imagem, enquanto na região inferior são expostas as características extraídas para cada uma das dez ROIs.



Fonte: Thirunavukkarasu et al. (2020)

Para o sistema proposto por Acharya et al. (2012) foi desenvolvida uma interface gráfica (Figura 8) na plataforma Matlab®. O botão “Load Image” permite que o usuário carregue a imagem termográfica na tela da própria interface e, logo em seguida, são apresentadas a imagens em tons de cinza resultantes da segmentação. No campo “Patient Information” contém as informações do paciente (nome, idade e gênero) correspondentes à imagem selecionada. Por fim, ao clicar o botão “Support Vector Machine (SVM)” as características da imagem são extraídas automaticamente, depois são utilizadas como entrada do classificador SVM e o resultado desse processo é exposto no campo “Result”. Na figura, observa-se que a classe final do termograma foi a Maligna.

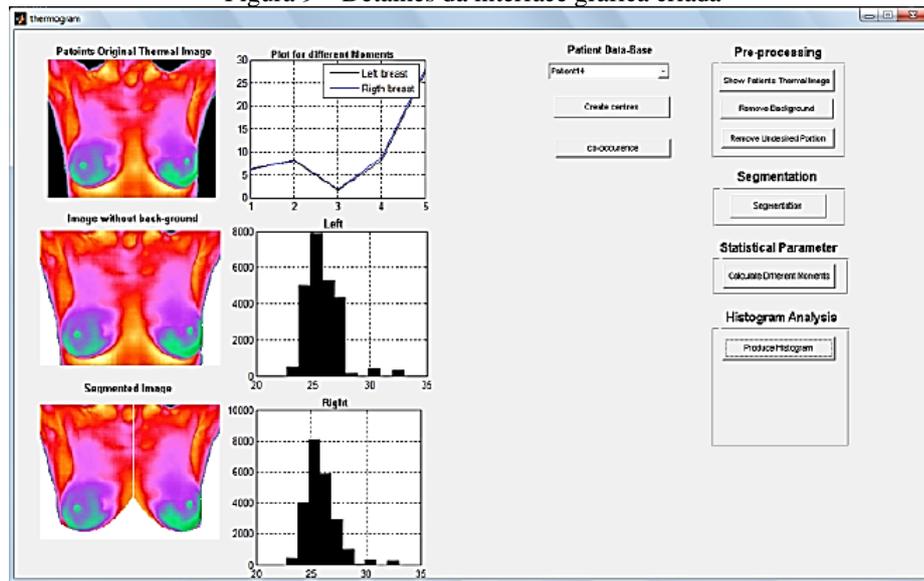
Figura 8 – GUI do sistema proposto



Fonte: Acharya et al. (2012)

Kapoor, Prasad e Patni (2012) criaram uma interface gráfica no Matlab<sup>®</sup> que inclui técnicas de segmentação automática da ROI e de análise de assimetria entre as mamas de imagens termográficas. Através dessa interface (Figura 9) o usuário pode escolher a imagem e selecionar a opção de pré-processamento dela, que consiste na remoção do fundo e no redimensionamento da imagem para remover partes do corpo indesejadas. Apertando o botão de segmentação, as áreas das mamas são segmentadas automaticamente através dos operadores Canny de detecção de bordas e de gradiente. A análise da assimetria é feita no acionamento do botão que, inicialmente, extrai os parâmetros de alta ordem estatística (do inglês, *High Order Statistical*) e calcula o centro de gravidade de regiões da mama, e do botão de geração de histograma que possibilita a análise da distribuição de temperatura sobre cada ROI. A interface permite ainda a visualização da imagem escolhida, da imagem pós-processamento e da imagem segmentada. No geral, os resultados mostraram que o estudo da assimetria auxilia no diagnóstico médico de anormalidades mamárias. Além disso, o uso da GUI permite uma abordagem mais eficaz, independente do sistema operacional. Porém, a desvantagem da análise realizada apenas por meio de uma segmentação é subjetividade e a baixa confiabilidade relacionada à ausência de um sistema de classificação, por exemplo.

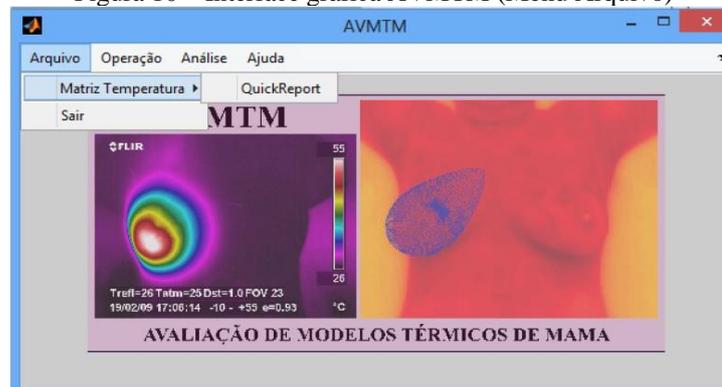
Figura 9 – Detalhes da interface gráfica criada



Fonte: Kapoor, Prasad, Patni (2012)

Santos (2014) desenvolveu um *framework*, nomeada de AVMTM (Avaliação de Modelos Térmicos de Mama), para análise de um modelo térmico da mama, a partir de ferramentas numéricas agrupadas numa interface gráfica. A interface criada (Figura 10) possui quatro menus: Arquivo, Operação, Análise e Ajuda. No menu Arquivo é possível abrir a imagem por infravermelho através do programa *QuickReport* e exportar a matriz de temperaturas como um arquivo *.csv*. No menu Operação, duas técnicas podem ser realizadas: mapeamento (frontal/lateral) ou triangulação (interpolação) e ainda a operação “Pontos comuns” a qual encontra as temperaturas dos pontos em comum nos dois mapeamentos, salvando-os em um arquivo *.txt*. No menu Análise podem-se escolher dois modelos de tumor (esférico/cilindro com bases elípticas) e realizar simulações com várias posições, tipos de tumor e tamanhos. Por fim, o menu Ajuda proporciona ao usuário uma explicação sobre a interface e suas funcionalidades.

Figura 10 – Interface gráfica AVMTM (Menu Arquivo)



Fonte: Santos (2014)

Bayas (2015) desenvolveu uma interface destinada a segmentação e análise de imagens termográficas. Os comandos foram programados no Matlab® e o menu principal (Figura 11) da ferramenta consiste em, basicamente, três opções: Pré-processamento, Características e Sair. Ao ser acionado o botão de Pré-processamento, a imagem original selecionada é apresentada e uma janela, com opções de ajuste de realce, filtragem e segmentação, é aberta. O botão Características fornece informações sobre os valores das características estatísticas e de textura extraídas nas imagens segmentadas (ROI completa antes da divisão entre as mamas, mama esquerda e mama direita). Além disso, é possível obter a classificação, através do algoritmo kNN, da imagem selecionada como sendo normal ou anormal.

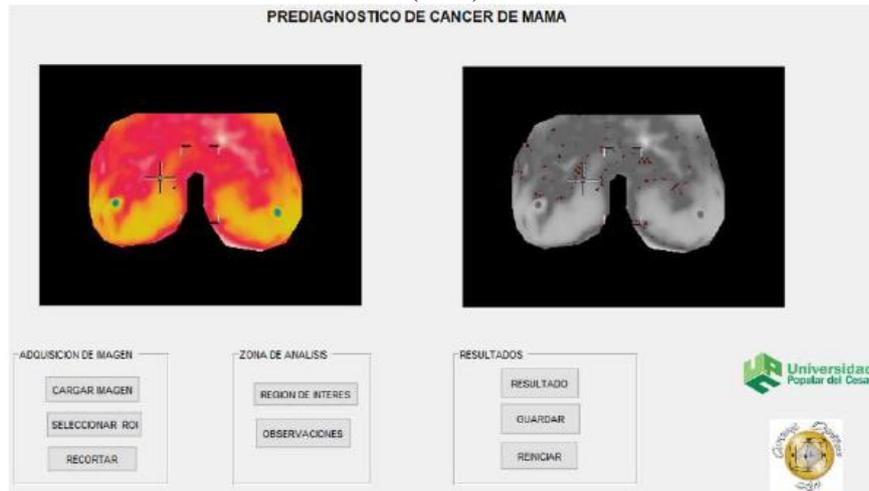
Figura 11 – Tela inicial da interface gráfica desenvolvida por Bayas (2015)



Fonte: Bayas (2015)

Em seu estudo, Granadillo et al. (2019) desenvolveram um algoritmo para detecção do câncer de mama por meio do processamento de imagens termográficas da mama e associaram o algoritmo desenvolvido a uma interface gráfica para facilitar seu uso. A interface (Figura 12) projetada permite a aquisição da imagem do computador com o botão “Cargar imagen” e a seleção manual da área de estudo ou ROI, por meio dos botões de “Selección de ROI” e “Recorte”. A imagem é então processada capturando apenas as áreas de estudo que apresentam componentes vermelhos, e esse processo é mostrado com o botão “Región de interés”. Por fim, quando o botão “Resultados” é pressionado, o *software* utiliza as ROI geradas (áreas cujos valores apresentam maior média de valores vermelhos) e classifica a imagem como doente ou saudável.

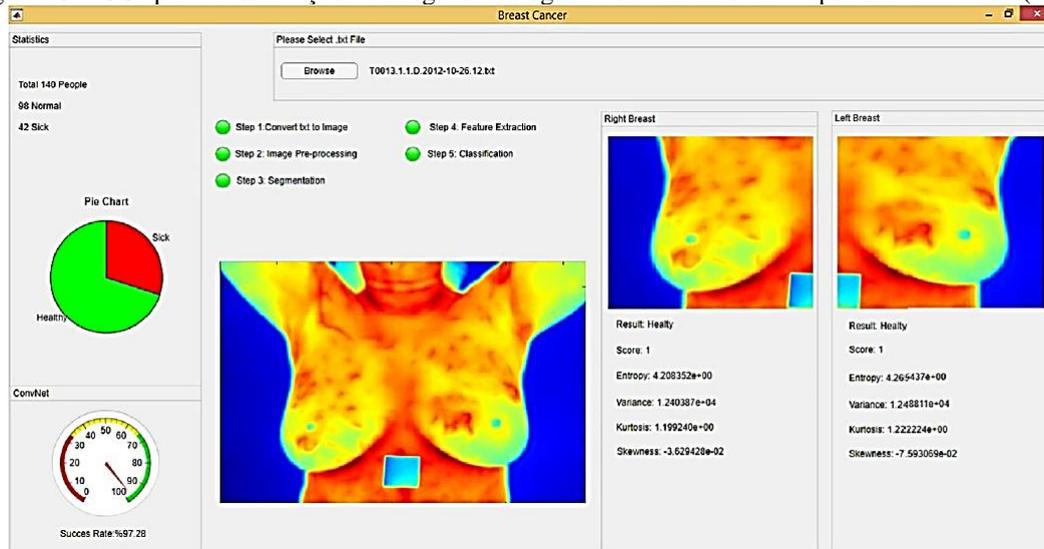
Figura 12 – Interface para auxiliar o diagnóstico precoce do câncer de mama desenvolvida por Granadillo et al (2019)



Fonte: Granadillo et al. (2019)

Ekici e Jawzal (2020) em seu estudo, também produziram uma interface para classificação de termogramas mamários (Figura 13). Para utilização dessa ferramenta computacional, cinco etapas são necessárias. Na Etapa 1 os arquivos de texto são convertidos em imagens para análise visual. Na Etapa 2 é realizado o pré-processamento da imagem, a fim de obter-se imagens claras e distinguíveis. As Etapas 3 e 4 são, respectivamente, segmentação da imagem e extração de recursos. Na Etapa 5 é realizada de fato a classificação da imagem, utilizando redes neurais convolucionais.

Figura 13 – GUI para classificação de imagens termográficas de mamas criada por Ekici e Jawzal (2019)



Fonte: Ekici e Jawzal (2020)

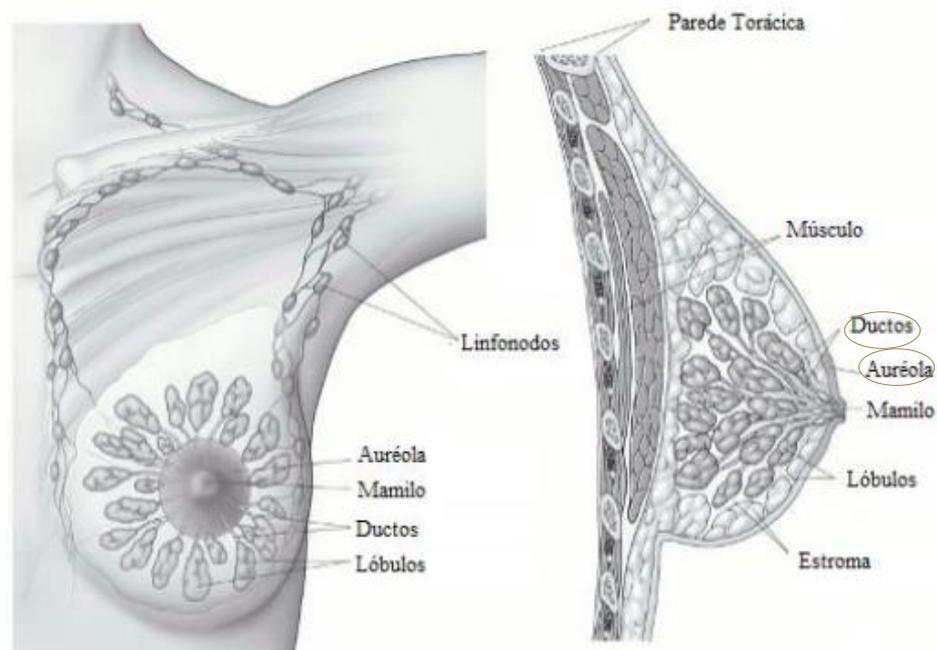
### 3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os conceitos e os modelos matemáticos que fundamentaram este trabalho.

#### 3.1 O CÂNCER DE MAMA

Os carcinomas mamários, na sua maioria, são originados no epitélio dos lóbulos e ductos da glândula mamária (POPOLIN, 2016). Esses subtipos (ductal e lobular) podem se apresentar na forma de carcinoma *in situ*, em que o tumor não possui características invasivas, permanecendo associado ao tecido mamário, e na forma invasiva, ocupando os tecidos adjacentes e disseminando-se para outros órgãos (YODER; WILKINSON; MASSOLL, 2007). A Figura 14 apresenta uma vista anatômica da mama feminina com os locais frequentemente comuns de incidência em destaque.

Figura 14 - Anatomia feminina da mama com regiões de maior incidência de câncer destacadas



Fonte: Popolin (2016)

O desenvolvimento do câncer de mama pode estar associado a fatores de risco que incluem a idade da mulher, a menstruação precoce, a terapia de reposição hormonal, a menopausa tardia, a alimentação e ao histórico familiar. Esses fatores estão relacionados a uma maior pré-

disposição da mulher em desenvolver o câncer, porém a sua formação está associada com uma proliferação celular descontrolada, anormal e autônoma (BRASILEIRO FILHO, 2016).

A autonomia de sobrevivência das células neoplásicas é possibilitada, entre outros, pela neoformação vascular ou angiogênese. As células malignas, para garantir o suprimento sanguíneo, induzem a formação de novos vasos a partir de vasos vizinhos. Na maioria dos tumores, a angiogênese e a malignidade estão correlacionadas: quanto maior a atividade angiogênica, maior é a potência de metastização do câncer e mais rápido é o seu progresso (BRASILEIRO FILHO, 2016).

### 3.2 TRANSFERÊNCIA DE CALOR NO CORPO HUMANO

Os organismos vivos, desde o mais simples ao mais complexo como os seres humanos, necessitam trocar matéria e energia com o ambiente circundante para sobreviverem. O homem, em sua complexidade, atua como sistema termodinâmico aberto e como produtor do que se pode chamar de energia térmica

A energia térmica é produzida pelas diversas reações químicas que ocorrem no corpo humano cujo intuito é fornecer energia interna para o sistema do corpo. Para manter a função corporal saudável, o ser humano deve manter uma temperatura interna estável.

As mudanças na temperatura do corpo são gerenciadas, em certa medida, pelas diversas atividades metabólicas no indivíduo. O calor gerado no corpo pelo metabolismo é dissipado para o meio através da condução, da convecção, da radiação e da evaporação (KAMANGAR et al., 2019; RAVANELLI; BONGERS; JAY, 2019).

O primeiro modelo térmico do corpo humano foi proposto por Burton no ano de 1934. Ele representou o corpo humano como um cilindro homogêneo e com geração uniforme de calor metabólico (BURTON, 1934). Foi demonstrado que o perfil de temperatura no estado estacionário possui um perfil parabólico, comprovando as observações experimentais de Bazett e McGlone (BAZZET & MCCLONE, 1927 apud SILVA, 2012; KAMANGAR et al., 2019).

Posterior a este trabalho, modelos mais complexos foram desenvolvidos para analisar a transferência de calor em seres vivos, dentre estes os principais são: o modelo de Pennes, o modelo de Wulff, o modelo de Klinger, o modelo de Chen e Holmes, o modelo dos vasos emparelhados e o modelo de Weinbaum, Jiji e Lemons (SILVA, 2012) descreveu em seu trabalho alguns destes modelos.

O modelo desenvolvido por Harry H. Pennes, em 1948, é referência dos modelos térmicos do corpo humano (ALBUQUERQUE NETO, 2010) devido à sua simplicidade e à menor necessidade de parâmetros termofísicos que são, muitas vezes, desconhecidos ou não muito bem determinados (BEZERRA, 2013). Pennes realizou uma modificação na equação da condução de calor, que passou a ser conhecida como “Equação da Biotransferência de Calor” (BHTE - *Bioheat Transfer Equation*) ou como “Equação de Pennes”, para levar em conta a geração de calor metabólica e a troca de energia térmica entre o sangue em escoamento e o tecido circundante (INCROPERA et al., 2007).

Dessa forma, realizando o balanço total de energia para um determinado tecido e desprezando os efeitos químicos, elétricos e nucleares obtém-se a Equação 1 abaixo:

$$\rho_t c_t \frac{\partial T_t}{\partial t} = \nabla \cdot (k_t \nabla T_t) + Q_m + Q_p + Q \quad (1)$$

onde  $\rho_t$  é massa específica do tecido ( $kg/m^3$ );  $c_t$  é o calor específico do tecido ( $J/(kg.K)$ );  $k_t$  é a condutividade térmica do tecido ( $W/(m.K)$ );  $T_t$  é a temperatura do tecido ( $K$ );  $t$  é a variável tempo ( $s$ );  $Q_m$  é a taxa volumétrica de geração de calor metabólico ( $W/m^3$ );  $Q_p$  é a taxa volumétrica de geração/remoção de calor devido à perfusão sanguínea ( $W/m^3$ ) e  $Q$  é a taxa volumétrica de geração/remoção de calor, incluindo outras fontes externas ( $W/m^3$ ).

A presença de um tecido canceroso influencia na alteração da distribuição de temperatura. Isso porque o rápido crescimento das células cancerígenas exige um aumento do número de vasos (angiogênese) para transportar nutrientes. Desse modo, a angiogênese aumenta a taxa metabólica dos tecidos circundantes, induzindo alterações nas temperaturas de superfície por meio da condução nos tecidos e da convecção nos vasos sanguíneos.

A fonte de calor devido à perfusão sanguínea pode ser modelada pela seguinte expressão

$$Q_p = \omega \rho_s c_s (T_a - T_v) \quad (2)$$

onde  $\omega$  é a taxa de perfusão sanguínea ( $m^3/s$ );  $\rho_s$  é a massa específica do sangue ( $kg/m^3$ );  $c_s$  é o calor específico do sangue ( $J/(kg.K)$ );  $T_a$  é a temperatura do sangue arterial entrando no volume de controle do tecido ( $K$ ) e  $T_v$  é a temperatura do sangue venoso saindo volume de controle do tecido ( $K$ ).

Para fins de simplificação, a temperatura do sangue venoso  $T_v$  pode ser aproximada a temperatura local do tecido  $T_t$  (CHARNY, 1992). A aproximação considera que a troca de calor

é completa, isto é, assume-se que a temperatura do sangue se iguala à temperatura do tecido. Também se considera que o efeito da perfusão é homogêneo e isotrópico (CHARNY, 1992). Assim, a Equação 2 pode ser reescrita como

$$Q_p = w\rho_s c_s (T_a - T_t) \quad (3)$$

A geração de calor metabólico  $Q_m$  é homogeneamente distribuída em todo tecido de interesse como uma taxa de deposição de energia por unidade de volume. Ela possui valores diferentes para o tecido do tumor e da mama. O calor  $Q_m$  do nódulo maligno está relacionado com seu tempo de duplicação  $\tau$  (dias) através de uma função hiperbólica (Equação 4). O diâmetro  $D$  (cm) do tumor varia segundo  $\tau$  e pode ser calculado pela Equação (3.5).

$$Q_m\tau = C, \text{ com } C = 3,27 \times 10^{26} \text{ W} \cdot \text{dia} \cdot \text{m}^3 \quad (4)$$

$$D = 0,01e^{[0,002134(\tau-50)]} \quad (5)$$

A troca total de energia entre o sangue e o tecido é diretamente proporcional à massa específica, ao calor específico e a taxa de perfusão do sangue através do tecido, e é descrita em termos da troca de calor sensível do sangue, como se pode observar na Equação 2. A principal vantagem do modelo de Pennes (Equação 1) sobre os demais modelos, é que o termo que leva em consideração a perfusão sanguínea varia linearmente com a temperatura (BEZERRA, 2013).

### 3.3 MÉTRICAS DE ANÁLISE DE PERFORMANCE DA SEGMENTAÇÃO DE IMAGENS DIGITAIS

O processo de segmentação de uma imagem consiste na identificação e separação de regiões que compartilham entre si características semelhantes, tais como cor, textura, intensidade ou regiões contralaterais (JACKMAN; SUN; ALLEN, 2009; DINIZ, 2017). Consequentemente, a segmentação fornece algumas regiões na imagem de saída que podem ser interpretadas com mais facilidade.

Para a extração da região de interesse, tanto a segmentação automática quanto a manual pode ser utilizada no suporte ao diagnóstico do câncer mamário. A segmentação automática, sem intervenção humana, em geral é mais rápida e proporciona uma maior objetividade na análise final. Normalmente, a opção pela segmentação manual surge pela necessidade de reduzir os erros de segmentação que podem surgir na forma automática devido à assimetria natural entre as mamas e às variações anatômicas entre pessoas diferentes no corpo humano.

A técnica de segmentação escolhida e a sua correta implementação são fundamentais para a extração de recursos e influenciam no desempenho dos métodos de tomada de decisão. Porém, a segmentação de termogramas torna-se difícil pela ausência de limites nítidos de temperatura em alguns locais e por suas estruturas disformes.

A confiabilidade de um método de segmentação automática de imagens termográficas de mamas pode ser validada através da utilização de *ground-truth* ou imagem de referência. A imagem de um modelo referência representa, neste caso, a segmentação correta da região de interesse na imagem adquirida (YANG et al., 1995; CHABRIER et al., 2004). Devido à ausência de *ground-truths*, o objetivo de alguns métodos de validação é mostrar que os resultados da segmentação automática estão dentro da variabilidade entre as diferentes segmentações manuais (WEI et al., 2002; KLEIN, 2016).

A validação deve tanto avaliar a precisão e as limitações do algoritmo quanto deixar claro se o método pode ser aplicado. Na área de processamento de imagens médicas, a validação completa de um método é um passo necessário quando se tem por objetivo aplicá-lo à rotina médica.

A avaliação da eficácia de algoritmos de segmentação pode ser feita comparando os resultados deles com uma imagem manualmente segmentada como referência, através de métricas como os coeficientes de similaridade Dice e Jaccard, bem como os índices de Falso Positivo e de Falso Negativo.

Estas métricas são baseadas no conceito de sobreposição das mamas, as quais avaliam o quão próximo o resultado da segmentação automática (S) está do que se pode considerar a demarcação do *ground-truth* (G). Para isso, é feita uma comparação de cada pixel da segmentação ao seu correspondente no *ground-truth*. O Quadro 3 apresenta as fórmulas para estas métricas.

Quadro 3 - Equações das métricas de análise de segmentação

Métricas	Equações
Índice Jaccard (IJ)	$IJ(G, S) = \frac{ G \cap S }{ G \cup S } \quad (6)$
Índice Dice (ID)	$ID(G, S) = \frac{2 G \cap S }{ G  +  S } = \frac{2 \times J(G, S)}{1 + J(G, S)} \quad (7)$
Falso Positivo (FP)	$FP(G, S) = \frac{ S  -  G \cap S }{ G } \quad (8)$
Falso Negativo (FN)	$FN(G, S) = \frac{ G  -  G \cap S }{ G } \quad (9)$

Fonte: A Autora (2022)

A taxa de Falso Positivo (FP) indica a quantidade de pixels segmentados como ROI, mas que não foram selecionados no *ground-truth*. A taxa de Falso Negativo (FN) indica o número de pixels que foram detectados no fundo da imagem segmentada, mas que fazem parte da ROI no *ground-truth*.

O coeficiente de similaridade de Dice é considerado a métrica mais utilizada na validação dos resultados da segmentação médica (HOSSAM; HARB; KADER, 2018). Assim como o coeficiente de similaridade de Jaccard, ele fornece um valor relativo à sobreposição espacial ou à similaridade entre dois conjuntos de *pixels* (DICE, 1945). Os valores dos coeficientes variam entre zero e um: resultados mais próximos de um significam que a segmentação é similar ao seu *ground-truth*; os resultados que tendem a zero são os que possuem diferenças significativas entre a segmentação automatizada e a imagem de referência correspondente (LEÎMATRE & YONG, 2010 apud XIE & BOVIK, 2013).

### 3.4 EXTRAÇÃO DE ATRIBUTOS

Formalmente, uma base de dados pode ser representada por uma matriz de  $\mathbf{X}_n^d$ , em que  $n$  é o número de dados ou objetos e  $d$  é o número de atributos ou características de cada objeto. O valor de  $d$  define a dimensionalidade dos dados. Cada elemento  $x_{ij}$  dessa matriz contém o valor da  $j$ -ésima característica para o  $i$ -ésimo dado (COPPIN, 2017).

Para a detecção de uma doença ou um achado suspeito em um número considerável de pacientes, as características devem ser simples de extrair, requerer a mínima intervenção do usuário e contribuir significativamente para sensibilidade do sistema (DOUGHERTY, 2009).

Os tipos básicos de características que podem ser extraídos das imagens são: escala de cinza e de cores, valores de temperatura, textura, conjunto de pontos, contornos ou curvas, superfícies e regiões especificadas (OLIVEIRA, 2015).

Atualmente, há uma constante geração e armazenamento de grande quantidade de dados. Apesar da evolução tecnológica dos processadores dos computadores, a tarefa de manipular estes conjuntos de dados ainda é um problema. Nesse contexto, surgiram os objetos simbólicos, também chamados de variáveis ou dados simbólicos, que tem como objetivo estender os métodos de análise de dados usuais (como por exemplo, métodos exploratórios e gráficos) para que possam tratar os dados mais complexos e de maior número (BOCK & DIDAY, 1999).

As variáveis simbólicas podem ser expressas, considerando um único indivíduo, através de conjuntos de intervalos, de histogramas, de categorias, dentre outros. Os tipos de variáveis simbólicas mais comuns são: multivaloradas, intervalares e modais (BOCK & DIDAY, 1999).

Considerando que  $O$  seja um conjunto de objetos simbólicos, uma variável simbólica  $Y$  é considerada intervalar se  $\forall k \in O$ , o subconjunto  $Y(k) = [a; b]$ , onde  $a \leq b$ . Por exemplo, seja  $O$  um grupo de jovens e  $Y$  o tempo semanal para lazer (em horas), para os indivíduos  $i, j \in E$  é possível ter:  $Y(i) = [3; 5]$  e  $Y(j) = [7; 9]$  (ARAÚJO; LIMA; SOUZA, 2014).

No estudo utilizando imagens termográficas, é possível definir um intervalo baseado nas temperaturas contidas em uma ROI como sendo  $Y = [T_{mín}; T_{máx}]$ , onde  $T_{mín}$  e  $T_{máx}$  são, nesta ordem, os limites inferior e superior do intervalo de temperatura. E representam a temperatura mínima e máxima da região de interesse, respectivamente (ARAÚJO; LIMA; SOUZA, 2014).

### 3.5 PRÉ-PROCESSAMENTO DE DADOS

As técnicas de pré-processamento de dados são usadas para melhorar a qualidade do conjunto de dados, minimizando ou eliminando problemas como diferentes características, dimensões ou formatos. O pré-processamento pode adequar os dados para sua utilização por um determinado algoritmo de aprendizagem de máquina. Além disso, pode gerar modelos mais fiéis à distribuição real da base de dados, reduzindo a sua complexidade computacional e possibilitando ajustes de parâmetros mais fáceis e rápidos (BECKMANN, 2010).

#### 3.5.1 Dados desbalanceados

Um conjunto de dados reais, geralmente, apresenta a particularidade de possuir uma sub-representação do número de amostras de uma determinada classe em comparação com outras classes. Uma base de dados desbalanceada compromete substancialmente o processo de aprendizado, visto que os algoritmos clássicos de aprendizagem de máquina esperam uma distribuição equilibrada de classes ou um custo igual de classificação incorreta (FERNANDEZ et al., 2013; ROUTH; KUHOO; MALLICK, 2018).

Quando um conjunto de dados é desequilibrado, os classificadores convencionais na maioria das vezes favorecem a classe majoritária, classificando incorretamente as observações minoritárias e resultando na perda de desempenho. O impacto do fator de desequilíbrio também pode ser observado na etapa de treinamento, na qual a classe minoritária pode não ser detectada

(NEKOOEIMEHR, 2016). A maioria dos pesquisadores tem voltado sua atenção a problemas de desbalanceamento entre classes binárias (HE & GARCIA, 2009; LEMAÎTRE; NOGUEIRA; ARIDAS, 2017), porém, quando há várias classes no conjunto de dados desbalanceado, as soluções para o problema binário podem não ter implementação direta nem êxito nos resultados.

Apesar de não existir um único método que funcione bem para todo os problemas de amostra com dados desbalanceados, os métodos de amostragem (em inglês, *sampling*) mostram potencial à medida que tentam melhorar o conjunto de dados em si, previamente à etapa de classificação. Os métodos de amostragem alteram a distribuição das instâncias das classes através da sobreamostragem (em inglês, *oversampling*) que replica exemplos da classe minoritária para obtenção de uma distribuição balanceada, e da subamostragem (do inglês, *undersampling*), que tem como objetivo balancear o conjunto de dados através da eliminação de instâncias majoritárias (LARA et al., 2017).

De forma geral, os métodos de subamostragem apresentam-se menos eficientes do que os métodos de sobreamostragem. O problema desse método encontra-se na possibilidade de remoção de instâncias majoritárias que carregam informações importantes da amostra, especialmente nos casos em que o conjunto de dados é pequeno (ZHOU, 2013 apud NEKOOEIMEHR, 2016).

Nos casos de sobreamostragem aleatória, o excesso de amostragem gera instâncias artificiais muito semelhantes aos originais, o que aumenta a probabilidade de ocorrer um sobreajuste (em inglês, *overfitting*). Neste caso, o modelo gerado na etapa de treinamento se ajusta muito bem ao conjunto de dados inicial, porém se mostra ineficiente na previsão de novos casos.

### *Métodos de sobreamostragem*

Para apresentação clara e objetiva dos métodos de sobreamostragem, foram adotadas algumas notações (HE & GARCIA, 2009) que serão utilizadas nesta seção.

Considerando a base de dados de treinamento  $B_{treino}$  com  $m$  instâncias tem-se que  $B_{treino} = \{(x_i, y_i), i = 1, \dots, m\}$ , onde  $x_i \in X$  é uma instância de dimensão  $n$  do espaço de características  $X = \{f_1, f_2, \dots, f_n\}$  e  $y_i \in Y = \{(1, \dots, C)\}$  é um rótulo da classe associada à instância  $x_i$ . Para os métodos que serão discutidos, será considerado  $C = 2$  representando o problema de classificação binária.

Além disso, define-se  $B_{treino,min} \subset B_{treino}$  e  $B_{treino,maj} \subset B_{treino}$ , onde  $B_{treino,min}$  são os exemplos da classe minoritária em  $B_{treino}$  e  $B_{treino,maj}$  são os exemplos da classe

majoritária em  $B_{treino}$ , de modo que  $B_{treino,min} \cap B_{treino,maj} = \{\emptyset\}$  e  $B_{treino,min} \cup B_{treino,maj} = \{B_{treino}\}$ .

### *Synthetic Minority Oversampling Technique (SMOTE)*

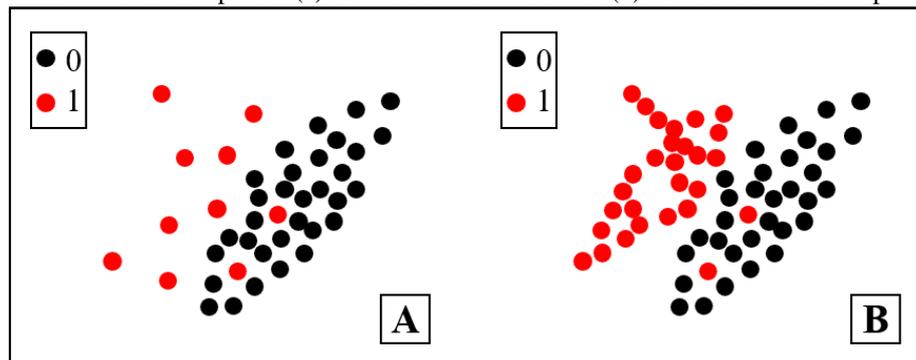
No que diz respeito a amostragem sintética, a técnica SMOTE (CHAWLA et al.,2012) é uma das mais conhecidas, tendo demonstrado sucesso em várias aplicações e servindo como base para outros métodos de sobreamostragem. O algoritmo SMOTE cria dados artificiais por meio de interpolação entre os exemplos minoritários.

O método cria instâncias sintéticas baseadas em uma instância minoritária e seus  $k$  vizinhos mais próximos. Um valor sintético é gerado com base em  $x_i$  e  $\hat{x}_i$ , sendo  $\hat{x}_i$  selecionado aleatoriamente entre os  $k$  vizinhos mais próximos de  $x_i$ , e em  $\delta$ , um número aleatório entre 0 e 1.

Um exemplo  $x_i$  é escolhido e o vetor obtido pela diferença entre o elemento em consideração e seu vizinho  $\hat{x}_i$  é multiplicado por  $\delta$ . O novo vetor é aplicado sobre o elemento em consideração, criando um ponto no espaço que fica na reta entre o elemento e seu vizinho. Este ponto representa a nova instância sintética  $x_{sintética}$  (Equação 10) da classe minoritária. O processo é repetido para cada elemento  $x_i$  do conjunto  $S_{min}$ , sendo finalizado ao atingir o número de elementos que se deseja gerar.

$$x_{sintético} = x_i + (\hat{x}_i - x_i) \times \delta \quad (10)$$

Figura 15 - Gráficos de dispersão (a) Amostra desbalanceada e (b) Amostra balanceada por SMOTE



Fonte: A Autora (2022)

A Figura 15 mostra uma distribuição binária de dados desbalanceados, onde as cores vermelha e preta representam exemplos minoritários e majoritários, respectivamente. As instâncias sintéticas aumentam o conjunto de dados original, melhorando significativamente o aprendizado (SÁEZ; KRAWCZYK; WOŹNIAK, 2016). Todavia, por desconsiderar a vizinhança entre as classes e replicar modelos muito específicos, pode gerar aumento da

sobreposição entre classes e problemas de sobreajustamento (HE & GARCIA, 2009; BECKMANN, 2010; LARA et al., 2017).

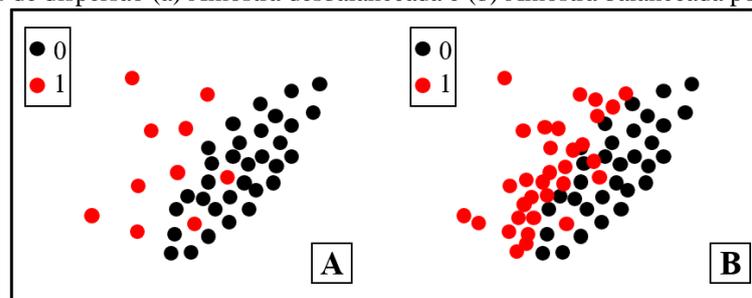
### *Borderline-SMOTE*

O algoritmo *Borderline-SMOTE* (HAN et al., 2005) é uma adaptação do SMOTE, desenvolvida para evitar ou minimizar os pontos fracos no algoritmo anterior. Pelo fato de os exemplos na fronteira ou próximos a ela serem mais suscetíveis à classificação incorreta (HAN et al., 2005), o objetivo do método é identificar os exemplos minoritários próximos à superfície de decisão e utilizá-los para gerar novos exemplos sintéticos (Figura 16).

O algoritmo classificará cada elemento  $x_i$  da amostra  $B_{treino}$  como sendo ruidoso – todos os vizinhos mais próximos diferem da classe de  $x_i$ , *danger* – pelo menos metade dos vizinhos mais próximos são da mesma classe de  $x_i$  e *safe* – todos os vizinhos mais próximos são da mesma classe de  $x_i$ . A geração de dados artificiais ocorre de forma similar ao SMOTE (interpolação), mas apenas as instâncias rotuladas como *danger* são sobreamostradas utilizando seus vizinhos mais próximos (HAN et al., 2005).

A imagem abaixo mostra que aqueles exemplos distantes da borda de decisão não foram considerados na geração de exemplos sintéticos. Isso inclui tanto os exemplos que são mais fáceis de classificar e os que são mais difíceis, dada à sobreposição de classe.

Figura 16 - Gráficos de dispersão (a) Amostra desbalanceada e (b) Amostra balanceada por *Borderline-SMOTE*



Fonte: A Autora (2022)

### *Amostragem Sintética Adaptativa - ADASYN*

O algoritmo ADASYN (HE et al., 2008) é baseado no SMOTE e busca minimizar o viés introduzido pelo desbalanceamento da amostra, deslocando de forma adaptativa a fronteira de decisão. O método pondera a quantidade de exemplos sintéticos gerados pela dificuldade no aprendizado dos exemplos da classe minoritária (HE & GARCIA, 2009). Para isso, primeiramente calcula-se a quantidade de dados artificiais ( $n_{x,sintética}$ ) que precisam ser gerados para a classe minoritária (Equação 11).

$$n_{x,sintética} = (|B_{treino,maj}| - |B_{treino,min}|) \times \beta \quad (11)$$

onde  $\beta \in [0,1]$  é um parâmetro que especifica o nível de equilíbrio desejado após a criação dos dados sintéticos.

Em seguida, para cada exemplo  $x_i \in B_{treino,min}$ , encontra-se os  $k$  vizinhos mais próximos, de acordo com a distribuição euclidiana, e calcula-se a distribuição de densidade  $\Gamma_i$  (Equação 12)

$$\Gamma_i = \frac{\Delta_i/k}{Z} \quad (12)$$

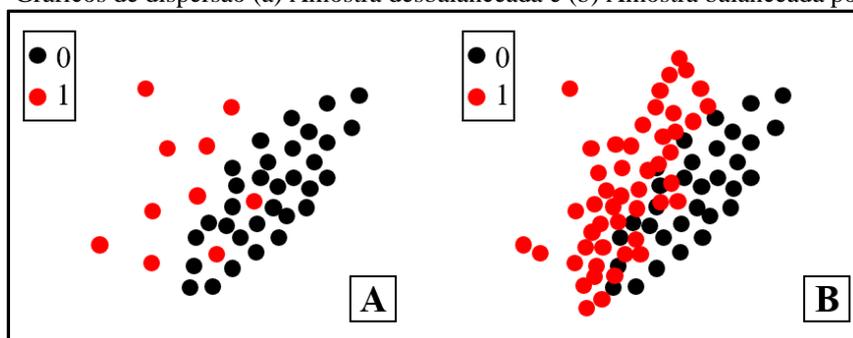
onde  $\Delta_i$  é o número de exemplos nos  $k$ -vizinhos de  $x_i$  que pertencem a  $B_{treino,maj}$ , e  $Z$  é uma constante de normalização para que  $\sum \Gamma_i = 1$ . Agora, é possível determinar o número de exemplos sintéticos  $g_i$  que precisam ser gerados para cada  $x_i \in B_{treino,min}$  (Equação 13).

$$g_i = \Gamma_i \times n_{x,sintética} \quad (13)$$

Finalmente, para cada exemplo da classe minoritária, é criado um  $g_i$  de acordo com a Equação 10. A ideia principal do algoritmo ADASYN é usar a distribuição de densidade  $\Gamma_i$  como critério para decidir, automaticamente, o número de exemplos sintéticos que serão criados para cada exemplo da classe minoritária (LARA et al., 2017). Assim, quanto maior o número de exemplos próximos entre si da classe majoritária, mais instâncias sintéticas serão criadas (Figura 17).

Ao contrário do método *Borderline*, observa-se que o foco do algoritmo ADASYN é, em grande parte, nos exemplos minoritários sobrepostos à classe majoritária. Nos casos em que esses exemplos de baixa densidade apresentarem valores discrepantes, o desempenho do modelo pode ser afetado pelo fato de a abordagem ADASYN visar o preenchimento das áreas entre eles.

Figura 17 - Gráficos de dispersão (a) Amostra desbalanceada e (b) Amostra balanceada por ADASYN



Fonte: A Autora (2022)

### *Cluster Based Oversampling (CBO)*

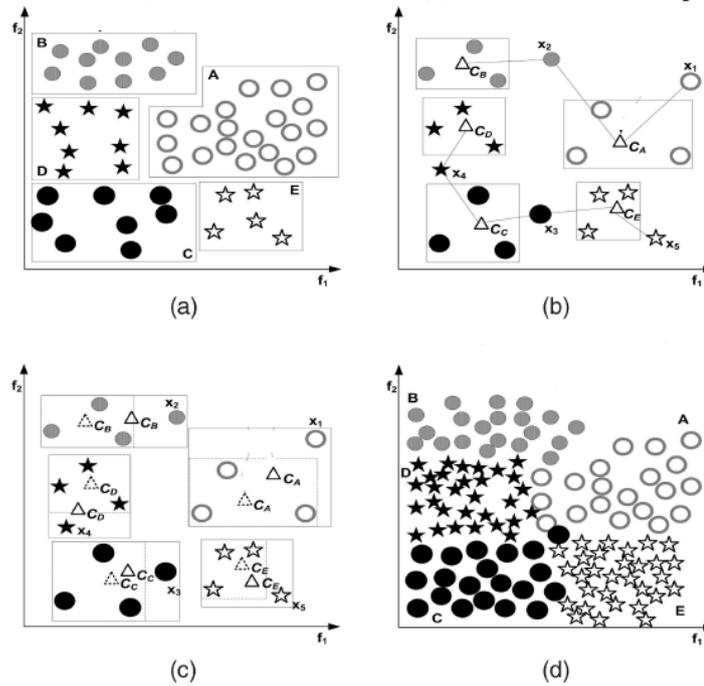
O método CBO (JO & JAPKOWICZ, 2004), em português “Sobreamostragem baseada em agrupamento”, foi proposto para lidar com problema de desbalanceamento entre classes e dentro da classe. O desbalanceamento entre classes se refere à desproporção da quantidade de exemplos entre cada classe, já o desequilíbrio dentro da classe está relacionado à desproporção do número de exemplos entre os subconjuntos que formam a classe (BARELLA, 2016).

O algoritmo *k-means* é utilizado como técnica de agrupamento. Nele,  $k$  exemplos são escolhidos aleatoriamente como representante de cada grupo. Os demais exemplos de treinamento são selecionados um por vez e, para cada exemplo, é calculada a distância euclidiana entre ele e cada centro de *cluster*. Cada exemplo de treinamento será atribuído ao *cluster* que exibe a menor distância. Por fim, os representantes dos *clusters* são atualizados (média dos exemplos do grupo) e o processo é repetido até que todos os exemplos se esgotem.

Para ilustrar as etapas do algoritmo, considera-se uma amostra com classes binárias distribuída de forma desbalanceada no espaço de características. A Figura 18a mostra que a classe majoritária foi dividida em três *clusters* (A, B e C) com 20, 10 e 8 exemplos, respectivamente. E a classe minoritária foi dividida em dois *clusters* (D e E) com 8 e 5 exemplos cada. A Figura 18b mostra a média do *cluster* (representada pelos triângulos) para três exemplos aleatórios ( $k = 3$ ) de cada *cluster*. Nesta figura, também apresenta o vetor de distância para cinco exemplos introduzidos individualmente ( $x_1, x_2, x_3, x_4$  e  $x_5$ ). As bordas e os centros dos agrupamentos são atualizados pela introdução dos cinco exemplos (Figura 18c).

Quando todos os exemplos forem analisados, o algoritmo CBO preenche os *clusters* de classe majoritária até que tenham o mesmo tamanho que o maior (A), ou seja, os *clusters* B e C terão 20 exemplos cada. Neste exemplo, o número de exemplos da classe majoritária  $N_{CBO}$  após a sobreamostragem é igual a sessenta (60). Em seguida, a sobreamostragem é feita nos *clusters* de classe minoritária (D e E) até que cada um tenha um número total de  $\frac{60}{2} = 30$  exemplo. A Figura 18d mostra o conjunto sobreamostrado pelo método CBO.

Figura 18 - Etapas do algoritmo CBO (a) Amostra desbalanceada, (b) Vetor de distância entre exemplos e centros dos clusters, (c) Clusters atualizados e (d) Amostra balanceada por CBO



Fonte: He e Garcia (2009)

### Safe Level - SMOTE

O algoritmo SMOTE sintetiza aleatoriamente as instâncias minoritárias ao longo do segmento de reta que une uma instância minoritária ao seu vizinho mais próximo, gerando dados sintéticos em regiões sobrepostas. A técnica *Safe Level* – SMOTE leva em conta as instâncias minoritárias ao longo desse segmento de reta, considerando diferentes graus de peso ou níveis de segurança (BUNKHUMPORNPAT; SINAPIROMSARAN; LURSINSAP, 2013).

Para cada instância minoritária, uma instância entre os  $k$ -vizinhos mais próximos é escolhida aleatoriamente e seu nível de segurança ( $SL$  – *Safe Level*) é calculado através da Equação 14. Se o  $SL$  for próximo de zero, a instância minoritária é considerada um ruído. Se o  $SL$  for próximo de seu  $k$ -vizinho mais próximo, a instância é considerada segura.

$$SL = n^{\circ} \text{ de instâncias minoritárias na vizinhança} \quad (14)$$

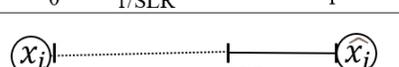
São calculados os níveis de segurança da instância minoritária e de cada vizinho mais próximo selecionado para, em seguida, se calcular uma taxa de nível de segurança (Equação 15). A taxa de nível de segurança  $TSL$  é usada para determinar uma região segura para a amostragem, de modo que quanto maior o seu valor, mais segura é a região.

$$TSL = \frac{\text{nível de segurança de } x_i}{\text{nível de segurança de } \hat{x}_i} \quad (15)$$

Há cinco casos correspondentes a uma taxa de nível de segurança ilustrados na Figura 19 (BUNKHUMPORNPAT; SINAPIROMSARAN; LURSINSAP, 2013). Na figura,  $x_i$  é

considerada uma instância minoritária,  $\hat{x}_i$  é um vizinho mais próximo dessa instância,  $SL_{x_i}$  é um nível seguro de  $x_i$  e  $SL_{\hat{x}_i}$  é um nível seguro de  $\hat{x}_i$ . O segmento de reta entre  $x_i$  e  $\hat{x}_i$  pode variar de 0 a 1.

Figura 19 - Casos correspondentes a taxas de nível de segurança

	1º caso $SL_{x_i} = SL_{\hat{x}_i} = 0$
	2º caso $SL_{x_i} \neq 0$ e $SL_{\hat{x}_i} = 0$
	3º caso $SL_{x_i} = SL_{\hat{x}_i} \neq 0$
	4º caso $SL_{x_i} > SL_{\hat{x}_i}$
	5º caso $SL_{x_i} < SL_{\hat{x}_i}$

Fonte: Bunkhumpornpat, Sinapiromsaran e Lursinsap (2013)

No primeiro caso, as instâncias  $x_i$  e  $\hat{x}_i$  são consideradas ruído, pois todos os  $k$  vizinhos mais próximos são instâncias majoritárias. Dessa forma, não há geração de dados sintéticos. O segundo caso mostra que apenas  $\hat{x}_i$  é uma instância de ruído, então o dado sintético é gerado longe dessa instância e pela duplicação de  $x_i$ .

No terceiro caso, tanto  $x_i$  quanto  $\hat{x}_i$  são seguros. Logo, uma instância sintética pode ser gerada em qualquer ponto do segmento de reta, através da metodologia SMOTE. Para o quarto caso, uma instância sintética é gerada mais próxima de  $x_i$  por ela ser mais segura, em relação a  $\hat{x}_i$ . Neste caso, a região de sobreamostragem se enquadra no segmento de reta  $\overline{0 (1/SLR)}$ .

O último caso mostra que uma instância sintética foi gerada mais próxima de  $\hat{x}_i$ , por ser mais segura em relação a  $x_i$ . A região de sobreamostragem se enquadra no segmento de reta  $\overline{(1 - 1/SLR) 1}$ .

O método *Safe Level-SMOTE* minimiza moderadamente o problema de sobreposição (PRATI; BATISTA; MONARD, 2004; BUNKHUMPORNPAT & SUBPAIBOONKIT, 2013), pois as instâncias sintéticas em regiões sobrepostas são criadas mais próximas da classe minoritária do que a classe majoritária, auxiliando no processo de aprendizagem para classificação. Porém, essa técnica ignora regiões densas com instâncias positivas, não sendo elas reconhecidas no processo de classificação.

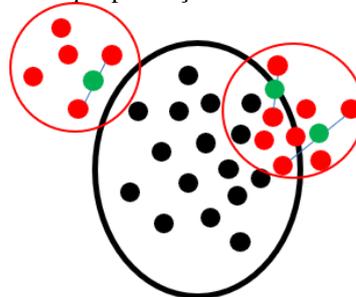
### *Adaptive Semi-Unsupervised Weighted Oversampling (ASUWO)*

A técnica A-SUWO (NEKOOEIMEHR & LAI-YUEN, 2016) foi desenvolvida para classificação de conjuntos desbalanceados de dados binários. O ASUWO agrupa as instâncias minoritárias e atribui pesos mais altos a cada *subcluster* mais próximo da classe majoritária. Essa abordagem permite a identificação de todas as instâncias próximas à borda de decisão e considera todos os *subclusters* para superamostragem (Figura 20).

O algoritmo consiste em três etapas principais: agrupamento não-supervisionado, dimensionamento adaptável de *subclusters* e geração de exemplos sintéticos. Na primeira etapa as instâncias minoritárias são agrupadas através da abordagem de agrupamento hierárquico semi-supervisionado, que constrói iterativamente *subclusters* minoritários evitando *subclusters* majoritários em seu meio.

Em seguida, o dimensionamento de cada *subcluster* minoritário é determinado com base no erro de classificação e na validação cruzada. O subconjunto com maior erro de classificação terá um tamanho maior, enquanto os subconjuntos com menor erro terão um tamanho menor.

Figura 20 - Abordagem de sobreamostragem entre instâncias selecionadas e seus 4 vizinhos mais próximos, considerando que pertençam ao mesmo *cluster*



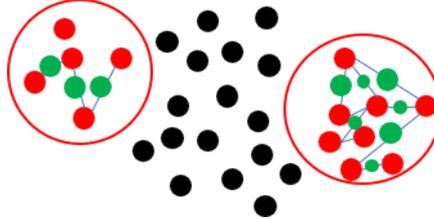
Fonte: A Autora (2022)

Finalmente, na etapa de geração de instâncias sintéticas, são atribuídos pesos às instâncias minoritárias com base na sua distância euclidiana média em relação aos  $k$ -vizinhos mais próximos da classe majoritária. A razão para atribuir pesos a essas instâncias está no fato de que exemplos minoritários mais próximos dos majoritários são mais propensos ao erro de classificação, logo são as mais importantes para o processo. Um dos vizinhos mais próximos  $\hat{x}_i$  da instância  $x_i$  é selecionado aleatoriamente, desde que pertença ao mesmo subconjunto, e uma nova instância  $x_{sintética}$  é gerada (Equação 16).

$$x_{sintética} = \beta x_i + (1 - \beta) \hat{x}_i \quad (16)$$

onde  $\beta$  é um número aleatório entre 0 e 1. No conjunto sobreamostrado (Figura 21), os *subclusters* minoritários possuem instâncias sintéticas geradas com maior proximidade das instâncias majoritárias.

Figura 21 - Criação de instâncias sintéticas, geradas mais próximas das instâncias majoritárias



Fonte: A Autora (2022)

### 3.5.2 Transformação de atributos numéricos

Diversas técnicas de aprendizado de máquina são influenciadas pelo tipo ou pelos valores dos dados que compõem uma base de dados. Dentre as técnicas que abordam esse problema, a transformação de atributos numéricos não altera o tipo de atributo, mas pode envolver, por exemplo, mudança de escala ou do intervalo de valores. A transformação é geralmente realizada para evitar que um atributo predomine sobre o outro, de forma errônea.

A normalização de dados é um tipo de transformações bastante utilizada, e é recomendável em casos de diferença considerável entre os limites inferior e superior dos atributos, o que leva a uma grande variação (HAN et al., 2004; GAMA NETO, 2018). As variáveis podem ser normalizadas segundo a amplitude ou segundo a distribuição, ambas variando de acordo com as funções de normalização utilizadas.

A normalização por reescala, conhecida como normalização Min-Max, realiza uma transformação linear dos dados adequando-os a uma dada escala (FACELI et al., 2011). Para isso, são definidos os valores mínimo (mín) e máximo (max) para os novos valores de cada atributo. A Equação 17 é utilizada para cálculo da observação normalizada  $v_{novo}$ . Para que os limites superior e inferior da nova escala sejam 1 e 0, respectivamente, deve-se considerar  $\max = 1$  e  $\mín = 0$ .

$$v_{novo,i} = \mín + \frac{x_i - \mín(x)}{\max(x) - \mín(x)} \quad (17)$$

onde  $x = (x_1, x_2, \dots, x_n)$ ,  $v_{novo,i}$  é a  $i$ -ésima observação normalizada e os valores  $\min(x)$  e  $\max(x)$  correspondem, respectivamente, ao menor e maior valor do atributo.

A normalização por padronização, ou *Z-score*, adiciona ou subtrai uma medida de localização para cada valor do atributo a ser normalizado e o valor resultante é multiplicado ou dividido por uma medida de escala. Assim, diferentes instâncias podem ter limites diferentes, mas terão os mesmos valores para as medidas de escala e espalhamento (FACELI et al., 2011).

Considerando a média ( $\mu$ ) como medida de localização e o desvio padrão ( $\sigma$ ) como medida de escala, o *Z-score* (Equação 18) padroniza os dados através do cálculo da distância entre um atributo  $x_i$  e a média, em termos do desvio padrão. O novo conjunto de valores possuirá média 0 e desvio padrão 1, mantendo as propriedades de forma (assimetria e curtose) do conjunto original.

$$v_{novo,i} = \frac{x_i - \mu}{\sigma} \quad (18)$$

A normalização por distribuição altera a escala de valores de uma instância. A técnica *Scale*, por exemplo, altera a escala dos dados em função do seu desvio padrão. Já a técnica *Centro* adequa os valores para centro com média zero. A função de ordenação de valores pode ser aplicada nessa técnica, fazendo com que os atributos a serem normalizados sejam ordenados e seus valores numéricos são substituídos pela posição ocupada no *ranking*.

### 3.5.3 Redução de dados

Um número elevado de atributos em técnicas de aprendizagem de máquinas pode inviabilizar a geração de modelos de predição, devido às restrições de espaço em memória ou tempo de processamento (MACEDO, 2012). Se cada atributo for considerado uma coordenada em um espaço  $d$ -dimensional, em que  $d$  é o número de atributos, o hipervolume que representa esse espaço cresce exponencialmente com a adição de novos atributos (FACELI et al., 2019).

Uma estratégia para minimizar o impacto do problema da quantidade de atributos é combinar ou eliminar partes irrelevantes do grupo. As técnicas de agregação, ao combinar grupos de atributos, substituem os valores originais por novos. As técnicas de seleção mantêm uma parte dos atributos originais e descartam os demais atributos (RESMINI et al., 2012; FACELI et al., 2019).

A seleção de características é uma opção para identificar quais são as mais significativas para uma posterior classificação, permitindo descartar aquelas que não contribuem para a distinção dos objetos entre as classes. A redução da dimensionalidade possibilita uma diminuição do custo computacional e de armazenamento, e um melhor entendimento dos dados e do modelo

criado para descrever os objetos (WEBB & COPSEY, 2010). Uma seleção de atributos pode ser feita por técnicas de ordenação ou por seleção, propriamente dita, de subconjunto.

As técnicas de ordenação (*ranking*) são metodologias simples de seleção, nas quais os atributos são ordenados segundo a sua relevância para a classificação de objetos em diferentes classes, por exemplo. A literatura expõe várias abordagens para o cálculo do escore associado ao *ranking*, dentre elas as mais usadas são as medidas de correlação e de distância entre vetores. Esse cálculo pode ou não considerar a informação sobre a classe (RESMINI, 2016).

O processo de seleção de um subconjunto de atributos é computacionalmente mais custoso que o *ranking*. Esse problema torna-se mais evidente com o crescimento do número de atributos, fazendo da seleção de subconjuntos uma abordagem intratável. As técnicas baseadas em *wrapper*, uso de algoritmo de aprendizado para análise do subconjunto de atributos na redução da taxa de erro de classificação, são utilizadas na avaliação da relevância do atributo.

