



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

TAIRINE ELLEN DA SILVA FREITAS

**USO CONJUNTO DE VARIÁVEIS INTERVALARES E VARIÁVEIS CLÁSSICAS
PARA CLASSIFICAÇÃO DE TERMOGRAMAS DE MAMA POR MEIO DA
DISTÂNCIA DE MAHALANOBIS**

Recife

2021

TAIRINE ELLEN DA SILVA FREITAS

**USO CONJUNTO DE VARIÁVEIS INTERVALARES E VARIÁVEIS CLÁSSICAS
PARA CLASSIFICAÇÃO DE TERMOGRAMAS DE MAMA POR MEIO DA
DISTÂNCIA DE MAHALANOBIS**

Dissertação 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 grau de Mestre em Engenharia Mecânica.

Área de concentração: Processos e Sistemas Térmicos.

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

Coorientador: Prof. Dr. Marcus Costa de Araújo.

Recife

2021

TAIRINE ELLEN DA SILVA FREITAS

**USO CONJUNTO DE VARIÁVEIS INTERVALARES E VARIÁVEIS CLÁSSICAS
PARA CLASSIFICAÇÃO DE TERMOGRAMAS DE MAMA POR MEIO DA
DISTÂNCIA DE MAHALANOBIS**

Dissertação 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 grau de Mestre em Engenharia Mecânica. Área de concentração: Processos e Sistemas Térmicos.

Aprovada em: 26/02/2021.

BANCA EXAMINADORA

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

Prof. Dr. Alvaro Antonio Ochoa Villa (Examinador Interno)
Universidade Federal de Pernambuco

Profa. Dra. Renata Maria Cardoso Rodrigues de Souza (Examinadora Externa)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Agradeço a meus pais por sempre me apoiarem e acreditarem em mim, mesmo quando eu não acredito. A meu namorado Nilton Pontes por me acompanhar durante essa jornada e ser meu porto seguro. A meu amigo de infância José Luna que me acolheu em Recife e, junto com Raquel Santos, tornaram-se minha pequena família recifense. Aos meus amigos da engenharia química e da vida, Vitor Renan, Tuana Medeiros, Jislane Santana e Lucas Medeiros, por estarem sempre me encorajando e torcendo por mim.

Agradeço à professora Dra. Rita de Cássia Fernandes de Lima pela orientação, conselhos e bom humor durante os anos de mestrado.

Agradeço ao professor Drº Marcus Costa de Araújo pela orientação, ajuda, paciência e disponibilidade e por estar sempre disposto a me atender quando eu precisei.

Agradeço à CAPES pelo apoio financeiro.

E agradeço a Deus por caminhar comigo e estar presente em tudo que faço.

RESUMO

O câncer de mama está entre os tipos de câncer mais relatados em mulheres e o diagnóstico precoce torna-se uma tarefa desafiadora nos estágios iniciais de desenvolvimento. Técnicas eficientes e automatizadas, como a termografia, são cada vez mais necessárias para detectar as células cancerosas das mamas. O objetivo deste trabalho é investigar a influência da combinação de variáveis intervalares e variáveis contínuas no processo de classificação de termogramas de mama para detecção do câncer de mama. Na metodologia aplicada, propôs-se uma modificação na matriz de covariância parametrizada de cada classe e ajustou-se o classificador de Mahalanobis para aceitar os dois tipos de variáveis a fim de classificar os indivíduos em benigno, maligno e cisto. Três abordagens do classificador proposto foram aplicadas a duas bases de dados diferentes (50 amostras e 97 amostras) e ajustadas para a classificação binária (câncer e não-câncer). Na base de dados estendida, a abordagem 3 obteve 46,39% de erro global de classificação e 67,86% de sensibilidade à classe maligno. Na classificação binária, a abordagem 3 apresentou o melhor conjunto de resultados: 30,93% de erro global de classificação, 64,28% de sensibilidade à classe câncer e 71% de sensibilidade à classe não-câncer.

Palavras-chave: câncer de mama; variáveis intervalares; distância de Mahalanobis intervalar; imagens termográficas; classificadores.

ABSTRACT

Breast cancer is among the most reported types of cancer in women, and early diagnosis becomes a challenging task in the early stages of development. Efficient and automated techniques, such as thermography, are increasingly needed to detect breast cancer cells. This work aims to investigate the influence of the combination of interval variables and continuous variables in the breast thermogram classification process for detecting breast cancer. In the applied methodology, a modification in the parameterized covariance matrix of each class was proposed and the Mahalanobis classifier was adjusted to accept both types of variables to classify individuals into benign, malignant, and cyst. Three approaches of the proposed classifier were applied to two different databases (50 samples and 97 samples) and adjusted for the binary classification (cancer and non-cancer). In the extended database, approach three obtained 46.39% of global classification error and 67.86% of sensitivity to the malignant class. In the binary classification, approach 3 presented the best set of results: 30.93% of global classification error, 64.28% of sensitivity to the cancer class, and 71% of sensitivity to the non-cancer class.

Keywords: breast cancer; interval variables; Mahalanobis distance interval; thermographic images; classifiers.

LISTA DE FIGURAS

Figura 1 –	Mamografias. a) tecido denso; b) tecido menos denso	18
Figura 2 –	a) imagem digital térmica; b) matriz de temperaturas.	20
Figura 3 –	Segmentação manual. a) curva selecionada manualmente; b) máscara gerada; c) região segmentada.	21
Figura 4 –	Segmentação automática realizada com a ferramenta desenvolvida por Dourado Neto (2014). a) imagem original; b) imagem segmentada.	21
Figura 5 –	Fluxograma simplificado da metodologia	37
Figura 6 –	a) Processamento morfológico sobre as temperaturas na mama; b) Imagem termográfica segmentada da mama.	38
Figura 7 –	Representação gráfica dos intervalos de temperatura para a mama direita e mama esquerda	40
Figura 8 –	Histogramas do erro de classificação geral. a) considerando $S^1_k(\lambda_k)$ como matriz de covariâncias; b) considerando $S^2_k(\lambda_k)$ como matriz de covariâncias; c) considerando $S^3_k(\lambda_k)$ como matriz de covariâncias.	50
Figura 9 –	Histogramas da sensibilidade para a classe maligno. a) considerando $S^1_k(\lambda_k)$ como matriz de covariâncias; b) considerando $S^2_k(\lambda_k)$ como matriz de covariâncias; c) considerando $S^3_k(\lambda_k)$ como matriz de covariâncias.	52
Figura 10 –	Histogramas do erro de classificação geral. a) considerando $S^1_k(\lambda_k)$ como matriz de covariâncias; b) $S^2_k(\lambda_k)$ como matriz de covariâncias; c) considerando $S^3_k(\lambda_k)$ como matriz de covariâncias.	56