A seleção de um subconjunto de atributos pode ser vista como um problema de busca, em que cada ponto no espaço de busca é visto como um possível subconjunto de atributos (BLUM & LANGELY, 1997).

### 3.6 MODELOS PREDITIVOS

O aprendizado de máquina preditivo é uma função que gera um estimador para um conjunto de exemplos rotulados. Se o domínio dos rótulos (classes) é formado por valores nominais conhecidos, tem-se um problema de classificação (aprendizado de conceitos), e o estimador gerado é um classificador (FACELI et al., 2019).

A classificação trata-se da parte do processo em que são estabelecidos critérios que podem ser usados para distinguir e identificar as imagens em diferentes classes (RUSS, 2001 apud DOURADO, 2014).

Na construção de um classificador é fundamental a definição de um grupo de treino, utilizado no processo de aprendizagem de máquina. Neste processo, o classificador “aprende” através da amostra de treinamento, gerando a regra de classificação que será aplicada nos casos futuros (ARAÚJO, 2014).

O aprendizado pode ser supervisionado ou não-supervisionado. Ele é dito supervisionado quando as classes dos objetos do grupo de treino são conhecidas, e não supervisionado (*clustering*) quando as classes não são conhecidas e as mais adequadas devem ser identificadas (DOURADO, 2014).

A quantificação e a qualificação do modelo preditivo gerado pelo classificador são feitas por técnicas de avaliação como, por exemplo, a validação cruzada. Entre os métodos de validação cruzada existentes, os mais comuns são o *k-fold* e o *leave-one-out*. No método *k-fold* a amostra inicial de  $n$  indivíduos é dividida aleatoriamente em  $k$  subconjuntos, dos quais  $k - 1$  são usados para treinamento e um para teste. A rotina é repetida  $k$  vezes até que todos os subconjuntos sejam usados como grupo de teste (ARAÚJO, 2014; DOURADO, 2014).

A técnica *leave-one-out* é computacionalmente custosa e é equivalente ao método de validação cruzada *k-fold*, onde  $k = n$ . Neste caso, um indivíduo serve de teste e  $n - 1$  são usados na construção da regra de classificação. O método é repetido  $n$  vezes até que todos os elementos do conjunto de dados sejam utilizados como teste, a partir de  $n$  conjuntos distintos de treino (ARAÚJO, 2014).

A seguir serão apresentados os métodos de aprendizado de máquina baseados em distância (algoritmo kNN), métodos probabilísticos (Naive Bayes), métodos baseados em busca (árvores de decisão), métodos baseados em otimização (SVM) e a análise discriminante linear.

### 3.6.1 Algoritmo do k vizinho mais próximo

O modelo de predição do  $k$  vizinhos mais próximos (kNN, do inglês, *k-Nearest Neighbour*), por ser intuitivo e simples, é um dos métodos considerado como referência na literatura (WEBB & COPSEY, 2011). As predições são realizadas por meio da proximidade entre os dados: dados similares tendem a estar concentrados em uma mesma região no espaço e, de maneira alternativa, dados diferentes estarão distantes entre si.

Em problemas de classificação utilizando o kNN, é necessária a pré-especificação do número de vizinhos  $k$  do objeto teste  $x_t = \{x_1, x_2, \dots, x_d\}$ , onde  $d$  é o número de dimensões (atributos) dos objetos de um conjunto de dados; do conjunto de dados de treinamento e da métrica de distância. Na escolha de  $k$  é necessário determinar um valor grande o suficiente para reduzir a sensibilidade ao ruído e, ao mesmo tempo, pequeno o suficiente para que a vizinhança não se estenda ao domínio de outras classes (FACELI et al., 2019).

O grupo de treino  $B_{treino}$  é composto por pares  $(x_i, y_i)$ ,  $i = 1, \dots, m$ , onde  $x_i$  é a  $i$ -ésima instância de treinamento e  $y_i$  é o indicador de classe correspondente. As medidas de distância mais usuais são as baseadas na métrica de Minkowski, dada pela Equação 19, em que  $x_t^l$  e  $x_i^l$  são elementos dos vetores  $x_t$  e  $x_i$ , respectivamente, e correspondem aos valores da coordenada  $l$  (atributos).

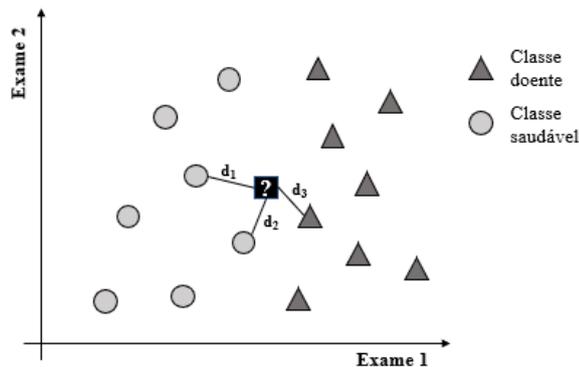
$$\text{distância}(x_t, x_i) = \sqrt[s]{\sum_{l=1}^d |x_t^l - x_i^l|^s} \quad (19)$$

Os valores de  $s$ , com  $1 \leq s < \infty$ , definem variações da métrica. Quando  $s = 2$ , por exemplo, a métrica de Minkowski é conhecida como distância Euclidiana (Equação 20), medida de distância mais popular e apropriada para os casos em que os componentes da base de dados possuem escalas semelhantes (WEBB & COPSEY, 2011).

$$\text{distância}(x_t, x_i) = \sqrt{\sum_{l=1}^d |x_t^l - x_i^l|^2} \quad (20)$$

Na Figura 22 é ilustrado um exemplo simples de um problema de classificação binário. Nesse exemplo, considera-se uma base de dados formada por indivíduos que podem ser classificados como saudáveis (círculos) ou doentes (triângulos), e o espaço de entrada é formado por dois atributos que representam o resultado de dois exames. Considerando o algoritmo  $k$ NN, com  $k = 3$  e medida de distância Euclidiana, o indivíduo teste (quadrado) seria classificado, de acordo com a votação majoritária dos seus vizinhos mais próximos, como saudável.

Figura 22 – Exemplo ilustrativo do algoritmo  $k$ -NN, para  $k = 3$



Fonte: A Autora (2022)

### 3.6.2 Naive Bayes

Uma forma alternativa de resolver problemas preditivos em aprendizagem de máquinas é através do uso de algoritmos baseados no teorema de Bayes ou métodos probabilísticos. Estes métodos admitem que a probabilidade de um evento  $A$  dado um evento  $B$  não depende apenas da relação entre os eventos, mas também de a probabilidade de  $A$  acontecer independentemente de  $B$ .

A probabilidade de o evento  $B$  acontecer pode ser estimada pela contabilização da frequência com que esse evento ocorre. Analogamente, é possível estimar a probabilidade de ocorrência de um evento  $B$ , para cada classe ou evento  $A$ , isto é,  $P(B|A)$ . No entanto, para se

calcular a probabilidade de um evento A acontecer, quando for observado um evento B,  $P(A|B)$ , é necessário aplicar o teorema de Bayes (WEBB & COPSEY, 2011)

O teorema de Bayes é usado para calcular a probabilidade *a posteriori* de um evento ou objeto pertencer a uma classe,  $P(A|B)$ , utilizando a probabilidade *a priori* da classe,  $P(A)$ , a probabilidade de observar vários objetos que pertencem à classe,  $P(B|A)$ , e a probabilidade de ocorrência desses objetos,  $P(B)$  (FACELI et al., 2019).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (21)$$

Em alguns cenários complexos, como o problema da dimensionalidade no aprendizado de classificadores, a probabilidade será analítica ou computacionalmente complexa ou intratável (WEBB & COPSEY, 2011).

Há diversas propostas relatadas na literatura para lidar com o problema de aprendizado em um cenário probabilístico (DUDA; HART; STORK, 2000). Uma estratégia é adotar hipóteses simplificadoras sobre a forma da função discriminante, utilizada para calcular as probabilidades condicionadas  $P(A|B)$ , levando a diferentes classificadores. Um exemplo desta estratégia é o classificador denominado Naive Bayes cuja função discriminante é expressa através da Equação 22.

$$P(y_k|\mathbf{x}) \propto \prod_{i=1}^d P(x_i|y_k) \quad (22)$$

onde  $P(y_k|\mathbf{x})$  denota a probabilidade de um exemplo  $\mathbf{x}$  pertencer à classe  $y_k$ , e  $x_i$  é o  $i$ -ésimo atributo do exemplo  $\mathbf{x}$ .

O classificador Naive Bayes adota a hipótese de que os valores dos atributos são independentes, condicionalmente à observação da classe (FACELI et al., 2019). Além disso, obedece a regra de decisão dada pela expressão:

$$\psi(\mathbf{x}) = \arg \max_k P(y_k|\mathbf{x}) \quad (23)$$

na qual  $\arg \max_k$  retorna a classe  $y_k$  com maior probabilidade de estar associada a  $\mathbf{x}$ , que é aquela que possui valor máximo para a  $P(y_k|\mathbf{x})$ .

Sendo assim, a fórmula do classificador Naive Bayes pode ser apresentada em uma forma aditiva por

$$\psi(\mathbf{x}) = \arg \max_k \sum_{i=1}^d \log P(x_i|y_k) + \log P(y_k) \quad (24)$$

Na prática, a Equação (3.24), também chamada de regra de Bayes ingênua, ignora interações entre os atributos, tanto na etapa de aprendizado quanto no seu uso efetivo como classificador (DUDA; HART; STORK, 2000; FACELI et al., 2019).

Ainda assim, esse método é bastante efetivo, apresentando bom desempenho em uma grande variedade de domínios (DOMINGOS, PAZZANI, 1997 apud FACELI et al., 2019) e robustez à presença de ruídos e atributos irrelevantes (KONONENKO, 1991 apud WEBB & COPSEY, 2011).

### 3.6.3 Árvore de decisão

Uma árvore de classificação ou árvore de decisão é um exemplo de processo de decisão que usa a estratégia dividir para conquistar. O conjunto de atributos é dividido em diferentes subconjuntos, aos quais recursivamente é aplicada a mesma estratégia. Forma-se, assim, agrupamentos para os quais a variável resposta apresenta comportamento homogêneo dentro desses grupos e heterogêneo entre eles (WEBB & COPSEY 2011; DANTAS & DONADIA, 2013; FACELI et al., 2019).

A árvore de decisão é um grafo acíclico direcionado representada como uma hierarquia de nós e arestas (Figura 23). Ela é formada, basicamente, pelo nó raiz ou inicial, nó intermediário ou interno e pelo nó folha ou final.

Os nós raiz e interno podem, também, ser denominados de nós não-terminais. E associado a cada um destes está uma das características (atributos). Um nó não-terminal contém um teste condicional baseado nos valores das características e as arestas que saem dele representam um valor, ou um conjunto de valores, que essa característica pode assumir (DANTAS & DONADIA, 2013).

Figura 23 – Representação de uma árvore como um tipo de grafo



Fonte: A Autora (2022)

Ao nó folha é atribuída uma função e associado a ele está um rótulo de classe. No caso mais simples, a função é a constante que minimiza a função de custo. Dessa forma, assumindo a função de custo 0-1, essa constante é representada pela moda estatística (FACELI et al., 2019).

Uma árvore de classificação simples prevê classificações com base em duas classes  $x_1$  e  $x_2$ . A predição começa no nó raiz e, dependendo da função usada pelo nó, a árvore se direciona para os nós internos ou para os nós folhas. Não atingindo o nó folha (terminal), o método é repetido até alcançá-lo. Assim, a classe que rotula o nó folha é atribuída ao exemplo usado como teste.

Cada nó da árvore corresponde a uma região no espaço de características. As regiões estabelecidas pelas folhas são mutuamente excludentes. Desse modo, a interseção das regiões abrangidas por quaisquer duas folhas é vazia. Então, uma árvore de decisão compreende todo universo de atributos, podendo fazer predições para qualquer exemplo de entrada (FACELI et al., 2019).

Alguns algoritmos de indução foram desenvolvidos para construção de uma árvore de decisão, como: *Interactive Dichotomizer* – ID3 (QUINLAN, 1979 apud MAIA, GOMES, CHAGAS, 2017), *Classification and Regression Trees* – CART (BREIMAN et al., 1984 apud MAIA, GOMES, CHAGAS, 2017) e C4.5 (QUINLAN, 1993 apud MAIA, GOMES, CHAGAS, 2017). O algoritmo usado neste trabalho foi o CART.

Para tornar os subconjuntos de dados mais homogêneos em relação a variável resposta, o algoritmo CART divide a base de dados sempre em dois nós segundo alguma função de impureza, como a entropia. Ao invés de armazenar, para cada nó, apenas a característica que minimiza a função de impureza, CART armazena as características que geram uma divisão similar, ordenadas pelo critério de impureza (DANTAS & DONADIA, 2013; FACELI et al., 2019).

A expansão da árvore até os nós finais respeita um critério de parada pré-estabelecido que pode ser, por exemplo, a convergência em relação aos ganhos na medida de impureza ou um número mínimo de indivíduos em um nó folha. As partições que pouco contribuem no aprendizado do método, representando ruídos ou erros, são descartadas através do processo de poda.

A classificação dos nós finais pode acontecer de acordo com a frequência predominante da variável resposta no nó avaliado ou pode-se incorporar custos de má classificação (DANTAS & DONADIA, 2013). Quando se procede a classificação de uma instância, se nessa instância o valor da característica testada no nó da árvore é desconhecido, o CART irá procurar, na lista ordenada de características alternativas, o primeiro atributo cujo valor é conhecido na instância (FACELI et al., 2019).

### **3.6.4 Máquinas de vetores de suporte**

As máquinas de vetores de suporte (SVM - *Support Vector Machine*) consistem em uma técnica computacional de aprendizagem de máquina que é capaz de resolver problemas de classificação e de regressão, adquirindo a habilidade de generalização, com o aprendizado na etapa de treinamento. A essência do SVM é a construção de um hiperplano ótimo, de modo que ele possa separar diferentes classes de dados com elevada margem possível. A margem de um classificador é definida como a menor distância entre as variáveis do conjunto de treinamento e o hiperplano usado na separação desses dados em classes (LORENA & CARVALHO, 2007; BORCHARTT, 2013).

Para encontrar a superfície de decisão ótima, o algoritmo de treinamento do SVM tenta separar da melhor forma possível os pontos relacionados aos dados das classes analisadas. Os pontos mais próximos do limite entre as diferentes classes são selecionados, por serem mais significativos na solução do que os pontos mais distantes e por ajudar a definir a forma da melhor superfície de decisão (LORENA & CARVALHO, 2007).

Alguns casos não conseguem gerar um bom hiperplano para ser usado como critério de separação. Para suprir esta dificuldade, utilizam funções Kernel ou Núcleo. Estas funções são menos custosas que outros métodos e permitem estender as aplicações sobre praticamente qualquer estrutura de dados.

De forma geral, um kernel  $K$  é uma função que recebe, por exemplo, dois pontos  $x_i$  e  $x_j$  no espaço de entradas e computa o produto escalar (Equação 25) desses dados no espaço de características, a partir de uma função  $\varphi$  de mapeamento. Esta função possibilita a construção de máquinas de aprendizagem com diferentes tipos de superfícies de decisão não-linear no espaço de entrada (BORCHARTT, 2013).

$$K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle \quad (25)$$

Alguns dos kernels (Quadro 4) mais utilizados na prática são os polinomiais, os de função base radial (*radial basis function* – RBF) e os sigmodais. O kernel polinomial com  $d = 1$ ,  $\delta = 1$  e  $\kappa = 0$  também é denominado linear, e seu uso implica na não realização de um mapeamento de dados (FACELI et al., 2019).

Quadro 4 - Funções kernel mais comuns

Tipo de kernel	Função $K(x_i, x_j)$	Parâmetros
Polinomial	$(\delta(x_i \cdot x_j) + \kappa)^d$	$\delta, \kappa, d$

RBF	$\exp(-\sigma( x_i - x_j ^2))$	$\sigma$
Sigmoidal	$\tanh((\delta(x_i \cdot x_j) + \kappa))$	$\delta, \kappa$

Fonte: Adaptada de Faceli et al. (2019)

### 3.6.5 Análise discriminante linear

A análise discriminante é comumente usada em problemas de classificação. É uma técnica usada para caracterizar as diferenças entre grupos e, a partir da análise de diferenças entre eles, obter a classificação de novas instâncias alocando-as em um dos grupos existentes.

Supondo a existência de  $n$  diferentes grupos, com  $n \geq 2$ . Para cada grupo  $\Pi_j$ ,  $j = 1, \dots, n$ , há uma função densidade de probabilidade  $f_j$  em  $\mathbb{R}^p$ , onde  $p$  indica o número de preditores, tal que se um objeto pertence a  $\Pi_j$ , então ele possui uma função  $f_j$  (WEBB & COPSEY, 2011). É necessário definir uma função de densidade probabilidade, ou função discriminante, que indique a qual grupo  $\Pi_j$  uma determinada observação  $X$  é mais provável pertencer.

Assumindo que observações  $X = x_1, x_2, \dots, x_p$  são extraídas de uma distribuição normal multivariada, com um vetor de média específico para cada classe e uma matriz de covariância comum. A distribuição normal multivariada admite que cada classe segue uma distribuição normal unidimensional, com uma dada correlação entre cada par de classes (GARETH et al., 2013). Formalmente, a função densidade de probabilidade da normal multivariada é definida como

$$f(x) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right) \quad (26)$$

onde  $\mu$  é a média de  $X$  (um vetor com  $p$  componentes) e  $\text{Cov}(X) = \Sigma$  é a matriz de covariância  $p \times p$  de  $X$ .

No caso de multiclass, o classificador por análise discriminante linear (*Linear Discriminant Analysis* – LDA) assume que as observações na  $k$  – ésima classe são retiradas de uma distribuição normal multivariada, na qual o vetor média ( $\mu_k$ ) é específico da classe e matriz de covariância ( $\Sigma$ ) é comum a todas as classes  $K$  (WEBB & COPSEY, 2011; GARETH et al., 2013). Aplicando a função densidade para a  $k$  – ésima classe,  $f_k(X = x)$ , no Teorema de Bayes obtém-se que o classificador atribui uma observação  $X = x$  à classe para qual a função discriminante  $\delta_k(x)$  (Equação 27) é maior.

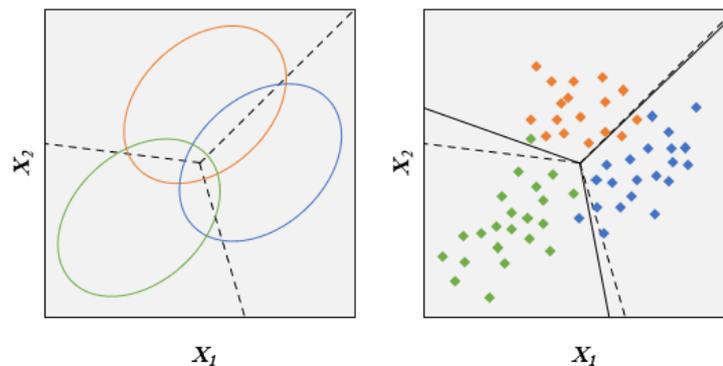
$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log \pi_k \quad (27)$$

Observa-se que  $\delta_k(x)$  é uma função linear de  $x$ . Ou seja, a regra de decisão da LDA depende apenas de  $x$  por meio de uma combinação linear de seus elementos. A função discriminante  $\delta_k(x)$  é usada para determinar o escore discriminante das observações avaliadas.

A Figura 31 ilustra um exemplo da metodologia de classificação por LDA. Para as três classes representadas, consideram-se as hipóteses de que as variáveis discriminantes possuem distribuição normal multivariada; as matrizes de covariância dos agrupamentos são iguais; os agrupamentos são do mesmo tamanho e apresentam diferenças quanto à média. As três elipses, representadas no quadro esquerdo da figura, caracterizam regiões que contêm 95% da probabilidade de cada uma das três classes (GARETH et al., 2013). As linhas tracejadas são os limites de decisão de Bayes.

Vinte observações tiradas de cada uma das três classes são exibidas no quadro esquerdo da Figura 24, e os limites de decisão LDA resultantes são mostradas como linhas pretas sólidas. Geralmente, os limites de decisão do LDA são muito próximos dos limites de decisão de Bayes, o que significa que a classificação por análise discriminante linear pode apresentar um bom desempenho (GARETH et al., 2013). Para atribuir uma nova instância  $X = x$ , estima-se os parâmetros desconhecidos de cada classe ( $\mu_1, \mu_2, \mu_3, \pi_1, \pi_2, \pi_3$  e  $\Sigma$ ), LDA insere essas estimativas na Equação 27 e classifica na classe para a qual  $\delta_k(x)$  é o maior.

Figura 24 – Exemplo LDA para três classes



Fonte: A Autora (2022)

### 3.7 AVALIAÇÃO DE MODELOS PREDITIVOS

A avaliação de desempenho de um classificador é um aspecto importante das etapas relacionadas ao aprendizado de máquina. De maneira geral, não é possível estabelecer,

previamente, uma técnica de aprendizagem de máquina que apresentará desempenho melhor na resolução de determinado tipo de problema (WEBB & COPSEY, 2011; FACELI et al., 2019).

O projeto de um classificador envolve ajustes de seus parâmetros livres, o que leva à obtenção de vários modelos para os mesmos dados. A experimentação é uma característica particular do domínio de aprendizagem de máquina. De fato, para qualquer modelo gerado, a realização de experimentos e a sua validação são importantes para demonstrar sua efetividade e reprodutibilidade (FACELI et al., 2019). Considerando modelos preditivos, a avaliação experimental pode ser feita através de medidas relacionadas ao desempenho obtido nas previsões realizadas.

### 3.7.1 Matriz de confusão

A matriz de confusão fornece a quantidade de erros e acertos segundo a classificação do objeto em análise. O número de acertos de cada classe encontra-se na diagonal principal, enquanto os demais elementos representam erros na classificação (VILA NOVA, 2017). Cada coluna da matriz descreve a distribuição real dos objetos e cada linha descreve as previsões do algoritmo para cada classe.

Por simplicidade, seja um problema binário com a Classe A denotando a classe positiva e a Classe B denotando a classe negativa. Tem-se então a matriz de confusão ilustrada no Quadro 5, em que

- VP corresponde ao número de verdadeiros positivos, ou seja, a quantidade de instâncias de classe positiva classificadas corretamente;
- VN corresponde ao número de verdadeiros negativos, ou seja, a quantidade de instâncias de classe negativa classificadas corretamente;
- FP corresponde ao número de falsos positivos, ou seja, a quantidade de instâncias pertencentes originalmente à classe negativa que foram incorretamente preditas como da classe positiva;
- FN corresponde ao número de falsos negativos, ou seja, a quantidade de instâncias cuja classe original é a positiva, mas que foram classificadas incorretamente como pertencente à classe negativa.

Quadro 5 - Matriz de confusão para um problema binário

		Classe verdadeira	
		Classe A	Classe B
Classe predita	Classe A	VP	FN

	Classe B	FP	VN
--	----------	----	----

Fonte: A Autora (2022)

A partir da matriz de confusão, outras medidas de desempenho podem ser derivadas (Quadro 6). Entre elas, tem-se: valor preditivo positivo (VPP), valor preditivo negativo (VPN), taxa de falsos positivos (TFP), taxa de falsos negativos (TFN), sensibilidade (SENS) e especificidade (ESP).

Quadro 6 - Demais medidas baseadas na matriz de confusão

		Condição Real		Prevalência
		Teste positivo	Teste negativo	
Previsto no teste	Teste positivo	VP	FN	VPP
	Teste negativo	FP	VN	VPN
Acurácia		SENS	TFP	
		TFN	ESP	

Fonte: A Autora (2022).

### *Acurácia*

A acurácia ou taxa de acerto (Equação 28) é medida em função dos valores verdadeiros positivos e negativos do teste. É calculada pela soma dos valores da diagonal principal da matriz de confusão, dividida pelo total de elementos da matriz ou total de casos da amostra ( $n$ ). Um teste com acurácia (ACC) alta significa que ele classificou corretamente grande parte das instâncias (RESMINI, 2016).

$$ACC = \frac{VP+VN}{n} = \frac{VP+VN}{VP+FP+FN+VN} \quad (28)$$

### *Sensibilidade e especificidade*

A avaliação de quão bom é um teste é feita, principalmente, através dos índices de sensibilidade e especificidade. A sensibilidade ou taxa de verdadeiro positivo (TVP) (Equação 29) corresponde à quantidade de acertos na classe positiva (FACELI et al., 2019). Ela representa a proporção de indivíduos com o problema em questão identificados corretamente pelo sistema, ou seja, informa o quão bom é o sistema em identificar tal problema.

Já a especificidade ou taxa de verdadeiro negativo (TVN) (Equação 30) corresponde à taxa de acerto na classe negativa (FACELI et al., 2019). É a proporção de indivíduos sem o problema que são apontados corretamente pelo sistema, isto é, informa o quão bom é um sistema em identificar o indivíduo sem o problema em questão.

O Quadro 7 abaixo faz uma síntese dos principais pontos acerca dos índices de sensibilidade e especificidade considerados pelos profissionais da área da saúde.

Quadro 7 - Aspectos importantes sobre os índices de sensibilidade e especificidade

	<b>Sensibilidade</b>	<b>Especificidade</b>
<b>Definição</b>	$\frac{VP}{VP+FN}$ (29)	$\frac{VN}{VN+FP}$ (30)
<b>Utilização</b>	Identifica pacientes doentes	Identifica pacientes saudáveis
<b>Pode ocasionar</b>	Falso Positivo	Falso Negativo
<b>Útil para</b>	Triagem	Confirmação de diagnóstico
<b>Melhor resultado</b>	Verdadeiro Positivo	Verdadeiro Negativo

Fonte: Adaptado de Soares e Siqueira (2002)

Observa-se, então, que um teste adequado de triagem é aquele que apresenta sensibilidade alta, fornecendo confiabilidade para uso em rastreamento (excluir doenças). Quanto mais sensível um teste, maior a segurança do especialista de que o indivíduo com resultado negativo não tem, de fato, a patologia.

#### *Prevalência e valores preditivos*

A prevalência do problema na população que está sendo testada é um aspecto relevante na predição de um classificador. É uma medida estatística usada para representar o quão comum ou raro é um determinado problema. A prevalência pode ser representada como a proporção entre os indivíduos de uma população que possuem o problema buscado e toda população analisada (RESMINI, 2016).

O Valor Preditivo Positivo (VPP) (Equação 31) representa a proporção de verdadeiros positivos entre todos os indivíduos com teste positivo. Antagonicamente, o Valor Preditivo Negativo (VPN) (Equação 32) é a proporção de verdadeiros negativos entre todos os indivíduos com teste negativo.

$$VPP = \frac{VP}{VP+FP} \quad (31)$$

$$VPN = \frac{VN}{VN+FN} \quad (32)$$

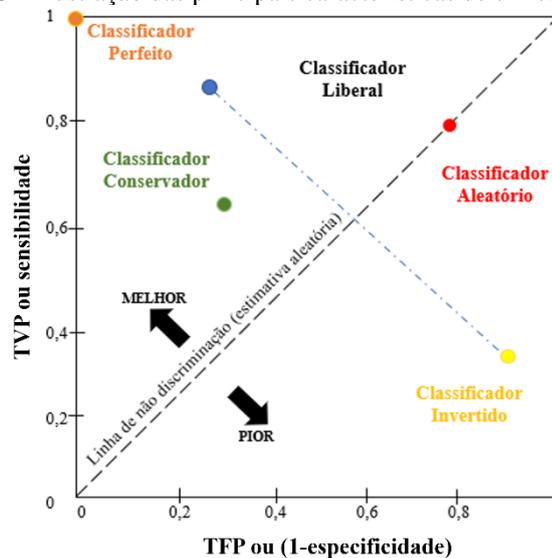
Em um mesmo teste, tem-se que quanto maior a prevalência, maior o VPP e menor o VPN. E, também, quanto mais sensível o teste, melhor o VPN; bem como, quanto mais específico o teste, melhor o VPP. Enquanto a sensibilidade e a especificidade do teste são inerentes ao teste e não variam, excluindo casos de erros técnicos, os valores preditivos dependem da prevalência do problema na população analisada (RESMINI, 2016).

### 3.7.2 Análise ROC

A curva característica de operação do receptor (*Receiving Operating Characteristics – ROC*) é um gráfico bidimensional com eixos  $X$  e  $Y$  representando os índices de TFP (Taxa de Falsos Positivos) e TVP (Taxa de Verdadeiros Positivos), respectivamente. O desempenho de um classificador pode ser plotado nessa curva, equivalendo a um ponto no espaço bidimensional (FACELI et al., 2019).

A Figura 25 apresenta os principais aspectos de um espaço ROC. A linha diagonal tracejada representa classificadores que realizam previsões aleatórias. Classificadores abaixo dessa linha são considerados piores que os aleatórios. O ponto (0,1) representa classificações perfeitas, em que todos os exemplos são classificados corretamente. O ponto (1,0), por outro lado, representa o desempenho de um classificador cujas previsões foram incorretas. O ponto (1,1) representa classificações sempre positivas, e o ponto (0,0) classificações sempre negativas (GARETH et al., 2013; FACELI et al., 2019).

Figura 25 – Ilustração das principais características de um espaço ROC



Fonte: A Autora (2022)

Para se obter a acurácia de um modelo de classificação, a área sob a curva ROC (*AUC - Area Under ROC Curve*) é verificada. A medida AUC produz valores entre 0 e 1. De modo que, valores mais próximos de 1 u.a. (unidade de área) são considerados melhores, quando comparado a outro algoritmo. A proximidade da diagonal implica em uma área em torno de 0,5 u.a., indicando uma previsão aleatória conforme assinalado no Quadro 8.

Quadro 8 - Valores de referência para avaliação segundo a medida AUC

Valor AUC	Interpretação
1,0	Perfeito
$0,9 \leq AUC < 1,0$	Excelente
$0,8 \leq AUC < 0,9$	Bom
$0,7 \leq AUC < 0,8$	Razoável
$0,6 \leq AUC < 0,7$	Ruim
$0,5 \leq AUC < 0,6$	Péssimo
$AUC < 0,5$	Reprovado

Fonte: Gareth et al. (2013)

### 3.7.3 Teste de concordância Kappa

O teste de concordância *Kappa* ( $\kappa$ ), também conhecido por coeficiente *Kappa*, é uma métrica que compara a acurácia observada com a acurácia esperada. Ou seja, avalia o nível de concordância entre dois conjuntos de dados (SANTOS, 2017 apud VASCONCELOS, 2017). O coeficiente de *Kappa* leva em consideração todos os elementos da matriz de confusão, assumindo que as classes são independentes e mutuamente excludentes, e o classificador e os pontos de referência operam de maneira independente (COHEN, 1960; VASCONCELOS, 2017; VILA NOVA, 2017).

O cálculo do coeficiente *Kappa* é dado por:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (33)$$

onde  $P_o$  é a acurácia observada, isto é, a proporção de acertos realizada pelo classificador. E  $P_e$  representa a acurácia esperada ou acertos esperados. Estas variáveis podem ser calculadas pelas Equações 34 e 35.

$$P_o = \frac{VP + VN}{VP + FN + FP + VN} \quad (34)$$

$$P_e = \frac{[(VP + FN)(VP + FP)] + [(VN + FP)(VN + FN)]}{(VP + FN + FP + VN)^2} \quad (35)$$

Assim sendo, o índice *Kappa* é considerado uma medida de concordância que permite avaliar se a concordância está além do esperado e, também, avaliar o grau dessa concordância (BEN-DAVID, 2008).

O coeficiente  $\kappa$  apresenta valores entre  $(-P_e / 1 - P_e)$  e 1 (BEN-DAVID, 2008). Quanto mais próximo do valor unitário, maior é o indicativo de que existe uma acurácia significativa. Os valores próximos de zero indicam que a concordância é puramente aleatória. Valores negativos são raros e sugerem nenhuma concordância, com acurácia menor do que a esperada pelo acaso.

Landis e Koch (1977) sugerem uma interpretação (Quadro 9) que permite ao pesquisador avaliar se a medida obtida é ou não satisfatória. Os autores ressaltam que apesar dos intervalos serem arbitrários, eles podem fornecer uma referência útil para análise dos resultados.

Quadro 9 - Interpretação do índice de Kappa

Valores $\kappa$	Interpretação
< 0,00	Ausência de concordância
0,00 – 0,20	Concordância pobre
0,21 – 0,40	Concordância leve
0,41 – 0,60	Concordância moderada
0,61 – 0,80	Concordância substantiva
0,81 – 0,99	Concordância ótima
1,00	Concordância perfeita

Fonte: Landis e Koch (1977)

Cumprido destacar que o valor  $\kappa$  depende da prevalência do problema em estudo. Uma grande prevalência resulta em um alto nível de concordância esperada pelo acaso, o que resultará num valor de  $\kappa$  mais baixo. Da mesma forma, um problema de baixa prevalência resultará em valores de  $\kappa$  mais altos (SANTOS, H., 2015).

Embora as medidas relatadas ao longo desse tópico estejam relacionadas com um teste de classificação binária, é possível generalizar este conceito para um problema com mais de duas classes.

### 3.8 ALGORITMOS DE OTIMIZAÇÃO GLOBAL

Os algoritmos de otimização global procuram calcular e caracterizar as soluções cujos valores representam o máximo e o mínimo de um conjunto. Diferentemente das técnicas de otimização local, esses métodos não exigem o uso de um bom modelo inicial, mas apenas o conhecimento dos limites do espaço de busca, onde está incluída a solução. Neste tipo de estratégia, a resposta sintética do modelo é calculada iterativamente e é atualizada por processos específicos até que o dado sintético convirja para o dado observado (VANZELER, 1999). Os algoritmos de otimização global que serão descritos, a seguir, são o Algoritmo Genético (AG) e a Evolução Diferencial (ED).

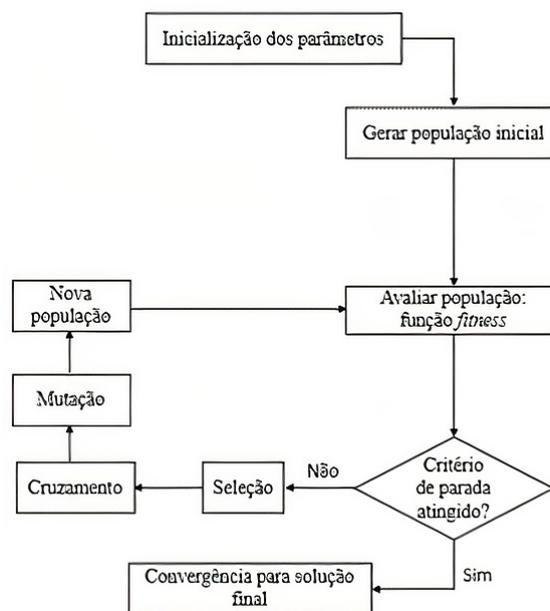
#### *Algoritmo Genético*

O Algoritmo Genético (Figura 34) é uma técnica de otimização inspirada no princípio da sobrevivência e reprodução dos indivíduos mais aptos, proposto por Darwin (LUCENA,

2013 apud VILA NOVA, 2017). Ele simula processos da evolução biológica de uma população de modelos gerados aleatoriamente. Os dados observados são comparados com os dados sintéticos, que compõem uma população aleatória, através de uma função-objetivo, denominada função de aptidão (função *fitness*). A população é atualizada com os melhores indivíduos até se atingir o ajuste ótimo. Tais atualizações são alcançadas através dos processos de codificação binária, seleção, cruzamento e mutação (VANZELER, 1999; VILA NOVA, 2017).

Em AGs, a solução é chamada de indivíduo ou cromossomo. Um cromossomo é uma estrutura de dados, frequentemente, representada por vetores binários (FACELI et al., 2019). Nesse caso, cada elemento (gene) representa a presença (valor binário 1) ou a ausência (valor binário 0) de alguma característica (FACELI et al., 2019). O tipo de valores (inteiro, real ou binário) depende da natureza do problema estudado.

Figura 26 – Fluxograma do método Algoritmo Genético



Fonte: Adaptado de Vila Nova (2017) e Falcone (2004)

No fluxograma (Figura 26), a população inicial é normalmente gerada pela escolha aleatória de indivíduos que representam uma solução do problema que se deseja resolver. Em seguida, tais indivíduos são avaliados através da função *fitness*. A função *fitness* mede a qualidade de cada solução. Ao longo do treinamento, é atribuído para cada indivíduo uma pontuação relacionada ao desempenho avaliado pela função de aptidão. Os indivíduos mais aptos (pais) são selecionados para originarem as próximas gerações (filhos), com base no valor de sua pontuação: quanto maior for, maior a probabilidade de serem escolhidos.

Há vários métodos para seleção dos indivíduos, sendo os mais comuns a seleção por torneio e a seleção por roleta, sendo esse último o adotado no presente trabalho.

*Seleção por roleta:* cada indivíduo é representado em uma roleta e sua fatia é proporcional ao seu índice de aptidão. A roleta é girada  $n$  vezes, onde  $n$  representa o número de indivíduos que se deseja selecionar, e a cada rodada sorteia-se um número entre 0 e 100 (correspondente à pontuação em porcentagem de cada indivíduo) ou entre 0 e 360 (correspondente a angulação na roleta). O indivíduo que possuir em seu intervalo o número sorteado, será selecionado (LINDEN, 2005).

Na reprodução são aplicados os operadores genéticos sobre os indivíduos selecionados, criando uma população. Os principais operadores são os de cruzamento (*crossover*) e de mutação (FACELI et al., 2019).

*Cruzamento:* é um processo aleatório aplicado a pares de indivíduos selecionados e ocorre com probabilidade fixa (taxa de cruzamento), que deve ser especificada pelo usuário. No cruzamento de um ponto, é definido um ponto de corte que divide cada pai em duas partes. Os filhos receberão uma parcela dos genes de cada um dos pais (Figura 27a) (LINDEN, 2005; FACELI et al., 2019).

*Mutação:* consiste na alteração de um ou mais genes do indivíduo. A aleatoriedade é definida de forma probabilística através da taxa de mutação. A operação de mutação permite a introdução e a manutenção da heterogeneidade de indivíduos na população (Figura 27b) (LINDEN, 2005; FACELI et al., 2019).

Figura 27 – Operadores genéticos



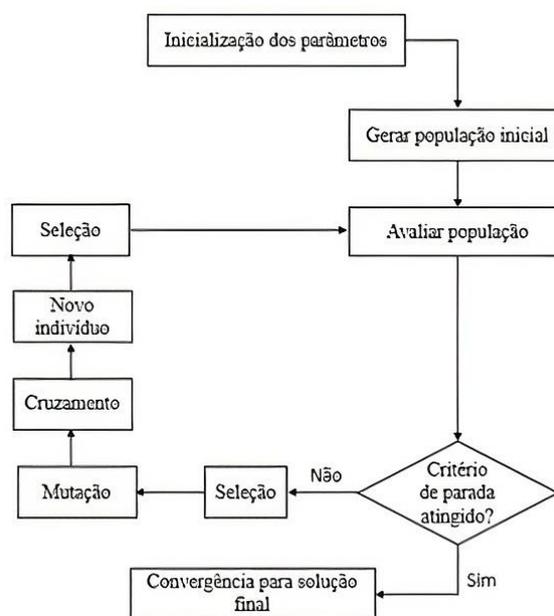
Fonte: Faceli et al. (2019)

A sequência de operações seleção-reprodução é repetida até que o critério de parada seja atingido. Diferentes critérios de parada podem ser utilizados, como a quantidade máxima de gerações (ou repetição do ciclo) ou enquanto não atingir uma solução satisfatória (VANZELER, 1999; LINDEN, 2006; FACELI et al., 2019).

### Evolução Diferencial

O Algoritmo de Evolução Diferencial (Figura 28) foi criado como uma alternativa para o Método de Algoritmo Genético, resolvendo de forma mais eficaz os principais problemas de otimização, agora em espaço contínuo. A Evolução Diferencial simula a evolução biológica semelhante ao Algoritmo Genético: há a manipulação dos indivíduos que representam as soluções candidatas e, no decorrer das gerações, essas soluções são modificadas através dos processos de mutação e cruzamento. Nesses processos são geradas novas possíveis soluções e, em seguida, é feita a seleção. O ciclo se repete até que um valor ideal seja encontrado (COLAÇO et al., 2006; BEZERRA, 2013; ARAUJO, 2016).

Figura 28 – Fluxograma do método Evolução Diferencial



Fonte: Adaptado de Falcone (2004)

O algoritmo cria uma população inicial com indivíduos escolhidos aleatoriamente. Cada indivíduo possui  $n$  componentes representados por valores reais, onde  $n$  é o número de variáveis de projeto. A população segue uma evolução natural, em que o número de indivíduos é constante em todas as gerações (ARAUJO, 2016). No decorrer das gerações, os indivíduos que representam as soluções candidatas sofrem alterações de mutação e cruzamento, pelas quais são geradas novas soluções candidatas.

**Mutação:** os indivíduos são alterados pela adição da diferença vetorial ponderada entre dois indivíduos aleatórios da população a um terceiro indivíduo (Equação 36) (BRANDÃO, 2014). Os novos indivíduos ( $v_{novo}$ ) são denotados vetores doadores ou modificados.

$$v_{novo} = v_{base} + F(v_1 - v_2), 0 < F < 2 \quad (36)$$

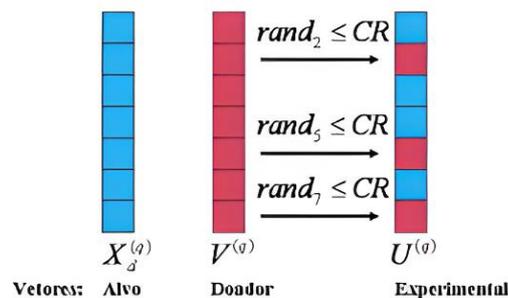
onde  $F$  é o fator de escala ou taxa de perturbação que determina a ponderação da diferença  $v_1 - v_2$ . A parcela do produto é usada para perturbar o terceiro vetor  $v_{base}$  ou o melhor vetor da população.

*Cruzamento*: usado para aumentar a diversidade dos indivíduos que sofreram mutação (PRICE, 1999 apud ARAUJO, 2016). Nesta operação, os componentes do vetor doador e do vetor alvo trocam atributos para formar o vetor modificado. O vetor alvo deve ser diferente daqueles já usados anteriormente e a sua escolha aleatória é dada segundo a probabilidade de cruzamento (CR), com CR compreendida entre 0 e 1 (BRANDÃO, 2014). O vetor tentativa ou experimental  $u_i^j$  é dado por

$$u_i^j = \begin{cases} v^j + F(v_{r1}^j - v_{r2}^j), & \text{se } r_i \leq CR. \\ v_i^j, & \text{caso contrário} \end{cases} \quad (37)$$

onde  $r_i$  é um número randômico entre 0 e 1 e  $v_i^j$  são as componentes do vetor alvo que competirá com o novo vetor gerado. A CR representa a probabilidade de o vetor experimental herdar os valores das variáveis do vetor doador, sendo CR um valor fornecido pelo usuário (BRANDÃO, 2014; ARAUJO, 2016). A Figura 29 representa o funcionamento do operador *crossover* dado pela Equação 37.

Figura 29 – Operador *crossover* para Evolução Diferencial



Fonte: Brandão (2014)

Se o vetor experimental apresentar um valor de aptidão, calculado através da função objetivo, menor que o do vetor alvo, então o vetor experimental substitui o vetor alvo na geração seguinte (ARAUJO, 2016). Caso contrário, o vetor alvo avança para a próxima geração. O processo iterativo continua até que algum critério de parada seja obedecido, sendo que um número máximo de gerações deve ser estabelecido.

#### **4 METODOLOGIA**

O presente trabalho tem como foco o desenvolvimento de ferramentas computacionais de CAD para análise de imagens termográficas de mamas que apresentem performance confiável, que sejam fáceis de usar, e que forneçam aos profissionais de saúde um suporte adicional para o diagnóstico do câncer mamário.

Os termogramas utilizados neste trabalho foram adquiridos no Ambulatório de Mastologia do Hospital das Clínicas da Universidade Federal de Pernambuco, seguindo o protocolo

descrito em Oliveira (2012). A câmera utilizada foi a FLIR S45 com resolução de 240x320 *pixels* e sensibilidade de 0,08 °C.

A amostra selecionada é composta por mulheres com diagnóstico concluído através de exames clínicos tradicionais, ultrassonografia, mamografia e biópsia quando necessária. Cada uma destas mulheres concordou em participar da pesquisa por meio da assinatura do Termo de Consentimento Livre e Esclarecido (TCLE), documento associado ao cadastro do projeto no Ministério de Saúde, sob o Registro CEP/CCS/UFPE N° 279/05 e aprovado pelo Comitê de Ética da UFPE.

As imagens por infravermelho têm formato *.jpeg*. Para a obtenção das matrizes de temperatura a partir das imagens termográficas, foi usado o *software* FLIR Tools que é disponibilizado pelo fornecedor da câmera. Neste programa, mesmo depois da aquisição das imagens e/ou quando for necessário, o usuário pode fazer ajustes no valor da emissividade, da temperatura refletida, da temperatura e da umidade relativa do ambiente e da distância entre a câmera e o objeto.

É possível exportar os valores de temperatura dos *pixels* para o programa Excel® da Microsoft. As planilhas são salvas no formato *.csv*, após trocar as vírgulas dos números decimais por pontos. As matrizes obtidas têm dimensão de 320x240 e são processadas como uma imagem térmica digital.

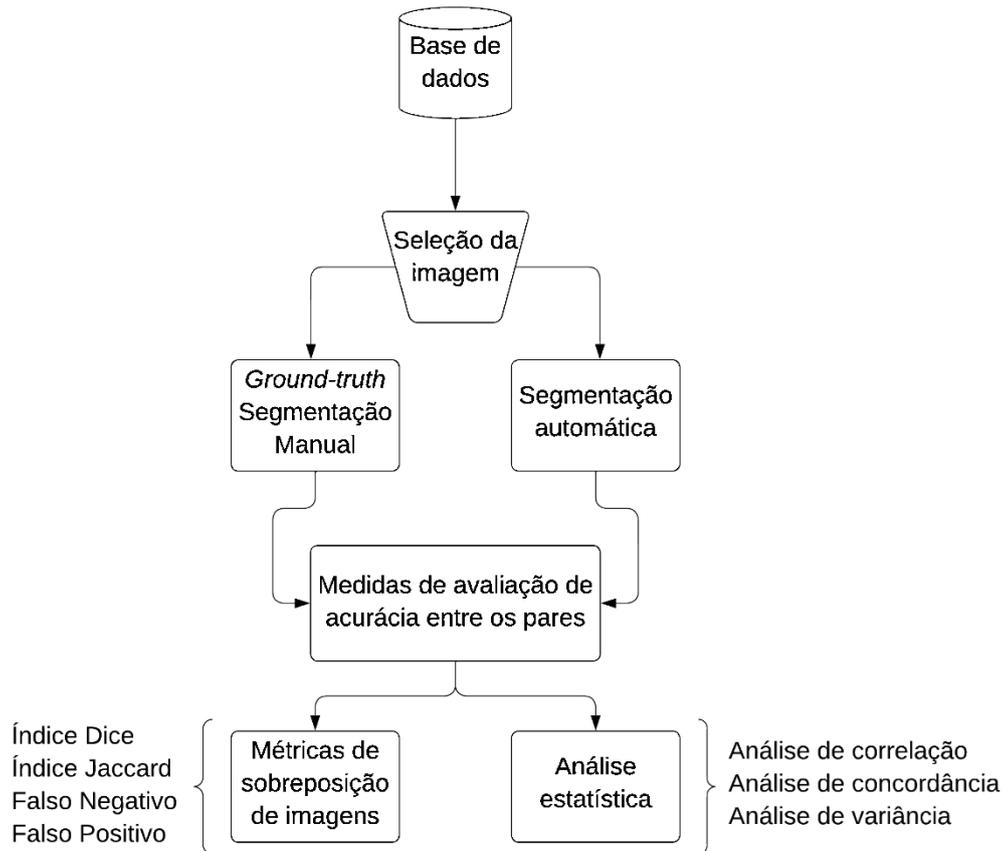
#### 4.1 O PROBLEMA DA VALIDAÇÃO DA SEGMENTAÇÃO AUTOMÁTICA

A segmentação automática da ROI, em termogramas de mamas, utilizada foi desenvolvida por Dourado (2014). O autor desenvolveu um método prático e de processamento rápido de extração de área mamária. Por esses motivos, a técnica foi adotada no presente trabalho.

Para garantir a confiabilidade dos resultados obtidos, o método de segmentação automática (DOURADO, 2014) foi avaliado, comparando-o com sua imagem de referência ou *ground-truth* correspondente. Essas imagens representam a segmentação de referência da ROI e foram obtidas manualmente através do método proposto por Araújo (2014). A validação do método automático foi feita de formas qualitativa e quantitativa, aplicando métricas de avaliação presentes na literatura.

A metodologia completa proposta para validação da segmentação automática é apresentada na Figura 30. Todos os processos descritos são aplicados a cada mama.

Figura 30 - Fluxograma da metodologia adotada para validação da segmentação automática



Fonte: A Autora (2022)

### *Base de dados*

Nesta etapa, foi considerada uma base de 90 imagens termográficas, incluindo pacientes com mamas de diferentes padrões anatômicos (pequena, média e grande) representando ao máximo a heterogeneidade encontrada em pacientes, permitindo assim ampliar o escopo de aplicação da técnica proposta por Dourado (2014).

### *Segmentação automática*

O estágio de segmentação visa a separação da região das mamas do restante da imagem, de forma totalmente automática. A segmentação automática reduz o tempo consumido se comparada com a segmentação manual, padroniza e torna os resultados repetíveis.

A matriz de temperaturas  $\mathbf{T}$  originalmente obtida a partir da imagem termográfica tem forma de uma planilha Excel<sup>®</sup> e foi reconstruída como uma imagem digital através do Matlab<sup>®</sup>.

Assim, atribuiu-se uma cor a cada valor de temperatura, e sobre essa nova imagem foi feita a segmentação.

No caso da segmentação automática, as matrizes de temperatura foram tratadas como imagens em tons de cinza, com os tons mais claros indicando as temperaturas mais altas, e o processo de extração da região de interesse ocorre sem nenhuma intervenção humana. A Figura 29 apresenta a imagem da paciente e o resultado desta segmentação. A metodologia de desenvolvimento da segmentação automática está descrita em Dourado (2014).

Figura 31 – Resultado da segmentação automática



(a) Imagem original

(b) Imagem segmentada

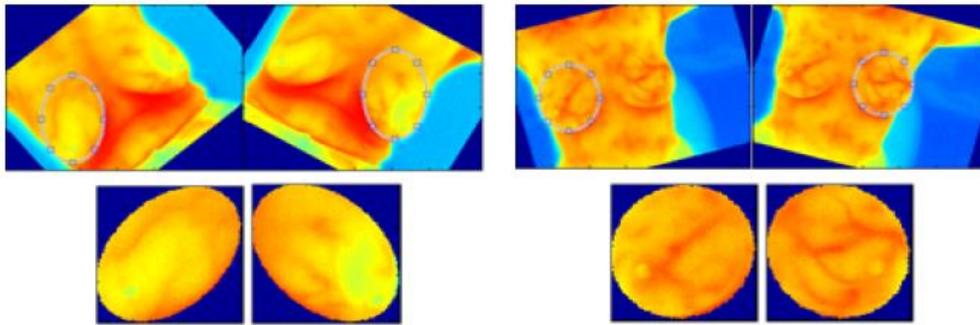
Fonte: Dourado (2014)

#### *Geração do ground-truth*

Para validar um método, é necessário obter imagens *ground-truth* que permitirão a medição da eficiência da segmentação automática proposta. Devido à ausência de *ground-truth* de imagens termográficas, foi adotado como padrão os resultados obtidos através da segmentação manual (ARAÚJO, 2014). É importante ressaltar, contudo, que mesmo utilizando mais de uma segmentação manual não se chega necessariamente a um *ground-truth* definitivo, em decorrência das variações existentes entre os resultados das diferentes segmentações.

Após o processamento da matriz de temperaturas  $\mathbf{T}$  na plataforma Matlab<sup>®</sup>, o usuário seleciona a região de interesse a partir de elementos elipsoidais, gerados de forma manual sobre a matriz  $\mathbf{T}$  e, após a seleção da área de interesse, despreza-se o restante da imagem. Ao final do processo, são obtidas duas matrizes de temperaturas independentes, uma referente à mama esquerda e outra à mama direita (Figura 32).

Figura 32 – Exemplo de resultados da segmentação manual



Fonte: Araújo (2014)

### *Métricas de sobreposição de imagens*

A avaliação da segmentação automática é feita comparando a similaridade entre as regiões extraídas pela técnica avaliada e seu *ground-truth* correspondente. Para a análise do desempenho da segmentação automática, foi desenvolvido um algoritmo capaz de extrair os coeficientes de similaridade Dice e Jaccard, bem como os índices de Falso Positivo e de Falso Negativo (Quadro 4).

### *Análise estatística*

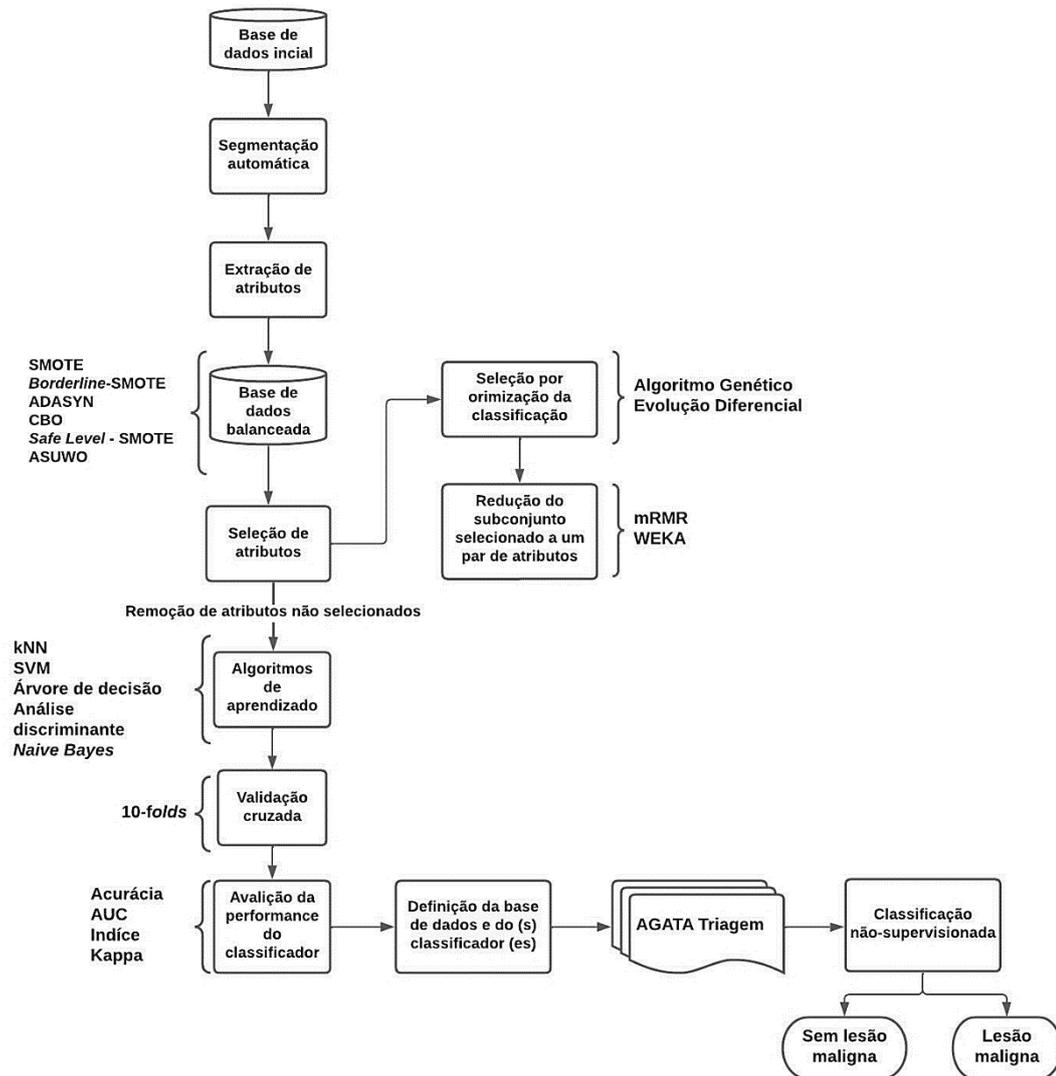
Os dados referentes às áreas segmentadas foram analisados no *software* Minitab<sup>®</sup> - versão 18. A aderência dos dados aos parâmetros de normalidade foi testada por meio do teste de Kolmogorov-Sminov. As variáveis com distribuição normal foram descritas como média, desvio padrão (DP) e seu intervalo de confiança. As variáveis com distribuição não-paramétrica foram descritas como mediana e intervalo interquartil. A concordância entre as áreas *ground-truth* e as obtidas pelo método de segmentação automática foi medida usando o gráfico de Bland-Altman (HIRAKATA & CAMEY, 2009), e o teste de correlação de Pearson. O objetivo foi verificar se as áreas são equivalentes e se um método poderia substituir o outro. Os limites de concordância do gráfico de Bland-Altman foram calculados para um intervalo de confiança de 95%, definidos com diferença média de  $\pm 1,96DP$ .

A influência do padrão anatômico das mamas na segmentação automática foi avaliada pela análise de variância de Kruskal-Wallis. Um valor de  $p < 0,05$  foi considerado para apresentar significância estatística. Todos os valores de  $p$  foram bicaudais.

## 4.2 ELABORAÇÃO DO SISTEMA DE CLASSIFICAÇÃO BINÁRIA

A metodologia proposta para elaboração do sistema CAD de classificação binária é apresentada na Figura 33.