LISTA DE TABELAS

Tabela 1 –	Resumo dos resultados de classificação com base de dados contendo 50 amostras	49
Tabela 2 –	Observações que atingiram a menor taxa de erro geral – Equação 4.16a	51
Tabela 3 –	Observações que atingiram a menor taxa de erro geral – Equação 4.16b	51
Tabela 4 –	Observações que atingiram a menor taxa de erro geral – Equação 4.16c	51
Tabela 5 –	Observações que atingiram o maior índice de sensibilidade para classe Maligno – Equação 4.16b	53
Tabela 6 –	Resumo de resultados para a metodologia proposta	54
Tabela 7 –	Resultados de classificação de Araújo (2014) com a base de dados estendida.	54
Tabela 8 –	Taxa de erro global e Sensibilidade da Classe Maligno para modelos clássicos da literatura para classificação com a base estendida para 3 classes (Brasileiro, 2020).	55
Tabela 9 –	Observações que atingiram a menor taxa de erro geral – Equação 4.16a	57
Tabela 10 –	Observações que atingiram a menor taxa de erro geral – Equação 4.16b	57
Tabela 11 –	Observações que atingiram a menor taxa de erro geral – Equação 4.16c	57
Tabela 12 –	Resumo dos resultados de classificação binária com a base de dados estendida – Menor taxa de erro apresentada para a metodologia proposta e a metodologia de Araújo (2017).	57
Tabela 13 –	Taxa de erro global e Sensibilidade da Classe Maligno para os modelos clássicos da literatura usados para classificação binária: câncer x não câncer	58

LISTA DE ABREVIATURAS E SIGLAS

ADS	Análise de Dados Simbólicos
BV	Block Variance
ELM	Aprendizagem de Máquina Extrema
FN	Falso Negativo
FP	Falso Positivo
KNN	K-Nearest Neighbors
MLP	Perceptron Multicamadas
MSPSF	Multiscale Spatially-Weighted Pixel-Contribution and Shape-Feature-Embedded Force
RBFN	Radial Basis Function Network
ROC	Receiver Characteristic Curve
ROI	Region of Interest
SDA	Symbolic Data Analysis
SMO	Sequential Minimal Optimization
SVM	Support Vector Machine
TCLE	Termo de Consentimento Livre e Esclarecido
UFPE	Universidade Federal de Pernambuco
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

LISTA DE SÍMBOLOS

$\dot{y}$	Variável intervalar
C	Número de classes
I	Comprimento referente à intersecção entre duas variáveis intervalares
$I[X]$	Vetor de variáveis intervalares
I_M	Comprimento referente à intersecção entre duas variáveis intervalares obtidas da matriz morfológica
Inf	Límite inferior de uma variável intervalar
K	Comprimento máximo compreendido por duas variáveis intervalares
K_M	Comprimento máximo compreendido por duas variáveis intervalares obtidas da matriz morfológica
MD	Intervalo de temperaturas para a mama direita
ME	Intervalo de temperaturas para a mama esquerda
Morf	Intervalo de dados morfológicos
n	Número de elementos da amostra
n_i	Número de elementos na classe ω_i
Sup	Límite superior de uma variável intervalar
T	Intervalo de temperaturas envolvendo as mamas direita e esquerda
$T_{\max}$	Temperatura máxima
$T_{\min}$	Temperatura mínima
u	Protótipo de da classe
x^c	Centro do intervalo
x^r	Raio do intervalo
y	vetor de intervalos
α	Distância entre os limites inferiores de duas variáveis intervalares
β	Distância entre os limites superiores de duas variáveis intervalares
β_M	Distância entre os limites superiores de duas variáveis intervalares obtidas da matriz morfológica
λ	Parâmetro de controle
ω	Classe

SUMÁRIO

1	INTRODUÇÃO	13
1.1	OBJETIVOS	15
1.1.1	Objetivos específicos	15
1.2	ORGANIZAÇÃO DO TRABALHO	16
1.3	METODOLOGIA EMPREGADA	16
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	TERMOGRAFIA APLICADO AO CÂNCER DE MAMA	18
2.2	PROCESSAMENTO DA IMAGEM TÉRMICA	19
2.2.1	Segmentação	20
2.2.2	Extração de características	21
2.3	ANÁLISE DE DADOS SIMBÓLICOS	22
2.3.1	Variáveis intervalares	23
2.4	CLASSIFICAÇÃO	24
2.4.1	Classificação das imagens termográficas	25
2.4.2	Classificadores de distância mínima	25
2.4.3	Validação cruzada Leave-One-Out	27
3	REVISÃO BIBLIOGRÁFICA	28
3.1	TERMOGRAFIA NA DETECÇÃO DE CÂNCER DE MAMA	28
3.2	SEGMENTAÇÃO	29
3.3	EXTRAÇÃO DE CARACTERÍSTICAS E CLASSIFICAÇÃO	32
3.4	ANÁLISE DE DADOS SIMBÓLICOS	35
3.5	CONTRIBUIÇÃO CIENTÍFICA DO TRABALHO	36
4	METODOLOGIA	37
4.1	BASE DE DADOS	38
4.2	EXTRAÇÃO DE CARACTERÍSTICAS	39
4.2.1	Obtenção de dados contínuos	39
4.2.2	Etapas de classificação	42
4.3	CLASSIFICAÇÃO	43
4.3.1	Distância de Mahalanobis parametrizada para intervalos	43
5	RESULTADOS	48
5.1	RESULTADOS DA METODOLOGIA PROPOSTA	48

5.1.1	Comparação com Araújo (2014) na base original	48
5.1.2	Aplicação da metodologia à base estendida	49
5.2	CLASSIFICAÇÃO BINÁRIA	55
6	CONCLUSÕES	59
	REFERÊNCIAS	61
	APÊNDICE A – OBSERVAÇÕES QUE ATINGIRAM O MAIOR ÍNDICE DE SENSIBILIDADE PARA CLASSE MALIGNO – EQUAÇÃO 4.24^a	67
	APÊNDICE B – OBSERVAÇÕES QUE ATINGIRAM O MAIOR ÍNDICE DE SENSIBILIDADE PARA CLASSE MALIGNO – EQUAÇÃO 4.24^c	68

1 INTRODUÇÃO

O aumento na geração de grandes volumes de dados nos últimos anos impulsionou a utilização dos computadores em tarefas amplamente complexas, gerando resultados cada vez mais precisos. Aprendizagem de máquina é o nome dado ao conjunto de técnicas que permitem aos computadores a realização dessas tarefas (SIDEY-GIBBONS et al., 2019).

No campo médico, a aplicação das técnicas de aprendizagem de máquina pode beneficiar as áreas de diagnóstico e de previsão de resultados. Algoritmos foram recentemente empregados com sucesso para classificar o câncer de pele e para prever a progressão de pré-diabetes para diabetes do tipo 2. Nos últimos anos, a união da aprendizagem de máquina junto com a termografia vem sendo estudada, tornando-se uma ferramenta poderosa no diagnóstico do câncer de mama (ANDERSON et al., 2016; ESTEVA et al., 2017).

O câncer de mama está entre os tipos de câncer mais relatados em mulheres. Alguns dos sintomas são dores em determinadas regiões próximas às mamas, alteração no tom da pele, sensibilidade no mamilo seguida por vazamento de fluido que pode conter sangue e alterações gerais no tamanho ou forma da mama (EKICI et al., 2020; TRAVIS et al., 2016). A maioria desses sintomas não aparece nos estágios iniciais de desenvolvimento, tornando o diagnóstico precoce uma tarefa desafiadora, porém fundamental para aumentar as taxas de sobrevivência (ALFAYES et al., 2020).

Os procedimentos padrões de detecção - como mamografia -, além de demandarem tempo e serem sujeitas a erro, não são eficazes em pacientes com mamas densas ou com tumores muito pequenos. Logo, técnicas eficientes e automatizadas para detectar as células cancerosas das mamas são cada vez mais necessárias, sendo a termografia uma delas (ALFAYES et al., 2020).

A termografia é uma técnica de diagnóstico por imagem que registra a variação da temperatura da superfície do corpo humano baseado na quantidade de radiação térmica que ele emite, visto que todo corpo com temperatura acima de zero absoluto emite radiação. A imagem é capturada com câmeras de infravermelho altamente sensíveis e reproduzida em um software que permite visualizá-la em alta resolução (GONZALEZ-HERNANDEZ et al., 2019; NG, 2009; KAKILETI et al., 2017).

A termografia aplicada à detecção de câncer de mama consiste em medir a temperatura das superfícies das mamas com a câmera de infravermelho para identificar as regiões com temperaturas atípicas após o processamento da imagem térmica. A técnica pode ser usada como instrumento de prevenção e de monitoramento de anomalias mamárias e tem

como vantagens ser rápida, indolor, sem contato, não invasiva, não ionizante e com custo moderado (GONZALEZ-HERNANDEZ et al., 2019; SINGH et al., 2020; DE VASCONCELOS et al., 2018).

As alterações na temperatura devido aos tumores são causadas pelo aumento da atividade metabólica das células cancerígenas, pela formação de novos vasos sanguíneos para nutrir o tumor, pela dilatação dos vasos sanguíneos e pelo calor gerado ao se liberar ácido nítrico no sangue (KAKILETI et al., 2017). Em comparação com a mamografia, a termografia permite detectar os sintomas precoces de câncer entre 8 a 10 anos mais cedo, podendo ser utilizada em mulheres jovens, com implantes e mulheres com mamas densas (ETEHADTAVAKOL et al., 2013; NG, 2009).

A temperatura superficial das mamas saudáveis é praticamente simétrica, logo as alterações das temperaturas oriundas dos tumores ou anormalidades causam a perda dessa simetria, sendo um dos indicadores térmicos do câncer (GONZALEZ-HERNANDEZ et al., 2019). Outros indicadores são: vascularidade hipertérmica, regiões com altas temperaturas, padrões vasculares complexos, padrões de calor aureolar e periareolar e diferenças de temperatura maior que 2°C entre os dois seios (KENNEDY et al., 2009; NG, 2009).

Durante o procedimento são obtidos termogramas, imagens térmicas ou infravermelhas que fornecem um mapa de temperaturas das mamas da paciente. Cada mama é representada por uma matriz de temperaturas, onde cada pixel da imagem corresponde a um valor. A região da mama é então formada por um conjunto de pixels, isto é, de temperaturas, onde a temperatura global da mama pode ser interpretada como que apresentando um comportamento de natureza intervalar.

As variáveis intervalares são abordadas na análise de dados simbólicos (ADS), área que abrange o estudo das variáveis simbólicas que se baseiam na ideia de que inferências estatísticas são geralmente necessárias a nível de grupo e não individual (BILLARD et al., 2006).

O agrupamento de dados permite que as observações sejam descritas por listas, intervalos, histogramas e afins - dados simbólicos -, em vez de valores pontuais únicos, os chamados dados clássicos tradicionais. De forma simplificada, os dados simbólicos podem ser definidos como a tomada de valores como hipercubos ou produtos de distribuição Cartesiana no espaço p -dimensional $\mathbb{R}^p$, em contraste com observações clássicas cujos valores são pontos em $\mathbb{R}^p$ (BILLARD et al., 2019).

Araújo (2017) desenvolveu um procedimento de classificação de imagens termográficas de anomalias mamárias com base na distância mínima de Mahalanobis para

dados intervalares das mamas. A distância foi definida através da matriz de covariância comum parametrizada para cada classe C_k do algoritmo de classificação. O autor definiu um parâmetro de controle λ_k para classe, que possui influência direta na acurácia do classificador.

Neste trabalho será investigada a influência da combinação de variáveis intervalares e variáveis clássicas aplicadas ao problema de classificação de termogramas de mama. Durante o desenvolvimento desta dissertação, serão propostas a modificação da matriz de covariância parametrizada de cada classe e o ajuste do classificador de Mahalanobis para que possa aceitar variáveis intervalares e variáveis contínuas. A partir das modificações e da inclusão do novo conjunto de variáveis, será realizado um estudo analisando a melhor combinação dos parâmetros λ_k . A inclusão das variáveis contínuas tem como finalidade tornar o classificador mais completo, utilizando o máximo de informações resultantes da etapa de extração de características.