Figura 33 - Fluxograma da metodologia adotada para desenvolvimento do sistema de classificação binária

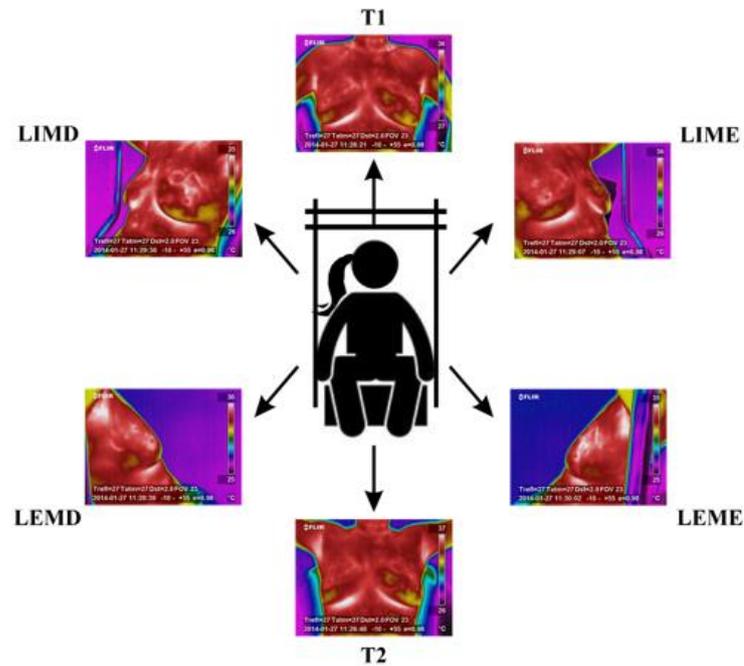


Fonte: A Autora (2022)

#### 4.2.1 Base de dados

As pacientes voluntárias do projeto foram submetidas à aquisição de imagens termográficas em, basicamente, duas posições diferentes: frontal e lateral. Cada uma delas envolve uma série padronizada da seguinte maneira: T1 (frontal com as mãos na cintura); T2 (frontal com as mãos levantadas segurando uma barra localizada acima da cabeça); T2 com grade; Lateral interna da mama direita (LIMD); Lateral interna da mama esquerda (LIME); Lateral externa da mama direita (LEMD) e Lateral externa da mama esquerda (LEME). A Figura 34 apresenta exemplo de cada uma das posições citadas.

Figura 34 – Exemplos das posições frontais e laterais realizadas na aquisição das imagens termográficas



Fonte: A Autora (2022)

A amostra utilizada neste trabalho é composta, inicialmente, por 233 imagens termográficas, do tipo T2, de pacientes sem lesão maligna e com lesão maligna. A distribuição diagnóstica das imagens utilizadas é apresentada na Tabela 1.

Tabela 1 – Amostra desbalanceada utilizada na classificação binária

		Análise binária			
		Sem lesão maligna		Com lesão maligna	
		Frequência	Percentual	Frequência	Percentual
Análise multi-classe	<b>Lesão maligna</b>	-	-	43	18,45%
	<b>Lesão benigna</b>	83	35,62%	-	-
	<b>Cisto</b>	43	18,45%	-	-
	<b>Saudável</b>	64	27,47%	-	-
	<b>TOTAL</b>	190	81,54%	43	18,45%

Fonte: A Autora (2022)

O desbalanceamento entre as classes da base de dados pode interferir no processo de classificação, tornando o classificador mais tendencioso para a classe majoritária (sem lesão maligna). No entanto, a classe minoritária (com lesão maligna) possui seu grau de importância no quesito do diagnóstico do câncer mamário, visto que uma pessoa com câncer ser diagnosticada como saudável ou sem lesão maligna pode interferir no prognóstico médico.

Para lidar com o problema do desbalanceamento entre as classes, foram adotadas estratégias de ajuste a nível de dados, adicionando instâncias à classe minoritária. Nesse sentido,

foram utilizadas técnicas de sobreamostragem definidas pelos métodos SMOTE, *Borderline-SMOTE*, ADASYN, CBO, *Safe Level – SMOTE* e *ASUWO* (ver Seção 3.5.1).

#### 4.2.2 Extração das características

Foram utilizadas as características apresentadas no Quadro 10 e descritas detalhadamente no Apêndice E. A escolha dessas características foi baseada majoritariamente nos trabalhos de (ACHARYA et al., 2012; BORCHARTT, 2012; DOURADO, 2014; ARAÚJO et al., 2016; MISHRA & RATH, 2020).

Com intuito de se determinar o conjunto mais apropriado para representar as regiões segmentadas (mama direita e mama esquerda), foram extraídas um total de 29 características. A partir da seleção dos atributos mais adequados à base de dados utilizada, foram testadas diversas combinações de diferentes características e cada uma delas serviu como entrada no processo de classificação.

A seguir, cada uma das características usada é descrita resumidamente. Para o cálculo delas, utilizou-se, em sua grande maioria, funções já disponibilizadas pelo *software* Matlab®.

As características  $c_{11}$  e  $c_{18}$  fazem parte de um conjunto que tiveram um critério de robustez aplicado (DOURADO, 2014). O critério de robustez consiste em considerar *pixels* que apresentam temperatura maior/menor ou igual aos *pixels* vizinhos que não estão localizados na borda da região segmentada. O conjunto de atributos foi definido para verificar se o critério de robustez utilizado na busca dos máximos e mínimos de temperatura interfere ou não nas medidas de cada característica.

Neste trabalho, foram utilizados os momentos centrais de segunda, terceira e quarta ordem, os quais caracterizam, respectivamente, a dispersão, a assimetria e a curtose de uma variável aleatória (MORETO, 2008).

Quadro 10 – Lista de características extraídas dos termogramas de mamas

Grupo	Característica	Definição
Medidas estatísticas	$c_1$	$\max(T_{ME,MAX}, T_{MD,MAX})$
	$c_2$	$\min(T_{ME,MAX}, T_{MD,MAX})$
	$c_3$	$\max(T_{ME,MIN}, T_{MD,MIN})$
	$c_4$	$\min(T_{ME,MIN}, T_{MD,MIN})$
	$c_5$	$c_1 - c_2$

	$c_6$	$c_2 - c_3$
	$c_7$	$c_3 - c_4$
	$c_8$	$c_1 - c_4$
	$c_9$	Média ( $\mu$ )
	$c_{10}$	Desvio padrão ( $\sigma$ )
	$c_{11}$	$\max(T_{ME,MAX}, T_{MD,MAX})_R$
	$c_{12}$	$\min(T_{ME,MAX}, T_{MD,MAX})_R$
	$c_{13}$	$\max(T_{ME,MIN}, T_{MD,MIN})_R$
	$c_{14}$	$\min(T_{ME,MIN}, T_{MD,MIN})_R$
	$c_{15}$	$c_{11} - c_{12}$
	$c_{16}$	$c_{12} - c_{13}$
	$c_{17}$	$c_{13} - c_{14}$
	$c_{18}$	$c_{11} - c_{14}$
	$c_{19}$	Obliquidade
	$c_{20}$	Curtose
	$c_{21}$	Mediana
	$c_{22}$	Entropia ( $H$ )
Medidas estatísticas de segunda ordem	$c_{23}$	Contraste ( <i>con</i> )
	$c_{24}$	Correlação ( <i>corr</i> )
	$c_{25}$	Homogeneidade ( <i>homog</i> )
	$c_{26}$	Energia ( <i>energ</i> )
Momentos estatísticos centrais	$c_{27}$	Momento 2
	$c_{28}$	Momento 3
	$c_{29}$	Momento 4

Fonte: A Autora (2022)

\*ME = mama esquerda; MD = mama direita; R = atributo com critério de robustez;  $T_{ME,MAX}$  = temperatura máxima da mama esquerda;  $T_{MD,MAX}$  = temperatura máxima da mama direita;  $T_{ME,MIN}$  = temperatura mínima da mama esquerda;  $T_{MD,MIN}$  = temperatura mínima da mama direita.

#### 4.3 SOBREAMOSTRAGEM POR GERAÇÃO DE DADOS SINTÉTICOS

O desbalanceamento da base de dados em relação à classe minoritária (pacientes com lesão mamária maligna) pode interferir negativamente no processo de classificação. Para minimizar este problema, após a extração de características, foi adotada a estratégia de sobreamostragem a fim de balancear a amostra utilizada na classificação binária de termogramas mamários.

Os métodos de sobreamostragem são conceitualmente simples e intuitivos, dado que se fundamentam, basicamente, na alteração da estrutura do conjunto de dados a partir da adição de amostras sintéticas. Além disso, como a base de imagens termográficas de mamas com diagnóstico concluído é reduzida, é preferível aumentar a quantidade de exemplos da classe minoritária, evitando possíveis perdas de informações discriminativas que podem ocorrer ao excluir imagens da base de dados.

Neste trabalho, foram utilizadas as seis técnicas de sobreamostragem por geração de dados sintéticos mais comuns descritas na literatura: SMOTE, *Borderline-SMOTE*, ADASYN, CBO, *Safe Level-SMOTE* e ASUWO (ver Seção 3.5.1). O objetivo foi verificar qual técnica melhor se adaptaria às particularidades do conjunto de dados para, com isso, gerar melhores resultados.

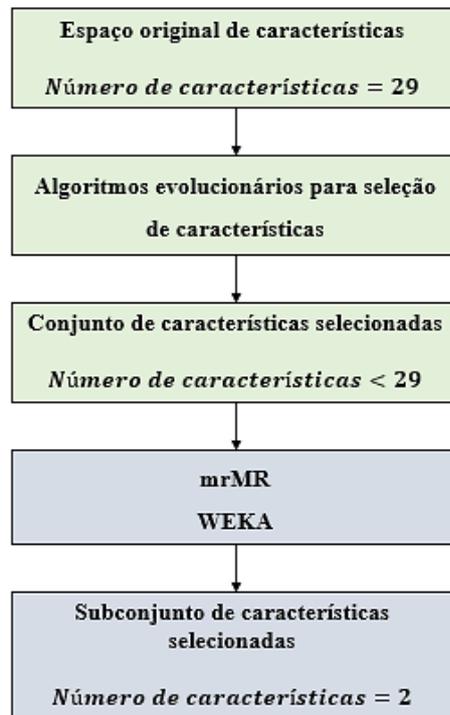
#### 4.4 SELEÇÃO DE CARACTERÍSTICAS

A utilização de recursos de alta dimensão pode reduzir o desempenho dos sistemas de classificação, uma vez que alguns atributos podem ser redundantes e não informativos (BRUZZONE & PERSELLO, 2008). Para encontrar a melhor precisão no processo de classificação, pode-se combinar diferentes conjuntos de atributos até encontrar aquela que seja a ideal. A seleção de características é o processo de obtenção de um subconjunto de atributos significativos para uso na construção de modelos de aprendizado de máquina.

Foram aplicados quatro métodos para a seleção das características extraídas dos termostogramas das mamas esquerda e direita: Algoritmo Genético (AG), Evolução Diferencial (ED), Mínima Redundância e Máxima Relevância (mRMR) e a biblioteca WEKA (*Waikato Environment for Knowledge Analysis*). A Figura 35 ilustra, de forma resumida, a metodologia adotada para seleção de características.

Os algoritmos evolucionários, AG e ED, foram utilizados para otimizar o resultado de classificação, a partir da seleção do subconjunto de características que minimizaram o erro de cada classificador analisado. A técnica mRMR e o WEKA foram usados para redução da dimensionalidade do conjunto selecionado pelo AG e ED, possibilitando a comparação entre os resultados obtidos anteriormente pelo grupo de pesquisa.

Figura 35 – Fluxograma da metodologia aplicada à seleção de características, onde a cor verde representa a seleção por otimização da classificação e a cor cinza representa a redução do subconjunto selecionado a um par de atributos



Fonte: A Autora (2022)

#### 4.4.1 Seleção de características baseada no Algoritmo Genético

O Algoritmo Genético possui uma configuração padrão indicada no Quadro 11. A população inicial, para este trabalho, foi randômica e é uma matriz de dimensão “*Tamanho da população × Comprimento do genoma*”. O tamanho da população equivale a quantidade inicial de características extraídas e o comprimento do genoma é o número de genes (bits) em cada cromossomo.

Quadro 11 – Parâmetros adotados na configuração do AG

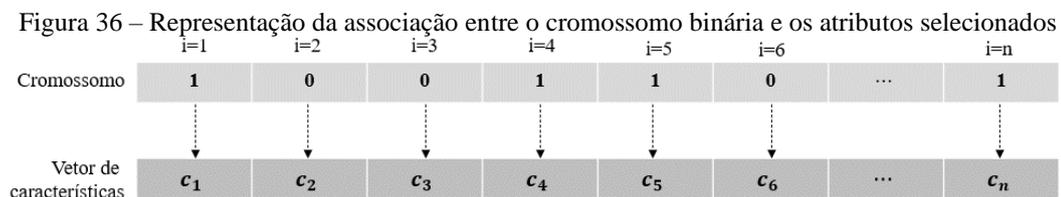
Parâmetros	Valores
População	29
Geração	População × 100
Taxa de cruzamento	0,8
Taxa de mutação	0,01
Operador de seleção	Roleta

Fonte: A Autora (2022)

A aptidão de cada cromossomo na população atual foi avaliada através de uma função *fitness* que levou em consideração o erro de classificação de um determinado algoritmo de aprendizado de máquina. Foram utilizados cinco algoritmos de aprendizado de máquina para

avaliação dos indivíduos (conjunto de combinação de recursos). Os classificadores utilizados na seleção de atributos são os mesmos da etapa seguinte.

A codificação do indivíduo é feita através de um cromossomo binário (*string* binário) de tamanho igual ao número de características. A cada geração do AG, o cromossomo binário especifica as características que serão consideradas no processo de classificação. Para isso, cada posição  $i$  do cromossomo indica se o atributo foi selecionado (valor 1) ou não selecionado (valor 0). A Figura 36 exemplifica a relação entre o cromossomo binário e as características selecionadas. Nela observa-se que as características  $c_2$ ,  $c_3$  e  $c_6$  estão associadas ao valor 0 no cromossomo, logo elas não seriam consideradas para obtenção do erro de classificação (função *fitness*).



Fonte: A Autora (2022)

Os indivíduos que formaram as novas gerações foram selecionados aleatoriamente pelo método da roleta. Na reprodução ou criação da nova população foram usados os operadores genéticos *crossover* e mutação.

A fração de cruzamento utilizada neste trabalho foi de 0,8, o que indica que 80% da população da próxima geração foi formada por indivíduos criados por cruzamento. Foi utilizado um ponto de *crossover* em uma posição  $p$  escolhida de forma aleatória. Dessa forma, por exemplo, o filho F1 recebe os genes do pai P1 até o ponto  $p$  e os genes do pai P2 do ponto  $p$  em diante. Já o filho F2, ao contrário, recebe os genes do pai P2 até o ponto  $p$  e do pai P1 recebe os genes do ponto  $p$  em diante.

Antes de a nova população ser finalizada, após o processo de *crossover* os genes dos indivíduos passaram por mutações implementadas randomicamente. A população utilizada na geração seguinte foi, portanto, formada pelos indivíduos da geração atual que apresentaram melhores valores de aptidão, sobrevivendo automaticamente até a nova geração, e pelos indivíduos que passaram pelas técnicas de cruzamento e mutação. Por fim, o critério de parada adotado foi número de evoluções da população.

#### 4.4.2 Seleção de características baseada na Evolução Diferencial

A evolução diferencial possui uma configuração padrão indicada no Quadro 12. A metodologia de geração da população inicial, da análise da aptidão (função *fitness*) e a codificação do indivíduo (subconjunto de características) para determinação dos atributos selecionados foi análoga àquela utilizada no AG.

Quadro 12 – Parâmetros adotados na configuração da ED

Parâmetros	Valores
População	29
Geração	População × 100
Taxa de cruzamento	0,8

Fonte: A Autora (2022)

Os estágios de mutação e cruzamento foram aplicados aos genes dos indivíduos para criação de um cromossomo teste binário  $t_{id}$  que foi usado na etapa de seleção. Nesta etapa,  $i$  representa a ordem do cromossomo na população e  $d$  é a dimensão do cromossomo que equivale à quantidade de características.

Na mutação, para um determinado cromossomo binário  $x_{id}$  da população, três cromossomos  $x_p$ ,  $x_q$  e  $x_r$  são selecionados aleatoriamente de modo que  $p \neq q \neq r \neq i$ . Foi calculado um cromossomo binário  $v_{id}$ , baseado nas diferenças entre  $x_p$  e  $x_q$  (Equação 38) (TOO et al., 2019).

$$v_{id} = \begin{cases} 0, & \text{se } x_p = x_q \\ x_p, & \text{caso contrário} \end{cases} \quad (38)$$

Em seguida, o cromossomo com o operador mutação incluído foi calculado conforme apresentado na Equação (39).

$$m_{id} = \begin{cases} 1, & \text{se } v_{id} = 1 \\ x_r, & \text{caso contrário} \end{cases} \quad (39)$$

No *crossover*, o cromossomo teste  $t_{id}$  foi gerado da seguinte forma:

$$t_{id} = \begin{cases} m_{id}, & \text{se } \delta \leq CR \text{ ou } d = d_{\text{randômico}} \\ x_{id}, & \text{caso contrário} \end{cases} \quad (40)$$

onde  $CR \in (0,1)$  é a taxa de cruzamento,  $d_{\text{randômico}} \in (0, N)$  é um número selecionado aleatoriamente e  $\delta$  é um número randômico distribuído entre 0 e 1. Neste trabalho,  $N = 29$  (número de características) e  $CR = 0,8$ .

A seleção do indivíduo para a próxima geração foi feita através da análise da função de aptidão. Quando o valor da função de aptidão de  $t_{id}$  foi melhor, o indivíduo atual  $x_{id}$  foi substituído pelo indivíduo teste. Quando o contrário ocorria, o indivíduo  $x_{id}$  foi mantido.

Na abordagem adotada, as iterações finalizaram com o número de evoluções determinado previamente. O indivíduo selecionado continha um subconjunto de características que foi considerado o ideal para cada classificador utilizado na função aptidão.

#### 4.4.3 Seleção de um par de características – mRMR e WEKA

Com o objetivo de reduzir para dois o número de características selecionadas pelo método evolutivo que forneceu maior acurácia, foi utilizada a técnica de mínima redundância e máxima relevância (mRMR). Isto foi feito para possibilitar a comparação com os resultados obtidos por trabalhos anteriores do grupo de pesquisa, que utilizaram um espaço de características selecionadas de dimensão dois.

Para a técnica mRMR, foi adotada a abordagem que envolve a análise dos atributos e rótulos em termos da correlação ou da informação mútua. Neste caso, o método foi desenvolvido para encontrar um par de características cuja correlação, entre elas e os rótulos de classe, fosse máxima e a correlação entre si fosse mínima.

Para determinar a informação mútua, foi analisada a qualidade da informação de cada atributo na caracterização adequada das classes. Isto foi feito através do cálculo da entropia (Equação 41) da classe de saída, chamada, também, de incerteza inicial (MANDAL & MUKHOPADHYAY, 2013).

$$H(X) = - \sum_{s=1}^{N_x} \sum P_x(x) \log(P_x(x)) \quad (41)$$

onde  $P_x$  é a densidade de probabilidade para a classe  $x$ , com  $x = 1, 2, \dots, N_x$ .

Em seguida, foi calculada entropia condicional para obtenção da incerteza média relacionada ao conjunto de características de entrada (Equação (42) (MANDAL & MUKHOPADHYAY, 2013). Na Equação 42,  $s$  é o conjunto de  $N_s$  características de entrada e  $P_x(X|S)$  é a probabilidade condicional para a classe  $x$  dado o conjunto de entrada  $S$ .

$$H(X|S) = \sum_{s=1}^{N_s} (\sum_{s=1}^{N_x} P_x(x|s) \log(P_x(x|s))) \quad (42)$$

No caso dos valores de entropia condicional e de entropia inicial serem iguais, então há uma total independência entre os atributos e a classe de saída. Dessa forma, a informação mútua  $I(X, S)$  entre as variáveis  $x$  e  $s$  foi calculada pela incerteza reduzida (Equação 43).

$$I(X, S) = H(X) - H(X|S) \quad (43)$$

A relevância das características foi determinada pela análise das suas distribuições de probabilidade nas diferentes classes, através da Equação 44. Se a distribuição foi aleatória ou uniforme em classes distintas, então a informação mútua com essas classes é zero. Se a distribuição de probabilidade dos atributos foi consideravelmente diferente entre as classes, então sua informação mútua foi grande.

$$relev_{max} = \frac{1}{|s^2|} \sum_{i \in s} I(h, s_i) \quad (44)$$

onde  $h$  é a classe em que está se analisando a correlação com as características.

A redundância mínima foi calculada pela Equação 45. O intuito foi selecionar as características com a máxima diferença mútua entre si.

$$redun_{min} = \frac{1}{|s^2|} \sum_{i,j \in s} I(i, j) \quad (45)$$

Como as duas condições dadas pelas Equações (4.21) e (4.22) são igualmente importantes, elas foram combinadas para seleção do par de características com mRMR (MANDAL & MUKHOPADHYAY, 2013).

A biblioteca WEKA utilizada, versão 3.8.4, é um ambiente de acesso livre que foi desenvolvido na *University of Waikato*, Nova Zelândia. Dentre as 29 características extraídas, um par foi selecionado usando o WEKA *ranker (Information Gain)* para, também, possibilitar a comparação com trabalhos anteriores e para avaliar se houve melhoras no desempenho dos classificadores.

#### 4.5 CLASSIFICAÇÃO DA AMOSTRA

Após as etapas mencionadas, as variáveis selecionadas serviram de entrada para a classificação da base de dados. Nessa etapa, foram testados os classificadores comumente utilizados: kNN, SVM, Árvore de Decisão, Análise Discriminante e Naive Bayes. Previamente a etapa de aprendizado de máquina propriamente dita, foi feita uma seleção da melhor configuração dos hiperparâmetros para construção dos modelos. Foi utilizado o aplicativo *Classification Learner*, disponível no Matlab<sup>®</sup>, para automatizar a seleção dos valores dos hiperparâmetros dos classificadores.

Para cada modelo, o aplicativo realiza diferentes combinações de hiperparâmetros e, por meio de um esquema de otimização bayesiana, minimiza o erro de classificação. Por fim, retorna os hiperparâmetros que maximizam o desempenho preditivo do seu modelo final de aprendizado de máquina. As configurações para cada um dos classificadores estão descritas no Quadro 13.

Devido à limitação no quantitativo de amostras na base de dado, optou-se pela utilização da validação cruzada, ao invés de particionar a amostra em grupo de treino e de validação. O procedimento *k-fold* de validação cruzada foi utilizado para avaliar a capacidade de generalização dos classificadores. A amostra foi particionada em 10 subconjuntos mutuamente excluídos do mesmo tamanho e cada subconjunto *k* foi usado para teste, enquanto os *k* – 1 conjuntos restantes foram utilizados na construção da regra de classificação.

Quadro 13 – Parâmetros adotados na configuração dos classificadores

Classificadores	Parâmetros
KNN	Vizinhos: 5; Distância: euclidiana
SVM	Kernel: polinomial (2)
Árvore de decisão	Observações no nó folha: 1; Número de nós: 1; Número de ramificações: 10
Análise discriminante	Discriminação: diagonal linear; Delta: 0
Naive Bayes	Distribuição: normal; Kernel: gaussiano

Fonte: A Autora (2022)

As medidas de performance utilizadas foram a acurácia (%), a sensibilidade (%), a especificidade (%), o coeficiente *Kappa* e o valor AUC, com seu respectivo intervalo de confiança. Foram usadas diferentes medidas de desempenho com o intuito de melhorar a confiabilidade dos resultados obtidos. Uma vez que, por exemplo, a análise isolada de um resultado alto de acurácia para uma base desbalanceada não indica que o método é adequado, pois ele pode ter previsto apenas a classe majoritária da base.

#### 4.6 CLASSIFICAÇÃO INDIVIDUAL NÃO SUPERVISIONADA

Os processos de classificação utilizados foram desenvolvidos para avaliar a eficiência de cada classificador no estudo da amostra específica. Para os fins da interface proposta, foi necessário ajustar os classificadores selecionados para possibilitar a análise individual de determinada paciente quanto à presença ou não de lesão maligna.

Foram definidos três grupos diferentes: grupo de treino, grupo de validação e de teste, baseados na amostra que compõe a nova base de dados. O grupo de treino, formado pelas características das imagens termográficas da base, foi utilizado para aprendizagem do algoritmo. O grupo de validação, formado pelas classes de cada paciente, foi usado para verificar a generalização do algoritmo, a fim de ajustar os parâmetros de classificação. Após o processo de aprendizagem do programa, o grupo de teste foi utilizado para avaliar se os parâmetros definidos funcionam bem para qualquer amostra.

A partir deste ponto, foram fixados os parâmetros dos classificadores escolhidos como os mais adequados e as características selecionadas a partir das bases de treino que geraram as melhores medidas de performance. Fazendo isto, cada paciente foi usada como “teste” e o resultado da classificação foi gerado apenas para ela.

Duas abordagens de classificação individual foram adotadas: uma baseada na criação de um teste múltiplo, e a outra na escolha do classificador que apresentou melhor desempenho nas taxas de acurácia, de sensibilidade e de especificidade.

Para construção de um teste múltiplo, a partir de dois métodos distintos, foi adotada a associação de testes em série ou sequenciais. Esta associação é bastante utilizada para triagens, principalmente nas situações em que não há urgência no atendimento e quando o paciente pode ser acompanhado ao longo do tempo (SOARES & SIQUEIRA, 1999). O Quadro 14 apresenta o procedimento de indicação do diagnóstico de uma combinação em série. Nela observa-se que se um dos testes for negativo, o resultado do teste em série também será negativo.

Quadro 14 – Resultado do teste sequencial, a partir dos resultados dos testes individuais X e Y

Teste X	Teste Y	Teste sequencial
–	desnecessário	–
+	–	–
+	+	+

Fonte: Adaptado de Soares e Siqueira (1999)

Para triagem, é comum aplicar testes com o objetivo de indicar as pessoas que tem maior probabilidade de estarem com o problema em questão, nesse caso, com lesão maligna mamária. Sabendo que testes com alta sensibilidade possuem maior capacidade em identificar corretamente os doentes, foi adotado como teste inicial o classificador que apresentou o maior valor deste índice.

Sendo o resultado da triagem inicial positivo, um teste confirmatório mais específico foi aplicado. Testes específicos possuem a capacidade de fornecer resultado negativo quando os indivíduos não apresentam a doença. Nesta situação, foi adotado como teste confirmatório o classificador com maior valor de especificidade.

Os índices de sensibilidade ( $Sens_{final}$ ) e de especificidade ( $Esp_{final}$ ) para o teste combinado em série são calculados através das Equações 46 e 47 (SOARES & SIQUEIRA, 1999). Os subíndices representam os testes X e Y.

$$Sens_{final} = Sens_X \times Sens_Y \quad (46)$$

$$Esp_{final} = Esp_X + Esp_Y - Esp_X \times Esp_Y \quad (47)$$

No contexto da prática clínica, as taxas de sensibilidade e de especificidade, apesar de serem medidas que sintetizam adequadamente as qualidades gerais de um teste, não são utilizadas de maneira isoladas na tomada de decisão da equipe médica. Pois são índices provenientes de uma situação em que há certeza do diagnóstico, algo que não acontece em um consultório médico, por exemplo.

Nessas circunstâncias, é comum associar os índices acima aos valores preditivos positivo e negativo (VPP e VPN, respectivamente) que, basicamente, são determinados pela relação entre a sensibilidade e a especificidade do teste e a prevalência da doença na base de dados analisada. Os cálculos de VPP e VPN para o teste combinado em série foram feitos através das Equações 48 e 49, respectivamente (SOARES & SIQUEIRA, 1999).

$$VPP_{final} = \frac{Sens_{final} \times Prev}{Sens_{final} \times Prev + (1 - Esp_{final}) \times (1 - Prev)} \quad (48)$$

$$VPN_{final} = \frac{Esp_{final} \times (1 - Prev)}{Esp_{final} \times (1 - Prev) + (1 - Sens_{final}) \times Prev} \quad (49)$$

onde *Prev* representa a prevalência de pacientes com lesão mamária no conjunto de dados estudado.

#### 4.7 ELABORAÇÃO DO SISTEMA DE CLASSIFICAÇÃO MULTICLASSES

Para o desenvolvimento de uma ferramenta de classificação multiclases, foram adotados alguns métodos aplicados na classificação binária: extração de atributos, seleção de atributos pelos algoritmos evolucionários (AG e ED) e classificação da amostra através do kNN, SVM, Análise Discriminante, Árvore de Decisão e Naive Bayes. As medidas de performance adotadas foram as mesmas da metodologia anterior: acurácia (%) (média), a sensibilidade (%) (média), a especificidade (%) (média), o coeficiente *Kappa* e o valor AUC, com seu respectivo intervalo de confiança.

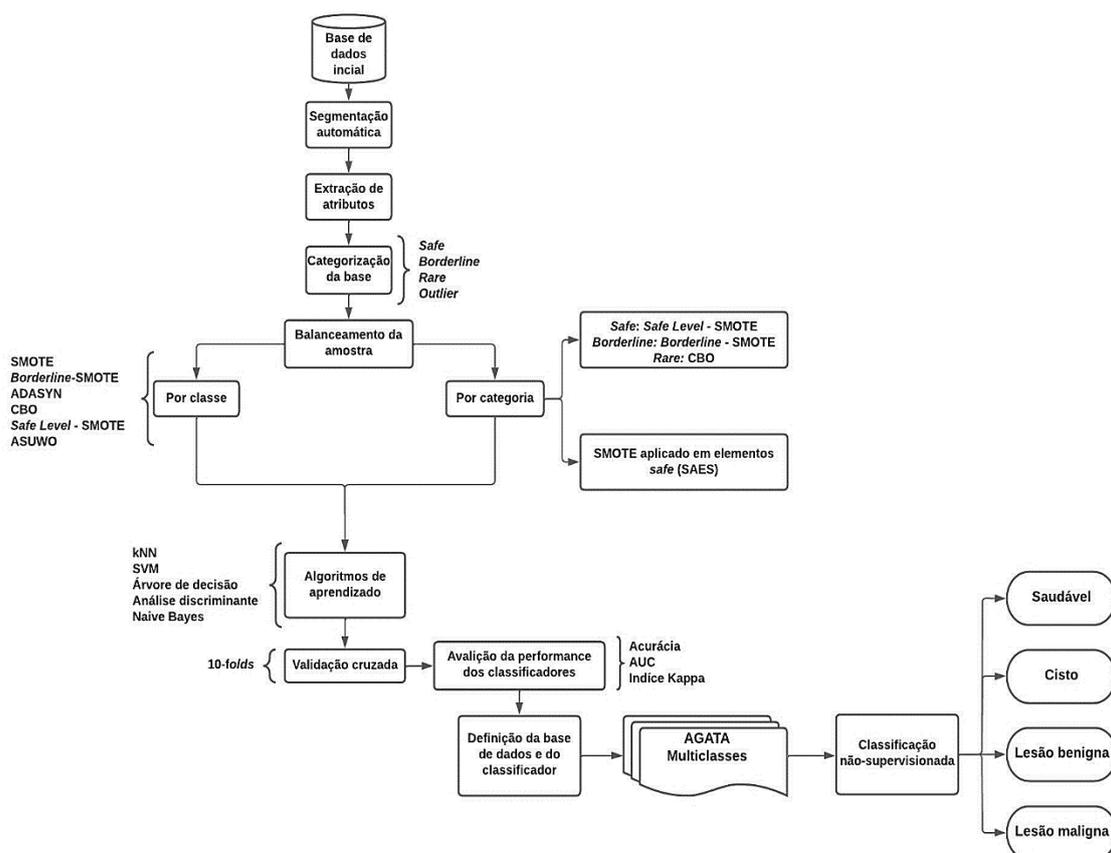
##### 4.7.1 Base de dados

A amostra utilizada nesta etapa corresponde a análise multiclases mostrada na Tabela 1. Analogamente à metodologia anterior, foram usadas imagens por infravermelho (IRs) de mamas do tipo T2 e analisados casos correspondentes à pacientes diagnosticadas como saudáveis (n = 64), com cisto (n=43), com lesão benigna (n=83) e com lesão maligna (n=43). Desta forma, a amostra inicial correspondeu a 233 IRs.

A base de dados caracteriza um problema multiclases de desbalanceamento. Por ser mais complexo, geralmente, as soluções adotadas para o problema binário não são aplicáveis

sem adaptações. No caso da base de dados deste trabalho, observa-se ainda que uma classe pode ser majoritária quando comparada a outras classes, mas pode ser minoritária ou equilibrada quando comparada com as demais. Desse modo, para lidar com este cenário foi feita a análise dos dados de cada classe para identificar suas propriedades individuais e, em seguida, aplicar as técnicas de balanceamento adequadamente (Figura 37).

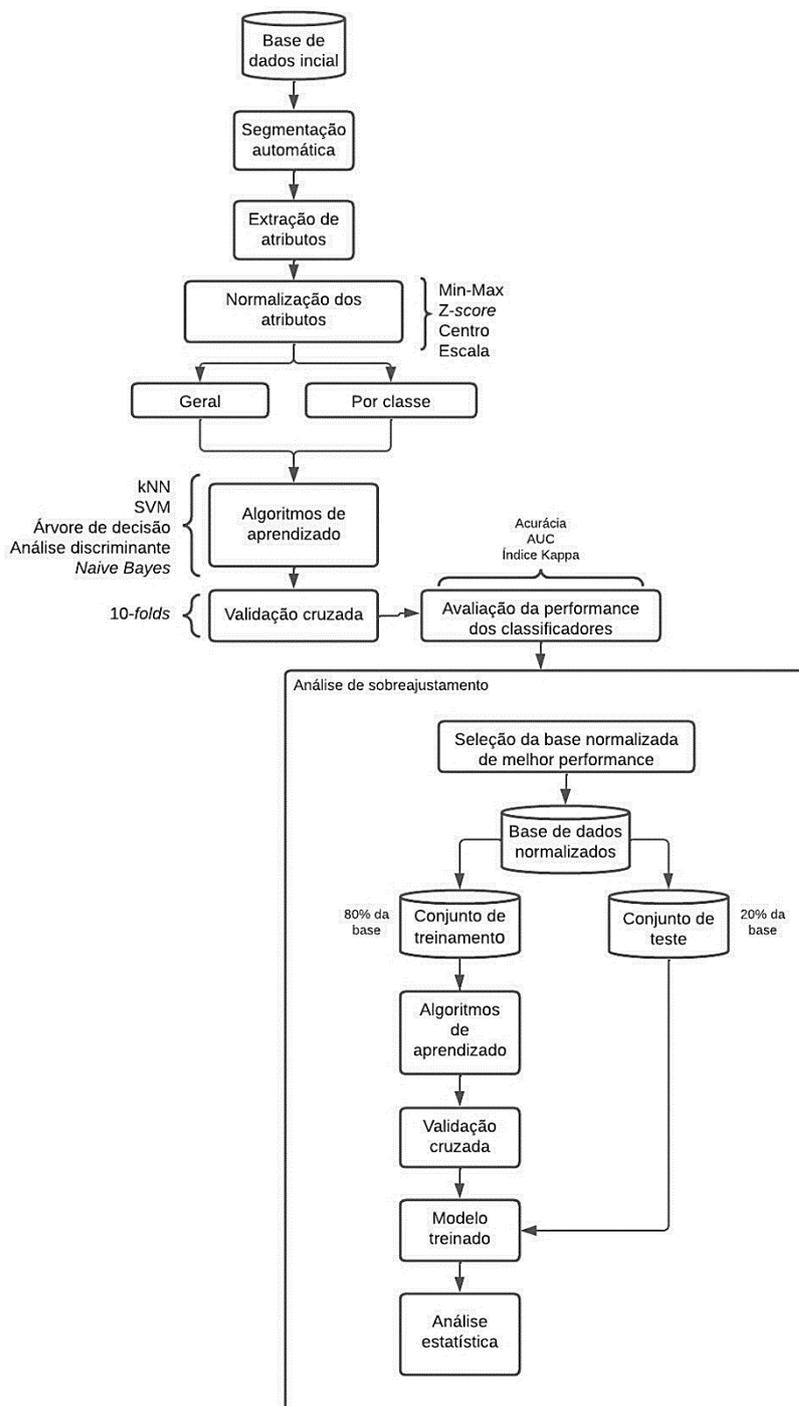
Figura 37 - Fluxograma da metodologia adotada para desenvolvimento do sistema CAD de multiclassificação, com base de dados balanceada



Fonte: A Autora (2022)

Outra metodologia adotada para lidar com a amostra desbalanceada foi a técnica de pré-processamento baseada na transformação de dados – normalização (Figura 38). Além do foco na melhora da performance dos classificadores, o uso da normalização buscou minimizar a sobreposição de atributos interclasses e a variabilidade de atributos intraclasses.

Figura 38 - Fluxograma da metodologia adotada para desenvolvimento do sistema CAD de multiclassificação, com base de dados desbalanceada e normalizada



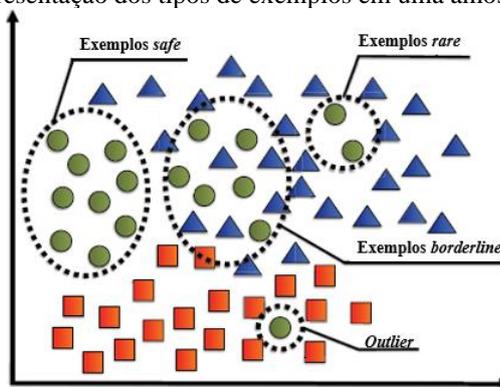
Fonte: A Autora (2022)

#### 4.7.2 Categorização e balanceamento da amostra

O exemplo de uma classe pode ser categorizado segundo sua vizinhança e seu posicionamento intra e interclasses, como sendo: *safe* (seguro), *borderline* (limítrofe), *rare* (raro) e *outlier* (atípico) (SÁEZ; KRAWCZYK; WOZNIAK, 2016) (Figura 39).

Os exemplos *safe* estão localizados em regiões homogêneas, ou seja, comporta pelos exemplos de uma única classe. Esses exemplos são considerados os mais fáceis de aprender por um classificador. A categoria *borderline* é atribuída a exemplos localizados nas regiões em torno do limite de decisão entre as classes ou localizados dentro de regiões sobrepostas de classes. Exemplos raros são conjuntos, normalmente de dois ou três exemplos de classes minoritárias, localizados na região de classes majoritárias, que estão distantes da fronteira de decisão, portanto, não são exemplos limítrofes e, ao mesmo tempo, não são exemplos singulares, portanto não são exatamente *outliers* (NAPIERALA & STEFANOWSKI, 2015).

Figura 39 – Representação dos tipos de exemplos em uma amostra desbalanceada



Fonte: Adaptada de Napierala e Stefanowski (2015)

Para especificar a categoria do exemplo, foi necessário analisar os rótulos de classe dos seus  $k$ -vizinhos mais próximos. A determinação do valor de  $k$  é importante para representação adequada do exemplo. De acordo com Napierala e Stefanowski (2016), a escolha de valores de  $k$  menores que 5 ou muito altos podem gerar uma má distinção da categoria do exemplo. Em consideração a esses aspectos, a partir da variação dos valores de  $k$  foi feita uma análise de sensibilidade, orientando a escolha do número de vizinhos que possibilite uma categorização mais precisa dos exemplos.

Foi considerado  $k$  assumindo valores inteiros de 5 a 15 (NAPIERALA & STEFANOWSKI, 2016) e para cada valor de  $k$ , a classe de um exemplo  $e$  foi comparada com os outros exemplos de sua vizinhança. Assim, a vizinhança foi analisada por meio de uma função  $nc(v)$ , que representa o número de vizinhos  $v$  que possuem o mesmo rótulo de classe do exemplo em questão. O critério adotado neste trabalho para categorização do exemplo foi o seguinte:

$$categoria(e) = \begin{cases} safe, & nc(v) \geq 80\% \\ borderline, & 40\% \leq nc(v) < 80\% \\ rare, & 0 < nc(v) < 40\% \\ outlier, & nc(v) = 0 \end{cases}$$

Após a análise da influência do valor de  $k$ , o número de vizinhos foi determinado para categorizar cada exemplo da amostra conforme as quatro categorias.

Previamente ao balanceamento por categoria, foram criados exemplos sintéticos para cada classe (saudável, cisto e lesão maligna) até atingir o tamanho da classe majoritária (lesão benigna). Este balanceamento foi feito por meio das técnicas de sobreamostragem utilizadas na metodologia de classificação binária e, cada uma das amostras balanceadas por elas foram avaliadas através do processo de classificação.

Para o balanceamento categórico, foram adotadas duas estratégias:

- Identificação no espaço de características dos exemplos das classes (saudável, cisto e lesão maligna) considerados *safe*, *borderline*, *rare* e *outlier*. Em seguida, aplicar os balanceamentos que apresentaram melhor performance na sobreamostragem de elementos *safe* e *rare*, e aplicar a técnica *Borderline-SMOTE* sobre os elementos limítrofes.
- Identificação dos elementos *safe* da amostra e sobreamostragem destes utilizando a técnica *SMOTE* e um percentual de geração de vetores sintéticos variando de 50 a 80%, em relação as demais categorias (NAPIERALA; STEFANOWSKI, 2015).

### 4.7.3 Normalização

Em problemas de aprendizado de máquina, a sobreposição de classes, associada ao desequilíbrio presente na amostra, pode prejudicar a etapa de aprendizagem, diminuindo o número de exemplos classificados corretamente. Os trabalhos realizados pelo grupo de pesquisa, mostraram que as características selecionadas podem não resultar em uma boa separação das classes, não apresentando fronteiras bem definidas entre elas e sobreposição de uma classe em relação às outras.

Sendo assim, com o intuito de melhorar os resultados obtidos foram usadas técnica de normalização por reescala, por padronização e por distribuição. Como mostra o Quadro 15, para cada técnica foram testadas opções específicas sobre os dados da base.

Quadro 15 – Técnicas de normalização testadas

Normalização	Opção testada	Descrição
Reescala	Min-Max	Adequação dos valores para um novo intervalo [0;1]
Padronização	Z-score	Adequação dos valores pelo cálculo da distância entre o atributo e a média, em termos do desvio padrão
Distribuição	Escala	Adequação dos valores em relação ao desvio padrão
	Centro	Adequação dos valores para centro com média zero

Fonte: A Autora (2022)

Para normalização da base desbalanceada, foram adotadas duas estratégias:

- Normalização dos atributos desconsiderando a divisão por classes;
- Normalização dos atributos por classe (HA et al., 2021), para qual foram testadas as cinco opções citadas anteriormente.

#### *Análise de superajustamento*

Na modelagem preditiva, o interesse é obter modelos que realizam as melhores previsões possíveis, dados os recursos computacionais disponíveis. Diante deste fato, foi feita uma análise estatística dos modelos de aprendizagem obtidos para identificar se algum deles superajustou o conjunto de treinamento.

O superajustamento (do inglês, *overfitting*) ocorre quando as previsões do modelo em observações futuras, que não foram usadas no treinamento, são fracas. Assim, quando há o ajuste excessivo dos dados de treinamento, a precisão do classificador para uma base teste será muito pequena. Isso porque o suposto modelo gerado pelo algoritmo de aprendizagem para os dados de treinamento, não é válido para os dados de teste/validação, sugerindo *overfitting*.

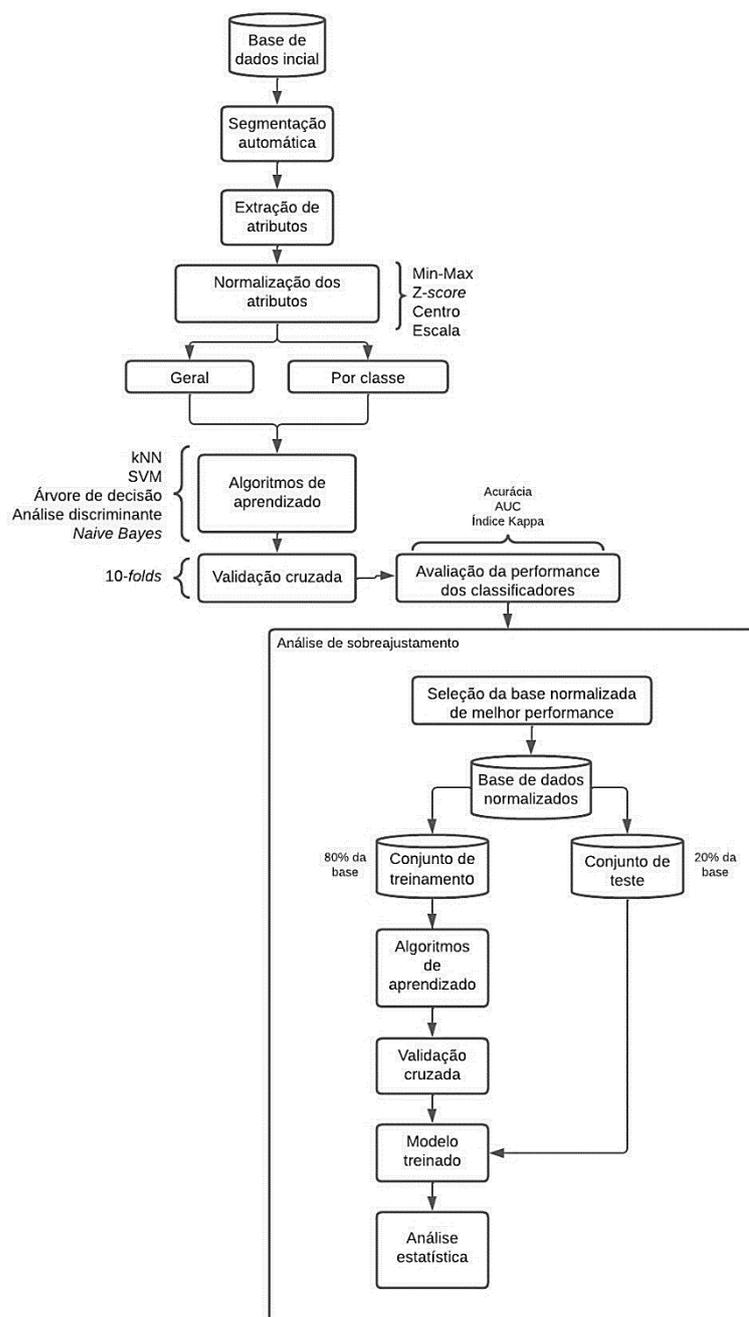
A presença de *overfitting* nos resultados obtidos pela classificação das bases normalizadas foi feita através de uma metodologia de análise estatística, além do uso da validação cruzada na base de treinamento. Todas as análises estatísticas foram realizadas com o *software* Minitab® - versão 18. As variáveis foram testadas quanto à sua normalidade através do teste de Kolmogorov-Smirnov. Para verificar diferenças entre médias ou medianas foi utilizado o teste t-Student ou o teste U de Mann-Whitney, respectivamente. A significância estatística foi definida para um valor de  $p < 0,05$ . Todos os valores de  $p$  são bicaudais.

Destaca-se que, anteriormente a análise estatística, os dados normalizados pelas técnicas que obtiveram melhor performance de classificação foram selecionados e divididos em conjuntos de treino e teste. Para isso, 80% dos dados foram usados para treinamento, com aplicação da validação cruzada para fornecimento das métricas de desempenho associadas a esse grupo, e 20% dos dados foram usados como teste no modelo gerado, com o intuito de verificar discrepâncias, ou não, nos desempenhos dos classificadores.

#### 4.8 ESTRATÉGIA PARA ESTIMATIVA DA LOCALIZAÇÃO DO TUMOR MALIGNO

A metodologia completa proposta para previsão do tamanho e da profundidade do tumor maligno é apresentada na Figura 40.

Figura 40 - Fluxograma da metodologia adotada para predição do tamanho e da profundidade da lesão maligna

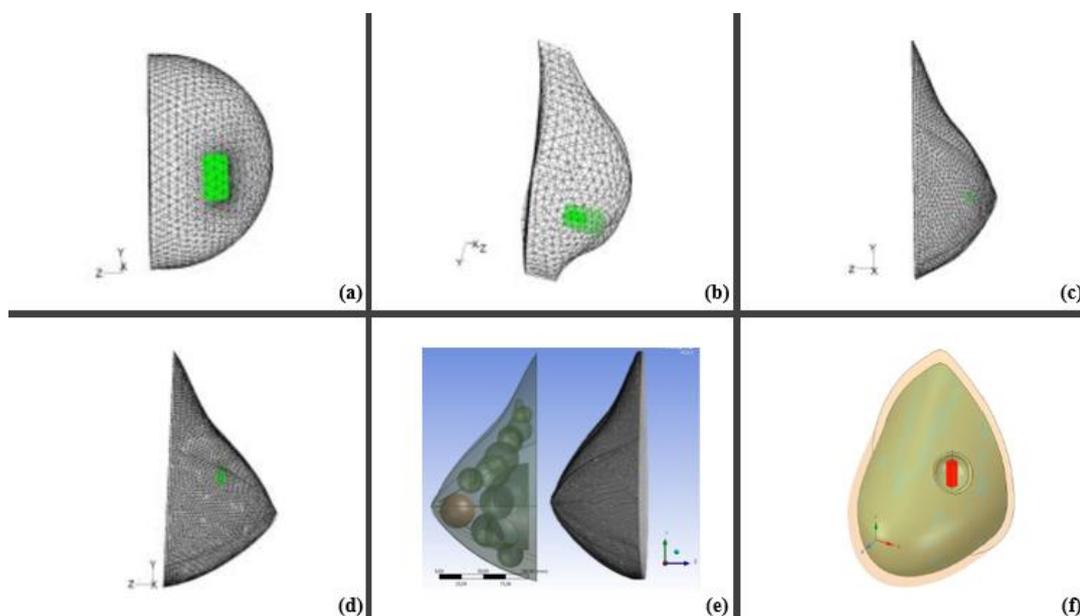


Fonte: A Autora (2022)

#### 4.8.1 Modelagem geométrica da mama

Em 2006, o grupo de pesquisa de Transferência de Calor em Sistemas Biológicos da UFPE iniciou seus estudos sobre geometrias substitutas para representar a anatomia da mama. Desde então, foram desenvolvidos seis modelos tridimensionais cuja evolução está representada na Figura 41.

Figura 41– Representação das mamas do grupo de pesquisa. (a) Semiesfera (b) Fantoma (c) Prótese de silicone (d) Prótese de silicone ajustada à paciente (e) Geometria com estruturas internas (f) Geometria de mama constituída por tecidos glandular e adiposo criada a partir de termograma



Fonte: Adaptado de Melo (2019) e de Viana (2016)

Inicialmente, foi utilizada uma semiesfera para representação tridimensional da mama (Figura 41a). Em seguida, a modelagem foi baseada em um fantoma do dorso feminino (Figura 41b), descrita em Santos et al. (2009). Viana (2010) desenvolveu as geometrias ilustradas nas Figuras 41c e 41d, sendo elas consideradas mais realistas quando comparadas com as anteriores.

Em parceria com a Universidade Federal Fluminense (UFF), a representação do volume mamário foi aprimorada através da extração da curva da prega inframamária e da curva do perfil da mama (SILVA, 2010). Viana (2016) incluiu a essa geometria esferas para representar os lóbulos mamários e cilindros para representar os ductos lactíferos (Figura 41e).

Melo et al. (2020) criou uma geometria 3D estruturada sobre a superfície plana posterior das próteses mamárias e ajustada a cada paciente a partir de termogramas (Figura 46f). Agora, o modelo mais realista do grupo de pesquisa foi construído a partir da curva de perfil da paciente extraída no *software SpaceClaim*<sup>®</sup>.

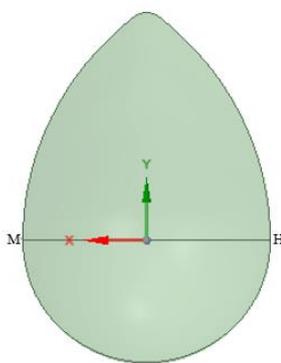
Como a composição da mama difere entre pessoas, seja por diferenças genéticas, raciais ou pelo processo de liposubstituição do parênquima mamário decorrente do envelhecimento, Melo et al. (2020) fez uma simplificação da estrutura interna da mama feminina através da criação de um modelo geométrico formado por tecido glandular, tecido adiposo e pela lesão mamária.

No presente trabalho, foram criados quatro modelos tridimensionais com diferentes composições de densidade mamária, a saber: modelo de mama Extremamente Densa (ExD),

modelo de mama Heterogeneamente Densa (HD), modelo de mama Predominantemente Gordurosa (PG) e modelo de mama Parcialmente Fibroglandular (PF). A metodologia completa para obtenção e desenvolvimento desses modelos é descrita em Melo et al. (2020).

De acordo com o método desenvolvido por Melo et al. (2020), o dimensionamento da parede torácica da geometria é determinado por um segmento de reta  $\overline{MH}$  que passa através do mamilo e possui extremidades os pontos limitantes da mama (Figura 42). Assim, baseada em descrições de literatura consagrada (NG & SUDHARSAN, 2001), foi usado um  $\overline{MH} = 14,4$  cm para representar a geometria média da mama feminina.

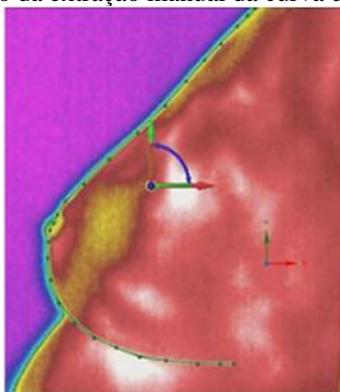
Figura 42 - Ajuste da base torácica com a medida da semirreta  $\overline{MH}$



Fonte: A Autora (2022)

O segmento de reta foi colocado sobre a superfície plana da geometria e a parede torácica foi dimensionada. Em seguida, a curva de perfil foi extraída de um termograma (Figura 43). Por fim, a curva de perfil foi adaptada à parede torácica, assumindo seu tamanho real proporcional a  $\overline{MH}$ .

Figura 43 – Representação da extração manual da curva de perfil de um termograma



Fonte: Queiroz Júnior e Lima (2020)

Os modelos de referência foram construídos baseados nas composições teciduais de acordo com o BI-RADS desenvolvido pela *American Cancer Research (ACR) EKPO*, (HOGG & MCENTEE, 2015; VIANA, 2016). Na modelagem de tais geometrias, foi utilizado um modelo padrão cujo tamanho corresponde à média do volume das mamas das mulheres brasileiras. Com isso e utilizando as porcentagens teciduais descritas no Quadro 16, foram obtidos os volumes de gordura e de tecido glandular para cada subtipo de mama. Como forma de simplificação anatômica, esses volumes foram inseridos na região central dos modelos. A Figura 44 contém a representação das geometrias produzidas.

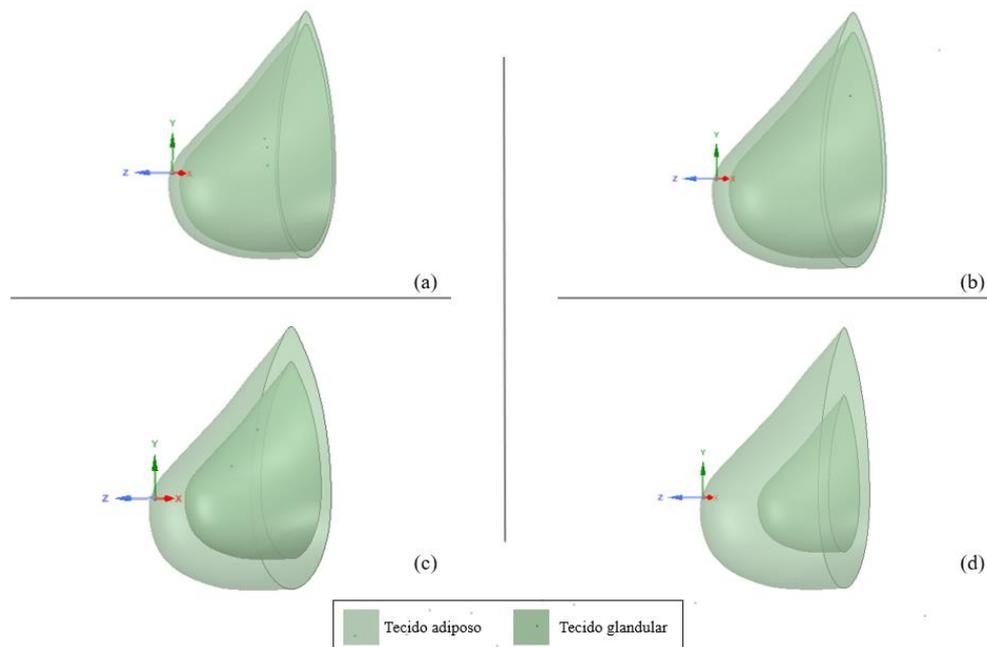
Quadro 16 – Composição percentual de tecido

<b>Composição tecidual</b>	<b>Tecido glandular</b>	<b>Tecido adiposo</b>
Extremamente densa	70 %	30 %
Heterogeneamente densa	60 %	40 %
Parcialmente fibroglandular	40 %	60 %
Predominantemente gordurosa	20 %	80 %

Fonte: Adaptado de Ekpo, Hogg e Mcentee (2015) e Viana (2016)

O tumor foi modelado como uma esfera. Além disso, as propriedades dos modelos foram estabelecidas para se obter a maior proximidade possível da fisiologia real da mama. A fim de minimizar variações significativas no perfil de temperatura da superfície, foi considerada uma distribuição espacial homogênea dos vasos sanguíneos em todos os componentes dos tecidos.

Figura 44 - Visão transversal de modelos de mama com diferentes composições mamárias. (a) ExD (b) HD (c) PF (d) PG



Fonte: A Autora (2022)

#### 4.8.2 Modelagem matemática utilizada

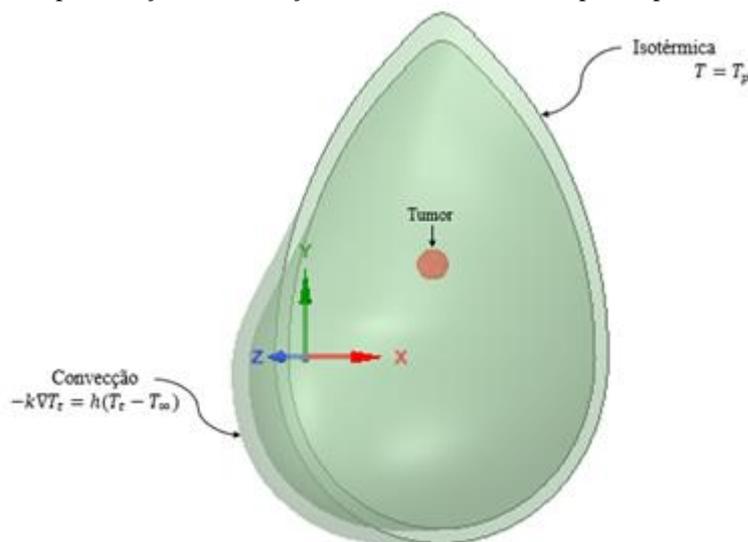
As geometrias construídas foram exportadas para o *software ANSYS Fluent*, no qual as simulações numéricas foram executadas juntamente com os módulos de geração de malha, de pré e pós-processamento da plataforma. As simulações foram realizadas através do Método dos Volumes Finitos (MVF), que divide a região de interesse em volumes de controle, nos quais a equação que descreve o problema é resolvida em cada um desses volumes, e posteriormente monta-se o sistema de equações algébricas que aproxima o problema em todo o domínio.

O *ANSYS Fluent* possui uma variação da equação de energia que leva em conta os mecanismos de transferência de calor especificados em cada domínio que se queira avaliar. Essa equação pode ser alterada através da inserção de *User-Defined Function (UDF)* que servem para implementar os termos fonte. Os termos fonte são constantes ou expressões que podem conter variáveis do sistema, como a temperatura (PINHEIRO, 2020). Os termos fontes do presente trabalho tratam da perfusão sanguínea e da geração de calor para cada região do domínio (tecido tumoral, tecido mamário e tecido adiposo). A UDF foi implementada seguindo o método descrito em Bezerra (2013).

O mecanismo de transferência de calor do problema estacionário foi modelado pela BHTE, descrita na Seção 3.2. As condições de contorno adotadas para todos os modelos foram especificadas na superfície da pele da mama, conforme mostrado no esquema da Figura 45. Na

extensão de toda parede torácica foi considerada a condição de contorno de Dirichlet, na qual a temperatura é representada por  $T_p = 37^\circ\text{C}$ . A condição de contorno de Cauchy foi aplicada à região cuja transferência de calor é por convecção da superfície da mama para o ar ambiente da sala, que mantém temperatura  $T_\infty$  com coeficiente de transferência de calor por convecção  $h = 13,5 \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-1}$  (NG & SUDARSHAN, 2001).

Figura 45 – Representação das condições de contorno adotadas para o problema em estudo



Fonte: A Autora (2022)

Além das considerações citadas na descrição das condições de contorno, a modelagem matemática foi baseada ainda em um conjunto de hipóteses simplificadoras (MELO et al., 2020), a saber:

- Temperatura do sangue ( $T_a$ ) igual a  $37^\circ\text{C}$ ;
- Temperatura constante da sala e igual a  $27^\circ\text{C}$ , onde foram feitas as aquisições das imagens por infravermelho;
- Propriedades termofísicas constantes em todas as regiões do domínio;
- Simulação numérica em regime permanente;
- Modelos de referência representados por três regiões bem definidas: tecido adiposo, tecido glandular e tecido tumoral.

#### 4.8.3 Propriedades termofísicas e simulação numérica

Um conjunto de parâmetros termofísicos, para cada tipo de tecido, foi usado na simulação numérica (Quadro 17). Posteriormente, foi realizada a discretização dos modelos de mamas

utilizando uma malha tetraédrica não-estruturada de volumes finitos. Para cada modelo, foi feito um estudo de convergência das malhas relacionadas ao conjunto mama-tumor. Esse estudo levou em consideração a temperatura máxima superficial ( $T_{max,supf}$ ) da região afetada pela lesão maligna e o critério de parada foi que a variação da  $T_{max,supf} \leq 0,1$  °C com o refinamento da malha.

Quadro 17 – Parâmetros termofísicos para todos os modelos de mamas

Parâmetros termofísicos	Tecido glandular	Tecido adiposo	Tecido tumoral	Sangue
$k$ ( $W \cdot m^{-1} \cdot K^{-1}$ )	0,48 <sup>a,b,d</sup>	0,21 <sup>a,f</sup>	0,48 <sup>a,b,d</sup>	-
$\rho$ ( $kg \cdot m^{-3}$ )	1.080 <sup>a,b,d</sup>	911 <sup>a</sup>	1.080 <sup>a,b,d</sup>	1.060 <sup>a,b,c</sup>
$c$ ( $J \cdot kg^{-1} \cdot K^{-1}$ )	3.000 <sup>b,d</sup>	3.100 <sup>a</sup>	3.500 <sup>b,d</sup>	4.200 <sup>b</sup>
$\omega$ ( $1 \cdot s^{-1}$ )	0,00018 <sup>b,d</sup>	0,0005 <sup>a</sup>	0,009 <sup>b,d</sup>	-
$Q_m$ ( $W \cdot m^{-3}$ )	450 <sup>b,e</sup>	400 <sup>f</sup>	*	-

Fonte: Adaptada de Melo et al. (2020)

$k$  = condutividade térmica,  $\rho$  = massa específica,  $c$  = calor específico,  $\omega$  = taxa de perfusão sanguínea,  $Q_m$  = taxa volumétrica de geração de calor metabólico.

<sup>a</sup> Hasgall (2015); <sup>b</sup> Bezerra (2013); <sup>c</sup> Okajima *et al.* (2009), <sup>d</sup> González (2007); <sup>e</sup> Ng e Sudharsan (2001); <sup>f</sup> Wahab et al (2015); \*Valor calculado de acordo com seu diâmetro e utilizando Equação 3.4 e Equação 3.5 (Seção 3.2)

Para cada tipo de densidade mamária, foram realizadas 33 simulações de casos hipotéticos. Foram simulados diferentes tamanhos e profundidades de tumor cujos valores foram definidos conforme a literatura (WAHAB et al., 2015) e usando valores existentes no banco de dados do grupo de pesquisa. Os tamanhos simulados variaram de 10 a 30 mm de diâmetro, com incremento de 0,62 mm. Já as profundidades variaram 5 a 19 mm, com incremento de 0,41 mm, medidas como sendo a distância normal em relação a superfície da mama. Foi fixada uma profundidade do tumor, variando-se apenas o tamanho do tumor, o que totalizou numa base de 132 casos simulados.

A posição adotada para a localização do tumor foi o Quadrante Superior Lateral (QSL), Segundo Pinheiro et al. (2013), o QSL é a região de maior prevalência de casos de lesões malignas, correspondendo a 50% dos casos notificados.

A temperatura ambiente adotada nas simulações numéricas foi de 27 °C. Esse valor representa a aproximação da média aritmética das temperaturas ambientes dos casos que compõe a base de dados do trabalho (233 pacientes).

#### 4.8.4 Modelo de aprendizado de máquina

A natureza do problema de determinação da profundidade e do tamanho de um tumor maligno envolve uma resposta quantitativa, o que levou a escolha da regressão numérica para tratá-lo. A regressão tem como objetivo a previsão de valores de dados numéricos e categóricos, no caso da regressão logística, através da construção de um modelo de previsão (IBAN & SENTÜRK, 2022).

Para fins de previsão neste trabalho, foi utilizada a SVM devido ao seu desempenho na etapa anterior de classificação das patologias mamárias. Além disso, foi analisada a colinearidade entre as variáveis independentes dos modelos e constatou-se que os conjuntos de características que representam cada geometria simulada possuem correlação moderada/baixa, sendo considerados dados não-lineares. Assim, a regressão por SVM mapeou os recursos de entrada para espaços dimensionais mais altos para reorganizar os dados de maneira linearmente solucionável.

Cada simulação possui um conjunto de dados contendo informações ponto a ponto dos valores de temperatura (em Kelvin, K) da malha da superfície da geometria avaliada. Desses dados, foram extraídas as características discutidas na Seção 4.2.2 (Quadro 12). Dentre elas, desconsideraram-se as características  $c_5$  a  $c_8$  (diferenças entre temperaturas) e as características  $c_{11}$  a  $c_{18}$  (extraídas segundo critério de robustez). Assim, foram extraído um total de 17 características nesta etapa do trabalho.

Foi feita uma análise estatística para identificar as mais significantes para expressar diferenças entre mamas saudável e com tumor maligno. Esses atributos serviram de entrada para criação de modelos preliminares de regressão por SVM. As saídas dos modelos foram, separadamente, os vetores com os valores simulados de profundidade e de tamanho do tumor maligno.

Anteriormente à construção final dos modelos de regressão para cada densidade mamária simulada, foi feita uma seleção dos atributos utilizando o algoritmo de relevância máxima de redundância mínima (mRMR). O mRMR indicou um conjunto ótimo de características que foram mutuamente e maximamente dissimilares e que poderiam representar a variável de resposta de forma adequada. Procurou-se fixar o menor número possível de características, pois modelos de regressão por aprendizado de máquina se treinados com muito vetores de entrada tendem ao sobreajustamento, o que pode levar a um alto erro de generalização e a um tempo excessivo de processamento (IBAN & SENTÜRK, 2022).

Uma vez definidos os parâmetros de entrada e de saída, o aplicativo *Regression Learner* do Matlab<sup>®</sup> foi usado para criar modelos de regressão por SVM. O *software* avalia modelos de

SVM com função *kernel* do tipo gaussiana, linear, quadrática e cúbica. A seleção do modelo mais adequado foi feita a partir da análise do gráfico dos valores preditos *vs.* valores observados, do coeficiente de determinação ( $R^2$ ) e da raiz quadrada do erro quadrático médio (REQM).

O melhor valor possível de  $R^2$  (Equação 50) é 1, o que indica que o modelo explica toda a variabilidade dos dados previstos em torno da sua média ou, simplesmente que os valores previstos são sempre iguais aos observados e todos esses pontos estão sobre a linha de regressão (KHAN et al., 2022).

$$R^2 = 1 - \frac{\sum(y - y_p)^2}{(y - \bar{y})^2} \quad (50)$$

onde  $y$  representa o valor observado,  $y_p$  é o valor previsto e  $\bar{y}$  é a média dos valores observados.

O REQM possui cálculo semelhante ao EQM, sendo sensível a erros elevados e, pela presença da raiz quadrada, apresenta valores do erro na mesma dimensão que o da variável analisada. A fórmula do REQM é demonstrada na equação abaixo, onde  $n$  denota o número de valores da base de dados, neste caso  $n = 132$ . Ressalta-se que foram selecionados para análise, os modelos que apresentaram um REQM  $< 4$ .

$$\text{REQM} = \sqrt{\frac{1}{n} \sum (y - \bar{y})^2} \quad (51)$$

Após a escolha dos modelos SVM de treinamento, foi feita a otimização dos hiperparâmetros: escala da função *kernel*, valor parâmetro C que visa adicionar uma penalidade a cada elemento previsto incorretamente. Um valor C alto fornece um modelo mais flexível, já um valor menor fornece um modelo mais rígido, menos sensível ao *overfitting*. Para otimizá-los, foi utilizado o método de otimização Bayesina cuja função objetivo foi o EQM do modelo.

O desempenho de cada modelo foi verificado em um conjunto de teste reservado previamente: 85% dos dados foram treinados e 15% foram usados para fins de teste. As métricas de treino e de teste foram comparadas para investigar a compatibilidade dos valores de REQM e a possível presença de um ajustamento excessivo dos modelos.

#### 4.8.5 AGATA Multiclasses – Módulo localização

Um módulo para previsão do tamanho e da profundidade do tumor maligno foi adicionado à interface AGATA Multiclasses. O intuito é fornecer ao usuário a possibilidade de estimar esses valores, dado que o *software* classificou a paciente como portadora de lesão maligna.

A IR selecionada para classificação é, então, retomada e o usuário segmenta manualmente as mamas atentando-se a dois aspectos principais: seleção de áreas, aproximadamente, equivalentes e não selecionar a região do osso esterno (entre mamas).

A simetria térmica foi usada para indicar a mama suspeita. Para isso foi utilizada a seguinte comparação entre as temperaturas médias ( $T_{med}$ ) das ROIs:

$$\begin{aligned} \text{se } T_{med,MD} > T_{med,ME} &\implies MD \text{ suspeita de possuir lesão maligna} \\ \text{caso contrário} &\implies ME \text{ suspeita de possuir lesão maligna} \end{aligned}$$

A escolha do modelo de regressão a ser utilizado para prever o tamanho e a profundidade foi associada à idade da paciente e sua composição tecidual, conforme mostrado no Quadro 18.

Quadro 18 – Classificação da densidade mamária de pacientes de acordo com a idade

<b>Idade</b>	<b>Densidade mamária</b>
≤ 25	Extremamente densa
26 – 45	Extremamente densa
46 – 60	Heterogeneamente densa
61 – 75	Parcialmente fibroglandular
> 75	Predominantemente gordurosa

Fonte: Adaptada de Spak et al. (2017)

Assim, com a informação de qual mama está a lesão e qual modelo de regressão será usado, no módulo de localização da GUI são apresentados os valores estimados de tamanho e profundidade e seus respectivos coeficientes de determinação.

#### 4.9 INFRAESTRUTURA COMPUTACIONAL

Para realização do presente trabalho, foi usado um computador com sistema operacional Windows 10, processador Intel Core i3 – 7200U (2,50 GHz), 8 GB de memória RAM e Sistema Operacional de 64 bits. As ferramentas computacionais utilizadas foram os *softwares* FLIR Tools<sup>®</sup> - versão 6.4.17317.1002, Matlab<sup>®</sup> - versão R2018a, Minitab<sup>®</sup> - versão 18, Microsoft Excel<sup>®</sup> - versão 2206 e Ansys<sup>®</sup> - versão 18.1, e a biblioteca WEKA – versão 3.8.4 (HALL et al., 2009). O FLIR Tools<sup>®</sup> foi usado para obtenção das matrizes de temperatura a partir das imagens termográficas. O Quadro 19 descreve, de formal geral, o uso das demais plataformas nas etapas descritas anteriormente.

Quadro 19 - Aplicações das plataformas computacionais utilizadas

<b>Validação da segmentação automática</b>	
<b>Infraestrutura computacional</b>	<b>Finalidade</b>
Matlab	<ul style="list-style-type: none"> <li>• Segmentação automática (DOURADO, 2014) <ul style="list-style-type: none"> <li>• <i>Ground-truth</i> (ARAÚJO, 2014)</li> </ul> </li> <li>• Implementação das medidas de sobreposição</li> </ul>
Microsoft Excel e Minitab	<ul style="list-style-type: none"> <li>• Análise estatística</li> </ul>
<b>Sistema de classificação binária</b>	
<b>Infraestrutura computacional</b>	<b>Finalidade</b>
Matlab	<ul style="list-style-type: none"> <li>• Segmentação automática (DOURADO, 2014)</li> <li>• Extração de atributos (funções disponíveis na biblioteca do <i>software</i>)</li> <li>• Implementação dos métodos de balanceamento</li> <li>• Implementação dos métodos de seleção de atributos (AG, ED e mRMR)</li> <li>• Classificação (aplicativo <i>Classification Learner</i> e funções disponíveis na biblioteca do <i>software</i>)</li> <li>• Implementação da avaliação da performance <ul style="list-style-type: none"> <li>• Interface AGATA Triagem (GUIDE)</li> </ul> </li> </ul>
WEKA	<ul style="list-style-type: none"> <li>• Seleção de um par de atributos (analisador <i>Information Gain</i>)</li> </ul>
<b>Sistema de classificação multiclases</b>	
<b>Infraestrutura computacional</b>	<b>Finalidade</b>
Matlab	<ul style="list-style-type: none"> <li>• Segmentação automática (DOURADO, 2014)</li> <li>• Extração de atributos (funções disponíveis na biblioteca do <i>software</i>)</li> <li>• Implementação dos métodos de balanceamento</li> <li>• Implmentação dos métodos de seleção de atributos (AG e ED)</li> <li>• Normalização (funções disponíveis na biblioteca do <i>software</i>)</li> <li>• Classificação (aplicativo <i>Classification Learner</i> e funções disponíveis na biblioteca do <i>software</i>)</li> <li>• Implementação da avaliação da performance <ul style="list-style-type: none"> <li>• Interface AGATA Multiclases (GUIDE)</li> </ul> </li> </ul>
Microsoft Excel e Minitab	<ul style="list-style-type: none"> <li>• Análise de sobreajustamento</li> </ul>
<b>Previsão do tamanho e da profundidade da lesão maligna</b>	
<b>Infraestrutura computacional</b>	<b>Finalidade</b>
Ansys	<ul style="list-style-type: none"> <li>• <i>SpaceClaim</i><sup>®</sup> (modelagem geométricas)</li> <li>• <i>Fluent</i><sup>®</sup> (simulação numérica)</li> </ul>
Matlab	<ul style="list-style-type: none"> <li>• Extração de atributos (funções disponíveis na biblioteca do <i>software</i>)</li> <li>• Modelos de regressão (aplicativo <i>Regression Learner</i>)</li> <li>• Avaliação dos modelos (aplicativo <i>Regression Learner</i>)</li> <li>• Módulo Localização – AGATA Multiclases (GUIDE)</li> </ul>
Microsoft Excel e Minitab	<ul style="list-style-type: none"> <li>• Análise da reprodutibilidade e da repetibilidade do método</li> </ul>

Fonte: A Autora (2022)

## 5 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados dos diversos estudos realizados para obtenção de sistemas CADs destinados à análise binária e multiclases de patologias mamárias em imagens por infravermelho. Além disso, na Seção 5.1 estão os resultados da metodologia adotada para validação da segmentação automática de termogramas de mamas proposta por Dourado (2014). E na Seção 5.4 são descritos os resultados obtidos com o procedimento para as estimativas de tamanho e de profundidade.

### 5.1 VALIDAÇÃO DA SEGMENTAÇÃO AUTOMÁTICA

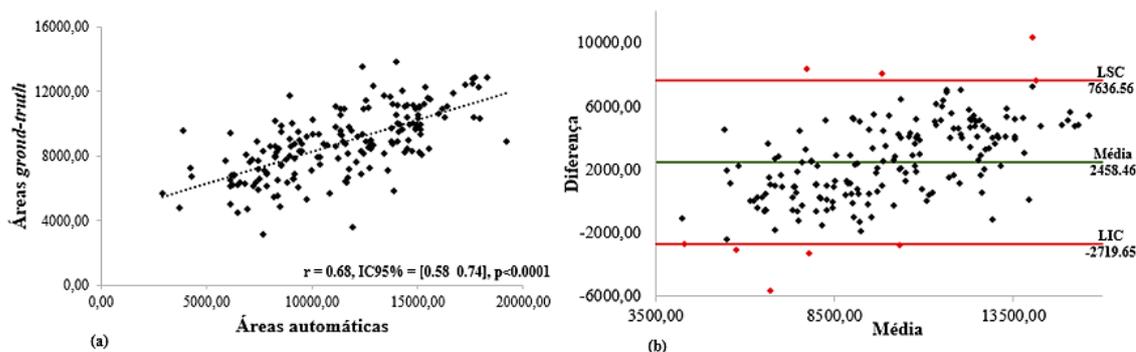
A validação de técnicas de segmentação, com base em imagens *ground-truth*, melhora a confiabilidade dos métodos utilizados para captura das imagens, além de permitir a medição da eficiência de diferentes algoritmos de extração de regiões de interesse. O método mais comum para gerar o *ground-truth* de imagens médicas é a utilização de segmentações manuais (Li et al., 2011) como proposto neste trabalho.

A média de idade das pacientes usadas nesta etapa foi de  $46,8 \pm 13,0$  com intervalo de confiança (IC95%) de 44,2 - 44,9 anos. O diagnóstico predominante foi lesão mamária benigna (33,3%). A aquisição das imagens por infravermelho foi realizada em ambiente controlado cujas medianas de temperatura ( $T_{amb}$ ) e umidade relativa (UR%) foram  $26,6 [25,7 - 27,7]$  °C e  $61,5 [58,0 - 64,0]$  %, respectivamente.

O tempo computacional para cada segmentação automática foi de aproximadamente 0,5 segundos por imagem. A metodologia de validação descrita anteriormente foi aplicada ao resultado da segmentação automática. Cada mama foi considerada individualmente. Assim, o conjunto de dados de validação foi  $n=180$ . As áreas *ground-truth* podem ser encontradas em: <https://bit.ly/30rTs65>.

Ao comparar o valor da área extraída por segmentação automática (S) com a área *ground-truth* (GT), verificou-se que as medidas possuem correlação moderada ( $r=0,68$ ) (Figura 46a). A concordância entre as áreas foi medida usando o gráfico de Bland-Altman (Figura 46b). O gráfico de Bland-Altman mostrou o viés médio de  $2.458,46 \pm 2.641,89$  pixels entre a segmentação automática e as áreas de verdade, e 95% dos dados estavam dentro dos limites de concordância (-2.719,65 e 7.636,56).

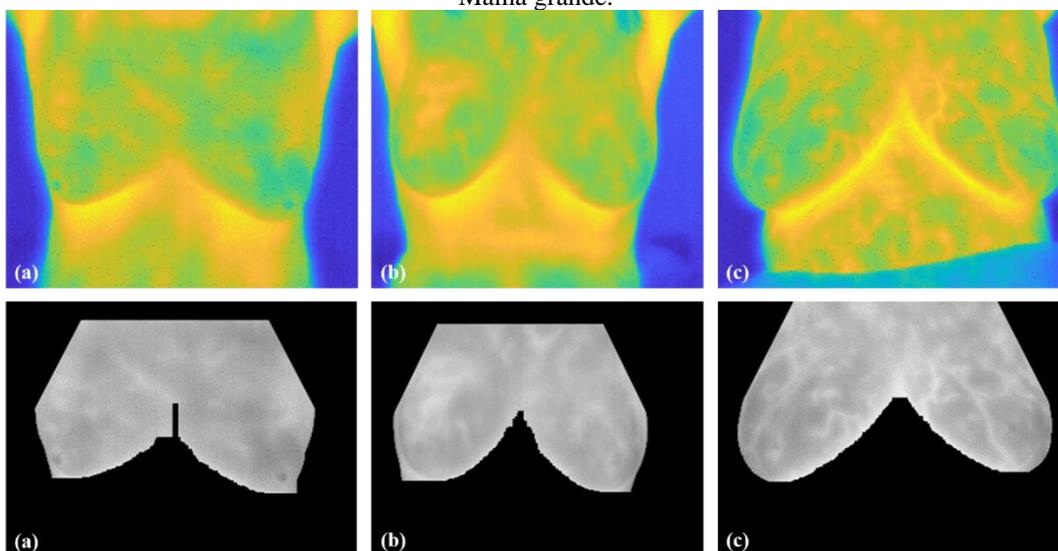
Figura 46 - (a) Gráfico de dispersão para as variáveis área S e área GT. (b) Gráfico de Bland-Altman para média e diferenças entre as áreas segmentadas. LSC: Limite Superior de Concordância; LIC: Limite Inferior de Concordância.



Fonte: A Autora (2022)

O sistema de segmentação automática produziu resultados satisfatórios para todas as imagens por infravermelho de mama analisadas. Na Figura 47 são apresentados exemplos dos resultados obtidos pelo método proposto em condições de diferentes corpos e a comparação com a imagem real.

Figura 47 - Resultados da segmentação para diferentes tipos de mama (a) Mama pequena (b) Mama média e (c) Mama grande.



Fonte: A Autora (2022)

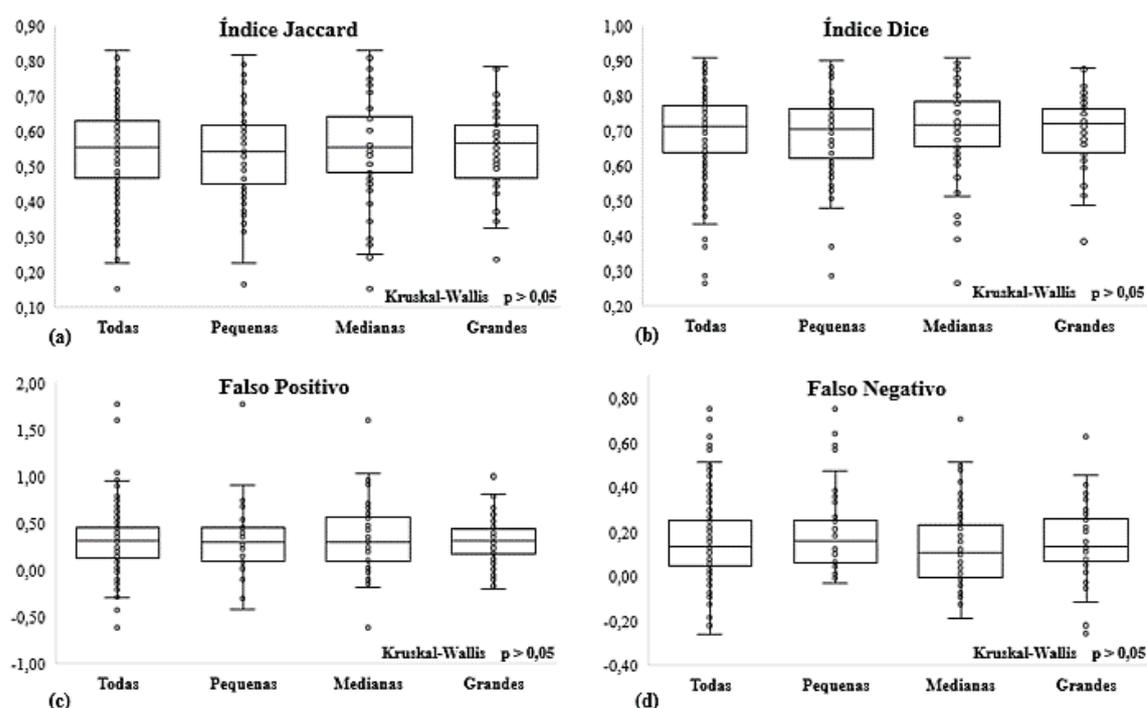
A concordância entre os métodos foi confirmada pelo gráfico de Bland Altman (Figura 46b), que indicou um bom nível de concordância entre as áreas extraídas. A diferença média obtida foi maior que zero mostrando que a segmentação automática selecionou mais pontos em relação à segmentação *ground-truth*.

Na literatura sobre o assunto, o GT é geralmente definido pelo especialista com um esboço como a forma final da segmentação automática proposta (Marques et al., 2016; Sathish et al., 2017; Adel et al., 2018). Por outro lado, para gerar o GT adotada neste estudo, o contorno já é pré-determinado por uma elipse que é ajustada manualmente ao tamanho da mama. Tal fato pode justificar os valores superestimados obtidos pela análise de concordância. Mesmo assim, a técnica de segmentação automática não mostrou diferença entre os diversos padrões anatômicos considerados, o que sugere adequação para todos os tipos de mamas.

Na Figura 48, são apresentados os resultados referentes às métricas de análise de segmentação automática. A análise dos valores da base com todas as imagens mostrou que os índices Jaccard e Dice tiveram valor médio de  $0,54 \pm 0,13$  e  $0,70 \pm 0,11$ , respectivamente. As taxas de Falso Positivo e Falso Negativo apresentaram, nesta ordem, valor médio de  $0,32 \pm 0,32$  e  $0,16 \pm 0,17$ .

Utilizando a escala de classificação de similaridade em função do índice Dice desenvolvido por Cabezas et al. (2014), houve concordância substancial entre as regiões segmentadas automaticamente e as regiões *ground-truth*. As taxas de falsos positivos e negativos tiveram valores relativamente próximos de zero, o que indica que a segmentação automática está selecionando corretamente o ROI, quando comparado com a respectiva GT.

Figura 48 - Métricas para avaliar a segmentação automática e sua variação em relação à verdade do terreno (a) Índice de Jaccard. (b) Dados de Índice. (c) Falso positivo. (d) Falso negativo



Fonte: A Autora (2022)

O resultado experimental (teste de Kurskal-Wallis  $p > 0,05$ ) mostrou que não houve diferenças estatisticamente significativas entre os grupos com todas as imagens e com especificação por tamanho anatômico.

Considerando o cenário em que o objetivo deste estudo foi validar a nova técnica de segmentação automática, os valores encontrados na análise de áreas sobrepostas não foram considerados uma limitação. Foi encontrada uma boa concordância entre os dados e este foi considerado o fator mais importante.

## 5.2 ANÁLISE DA CLASSIFICAÇÃO BINÁRIA

Neste trabalho, 29 características foram extraídas e submetidas a diferentes técnicas de pré-processamento e de classificação. No Apêndice F encontram-se os gráficos *boxplot* para cada atributo extraído da base de dados ( $n = 233$  termogramas de mamas). Neles observou-se que a distribuição das características nas classes “Sem lesão maligna” e “Com lesão maligna” possuem valores de mediana próximos. Tal fato indica que, apesar da grande quantidade de características extraídas, não foi possível estabelecer uma que maximizasse a separabilidade entre as classes. Assim, foi necessário realizar experimentos com diversas combinações desses atributos, determinadas pela seleção da combinação que otimizou a acurácia de cada técnica de aprendizado de máquina utilizada.

### *Utilização de todas as características*

Inicialmente, foram aplicadas todas as características para avaliação do desempenho dos classificadores (Apêndice G). Os dados mostram que o classificador Análise Discriminante obteve uma performance mais uniforme em relação às taxas de acurácia, sensibilidade e especificidade, apesar de o índice *kappa* indicar uma concordância leve entre a acurácia obtida e a esperada. Os demais classificadores apresentaram baixa sensibilidade, o que pode estar atrelado ao fato de que há uma maior prevalência da Classe Sem Lesão Maligna na base de dados inicial.

Esses resultados, mostram que o desempenho obtido com a classificação da amostra desbalanceada, utilizando todas as características extraídas, foi inferior aos resultados da literatura vigente, considerando os autores (MADHAVI & BOBY, 2017; ARUL EDWIN RAJ; SUNDARAM; JAYA, 2019; EKICI & JAWZAL, 2020; MISHRA & RATH, 2020) que estudaram o problema binário em uma base de dados desbalanceada em seus testes. Entretanto, destaca-se que a base do presente trabalho foi, em média, 278% maior do que aquelas descritas nos trabalhos citados, com um grau de desbalanceamento elevado (classe majoritária é 3,41 vezes

maior que a minoritária) e com sobreposição entre classes que pode ocasionar problemas na geração de um bom modelo de classificação.

#### *Utilização das características selecionadas*

A eventual presença de características redundantes e irrelevantes, que não contribuem com efeito na definição de uma determinada classe, pode interferir negativamente na performance dos classificadores. Sendo assim, nesta etapa, foram utilizados métodos evolucionários (Algoritmo Genético e Evolução Diferencial) para seleção da melhor combinação de atributos que fornecesse uma maior taxa de acurácia para cada classificador, utilizando o menor número possível de atributos (Tabela 2). As taxas de acurácia (ACC) apresentadas foram calculadas a partir do erro de classificação, adotado como função objetivo do AG e ED.

Tabela 2 – Resultados da seleção de características para base de dados binária desbalanceada

Classificadores	Seleção de atributos AG		Seleção de atributos ED	
	ACC (%)	Quant.	ACC (%)	Quant.
<b>Análise discriminante</b>	88,82	9	88,79	9
<b>Árvore de decisão</b>	88,79	14	91,81	17
<b>kNN</b>	90,62	13	90,09	20
<b>Naive Bayes</b>	89,78	6	90,09	13
<b>SVM</b>	91,10	13	87,93	11

Fonte: A Autora (2022)

\*Quant. – quantidade de características selecionadas

Os parâmetros definidos, após realização de testes, para o AG e o ED foram apresentados nos Quadros 11 e 12 (ver Seções 4.4.1 e 4.4.2). A utilização de um valor de gerações muito baixo, diminuiria o desempenho dos algoritmos. Dessa forma, apesar do alto tempo de processamento, foi empregado um valor de geração (2.900) que cobrisse o domínio do problema, evitando a convergência para soluções locais (CHAMBERS, 2019).

A taxa de mutação utilizada (0,01) foi embasada pela literatura vigente. Chambers (2019) ressalta a importância do uso de taxas abaixo de 0,1, a fim de se evitar uma taxa de mutação muito alta, já que ela pode tornar a busca aleatória, prejudicando, também, a convergência dos algoritmos.

Comparada com os resultados obtidos com o uso de todas as características, a performance dos classificadores treinados por subconjuntos selecionados melhorou, com destaque para a acurácia e para o índice *kappa*. A acurácia média saiu de 77,83 [69,83;84,48] % para 87,33 [85,34; 88,79] % (seleção AG) / 85,35 [83,019; 87,50] % (seleção ED). Já o índice *kappa* indicou uma alteração da concordância leve para a moderada.

#### *Balanceamento da amostra*

Com o intuito de melhorar os resultados obtidos, por meio da minimização da influência da classe majoritária, o conjunto de dados foi submetido ao processo de balanceamento através da aplicação de métodos de sobreamostragem (SMOTE, *Borderline*- SMOTE, CBO, *Safe level* - SMOTE, ASUWO e ADASYN). Após cada balanceamento, a base de dados formada é mostrada na Tabela 3.

Tabela 3 – Resultados do balanceamento da base de dados binária

<b>Métodos de sobreamostragem</b>						
	<b>SM</b>		<b>B-SM</b>		<b>CBO</b>	
	Quantidade	Percentual (%)	Quantidade	Percentual (%)	Quantidade	Percentual (%)
<b>SLM</b>	190	47,50	190	50,00	190	50,00
<b>CLM</b>	210	52,50	190	50,00	190	50,00
<b>Total</b>	400	100,00	380	100,00	380	100,00
	<b>SL-SM</b>		<b>AW</b>		<b>AD</b>	
	Quantidade	Percentual (%)	Quantidade	Percentual (%)	Quantidade	Percentual (%)
<b>SLM</b>	190	50,00	188	49,87	190	50,80
<b>CLM</b>	190	50,00	189	50,13	184	49,20
<b>Total</b>	380	100,00	377	100,00	374	100,00

Fonte: A Autora (2022)

SLM: Sem lesão maligna; CLM: Com lesão maligna  
SMOTE: SM; *Borderline*-SMOTE: B-SM; *Safe level* SMOTE: SL-SM; ASUWO: AW; ADASYN: AD

A performance dos classificadores adotados neste trabalho foi avaliada para cada conjunto de dados obtido, considerando os diferentes métodos de balanceamento por sobreamostragem (Apêndice H). Constatou-se que a utilização de tais métodos aperfeiçoou a performance dos classificadores, quando comparada com a obtida na avaliação da amostra desbalanceada (Apêndice G). O processo de balanceamento favoreceu a taxa de sensibilidade, equilibrando-a com a especificidade, com valores AUC variando de bom a excelente. Na escolha de testes diagnósticos, idealmente, deve-se optar por aqueles que possuem altas sensibilidade e especificidade, o que minimizaria resultados falsos positivo e negativo.

Na análise dos valores  $\kappa$ , destaca-se que a implementação do balanceamento entre as classes melhorou a confiabilidade dos algoritmos. Os testes tiveram sua faixa alterada para um nível de concordância que variou de substantiva a ótima.

#### *Seleção de características na amostra balanceada*

Para seleção de atributos, foram escolhidas bases de dados balanceadas por três metodologias distintas (SMOTE, *Borderline*-SMOTE e ASUWO). A técnica SMOTE foi escolhida por ser uma das mais conhecidas e que se tem mostrado eficaz em várias aplicações na literatura

(LARA et al., 2017). Já o *Borderline-SMOTE* e o ASUWO foram selecionados por terem apresentado os melhores índices de classificação.

Na Tabela 4 contém o quantitativo de atributos selecionados pelos algoritmos evolucionários e as taxas de acurácia para cada classificador. Desses atributos, a técnica mRMR selecionou as duas características consideradas mais relevantes e com menor redundância. Ressalta-se que na escolha do subconjunto para aplicação da mRMR considerou-se aquele que forneceu menor quantitativo de atributos selecionados, quando a acurácia entre os algoritmos evolucionários foi igual.

A metodologia baseada no WEKA guiou, dentre os atributos que formam o conjunto inicial, a escolha dos seguintes pares:  $c_3$  e  $c_6$  (SMOTE),  $c_3$  e  $c_{21}$  (SMOTE *Borderline*) e  $c_3$  e  $c_{13}$  (ASUWO).

A técnica mRMR apresentou medidas de desempenho inferiores para todos os classificadores e para as bases de dados balanceadas pelas três técnicas escolhidas anteriormente. Já na avaliação comparativa do desempenho obtido entre os resultados desta etapa (amostras balanceadas + seletores de características) e os resultados utilizando todas as características extraídas das amostras balanceadas, especificamente pelos três métodos selecionados (SMOTE, SMOTE *Borderline* e ASUWO) (Tabela 4), foram adotados os índices AUC e  $\kappa$  (Kappa).

A AUC foi selecionada por levar em consideração os acertos (TVP) e os erros (TFP) dos classificadores e, por isso, é mais adequado numa análise preliminar do desempenho. Já o índice  $\kappa$  é considerado um dado conservador que fornece uma avaliação do nível de concordância entre as acurácias (previstas e esperadas). Para esta análise, consideraram-se as medidas relacionadas à aplicação das características extraídas pelas técnicas AG, ED e WEKA. Nos Apêndices I, J e K encontram-se mais detalhes sobre a performance dos classificadores para essa etapa.

Tabela 4 – Resultados da seleção de características para base de dados binária balanceada por SMOTE, SMOTE *Borderline* e ASUWO

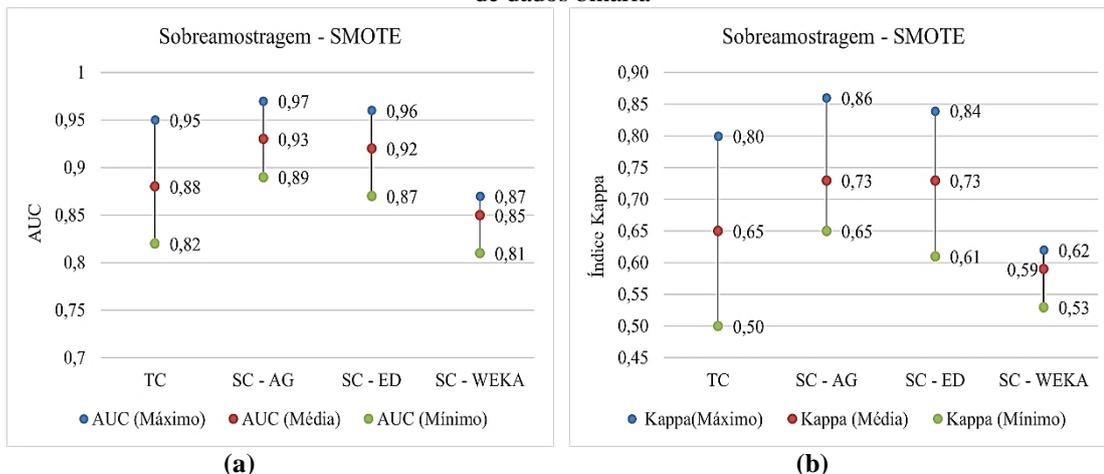
Classificadores	Método de balanceamento				
	SMOTE				
	Seleção de atributos AG		Seleção de atributos ED		Seleção mRMR
	ACC (%)	Quant.	ACC (%)	Quant.	Par selecionado
<b>Análise discriminante</b>	84,25	9	81,75	13	$c_{16}, c_{17}$
<b>Árvore de decisão</b>	91,00	17	89,75	14	$c_{19}, c_{28}$
<b>kNN</b>	91,75	19	91,00	25	$c_{10}, c_{13}$
<b>Naive Bayes</b>	84,24	11	83,25	14	$c_7, c_{17}$
<b>SVM</b>	96,50	17	96,50	22	$c_9, c_{10}$
Classificadores	Método de balanceamento				
	SMOTE <i>Borderline</i>				
	Seleção de atributos AG		Seleção de atributos ED		Seleção mRMR
	ACC (%)	Quant.	ACC (%)	Quant.	Par selecionado
<b>Análise discriminante</b>	91,05	10	88,42	21	$c_{16}, c_{17}$
<b>Árvore de decisão</b>	92,89	19	91,84	16	$c_{17}, c_{29}$
<b>kNN</b>	93,16	14	93,16	20	$c_{17}, c_{29}$
<b>Naive Bayes</b>	91,32	8	90,53	12	$c_7, c_{20}$
<b>SVM</b>	94,47	18	94,47	29	$c_{17}, c_{29}$
Classificadores	Método de balanceamento				
	ASUWO				
	Seleção de atributos AG		Seleção de atributos ED		Seleção mRMR
	ACC (%)	Quant.	ACC (%)	Quant.	Par selecionado
<b>Análise discriminante</b>	90,45	15	89,66	19	$c_{10}, c_{17}$
<b>Árvore de decisão</b>	93,10	13	93,10	13	$c_5, c_{28}$
<b>kNN</b>	95,23	14	95,49	17	$c_{10}, c_{15}$
<b>Naive Bayes</b>	90,72	12	88,86	23	$c_{10}, c_{17}$
<b>SVM</b>	98,41	17	98,94	19	$c_{15}, c_{28}$

Fonte: A Autora (2022)

\*Quant. – quantidade de características selecionadas

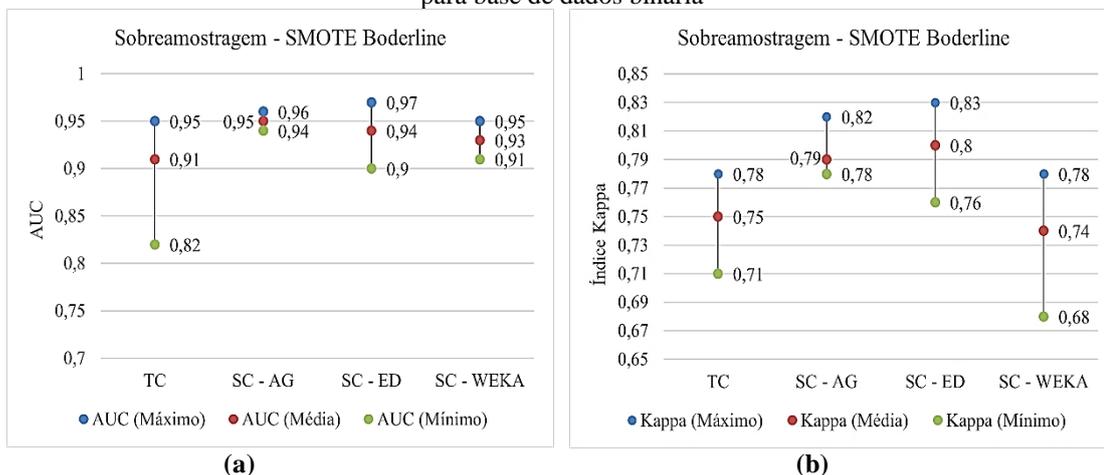
Os intervalos representados nos gráficos das Figuras 49-51 foram construídos a partir da média dos índices AUC e  $\kappa$  de todos os classificadores, para cada método de *oversampling* e de seleção de características escolhidos anteriormente. A determinação dos valores máximo e mínimo apresentados, também, foi feita através da análise do desempenho obtido pelos classificadores em cada amostra balanceada. Nesses gráficos, TC, SC-AG, SC-ED e SC-WEKA representam todas características, seleção de características por Algoritmo Genético (AG), por Evolução Diferencial (ED) e pelo WEKA, respectivamente.

Figura 49 – Resultados da (a) AUC e do (b) Índice Kappa utilizando sobreamostragem por SMOTE, para base de dados binária



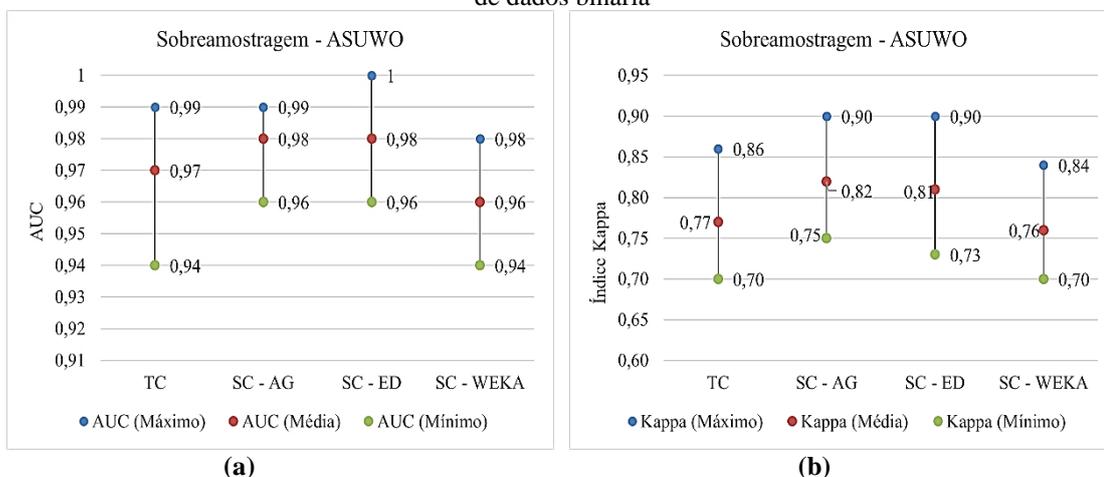
Fonte: A Autora (2022)

Figura 50– Resultados da (a) AUC e do (b) Índice Kappa utilizando sobreamostragem por SMOTE Borderline, para base de dados binária



Fonte: A Autora (2022)

Figura 51 - Resultados da (a) AUC e do (b) Índice Kappa utilizando sobreamostragem por ASUWO, para base de dados binária



Fonte: A Autora (2022)

A partir da análise gráfica dos índices AUC e  $\kappa$ , nota-se que, para as bases balanceadas por SMOTE e *Borderline*- SMOTE, o uso dos atributos selecionados pelos métodos evolucionários proporcionou, de forma geral, melhorias no desempenho dos classificadores. Quando comparados com a utilização de todas as características, também se observou uma diminuição dos intervalos que pode ser associada a um aumento da homogeneidade da performance dos classificadores treinados com tais atributos.

A aplicação dos atributos selecionados através do WEKA na base sobreamostrada pelo SMOTE gerou uma redução intervalar nos índices AUC e  $\kappa$ . Porém, para o AUC não houve melhorias nos valores extremos do intervalo, o que sugere que apesar da diminuição do tempo de processamento computacional pelo uso de um par de características, não houve otimização considerável dos valores. Na generalidade, não houve aumento dos valores de  $\kappa$ .

Para amostra balanceada pela técnica ASUWO, houve aumento da precisão provocado pela diminuição dos intervalos associados ao índice AUC dos classificadores treinados e validados com atributos selecionados. O impacto foi mínimo na performance obtida com o uso do par selecionado pelo *software* WEKA, comparada àquela alcançada com a aplicação de todas as características.

O coeficiente  $\kappa$ , ainda para base sobreamostrada por ASUWO, não apresentou diminuição do intervalo em nenhuma das metodologias adotadas. No entanto, os métodos SC-AG e SC-ED proporcionaram respostas superiores, sugerindo que o uso das características selecionadas por eles gera maior concordância entre os resultados.

Avaliando a metodologia final, os melhores resultados foram obtidos através do emprego das técnicas de sobreamostragem e de seleção de características por algoritmos evolucionários. Assim sendo, das tabelas e gráficos anteriores, destacam-se positivamente os classificadores SVM (seleção AG + base SMOTE), com acurácia de 93,00% (ver Apêndice I); kNN (seleção ED + base *Borderline*- SMOTE), com acurácia de 91,32% (ver Apêndice J). e SVM (seleção AG + base ASUWO), com acurácia de 95,23% (ver Apêndice K). Apesar de performances similares do classificador SVM utilizando seleção por AG e ED na base ASUWO, a seleção por algoritmo genético foi escolhida pela menor quantidade de características selecionadas.

Quanto a comparação entre os melhores resultados obtidos nos experimentos descritos anteriormente (destacado em **negrito**) e os resultados apresentados na literatura e nos trabalhos anteriores do grupo de pesquisa (destacados de *itálico*), nota-se que a performance do classificador SVM (seleção AG + base ASUWO) do presente trabalho foi, de maneira geral, superior à de trabalhos encontrados (Quadro 20). E, mesmo entre aqueles cujo resultado da metodologia

adotada não foi o melhor, ainda pode-se destacar o emprego de uma base ampliada, sendo mais representativa para o aprendizado do modelo. A comparação é imprescindível para fundamentar se os resultados obtidos foram propícios ou não.

Ressalta-se, no entanto, que a comparação entre os trabalhos do mesmo grupo não é totalmente adequada, pois a base de dados é atualizada à medida que novos casos são concluídos. Portanto, raramente serão encontrados trabalhos que utilizarão o mesmo conjunto de imagens. Isso vale para os trabalhos de autores externos ao grupo de pesquisa, que além de utilizarem base com tamanhos distintos, adotam outros protocolos de aquisição.

Quadro 20 - Síntese dos resultados de classificação binária de imagens termográficas mamárias de trabalhos encontrados na literatura

<b>Autor</b>	<b>Base</b>	<b>Classificador</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>
<i>Dourado (2014)</i>	234	SVM	79,49	67,44	82,20
<i>Queiroz (2016)</i>	98	SVM	74,74	51,85	83,82
Bhowmik et al. (2016)	20	Assimetria térmica	90,00	90,00	90,00
Lashkari, Pak e Firouzmand (2016)	–	AdaBoost	86,38	–	–
Silva et al. (2016)	80	Redes Bayesianas ANN	– –	100,00 92,59	100,00 96,30
Madhavi e Bobby (2017)	67	LSSVM	89,00	86,00	100,00
<i>Vila Nova (2017)</i>	380	<i>Combinação por Voting</i>	94,21	95,23	93,16
<i>Vasconcelos, Santos e Lima (2018)</i>	380	SVM	93,42	94,73	92,10
Pramanik et al. (2019)	50	ANN	89,40	86,00	90,00
Arul Edwin Raj, Sundaram e Jaya (2020)	40	SVM	90,00	91,60	87,50
Ekici e Jawzal (2020)	140	CNN	98,95	–	–
Lennox e Haskins (2020)	60	Random Forest	90,00	–	–
Mishra e Rath (2020)	56	Random Forest	95,45	99,17	88,07
Silva et al. (2020)	64	SVM	100,00	100,00	100,00
<b>Metodologia de Classificação Binária</b>	<b>377</b>	<b>SVM</b>	<b>95,23</b>	<b>93,65</b>	<b>96,81</b>

Fonte: A Autora (2022)

### 5.2.1 AGATA Triagem - proposta de um sistema CAD de classificação binária

O AGATA Triagem (Figura 52) é um sistema que foi desenvolvido usando as técnicas descritas na presente tese. A ferramenta computacional fornece ao usuário um auxílio automatizado no diagnóstico precoce do câncer de mama, a partir de matrizes de temperatura provenientes de imagens térmicas. O *software* foi desenvolvido no Matlab<sup>®</sup> a partir da associação dos sistemas CADs analisados a uma interface gráfica de usuário (GUI, do inglês *Graphical User Interface*).

As interfaces gráficas do Matlab® foram criadas através da função *GUIDE* (*Graphical User Interface Development Environment*). Nesta ferramenta é possível projetar a GUI, selecionando e alinhando os componentes inseridos na mesma. Ao gravar a GUI, é gerado um sub-programa que inclui funções de esqueleto que podem ser modificadas para implementar o código que define o comportamento da interface (CHAPMAN, 2003).

Figura 52– Interface gráfica AGATA Triagem

Fonte: A Autora (2022)

A interface gráfica desenvolvida possui três menus (**Abrir**, **Ajuda** e **Sair**) e três painéis (**Informações do Paciente**, **Resultado** e **Estatística**). No menu **Abrir** o usuário tem a opção utilizar o programa FLIR QuickReport e exportar a matriz de temperaturas, no formato *.csv*, da imagem termográfica selecionada. No menu **Ajuda** encontra-se uma breve explicação sobre a interface e suas funcionalidades, com o intuito de que o usuário possa tirar suas dúvidas acerca dela.

No painel **Informações do Paciente** o médico/pesquisador pode inserir os dados do paciente e, da mesma forma que o resultado de classificação gerado, estes dados serão armazenados em um arquivo *.txt* para possibilitar o seu uso no painel **Estatística**. O painel **Estatística** serve para produção de gráficos (setores – **Casos clínicos**, e histogramas – **Casos clínicos×Idade**) dos dados registrados em determinado período.

Duas versões do AGATA para triagem foram desenvolvidas: uma associada à implementação de testes combinados em série, e a outra versão implementada com o classificador que apresentou, de maneira geral, a melhor performance. Para esta última versão, foi selecionado o classificador SVM (seleção AG + base ASUWO), destacado anteriormente, cujos valores preditivos foram de 93,81% (VPP) e 96,72% (VPN).

Na versão AGATA- Testes combinados, foi implementado como teste de triagem inicial o classificador com maior índice de sensibilidade – k-NN (seleção AG + base SMOTE), que apresentou sensibilidade de 100,00% e especificidade de 78,95%. Para o teste de confirmação de diagnóstico, foi utilizado o classificador com maior especificidade – k-NN (seleção ED + base ASUWO), que apresentou sensibilidade de 82,54% e especificidade de 100,00%. Por consequência, os valores finais de sensibilidade e especificidade do teste sequencial foram de, respectivamente, 82,54 % e 100,00 %. Já os valores preditivos foram de 100,00% (VPP) e de 96,20% (VPN).

A fim de padronizar e uniformizar os resultados para apresentação na interface, foi definido um novo índice ICT (Índice de Classificação Térmica) baseado nas imagens termográficas e nos classificadores estatísticos, para indicar a existência de possíveis patologias mamárias. A proposta de tal índice é ser similar ao BI-RADS que é usado nos exames de mamografia e ultrassonografia. O Quadro 21 mostra os valores atribuídos ao ICT e suas respectivas definições.

Quadro 21 – Classificação ICT

Classificação	Avaliação
ICT 0	Sem lesão maligna
ICT 1	Saudável
ICT 2	Cisto
ICT 3	Lesão benigna
ICT 4	Lesão maligna

Fonte: A Autora (2022)

O índice ICT 0 foi incluído os casos de triagem utilizando a avaliação binária. Neste cenário, a paciente que apresentar ICT 0 (Sem lesão maligna) pode ser portadora de cisto, lesão benigno ou não possuir anormalidades.

O resultado associado ao ICT é apresentado no painel **Resultado**. No primeiro espaço aparecerá um número referente ao índice, e no segundo espaço aparecerá a avaliação. No caso de a classificação binária ser negativa para lesão maligna (ICT 0), no local destinado ao valor preditivo aparecerá o VPN, caso contrário aparecerá o VPP.

### 5.3 ANÁLISE DA CLASSIFICAÇÃO MULTICLASSES

Neste tópico serão apresentados e discutidos os resultados obtidos na classificação multiclases da base de dados de imagens termográficas de mamas.

#### 5.3.1 Base de dados desbalanceada

Seguindo, inicialmente, a mesma linha metodológica utilizada para classificação binária, foi feita a análise da amostra multiclases (Tumor Maligno, Tumor Benigno, Cisto e Saudável) desbalanceada utilizando todas as características (os resultados encontram-se no Apêndice L). Posteriormente, foi feita a classificação após seleção de atributos por AG e ED cujos parâmetros foram apresentados nos Quadros 11 e 12 (Ver Seções 4.4.1 e 4.4.2). A Tabela 5 mostra a quantidade de atributos selecionada para cada algoritmo de aprendizagem e as acurácias calculadas a partir do erro de classificação, que foi usado como função objetivo dos métodos evolucionários de seleção de características.

Tabela 5 – Resultados da seleção de características para base de dados multiclases desbalanceada

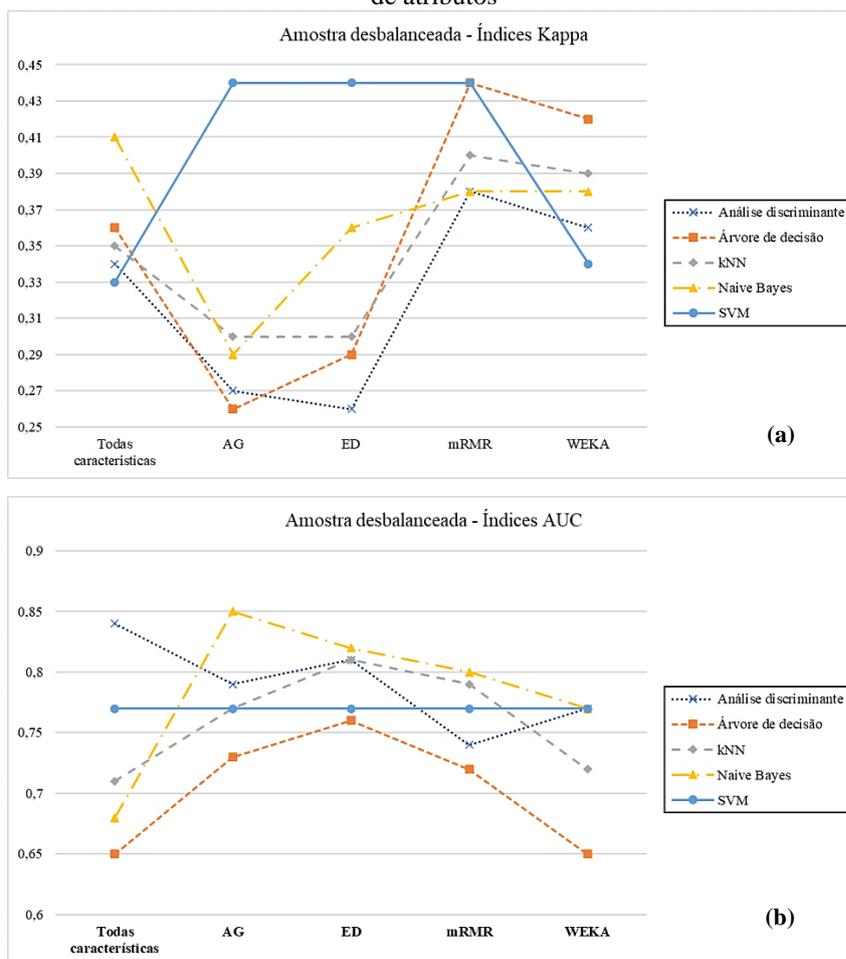
Classificadores	Seleção de atributos AG		Seleção de atributos ED	
	ACC (%)	Quant.	ACC (%)	Quant.
<b>Análise discriminante</b>	52,16	12	50,86	17
<b>Árvore de decisão</b>	49,14	16	46,12	16
<b>kNN</b>	56,47	10	50,86	18
<b>Naive Bayes</b>	50,00	7	47,41	17
<b>SVM</b>	53,00	11	53,00	17

Fonte: A Autora (2022)

\*Quant. – quantidade de características selecionadas

Os pares de características selecionadas pelo algoritmo mRMR para cada classificador foram:  $c_{13}$  e  $c_{10}$  (Análise discriminante),  $c_{14}$  e  $c_{11}$  (Árvore de decisão),  $c_5$  e  $c_3$  (kNN),  $c_7$  e  $c_6$  (Naive Bayes) e  $c_9$  e  $c_4$  (SVM). Já com o uso do WEKA, foi obtido o par  $c_6$  e  $c_8$  de características. A Figura 53 apresenta, de forma simplificada, a variação dos resultados obtidos para esta etapa do problema multiclases.