O presente trabalho faz parte do projeto de pesquisa “**Análise da viabilidade do uso de câmera termográfica como ferramenta auxiliar no diagnóstico de câncer de mama em hospital público localizado em clima tropical**” e aprovado pelo Comitê de Ética da Universidade Federal de Pernambuco (UFPE) – Brasil, registrado no Ministério da Saúde sob o nº CEP/CCS/UFPE Nº 279/05. O projeto de pesquisa se encontra em andamento desde novembro de 2005.

1.1 OBJETIVOS

O presente trabalho tem como objetivo geral ajustar um classificador de distância mínima de Mahalanobis de forma que aceite uma combinação de variáveis intervalares e variáveis contínuas, para ser utilizado na classificação de termogramas de mama como ferramenta de detecção do câncer de mama.

1.1.1 Objetivos específicos

- Levantamento bibliográfico sobre o tema;
- Modificar a matriz de covariâncias parametrizada de cada classe no classificador desenvolvido por Araújo (2017) para aceitar intervalos e variáveis contínuas;
- Definir um conjunto de variáveis clássicas para acrescentar ao classificador desenvolvido por Araújo (2017);
- Utilizar o classificador modificado em uma nova base de dados estendida;

- Comparar os resultados obtidos com os resultados do classificador original de Araújo (2017).

1.2 ORGANIZAÇÃO DO TRABALHO

Além desse capítulo de introdução, essa dissertação é constituída por mais cinco capítulos, descritos a seguir:

- O Capítulo 2 apresenta a fundamentação teórica da pesquisa trazendo o conceito de termografia e seu papel na detecção do câncer de mama, o pré-processamento da imagem térmica e o processo de classificação das imagens, apresentando o conceito de classificadores de distância mínima. O capítulo também abordará a análise de dados simbólicos com destaque para as variáveis intervalares;
- O Capítulo 3 apresenta uma revisão da literatura com relação à utilização da termografia na detecção de câncer de mama, ao pré-processamento das imagens térmicas e à análise de dados simbólicos, destacando os trabalhos desenvolvidos principalmente nos últimos dez anos;
- O Capítulo 4 descreve a metodologia adotada para o desenvolvimento desse trabalho;
- O Capítulo 5 exhibe os resultados obtidos e as discussões pertinentes a eles;
- O Capítulo 6 apresenta as conclusões resultantes do presente trabalho, as considerações finais e sugestões para trabalhos futuros.

1.3 METODOLOGIA EMPREGADA

O presente trabalho é um processo de simulação de uma pesquisa do tipo quantitativa, no qual foi utilizado o software Matlab® versão 2013 para a simulação.

A metodologia está dividida nas seguintes etapas:

- processamento do conjunto de dados original e obtenção das variáveis contínuas;
- extração de características que compõem o vetor de entrada no processo de classificação, ou seja, os dados intervalares e os dados contínuos;
- classificação, na qual o vetor de entrada é utilizado em um classificador de distância mínima que usa uma adaptação da distância de Mahalanobis baseada em intervalos.

A metodologia proposta foi aplicada a uma base de dados contendo 50 amostras e a uma base de dados estendida contendo 97 amostras. Os resultados foram avaliados em relação

à taxa de erro global de classificação e aos índices de sensibilidade e especificidade para cada classe em dois estudos: um com três classes (benigno, maligno e cisto) e outro com duas classes (câncer e não câncer).

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os conceitos e estudos relacionados a imagens termográficas e a análise de suas características, assim como a extração de dados e métodos de classificação.

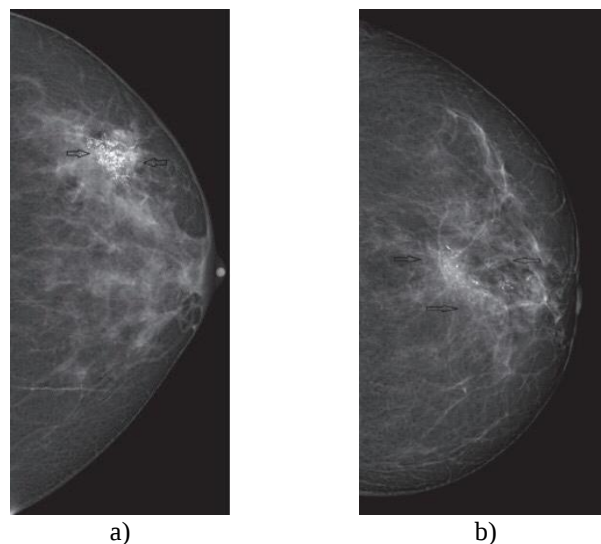
2.1 TERMOGRAFIA APLICADA AO CÂNCER DE MAMA

A detecção precoce do câncer de mama é muito eficaz para aumentar as taxas de sobrevivência dos pacientes e diminuir as taxas de mastectomia. Por dependerem da experiência e análise a olho nu de médicos e oncologistas, as técnicas de diagnóstico tradicionais são demoradas e propensas a erros; assim sendo, são necessárias técnicas eficientes e automatizadas para detectar as células lesionadas (ALI et al., 2015).

Mamografia, ressonância magnética, tomografia e ultrassom são algumas das técnicas de diagnóstico por imagem que podem identificar os sinais precoces do câncer de mama, sendo eles distorção arquitetônica, massa, microcalcificação e assimetrias nas mamas (NOVER et al., 2009).

A técnica de detecção mais conhecida e utilizada é a mamografia. Porém, em mulheres jovens e/ou com mamas densas o procedimento não funciona bem, visto que o tecido denso pode esconder o tumor (ETEHADTAVAKOL et al., 2013; LEE et al., 2015). A diferença pode ser vista na Figura 1. Diante dessa limitação, surge o uso da termografia na detecção do câncer de mama.

Figura 1 - Mamografias. a) tecido denso; b) tecido menos denso



Fonte: AYDINER et al., 2019

A termografia detecta as lesões de câncer em seu estágio inicial, antes que mudanças na estrutura da mama sejam percebidas, podendo ser usada em mulheres jovens, mulheres com implantes mamários e mulheres com altas densidades mamárias (ARAKERI et al., 2015; ETEHADTAVAKOL et al., 2013). Outras vantagens da termografia englobam ser uma técnica de diagnóstico por imagem indolor, não invasiva, sem contato e totalmente segura para a paciente, pois, ao contrário da mamografia, não há exposição à radiação ionizante (raios X) (KENNEDY et al., 2009).