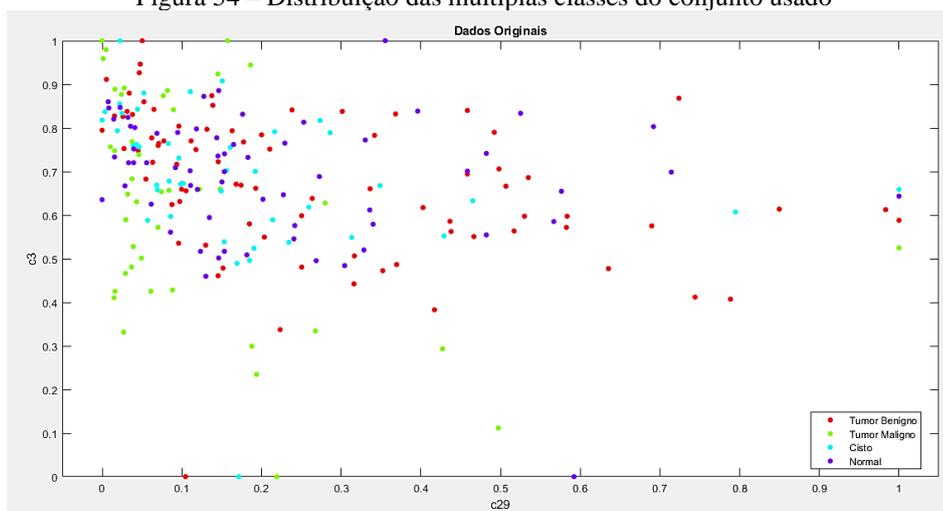
Figura 53 – Variação dos índices (a) Kappa e (b) AUC de multiclassificação em função das técnicas de seleção de atributos



Fonte: A Autora (2022)

Dentre os problemas encontrados para o aprendizado de modelos em conjuntos com múltiplas classes, destacam-se o desequilíbrio das quantidades representantes de cada classe e a separabilidade das classes em questão. Tais problemas são observados no conjunto multiclases utilizado neste trabalho (Figura 54), bem como a sobreposição entre as classes em diversas regiões do espaço de atributos. Assim, os desempenhos apresentados na Figura 54 podem ser justificados por esses fatores. Resultados similares também foram obtidos por outros pesquisadores (ARAÚJO, 2014; VILA NOVA, 2017; ALFAYEZ; EL-SOUD; GABER, 2020).

Figura 54 – Distribuição das múltiplas classes do conjunto usado



Fonte: A Autora (2022)

### 5.3.2 Base de dados balanceada

Inicialmente, a proposta direta para tratar o problema dos dados foi o balanceamento da amostra multiclasse por sobreamostragem. Diante disso, um número pequeno de instâncias por classes dificulta ainda mais o aprendizado do conceito e torna o conjunto mais suscetível aos efeitos do desbalanceamento justificando, assim, a escolha da sobreamostragem a subamostragem. Ressalta-se que o problema de multiclassificação de amostras balanceadas foi abordado com o uso de todos os atributos definidos, uma vez que das análises anteriores constatou-se que, apesar de diminuir o custo computacional e as possíveis redundâncias, a seleção não teve impacto significativo nos índices de performance dos classificadores.

#### *Estratégia A*

Nesta etapa, foram criados exemplos sintéticos para cada classe (saudável - SDV, cisto - C e lesão maligna - LM) até atingir o quantitativo de dados da classe majoritária (lesão benigna - LB), através das técnicas SMOTE, *Safe level*, *Borderline*, CBO, ASUWO e ADASYN. As amostras resultantes das sobreamostragens foram de 332 (83 LB/ 83 LM/ 83 C/ 83 SDV) para SMOTE, *Safe level*, *Borderline* e CBO; 328 (83 LB/ 81 LM/ 83 C/ 81 SDV) para ASUWO e 331 (83 LB/ 87 LM/ 86 C/ 75 SDV) para ADASYN.

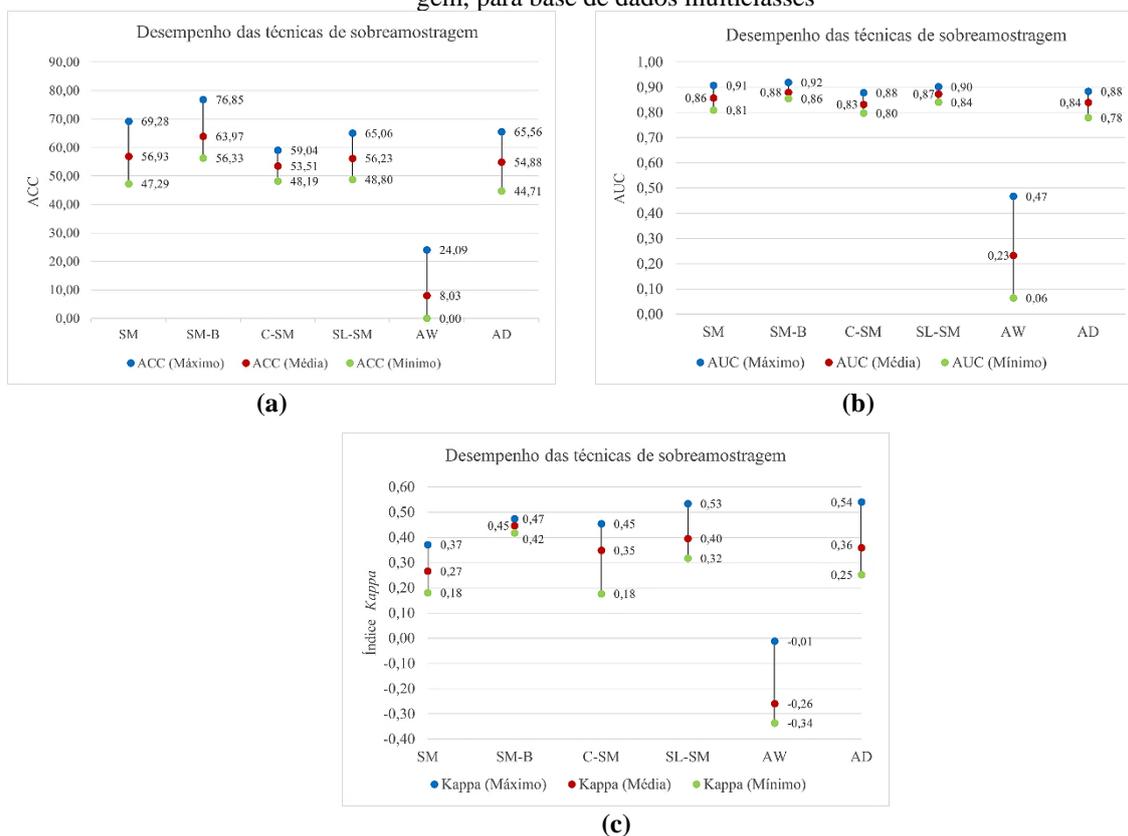
Comparados aos resultados da amostra desbalanceada, observa-se que além de não ter ocorrido aumento significativo das taxas de desempenho (Figura 55), também houve acurácias extremamente baixas (0 a 24,09) para o método de balanceamento ASUWO. Assim, apesar do aumento do número de atributos e da eliminação de um dos problemas (desbalanceamento), as

estratégias *oversampling* comumente usadas neste trabalho não demonstraram resultados satisfatórios.

Uma vez que a Estratégia A não caracterizou bem o espaço global e as áreas de vizinhança, o baixo desempenho apresentado pelos classificadores propostos pode estar associado à presença de um pequeno número de exemplos *safe*. Os resultados obtidos nessa etapa também se devem à existência da sobreposição de exemplos e de uma fronteira complicada entre classes. A performance completa dos classificadores para a base de dados balanceada pode ser encontrada no Apêndice M.

Na sequência serão discutidas possíveis soluções para o problema do pequeno número de exemplos *safe* (Estratégias B e C).

Figura 55 – Resultados da (a) ACC, da (b) AUC do (c) Índice *Kappa* utilizando os métodos de sobreamostragem, para base de dados multiclass



Fonte: A Autora (2022)

### Categorização da amostra

A fim de melhorar a técnica de balanceamento para múltiplas classes, foi feita a análise dos dados de cada classe para identificar suas propriedades individuais. Para isso, foi utilizado o método de análise dos  $k$ -vizinhos mais próximos para categorizar os exemplos das classes

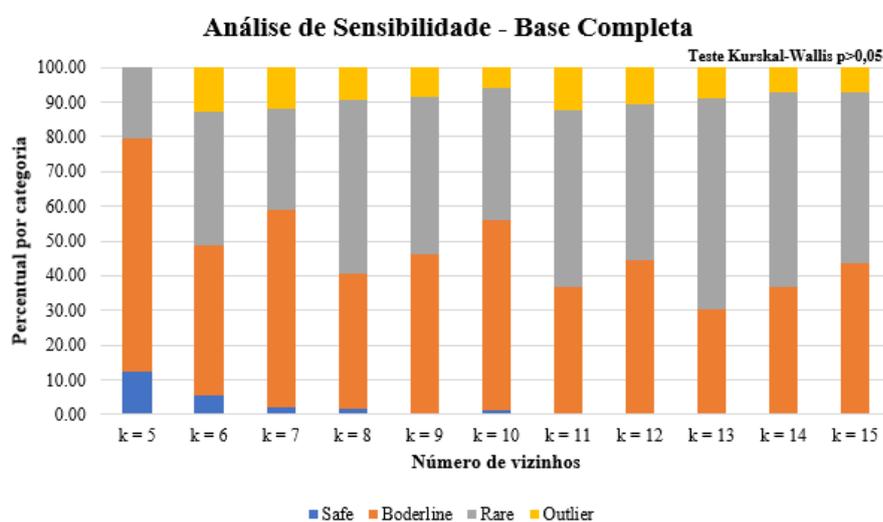
(saudável, cisto e lesão maligna) como sendo *safe*, *borderline*, *rare* ou *outlier*. As definições estabelecidas no subtópico 4.7.1.1 foram usadas para determinar a categorização.

Primeiro, foi feita uma análise de sensibilidade do método de categorização dos atributos ao tamanho da vizinhança (Figura 56). A escolha de valores de  $k$  menores que 5 ou muito altos pode gerar uma má distinção da categoria do exemplo. Em consideração a esses aspectos, foi utilizado  $k$  variando de 5 a 15.

Cada barra do gráfico (Figura 56) foi estratificada em quatro partes, representando o percentual de determinado tipo de categoria. Nota-se que para cada variação do número de vizinhos, as diferenças entre as porcentagens não são significativas (teste de Kurskal-Wallis  $p > 0,05$ ). Para alguns conjuntos de dados, por exemplo  $k = 7$  (alterações entre categorias *safe* e *borderline*) e  $k = 13$  (alterações entre categorias *rare* e *outlier*), as diferenças são maiores.

No entanto, apesar dessas diferenças entre categorias particulares, observou-se que as tendências gerais das distribuições de rótulos foram estáveis para toda base de dados. Como os resultados do método de categorização adotado não dependem significativamente do número de vizinhos  $k$ , foi escolhido  $k = 5$ .

Figura 56 – Distribuição categórica dos exemplos em função do tamanho da vizinhança



Admitindo  $k = 5$ , a análise dos exemplos de cada classe resultou na categorização da Tabela 6. Os exemplos *outliers* foram inexistentes nessa configuração. As quatro classes estudadas apresentaram grandes quantidades de exemplos difíceis (*borderline* e *rare*) em comparação a quantidade de exemplos *safe*.

Nesse cenário com poucos exemplos considerados seguros, as classes são representadas por exemplos difíceis, sendo necessário fortalecer suas regiões a partir da sua sobreamostragem

(Estratégia B). Uma outra abordagem seria a sobreamostragem da categoria *safe*, visto que a aprendizagem dos classificadores é beneficiada nos conjuntos em que exemplos seguros prevalecem (Estratégia C).

Tabela 6 – Distribuição categórica dos exemplos em função do tamanho da vizinhança

Classes	Categorização			
	<i>Safe</i>	<i>Borderline</i>	<i>Rare</i>	<i>Outlier</i>
<b>Tumor Maligno</b>	12	18	12	0
<b>Tumor Benigno</b>	9	61	14	0
<b>Cisto</b>	0	28	15	0
<b>Saudável</b>	9	49	6	0

Fonte: A Autora (2022)

### *Estratégia B*

Aqui o objetivo foi avaliar a sobreamostragem através do fortalecimento da fronteira de decisão, do aumento da quantidade de exemplos *safe* e do aumento da quantidade de exemplos *rare*, na tentativa que esses exemplos (*rare*) possam ser categorizados como *safe*.

Para isso, foram aplicados os balanceamentos que tiveram melhor performance – *Safe level* SMOTE e CBO - na sobreamostragem dos elementos *safe* e *rare*, respectivamente. Os elementos de fronteira foram sobreamostrados através da técnica *Borderline*-SMOTE. Após esse procedimento, a base de dados totalizou em 334 amostras (83 LB/ 84 LM/ 83 C/ 84 SDV). A Tabela 7 mostra os resultados encontrados após a classificação da base de dados.

Tabela 7– Resultados da classificação multiclassificada da amostra balanceada (aumento proporcional dentro de cada categoria)

Medidas de performance	Métodos de aprendizagem				
	Análise discriminante	Árvore de Decisão	kNN	Naive Bayes	SVM
Acurácia (%)	50,00	42,22	68,56	42,22	56,89
Sensibilidade (%)	49,95	42,16	68,52	42,12	56,79
Especificidade (%)	83,34	80,74	89,52	80,73	85,62
Coefficiente kappa	0,33	0,23	0,58	0,23	0,42
AUC	0,87	0,80	0,91	0,79	0,87

Fonte: A Autora (2022)

A Estratégia B apresentou desempenho similar às técnicas tradicionais de balanceamento, sendo inferior, de modo geral, ao obtido pela base balanceada por SMOTE e *Borderline* (ver Apêndice N). O teste  $\kappa$  de concordância permaneceu na faixa de concordância leve, com destaque para o classificador kNN que atingiu concordância moderada. Esse último dado já era previsível pelo fato do classificador kNN possuir a maior acurácia entre os métodos (68,56%), sendo considerada razoável.

Apesar do aumento de exemplos *safe* e do fortalecimento dos limites de decisão a partir da sobreamostragem dos exemplos limítrofes, a performance dos classificadores não foi otimizada. O que pode ser justificado pelo fato de a amostra final não ter gerado porcentagem considerável, maior que 50%, para os elementos *safe* (14,37%). A nova configuração possui uma quantidade de elementos *borderline* (64,37%) que pode ter sido gerada pela mudança de categorização de elementos importantes.

### *Estratégia C*

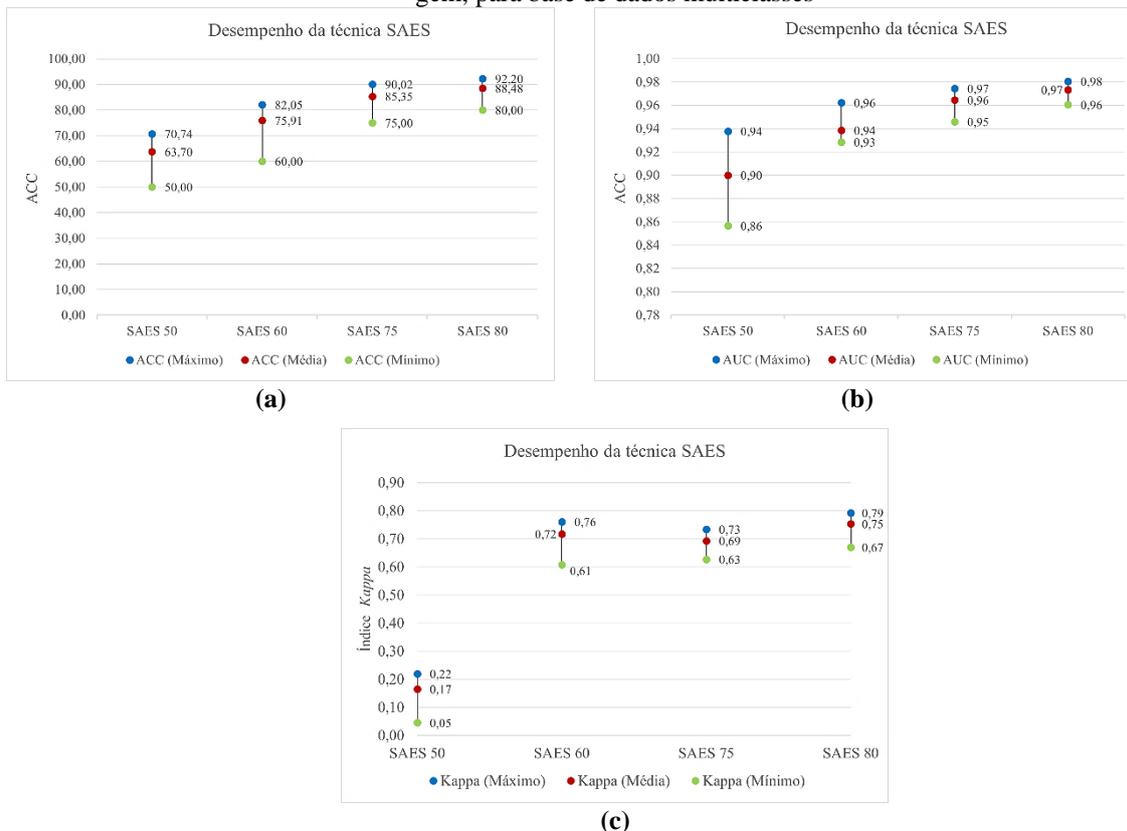
Identificação dos elementos *safe* da amostra e sobreamostragem utilizando a técnica SMOTE e um percentual de geração de vetores sintéticos seguros de 50, 60, 75 e 80%, em relação as demais categorias. A técnica foi denominada de SAES - SMOTE Aplicado em Elementos *Safe*. No Apêndice N encontram-se todas as taxas calculadas para os classificadores.

À medida que a porcentagem de elementos *safe* aumentou dentro da classe, o desempenho dos classificadores melhorou (Figura 57). O classificador SVM, por exemplo, alcançou acurácia de 92,20% com índice ótimo de concordância  $\kappa$ . Quando comparados com os resultados das estratégias anteriores, também se observaram melhorias no desempenho dos algoritmos de aprendizado de máquina.

Foi estudada a quantidade de acertos pelo classificador SVM (melhor desempenho) de elementos *borderline* e *rare* antes e depois da aplicação da técnica SAES 80%.

Assim sendo, verificou-se que a amostra original desbalanceada (233 dados) apresentou acurácias de 31,91% (15/47) e de 36,54% (57/156) para os elementos *rare* e *borderline*, respectivamente. Após o balanceamento (1.552 dados), a acurácia para o *rare* foi de 33,22% (99/298) e para o *borderline* 66,20% (333/503). Além disso, considerando os elementos da base original, observaram-se acertos de 61,70% (29/47) e de 72,44% (113/156) para os elementos *rare* e *borderline*, respectivamente. Esses valores confirmaram que a sobreamostragem de elementos *safe*, melhoraram o aprendizado global, incluindo outras categorias consideradas de difícil aprendizado.

Figura 57 – Resultados da (a) ACC, da (b) AUC do (c) Índice *Kappa* utilizando os métodos de sobreamostragem, para base de dados multiclases



Fonte: A Autora (2022)

### 5.3.3 Base de dados normalizada

Os testes realizados com os diferentes métodos de balanceamento geraram resultados abaixo do esperado, quando comparados aos resultados da mesma etapa na classificação binária dos objetos. Além da complexidade inerente ao problema de classificação multiclases, o balanceamento não minimizou o problema da separabilidade das classes.

A distribuição dos valores para cada atributo foi representada em histogramas (APÊNDICE O). Da análise gráfica, observou-se que os atributos possuem escalas e magnitudes da escala distintas, o que pode ser um problema para a aprendizagem dos classificadores, principalmente para aqueles que dependem da distância (kNN e SVM). Isso porque há uma certa tendência de que um peso maior seja dado aos atributos com maior magnitude, tornando os algoritmos de aprendizado citados enviesados para eles. Além disso, a maioria dos fluxos de trabalho dos modelos de aprendizado de máquina assume que as distribuições dos dados de entradas são gaussianas (HA et al., 2021). Porém, observou-se que quase todos os atributos utilizados no presente trabalho possuem distribuições distorcidas, sugerindo necessidade de remodelagem na distribuição de dados.

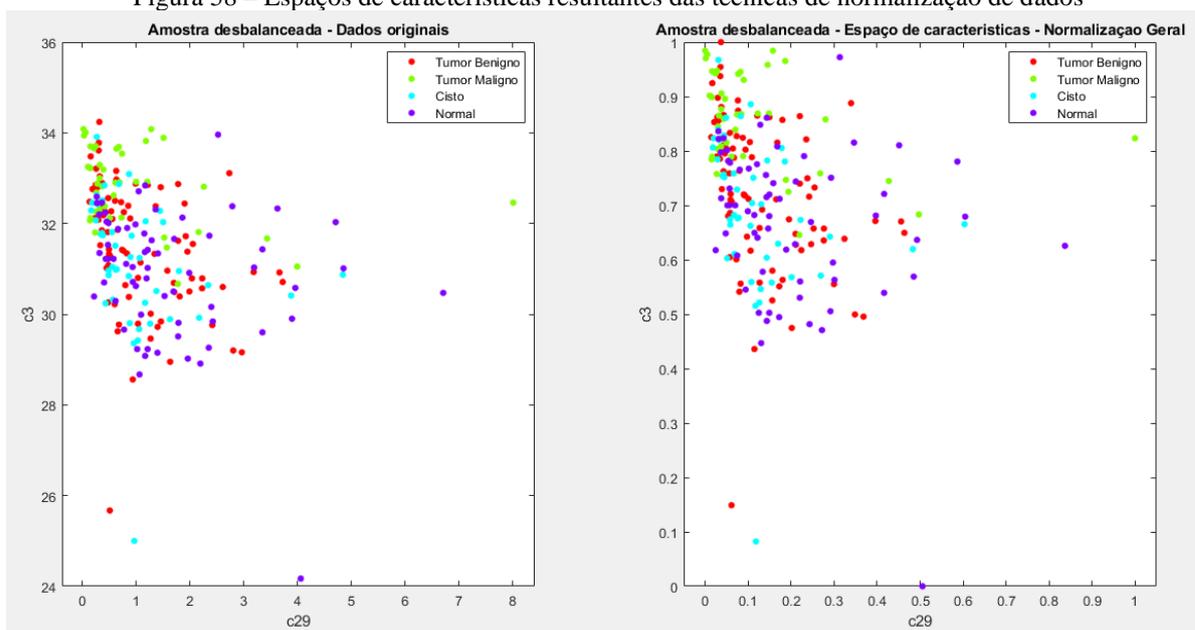
A normalização dos dados foi aplicada com o intuito de possibilitar uma comparação adequada entre os métodos, uma melhor separabilidade das classes e, conseqüentemente, uma otimização do reconhecimento de padrões. Neste trabalho, foram implementados os métodos de normalização descritos no Quadro 15. E cada método foi aplicado à base de dados completa e a cada classe da amostra.

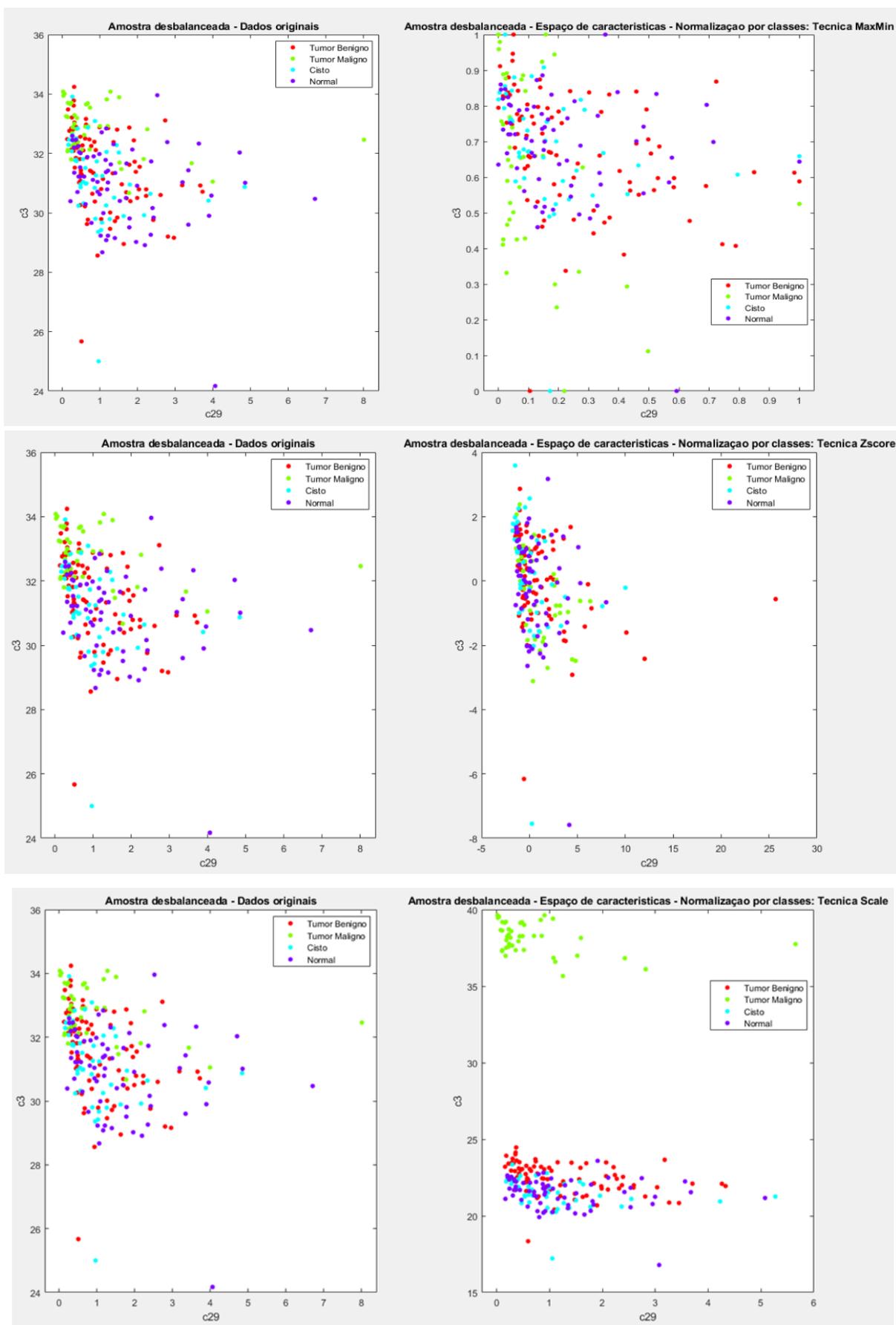
Inicialmente, os métodos de normalização adotados foram avaliados através da comparação entre o espaço de características original e os espaços de características resultante da aplicação de cada uma das metodologias. Os resultados são mostrados na Figura 58.

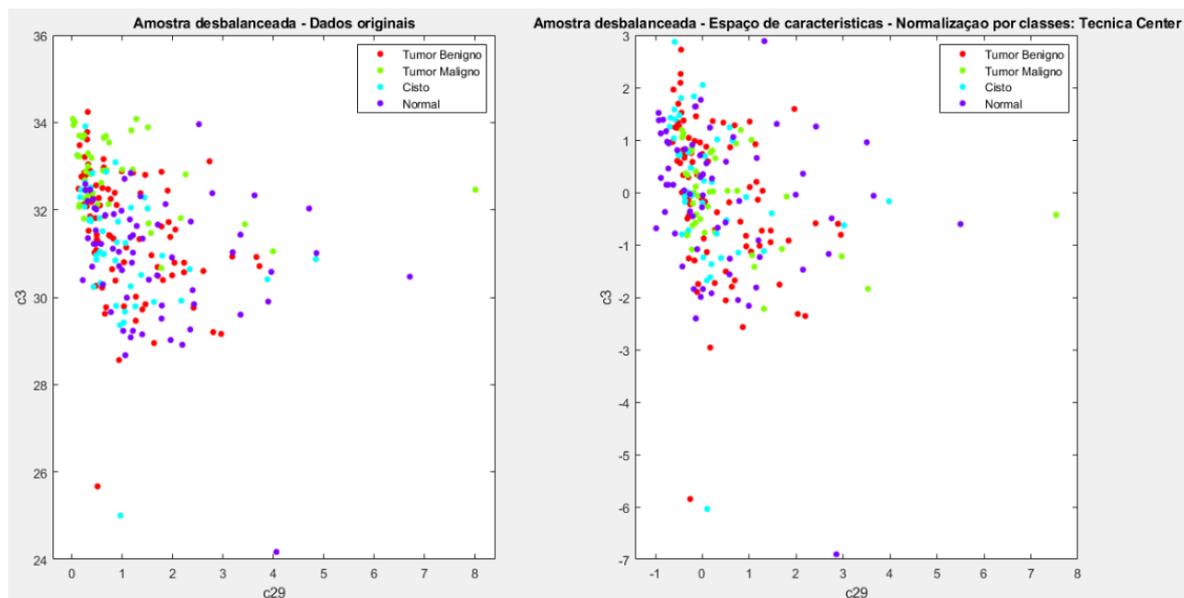
A normalização aplicada a base de dados completa (“Normalização geral”) não apresentou diferenças na distribuição deles quando comparada com a distribuição da base original (“Sem normalização”). Isso ocorreu, uma vez que, ao realizar a normalização utilizando a base completa, a escala foi redimensionada, porém as diferenças de magnitudes entre as características permaneceram iguais.

As normalizações por classe com uso das técnicas *Zscore* e *Center* aproximaram os *clusters* das diferentes classes, o que reduz a confiança de um algoritmo supervisionado de aprendizado de máquina para distinguir entre as classes. Mais à frente será vista a influência da diminuição da distância entre os grupos na performance dos classificadores a essas bases. Seguindo o mesmo raciocínio, as técnicas de normalização *Max/Min* e *Scale* melhoraram a separabilidade entre as classes, com destaque para o método *Scale*. Isso reflete na performance dos classificadores, como será visto adiante.

Figura 58 – Espaços de características resultantes das técnicas de normalização de dados



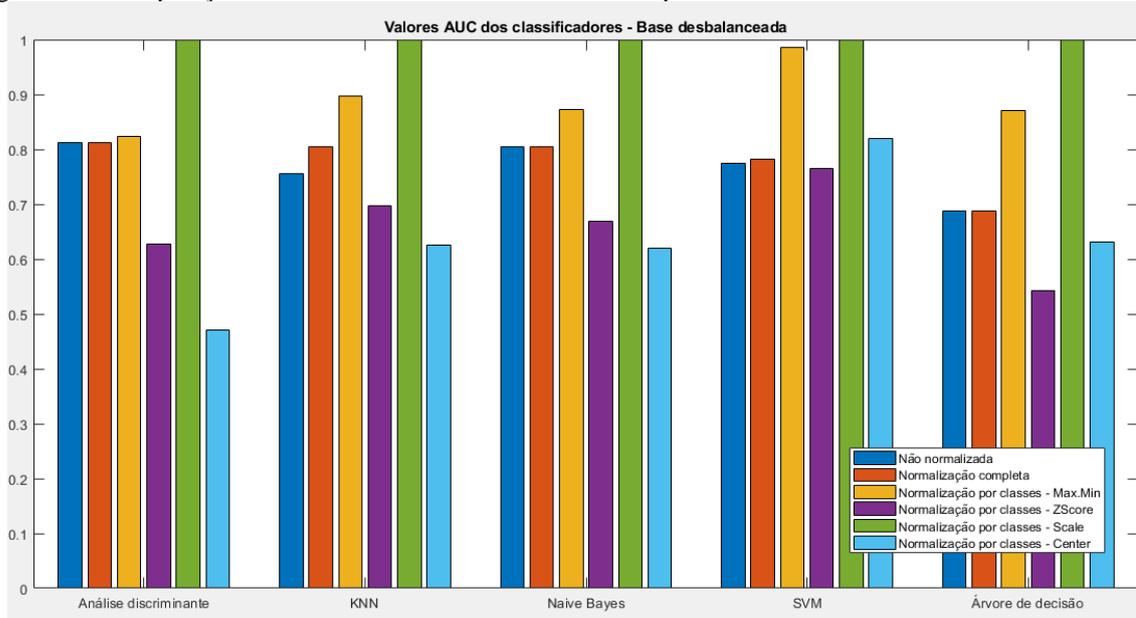




Fonte: A Autora (2022)

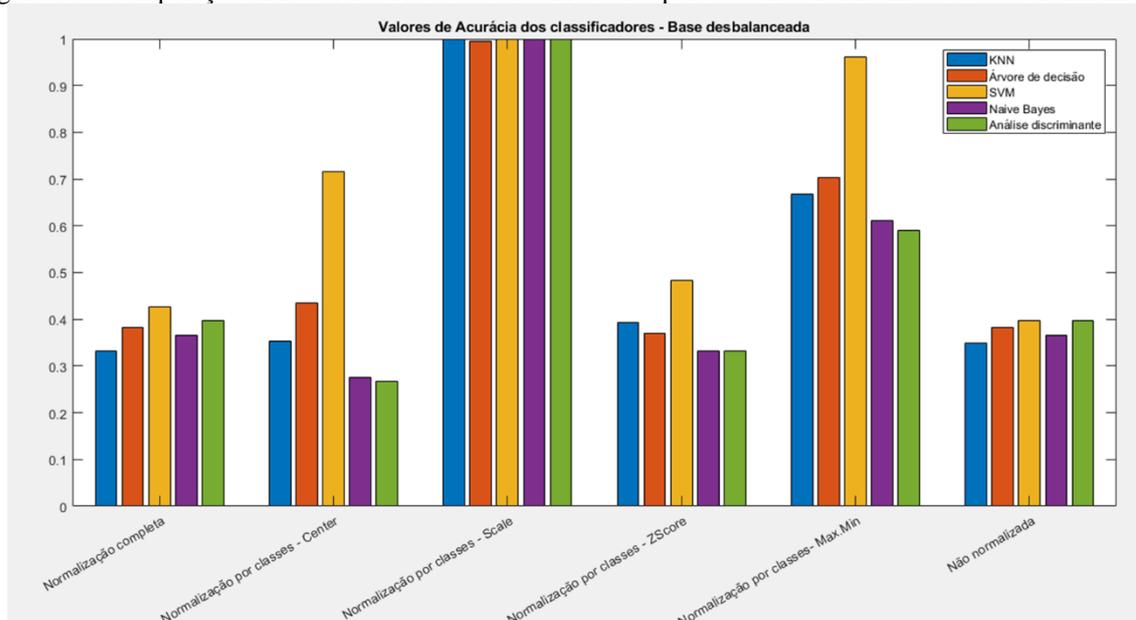
Em seguida, foram aplicados a cada uma das bases normalizadas os métodos de classificação, com validação cruzada de *10-folds*, adotados nesta pesquisa. As métricas de performance podem ser encontradas detalhadamente no Apêndice O. As Figuras 59 e 60 mostram, respectivamente, uma comparação entre os valores AUC e de acurácia obtidos. As bases cujo problema da separabilidade foi reforçado apresentaram pior desempenho, com acurácias consideradas insignificantes, segundo os valores  $\kappa$  de concordância. O método Max/Min apresentou uma acurácia de 96,12% para o classificador SVM, no entanto para os demais classificadores houve uma diminuição deste valor com índices  $\kappa$  que indicam uma concordância aleatória. O resultado do classificador SVM pode estar mais relacionado a um sobreajustamento ao invés de valores de performance verdadeiramente confiáveis, uma vez que para os demais classificadores houve consistência quanto a um  $\kappa$  baixo.

Figura 59 – Comparação entre valores AUC dos classificadores para base não normalizada e bases normalizadas



Fonte: A Autora (2022)

Figura 60 – Comparação entre valores ACC dos classificadores para base não normalizada e bases normalizadas



Fonte: A Autora (2022)

Como a base desbalanceada e normalizada através da metodologia por Escala (Scale) apresentou valores de 100% (concordância perfeita) para a ACC e para a AUC (Tabela 8), foi realizada análise da presença ou não de *overfitting* no resultado. Foram avaliados os classificadores com os hiperparâmetros descritos no Quadro 13 e estes mesmos classificadores com os hiperparâmetros otimizados para a base de dados normalizadas por *Scale*. A existência de *overfitting* indica uma adequação excessiva do classificador à base de treino, não respondendo adequadamente a exemplos fora desse universo. Assim, para os classificadores cujo teste foi

positivo, houve exclusão desses. Ressalta-se que o uso dos classificadores com hiperparâmetros otimizados foi feito como estratégia de dupla verificação. Ou seja, o uso desses classificadores foi feito com o objetivo de avaliar a consistência das diferenças já analisadas pelos testes estatísticos, confirmando ou não o *overfitting*.

Tabela 8 – Teste *overfitting* em amostra desbalanceada e normalizada pelo método Escala (*Scale*)

Classificador	Teste Kolmogorov-Smirnov		Teste de hipóteses		
	Valor-p	Distribuição normal?	Teste de Wilcoxon amostra pareada	Teste t amostra pareada	Possibilidade <i>overfitting</i> ?
<b>kNN</b>	<0,010	Não	0,755	-	Não
<b>kNN otimizado</b>	<0,010	Não	1,000	-	Não
<b>Árvore de decisão</b>	<0,010	Não	0,588	-	Não
<b>Árvore de decisão otimizado</b>	<0,010	Não	0,588	-	Não
<b>SVM</b>	>0,100	Sim	-	0,330	Não
<b>SVM otimizado</b>	>0,100	Sim	-	0,330	Não
<b>Naive Bayes</b>	>0,100	Sim	-	0,330	Não
<b>Naive Bayes otimizado</b>	>0,100	Sim	-	0,330	Não
<b>Análise discriminante</b>	>0,100	Sim	-	0,330	Não
<b>Análise discriminante otimizado</b>	>0,100	Sim	-	0,330	Não

Fonte: A Autora (2022)

Foi possível confirmar que o procedimento de normalização *Scale* é eficaz na separabilidade das classes, comprovando sua capacidade de prever a presença ou a ausência de patologias mamárias em imagens IRs com altíssima precisão e sem *overfitting*. No entanto, esse resultado não pode ser estendido à classificação não supervisionada (análise individual de pacientes), visto que a normalização é feita tomando como referência valores de um grupo previamente conhecido.

Ressalta-se também que os algoritmos evolutivos (Algoritmo Genético e Evolução Diferencial) foram usados, anteriormente, como ferramentas de busca do conjunto de características que minimizassem o erro de generalização ou, com outras palavras, maximizassem a acurácia dos classificadores. No problema multiclases, estes algoritmos foram dispensáveis para a base com classes individualmente normalizadas através do método *Scale*, pois a acurácia da classificação, para este grupo, foi máxima.

Avaliando as metodologias para classificação multiclases, os melhores resultados foram obtidos através do emprego da técnica SAES (balanceamento por sobreamostragem de elementos seguros) e da técnica de normalização *Scale* na amostra desbalanceada. Assim sendo, das tabelas e gráficos anteriores, para a base balanceada, destacaram-se positivamente os classificadores SVM e Naive Bayes ambos com acurácia de 92,20%. E para a base normalizada com a técnica *Scale*, destacaram-se todos os classificadores com acurácia 100%. Apesar dessas performances idênticas, para ambas as metodologias de pré-processamento foi escolhido o classificador SVM por ser comumente utilizado em problemas de avaliação de imagens termográficas de mamas.

Em relação à comparação entre os melhores resultados obtidos nos experimentos descritos anteriormente (destacado em **negrito**) e os resultados apresentados na literatura e nos trabalhos anteriores do grupo de pesquisa (destacados de *itálico*), nota-se que a performance do classificador SVM para a base balanceada foi superior à dos trabalhos encontrados, com exceção do trabalho de Alfayez, El-Soud e Gaber (2020) com a técnica de classificação ELM. Para a base desbalanceada e normalizada, a performance foi máxima e superior a todas listadas no Quadro 22.

Quadro 22 – Síntese dos resultados de classificação multiclases de imagens termográficas mamárias de trabalhos desenvolvidos no grupo de pesquisa e encontrados na literatura

<b>Autor</b>	<b>Base</b>	<b>Classificador</b>	<b>ACC (%)</b>	<b>Sens. (%)</b>	<b>Esp. (%)</b>
<i>Araújo (2014)</i>	50	<i>Distância Mahalanobis</i>	84,00	85,70	86,50
<i>Dourado (2014)</i>	234	<i>SVM</i>	89,36	87,50	90,32
<i>Vila Nova (2017)</i>	87	<i>Distância Mahalanobis Multiclassificadores</i>	72,71 71,15	96,30 87,18	- -
Santana et al. (2018)	825	ELM MLP	71,22 76,01	-	-
<i>Vasconcelos, Santos e Lima (2018)</i>	175	<i>SVM</i>	63,46	80,77	86,54
Gogoi et al. (2019)	60	SVM	83,22	85,56	73,23
Alfayez, El-Soud e Gaber (2020)	1.345	MLP ELM	80,04 99,10	84,00 98,05	61,60 97,03
<b>Metodologia de Classificação Multiclases (Base Balanceada)</b>	<b>1.552</b>	<b>SVM</b>	<b>92,20</b>	<b>92,32</b>	<b>97,40</b>
<b>Metodologia de Classificação Multiclases (Base Normalizada)</b>	<b>232</b>	<b>SVM</b>	<b>100</b>	<b>100</b>	<b>100</b>

Fonte: A Autora (2022)

#### 5.4 ANÁLISE DA PREVISÃO DO TAMANHO E PROFUNDIDADE DO TUMOR MALIGNO

As simulações dos diferentes tipos de mamas (Extremamente Densa-ExD, Heterogeneamente Densa- HD; Parcialmente Fibroglandular-PF e Predominantemente Gordurosa-PG) foram usadas para gerar modelos de regressão. Os desempenhos dos modelos foram determinados pelos coeficientes  $R^2$  e REMQ (raiz do erro médio quadrático). As imagens termográficas do banco de dados cujas pacientes possuem exames com informações de tamanho e localização da lesão foram usadas para verificar a eficácia do modelo. Para ampliar os testes de eficácia, também foram usadas imagens IRs de pacientes do banco de dados da Universidade Federal Fluminense (UFF). É importante ressaltar que da base de dados da UFF foram selecionadas as imagens adquiridas através do protocolo estático, semelhante ao que é adotado pelo grupo de pesquisa da UFPE.

##### 5.4.1 Avaliação das características

Foram extraídas 17 características da base de imagens IRs formadas por pacientes com mamas saudáveis (64 imagens) e com mamas acometidas por tumor maligno (42 imagens). Todo estudo estatístico (Tabela 9) foi realizado no *software* Microsoft Excel<sup>®</sup>, disponível no Office365. As características foram testadas quanto à normalidade por meio do teste Kolmogorov-Smirnov. Para análise da diferença estatística, foram usados os testes  $t$  (distribuição

paramétrica) e Wilcoxon (distribuição não-paramétrica). O nível de significância foi definido para um valor  $p < 0,05$ . Todos os valores  $p$  apresentados são bicaudais.

Tabela 9 – Estudo da diferença estatística entre as características extraídas de mamas saudáveis e de mamas com tumor maligno

Característica	Média para mama saudável	Média para mama com TM	Valor de p
$c_1$	35,528	34,700	$p < 0,0001^a$
$c_2$	35,197	37,088	$p < 0,0001^b$
$c_3$	32,801	30,908	$p < 0,0001^b$
$c_4$	32,105	30,405	$p < 0,0001^a$
$c_9$	33,978	32,435	$p < 0,0001^a$
$c_{10}$	0,604	0,754	$p < 0,0001^a$
$c_{19}$	-0,544	-0,008	$p < 0,005^a$
$c_{20}$	5,029	4,408	$p = 0,5352^a$
$c_{21}$	33,939	32,451	$p = 0,0001^a$
$c_{22}$	2,749	2,397	$p = 0,005^a$
$c_{23}$	0,183	0,155	$p < 0,0001^b$
$c_{24}$	0,986	0,984	$p < 0,0001^a$
$c_{25}$	0,979	0,984	$p < 0,0001^a$
$c_{26}$	0,515	0,575	$p = 0,9404^a$
$c_{27}$	0,396	0,593	$p = 0,9913^a$
$c_{28}$	-0,147	0,056	$p = 0,9382^a$
$c_{29}$	0,975	1,646	$p = 0,7948^a$

Fonte: A Autora (2022)

a – Teste *t de Student*; b – Teste Wilcoxon

Com exceção das características  $c_{19}$  (assimetria),  $c_{20}$  (curtose),  $c_{26}$  (energia) e  $c_{27}$  a  $c_{29}$  (momentos estatísticos), foi constatada uma diferença estatisticamente significativa entre os demais atributos avaliados. Numa segunda análise, foram desconsideradas as características  $c_1$  a  $c_4$  e  $c_{22}$  a  $c_{25}$  pela inviabilidade de extração dos dados simulados. Essa decisão foi tomada pois o cálculo do primeiro grupo de características envolve a comparação entre as temperaturas máxima e mínima de mamas direita e esquerda, e o no cálculo do segundo grupo são feitas combinações de diferentes níveis de cinza.

Assim, o grupo foi reduzido para 4 características:  $c_9$  (média),  $c_{10}$  (desvio padrão),  $c_{21}$  (mediana) e, incluiu-se, a temperatura máxima ( $c$ ). Além de significantes estatisticamente, esses atributos carregam por definição valores que representam adequadamente um conjunto de dados quanto à uniformidade de distribuição, a presença de valores mais altos de temperatura e à simetria térmica.

### 5.4.2 Previsão do tamanho do tumor maligno

Os modelos de regressão foram treinados no aplicativo *Regression Learner* (Matlab®). Foram definidos como entrada e saída dos modelos, respectivamente, a matriz  $X = [c, c_{10}, c_{21}, c_{22}]$  e vetor  $Y_t$  de tamanhos simulados, escolhendo-se os modelos que apresentaram, para cada tipo de densidade mamária, o maior  $R^2$  e o menor REQM. Os modelos gerados a partir de todas as variáveis independentes serão chamados de modelos completos.

Na análise inicial da regressão SVM para a mama com densidade ExD foi verificado  $REQM > 4$  para os diferentes tipos de kernel avaliados, o que não atende ao critério de seleção determinado anteriormente. Por esse motivo, foram testados outros modelos disponíveis no aplicativo *Regression Learner* (Matlab®) e a técnica que apresentou resultados adequados foi a de Regressão por Processo Gaussiano (RPG).

De forma geral, processos gaussianos são caracterizados por gaussianas multivariadas. Uma função  $f$  é um processo gaussiano se qualquer conjunto finito  $\{f(x_1), f(x_2), \dots, f(x_n)\}$  segue uma distribuição normal multivariada, onde  $x$  é um vetor de dados observados. Igualmente a uma gaussiana, o processo gaussiano é definido por uma função média, assumida como zero, e por uma função *kernel* (função de covariância).

Para esse caso, duas funções RPG, com todas as variáveis preditoras, apresentaram  $RMEQ \leq 4$ : RPG com função base exponencial e RPG com função base constante (gerada pela otimização de hiperparâmetros). Os modelos RPG com duas variáveis preditoras foram descartados por apresentarem desempenhos inferiores aos modelos completos. Com aplicação da base teste às funções preditoras iniciais (com quatro variáveis), verificou-se um  $R^2 = 1$  com  $RMSE = 0,34$  (RPG função base exponencial) e com  $RMSE = 0,21$  (RPG função base constante), o que certificou os modelos preditores. Assim, por apresentar menor RMSE, o modelo adotado para a mama tipo ED foi o de quatro variáveis preditoras e com os hiperparâmetros otimizados (Tabela 10).

Para a mama com densidade HD, foram avaliados modelos de regressão baseados em SVM e foram encontrados coeficientes de determinação e REQM com variação entre 0,81–0,99 e 0,71–2,74, respectivamente. Os modelos com dois preditores não obtiveram performance tão satisfatória quanto os que usaram todos os preditores, pois apresentaram erros maiores que quatro. Com aplicação da base teste, foram encontrados valores de  $R^2 \geq 0,94$  e de  $RMSE < 1,88$ , evidenciando ajuste adequado. Com isso, o modelo SVM selecionado foi o de *kernel* quadrático (Tabela 10) gerado com as quatro variáveis independentes.

Considerando a mama com densidade PF, o coeficiente de determinação e o REQM variaram, respectivamente, entre 0,78 – 0,88 e 2,35 – 3,13 (modelo completo) e 0,74 – 0,79 e 3,08 – 3,44 (modelo com variáveis selecionadas pelo mRMR). Para ambos os grupos, destacam-se os modelos com hiperparâmetros otimizados cujas funções kernel foram a quadrática (modelo completo) e a linear (modelo com duas variáveis). Esses modelos foram avaliados na base teste de 20% e apresentam  $R^2 > 0,90$  e  $REQM < 1,50$ , certificando um bom ajuste dos modelos preditivos. Dessa forma, foi determinado o modelo gerado com quatro variáveis preditoras e com os hiperparâmetros otimizados (Tabela 10).

A mama com densidade PG apresentou coeficiente de determinação variando entre 0,76 e 0,82 (modelo completo), e 0,81 e 0,86 (modelo com variáveis selecionadas pelo mRMR). O REQM variou entre 2,48 e 3,45 (modelo completo); e 2,63 e 3,07 (modelo com variáveis selecionadas por mRMR). Novamente, os modelos com hiperparâmetros otimizados possuíram melhor performance.

Na aplicação da base teste os modelos completos apresentaram  $R^2 > 0,83$  e  $REQM < 2,98$ , confirmando modelos preditivos com ajustes apropriados. Entretanto, nenhum modelo de duas variáveis apresentou REQM (teste) tão baixo quanto seu REQM de treino. Por esses motivos, foi escolhido para mama tipo PG o modelo com quatro variáveis e com os hiperparâmetros otimizados (Tabela 10).

Tabela 10 – Características dos modelos preditivos para o tamanho do tumor maligno e seus valores de performance

Tipo de mama	Modelo preditivo	C	$\epsilon$	$REQM_{\text{treino}}$	$REQM_{\text{teste}}$	$R^2_{\text{treino}}$	$R^2_{\text{teste}}$
<b>ED</b>	RPG	-	-	3,22	0,21	0,77	1,00
<b>HD</b>	SVM Quadrático	7,88	0,69	0,71	0,58	0,99	0,99
<b>PF</b>	SVM Quadrático	0,06	0,65	2,35	1,37	0,88	0,93
<b>PG</b>	SVM Linear	1,72	0,02	2,76	2,48	0,82	0,85

Fonte: A Autora (2022)

### *Análise de casos – tamanho do tumor maligno*

Foram avaliadas 30 imagens termográficas de pacientes com lesão maligna, das quais 9 foram caracterizadas com mamas ExD, 12 com mamas HD, 5 com mamas PF e 4 com mamas PG. A fim de estudar a reprodutibilidade e a repetibilidade (R&R) do método, dois usuários realizaram o procedimento para predição do tamanho da lesão no grupo teste.

Considerando o universo amostral, a taxa de acerto da mama acometida pela lesão maligna foi de 79% (Usuário 1) e de 93% (Usuário 2), mostrando que a análise da temperatura média pode indicar a presença de anormalidade. Em relação as estimativas do diâmetro do tumor maligno, os resultados são apresentados na Tabela 11.

A reprodutibilidade e a repetibilidade são dois componentes de precisão de um dado sistema. O estudo R&R foi feito no *software* Minitab® - versão 18 e, como resposta o *software* fornece valores percentuais referentes aos componentes de variância responsáveis por avaliar a contribuição da quantidade de variação por cada fonte de erro do procedimento.

Sendo assim, as componentes de variância da repetibilidade e da reprodutibilidade foram, respectivamente, 98,75% e 1,25%. Esse resultado sugere que o sistema de predição apresenta capacidade de fornecer resultados consistentes independente do usuário. No entanto, há uma variabilidade alta na seleção da ROI para estimativa, o que indica uma necessidade de aperfeiçoamento da técnica de segmentação por parte dos usuários.

A importância em se estimar o tamanho de um tumor maligno está na determinação do estágio da doença e se a paciente precisará de terapias neoadjuvantes. Por exemplo, se um tumor possui diâmetro de 10 mm e a estimativa fornece um valor de 5 mm, há um comprometimento do planejamento terapêutico ou do prognóstico da paciente. Tal problema não é inerente apenas à termografia, visto que as medidas originais utilizadas neste trabalho foram obtidas de imagens de ultrassonografia e os operadores podem cometer erro de sub ou superestimativas dos diâmetros. Nesse ponto ressalta-se que uma comparação correta entre valores (observados *versus* estimados) só aconteceria se fossem fornecidas as medidas da peça pós-cirurgia de retirada.

Pelos dados da ultrassonografia, a maioria das pacientes (53,33%) se encontravam no Estágio I de desenvolvimento (tumores com diâmetro < 20 mm), o que altera os padrões de temperatura por possuir um baixo contraste na distribuição de temperatura da superfície. Em seus trabalhos, Amri e Saidane (2011) e Ng e Sudharsan (2001) concluíram que o tamanho do tumor não tem muita influência em profundidades maiores que 20 mm, e que, inclusive, profundidades acima de 38 mm podem não ser detectadas pela câmera termográfica. Isso confirma o fato de que a profundidade do tumor tem impacto maior na distribuição da temperatura superficial em comparação com o diâmetro dele.

Tabela 11– Resultados das estimativas de tamanho de tumores malignos usando imagens por infravermelho

<b>Tipo de densidade mamária: Extremamente densa</b>		
<b>Original (mm)</b>	<b>Usuário 1 (mm)</b>	<b>Usuário 2 (mm)</b>
9,00	11,62	14,14
10,00	11,31	10,24
12,00	25,34	27,39
20,00	14,39	17,79
20,00	28,42	28,97
30,00	15,87	17,80
30,00	14,70	17,78
30,00	15,95	27,80
60,00	20,01	21,99
<b>Tipo de densidade mamária: Heterogeneamente densa</b>		
<b>Original (mm)</b>	<b>Usuário 1 (mm)</b>	<b>Usuário 2 (mm)</b>
6,00	2,99	6,31
15,00	13,16	9,12
20,00	29,81	14,92
20,00	40,79	50,83
28,00	26,37	30,04
40,00	80,57	58,02
45,00	94,22	78,60
50,00	28,71	4,19
50,00	23,09	27,94
50,00	55,16	28,33
60,00	79,29	52,72
70,00	78,20	53,06
<b>Tipo de densidade mamária: Parcialmente fibroglandular</b>		
<b>Original (mm)</b>	<b>Usuário 1 (mm)</b>	<b>Usuário 2 (mm)</b>
7,00	18,42	18,61
10,00	3,42	6,34
15,00	14,56	6,48
20,00	32,40	26,45
50,00	38,25	6,64
<b>Tipo de densidade mamária: Predominantemente gordurosa</b>		
<b>Original (mm)</b>	<b>Usuário 1 (mm)</b>	<b>Usuário 2 (mm)</b>
7,00	5,47	4,79
20,00	31,20	25,48
20,00	40,85	33,61
60,00	43,34	42,83

Fonte: A Autora (2022)

### 5.4.3 Previsão da profundidade do tumor maligno

Seguindo a mesma metodologia adotada para estimativa da lesão maligna, os modelos de regressão para a profundidade do tumor foram treinados no aplicativo *Regression Learner*

(Matlab<sup>®</sup>), tendo a matriz  $X$  como entrada e como saída o vetor  $Y_p$  de profundidades simuladas. Foi utilizado o mesmo critério para avaliação da performance e seleção dos modelos.

A partir dos dados simulados da mama com densidade ExD, foram avaliados os modelos completos de regressão baseados em SVM e os melhores resultados foram para as funções de *kernel* cúbico ( $R^2 = 0,45$  e  $REQM = 3,21$ ) e de *kernel* gaussiano ( $R^2 = 0,37$  e  $REQM = 3,44$ ). Apenas um modelo com dois preditores apresentou performance adequada: SVM com função *kernel* gaussiana. Com aplicação da base teste, o modelo completo com função *kernel* cúbica foi descartado por apresentar um  $REQM_{teste} \ll REQM_{treino}$ . Dessa forma, por apresentar melhor desempenho, foi escolhido para a mama ED o modelo SVM com duas variáveis selecionadas ( $T_{máx}$  e  $T_{méd}$ ) com valores de  $R^2 \geq 0,94$  e de  $RMSE < 1,88$ , evidenciando ajuste adequado. Com isso, o modelo SVM selecionado foi o de *kernel* quadrático (Tabela 12) gerado com duas variáveis independentes.