Qualquer desenvolvimento canceroso está associado à formação de inflamação e de vasos sanguíneos, causando um aumento no perfil de temperatura da região, que será identificado pela captura da imagem térmica (EKICI et al., 2020).

Na detecção do câncer de mama, a imagem térmica ou infravermelha fornece um mapa de temperaturas das mamas da paciente, denominado termograma. A existência de tumor nas mamas aumenta a temperatura da superfície circundante, causando uma assimetria térmica entre o seio saudável e o seio não saudável. A termografia é então capaz de capturar a diferença de temperatura existente, permitindo identificar anormalidades térmicas (EKICI et al., 2020).

Como a termografia infravermelha é inerentemente uma medição de superfície, o transporte térmico entre o tumor e a superfície da pele é um fator crítico na detecção termográfica; logo, tumores pequenos, profundos e/ou difusos podem ser indetectáveis durante o procedimento. Sendo assim, a capacidade da termografia para detecção do câncer de mama depende diretamente das propriedades térmicas do tecido, do tamanho, da profundidade, da taxa de crescimento e da concentração do tumor (GAUTHERIE, 1980 apud LOZANO et al., 2019).

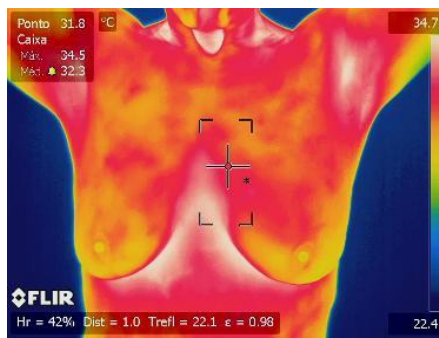
Os termogramas são capturados através de uma câmera de infravermelho, que consiste em lentes transparentes a infravermelho, um sistema de escaneamento e um detector de infravermelho (LAU, 1977 apud LOZANO et al., 2019). Após a captura, as imagens são processadas em um software adequado, para que a extração de características possa ser realizada posteriormente.

2.2 PROCESSAMENTO DA IMAGEM TÉRMICA

Uma imagem digital pode ser definida como uma função bidimensional $f(x,y)$, na qual x e y são coordenadas espaciais e f é chamado de intensidade da imagem no ponto (x,y) ; ela é composta de um número finito de elementos, os chamados pixels (GONZALES et al., 2009).

Um termograma é representado por uma imagem pseudocolor, imagem digital indexada que contém dois componentes: uma matriz de temperaturas, T , e um mapa de cor (Figura 2). A função $f(x,y)$ está associada à matriz de temperaturas, onde o valor $T(x,y)$ corresponde à temperatura no ponto (x,y) do plano e funciona como ponteiro no mapa de cor (ARAÚJO, 2014).

Figura 2 - a) imagem digital térmica; b) matriz de temperaturas.



a)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
2	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
3	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
4	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
5	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
6	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
7	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
8	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
9	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
10	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
11	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
12	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
13	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
14	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
15	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
16	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
17	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
18	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
19	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
20	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
21	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
22	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
23	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
24	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
25	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
26	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
27	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
28	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301
29	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302
30	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301	30.302	30.301

b)

As cores dos pixels serão determinadas pelo mapa de cor utilizado e pelos limites inferior e superior de temperatura definidos pela escala do termograma. Ainda que haja mudança da escala de temperatura ou do mapa de cor, os valores de temperatura permanecem constantes, pois cada pixel representa o valor de temperatura definido pela matriz T (QUEIROZ, 2016).

2.2.1 Segmentação

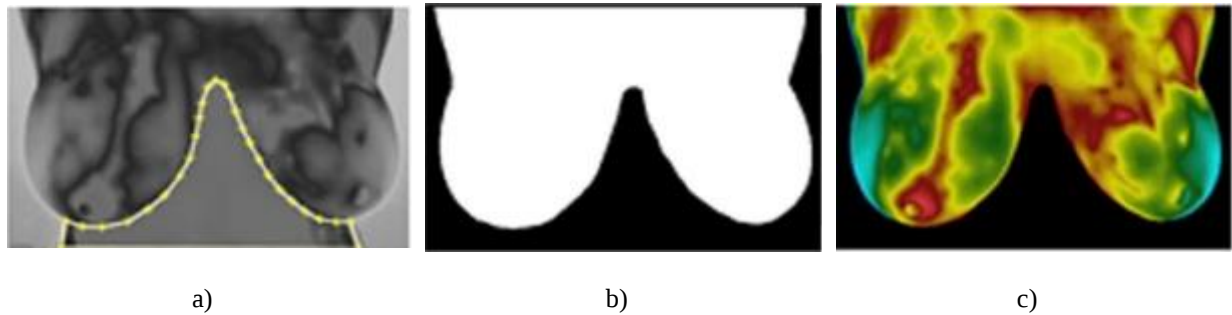
O processo de segmentação da região de interesse (ROI - Region of Interest) consiste em identificar a área da mama que será processada, separando-a do plano de fundo da imagem (SINGH et al., 2020).

Normalmente, um tumor maligno apresenta forma irregular, então o tamanho e as fronteiras das regiões mais quentes contidas na imagem são fundamentais para determinar a melhor maneira de detectar as anormalidades. A segmentação de uma imagem médica é uma etapa importante que deve ser realizada precisamente, pois é imprescindível que a anormalidade esteja dentro da região de interesse (GOLESTANI et al., 2014).

A extração da região de interesse da termografia da mama pode ser totalmente automática ou totalmente manual (MACHADO et al., 2013) e deve incluir todos os tecidos

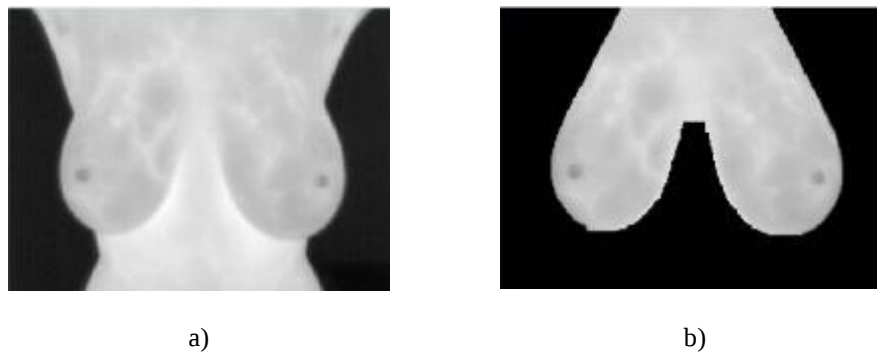
mamários e os grupos de gânglios próximos; porém, por ser uma tarefa desafiadora, alguns autores preferem a extração manual ou semiautomática devido à natureza amorfa e limites irregulares nas imagens (ZHOU et al., 2004). Exemplos de segmentação manual e totalmente automática podem ser vistos nas Figuras 3 e 4, respectivamente.

Figura 3 - Segmentação manual. a) curva selecionada manualmente; b) máscara gerada; c) região segmentada



Fonte: GOGOI et al., 2018

Figura 4 - Segmentação automática realizada com a ferramenta desenvolvida por Dourado Neto (2014). a) imagem original; b) imagem segmentada.



Fonte: DE VASCONCELOS et al., 2018

Ao concluir a separação da região da mama, alguns recursos ou características são extraídos para alimentar algoritmos de classificação, que serão aplicados para classificar as possíveis anormalidades identificadas nos termogramas (SUGANTHI et al., 2014).

2.2.2 Extração de características

As técnicas de extração ajudam a identificar um aumento na temperatura devido a outras razões além do tumor e ajudam a superar algumas desvantagens da termografia como a inabilidade em diagnosticar tumores pequenos e diferenças na interpretação das imagens por mais de um médico (SINGH et al., 2020).

Nessa etapa, determinados parâmetros são extraídos do termograma para compor os dados de entrada da ferramenta de classificação. Os tipos básicos de extração consideram uma ou mais das características a seguir: escala de cinza e de cores, valores de temperatura, textura, conjunto de pontos, contornos ou curvas, superfícies e regiões específicas (OLIVEIRA, 2015).

Dentre as características escolhidas, é fundamental identificar as informações mais relevantes para a classificação, descartando aquelas que não têm papel na distinção entre as classes (QUEIROZ, 2016).

Apesar dos métodos de seleção de características ajudarem a reduzir a complexidade computacional, a manipulação e o armazenamento de grandes quantidades de dados ainda são considerados um problema. Diante disso, surgiram os dados ou variáveis simbólicas, cujo objetivo é estender os métodos de análise de dados comuns para que possam ser aplicados a dados complexos e de maior número (BOCK et al., 2000).

2.3 ANÁLISE DE DADOS SIMBÓLICOS

A análise de dados simbólicos (SDA - Symbolic Data Analysis) serve para representar grandes conjuntos de dados de maneira compacta, abrangendo os métodos para manipular tais representações (BILLARD et al., 2006; DIDAY et al., 2008). O desenvolvimento recente de uma maneira mais barata e mais gerenciável de armazenar as grandes quantidades de dados digitais ajuda a registrar dados complexos em quase todos os campos de estudo (BOCK et al., 2000). Como exemplo, pode-se citar o uso de intervalos para representar dados imprecisos ou incertos (BILLARD et al., 2006; DIDAY et al., 2008).

A análise de dados simbólicos se baseia na noção de que inferências estatísticas são geralmente necessárias em nível de grupo e não em nível individual. A SDA basicamente agrega os dados individuais em grupos (isto é, símbolos) para construir modelos que possam inferir diretamente nos grupos criados (BILLARD, 2011 apud BERANGER et al., 2018; BILLARD et al., 2006).

Variáveis intervalares ocorrem naturalmente e são muito comuns atualmente, logo é necessário desenvolver métodos estatísticos apropriados para esse tipo de variáveis. Além disso, em muitas aplicações do mundo real, as informações disponíveis são intrinsecamente imprecisas, incertas e ambíguas; portanto, não podem ser expressas por um único valor numérico, como nos dados clássicos (ROY et al., 2019).

A escolha mais comum de agrupamentos é o intervalo aleatório, no qual, para as observações a nível individual $X_1, \dots, X_n \in \mathbb{R}$, o intervalo é tipicamente construído como $S = (\min_i X_i, \max_i X_i) \subseteq \mathbb{R}$ (BERANGER et al., 2018).

Os intervalos $S_1, \dots, S_m \in \mathcal{S}$ são considerados os novos "pontos" de dados, onde cada dado é uma distribuição de algum tipo com uma estrutura de distribuição interna. Então a inferência estatística será realizada diretamente nos intervalos, com referência a suas formas de distribuição e sem levar em conta os dados a nível individual (BERANGER et al., 2018).

2.3.1 Variáveis intervalares

Seja $\mathbf{I}[\mathbf{X}]$ a representação do vetor de variáveis intervalares de dimensão $n \times 1$, onde n denota o número de elementos da amostra. Então, pode-se escrever $\mathbf{I}[\mathbf{X}]$ como:

$$\mathbf{I}[\mathbf{X}] = \begin{pmatrix} [x_1^-, x_1^+] \\ \vdots \\ [x_n^-, x_n^+] \end{pmatrix} \quad (2.1)$$

Onde a i -ésima linha de $\mathbf{I}[\mathbf{X}]$ pertence a i -ésima variável e cada componente é um intervalo representado por $I[x]_i \equiv [x_i^-, x_i^+]$, onde $x_i^- < x_i^+$ e $i = 1, \dots, n$. O intervalo genérico $I[x]_i$ também pode ser expresso pelo par $\{x_i^c, x_i^r\}$, onde x_i^c representa o centro do intervalo e x_i^r representa o raio do intervalo, dados pelas Equações 2.2 e 2.3, respectivamente:

$$x_i^c = \frac{x_i^+ + x_i^-}{2} \quad (2.2)$$

$$x_i^r = \frac{x_i^+ - x_i^-}{2} \quad (2.3)$$

Ou seja, transforma-se ortogonalmente a variável intervalar em duas variáveis não correlacionadas. Portanto, a matriz de dados intervalares pode ser representada unicamente por $\mathbf{I}[\mathbf{X}] \equiv \{X^c, X^r\}$, onde:

$$X^c = \begin{pmatrix} x_1^c \\ \vdots \\ x_n^c \end{pmatrix} \text{ e } X^r = \begin{pmatrix} x_1^r \\ \vdots \\ x_n^r \end{pmatrix} \quad (2.4)$$

Na imagem termográfica, pode-se agrupar as temperaturas dos pixels contidos na região de interesse no intervalo $y = [T_{min}, T_{max}]$, onde T_{min} representa a temperatura mínima obtida na região de interesse e T_{max} , a temperatura máxima obtida na região de interesse, ou seja, os limites inferior e superior do intervalo de temperatura.