Considerando a mama com densidade HD, o coeficiente de determinação e o  $REQM$  variaram, respectivamente, entre 0,24 e 0,66, e entre 2,35 e 3,50 (modelo completo) e 0,35 – 0,71 e 2,18 – 3,24 (modelo com variáveis selecionadas pelo mRMR). Os modelos que apresentaram  $R^2$  baixos (menores que 0,5) foram desconsiderados antes da aplicação do grupo teste. Na avaliação dos modelos com a base teste de 20% observou-se que muitos deles apresentaram uma  $RMSE_{teste}$  maiores que os de  $RMSE_{treino}$ , invalidando-os. Aqueles que certificaram bom ajuste dos modelos preditos apresentaram, de forma geral,  $R^2 \geq 0,50$  e  $REQM \leq 2,93$ . Como o modelo completo foi validado e apresentou melhor performance, ele foi selecionado para as predições de mamas com densidade HD (Tabela 12).

Para a mama com densidade PF, os modelos SVM com todas as variáveis predictoras e função *kernel* gaussiana com escalas 0,5 e 8,0 apresentaram  $RMEQ \leq 4$ . No entanto, os valores de  $R^2$  foram baixos (0,36 e 0,24 respectivamente), sugerindo uma baixa correlação entre os modelos criados e os dados. Os modelos SVM com variáveis selecionadas pelo método mRMR não foram considerados, pois apresentaram  $RMEQ$  entre 4,20 e 4,40. Por esses motivos, foram testados outros modelos disponíveis no aplicativo e a técnica que apresentou resultados adequados foi a de Regressão por Processo Gaussiano (RPG).

As funções RPG com função *kernel* exponencial e função *kernel* quadrática (gerada pela otimização de hiperparâmetros) apresentaram  $RMEQ \leq 4$ , tanto para o modelo completo quanto para o modelo com duas variáveis predictoras. Com aplicação da base teste aos modelos, verificou-se  $R^2 > 0,47$  e  $RMSE < 2,18$  para ambos os modelos (completo e de duas variáveis predictoras). Por apresentar valores  $RMEQ$  de treino e de teste próximos, indicando

um ajuste adequado do modelo gerado, foi determinado o modelo com duas variáveis preditoras ( $T_{máx}$  e mediana) e com os hiperparâmetros otimizados (Tabela 12).

Em relação a mama com densidade PG, apesar de alguns fornecerem  $REQM \leq 4$ , os modelos de regressão SVM (completo e com duas variáveis preditoras) foram rejeitados pelo fato do  $R^2$  apresentarem valores menores que 0,36 indicando uma baixa correlação entre os modelos propostos e os dados. Por esse motivo, foram testadas outras técnicas disponíveis no aplicativo utilizado. Os modelos com variáveis selecionadas não obtiveram desempenho adequado, com erros maiores que quatro e coeficientes de determinação baixos. A melhor performance foi a do modelo de regressão por árvore de decisão, treinado com todas as variáveis independentes. Esse modelo foi validado com a base teste de 20% ( $REQM = 2,76$  e  $R^2 = 0,61$ ) e foi selecionado para predição da profundidade de tumores malignos em mamas PG.

Tabela 12– Características dos modelos preditivos para a profundidade do tumor maligno e seus valores de performance

Tipo de mama	Modelo preditivo	C	$\epsilon$	$REQM_{treino}$	$REQM_{teste}$	$R^2_{treino}$	$R^2_{teste}$
<b>ExD</b>	SVM Gaussiana	5,1736	0,5174	3,3359	2,4913	0,41	0,53
<b>HD</b>	SVM Cúbico	4,6692	0,4669	2,8428	2,4191	0,50	0,71
<b>PF</b>	GPR	-	-	2,6967	2,1225	0,67	0,44
<b>PG</b>	Árvore	-	-	2,9153	2,76	0,54	0,61

Fonte: A Autora (2022)

#### *Análise de casos – profundidade do tumor maligno*

Da base de dados utilizada como teste, apenas duas pacientes possuíam informações quanto à profundidade da lesão maligna. A Paciente #1 possuía 49 anos e através de biópsia foi confirmada a presença de um nódulo sólido maligno localizado na mama esquerda, com profundidade de 16 mm da pele. A Paciente #2 possuía 54 anos e foi diagnosticada com *carcinoma ductal* invasivo em sua mama esquerda, com profundidade de 8 mm da pele. Pela classificação segundo a faixa etária, ambas as pacientes possuem mamas heterogeneamente densas. As dimensões dos tumores foram extraídas dos exames de ultrassonografia das pacientes.

A Tabela 13 apresenta as estimativas feitas pelos usuários através da aplicação da metodologia desenvolvida. Como a base para este estudo de casos é pequena, foi calculado apenas o erro (CHAPRA, 2013) entre a estimativa e valor real da profundidade do tumor,

considerando que o valor fornecido pela ultrassonografia é a melhor estimativa disponível do valor verdadeiro. A diferença entre as estimativas geradas pelos usuários para Paciente #1 foram de 0,34 e 0,55, e para Paciente #2 foram de 1,38 e 0,51. Apesar dos erros absolutos terem sido baixos, a estimativa da profundidade para a Paciente #2 gerou erros relativos percentuais de até 17,25% (Usuário 1). Comparando os resultados obtidos pelos Usuários 1 e 2 para a paciente em questão, é possível notar uma diferença de cerca de 11 pontos percentuais quanto aos erros relativos calculados a partir das estimativas de cada profissional. Isso evidencia que a geração de resultados mais acurados possui dependência com a seleção da ROI por parte do operador.

Tabela 13 – Resultados das estimativas da profundidade de tumores malignos usando imagens infravermelhas

Paciente	Original (mm)	Usuário 1 (mm)	Usuário 2 (mm)
Paciente #1	16,00	15,66	15,45
Paciente #2	8,00	6,62	7,49

Fonte: A Autora (2022)

## 5.5 AGATA MULTICLASSES - PROPOSTA DE UM SISTEMA CAD DE MULTICLASSIFICAÇÃO

Foi criada uma GUI com diversos controles para a classificação multiclasses, a estimativa da profundidade e do tamanho de uma possível lesão maligna, bem como a produção de gráficos dos resultados registrados no *software*. A Figura 61 mostra a janela principal da interface desenvolvida.

A interface AGATA Multiclasses (Figura 60) possui três menus (**Abrir**, **Ajuda** e **Sair**) e quatro painéis (**Informações do Paciente**, **Histórico pessoal e familiar**, **Resultado** e **Localização**). Os menus possuem as mesmas funções descritas no AGATA Triagem (Tópico 5.2.1), exceto o menu **Abrir** que, além do QuickReport, permite a abertura do *software* Fluent, reunindo os programas essenciais para análise e simulação de casos.

No painel **Informações do Paciente** foi incluída a altura e o peso para que associados à idade e aos dados solicitados no painel **Histórico pessoal e familiar** possam ser avaliados e registrados fatores de risco de desenvolvimento do câncer de mama (por exemplo, idade avançada, características reprodutivas, hábitos de vida e antecedentes familiar e pessoal). Como o objetivo é usar o AGATA para auxiliar no rastreamento de patologias mamárias em diferentes hospitais e localidades, é fundamental garantir que tais informações sejam

coletadas para melhor acompanhamento da paciente e para posterior criação de modelos que levem características clínicas em consideração.

Figura 61 – Interface gráfica AGATA Versão Multiclasses

The interface is titled "Versão Multiclasses" and is divided into several sections:

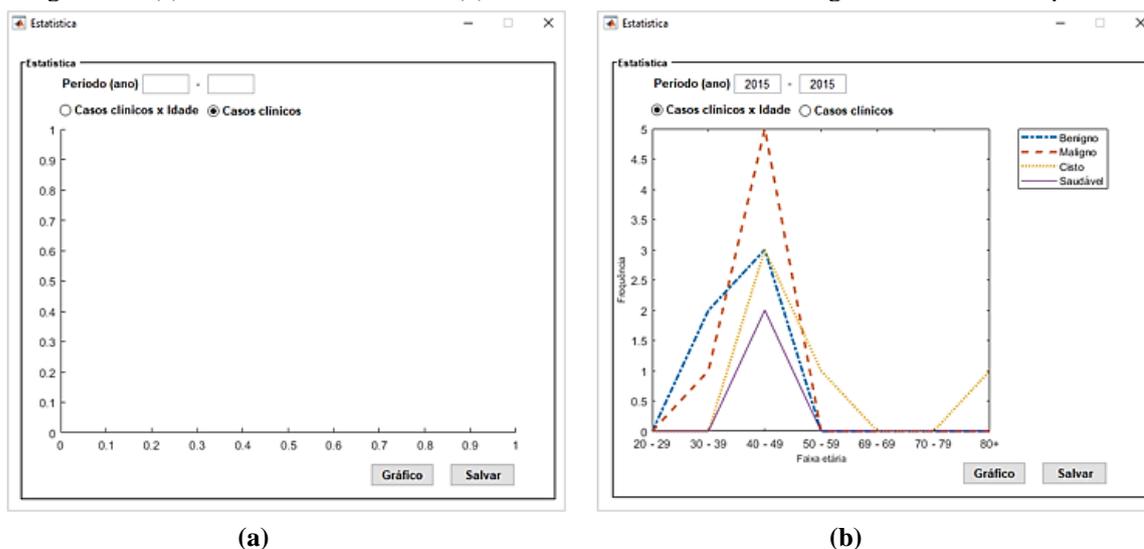
- Informações do Paciente:** Includes input fields for "Número do Prontuário", "Gênero", "Idade" (with "anos" unit), "Altura (cm)", and "Peso (g)".
- Histórico pessoal e familiar:** Contains ten dropdown menus for:
  1. Atividade física
  2. Etilismo
  3. Tabagismo
  4. Faixa etária da 1ª menstruação
  5. Nuligesta
  6. Idade da primeira gestação (with "anos" unit)
  7. Amamentou?
  8. Menopausa?
  9. História pessoal de câncer de mama e/ou ovário?
  10. Quantos parentes de primeiro grau tiveram câncer de mama e/ou ovário?
- Resultado:** Features "ICT" and "Valor Preditivo" input fields.
- Localização:** Includes fields for "Mama acometida", "Profundidade (mm)", and "Tamanho (mm)".

At the bottom of the interface are three buttons: "Iniciar", "Estatística", and "Salvar".

Fonte: A Autora (2022)

A produção de gráficos através do botão **Estatística** é similar ao painel **Estatística** do AGATA Triagem. Agora, ao clicar no botão uma sub-interface (Figura 62a) é aberta e, além disso, os **Casos clínicos** × **Idade** dos dados registrados em determinado período são apresentados através do gráfico da Figura 62b.

Figura 62 - (a) Sub-interface Estatística (b) Gráfico Casos clínicos x Idade gerado com dados hipotéticos



Fonte: A autora (2022)

No AGATA Multiclasses foi implementado o classificador SVM (base balanceada por SAES 80%) para análise não supervisionada e individual das imagens termográficas de mamas. Para este caso, o classificador foi também validado através da técnica *leave-one-out* na qual um indivíduo serve de teste e  $n-1$  são usados na construção da regra de classificação. O método é repetido  $n$  vezes até que todos os elementos do conjunto de dados sejam utilizados como teste, a partir de  $n$  conjuntos distintos de treino. Ao final da validação, o classificador apresentou acurácia de 92,46%, sensibilidade de 92,56% e especificidade de 97,48%.

No painel **Resultados** aparecerá os índices ICT 1 – 4 cujas descrições foram apresentadas no Quadro 22. Caso o resultado da análise for ICT 4 (Lesão Maligna), o operador tem a opção de estimar a profundidade e o tamanho da lesão (painel **Localização**).

## 5.6 ESTUDO DE CASOS

O Quadro 23 apresenta os resultados obtidos no estudo de casos de pacientes com diagnósticos concluídos através de exame clínico, de ultrassonografias, de mamografias e biópsias. Foram selecionadas, aleatoriamente, imagens de pacientes que não compõem a base de dados utilizada no desenvolvimento das diferentes versões da ferramenta AGATA.

Observa-se que na classificação da Paciente 4 houve a geração de falsos negativos com o uso das versões de Triagem avaliadas. Este erro deve ser tratado com bastante cuidado, uma vez que pode acarretar problemas mais graves como, por exemplo, a procrastinação do tratamento do câncer, uma vez que poderia adiar a correta investigação diagnóstica. Por isso,

não só este como os demais métodos de detecção do câncer de mama são analisados em conjunto, a fim de que a probabilidade de se ter um falso negativo ou um falso positivo diminua seja ainda menor.

Quadro 23– Resultados obtidos na classificação das pacientes estudadas

			Versões da ferramenta AGATA		
Pacientes	Idade	Diagnóstico	Triagem Série	Triagem	Multiclassificação
Paciente 1	60	Saudável	ICT 0	ICT 4	ICT 4
Paciente 2	63	Lesão Benigna	ICT 4	ICT 0	ICT 0
Paciente 3	73	Cisto	ICT 0	ICT 0	ICT 4
Paciente 4	82	Lesão Maligna	ICT 0	ICT 0	ICT 4
Paciente 5	57	Lesão Maligna	ICT 0	ICT 4	ICT 4

Fonte: A Autora (2022)

## 5.7 SÍNTESE DAS CONTRIBUIÇÕES DO TRABALHO

Podem-se pontuar as seguintes contribuições:

- Validação da segmentação automática, desenvolvida Dourado (2014), para maior confiabilidade dos sistemas CAD criados;
- Criação de banco com imagens *ground-truth* de termogramas mamários que podem ser usadas por outros pesquisadores para validação de metodologias de segmentação automática;
- Construção de base de dados incluindo pacientes saudáveis;
- Análise da base de dados (multiclasses) com respeito a quantidade de exemplos, relacionando essa categorização com desempenho dos classificadores utilizados;
- Desenvolvimento de estratégia de sobreamostragem baseada em elementos *safe*;
- Desenvolvimento de sistemas CAD (aprendizado supervisionado) para análises binária e multiclasses com taxas de acurácia superiores a 90,00%, incluir sensibilidade e especificidade;
- Análise sistemática de sobreajustamento através de testes estatísticos;
- Construção de banco de perfil de temperaturas simuladas, com 132 elementos, considerando os diferentes tipos de composição tecidual das mamas;

- Elaboração de sistema preditivo de tamanho e profundidade tumoral de lesões malignas, baseadas em características térmicas;
- Produção de uma ferramenta de análise individual para aplicação clínica real no auxílio ao diagnóstico do câncer de mama.

## 6 CONCLUSÕES

Os estudos apresentados na presente tese forneceram embasamento teórico para o desenvolvimento de um produto: a plataforma computacional AGATA, versão Triagem e versão Multiclasses. A plataforma é uma ferramenta amigável que foi desenvolvida para integrar os programas criados neste trabalho e para facilitar a aplicação da termografia no meio clínico. A GUI AGATA não requer necessariamente um especialista em sistemas de dados para ser operada, podendo ser facilmente usada por engenheiros, médicos, especialistas, pesquisadores e acadêmicos. As duas versões são destinadas a avaliação individual (não-supervisionada) de pacientes.

O referido estudo apresentou novos sistemas CAD capazes de auxiliar na triagem de pacientes e na detecção de patologias mamárias através de imagens termográficas. As imagens IRs normalmente têm baixo contraste. Por isso, procurou-se segmentar a região mamária para auxiliar no processo de detecção do câncer de mama. Foi adotada uma metodologia para verificação e validação da técnica de segmentação automática de termogramas mamários em relação ao *ground-truth*.

Através da metodologia desenvolvida, concluiu-se que os métodos de extração das áreas mamárias analisadas concordaram. Isso sugere que a segmentação automática avaliada pode ser utilizada no processo de diagnóstico auxiliado por computador, principalmente quando o objetivo é obter uma classificação com menor tempo de processamento. Os cálculos estatísticos foram feitos com base no número de *pixels* semelhantes e diferentes entre as duas ROIs. Os resultados obtidos nas medidas sobrepostas mostraram que houve 69% de intersecção entre as regiões segmentadas e as regiões *ground-truth*. Este resultado equivale a uma concordância substancial, que atesta a conformidade do método de segmentação automática usado.

A seleção de atributos com os métodos evolutivos trouxe otimização para as medidas de performance dos classificadores utilizados e mostrou que o uso de pares de atributos, metodologia até então adotada pelo grupo de pesquisa, não fornece os melhores resultados.

Para a classificação binária, o melhor desempenho foi obtido usando o classificador SVM combinado a seleção de atributos por Algoritmo Genético e a sobreamostragem ASUWO: 95,23% de acurácia, 93,65% de sensibilidade e 96,81% de especificidade. Também em relação a este sistema, o índice Kappa foi de 0,90 e a AUC foi de 0,99. A proximidade de ambos os índices ao valor 1 mostra uma concordância significativa entre os valores previstos e esperados.

A utilização de uma base com classes variadas (lesão maligna, lesão benigna, cisto e saudável) é um diferencial do presente trabalho. Essa escolha procurou retratar a variedade de

diagnósticos que podem ser feitos pelos especialistas, bem como o desbalanceamento inerente aos problemas da área médica. No entanto, abordar a variabilidade clínica pela análise de temperaturas gera um espaço de características com classes altamente sobrepostas, dispersas e com um número alto de exemplos inseguros (limítrofes, raros e *outliers*).

Em relação ao problema do desbalanceamento multiclases, o método de sobreamostragem de vetores sintéticos baseado na análise de rótulos de exemplos pertencentes a uma dada classe gerou os melhores desempenhos. O destaque foi para o método SAES 80% cujas métricas de diagnóstico, para validação cruzada 10-folds, foram 92,20% para acurácia, 92,32% para a sensibilidade e 97,40% para especificidade.

Quanto à sobreposição das classes, foram avaliados métodos de normalização dos dados da base desbalanceada. A normalização por classe gerou uma maior separação entre as classes, com destaque para o método *Scale* que alcançou taxas de 100% para principais métricas dos classificadores.

Os resultados obtidos, tanto para a classificação binária quanto a multiclases, atestam que a termografia associada a técnicas de inteligência computacional pode ser usada como ferramenta complementar na triagem do câncer de mamas. Ressalta-se ainda que, na base avaliada, a temperatura média mostrou ser um parâmetro eficaz na identificação da mama com patologia.

Em complementação às técnicas de diagnóstico auxiliado por computador, a partir da distribuição da temperatura superficial de geometrias mamárias simuladas foram desenvolvidos modelos de regressão para estimativa da profundidade e do tamanho de tumor maligno, usando casos assim diagnosticados. Os casos apresentados foram uma primeira tentativa de fornecer esses valores. De forma geral, foram obtidos modelos de regressão com  $R^2 > 0,77$  significando relações fortes entre as variáveis correlacionadas. Nos estudos de casos realizados, os resultados demonstraram que tumores com profundidades maiores ou iguais a 20 mm apresentam grandes diferenças entre o valor real e o valor estimado através da técnica aplicada pelos usuários, reforçando que tumores profundos geram padrões de temperatura com baixo contraste nas distribuições de temperatura da superfície mamária. Neste ponto, há uma necessidade de aprimorar os modelos de regressão para tumores pequenos e profundos, além de gerar dados em larga escala para obtenção de modelos mais bem ajustados.

### *Trabalhos futuros*

Como trabalhos futuros, pode-se sugerir:

- Aumentar a base de dados utilizada tanto para classificação supervisionada quanto para avaliação individual implementada na interface gráfica;

- Coletar informações clínicas dos prontuários das pacientes para criação de modelos de classificação baseado nos fatores de risco de desenvolvimento de câncer de mama;
- Aperfeiçoar os modelos de regressão por tipo de mama a partir da simulação de novos casos (tumores localizados no quadrante inferior, por exemplo) e da utilização de outras técnicas de regressão por aprendizado de máquina, como Rede Neural e *Random Forest*;
- Ampliar a estimativa da localização para outras patologias das mamas, como tumor benigno e cisto;
- Tornar a ferramenta computacional AGATA disponível para realização de testes em consultas médicas, como possível realização de triagem durante a espera dos pacientes em consultório ou ambulatórios do serviço público;
- Gerenciar a execução de simulações numéricas de perfis de temperatura e estimativa de parâmetros termofísicos.

## REFERÊNCIAS

- ACHARYA, U. R. et al. Higher order spectra analysis of breast thermograms for the automated identification of breast cancer. **Expert Systems**, v. 31, n. 1, p. 37-47, 2014.
- ACHARYA, U. R. et al. Thermography based breast cancer detection using texture features and support vector machine. **Journal of medical systems**, v. 36, n. 3, p. 1503-1510, 2012.
- ADEL, M.; ABDELHAMID, B.; EL-RAMLY, S. Automatic image segmentation of breast thermograms. In: **Proceedings of the 2018 7th International Conference on Bioinformatics and Biomedical Science**, p. 88-94, 2018.
- ALBUQUERQUE NETO, C. **Modelo integrado dos sistemas térmico e respiratório do corpo humano**. Tese (Doutorado) - Universidade de São Paulo, 2010.
- ALFAYEZ, F.; EL-SOUD, M. W. A.; GABER, T. Thermogram Breast Cancer Detection: a comparative study of two machine learning techniques. **Applied Sciences**, v. 10, n. 2, p. 551, 2020.
- ALI, M. A. S. et al. Detection of breast abnormalities of thermograms based on a new segmentation method. In: **2015 Federated conference on computer science and information systems (FedCSIS)**. IEEE, p. 255-261, 2015.
- POMALAZA, M. A. A. **Segmentación automática de la planta de pie en imágenes termográficas en entorno ruidoso para el diagnóstico de pie diabético**. Tese (Doutorado) - Pontificia Universidad Católica del Perú, Facultad de Ciencias e Ingeniería. Mención: Ingeniería Informática, 2016.
- AMMER, K.; RING, E. F. J. Standard procedures for infrared imaging in medicine. **Biomedical Engineering Handbook**, CRC Press, v. 1, 2006.
- AMORIM, A. M. A. M. et al. Termografia infravermelha na Odontologia. **HU rev**, p. 15-22, 2018.
- ARAUJO, R. L. **Evolução diferencial para problemas de otimização com restrições lineares**. Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, 2016.
- ARAÚJO, M. C. d. **Uso de imagens termográficas para classificação de anormalidades de mama baseado em variáveis simbólicas intervalares**. Tese (Doutorado) - Universidade Federal de Pernambuco, 2014.
- ARAÚJO, M. C. d.; LIMA, R. C. F. d.; SOUZA, R. M. C. R. d. Interval symbolic feature extraction for thermography breast cancer detection. **Expert Systems with Applications**, v. 41, n. 15, p. 6728-6737, 2014.
- ARAÚJO, M. C. d. et al. The Evolution of New Trends in Breast Thermography. In: **Biomedical Computing for Breast Cancer Detection and Diagnosis**. IGI Global, p. 128-171, 2021.

ARORA, N. et al. Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. **The American Journal of Surgery**, v. 196, n. 4, p. 523-526, 2008.

ARUL EDWIN RAJ, A. M.; SUNDARAM, M.; JAYA, T. Thermography based breast cancer detection using self-adaptive gray level histogram equalization color enhancement method. **International Journal of Imaging Systems and Technology**, v. 31, n. 2, p. 854-873, 2021.

BARBOSA, A. Câncer, direito e cidadania: como a lei pode beneficiar mulheres e familiares. **São Paulo: Atlas**, 2010.

BARELLA, V. H. **Técnicas para o problema de dados desbalanceados em classificação hierárquica**. Dissertação (Mestrado) - Universidade de São Paulo, 2015.

BARROS, A. et al. Thermographic study of the orofacial structures involved in clarinetists musical performance. **Dentistry Journal**, v. 6, n. 4, p. 62, 2018.

BATISTA, G. E. A. P. A.; PRATI, R. C.; MONARD, M. C. A study of the behavior of several methods for balancing machine learning training data. **ACM SIGKDD explorations newsletter**, v. 6, n. 1, p. 20-29, 2004.

BAYAS, T. M. M. **Segmentación y análisis de imágenes termográficas mediante descriptores de textura simples**. Trabalho de Conclusão de Curso (Engenharia Eletrônica e Comunicações) – Universidad Técnica de Ambato, 2015.

BECKMANN, M. **Algoritmos genéticos como estratégia de pré-processamento em conjuntos de dados desbalanceados**. Dissertação (Mestrado) - Universidade Federal do Rio de Janeiro, 2010.

BEN-DAVID, A. Comparação da precisão da classificação usando o Kappa ponderado de Cohen. **Sistemas Especialistas com Aplicações**, v. 34, n. 2, pág. 825-832, 2008.

BERG, W. A. et al. Cystic breast masses and the ACRIN 6666 experience. **Radiologic Clinics**, v. 48, n. 5, p. 931-987, 2010.

BEZERRA, L. A. **Estimativa de parâmetros termofísicos da mama e de distúrbios mamários a partir de termografia por infravermelho utilizando técnicas de otimização**. Tese (Doutorado) - Universidade Federal de Pernambuco, 2013.

BEZERRA, L. A. et al. An empirical correlation to estimate thermal properties of the breast and of the breast nodule using thermographic images and optimization techniques. **International Journal of Heat and Mass Transfer**, v. 149, p. 119215, 2020.

BHOWMIK, M. K. et al. Standardization of infrared breast thermogram acquisition protocols and abnormality analysis of breast thermograms. In: **Thermosense: Thermal Infrared Applications XXXVIII**. SPIE, p. 269-286, 2016.

BLUM, Avrim L.; LANGLEY, Pat. Selection of relevant features and examples in machine learning. **Artificial intelligence**, v. 97, n. 1-2, p. 245-271, 1997.

BOCK, H.; DIDAY, E. **Analysis of symbolic data: exploratory methods for extracting statistical information from complex data**. Springer Science & Business Media, 1999.

BRASILEIRO FILHO, G. **Bogliolo Patologia**. 6. ed. Brasil: Guanabara Koogan, 2016. ISBN 8527706113.

BORCHARTT, T. B. **Análise de imagens termográficas para a classificação de alterações na mama**. Tese (Doutorado) - Universidade Federal Fluminense, 2013.

BRANDÃO, M. A. L. **Estudo de alguns métodos determinísticos de otimização irrestrita**. Dissertação (Mestrado) – Universidade Federal de Uberlândia, 2010.

BRIOSCHI, M. L.; OKIMOTO, M. L. L. R.; VARGAS, J. V. C. The utilization of infrared imaging for occupational disease study in industrial work. **Work**, v. 41, n. Supplement 1, p. 503-509, 2012.

BRITO, C.; PORTELA, M. C.; VASCONCELLOS, M. T. L. d. Assistência oncológica pelo SUS a mulheres com câncer de mama no Estado do Rio de Janeiro. **Revista de Saúde Pública**, v. 39, p. 874-881, 2005.

BRUZZONE, L.; PERSELLO, C. Approaches based on support vector machine to classification of remote sensing data. En: Chen, CH (Ed.), **Handbook of Pattern Recognition and Computer Vision**. 2009.

BRZEZINSKI, R.Y. et al. Automated thermal imaging for the detection of fatty liver disease. **Scientific reports**, v. 10, n. 1, p. 1-11, 2020.

BUNKHUMPORNPAT, C.; SINAPIROMSARAN, K.; LURSINSAP, C. Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. In: **Pacific-Asia conference on knowledge discovery and data mining**. Springer, Berlin, Heidelberg, p. 475-482, 2009.

BURTON, A. C. The application of the theory of heat flow to the study of energy metabolism: Five figures. **The Journal of Nutrition**, v. 7, n. 5, p. 497-533, 1934.

CABEZAS, M. et al. Automatic multiple sclerosis lesion detection in brain MRI by FLAIR thresholding. **Computer methods and programs in biomedicine**, v. 115, n. 3, p. 147-161, 2014.

CALDERÓN-CONTRERAS, J. d. et al. A fuzzy computer aided diagnosis system using breast thermography. In: **2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)**. IEEE, p. 105-108, 2015.

CALISTO, F. M.; NUNES, N.; NASCIMENTO, J. C. BreastScreening: on the use of multi-modality in medical imaging diagnosis. In: **Proceedings of the international conference on advanced visual interfaces**, p. 1-5, 2020.

CASTRO, D. L.; TEGOLO, D.; VALENTI, C. A visual framework to create photorealistic retinal vessels for diagnosis purposes. **Journal of Biomedical Informatics**, v. 108, p. 103490, 2020.

CESARIS, R. et al. Changes in blood pressure and thermographic values resulting from use of a beta-blocker plus diuretic and of an alpha-beta-blocker plus diuretic. **Drugs under experimental and clinical research**, v. 11, n. 10, p. 725-729, 1985.

CHABRIER, S. et al. A comparative study of supervised evaluation criteria for image segmentation. In: **2004 12th European Signal Processing Conference**. IEEE, p. 1143-1146, 2004.

CHAPMAN, S. J. **Programação em MATLAB® para engenheiros**. Thomson. São Paulo, 2003.

CHAPRA, S. C. **Métodos numéricos para engenharia**. 5. ed. São Paulo, SP: McGraw-Hill, 2008. ISBN 9788586804878.

CHARNY, C. K. Mathematical models of bioheat transfer. **Advances in heat transfer**, v. 22, p. 19-155, 1992.

CHAWLA, N. V. et al. SMOTE: synthetic minority over-sampling technique. **Journal of artificial intelligence research**, v. 16, p. 321-357, 2002.

CHERKAS, L. F. et al. Use of thermographic criteria to identify Raynaud's phenomenon in a population setting. **The Journal of rheumatology**, v. 30, n. 4, p. 720-722, 2003.

COLAÇO, M. J.; ORLANDE, H. R. B.; DULIKRAVICH, G. S. Inverse and optimization problems in heat transfer. **Journal of the Brazilian Society of Mechanical Sciences and Engineering**, v. 28, p. 1-24, 2006.

CONCI, A.; AZEVEDO, E.; LETA, F. R. **Computação Gráfica**. Editora Campus/Elsevier, Rio de Janeiro, v. 2., p. 432, 2008.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and psychological measurement**, v. 20, n. 1, p. 37-46, 1960.

COPPIN, B. **Inteligência artificial**. 1. ed. Rio de Janeiro: LTC, 2017. ISBN 978-85-216-1729-7

DANTAS, D.; DONADIA, E. **Comparação entre as técnicas de regressão logística, árvore de decisão, bagging e random forest aplicadas a um estudo de concessão de crédito**. Trabalho de conclusão de curso (Bacharelado em Estatística) - Universidade Federal do Paraná, 2013.

DIAKIDES, M.; BRONZINO, J. D. **Medical Infrared Imaging**, CRC Press, Boca Raton, Fla, USA, 2007.

DIBAI-FILHO, A. V. et al. The chronicity of myogenous temporomandibular disorder changes the skin temperature over the anterior temporalis muscle. **Journal of bodywork and movement therapies**, v. 18, n. 3, p. 430-434, 2014.

DICE, L. R. Measures of the Amount of Ecologic Association Between Species. **Ecology**, v. 26, n. 3, p. 297-302, 1945.

DINIZ, J. B. **Análise de algoritmos de segmentação de melanoma em imagens dermatoscópicas**. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) - Universidade Federal Rural de Pernambuco, 2017.

DOUGHERTY, G. **Digital Image Processing for Medical Applications**. 1. ed. Cambridge University Press, New York, 2009.

DOURADO, H. M. **Segmentação e análise automáticas de termogramas: um método auxiliar na detecção do câncer de mama um método auxiliar na detecção do câncer de mama**. Dissertação (Mestrado) - Universidade Federal de Pernambuco, 2014.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. 2 ed. New York: John Wiley and Sons, 2000.

EKICI, S.; JAWZAL, Hu. Breast cancer diagnosis using thermography and convolutional neural networks. **Medical hypotheses**, v. 137, p. 109542, 2020.

EKPO, E. U.; HOGG, P.; MCENTEE, M. F. Breast composition: Measurement and clinical use. **Radiography**, online, v. 21, p. 324-333, June 2015.

ETEHADTAVAKOL, M.; NG, E. Y. K. Breast thermography as a potential non-contact method in the early detection of cancer: a review. **Journal of Mechanics in Medicine and Biology**, v. 13, n. 02, p. 1330001, 2013.

FACELI, K. et al. **Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina**. Rio de Janeiro: LTC, 2011. ISBN 978-85-216-1880-5.

FALCONE, M. A. G. **Estudo comparativo entre algoritmo genético e evolução diferencial para otimização de um modelo de cadeia de suprimento simplificada**. Dissertação (Mestrado) - Pontifícia Universidade Católica do Paraná, 2004.

FARIAS, T. M. T. et al. Automatic Segmentation of Neonates Thermal Imaging for Evaluation of Trunk Thermal Asymmetry. In: **2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)**. IEEE, p. 989-992, 2018.

FAROOQ, M. A.; CORCORAN, P. Infrared imaging for human thermography and breast tumor classification using thermal images. In: **2020 31st Irish Signals and Systems Conference (ISSC)**. IEEE, p. 1-6, 2020.

FERNÁNDEZ, A. et al. Analysing the classification of imbalanced datasets with multiple classes: Binarization techniques and ad-hoc approaches. **Knowledge-based systems**, v. 42, p. 97-110, 2013.

FLORES-SAHAGUN, J. H.; VARGAS, J. V. C.; MULINARI-BRENNER, F. A. Analysis and diagnosis of basal cell carcinoma (BCC) via infrared imaging. **Infrared Physics & Technology**, v. 54, n. 5, p. 367-378, 2011.

FRANCO, N. H. et al. ThermoLabAnimal—A high-throughput analysis software for non-invasive thermal assessment of laboratory mice. **Physiology & behavior**, v. 207, p. 113-121, 2019.

FRIZE, M.; HERRY, C.; ROBERGE, R. Processing of thermal images to detect breast cancer: Comparison with previous work. In: **Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society [Engineering in Medicine and Biology]**. IEEE, p. 1159-1160, 2002.

FURUICHI, M. et al. The laboratory-confirmed dengue fever and chikungunya fever cases at the Narita Airport Quarantine Station in 2013. **Japanese Journal of Infectious Diseases**, p. JJID. 2014.242, 2014.

GAMA NETO, M. V. d. **O processo CRISP-DM aplicado na construção de uma solução para Análise de Risco de Crédito**. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) - Universidade Federal de Pernambuco, 2018.

GARETH, J. et al. **An introduction to statistical learning: with applications in R**. Springer, 2013.

GOGOI, U. R. et al. Discriminative feature selection for breast abnormality detection and accurate classification of thermograms. In: **2017 international conference on innovations in electronics, signal processing and communication (IESC)**. IEEE, p. 39-44, 2017.

GORBACH, A. M. et al. Intraoperative infrared imaging of brain tumors. **Journal of neurosurgery**, v. 101, n. 6, p. 960-969, 2004.

GONÇALVES, C. B. **Detecção de câncer de mama utilizando imagens termográficas**. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) - Universidade Federal de Uberlândia, 2017.

GONZÁLEZ, F. J. Thermal simulation of breast tumors. **Revista Mexicana de Física**, v. 53, n. 4, p. 323-326, 2007.

GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. L. **Digital Image Processing Using MATLAB®**. Gatesmark Publishing, 2009.

GONZALEZ, J. R. et al. An approach for thyroid nodule analysis using thermographic images. In: **Application of infrared to biomedical sciences**. Springer, Singapore, p. 451-475, 2017.

GRANADILLO, D.B.; MORALES, Y.; TORRES, C.M.; BENJUMEA, E. Segmentación y análisis de imágenes termográficas para la detección de cáncer de mama. **Desarrollo e Innovación en Ingeniería**, v. 4, p. 316-323, 2019.

HA, T. N. et al. An in-depth analysis of logarithmic data transformation and per-class normalization in machine learning: Application to unsupervised classification of a turbidite system in the Canterbury Basin, New Zealand, and supervised classification of salt in the Eugene Island minibasin, Gulf of Mexico. **Interpretation**, v. 9, n. 3, p. T685-T710, 2021.

HADDAD, D. S.; BRIOSCHI, M. L.; ARITA, E. S. Thermographic and clinical correlation of myofascial trigger points in the masticatory muscles. **Dentomaxillofacial Radiology**, v. 41, n. 8, p. 621-629, 2012.

HADDAD, D. S. et al. Thermographic characterization of masticatory muscle regions in volunteers with and without myogenous temporomandibular disorder: preliminary results. **Dentomaxillofacial Radiology**, v. 43, n. 8, p. 20130440, 2014.

HALL, M. A. Correlation-based feature selection for machine learning. University of Waikato Hamilton, 1999.

HAN, H.; WANG, W; MAO, B. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In: **International conference on intelligent computing**. Springer, Berlin, Heidelberg, p. 878-887, 2005.

HAN, J. et al. Mining frequent patterns without candidate generation: A frequent-pattern tree approach. **Data mining and knowledge discovery**, v. 8, n. 1, p. 53-87, 2004.

HARDY, J. D. et al. Radiation of heat from the human body. V. The transmission of infra-red radiation through skin. **The Journal of Clinical Investigation**, v. 15, n. 1, p. 1-9, 1936.

HASGALL, P. A. IT'IS Database for thermal and electromagnetic parameters of biological tissues, versão 3.0. IT'IS - The Foundation for Research on Information Technologies in Society. Zurich, Suíça. 2015.

HASTINGS, R. C. et al. Bacterial density in the skin in lepromatous leprosy as related to temperature. **Leprosy review**, v. 39, n. 2, p. 71-4, 1968.

HE, H. et al. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In: **2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)**. IEEE, p. 1322-1328, 2008.

HE, H.; GARCIA, E. A. Learning from imbalanced data. **IEEE Transactions on knowledge and data engineering**, v. 21, n. 9, p. 1263-1284, 2009.

HEAD, J. F. et al. Image analysis of digitized infrared images of the breasts from a first-generation infrared imaging system. In: **Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering'**(Cat. No. 97CH36136). IEEE, p. 681-684, 1997.

HEIMBECHER, C. T.; ULBRICHT, L. Termografia aplicada ao fenômeno de Raynaud: artigo de revisão bibliométrica. **Pan American Journal of Medical Thermology**, v. 4, p. 15-23, 2018.

HILDEBRANDT, C.; RASCHNER, C.; AMMER, K. An overview of recent application of medical infrared thermography in sports medicine in Austria. **Sensors**, v. 10, n. 5, p. 4700-4715, 2010.

Hirakata, V. N.; Camey, S. A. Análise de concordância entre métodos de Bland-Altman. **Revista HCP**, v. 29, n. 3, p. 261-268, 2009.

HOSAKI, Y. et al. Non-invasive study for peripheral circulation in patients with diabetes mellitus. **Annual reports of Misasa Medical Branch**, v. 72, p. 31-37, 2002.

HOSSAIN, S.; MOHAMMADI, F. A. Tumor parameter estimation considering the body geometry by thermography. **Computers in biology and medicine**, v. 76, p. 80-93, 2016.

HOSSAM, A.; HARB, H. M.; ABD EL KADER, H. M. Automatic image segmentation method for breast cancer analysis using thermography. **JES. Journal of Engineering Sciences**, v. 46, n. 1, p. 12-32, 2018.

IBAN, M. C.; ŞENTÜRK, E. Machine learning regression models for prediction of multiple ionospheric parameters. **Advances in Space Research**, v. 69, n. 3, p. 1319-1334, 2022.

INCA. Instituto Nacional de Câncer José Alencar Gomes da Silva. **Estimativas 2020: Incidência de Câncer no Brasil**. Rio de Janeiro, Brazil: Ministério da Saúde, 2019.

INCROPERA, F. P. et al. Fundamentals of heat and mass transfer. 6th. ed. [s.l.] John Wiley & Sons, Inc., 2007.

JACKMAN, P.; SUN, D.-W.; ALLEN, P. Automatic segmentation of beef longissimus dorsi muscle and marbling by an adaptable algorithm. **Meat Science**, v. 83, n. 2, p. 187-194, 2009.

JEYANATHAN, J.; JEYASHREE, P.; SHENBAGAVALLI, A. Transform based classification of breast thermograms using multilayer perceptron back propagation neural network. **Int. J. Pure Appl. Math**, v. 118, n. 20, p. 1955-1961, 2018.

JO, T.; JAPKOWICZ, N. Class imbalances versus small disjuncts. **ACM Sigkdd Explorations Newsletter**, v. 6, n. 1, p. 40-49, 2004.

KAMANGAR, S. et al. Numerical analysis of heat transfer in human head. **Journal of Mechanical Science and Technology**, v. 33, n. 7, p. 3597-3605, 2019.

KANDLIKAR, Satish G. et al. Infrared imaging technology for breast cancer detection—Current status, protocols and new directions. **International Journal of Heat and Mass Transfer**, v. 108, p. 2303-2320, 2017.

KAPOOR, P. et al. Image segmentation and asymmetry analysis of breast thermograms for tumor detection. **International Journal of Computer Applications**, v. 50, n. 9, 2012.

KELEŞ, A.; KELEŞ, A.; YAVUZ, U. Expert system based on neuro-fuzzy rules for diagnosis breast cancer. **Expert systems with applications**, v. 38, n. 5, p. 5719-5726, 2011.

KHAN, M. A. et al. Performance evaluation of regression models for COVID-19: A statistical and predictive perspective. **Ain Shams Engineering Journal**, v. 13, n. 2, p. 101574, 2022.

KLEIN, P. C. **Detecção de lesões de esclerose múltipla em imagens de ressonância magnética do tipo Fluid Attenuated Inversion Recovery (FLAIR)**. Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio Grande do Sul, 2016.

LAHIRI, B. B. et al. Medical applications of infrared thermography: a review. **Infrared Physics & Technology**, v. 55, n. 4, p. 221-235, 2012.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **biometrics**, p. 159-174, 1977.

LARA, D. F. et al. Estudos empíricos dos métodos de balanceamento para a classificação. In: **XVIII Simposio Argentino de Inteligencia Artificial (ASAI)-JAIIO 46 (Córdoba, 2017)**, 2017.

LASHKARI, A.; PAK, F.; FIROUZMAND, M. Full intelligent cancer classification of thermal breast images to assist physician in clinical diagnostic applications. **Journal of medical signals and sensors**, v. 6, n. 1, p. 12, 2016.

LEMAÎTRE, G.; NOGUEIRA, F.; ARIDAS, C. K. Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. **The Journal of Machine Learning Research**, v. 18, n. 1, p. 559-563, 2017.

LENNOX, N.; HASKINS, B. Contrasting classifiers for the detection of breast cancer using thermographic images. In: **Proceedings of the 2nd International Conference on Intelligent and Innovative Computing Applications**, p. 1-9, 2020.

LESSA, V.; MARENGONI, M. Applying artificial neural network for the classification of breast cancer using infrared thermographic images. In: **International Conference on Computer Vision and Graphics**. Springer, Cham, p. 429-438, 2016.

LI, X. et al. Estimating the ground truth from multiple individual segmentations with application to skin lesion segmentation. In: **Proc. Medical Image Understanding and Analysis Conference, UK**. p. 101-106, 2010.

LINDEN, R. Algoritmos Genéticos: Teoria e Implementação. **Editora Ciência Moderna, 3ª Edição**, 2005.

LIU, C. et al. Infrared dermal thermography on diabetic feet soles to predict ulcerations: a case study. In: **Advanced Biomedical and Clinical Diagnostic Systems XI**. SPIE, 2013. p. 102-110.

LORENA, A. C.; CARVALHO, A. C. P. L. F. d. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43-67, 2007.

MACEDO, D. C. d. **Comparação da redução de dimensionalidade de dados usando seleção de atributos e conceito de framework: um experimento no domínio de clientes**. Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná, 2012.

MADHAVI, V.; BOBBY, T. C. Thermal imaging-based breast cancer analysis using BEMD and uniform RLBP. In: **2017 Third International Conference on Biosignals, Images and Instrumentation (ICBSII)**. IEEE, p. 1-6, 2017.

MAGAS, V. et al. Evaluation of thermal imaging for the diagnosis of repetitive strain injuries of the wrist and hand joints. **Research on Biomedical Engineering**, v. 35, n. 1, p. 57-64, 2019.

MAIA, C. M.; GOMES, J. C. M.; CHAGAS, L. D. Estudo sobre o uso de árvores de decisão na área da saúde. **Anais do Encontro de Computação do Oeste Potiguar ECOP/UFERSA (ISSN 2526-7574)**, n. 1, 2017.

MAJEED, R. F. et al. Skin Tumors Diagnosis Utilizing Case Based Reasoning and The Expert System. **Kurdistan Journal of Applied Research**, v. 5, n. 1, p. 96-114, 2020.

MANDAL, M.; MUKHOPADHYAY, A. An improved minimum redundancy maximum relevance approach for feature selection in gene expression data. **Procedia Technology**, v. 10, p. 20-27, 2013.

MANGINAS, A. et al. Right ventricular endocardial thermography in transplanted and coronary artery disease patients: first human application. **The Journal of invasive cardiology**, v. 22, n. 9, p. 400-404, 2010.

MARQUES, R. S. et al. An approach for automatic segmentation of thermal imaging in Computer Aided Diagnosis. **IEEE Latin America Transactions**, v. 14, n. 4, p. 1856-1865, 2016.

MASON, B. R.; GRAFF, A. J.; PEGG, S. P. Colour thermography in the diagnosis of the depth of burn injury. **Burns**, v. 7, n. 3, p. 197-202, 1981.

MELO, J. R. F. et al. Development of a three-dimensional surrogate geometry of the breast and its use in estimating the thermal conductivities of breast tissue and breast lesions based on infrared images. **International Communications in Heat and Mass Transfer**, v. 108, p. 104279, 2019.

MISHRA, V.; RATH, S. K. Detection of breast cancer tumours based on feature reduction and classification of thermograms. **Quantitative InfraRed Thermography Journal**, v. 18, n. 5, p. 300-313, 2021.

MORASIEWICZ, M. et al. Gait analysis in patients after lengthening and correction of tibia with Ilizarov technique. **Physiotherapy Quarterly**, v. 18, n. 1, p. 9, 2010.

MORETO, F. A. d. L. **Análise de componentes independentes aplicada à separação de sinais de áudio**. Dissertação (Mestrado) - Universidade de São Paulo, 2008.

MORGAN, P.B.; TULLO, A. B.; EFRON, N. Ocular surface cooling in dry eye—a pilot study. **Journal of the British Contact Lens Association**, v. 19, n. 1, p. 7-10, 1996.

NARAYANAMURTHY, V. et al. Skin cancer detection using non-invasive techniques. **RSC advances**, v. 8, n. 49, p. 28095-28130, 2018.

NEKOOEIMEHR, I.; LAI-YUEN, S. K. Adaptive semi-supervised weighted oversampling (A-SUWO) for imbalanced datasets. **Expert Systems with Applications**, v. 46, p. 405-416, 2016.

NG, E. Y. K. Is thermal scanner losing its bite in mass screening of fever due to SARS? **Medical physics**, v. 32, n. 1, p. 93-97, 2005.

NG, E.Y.K. A review of thermography as promising non-invasive detection modality for breast tumor. **International Journal of Thermal Sciences**, v. 48, n. 5, p. 849-859, 2009.

NG, E. Y. K.; KEE, E. C.; ACHARYA, U. Rajendra. Advanced technique in breast thermography analysis. In: **2005 IEEE Engineering in Medicine and Biology 27th Annual Conference**. IEEE, p. 710-713, 2006.

NG, E.Y. K.; SUDHARSAN, N. M. Effect of blood flow, tumor and cold stress in a female breast: A novel time-accurate computer simulation. Proceedings of the Institution of Mechanical Engineers, **Journal of Engineering in Medicine**, v. 215, n. 4, p. 393–404, 2001.

NG, E. Y. K; SUDHARSAN, N. M. Computer simulation in conjunction with medical thermography as an adjunct tool for early detection of breast cancer. **BMC cancer**, v. 4, n. 1, p. 1-6, 2004.

NOGUEIRA, C. V.; NOGUEIRA, C. F.; ELY, J. B. Termografia por Infravermelho em Cirurgia Plástica-Novos Horizontes. **Pan American Journal of Medical Thermology**, v. 1, n. 2, p. 81-87, 2015.

NGUYEN, A. V. et al. Comparison of 3 infrared thermal detection systems and self-report for mass fever screening. **Emerging infectious diseases**, v. 16, n. 11, p. 1710, 2010.

OHASHI, Y.; UCHIDA, I. Some considerations on the diagnosis of breast cancer by thermography in patients with nonpalpable breast cancer. In: **Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering'**(Cat. No. 97CH36136). IEEE, p. 670-672, 1997.

OHUCHI, N. et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial. **The Lancet**, v. 387, n. 10016, p. 341-348, 2016.

OKAJIMA, J.; MARUYAMA, S.; TAKEDA, H.; KOMIYA, A. Dimensionless solutions and general characteristics of bioheat transfer during thermal therapy. **Journal Thermal of Biology**, v. 34, p. 377-384, 2009.

OLIVEIRA, M. M. **Desenvolvimento de protocolo e construção de um aparato mecânico para padronização da aquisição de imagens termográficas de mama**. p. 104. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2012.

OLIVEIRA, L. F. M. **Avaliação de técnicas de processamento de imagem e extração de características em imagens de pulmão e de coluna vertebral**. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) - Universidade Estadual de Londrina, 2015.

OMRANIPOUR, R. et al. Comparison of the accuracy of thermography and mammography in the detection of breast cancer. **Breast Care**, v. 11, n. 4, p. 260-264, 2016.

OSELAME, G. B.; NEVES, E. B.; SANCHES, I. J. Desenvolvimento de hardware e software para acompanhamento de lesões dermatológicas. **XXIV Congresso Brasileiro de Engenharia Biomédica**. p. 110–113, 2014.

PINHO, S. R. A. **Apoio ao Diagnóstico de Imagens de Mamografia por Técnicas Computacionais**. Dissertação (Mestrado) - Universidade do Porto, 2015.

POPOLIN, C. P. **Avaliação da atividade antitumoral in vitro e in vivo de novos complexos de rutênio contendo ácido salicílico e derivados**. Dissertação (Mestrado) - Universidade Federal de São Carlos, 2016.

PRAMANIK, S. et al. A computer-aided hybrid framework for early diagnosis of breast cancer. In: **Advanced Computing and Systems for Security**. Springer, Singapore, p. 111-124, 2019.

PRATI, R.C.; BATISTA, G. E. A. P. A; MONARD, M. C. Class imbalances versus class overlapping: an analysis of a learning system behavior. In: **Mexican international conference on artificial intelligence**. Springer, Berlin, Heidelberg, p. 312-321, 2004.

QUEIROZ, K. F. F. C. **Desenvolvimento e implementação de uma ferramenta computacional de uso médico para análise de imagens termográficas**. Dissertação (Mestrado) - Universidade Federal de Pernambuco, 2016.

RAGHAVENDRA, U. et al. Computer-aided diagnosis for the identification of breast cancer using thermogram images: A comprehensive review. **Infrared Physics & Technology**, v. 102, p. 103041, 2019.

RAMÍREZ-TORRES, A. et al. The role of malignant tissue on the thermal distribution of cancerous breast. **Journal of Theoretical Biology**, v. 426, p. 152-161, 2017.

RAVANELLI, N.; BONGERS, C. C. W. G.; JAY, O. The biophysics of human heat exchange. In: **Heat Stress in Sport and Exercise**. Springer, Cham, p. 29-43, 2019.

RESMINI, R. et al. Diagnóstico precoce de doenças mamárias usando imagens térmicas e aprendizado de máquina. **Revista Eletrônica do Alto Vale do Itajaí**, v. 1, n. 1, p. 55-67, 2012.

RESMINI, R. **Classificação de doenças de mama usando imagens térmicas**. Tese (Doutorado) - Universidade Federal Fluminense, 2016.

RING, E. F. et al. Quantitation of thermography in arthritis using multi-isothermal analysis. II. Effect of nonsteroidal anti-inflammatory therapy on the thermographic index. **Annals of the rheumatic diseases**, v. 33, n. 4, p. 353, 1974.

RODRIGUES, A. L. et al. Identification of mammary lesions in thermographic images: feature selection study using genetic algorithms and particle swarm optimization. **Research on Biomedical Engineering**, v. 35, n. 3, p. 213-222, 2019.

ROUT, N.; KUHO, D. M.; MALLICK, M. K. Analysing the Multi-class Imbalanced Datasets using Boosting Methods and Relevant Information. **International Journal of Pure and Applied Mathematics**, v. 120, n. 6, p. 6171-6191, 2018.

SÁEZ, J. A.; KRAWCZYK, B.; WOŹNIAK, M. Analyzing the oversampling of different classes and types of examples in multi-class imbalanced datasets. **Pattern Recognition**, v. 57, p. 164-178, 2016.

SAID, M. A. Novel Vision-based Thermal People Counting Tool for Tracking Infected People with Viruses Like COVID-19. **Journal of Advanced Research in Dynamical and Control Systems**, v. 12, n. SP7, p. 1115–1119, 2020.

SANTANA, M. A. d. et al. Breast cancer diagnosis based on mammary thermography and extreme learning machines. **Research on Biomedical Engineering**, v. 34, p. 45-53, 2018.

SANTANA, M. A. d. **Sistemas inteligentes para apoio ao diagnóstico do câncer de mama usando imagens mamográficas e termográficas**. Dissertação (Mestrado) - Universidade Federal de Pernambuco, 2020.

SANTOS, E. B. d.; BIANCO, H. T.; BRIOSCHI, M. L. Thermography in assessing cardiovascular risk. **Pan American Journal of Medical Thermology**, v. 2, n. 1, p. 23-25, 2015.

SANTOS, H. T. A. d. Deficiências da estatística Kappa na concordância entre avaliadores e medidas alternativas. 2015.

SANTOS, L. C. **Plataforma computacional para análise de imagens termográficas visando a detecção e a caracterização de anomalias mamárias**. p. 128. Tese (Doutorado) - Universidade Federal de Pernambuco, Recife, 2014.

SATHISH, D. et al. Asymmetry analysis of breast thermograms using automated segmentation and texture features. **Signal, Image and Video Processing**, v. 11, n. 4, p. 745-752, 2017.

SAXENA, A. K.; WILLITAL, G. H. Infrared thermography: experience from a decade of pediatric imaging. **European journal of pediatrics**, v. 167, n. 7, p. 757-764, 2008.

SCHAEFER, G. ACO classification of thermogram symmetry features for breast cancer diagnosis. **Memetic Computing**, v. 6, n. 3, p. 207-212, 2014.

SCHAEFER, G.; NAKASHIMA, T. Strategies for addressing class imbalance in ensemble classification of thermography breast cancer features. In: **2015 IEEE Congress on Evolutionary Computation (CEC)**. IEEE, p. 2362-2367, 2015.

SCHAEFER, G.; ZÁVIŠEK, M.; NAKASHIMA, T. Thermography based breast cancer analysis using statistical features and fuzzy classification. **Pattern recognition**, v. 42, n. 6, p. 1133-1137, 2009.

SCHIAVENATO, M.; THIELE, R. G. Thermography detects subclinical inflammation in chronic tophaceous gout. **The Journal of rheumatology**, v. 39, n. 1, p. 182-183, 2012.

SELVARANI, A.; SURESH, G. R. Infrared thermal imaging for diabetes detection and measurement. **Journal of medical systems**, v. 43, n. 2, p. 1-11, 2019.

SILVA, A. M. M. Infra-estrutura de informática para sistemas de apoio ao diagnóstico aplicada a servidores PACS. **Anais do 3o. Simpósio de Instrumentação e Imagens Médicas, 2007, Brasil**, 2007.

SILVA, A. S. V.; ARAÚJO, M. C. d.; LIMA, R. d. C. F. d. Avaliação da utilização da termografia como ferramenta de triagem de câncer de mama em mulheres residentes em regiões com

poucos recursos médicos. **Encontro Nacional de Engenharia Biomecânica - ENEBI 2015**, 2015.

SILVA, J. D. d. **Modelagem computacional dos efeitos da termoterapia transpupilar em tumores oculares através do método dos volumes finitos em malhas não estruturadas**. Tese (Doutorado) - Universidade Federal de Pernambuco, 2012.

SILVA, L. F. et al. Hybrid analysis for indicating patients with breast cancer using temperature time series. **Computer methods and programs in biomedicine**, v. 130, p. 142-153, 2016.

SILVA, M. C. d. **Comparação de algoritmos de segmentação de imagens em cenários não típicos**. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) - Universidade Estadual de Londrina, 2017.

SILVA, R. C. F.; HORTALE, V. A. Rastreamento do câncer de mama no Brasil: quem, como e por quê? **Revista Brasileira de Cancerologia**, v. 58, n. 1, p. 67-71, 2012.

SILVA, S. V. **Reconstrução da geometria da mama a partir de imagens termográficas**. Tese (Doutorado) - Universidade Federal Fluminense, Niterói, 2010.

SILVA, T. A. E. d. et al. A computational method to assist the diagnosis of breast disease using dynamic thermography. **Sensors**, v. 20, n. 14, p. 3866, 2020.

SIVANANDAM, S. et al. Medical thermography: a diagnostic approach for type 2 diabetes based on non-contact infrared thermal imaging. **Endocrine**, v. 42, n. 2, p. 343-351, 2012.

SOARES, J. F.; SIQUEIRA, A. L. **Introdução à estatística médica**. COOPMED, ed. 2, Belo Horizonte, 2002.

SOLIZ, P. et al. Detection of diabetic peripheral neuropathy using spatial-temporal analysis in infrared videos. In: **2016 50th Asilomar Conference on Signals, Systems and Computers**. IEEE, p. 263-267, 2016.

SOUZA, M.V. **Termografia como exame complementar no diagnóstico diferencial de LER/DORT em trabalhadores bancários de Criciúma**. Trabalho de Conclusão de Curso (Bacharelado em Fisioterapia) - Universidade do Extremo Sul Catarinense, 2012.