2.4 CLASSIFICAÇÃO

A extração de características é seguida pela etapa de classificação, passo fundamental nas técnicas de reconhecimento de padrões, onde as termografias serão agrupadas a fim de apontar se a paciente possui uma mama saudável ou não saudável a partir das informações extraídas.

Segundo Duda et al (2009), reconhecer padrões é o ato de coletar dados e tomar uma ação com base na categoria do padrão. A classificação dos padrões geralmente é um processo supervisionado no qual uma regra de decisão é construída com base em um conjunto de amostras previamente classificadas para ser utilizada na classificação de novos objetos (FERREIRA et al., 2007; SCHAEFER et al., 2009).

Problema de classificação:

Seja $\mathbf{x} \in \mathbb{R}^p$ um conjunto de objetos definidos por p variáveis $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$, onde x_{ij} é o valor da j -ésima variável de $\mathbf{x}_i$ e T , o operador de transposição. Suponha que existem n observações divididas em C classes – n_1 na classe ω_1 , n_2 na classe ω_2 e assim por diante – de forma que $n = n_1 + n_2 + \dots + n_C$. A matriz $\mathbf{X}$ com n observações a p variáveis é definida por:

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \quad (2.5)$$

onde as linhas representam as observações individuais e as colunas representam as variáveis.

Um problema de classificação consiste em alocar uma observação individual $\mathbf{x}$, descrita por um vetor de atributos, em uma das C classes existentes. Para isso, uma regra de decisão divide o espaço de características em C regiões distintas Ω_i , $i = 1, 2, \dots, C$, definidas

com base na regra de classificação. O vetor de entrada $\mathbf{x}$ será alocado na classe ω_i quando ele pertencer à região Ω_i dentro do espaço de características (HASTIE et al., 2009).

2.4.1 Classificação das imagens termográficas

Em imagens digitais, o reconhecimento de padrões é feito por meio de medições estatísticas derivadas das características espectrais dos pixels. A imagem será então destinada à sua classe mais provável utilizando a regra de decisão (NOVA, 2017).

O sistema de classificação se inicia com a aquisição das imagens termográficas obtidas através da câmera termografia. Os termogramas passam pela etapa de pré-processamento para eliminar as informações irrelevantes de forma que as características sejam extraídas dos dados processados. Assim, tem-se a etapa de treinamento, na qual uma regra de classificação é construída com base em um conjunto de dados de treinamento. Após a regra de classificação ser devidamente treinada, ela pode ser usada em um grupo de dados independentes, o grupo de teste, constituindo a etapa de validação (ARAÚJO, 2014).

2.4.2 Classificadores de distância mínima

A formação de subgrupos ou subconjuntos em um conjunto de indivíduos é governada pelos índices de similaridade e dissimilaridade, funções de proximidade que representam, respectivamente, o grau de semelhança ou de diferença entre duas observações, ou seja, a distância entre elas (BILLARD et al., 2006).

Seja a e b dois objetos pertencentes ao conjunto E . Então, o índice de dissimilaridade $d(a,b)$ é a medida que satisfaz:

- i. $d(a,b) = d(b,a)$;
- ii. $d(a,a) = d(b,b) > d(a,b)$ para todo $a \neq b$;
- iii. $d(a,a) = 0$ para todo $a \in E$;

Uma medida de distância, também chamada de métrica, satisfaz ainda os seguintes itens:

- iv. $d(a,b) = 0$ implica que $a = b$;
- v. $d(a,b) \leq d(a,c) + d(c,b)$ para todo $a, b, c \in E$.

O classificador de distância mínima é um método de classificação supervisionada simples que utiliza o ponto central para representar uma classe no conjunto de dados de

treinamento (SATHYA et al., 2017). O classificador atribui um padrão desconhecido $\mathbf{x}$ a uma classe ω_i , $i = 1, 2, \dots, C$, com base na menor distância entre o padrão $\mathbf{x}$ e o protótipo da classe, representados pelas médias de cada classe.

Os protótipos da i -ésima classe $\mathbf{u}_i = (u_{i1}, u_{i2}, \dots, u_{ip})$, $i = 1, 2, \dots, C$ representam os indivíduos cujos atributos correspondem aos atributos médios de todos os indivíduos da classe.

$$\mathbf{u}_i = \left(\frac{\sum_{x \in C_i} x_{i1}}{n_i}, \frac{\sum_{x \in C_i} x_{i2}}{n_i}, \dots, \frac{\sum_{x \in C_i} x_{ip}}{n_i} \right) \quad (2.6)$$

O objeto $\mathbf{x}$ é atribuído à classe para qual a distância entre $\mathbf{x}$ e o protótipo da classe, u_i , analisada é mínima, ou seja,

$$d(\mathbf{x}, u_i) = \min_i d(\mathbf{x}, u_i), \quad i = 1, 2, \dots, C \quad (2.7)$$

Neste trabalho, foi utilizada a distância de Mahalanobis no classificador de distância mínima para dados intervalares definida por Souza et al. (2004).

Seja $\Gamma = \dot{y}_1, \dot{y}_2, \dots, \dot{y}_m$ um grupo de m objetos simbólicos descritos por p variáveis intervalares, no qual cada objeto $\dot{y}_i$ ($i = 1, 2, \dots, m$) é representado como um vetor de intervalos $\dot{y}_i = ([a_{i1}, b_{i1}], [a_{i2}, b_{i2}], \dots, [a_{ip}, b_{ip}])^T$. Seja P uma partição de Γ em K classes $C_1, C_2, \dots, C_K$, onde cada classe C_k possui um protótipo u_k também representado como um vetor de intervalos $\dot{g}_k = ([\alpha_{k1}, \beta_{k1}], [\alpha_{k2}, \beta_{k2}], \dots, [\alpha_{kp}, \beta_{kp}])^T$ que corresponde ao vetor de médias dos intervalos de C_k . Para $j = 1, 2, \dots, p$:

$$[\alpha_{kj}, \beta_{kj}] = \left[\frac{\sum_{i \in C_k} a_{ij}}{n_k}, \frac{\sum_{i \in C_k} b_{ij}}{n_k} \right] \quad (2.8)$$

onde n_k representa o número de elementos na classe C_k .