SOUZA, C. B.; ANDREÃO, R. V.; SEGATTO, M. V. Processamento de sinais de ECG para geração automática de alarmes. In: **VI Workshop de Informática Médica-WIM2006**. 2006.

SPAK, D. A. et al. BI-RADS® fifth edition: A summary of changes. **Diagnostic and interventional imaging**, v. 98, n. 3, p. 179-190, 2017.

SUNG, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: a cancer journal for clinicians**, v. 71, n. 3, p. 209-249, 2021.

TOO, J. et al. EMG feature selection and classification using a Pbest-guide binary particle swarm optimization. **Computation**, v. 7, n. 1, p. 12, 2019.

TEIXEIRA, B. C. et al. Marcadores inflamatórios, função endotelial e riscos cardiovasculares. **Jornal vascular brasileiro**, v. 13, p. 108-115, 2014.

THIRUNAVUKKARASU, U. et al. Human tongue thermography could be a prognostic tool for prescreening the type II diabetes mellitus. **Evidence-Based Complementary and Alternative Medicine**, v. 2020, 2020.

TRAFARSKI, A. et al. The quality of diagnosis by IR thermography as a function of thermal stimulation in chosen medical applications. In: **9th International Conference on Quantitative InfraRed Thermography July**. 2008. p. 2-5.

VANZELER, F. J. A. **Uma comparação entre dois algoritmos de otimização global (Algoritmo genético e evolução diferencial) para inversão de coeficientes de reflexão em meios fraturados**. Dissertação (Mestrado) – Universidade Estadual de Campinas, Campinas, 1999.

VARDASCA, R.; VAZ, L.; MENDES, J. Classification and decision making of medical infrared thermal images. In: **Classification in BioApps**. Springer, Cham, p. 79-104, 2018.

VARGAS, J. V. C. et al. Normalized methodology for medical infrared imaging. **Infrared Physics & Technology**, v. 52, n. 1, p. 42-47, 2009.

VASCONCELOS, J. H. **Investigação sobre métodos de classificação para uso em termografia de mama**. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2017.

VASCONCELOS, J. H.; DOS SANTOS, W. P.; LIMA, R. de Cássia Fernandes. Analysis of methods of classification of breast thermographic images to determine their viability in the early breast cancer detection. **IEEE Latin America Transactions**, v. 16, n. 6, p. 1631-1637, 2018.

VIEIRA, S. **Introdução à bioestatística**. 4ª edição. Elsevier Brasil, 2011.

VILA NOVA, R. d. L. et al. **Uso de imagens termográficas de mama para análise de patologias através da comparação entre diversos classificadores estatísticos**. p. 107. Dissertação (Mestrado), Universidade Federal de Pernambuco, Recife, Brasil, 2017.

WAHAB, A. A. et al. Thermal distribution analysis of three-dimensional tumor-embedded breast models with different breast density compositions. **Medical & biological engineering & computing**, v. 54, n. 9, p. 1363-1373, 2016.

WEBB, A. R.; COPSEY, K. D. **Statistical pattern recognition**. 3 ed., Wiley, 2011

WEI, X. et al. Quantitative analysis of MRI signal abnormalities of brain white matter with high reproducibility and accuracy. **Journal of Magnetic Resonance Imaging**, v. 15, n. 2, p. 203–209, 2002.

WU, C. L. et al. The application of infrared thermography in the assessment of patients with coccygodynia before and after manual therapy combined with diathermy. **Journal of manipulative and physiological therapeutics**, v. 32, n. 4, p. 287-293, 2009.

XIE, F.; BOVIK, A. C. Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm. **Pattern Recognition**, v. 46, n. 3, p. 1012–1019, 2013.

YANG, L. et al. A supervised approach to the evaluation of image segmentation methods. In: HLAVÁČ, V.; ŠÁRA, R. **Computer Analysis of Images and Patterns**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, v. 970, p. 759–765, 1995.

YODER, B. J.; WILKINSON, Edward J.; MASSOLL, Nicole A. Molecular and morphologic distinctions between infiltrating ductal and lobular carcinoma of the breast. **The breast journal**, v. 13, n. 2, p. 172-179, 2007.

ŽIVCÁK, J.; MADARASZ, L.; HUDAK, R. Application of medical thermography in the diagnostics of Carpal Tunnel Syndrome. In: **2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI)**. IEEE, p. 535-539, 2011.

## APÊNDICE A – TERMOGRAFIA APLICADA ÀS CIÊNCIAS DA SAÚDE

Desde 400 a.C., a temperatura da pele é considerada um aspecto fisiológico do corpo humano que indica um estado saudável ou patológico (HOUDAS; RING, 1982 apud LAHIRI et al., 2012). Em 1934, Hardy descreveu o papel fisiológico da emissão de radiação infravermelha pelo corpo humano e estabeleceu a importância do diagnóstico a partir da temperatura medida pela técnica de infravermelho (HARDY, 1934 apud HARDY; MUSCHENHEIM, 1936). Porém, apenas em 1956, Ray Lawson realizou a primeira aplicação médica da termografia por infravermelho na análise da distribuição de temperatura de uma mama com tumor maligno (VARDASCA; VAZ; MENDES, 2018). Desde então, esta técnica vem mostrando grande potencial nas aplicações clínicas e seu uso vem crescendo amplamente.

Atualmente, as imagens termográficas são geradas a partir de uma câmera que é sensível à radiação infravermelha. Os seus sensores captam a radiação térmica gerada pelo objeto, neste caso o corpo humano, que apresenta temperatura acima do zero absoluto (BORCHARTT, 2013). A quantidade de energia registrada é convertida em um sinal que é utilizado para calcular a temperatura real do corpo. Entretanto, além da radiação devido à temperatura do objeto, chega à câmera radiação emitida e a refletida pelo mesmo proveniente de fontes adjacentes e a radiação emitida pela atmosfera.

Nos últimos anos, as câmeras termográficas têm sido utilizadas com êxito no auxílio à detecção do câncer de mama, da neuropatia diabética e desordem vascular, no estudo da termorregulação, na triagem da febre, na termoencefaloscopia (imagens cerebrais), na odontologia, na dermatologia, no diagnóstico de doenças reumatológicas e osteomusculares, no diagnóstico da síndrome do olho seco, no tratamento de doenças hepáticas e na detecção de metástase no fígado, na isquemia intestinal, no transplante renal, no tratamento de doenças cardíacas e na ginecologia. As imagens termográficas têm sido usadas, também, no tratamento com acupuntura, na crioterapia, na medicina legal, nas cirurgias plásticas e na avaliação dos danos da radiação no corpo humano (AMMER & RING, 1995; JUNG et al., 2003 apud LAHIRI et al., 2012).

Na área da fisioterapia, Trafarski et al. (2008) usaram termografia para avaliar a eficácia do procedimento local de crioterapia em pacientes com artrite reumatoide. Para isso, eles determinaram a intensidade térmica durante o procedimento e a duração da resposta do corpo ao estímulo, que consistiu na aplicação de vapores de nitrogênio líquido ou ar fresco na mão. Uma comparação da distribuição de temperatura na superfície da mão foi feita em diferentes momentos: antes, imediatamente após e em determinados intervalos após a aplicação do estímulo.

A resposta do organismo à estimulação com vapores de nitrato líquido durou cerca de 2 horas e, no caso de ar frio, a duração do resfriamento foi de 90 minutos. A pesquisa mostrou que a termografia é uma técnica de medição viável para este procedimento e é útil na avaliação do progresso e monitoramento do tratamento.

Wu et al. (2009) estudaram em 53 pacientes com quadro de dor coccígea previamente diagnosticada a potencial utilidade da imagem térmica na avaliação das respostas fisiológicas locais antes e depois de terapias conservadoras. Os pacientes foram submetidos a modalidades terapêuticas (terapia manual + fisioterapia por diatermia por ondas curtas) realizada três vezes por semana, durante oito semanas. Escala numérica de classificação da dor e termografia antes do tratamento e em doze semanas foram utilizadas para avaliar os voluntários. Foram observadas diferenças significativas em ambos os critérios de avaliação: relato de dor na escala numérica de classificação (de 6,15 para 2,70 com valor de  $p < 0,05$ ) e temperatura superficial da pele na termografia (de 30,16 °C para 28,70 °C com valor de  $p < 0,05$ ). A correlação entre a melhora da dor e diminuição da temperatura foi positiva e estatisticamente significativa ( $r = 0,67$ ,  $p < 0,01$ ). Dessa forma, os achados desse estudo mostram que a termografia pode distinguir objetiva e quantitativamente o decréscimo das temperaturas da superfície, de modo a se correlacionar com as mudanças na intensidade da dor após o tratamento da síndrome da dor coccígea.

Na avaliação prognóstica, observou-se que a termografia pode distinguir objetiva e quantitativamente a diminuição da temperatura da superfície, de modo a se correlacionar com as alterações na intensidade da dor após a fisioterapia terapêutica, como nos casos de lesão por esforço repetitivo (MAGAS et al., 2012; BRIOSCHI; OKIMOTO; VARGAS, 2012; SOUZA, 2012).

Na odontologia, a termografia pode complementar o exame físico anatômico, pois se apresenta em condições fisiológicas em tempo real, de acordo com a temperatura (AMORIM et al., 2018). Tem-se mostrado uma técnica eficaz quando se trata do diagnóstico complementar na monitorização do complexo crânio-cervical-mandibular (BARROS et al., 2018). Assim, a inspeção térmica é capaz de auxiliar no diagnóstico e planejamento do tratamento de alterações orofaciais, como processos inflamatórios e infecciosos na região orofacial, acompanhamento pós-cirúrgico e na avaliação de rotina de pacientes com disfunção temporomandibular (DTM) (HADDAD; BRIOSCHI; ARITA, 2012; DIBAI-FILHO et al., 2014).

Gratt et al. (1994 apud HADDAD; BRIOSCHI; ARITA, 2012), desenvolveu um novo sistema de classificação de dor orofacial crônica usando termogramas. Os autores utilizaram um total de 80 pacientes: 50 com articulação temporomandibular com desarranjo interno ou osteoartrose e 30 pacientes sem alteração alguma. A classificação baseou-se na diferença

térmica entre os lados direito e esquerdo de cada indivíduo. Os termogramas foram identificados como normal, quando a variação térmica estava entre 0 e 0,25 °C, como quente (maiores que 0,35 °C) ou frio (menores que 0,35 °C). Os autores relataram que os termogramas quentes têm o diagnóstico clínico de dor mantida simpaticamente, mediada por nervos periféricos, artropatia da articulação temporomandibular (ATM) ou sinusite maxilar, enquanto os termogramas frios possuem o diagnóstico clínico de dor mediada por nervos periféricos ou dor simpaticamente independente. Esta nova metodologia foi considerada precisa em 92% dos casos.

Canavan et al. (1995 apud HADDAD; BRIOSCHI; ARITA, 2012) estudaram um grupo de indivíduos com desordem temporomandibular (DTM) leve a moderada e relataram que os valores da diferença térmica entre os lados direito e esquerdo estão correlacionados com o nível de dor nos indivíduos.

Em um estudo preliminar usando a termografia como diagnóstico complementar para DTM mio gênica, as regiões dos músculos mastigatórios de 23 mulheres foram avaliadas por projeções laterais e exame clínico, dividindo-se da seguinte forma: pacientes com e sem a patologia. Os autores descobriram que as regiões do masseter e dos músculos temporais anteriores apresentavam temperatura mais baixa na presença de DTM mio gênica quando comparadas ao grupo controle. Eles também afirmam que a termografia é um método útil, não-invasivo e não-ionizante que pode auxiliar no diagnóstico dessa disfunção (HADDAD et al., 2014). Outras áreas da odontologia também podem se beneficiar das pesquisas termográficas, como a endodontia, a periodontia, a estomatologia, a patologia oral, a implantologia e a cirurgia oral (AMORIM et al., 2018).

Na faixa etária pediátrica, principalmente em crianças, o diagnóstico clínico relacionado à temperatura é ainda mais difícil devido ao desconforto do paciente e às dificuldades de comunicação. Assim, a termografia pode ser de grande utilidade para esse grupo. Nessa perspectiva, Saxena e Willital (2008) revisaram 483 imagens térmicas de 285 crianças e identificaram que hemangiomas, malformações vasculares e amputações de dedos relacionadas a reimplantes e a queimaduras em recém-nascidos com estados clínicos na faixa etária pediátrica podem se beneficiar dos exames termográficos. Na sala de emergência, demonstrou ser uma ferramenta valiosa para o diagnóstico rápido de trombose de extremidades, varicoceles, inflamação, abscessos, gangrena e infecções de feridas.

No planejamento da cirurgia plástica deve-se tomar conhecimento da anatomia da circulação dos tecidos que serão mobilizados. Para este fim, Nogueira, Nogueira e Ely (2015) propuseram a integração da termografia por infravermelho à rotina de realização de imagens para diagnóstico e prognóstico. Os resultados obtidos sugerem que esta técnica possui atuação

muito positiva nas cirurgias plásticas, demonstrando ser um método prático, rápido e eficaz que pôde minimizar os índices de perdas teciduais e melhorar a qualidade dos resultados cirúrgicos e pós-cirúrgicos (por exemplo, o acompanhamento de úlceras de pressão).

Ainda dentro da cirurgia plástica, a avaliação da área e profundidade de queimaduras é fundamental para o tratamento e prognóstico dos pacientes. Mason, Graff e Pegg (1981), diante disso, utilizaram a termografia para estimar a profundidade da lesão de queimadura. Nos 12 pacientes avaliados, as queimaduras eram pequenas em termos de área de superfície corporal percentual, mas envolviam áreas de importância anatômica e funcional primordial. As mãos estiveram envolvidas em 8 dos 12 casos, os pés em 2, a axila em 3 e 1 outro caso envolveu uma pequena área de queimadura profunda na parte média da panturrilha. Quanto a estimativa da profundidade da lesão, o diagnóstico clínico apresentou uma precisão de 58% e a termografia aumentou para 83%. Os resultados desse estudo indicam que o diagnóstico termográfico possui um papel definitivo no diagnóstico rápido não-invasivo e preciso de queimaduras profundas de espessura parcial e total podendo, assim, auxiliar na decisão de fornecer a esses pacientes o benefício de uma excisão precoce e de um enxerto de pele. Isso pode, por sua vez, encurtar sua hospitalização e dar melhores resultados funcionais.

Os termogramas também são utilizados de forma complementar para avaliar distúrbios e lesões do sistema músculo-esquelético. Seu uso está relacionado à identificação de qual parte do sistema é afetada por uma doença ou lesão ou para avaliar a eficácia de um tratamento. Hildebrandt, Raschner e Ammer (2010), por exemplo, usaram termografia para avaliar lesões por esforço repetitivo no joelho e tendinopatia patelar em esquiadores alpinos. Um grupo de quinze atletas foi acompanhado. Desse grupo, sete atletas relataram sintomas de reações locais no joelho e os demais eram assintomáticos. Atletas sintomáticos apresentaram diferenças de temperatura média entre o joelho sintomático e o joelho contralateral de  $1,4 \pm 0,58$  °C, enquanto atletas não-sintomáticos apresentaram variação de temperatura média entre os dois joelhos de  $0,3 \pm 0,61$  °C.

A fase de regeneração óssea pós-fratura é caracterizada por um aumento no metabolismo e, conseqüentemente, na temperatura dos tecidos locais. Com base nesse fato, Morasiewicz et al. (2010) avaliaram a correlação da análise termográfica com a evidência radiográfica. Para isso, acompanharam o caso de 18 pacientes em tratamento de distração osteogênica para alongamento ósseo com fixador externo Ilizarov. Os resultados indicaram correlação estatisticamente significativa entre regeneração óssea e índices térmicos. Os valores médios do coeficiente de correlação de Spearman para a tibia e o fêmur foram de 0,925 ( $p < 0,01$ ) e 0,724 ( $p < 0,05$ ), respectivamente. A termografia provou ser um complemento valioso aos métodos tradicionais

de diagnóstico de fraturas e pode ser usada para monitorar e avaliar a formação e remodelação óssea em todas as etapas do tratamento.

Também dentro da ortopedia e da reumatologia, Zivcak, Madarasz e Hudak (2011) utilizaram a termografia para avaliar a distribuição da temperatura da pele do dorso da mão de portadores de síndrome do túnel do carpo em comparação com indivíduos saudáveis. O estudo indicou maior temperatura dos dedos comparada à temperatura dos punhos nos indivíduos portadores da síndrome ao contrário dos indivíduos saudáveis que apresentaram extremidades com temperaturas mais baixas do que as temperaturas dos punhos.

A reação inflamatória é conhecida há muito tempo. Na Antiguidade, os gregos já a definiam pelos seus sinais e sintomas típicos, considerados sinais cardinais: calor, rubor, tumor e dor, aos quais os médicos romanos acrescentaram as alterações funcionais, que muitas vezes acompanham as inflamações, sobretudo, aquelas de caráter agudo (BRASILEIRO FILHO, 2016). Seguindo essa concepção, diversos estudos empregaram as imagens por infravermelho na avaliação diagnóstica e no acompanhamento de tratamento nos casos de patologias inflamatórias agudas, como o fenômeno de Raynaud, a gota e a artrite (RING et al., 1974; CHERKAS et al., 2003; CHIKURA et al., 2010; SCHIAVENATO & THIELE, 2012; HEIMBECHER & ULBRICHT, 2018).

As doenças de pele, em geral, causam inflamação que por sua vez causa um padrão de temperatura anormal na superfície da pele. A termografia tem sido usada para o diagnóstico de pacientes com hanseníase e foi relatado que áreas mais frias (como orelha e região da borda nasal) são fortemente afetadas (HASTINGS et al., 1968). A avaliação termográfica também tem sido empregada na avaliação terapêutica e remissão de doenças da pele, mostrando-se superior quando comparada aos métodos convencionais. No caso da hanseníase, os resultados infravermelhos normalizados demonstraram o sucesso do tratamento desde a partir do 87º dia, enquanto a repigmentação foi avaliada visualmente somente após o dia 182 (VARGAS et al., 2009).

Diabetes é uma doença crônica não-transmissível que vem apresentando aumento da incidência e, conseqüentemente, da prevalência em todo o mundo. Em seus estágios mais graves podem causar paralisia e problemas nas pernas. As principais razões para esses problemas são o baixo fluxo sanguíneo, conhecido como distúrbio vascular, e a perda de sensibilidade ou fraqueza, também conhecida como neuropatia na terminologia médica. Neste tipo de doença, os pacientes normalmente apresentam temperatura anormal da pele, tornando a termografia uma ferramenta apropriada para o diagnóstico de doenças vasculares ou de neuropatologias (SELVARANI & SURESH, 2019; FAROOQ & CORCORAN, 2020).

Assim, a termografia infravermelha pode desempenhar um papel vital no diagnóstico inicial de diabetes no corpo humano, que acabará por ajudar os médicos e especialistas a fornecer o tratamento adequado aos seus pacientes.

Soliz et al. (2016) desenvolveram um sistema de imagem ótica térmica capaz de detectar os primeiros sintomas da neuropatia periférica diabética na sola do pé, que é responsável por cerca de 25% das internações em pacientes com diabetes. O sistema funciona gravando um vídeo com câmera térmica do pé do paciente durante a recuperação do *stress* pelo frio. O sistema funciona extraindo as pós-imagens, conhecidas como sinais funcionais, para detectar mudanças dinâmicas no fluxo sanguíneo microvascular. De acordo com sua pesquisa inicial, o sistema pode mostrar diferenças estatísticas visíveis e significativas entre pacientes normais e indivíduos que foram diagnosticados com neuropatia periférica.

Indivíduos com diabetes normalmente são hospitalizados devido a complicações tais como úlceras nas extremidades do corpo. Hosaki et al. (2002) investigaram a circulação periférica em um grupo de diabéticos por meio de imagens térmicas e observaram que gradientes de temperatura, indicando fluxo anormal nas regiões afetadas, estão correlacionados com achados clínicos. Por sua vez, Liu et al. (2013) integraram a câmera termográfica à rotina clínica para exames frequentes nos pés dos pacientes e desenvolveram uma ferramenta quantitativa para análise dos termogramas. Nesse caso, o uso da termografia possibilitou a identificação de riscos de úlceras ou outras complicações a partir do aumento/diminuição da temperatura nos membros avaliados.

Sivanandam et al. (2012) avaliaram a aplicação da termografia, tanto para diagnóstico como prevenção de diabetes tipo 2, em comparação com ensaios bioquímicos de hemoglobina glicada. No grupo de diabéticos, a hemoglobina glicada mostrou correlação negativa com a temperatura na região da carótida ( $r = -0,471$ ,  $p = 0,01$ ) e a temperatura média da pele foi mais baixa do que o grupo normal em regiões do corpo como joelho ( $p = 0,002$ ), tíbia ( $p = 0,003$ ), testa ( $p = 0,014$ ) e palma da mão ( $p = 0,019$ ). O estudo concluiu que com o aumento da hemoglobina glicada a temperatura da pele diminui, sendo que tal diminuição da temperatura da pele pode decorrer da diminuição da taxa metabólica basal, diminuição da perfusão sanguínea e da elevada resistência à insulina. Para esses autores, a termografia pode ser usada como uma ferramenta de diagnóstico, bem como de prognóstico para o diabetes tipo 2.

Baseando-se na medicina oriental que afirma que a língua humana é capaz de refletir condições fisiopatológica do corpo, Thirunavukkarasu et al. (2020) capturaram a imagem infravermelho da língua de 140 voluntários a fim de classificá-los como diabéticos ou saudáveis. O sistema desenvolvido pelos autores apresentou uma precisão diagnóstica de 94,28%. A

temperatura medida da superfície da língua do grupo de diabéticos foi maior do que o normal, apresentando uma correlação estatística positiva ( $r = 0,5688$ ) entre a hemoglobina glicada e a distribuição térmica na região da língua.

Além de ser utilizada em estudos envolvendo indivíduos com coração transplantado (MANGINAS et al., 2010), a termografia pode auxiliar também na detecção de sinais iniciais de aterosclerose e, assim, atuar como um indicador precoce de ataque cardíaco. Nas pesquisas de Santos, Branco e Brioschi, (2015) foi investigado o uso desta técnica na avaliação da função endotelial e a correlacionou com o risco de eventos cardiovasculares. Fisiologicamente, o endotélio desempenha papel protetor do vaso sanguíneo, sendo também responsável por regular a homeostase vascular. Uma parede endotelial íntegra é essencial para a saúde vascular. Alguns fatores são responsáveis pela integridade deste tecido, como o óxido nítrico que provoca a dilatação do vaso sanguíneo em resposta ao estresse de cisalhamento. A manutenção do sistema vascular é feita pela expressão da enzima óxido nítrico sintase endotelial, que produz óxido nítrico em resposta ao incremento do fluxo sanguíneo. Quando há inibição desta enzima, ocorre diminuição da produção de óxido nítrico, causando disfunção endotelial e eventos cardiovasculares (TEIXEIRA et al., 2014).

Algumas doenças, como hipertensão arterial, diabetes *mellitus*, insuficiência cardíaca e hipercolesterolemia, podem causar danos ao endotélio, gerando disfunção endotelial, que, muitas vezes, está relacionada à aterosclerose e aos eventos cardiovasculares. Os resultados do trabalho de Santos, Branco e Brioschi (2015) mostraram a existência de uma forte correlação entre os achados termográficos (por exemplo, hiperatividade simpática e temperatura máxima de isquemia) e o risco cardiovascular. Sendo assim, o estudo da função endotelial por termografia pode ser utilizado como um marcador complementar para melhorar a estratificação do risco cardiovascular e pode também ser útil para a avaliação das estratégias terapêuticas.

Os termogramas também podem ser usados no monitoramento da pressão arterial e na avaliação de terapias anti-hipertensivas mais eficazes. Dois grupos diferentes de hipertensos, um tratado com uma combinação de atenolol mais clortalidona e o outro tratado com labetalol mais clortalidona foram monitorados por Cesaris et al. (1985). Por meio da termografia, foi observada hipotermia nas mãos dos indivíduos em ambos os grupos antes do tratamento. Após 1 mês de tratamento, nenhuma mudança no fluxo vascular periférico pôde ser observada no primeiro grupo, enquanto melhorias significativas no fluxo vascular foram observadas no segundo grupo.

Existe a necessidade de técnicas de rastreamento eficazes para que as pandemias de doenças infecciosas sejam prevenidas rapidamente. Através de termogramas do pescoço e do

rosto, Nguyen et al. (2010) inferiram que além de sensível, a termografia é um método rápido e não necessita do contato com a pele do paciente para monitoramento da febre. A alta temperatura corporal é um dos sintomas mais comuns em diversas doenças, inclusive nas infecciosas. Portanto, o método citado torna-se uma ferramenta poderosa para realização de uma triagem inicial de determinado grupo de pessoas (NG, 2005). Em 2009, Nishiura e Kamyia utilizaram a câmera termográfica para avaliação de passageiros no aeroporto internacional de Nauta, no Japão, durante a crise na saúde gerada pelo vírus Influenza A (H1N1). Estratégia semelhante foi utilizada em numerosos aeroportos na atual pandemia da COVID-19 (SAID, 2020). Ainda nessa perspectiva de rastreamento, Furuichi et al. (2015) indicaram a utilização deste mesmo equipamento para auxílio do controle prévio da chikungunya e da dengue.

A termologia é usada no campo da oftalmologia humana para o diagnóstico de síndromes de olho seco e problemas oculares, observando a fisiologia do olho. O processo de termografia infravermelha funciona detectando os comportamentos anormais da temperatura do olho seco ( $32,38 \pm 0,69$  °C) que costuma ser ligeiramente mais elevada em comparação com a temperatura de um olho saudável ( $31,94 \pm 0,54$  °C) (MORGAN; TULLO; EFRON, 1996).

O exame termográfico também pode ser usado com eficácia para diferentes tipos de detecção de câncer em estágios iniciais, o que inclui detecção de câncer de mama (SANTANA et al., 2018), detecção de câncer de pele (FLORES-SAHAGUN; VARGAS; MULINARI-BRENNER, 2011; NARAYANAMURTHY et al., 2018), avaliação de tumores malignos da tireoide (DAMIÃO et al., 2021) e diagnóstico de tumor cerebral (GORBACH et al., 2004).

Quadro A1 - Resumo dos estudos que aplicaram termografia na área das ciências da saúde  
 Continua

Área da saúde	Autor	Ano	Objetivo
Fisioterapia	Magas et al.	2019	Avaliar a viabilidade do uso de imagens termográficas no diagnóstico das lesões por esforços repetitivos (LER) nas articulações do punho e das mãos.
	Brioschi, Okimoto e Vargas	2012	Analisar o potencial da imagem térmica como auxiliar na investigação dos efeitos térmicos que o trabalho pode gerar na pele humana que podem refletir condições ocupacionais fisiológicas e/ou patológicas.
	Souza	2012	Verificar as relações entre os achados da termografia em bancários e a condição funcional do membro superior determinada pela força escapular, força de preensão manual e de pinça média.
	Wu et al.	2009	Analisar o potencial da imagem termográfica na avaliação das respostas fisiológicas locais antes e depois de terapias conservadoras em quadros de dor occípica.
	Trafarski et al.	2008	Utilização da termografia para avaliar a eficácia do procedimento local de crioterapia em pacientes com artrite reumatóide.
Odontologia	Amorim et al.	2018	Revisão integrativa sobre as aplicações da termografia na odontologia.
	Barros et al.	2018	Analisar e registrar os padrões térmicos para avaliar anatomicamente e fisiologicamente certas estruturas orofaciais do complexo crânio-cervical-mandibular, como os músculos mastigatórios e a região da articulação temporomandibular.
	Dibai-Filho et al.	2014	Investigar a correlação entre a duração da DTM e possíveis alterações da temperatura da pele na região dos músculos masseter e temporal anterior.
	Haddad et al.	2014	Comparar os padrões térmicos de regiões selecionadas da musculatura mastigatória de voluntários com e sem DTM mio gênica, por meio da termografia infravermelha.
	Haddad, Brioschi e Arita	2012	Identificar e correlacionar pontos-gatilho miofasciais nos músculos mastigatórios, por meio da termografia e da dolorimetria de pressão.
Odontologia	Canavan et al.	1995	Caracterizar termicamente, através do uso da termografia, indivíduos assintomáticos (controle) e um grupo de pacientes com distúrbios da articulação temporomandibular leve a moderado.
	Gratt et al.	1994	Avaliar a termografia como alternativa diagnóstica para análise das disfunções temporomandibulares.
Pediatria	Saxena e Willital	2008	Revisar as imagens termográficas e identificar estados clínicos na faixa etária pediátrica que se beneficiaram desse exame. Além disso, o estudo também visa determinar a viabilidade e os benefícios do emprego dessa técnica na faixa etária neonatal e pediátrica.

Quadro A1 - Resumo dos estudos que aplicaram termografia na área das ciências da saúde  
Continuação

Área da saúde	Autor	Ano	Objetivo
Cirurgia Plástica	Nogueira, Nogueira e Ely	2015	Demonstrar a importância da introdução rotineira de realização de imagens por termografia infravermelha em cirurgia plástica no diagnóstico e prognóstico e seguimento da especialidade.
	Mason, Graff e Pegg	1981	Utilizar a termografia para estimar a profundidade da lesão de queimadura.
Ortopedia e Reumatologia	Schiavenato & Thiele	2012	Relato de caso de uma avaliação termográfica de tofos gotosos em um homem de 42 anos com história de gota tofácea crônica.
	Zivcak, Madarasz e Hudak	2011	Avaliar a distribuição da temperatura da pele do dorso da mão de portadores de síndrome do túnel do carpo em comparação com indivíduos saudáveis.
	Chikura et al.	2010	Realizar um estudo retrospectivo, com uso da termografia, para avaliar se o grau de envolvimento do polegar difere entre o fenômeno de Raynaud primário e o fenômeno de Raynaud secundário.
	Hildebrandt et al.	2010	Investigar uso da termografia para avaliar lesões por esforço repetitivo no joelho e tendinopatia patelar em esquiadores alpinos.
	Morasiewicz et al.	2010	Estabelecer uma correlação entre o regenerado ósseo, avaliado com base em evidências radiográficas, e as imagens termográficas do membro registradas de forma não-invasiva com uma câmera termográfica.
Ortopedia e Reumatologia	Cherkas et al.	2003	Avaliar o valor das medidas termográficas da pele após teste do frio para a classificação do fenômeno de Raynaud em uma população saudável.
	Ring et al.	1974	Avaliar o papel da termografia como ferramenta capaz de medir o progresso terapêutico de pacientes portadores de artrite reumatóide e/ou gota em uso de agentes anti-inflamatórios não esteroidais orais.
Dermatologia	Lahiri et al.	2015	Investigar o aquecimento local da temperatura da região da bochecha e da orelha devido ao uso de celulares.
	Vargas et al.	2009	Descrição de um relato de caso de paciente com hanseníase e hepatite C que fez uso da termografia para avaliar progressão clínica.
	Hastings et al.	1968	Aplicar a termografia para o diagnóstico de pacientes com hanseníase.
Endocrinologia	Thirunavukkarasu et al.	2020	Analisar e classificar pacientes como diabéticos ou não com base nas variações térmicas na língua humana.
	Soliz et al.	2016	Desenvolvimento de um sistema de imagem ótica térmica capaz de detectar os primeiros sintomas da neuropatia periférica diabética no pé plantar.
	Liu et al.	2013	Utilização da câmera termográfica, associada a um <i>software</i> de classificação, em rotinas clínicas para identificação de riscos de úlceras ou de amputações em pacientes diabéticos.

Quadro A1 - Resumo dos estudos que aplicaram termografia na área das ciências da saúde  
Conclusão

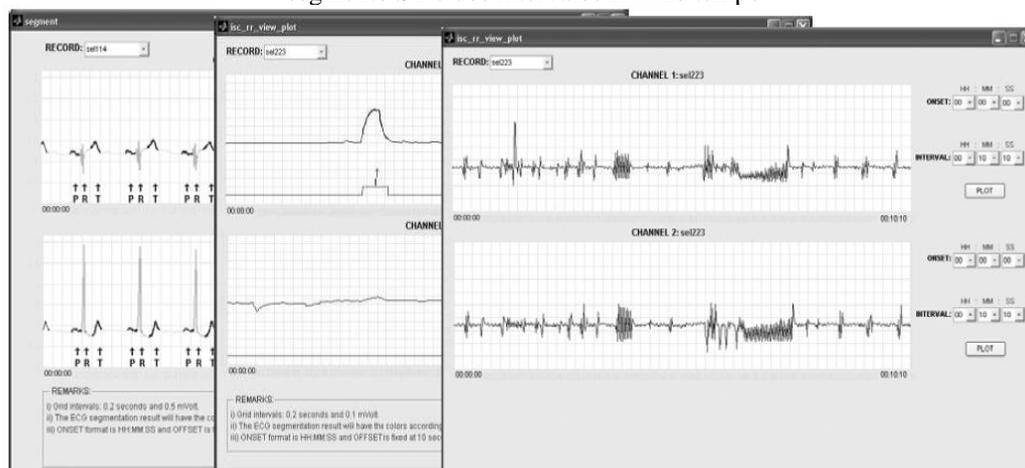
Área da saúde	Autor	Ano	Objetivo
Endocrinologia	Sivanandam et al.	2012	Avaliar a aplicação da termografia, tanto para diagnóstico como prevenção de diabetes tipo 2, em comparação com ensaios bioquímicos de hemoglobina glicada.
	Hosaki et al.	2002	Investigar a circulação periférica em pacientes diabéticos para identificação de gradientes de temperatura que sugerem achados clínicos por fluxo sanguíneo anormal.
Cardiologia	Santos, Branco e Brioschi	2015	Investigar o uso das imagens infravermelho na avaliação da função endotelial e a correlacionar com o risco de eventos cardiovasculares.
	Cesaris et al.	1985	Utilizar a termografia como ferramenta auxiliar na investigação de terapias anti-hipertensivas.
Infecologia	Said	2020	Utilização de câmeras termográficas para triagem de passageiros em aeroportos.
	Furuichi et al.	2015	Investigação do potencial auxílio da termografia no controle prévio da chikungunya e da dengue.
	Nguyen et al.	2010	Avaliação da febre por meio das imagens infravermelho.
	Nishiura e Kamyia	2009	Realização de triagem subclínica de infectados por H1N1 com uso de uma câmera termográfica.
Oftalmologia	Morgan, Tullo e Efron.	1996	Utilizar a termografia para medir a mudança na temperatura da superfície ocular na abertura do olho em um grupo de 11 pacientes com olho seco e sete controles.
Oncologia	Damião et al.	2021	Avaliar as diferenças de temperatura entre nódulos tireoidianos benignos e malignos através da termografia.
	Narayanamurthy et al.	2018	Elucidar sobre métodos não-invasivos de diagnóstico, incluindo a termografia, de lesões de pele
	Flores-Sahagun, Vargas e Mulinari-Brenner	2011	Estabelecer as faixas de normalidade esperadas e de pacientes com carcinoma basocelular previamente diagnosticados por meio de biópsias das regiões afetadas.
	Gorbach et al.	2004	Investigar os efeitos na fisiologia vascular decorrentes de um tumor cerebral.

Fonte: A Autora (2022)

## APÊNDICE B - INTERFACES GRÁFICAS NÃO-RELACIONADAS À TERMOGRAFIA

Souza, Andreão e Segatto (2006) implementaram e integraram rotinas de análise de eletrocardiograma (ECG), em ambiente Matlab®, para o monitoramento da atividade elétrica do coração. A interface gráfica desenvolvida proporciona a visualização do sinal de ECG, assim como os resultados dos algoritmos de segmentação por abordagem estatística Markoviana das ondas elementares do batimento cardíaco (P, QRS, T e linhas isométricas) e de classificação dos eventos cardíacos (batimentos ventriculares prematuros e isquemia). No uso da interface deve-se, primeiramente, executar a operação de inicialização das variáveis do sistema indicando o caminho onde estão salvos os registros, a pasta onde serão salvos os resultados e a base de dados escolhida. Em seguida, após a ação dos algoritmos citados, o usuário tem as seguintes opções de visualização dos resultados (Figura B1): sinal puro do ECG, sinal segmentado e sinal classificado. É possível, também, a exibição dos intervalos R-R e dos eventos de isquemia escolhendo-se o tamanho do bloco do sinal, de modo a facilitar a interpretação dos dados. Os autores concluíram que, apesar de não operar em tempo real, o aplicativo possibilitou a detecção de eventos importantes no monitoramento de pacientes com risco de infarto no miocárdio.

Figura B1 – Janelas da GUI com exemplos de visualização do sinal de ECG segmentado, do gráfico de desvio do segmento ST e dos intervalos R-R no tempo

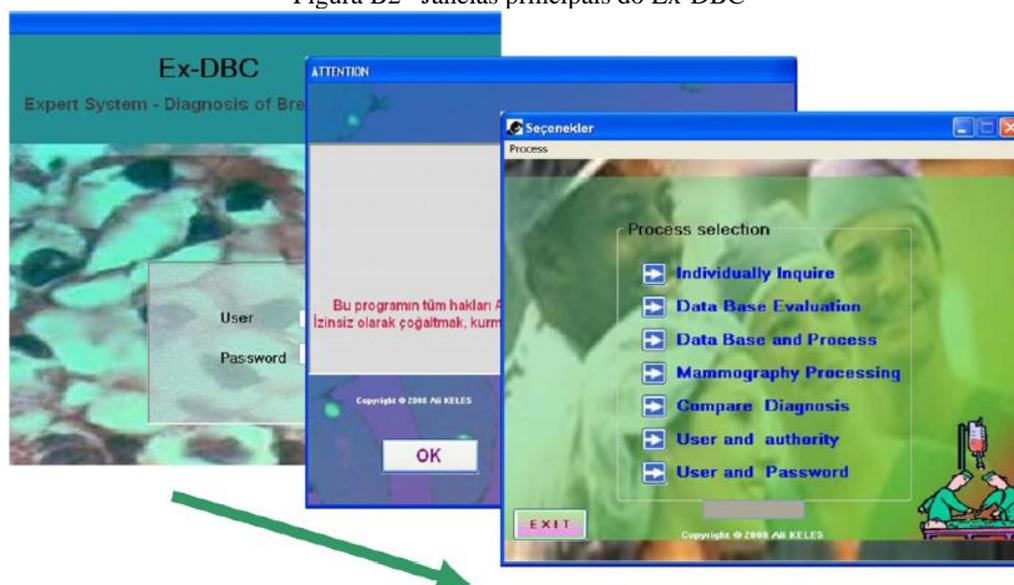


Fonte: Souza, Andreão, Segatto (2006)

Keleş, Keleş e Yavuz (2011) projetaram uma ferramenta computacional, Ex-DBC (*Expert system for Diagnosis of Breast Cancer*), capaz de diferenciar achados em imagens mamográficas em benignos e malignos usando o método neurofuzzy como máquina de inferência. O programa foi implementado em Visual C# e o SQL Server 2005 foi usado como uma plataforma de banco de dados. Na tela de entrada (Figura B2), o usuário pode inserir seu próprio nome de usuário e senha. Se o usuário não estiver cadastrado no sistema ou fizer uma entrada errada, o

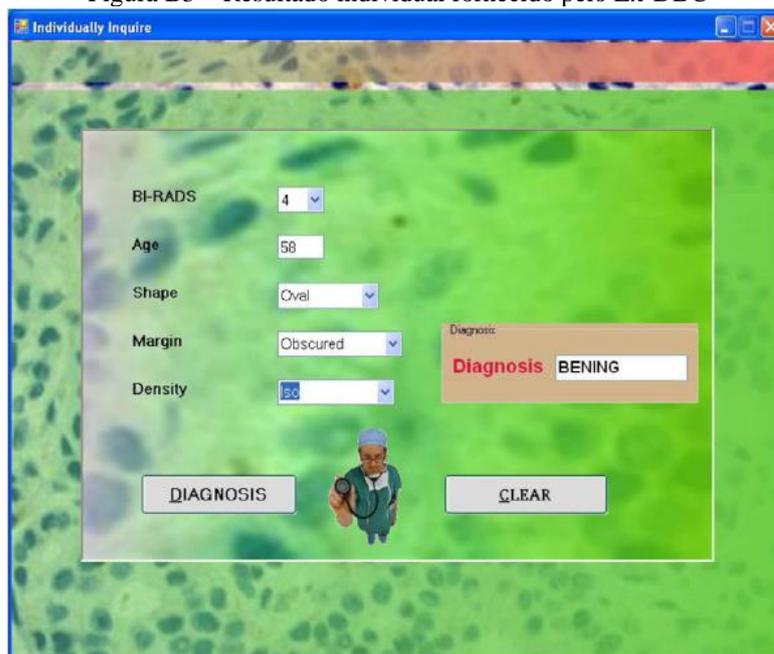
sistema não permite o acesso e retorna um aviso de segurança. Algumas informações explicativas relacionadas ao sistema aparecem na tela após a entrada. Para realização da análise da imagem mamográfica, são inseridas informações como idade da paciente, BI-RADS, forma, tipo de margem e densidade da lesão que são usadas para, por meio do método de classificação especificado, gerar o diagnóstico para esta paciente. Na opção “*Data base evaluation*” o usuário pode fazer uma avaliação diagnóstica para todas as pacientes registradas no banco de dados. É gerada uma lista de resultados para as pacientes desejadas. Além disso, informações estatísticas como número de casos benignos, malignos e o total de casos pode ser visto no topo desta tela. Apenas usuários autorizados podem usar o processo “*Data base and process*” que exibe uma lista de todas as pacientes na parte inferior da tela. A manutenção dos prontuários, a entrada de uma nova paciente, a busca de uma paciente já cadastrada ou a exclusão de uma paciente do cadastro são algumas das operações permitidas nesse módulo. Na opção “*Mammography processing*” é possível visualizar e exportar para processamento a imagem da mamografia de qualquer paciente. E a opção “*Compare diagnosis*” foi concebida com intuito educativo. Nela, o usuário pode selecionar pacientes para testar sua auto habilidade de diagnóstico de câncer de mama. Como resultado (Figura B3), o sistema Ex-DBC compara, estatisticamente, o diagnóstico do usuário com aquele gerado pelo sistema e o com o diagnóstico real. O botão Sair na tela do menu fecha todos os processos abertos no sistema.

Figura B2– Janelas principais do Ex-DBC



Fonte: Keleş, Keleş, Yavuz (2011)

Figura B3 – Resultado individual fornecido pelo Ex-DBC



Fonte: Keleş, Keleş, Yavuz (2011)

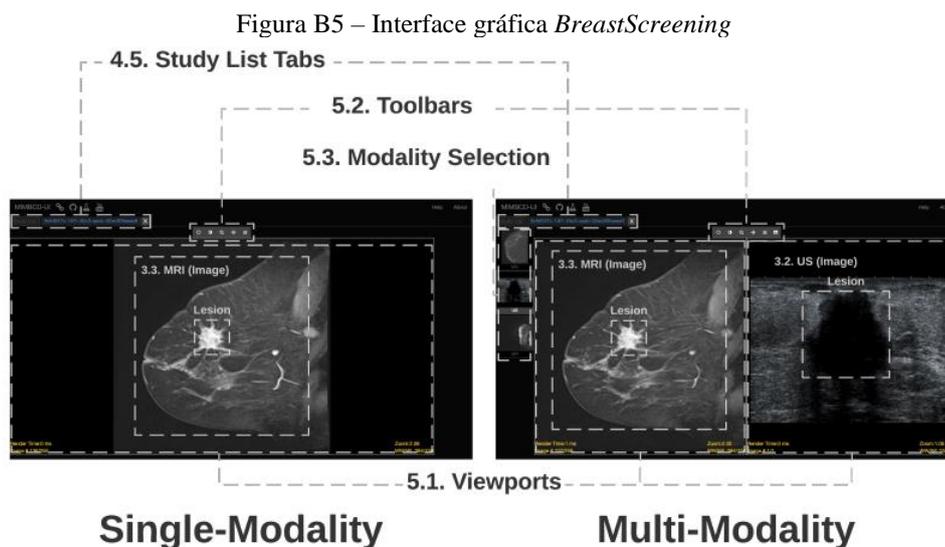
Oselame, Neves e Sanches (2014) propuseram um método para a medição de lesões dermatológicas por meio da análise de imagens captadas por um *hardware* desenvolvido para este fim. Este método estabelece o uso da fotogrametria, cuja finalidade é obter medidas da forma e das dimensões dos corpos ou de parte deles por imagens digitais. Lesões benignas costumam ter uma área inferior a  $10 \text{ mm}^2$ , enquanto melanomas usualmente são maiores. Utilizando o Matlab<sup>®</sup>, foi implementada uma interface gráfica para facilitar o uso da ferramenta computacional criada para medição da área de determinada lesão. Na interface é possível abrir a imagem desejada e, ao clicar no botão “Resultado em  $\text{cm}^2$ ”, converter a área calculada em *pixel* para centímetros quadrado. A Figura B4 representa o processo final realizado pelo *software* na GUI desenvolvida. Concluiu-se que a técnica proposta foi capaz de garantir precisão na medição das áreas das lesões, podendo associá-las ao fator de risco para melanomas, com baixo custo de utilização.

Figura B4 – Interface gráfica Perímetro



Fonte: Oselame, Neves, Sanches (2014)

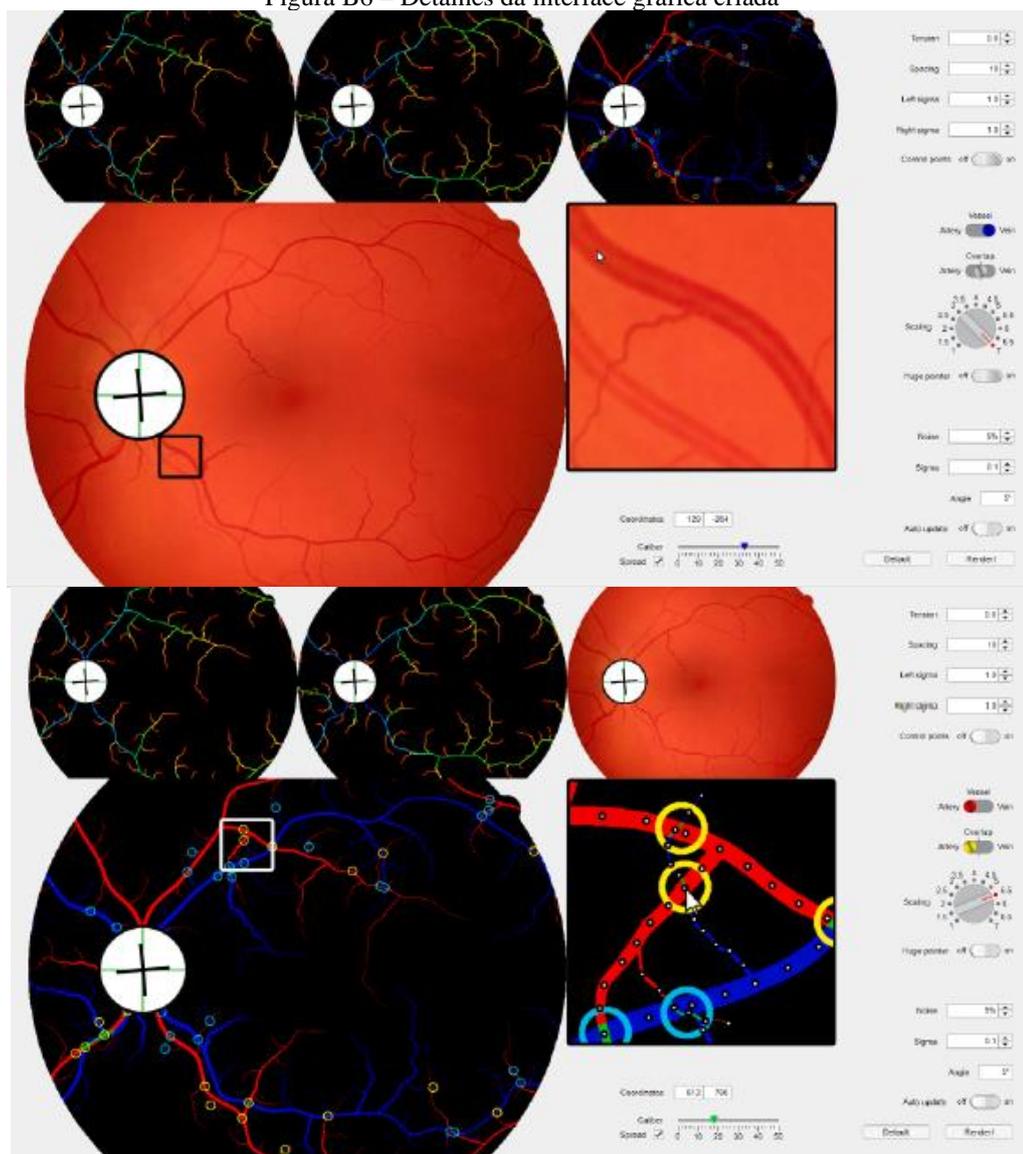
Calisto, Nunes e Nascimento (2020) descrevem o desenvolvimento e a implantação de uma interface de usuário (*BreastScreening*) de imagem médica multimodal para rastreamento de patologias da mama. O *BreastScreening* (Figura B5) foi desenhado para integrar informações de diferentes modalidades de imagem da mama – Mamografia, Ultrassonografia e Ressonância Magnética. De modo geral, o sistema desenvolvido pelos autores possui as seguintes funcionalidades: (1) uma interface para identificar (e registrar o *ground-truth*) até dois tipos de lesões (massas e calcificações) nas modalidades de imagem; (2) suporte para categorização dos tecidos mamários (denso e não denso); (3) um esquema de classificação e recomendação para a gravidade da lesão usando o sistema BI-RADS; (4) acesso imediato a covariáveis clínicas, como registros pessoais e familiares; e (5) visualizações adequadas para um diagnóstico e acompanhamento dos pacientes. De modo simplificado, a interface do usuário consiste em dois componentes principais: lista de visualizações do paciente e visualizações de diagnóstico por imagem médica. *BreastScreening* foi implementado usando CornerstoneJS (URBAN et al., 2017 *apud* CALISTO; NUNES; NASCIMENTO, 2020) com um servidor NodeJS, usado para executar código JavaScript fora do navegador. CornerstoneJS fornece funções essenciais, como renderização de imagem, suporte a ferramentas e recuperação e interpretação de DICOM (*Digital Imaging and Communications in Medicine*). Esse design também permite que o aplicativo seja facilmente estendido simplesmente adicionando novos pacotes.



Fonte: Calisto, Nunes, Nascimento (2020)

Lo Castro, Tegolo e Valenti (2020) desenvolveram uma metodologia mais simples, comparada as atuais existentes na literatura, para produção de imagens sintéticas do fundo de olho que permite a simulação de patologias e a previsão de doenças oculares. A abordagem proposta baseia-se no princípio de ação mínima e posiciona corretamente os vasos no fundo do olho simulado sem o uso de informações morfométricas reais. Uma análise a posteriori das características médias, como tamanho, comprimento, bifurcações e posicionamento do ponto final, confirmou a precisão substancial da abordagem proposta em comparação com dados reais. O método desenvolvido pelos autores foi associado a uma GUI (Figura B6), permitindo que o usuário faça quaisquer alterações em tempo real, controlando as posições dos pontos de controle. Esse sistema demonstrou fácil utilização e alta aplicabilidade na oftalmologia clínica, uma vez que existem poucas imagens retiniais de alta resolução de domínio público com *ground-truth* feitos à mão, especialmente aquelas que mostram a classificação de artérias e veias efetuada por médicos especialistas. Todo o *software* foi desenvolvido na linguagem Matlab<sup>®</sup>. O arcabouço computacional criado permite a escolha de quatro possíveis visualizações disponíveis para a imagem que vai ser passada pelo processamento (segmentação artéria/veia, sobreposições e calibração da imagem renderizada). O usuário pode criar vasos sanguíneos, além de poder mover cada ponto de controle usando o *mouse*. Também é possível ajustar o ruído a ser adicionado apenas aos vasos, para aumentar sua variabilidade e permitir que eles se misturem melhor com a imagem que representa o fundo.

Figura B6 – Detalhes da interface gráfica criada



Fonte: Lo Castro, Tegolo e Valenti (2020)

Majeed et al. (2020) motivados em construir um *software* que pudesse auxiliar os profissionais de saúde no diagnóstico precoce do câncer de pele, desenvolveram um sistema barato, fácil de usar e relativamente preciso para detectar câncer de pele em estágio inicial baseado em imagens digitais e técnicas de processamento de imagens. O sistema proposto para detecção de câncer de pele consiste em duas fases, a primeira fase é adicionar casos no arquivo treino e a segunda fase é testar o novo caso. O *software* proposto (Figura B7) compara a imagem da lesão cutânea de um novo caso (imagem alvo) com os casos antigos (imagem fonte) na base de dados, com cada caso representado com seu próprio peso e informações relacionadas à lesão cutânea. Assim, o sistema se propõe a encontrar o caso mais semelhante ao caso alvo, e reutilizar a solução do caso semelhante para o caso alvo. Por fim, após a conclusão do diagnóstico do novo caso, este é armazenado na base de dados a fim de ser utilizado em análises futuras. Na

utilização da GUI, o primeiro passo é preencher as informações clínicas do paciente e da lesão a ser investigada. Em seguida, é carregada a imagem da lesão alvo e, após a etapa de pré-processamento, são mostradas as lesões similares presentes na base de conhecimento. A quantidade de imagens similares a serem mostradas pode ser escolhida pelo operador. O resultado de casos semelhantes é mostrado na tabela com suas próprias imagens, o diagnóstico e a estatística de similaridade em relação à imagem alvo.

Figura B7 – Sistema proposto por Majeed et al. (2020)

The screenshot shows a software interface titled "skin tumor detection". It features several input fields for clinical data, a grid of 10 skin lesion images, and a table of search results. The input fields are organized into two rows:

- Row 1: Duration of the lesion (4), Diameter (4.3), Skin type (2), Border of the lesion (0), Nature of work (1), Smoking (0), Age (48), ID (106).
- Row 2: Color of lesion (1), Ulceration (1), Appearance of the lesion (2), Location of lesion (24), Large mole at birth (1), Gender (0).

Below the input fields, there are two large images of skin lesions. To the right of these images are buttons for "Check in KB cancer", "Check in KB normal", "Load picture", and "Number of similar cases" (set to 10). Further right are buttons for "Similarity cancer", "Store picture cancer", "Similarity normal", and "Store picture normal".

Below the input fields and buttons, there is a grid of 10 small images of skin lesions, numbered 1 through 10. To the right of this grid is a table with the following data:

targetid	carcoid	resultc	sc
106	97	66.666641	1
106	31	33.333321	0.78942237041
106	68	50	0.78920856439
106	21	33.333321	0.78203957096
106	54	50	0.75869574156
106	61	50	0.7286220971
106	76	66.666641	0.71855739390
106	74	33.333321	0.71594010875
106	75	50	0.69051522272
106	89	66.666641	0.68901314477

At the bottom of the interface, there are buttons for "Expect recommendation", "Exit", and "Information".

Fonte: Majeed et al. (2020)

## APÊNDICE C - APLICAÇÕES DE MÉTODOS DE OTIMIZAÇÃO

O processo de criação de algo que seja o mais eficiente possível é chamado de otimização (CHAPRA, 2013). Do ponto de vista matemático, a otimização trata da determinação dos mínimos e máximos de uma função através da escolha sistemática, dentro de um conjunto viável e respeitando as restrições estabelecidas, dos valores das variáveis da função em estudo (SANTANA, 2012).

As técnicas de otimização geralmente apresentam uma função objetivo, que pode ser maximizada ou minimizada e usada no cálculo dos valores ótimos; um espaço de busca onde estão as possíveis soluções do problema e um conjunto de restrições do problema.

Os problemas práticos de otimização, por sua complexidade e volume de dados, requerem soluções numéricas computacionais (CHAPRA, 2013). Nesse contexto, um algoritmo de otimização para resolução de problemas em engenharia pode ser utilizado, por exemplo, na estimativa de parâmetros desconhecidos, na seleção de atributos de classificação e no melhoramento de resultados de classificadores. Assim, o presente tópico foi organizado seguindo tais aplicações.

Bhowmik e Repaka (2016) aplicaram duas técnicas de otimização, Algoritmo Genético (AG) e Recozimento Simulado, para, por meio de análise inversa, estimar características de crescimento tumoral, como diâmetro do tumor e profundidade de invasão, e propriedades termofísicas, como perfusão sanguínea e geração de calor metabólico, em um modelo de melanoma de pele humana. O modelo geométrico utilizado foi composto por um cubo de 20x20 mm, com as diferentes camadas representando os diferentes níveis da pele, e um cilindro de bases circulares para representação do melanoma. A simulação numérica foi realizada usando o *software* de elementos finitos disponível comercialmente, COMSOL *Multiphysics* (versão 5.0), e validada com imagens termográficas. Os autores obtiveram um erro menor que 10% para as estimativas realizadas no estudo.

Hosain e Mohammadi (2016) aplicaram um método de otimização na técnica de AG para parametrizar um tumor mamário, minimizando uma dada função objetivo. Um modelo físico-anatômico da mama foi desenvolvido levando em consideração a geometria (deformação postural ou características específicas da mama em estudo, como ptose) e a heterogeneidade do órgão. Resolvendo a equação de Pennes em estado estacionário e com condições de contorno não-lineares, os autores utilizaram o método dos elementos finitos para calcular o perfil de temperatura da superfície da geometria, com um nódulo esférico inserido. Os dados obtidos foram aplicados ao otimizador AG, usado para estimar parâmetros térmicos (taxa metabólica)

e físicos (localização e tamanho do tumor). O resultado estimado foi comparado com simulação numérica e trabalhos anteriores e, em relação a este último, o estudo estima resultados com acurácia de 99%, no qual a profundidade (em *cm*) e o raio (em *mm*) foram obtidos com até cinco casas decimais.

Bezerra et al. (2020) utilizaram o método inverso para estimar propriedades termofísicas da mama e do tumor, utilizando a Programação Quadrática Sequencial (SQP – *Sequential Quadratic Programming*). Foi desenvolvido um modelo geométrico baseado em próteses externas de silicone e, de acordo com os dados de ultrassonografia das pacientes voluntárias, foi inserido um nódulo mamário de formato cilíndrico. Uma análise de sensibilidade de parâmetros, feita previamente à resolução do problema inverso, indicou que a temperatura superficial da mama pode sofrer influência da condutividade térmica e da perfusão sanguínea dos tecidos. O *software* de Dinâmica dos Fluidos Computacional (CFD – *Computational Fluid Dynamics*) Fluent® foi usado para calcular as temperaturas. Os dados de temperaturas máximas obtidos numericamente, por simulação, e coletados experimentalmente através de uma câmera termográfica compuseram a função objetivo do problema inverso. O otimizador SQP foi usado para minimizar tal função objetivo. Os autores analisaram diferentes casos e os resultados demonstraram que é possível estimar as propriedades (condutividade térmica e perfusão sanguínea) de forma satisfatória, embora apenas os dados de temperatura máxima tenham sido utilizados.

Vila Nova (2017) utilizou técnicas de otimização para determinar a melhor combinação de parâmetros de peso  $\lambda_k$ ,  $1 \leq k \leq 4$ , de um classificador de imagens termográficas de mamas baseado na distância parametrizada de Mahalanobis, desenvolvido por Araújo (2014). Inicialmente, para encontrar o valor ideal de cada  $\lambda_k$ , foi desenvolvida uma rotina iterativa no Matlab® para variação dos parâmetros em valores entre 0 e 1 com incrementos de 0,25 e de 0,125. A melhor combinação dos valores de todos  $\lambda_k$ , correspondente a cada classe (Tumor Maligno, Tumor Benigno, Cisto e Normal), foi feita através da técnica de Algoritmos Genéticos para múltiplas funções objetivo. As funções objetivo foram obtidas através de regressão múltipla, com as variáveis de entrada sendo os parâmetros de peso  $\lambda_k$  e as variáveis de saída sendo a taxa de acerto e a sensibilidade do classificador à Classe Maligno. O processo de otimização foi realizado no *software* Matlab®. Os valores ótimos das duas funções objetivo referentes à taxa de acerto e à sensibilidade à Classe Maligno foram, respectivamente, 72,41% e 96,30%. A melhor combinação dos parâmetros de peso  $\lambda_k$  foram:  $\lambda_1=0,4439$ ;  $\lambda_2=0,5127$ ;  $\lambda_3=0,97243$  e  $\lambda_4=0,88666$ .

Silva et al. (2016) em seu estudo utilizaram o módulo Auto-WEKA do *software* WEKA para otimizar parâmetros dos classificadores usados na tarefa de gerar diagnósticos para imagens termográficas de mamas com diferentes patologias. Este problema é conhecido como problema de otimização de hiperparâmetros ou seleção de modelos. A maximização do resultado do classificador, de acordo com alguma medida estabelecida, como a acurácia, é a função objetivo do método. O Auto-WEKA utiliza a otimização bayesiana e explora configurações de parâmetros para vários algoritmos de aprendizagem e recomenda ao usuário o método que provavelmente terá um bom desempenho. Dentre 39 algoritmos de classificação testados pelos autores, *K-Star* e *Bayes Net* apresentaram 100% de acurácia de classificação após o processo de otimização de hiperparâmetros.

Rodrigues et al. (2019) propuseram o uso do AG e da otimização por enxame de partículas (PSO – *Particle Swarm Optimization*) para seleção de características extraídas de imagens termográficas de mamas. A seleção de recursos visa otimizar o processo de aprendizado de máquina, a partir da escolha dos recursos mais relevantes e não redundantes. Os autores utilizaram diversos classificadores para avaliar a performance dos subconjuntos com recursos selecionados. Em relação aos métodos de seleção, foi observado que o AG apresentou uma eficiência ligeiramente superior quando comparado ao PSO. Usando o AG e o classificador SVM, foi possível reduzir o conjunto inicial de 169 recursos, com acurácia de 91,12%, para 57 recursos com acurácia de 86,16%. Dessa forma, a redução mostrou-se apropriada e com perda mínima de precisão, em relação ao conjunto completo de recursos.

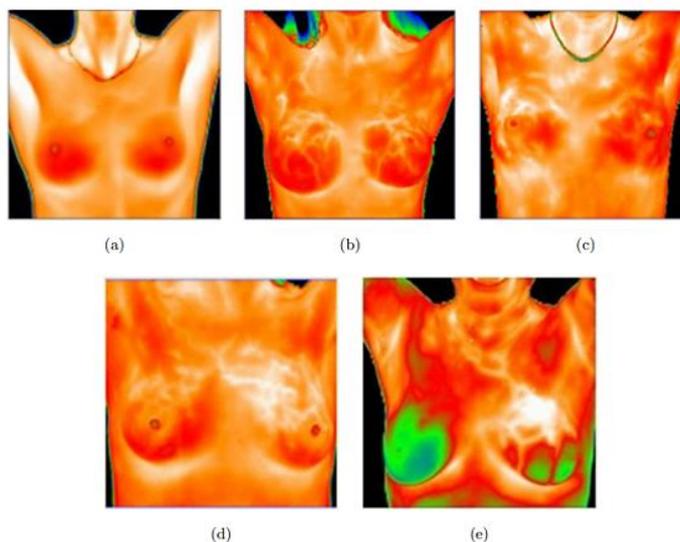
## APÊNDICE D - ÍNDICES DE AVALIAÇÃO TERMOGRÁFICA DE ANOMALIAS MAMÁRIAS

A termografia tem mostrado um grande potencial na triagem do câncer de mama, porém as interpretações das suas imagens dependem dos especialistas. Tal fato pode levar a erros sistêmicos e, a depender de quem analisa o termograma, a resultados desiguais.

Além de desenvolver sistemas CAD para minimizar tais fatos, alguns pesquisadores criaram índices relativos ao grau de risco de câncer de mama. Isso pode auxiliar os médicos especialistas a discriminar de forma mais padronizada os pacientes saudáveis dos patológicos. Os artigos listados neste tópico foram selecionados levando em consideração, principalmente, a criação e/ou utilização de índices para auxiliar na discriminação de classes durante a triagem, por exemplo.

O sistema de classificação de Marseille fornece critérios para avaliação de risco em imagens termográficas de mamas. Este método analítico fornece uma escala TH 1-TH 5 como um resumo baseado em características térmicas específicas e níveis diferenciais de energia infravermelha (Figura D1). Dessa forma, para os resultados da termografia mamária, o sistema Marseille categoriza as mamas da paciente como: TH 1, ausência de vascularização; TH 2, vascularização normal; TH 3, tecido observado apresenta pontos de calor com temperatura mais elevada; TH 4, tecido com anormalidades presentes e TH 5, tecido com função tecidual gravemente alterada. As figuras abaixo apresentam imagens de pacientes enquadradas em diferentes níveis da escala Marseille (ETEHADTAVAKOL & NG, 2013; OMRANIPOUR et al., 2016; RESMINI, 2016).

Figura D1 – Amostra com diferentes classificações da escala Marseille. (a) TH-1, (b) TH-2, (c) TH-3, (d) TH-4, e (e) TH-5



Fonte: Etehadtavakol e Ng (2013)

O Centro Multidisciplinar de Câncer de Mama Ville Marie estabeleceu uma escala de classificação das imagens termográficas baseada em informações clínicas pertinentes, comparando imagens de ambas as mamas com imagens anteriores. A escala infravermelha (IR) é dividida em cinco categorias: IR 1 (de ausência de qualquer padrão vascular à simetria vascular leve); IR 2 (padrão vascular significativo, mas simétrico ou assimetria leve); IR 3 (um sinal anormal); IR 4 (dois sinais anormais) e IR 5 (acima de três sinais anormais) (DIAKIDES & BRONZINO, 2007; RESMINI, 2016).

Os sinais anormais foram caracterizados através da comparação entre as mamas esquerda e direita, na qual foram avaliados: a presença de uma assimetria vascular significativa; de vasos tortuosos incomuns que formam laços, arborização anormal ou padrões irregulares; de um aumento focal de temperatura de  $1^{\circ}\text{C}$  ( $\Delta T$ ) quando comparado ao local contralateral e quando associado à área de anormalidade clínica; de um  $\Delta T$  focal de  $2^{\circ}\text{C}$  *versus* local contralateral; de um  $\Delta T$  focal de  $3^{\circ}\text{C}$  *versus* o restante da mama homolateral quando não presente suspeitas no local contralateral e de um  $\Delta T$  global da mama de  $1,5^{\circ}\text{C}$  *versus* a região contralateral. Então, para uma imagem ser considerada anormal exigia a presença de pelo menos um dos sinais listados (DIAKIDES & BRONZINO, 2007; RESMINI, 2016).

Ohashi e Uchida (2000) desenvolveram um índice (Quadro D1) baseado no resultado positivo ou negativo gerado pela análise de termogramas obtidos através da termografia por estado estacionário, sequencial e por subtração entre dois termogramas. A termografia em estado estacionário foi realizada a uma temperatura de  $21^{\circ}\text{C}$  com as mãos do paciente em sua cabeça. Em seguida, as mamas dos pacientes foram expostas à estresses térmicos para remover

o padrão vascular e, depois de encerrado o estresse, o padrão foi recuperado gradualmente. A recuperação térmica foi registrada a cada 15 segundos na visão anteroposterior. A termografia por subtração foi feita subtraindo o segundo termograma de um obtido sequencialmente primeiro.

Os critérios de diagnóstico para câncer de mama usados pelos autores foram: 1- ponto quente assimétrico; 2- assimetria anormal e exagerada; 3- diferenças significativas de temperatura (alto índice térmico); 4- padrões de calor positivo na termográfica sequencial e 5- existência de um ponto quente na termografia por subtração. A presença dos itens 1, 2 ou 3 foi considerada positiva em relação à presença de tumor maligno. Então, a partir dos resultados (positivo ou negativo) dos três métodos de aquisição de termogramas, os autores propuseram um índice que varia de 1 a 8 conforme o quadro a seguir.

Quadro D1 – Definição dos índices/grupos de acordo com Ohashi e Uchida (2000)

<b>Termografia por estado estacionário</b>	<b>Termografia sequencial</b>	<b>Termografia por subtração</b>	<b>Índice/Grupo</b>
Positivo	Positivo	Positivo	I
Positivo	Positivo	Negativo	II
Positivo	Negativo	Positivo	III
Positivo	Negativo	Negativo	IV
Negativo	Positivo	Positivo	V
Negativo	Positivo	Negativo	VI
Negativo	Negativo	Positivo	VII
Negativo	Negativo	Negativo	VIII

Fonte: Adaptada de Ohashi e Uchida (2000)

Frize, Herry e Roberge (2002) utilizaram a metodologia adotada por Head et al. (1997) para padronizar os resultados das análises das imagens termográficas adquiridas por eles. Após a anamnese, a paciente foi encaminhada para uma sala escura e com temperatura ambiente de 20 °C. As mamas foram levemente resfriadas por 20 minutos e, durante o restabelecimento da sua temperatura, foi feita a aquisição dos termogramas.

A definição do índice de padronização foi baseada em cálculo de pontuações: se a temperatura média de um quadrante foi de 0,5 a 1 °C maior que o mesmo quadrante da mama oposta, então foi atribuída uma pontuação de 0,5. Porém, se a temperatura média foi superior a 1 °C, a pontuação atribuída foi de 1. A partir do somatório das pontuações, foi criado o índice que tem valor entre 0 e 4 (FRIZE; HERRY; ROBERGE, 2002). Os autores não informaram o que cada valor representa, apenas que os índices maiores que 1 indicavam pacientes com anormalidades mamárias.

Raghavendra et al. (2016) propuseram o Índice de Risco de Câncer de Mama (BCRI – *Breast Cancer Risk Index*) baseado na combinação de recursos resultantes da projeção linear, através da função Kernel, das características extraídas dos termogramas. Essa projeção é aplicada para reduzir a dimensionalidade dos histogramas de gradientes orientados, selecionados por armazenar as mínimas variações internas das imagens termográficas. O BCRI (Equação (D.1)) foi formulado de tal maneira que o seu intervalo é distintamente diferente para classes normal e maligna.

$$BCRI = 3,5 - 1,489 \times KLPP6 - 1,252 \times KLPP9 \quad (D.1)$$

Dentre vinte e quatro coeficientes KLPP (*Kernel Locality Preserving Projection*) extraídos, os índices KLPP6 e KLPP9 foram escolhidos por apresentarem maior discriminação entre as classes citadas. Os valores constantes foram gerados usando regressão linear multivariada, através da qual os valores selecionados proporcionaram a máxima separação entre duas classes.

## APÊNDICE E – DESCRIÇÃO DOS ATRIBUTOS EXTRAÍDOS DOS TERMOGRAMAS DE MAMA

### Medidas estatísticas

As características que compõem este grupo são baseadas em medidas estatísticas simples, e que foram calculadas diretamente sobre os valores de temperaturas extraídas na etapa de segmentação das regiões das mamas direita e esquerda.

Para as equações apresentadas a seguir, foram adotadas as seguintes notações:

- $T_i$  – representa o valor de temperatura de cada elemento da matriz;
- $N$  – representa o total de elementos presentes na matriz de temperatura;
- $E$  – representa o valor médio esperado,

### *Média*

A média (Equação (E1)), medida de tendência central mais utilizada, é um valor que corresponde à região onde os dados mais se concentram (VIEIRA, 2011). Nas matrizes de temperatura, a média pode ser considerada uma representação simplificada da região de interesse, visto que os dados estão distribuídos de forma uniforme.

$$\mu = \frac{1}{N} \sum_{i=1}^N T_i \quad (E1)$$

### *Desvio padrão*

O desvio padrão (Equação (E2)) é uma medida de variabilidade que mede adequadamente a dispersão dos dados em torno da média  $\mu$  (VIEIRA, 2011). Um valor de desvio padrão mais alto indica maior dispersão nos dados, indicando uma maior desigualdade entre os dados da matriz de temperatura analisada. Já um valor mais baixo de desvio padrão indica que estes dados estão próximos da média e, portanto, a matriz possui valores de temperatura semelhantes.

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (T_i - \mu)^2} \quad (E2)$$

### *Mediana*

A mediana é definida como sendo o valor que ocupa a posição central do conjunto de dados ordenados (VIEIRA, 2011). Em aplicações da termografia, pacientes saudáveis possuem simetria térmica entre as duas mamas. Dessa forma, quanto maior a diferença entre as medianas

das mamas, maior a probabilidade de anomalias; o inverso também é verdadeiro (LESSA & MARENGONI, 2016).

Quando o número de dados da matriz resultante da segmentação é ímpar, após ordenação crescente destes dados, o valor da mediana é o valor na posição central. Para um número de dados par, existem dois valores na posição central. Neste caso, a mediana é a média desses dois valores.

### *Obliquidade*

A obliquidade (Equação (E3)) é uma medida de assimetria de uma distribuição de probabilidade em relação à média (GONZALEZ; WOODS; EDDINS, 2009). Na análise da matriz de temperatura, é uma medida que indica o distanciamento da temperatura em torno da média (FARIAS, 2017 apud GONÇALVES, 2017). No caso de o valor de obliquidade ser próximo de zero, significa que há presença de simetria na distribuição. Caso contrário, uma tendência à esquerda gera uma obliquidade menor que zero e uma tendência à direita uma obliquidade maior que zero.

$$S = \frac{E(T_i - \mu)^3}{\sigma^3} \quad (E3)$$

### *Curtose*

A curtose (Equação (E4)) caracteriza o grau de achatamento da curva de distribuição de dados. É uma medida de quão sujeita a *outliers* é uma distribuição (GONZALEZ; WOODS; EDDINS, 2009). No caso da normalidade o valor da curtose é três. As distribuições mais propensas a *outliers*, em relação à distribuição normal, tem curtose maior que três e àquelas menos propensas a *outliers* tem valor menor que três. Algumas definições de curtose subtraem três do valor calculado, de modo que a distribuição normal possua um valor zero de curtose (GONZALEZ; WOODS; EDDINS, 2009).

$$k = \frac{E(T_i - \mu)^4}{\sigma^4} \quad (E4)$$

### *Entropia*

A entropia (Equação (E5)) é uma medida estatística de aleatoriedade que pode ser usada para caracterizar a textura da imagem de entrada (GONZALEZ; WOODS; EDDINS, 2009). O valor da entropia fornece uma quantificação da randomicidade de uma imagem. Assim, um

valor elevado de entropia indica que a imagem analisada apresenta mais irregularidades ou padrões atípicos (GONZALEZ; WOODS; EDDINS, 2009).

$$H = -\sum_{i=1}^n p_i \times \log_2(p_i) \quad (\text{E5})$$

onde  $n$  é o número total de pixels da imagem em tons de cinza e  $p_i$  é a probabilidade do  $i$ -ésimo nível de cinza acontecer novamente (GONZALEZ; WOODS; EDDINS, 2009).

Na análise de termogramas, quanto mais simétrica for a distribuição de temperatura, menor será a entropia. Logo, termogramas que apresentam valores de entropia significativamente diferentes, quando se comparam as mamas direita e esquerda, possuem maior assimetria e maior probabilidade de anormalidade (LESSA & MARENGONI, 2016).

### Medidas estatísticas de segunda ordem

As medidas estatísticas de segunda ordem utilizadas neste trabalho são baseadas na Matriz de Coocorrência em Escala de Cinza (do inglês *Grey Level Cooccurrence Matrix*, GLCM).

A GLCM é uma matriz quadrada, com dimensões de  $256 \times 256$ , cujos valores correspondem às transições de níveis de cinza entre *pixels* que ocorrem em uma imagem. A obtenção de cada valor da matriz é fundamentada na análise da direção e da distância entre dois *pixels* adjacentes (RESMINI, 2016).

Matematicamente, uma matriz de coocorrência  $C$  (Equação (E6)) é calculada a partir de uma imagem  $I$  de dimensões  $N \times M$ , parametrizada por um deslocamento  $(\Delta x, \Delta y)$ .

$$C_{\Delta x, \Delta y}(j, k) = \sum_{p=1}^N \sum_{q=1}^M \begin{cases} 1, & \text{se } I(p, q) = j \text{ e } I(p + \Delta x, q + \Delta y) = k \\ 0, & \text{caso contrário} \end{cases} \quad (\text{E6})$$

onde  $j$  e  $k$  representam os valores de intensidade de cinza da imagem,  $p$  e  $q$  são as coordenadas na imagem  $I$  e o deslocamento  $(\Delta x, \Delta y)$  depende da direção  $\theta$  (horizontal, vertical ou diagonal) e da distância  $d$  em que a matriz é computada (RESMINI, 2016).

De forma geral, cada posição da matriz de coocorrência armazena a frequência com que o valor os *pixels*  $j$  e  $k$  ocorrem na imagem, distantes  $\Delta x$  colunas e  $\Delta y$  linhas (CONCI; AZEVEDO; LETA, 2008; RESMINI, 2016).

### *Contraste*

Trata-se de uma estimativa das variações locais ao quadrado das intensidades de cinza entre pares de *pixels* da matriz de coocorrência (ACHARYA et al., 2012; GONZALEZ;

WOODS; EDDINS, 2009). A comparação entre pixel vizinhos é realizada em toda a imagem. Os valores na diagonal da GLCM representam pouco contraste, aumentando proporcionalmente à medida que se distancia dela (GONZALEZ; WOODS; EDDINS, 2009). O contraste é calculado através da Equação (E7):

$$con = \sum_{j,k=1}^n |j - k|^2 P(k, j) \quad (E7)$$

onde  $j$  e  $k$  são posições dos *pixels* na matriz e  $P(j, k)$  é o valor do pixel na posição  $(j, k)$  da matriz GLCM.

### Correlação

Quanto à análise de textura, a correlação (Equação (E8)) é uma medida que indica o grau de dependência linear de *pixels* em relação a sua vizinhança (ACHARYA et al., 2012). Os valores variam numa faixa de -1 a 1, na qual os valores extremos correspondem a uma correlação perfeita positiva ou negativa, respectivamente (GONZALEZ; WOODS; EDDINS, 2009).

$$corr = \sum_{j,k=1}^n \frac{(j-\mu_j)(k-\mu_k)P(j,k)}{\sigma_j\sigma_k} \quad (E8)$$

### Homogeneidade

A homogeneidade (Equação (E9)) pode ser interpretada como uma medida inversa ao contraste. Isso significa que esta medida retorna valores que representam a proximidade da distribuição dos elementos em relação à diagonal da GLCM (GONZALEZ; WOODS; EDDINS, 2009). Assim, altos valores de homogeneidade indicam que imagem contém baixa variabilidade entre os níveis de cinza.

$$homog = \sum_{j,k=1}^n \frac{P(j,k)}{1+|j-k|} \quad (E9)$$

### Energia

Energia ou uniformidade (Equação (E10)) é uma medida também derivada da GLCM e retorna a soma de todos os elementos elevados à segunda potência. Trata-se de uma estimativa que pode assumir valores de intervalo [0,1], no qual o valor 1 indica uma imagem constante, com menos tom de cinza em toda a imagem (GONZALEZ; WOODS; EDDINS, 2009).

$$energ = \sum_{j,k=1}^n (P(i, j))^2 \quad (E10)$$

### Momento estatístico

O momento estatístico  $m_k$  é considerado um descritor de ordem mais alta de uma dada distribuição. Um momento de ordem  $k$  é definido como o valor médio esperado  $E(x^k)$  de uma variável aleatória discreta  $x^k$ . Matematicamente, tem-se que para cada número inteiro  $k$ , o  $k$ -ésimo momento de uma variável  $x$ , cuja função densidade de probabilidade é dada por  $f(x)$ , é definido como

$$m_k = E(x^k) \quad (\text{E11})$$

$$m_k = E(x^k) = \int_{-\infty}^{\infty} f(x)x^k dx \quad (\text{E12})$$

Há casos cujas distribuições possuem todos os seus momentos com valor infinito e, além disso, pode ocorrer de, mesmo conhecendo os momentos, não ser possível caracterizar unicamente a função  $f(x)$  (MORETO, 2008). Na prática, a maioria das distribuições possuem momentos com valores finitos e, conhecer estes valores é equivalente a conhecer a função  $f(x)$  (HYVÄRINEN et al., 2001 apud MORETO, 2008).

### *Momento estatístico central*

Os momentos mais importantes são os de primeira ao de quarta ordem (MORETO, 2008). O momento de ordem  $k = 1$  é também denominado de média ( $\mu$ ). A partir do momento estatístico de ordem  $k = 2$  é possível calcular momentos centrais. Os momentos estatísticos centrais são baseados nos desvios em relação à média  $\mu$  da amostra (Equação (E13)).

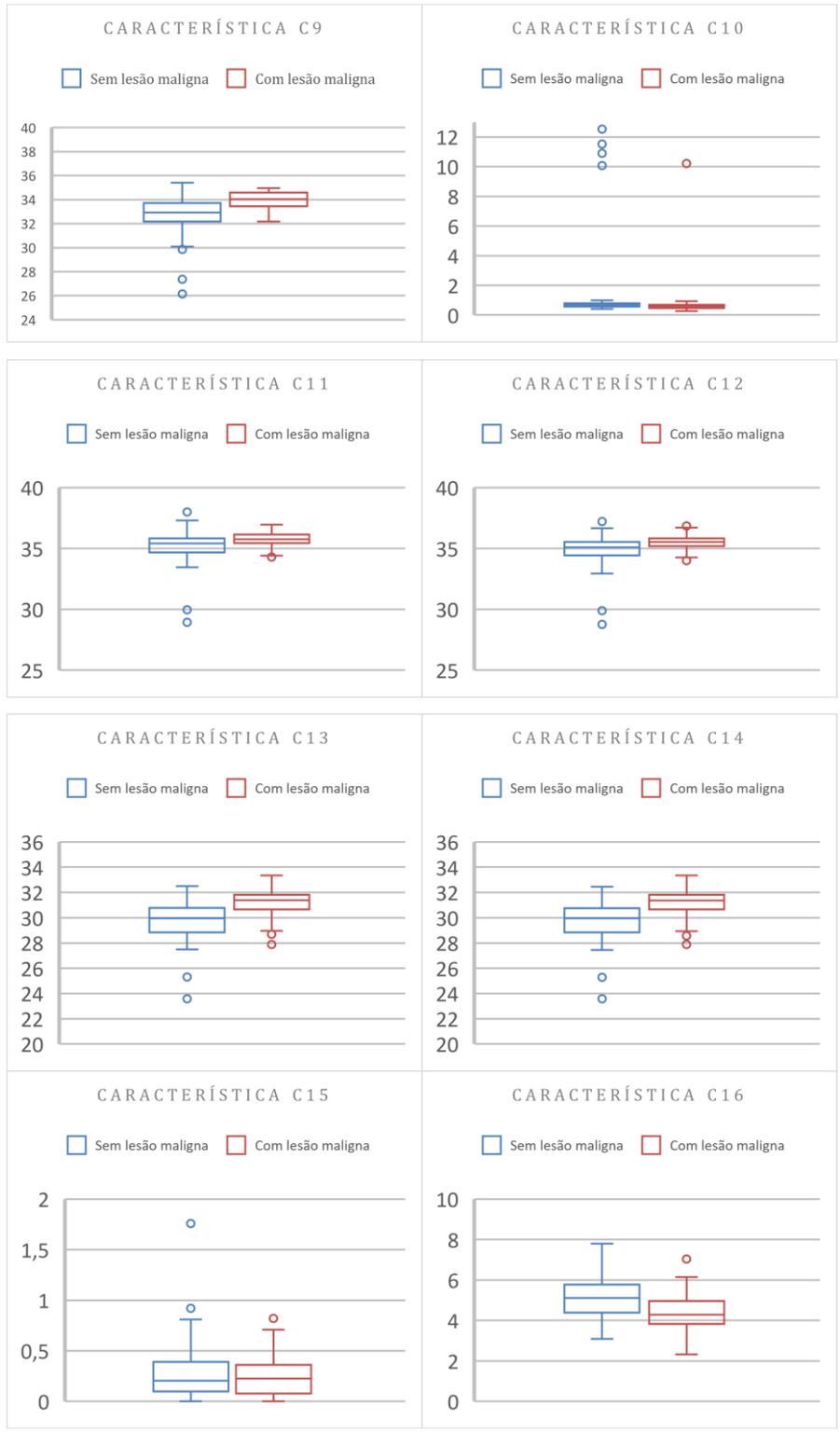
$$m_k = E[(x - \mu)^k] \quad (\text{E13})$$

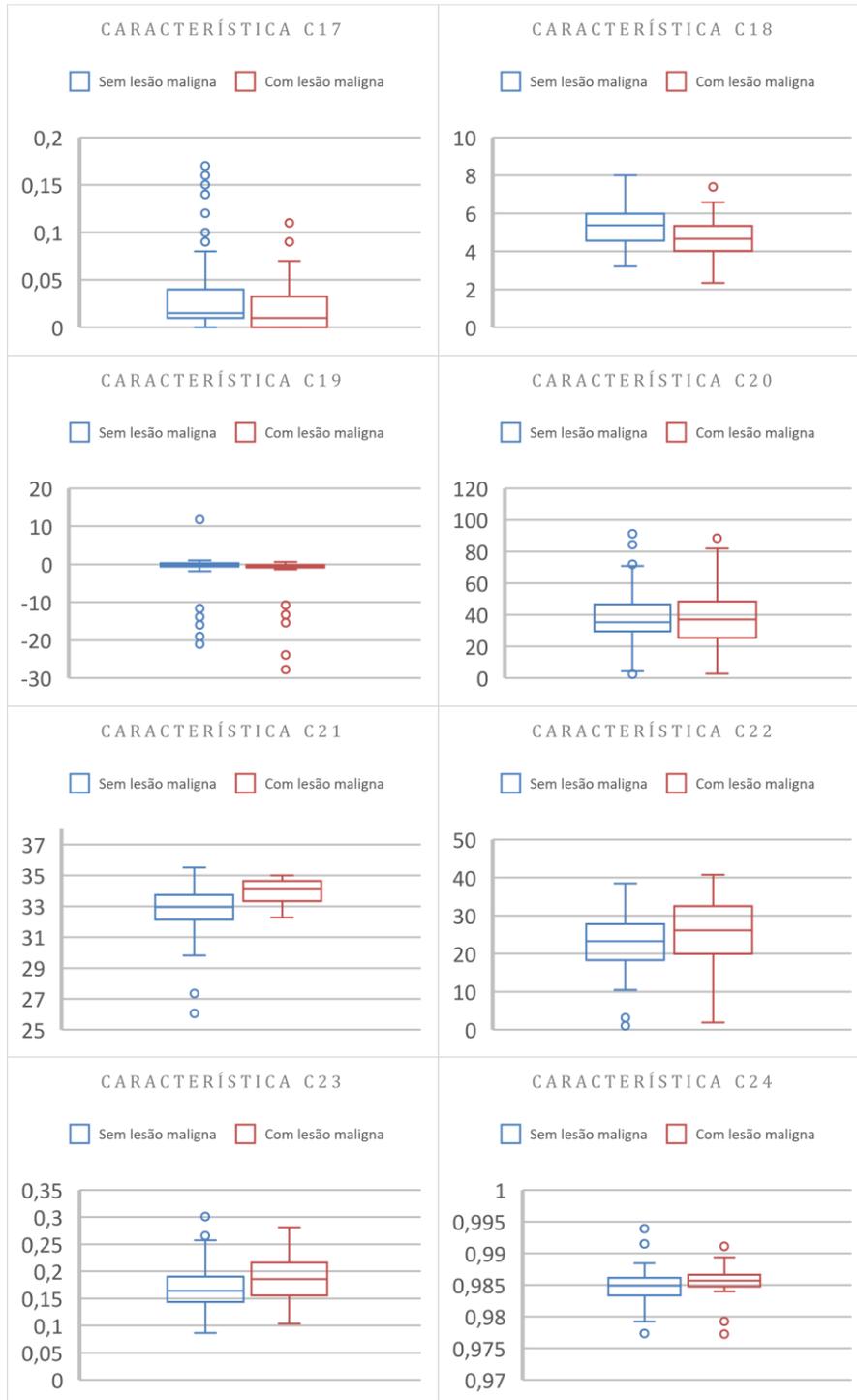
Como afirmado anteriormente, é possível utilizar os momentos para caracterizar a distribuição definida pela função  $f(x)$ . Para isso, é necessário estimar os valores esperados da amostra, uma vez que os seus dados são as únicas informações disponíveis (MORETO, 2008). Então, seja  $N$  o conjunto de dados da amostra, a estimativa do momento estatístico central de  $x$  (ou momento amostral) é calculada através da Equação ((E14)).

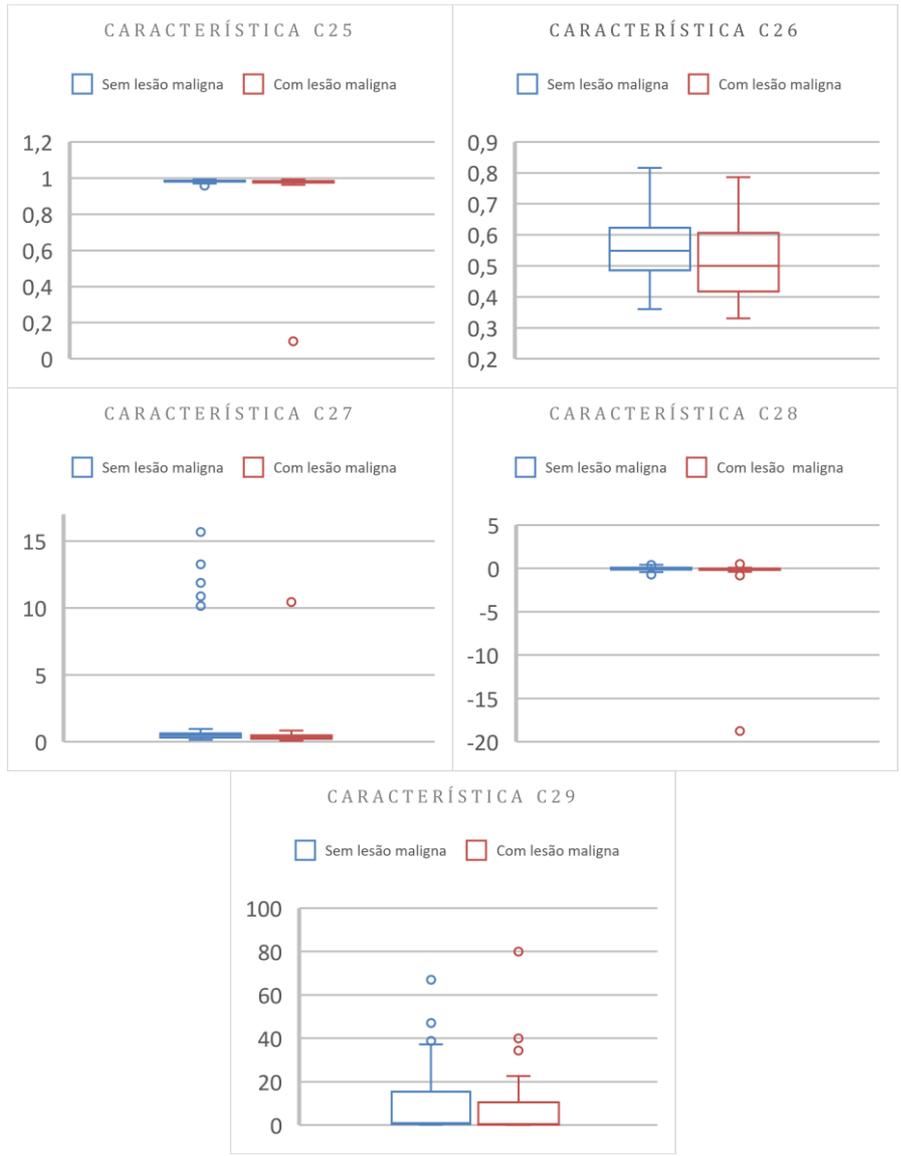
$$m_k = \frac{1}{N} \sum_{j=1}^N (x_j - \mu)^k \quad (\text{E14})$$

## APÊNDICE F – GRÁFICOS *BOXPLOT* DOS ATRIBUTOS EXTRAÍDOS DA BASE DE DADOS BINÁRIA









**APÊNDICE G – RESULTADOS DA CLASSIFICAÇÃO BINÁRIA PARA BASE DE DADOS DESBALANCEADA**

<b>Todas as características</b>					
<b>Classificadores</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>	<b>Índice Kappa</b>	<b>AUC</b>
<b>Análise discriminante</b>	74,14	71,43	74,74	0,35	0,76 [0,67 0,85]
<b>Árvore de decisão</b>	84,48	50,00	92,11	0,45	0,78 [0,44 0,87]
<b>kNN</b>	84,05	33,33	95,26	0,35	0,87 [0,80 0,94]
<b>Naive Bayes</b>	69,83	73,81	68,95	0,30	0,74 [0,66 0,84]
<b>SVM</b>	76,29	45,24	83,76	0,26	0,63 [0,53 0,73]
<b>Características selecionados pelo AG</b>					
<b>Classificadores</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>	<b>Índice Kappa</b>	<b>AUC</b>
<b>Análise discriminante</b>	87,50	61,90	93,16	0,57	0,86 [0,79 0,94]
<b>Árvore de decisão</b>	87,93	54,76	95,26	0,55	0,88 [0,82 0,95]
<b>kNN</b>	88,79	47,62	97,84	0,55	0,96 [0,92 1,00]
<b>Naive Bayes</b>	87,07	50,00	95,26	0,51	0,85 [0,77 0,92]
<b>SVM</b>	85,34	59,52	91,05	0,51	0,82 [0,74 0,90]
<b>Características selecionados pela ED</b>					
<b>Classificadores</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>	<b>Índice Kappa</b>	<b>AUC</b>
<b>Análise discriminante</b>	87,50	59,52	93,68	0,56	0,86 [0,78 0,93]
<b>Árvore de decisão</b>	86,21	50,00	94,21	0,49	0,82 [0,74 0,90]
<b>kNN</b>	86,21	38,09	96,84	0,43	0,89 [0,82 0,96]
<b>Naive Bayes</b>	83,19	50,00	90,53	0,42	0,75 [0,66 0,84]
<b>SVM</b>	83,62	61,90	88,42	0,48	0,89 [0,82 0,96]

Fonte: A Autora (2022)

**APÊNDICE H - RESULTADO DA CLASSIFICAÇÃO BINÁRIA PARA OS DIFERENTES MÉTODOS DE BALANCEAMENTO DA AMOSTRA, UTILIZANDO TODAS AS CARACTERÍSTICAS**

		<b>SM</b>	<b>SM-B</b>	<b>C-SM</b>	<b>SL-SM</b>	<b>AW</b>	<b>AD</b>	
<b>Classificador</b>	<b>Análise discriminante</b>	ACC (%)	75,75	85,53	78,16	82,63	85,41	70,05
		SENS. (%)	84,29	93,16	83,68	91,58	72,49	76,63
		ESP (%)	66,32	77,89	72,63	73,68	98,40	63,68
		Índice Kappa	0,51	0,71	0,56	0,65	0,71	0,40
		AUC	0,82 [0,77 0,86]	0,82 [0,89 0,95]	0,84 [0,79 0,88]	0,89 [0,86 0,93]	0,97 [0,95 0,99]	0,74 [0,69 0,79]
<b>Classificador</b>	<b>Árvore de decisão</b>	ACC (%)	83,50	88,95	86,58	89,74	90,45	84,76
		SENS. (%)	86,67	89,47	86,84	89,47	87,30	89,67
		ESP (%)	80,00	88,42	86,32	90,00	93,62	80,00
		Índice Kappa	0,67	0,78	0,73	0,79	0,81	0,70
		AUC	0,89 [0,86 0,92]	0,94 [0,92 0,97]	0,92 [0,90 0,95]	0,95 [0,93 0,97]	0,96 [0,94 0,98]	0,91 [0,88 0,94]
<b>Classificador</b>	<b>kNN</b>	ACC (%)	87,75	88,85	84,74	88,16	89,92	86,10
		SENS. (%)	100,00	87,89	89,47	94,21	83,07	99,46
		ESP (%)	74,21	90,00	80,00	82,11	96,81	73,16
		Índice Kappa	0,75	0,78	0,69	0,76	0,79	0,72
		AUC	0,94 [0,92 0,97]	0,95 [0,92 0,97]	0,91 [0,88 0,94]	0,94 [0,91 0,96]	0,97 [0,96 0,99]	0,94 [0,91 0,96]
<b>Classificador</b>	<b>Naive Bayes</b>	ACC (%)	75,50	85,79	77,37	81,58	84,88	71,12
		SENS. (%)	85,71	91,58	85,26	91,58	74,07	82,61
		ESP (%)	64,21	80,00	69,47	71,58	95,74	60,00
		Índice Kappa	0,50	0,72	0,55	0,63	0,70	0,42
		AUC	0,82 [0,78 0,86]	0,92 [0,89 0,95]	0,84 [0,80 0,88]	0,89 [0,85 0,92]	0,94 [0,91 0,96]	0,78 [0,73 0,82]
<b>Classificador</b>	<b>SVM</b>	ACC (%)	90,00	87,89	97,89	88,16	93,10	89,30
		SENS. (%)	97,14	90,53	94,77	95,26	98,83	97,83
		ESP (%)	82,11	85,26	81,05	81,05	98,40	81,05
		Índice Kappa	0,80	0,76	0,76	0,76	0,86	0,79
		AUC	0,95 [0,93 0,97]	0,93 [0,91 0,96]	0,94 [0,91 0,96]	0,94 [0,91 0,96]	0,99 [0,98 1,00]	0,95 [0,92 0,97]

Fonte: A Autora (2022)

**APÊNDICE I - RESULTADOS DA CLASSIFICAÇÃO BINÁRIA PARA AMOSTRA  
BALANCEADA POR SMOTE, COM CONJUNTO SELECIONADO DE ATRIBUTOS**

<b>SMOTE</b>				
<b>Análise discriminante</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	85,25	80,50	70,26	80,26
<b>SENS (%)</b>	87,62	88,57	75,26	87,37
<b>ESP (%)</b>	82,63	71,58	65,26	73,16
<b>Índice Kappa</b>	0,70	0,61	0,40	0,61
<b>AUC</b>	0,91 [0,88 0,94]	0,87 [0,84 0,90]	0,74 [0,69 0,79]	0,87 [0,83 0,90]
<b>Árvore de decisão</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	82,75	89,75	68,68	76,32
<b>SENS (%)</b>	89,05	89,52	68,95	77,37
<b>ESP (%)</b>	75,79	90,00	68,42	75,26
<b>Índice Kappa</b>	0,65	0,79	0,37	0,53
<b>AUC</b>	0,89 [0,86 0,92]	0,95 [0,93 0,97]	0,70 [0,65 0,76]	0,81 [0,76 0,85]
<b>kNN</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	90,00	89,25	72,11	79,47
<b>SENS (%)</b>	100,00	100,00	77,37	83,68
<b>ESP (%)</b>	78,95	77,37	66,84	75,26
<b>Índice Kappa</b>	0,80	0,78	0,44	0,59
<b>AUC</b>	0,95 [0,93 0,97]	0,95 [0,93 0,97]	0,76 [0,72 0,81]	0,85 [0,81 0,89]
<b>Naive Bayes</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	82,50	91,25	64,47	80,79
<b>SENS (%)</b>	90,95	88,09	71,58	87,37
<b>ESP (%)</b>	73,16	73,69	57,37	74,21
<b>Índice Kappa</b>	0,65	0,62	0,29	0,62
<b>AUC</b>	0,89 [0,85 0,92]	0,87 [0,84 0,91]	0,67 [0,62 0,73]	0,87 [0,83 0,91]
<b>SVM</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	93,00	92,00	71,58	80,26
<b>SENS (%)</b>	98,09	98,57	89,47	88,95
<b>ESP (%)</b>	97,36	84,74	53,68	71,58
<b>Índice Kappa</b>	0,86	0,84	0,43	0,61
<b>AUC</b>	0,97 [0,95 0,98]	0,96 [0,94 0,98]	0,81 [0,76 0,85]	0,87 [0,83 0,91]

Fonte: A Autora (2022)

**APÊNDICE J - RESULTADOS DA CLASSIFICAÇÃO BINÁRIA PARA AMOSTRA  
BALANCEADA POR *BORDERLINE*-SMOTE, COM CONJUNTO SELECIONADO  
DE ATRIBUTOS**

<i>Borderline</i> -SMOTE				
Análise discriminante				
	AG	ED	mRMR	WEKA
ACC (%)	89,21	87,89	78,16	83,95
SENS (%)	93,16	93,16	85,26	92,63
ESP (%)	85,26	82,63	71,05	75,26
Índice Kappa	0,78	0,76	0,56	0,68
AUC	0,94 [0,92 0,97]	0,94 [0,91 0,96]	0,84 [0,81 0,88]	0,91 [0,88 0,94]
Árvore de decisão				
	AG	ED	mRMR	WEKA
ACC (%)	88,42	91,84	83,16	87,63
SENS (%)	88,42	90,53	81,05	87,37
ESP (%)	88,42	93,16	85,26	87,89
Índice Kappa	0,77	0,84	0,66	0,75
AUC	0,94 [0,92 0,97]	0,97 [0,95 0,99]	0,88 [0,85 0,92]	0,93 [0,91 0,96]
kNN				
	AG	ED	mRMR	WEKA
ACC (%)	90,79	91,32	72,11	89,21
SENS (%)	88,42	91,58	70,53	90,00
ESP (%)	93,42	91,05	73,68	88,42
Índice Kappa	0,82	0,83	0,44	0,78
AUC	0,96 [0,94 0,98]	0,96 [0,94 0,98]	0,74 [0,70 0,79]	0,95 [0,92 0,97]
Naive Bayes				
	AG	ED	mRMR	WEKA
ACC (%)	88,68	89,47	60,79	86,05
SENS (%)	92,10	92,63	43,68	92,63
ESP (%)	85,26	86,32	77,89	79,47
Índice Kappa	0,77	0,79	0,22	0,72
AUC	0,94 [0,92 0,97]	0,95 [0,92 0,97]	0,57 [0,51 0,63]	0,92 [0,89 0,95]
SVM				
	AG	ED	mRMR	WEKA
ACC (%)	89,74	88,95	68,95	88,68
SENS (%)	90,53	91,58	74,74	88,42
ESP (%)	88,95	86,32	63,16	88,95
Índice Kappa	0,79	0,78	0,38	0,77
AUC	0,95 [0,93 0,97]	0,94 [0,92 0,97]	0,73 [0,68 0,78]	0,94 [0,92 0,97]

Fonte: A Autora (2022)

**APÊNDICE K – RESULTADOS DA CLASSIFICAÇÃO BINÁRIA PARA AMOSTRA  
BALANCEADA POR ASUWO, COM CONJUNTO SELECIONADO DE ATRIBUTOS**

<b>ASUWO</b>				
<b>Análise discriminante</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	87,53	88,33	75,33	84,08
<b>SENS (%)</b>	77,78	78,31	63,49	69,31
<b>ESP (%)</b>	97,34	98,40	87,23	98,94
<b>Índice Kappa</b>	0,75	0,77	0,51	0,70
<b>AUC</b>	0,97 [0,95 0,98]	0,98 [0,96 0,99]	0,80 [0,75 0,84]	0,97 [0,96 0,99]
<b>Árvore de decisão</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	90,72	91,51	85,14	91,25
<b>SENS (%)</b>	88,89	88,34	84,13	88,36
<b>ESP (%)</b>	92,55	94,68	86,17	94,15
<b>Índice Kappa</b>	0,81	0,83	0,70	0,83
<b>AUC</b>	0,96 [0,94 0,98]	0,97 [0,95 0,99]	0,91 [0,88 0,94]	0,96 [0,95 0,99]
<b>kNN</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	91,78	91,25	79,84	92,04
<b>SENS (%)</b>	84,13	82,54	78,31	88,89
<b>ESP (%)</b>	99,47	100,00	81,38	95,21
<b>Índice Kappa</b>	0,84	0,82	0,60	0,84
<b>AUC</b>	0,99 [0,98 1,00]	-	0,85 [0,81 0,89]	0,98 [0,96 0,99]
<b>Naive Bayes</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	89,12	86,47	72,68	84,88
<b>SENS (%)</b>	79,89	76,19	54,50	71,43
<b>ESP (%)</b>	98,40	96,81	90,96	98,40
<b>Índice Kappa</b>	0,78	0,73	0,45	0,70
<b>AUC</b>	0,98 [0,97 0,99]	0,96 [0,93 0,98]	0,79 [0,74 0,84]	0,97 [0,95 0,99]
<b>SVM</b>				
	<b>AG</b>	<b>ED</b>	<b>mRMR</b>	<b>WEKA</b>
<b>ACC (%)</b>	95,23	95,23	62,60	86,47
<b>SENS (%)</b>	93,65	92,06	46,56	78,84
<b>ESP (%)</b>	96,81	98,40	78,72	94,15
<b>Índice Kappa</b>	0,90	0,90	0,25	0,73
<b>AUC</b>	0,99 [0,98 1,00]	0,99 [0,98 1,00]	0,58 [0,52 0,64]	0,94 [0,91 0,96]

Fonte: A Autora (2022)

**APÊNDICE L – PERFORMANCE DOS CLASSIFICADORES EM BASE MULTI-  
CLASSES DESBALANCEADA**

<b>Todas as características</b>					
<b>Classificadores</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>	<b>Índice Kappa</b>	<b>AUC</b>
Análise discriminante	43,53	42,09	79,74	0,34	0,84
Árvore de decisão	41,38	39,70	79,56	0,36	0,65
kNN	42,24	40,90	79,59	0,35	0,71
Naive Bayes	36,21	29,58	76,31	0,41	0,68
SVM	44,40	42,52	80,53	0,33	0,77
<b>Características selecionadas pelo AG</b>					
<b>Classificadores</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>	<b>Índice Kappa</b>	<b>AUC</b>
Análise discriminante	48,28	48,59	82,10	0,27	0,79
Árvore de decisão	49,14	49,17	82,36	0,26	0,73
kNN	46,16	42,52	80,69	0,30	0,77
Naive Bayes	47,41	48,72	82,01	0,29	0,85
SVM	33,19	32,10	76,68	0,44	0,77
<b>Características selecionadas pela ED</b>					
<b>Classificadores</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>	<b>Índice Kappa</b>	<b>AUC</b>
Análise discriminante	49,57	49,65	82,55	0,26	0,81
Árvore de decisão	47,41	46,00	81,57	0,29	0,76
kNN	46,12	43,81	80,64	0,30	0,81
Naive Bayes	41,38	43,73	80,19	0,36	0,82
SVM	33,19	32,10	76,68	0,44	0,77
<b>Par de características selecionado pelo WEKA</b>					
<b>Classificadores</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>	<b>Índice Kappa</b>	<b>AUC</b>
Análise discriminante	41,81	38,33	79,09	0,36	0,77
Árvore de decisão	35,34	33,43	77,44	0,42	0,65
kNN	38,36	35,36	77,98	0,39	0,72
Naive Bayes	39,32	35,87	78,21	0,38	0,77
SVM	43,53	36,77	79,21	0,34	0,77
<b>Par de características selecionado pelo mRMR</b>					
<b>Classificadores</b>	<b>ACC (%)</b>	<b>SENS (%)</b>	<b>ESP (%)</b>	<b>Índice Kappa</b>	<b>AUC</b>
Análise discriminante	39,66	35,39	78,08	0,38	0,74
Árvore de decisão	33,19	31,72	76,67	0,44	0,72
kNN	37,50	37,68	78,08	0,40	0,79
Naive Bayes	39,22	31,15	77,08	0,38	0,80
SVM	32,76	25,41	74,72	0,44	0,77

Fonte: A Autora (2022)

**APÊNDICE M – PERFORMANCE DOS CLASSIFICADORES EM BASE MULTICLASSES BALANCEADA**

	<b>SMOTE</b>	<b>Borderline-SMOTE</b>	<b>CBO</b>	<b>Safe level - SMOTE</b>	<b>ASUWO</b>	<b>ADASYN</b>
<b>SVM</b>						
<b>Acurácia</b>	54,22	58,73	53,31	54,82	24,09	54,38
<b>Sensibilidade</b>	54,22	58,73	53,31	54,82	24,08	54,08
<b>Especificidade</b>	84,74	86,24	84,44	84,94	74,71	84,76
<b>Kappa</b>	0,3715	0,4739	0,3946	0,3976	-0,0119	0,54
<b>AUC</b>	0,8716	0,8988	0,8607	0,8971	0,4675	0,8645
<b>Árvore de decisão</b>						
<b>Acurácia</b>	47,29	56,33	48,19	48,8	0	44,71
<b>Sensibilidade</b>	47,29	56,33	48,19	48,8	0	44,48
<b>Especificidade</b>	82,43	85,44	82,73	82,93	66,71	81,58
<b>Kappa</b>	28,86	0,4177	0,4538	0,3173	-0,3316	0,3217
<b>AUC</b>	0,8099	0,861	0,8126	0,842	0,0719	0,8116
<b>k-NN</b>						
<b>Acurácia</b>	69,28	76,85	59,04	65,06	0	65,56
<b>Sensibilidade</b>	69,28	60,54	59,04	65,06	0	56,22
<b>Especificidade</b>	89,76	79,55	86,35	88,35	66,6	88,52
<b>Kappa</b>	0,1807	0,4498	0,3775	0,5341	-0,3358	0,3902
<b>AUC</b>	0,9076	0,8554	0,8126	0,8855	0,3529	0,8593
<b>Naive Bayes</b>						
<b>Acurácia</b>	48,8	56,63	38,25	54,52	0	47,13
<b>Sensibilidade</b>	48,8	56,63	38,25	54,52	0	46,91
<b>Especificidade</b>	82,93	85,54	79,42	84,84	66,73	82,37
<b>Kappa</b>	0,2676	0,4217	0,1767	0,3936	-0,3309	0,2907
<b>AUC</b>	0,8229	0,8693	0,7976	0,8407	0,0647	0,7795
<b>Análise discriminante</b>						
<b>Acurácia</b>	51,81	57,83	50,6	50,3	3,66	49,85
<b>Sensibilidade</b>	51,81	57,83	50,6	50,3	3,65	49,67
<b>Especificidade</b>	83,94	85,94	83,53	83,43	67,87	83,26
<b>Kappa</b>	0,2219	0,4719	0,3414	0,3373	-0,285	0,2523
<b>AUC</b>	0,8805	0,9196	0,8789	0,902	0,2114	0,8837

Fonte: A Autora (2022)

**APÊNDICE N – PERFORMANCE DOS CLASSIFICADORES EM BASE MULTI-  
CLASSES BALANCEADA PELA TÉCNICA SAES**

<b>SAES 50%</b>			<b>SAES 60%</b>			<b>SAES 75%</b>			<b>SAES 80%</b>		
Classe	Qtd.	%	Classe	Qtd.	%	Classe	Qtd.	%	Classe	Qtd.	%
LB	117	25,55	LB	188	25,00	LB	324	26,96	LB	407	26,22
LM	106	23,14	LM	188	25,00	LM	285	23,71	LM	374	24,10
CISTO	107	23,36	CISTO	188	25,00	CISTO	286	23,79	CISTO	375	24,16
SDV	128	27,95	SDV	188	25,00	SDV	307	25,54	SDV	396	25,52
<b>SVM</b>											
<b>ACC</b>	69,87		80,85		90,02		92,20				
<b>SENS</b>	70,12		80,85		90,23		92,32				
<b>ESP</b>	89,94		93,65		96,67		97,40				
<b><math>\kappa</math></b>	0,1965		0,7447		0,7338		0,7921				
<b>AUC</b>	0,9182		0,9621		0,9719		0,9788				
<b>Árvore de decisão</b>											
<b>ACC</b>	64,19		80,72		87,1		89,76				
<b>SENS</b>	64,22		80,72		87,22		89,85				
<b>ESP</b>	88,03		93,57		95,67		96,58				
<b><math>\kappa</math></b>	0,0451		0,7429		0,6561		0,7268				
<b>AUC</b>	0,8565		0,9362		0,9458		0,9604				
<b>k-NN</b>											
<b>ACC</b>	70,74		82,05		89,27		91,95				
<b>SENS</b>	71,11		82,05		89,43		92,05				
<b>ESP</b>	90,26		94,02		96,42		97,32				
<b><math>\kappa</math></b>	0,2198		0,7606		0,7138		0,7852				
<b>AUC</b>	0,8692		0,9371		0,9579		0,9678				
<b>Naive Bayes</b>											
<b>ACC</b>	69,87		70,61		90,02		92,20				
<b>SENS</b>	70,12		70,61		90,23		92,32				
<b>ESP</b>	89,94		90,2		96,67		97,40				
<b><math>\kappa</math></b>	0,1965		0,6082		0,7338		0,7921				
<b>AUC</b>	0,9182		0,9284		0,9719		0,9788				
<b>Análise discriminante</b>											
<b>ACC</b>	68,78		79,79		86,02		87,63				
<b>SENS</b>	68,89		79,79		86,17		87,68				
<b>ESP</b>	89,57		93,26		95,34		95,86				
<b><math>\kappa</math></b>	0,1674		0,7305		0,6273		0,6701				
<b>AUC</b>	0,9377		0,9284		0,9743		0,9805				

Fonte: A autora (2022).

\*Qtd – quantidade; % - percentual; LB – lesão benigna; LM – lesão maligna; SDV – saudável

**APÊNDICE O – PERFORMANCE DOS CLASSIFICADORES EM BASE MULTI-  
CLASSES DESBALANCEADA E NORMALIZADA**

	Normalização completa		Normalização por classe		
	Max/Min	Max/Min	Zscore	Scale	Center
<b>SVM</b>					
<b>Acurácia</b>	42,67	96,12	48,28	100,00	71,55
<b>Sensibilidade</b>	39,60	95,18	46,67	100,00	68,46
<b>Especificidade</b>	79,28	98,74	82,32	100,00	89,94
<b>Kappa</b>	0,35	0,90	0,27	1,00	0,24
<b>AUC</b>	0,78	0,99	0,76	1,00	0,82
<b>Árvore de decisão</b>					
<b>Acurácia</b>	38,36	70,26	37,07	100,00	45,53
<b>Sensibilidade</b>	38,45	68,84	35,82	100,00	41,50
<b>Especificidade</b>	78,47	89,79	78,28	100,00	80,71
<b>Kappa</b>	0,39	0,21	0,40	1,00	0,34
<b>AUC</b>	0,69	0,87	0,54	1,00	0,63
<b>kNN</b>					
<b>Acurácia</b>	33,19	66,81	39,22	100,00	35,34
<b>Sensibilidade</b>	33,01	60,76	37,27	100,00	32,07
<b>Especificidade</b>	76,04	87,98	78,88	100,00	76,86
<b>Kappa</b>	0,44	0,11	0,38	1,00	0,42
<b>AUC</b>	0,80	0,90	0,70	1,00	0,63
<b>Naive Bayes</b>					
<b>Acurácia</b>	36,64	61,21	33,19	100,00	27,59
<b>Sensibilidade</b>	39,66	57,82	32,92	100,00	29,89
<b>Especificidade</b>	78,56	86,87	77,62	100,00	75,95
<b>Kappa</b>	0,41	0,03	0,44	1,00	0,48
<b>AUC</b>	0,81	0,87	0,67	1,00	0,62
<b>Análise discriminante</b>					
<b>Acurácia</b>	39,66	59,05	33,19	100,00	26,72
<b>Sensibilidade</b>	41,53	57,17	31,55	100,00	23,05
<b>Especificidade</b>	79,45	86,46	76,90	100,00	74,14
<b>Kappa</b>	0,38	0,08	0,46	1,00	0,49
<b>AUC</b>	0,81	0,82	0,63	1,00	0,47

Fonte: A Autora (2022)

## APÊNDICE P – HISTOGRAMAS DAS CARACTERÍSTICAS EXTRAÍDAS DA BASE MULTICLASSES