O classificador de distância mínima aloca um objeto $\dot{v}_i$ para a classe C_k se

$$d^\lambda(\dot{v}_i, \dot{g}_k) \leq d^\lambda(\dot{v}_i, \dot{g}_r) \quad \forall r, r = 1, 2, \dots, K. \quad (2.9)$$

A função $d^\lambda(\dot{v}_i, \dot{g}_k)$ corresponde a uma distância de Mahalanobis parametrizada para intervalos. Seja $v_{i1} = (a_{i1}, a_{i2}, \dots, a_{ip})^T$ e $v_{is} = (b_{i1}, b_{i2}, \dots, b_{ip})^T$ dois vetores contendo os

valores de Inf e Sup de $\dot{v}_i$, respectivamente; e $g_{kl} = (\alpha_{k1}, \alpha_{k2}, \dots, \alpha_{kp})^T$ e $g_{ks} = (\beta_{k1}, \beta_{k2}, \dots, \beta_{kp})^T$ dois vetores contendo os valores de Inf e Sup de $\dot{g}_k$, respectivamente.

Na Equação, está representada a distância de Mahalanobis $\delta(\dot{v}_i, \dot{g}_k)$ entre a observações intervalares $\dot{v}_i$ e $\dot{g}_k$, que é definida como a combinação da distância entre os valores inferiores das observações intervalares e da distância entre os valores superiores das observações intervalares:

$$\delta(\dot{v}_i, \dot{g}_k) = d(v_{il}, g_{kl}) + d(v_{is}, g_{ks}) \quad (2.10)$$

sendo

$$d(v_{il}, g_{kl}) = (v_{il} - g_{kl})^T S_I (v_{il} - g_{kl}) \quad (2.11)$$

$$d(v_{is}, g_{ks}) = (v_{is} - g_{ks})^T S_S (v_{is} - g_{ks}) \quad (2.12)$$

onde $d(v_{il}, g_{kl})$ é a distância de Mahalanobis entre os vetores v_{il} e g_{kl} , $d(v_{is}, g_{ks})$ é a distância de Mahalanobis entre os vetores v_{is} e g_{ks} e S_I e S_S são as matrizes de covariância para os valores de Inf e Sup, respectivamente.

2.4.3 Validação cruzada Leave-One-Out

O desempenho de um classificador é avaliado a partir da divisão das amostras em dois grupos mutuamente excludente: o grupo de treino, no qual se baseia o classificador e um grupo de teste, onde o classificador treinado é aplicado (Araújo, 2017). Quando o número de amostras é pequeno ou moderado, a divisão do conjunto entre esses dois grupos não consegue fornecer informações suficiente para treinar o classificador. Nesses casos, utiliza-se técnicas específicas de validação cruzada, sendo uma delas a técnica *Leave-One-Out* (SOUZA, 1999).

No método de validação cruzada *Leave-One-Out*, cada um dos n indivíduos contido no grupo de amostras é utilizado como grupo de teste, enquanto os $n - 1$ indivíduos restantes são utilizados como grupo de treino para construir a regra de classificação. A validação é feita até que todos os elementos assumam o papel de grupo de teste (FERREIRA, 2007).

3 REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta, em ordem cronológica, estudos e trabalhos sobre a termografia aplicada à detecção de câncer de mama, técnicas de segmentação de imagens, extração de características, classificação e análise de dados simbólicos.

3.1 TERMOGRAFIA NA DETECÇÃO DE CÂNCER DE MAMA

Os primeiros estudos envolvendo câncer de mama e medição de temperatura surgiram desde a década de 1950. Lawson (1956) observou que 26 indivíduos com câncer de mama apresentaram um aumento de $1,3^{\circ}\text{C}$ na região próxima ao tumor em comparação a outras regiões e, em 1963, juntamente com Chutgai, Lawson descobriu que o aumento da temperatura na região tumoral podia chegar a $3,5^{\circ}\text{C}$, com média de 2°C .

Gautherie (1980) mediu a distribuição da temperatura interna em mamas saudáveis e com câncer utilizando uma agulha termopar. A região com tumor apresentou aumento de aproximadamente 3°C em relação às mamas saudáveis. O autor mediu a atividade metabólica e a perfusão nos tecidos (saudáveis e com tumor) de 128 pacientes. Após o estudo, ele concluiu que a atividade metabólica dos tumores é 15 vezes maior que a do tecido saudável (GONZALEZ-HERNANDEZ et al., 2019).

A partir dos anos 90, com a evolução dos equipamentos de aquisições de imagens térmicas, surgiram os primeiros trabalhos utilizando as câmeras termográficas. Estas eram acopladas a computadores com programas específicos que permitiam que as imagens obtidas fossem processadas e arquivadas com boa qualidade (DE MEIRA et al., 2014). Porém, devido à inconsistência dos resultados em decorrência de análises de formato qualitativo das imagens e à interpretação subjetiva dos diagnósticos, a termografia foi aos poucos sendo deixada de lado. O interesse no uso das imagens termográficas para diagnóstico do câncer de mama voltou a ganhar força após o amadurecimento dos estudos de aprendizagem de máquina. Unindo as ferramentas estatísticas e de aprendizagem de máquina à interpretação das imagens, foi possível realizar análises quantitativas.

Ng (2009) publicou a descrição sistemática dos padrões e protocolos necessários durante a aquisição das imagens infravermelhas. Kennedy et al (2009) propuseram um método de detecção de câncer de mama combinando termografia e mamografia. Os resultados mostraram que a detecção baseada em termograma alcançou 83% de sensibilidade, a baseada

em mamografia alcançou 90% e termograma combinado com mamografia alcançou 95% de sensibilidade.

Em 2013, Bezerra et al. desenvolveram um aparato mecânico e um protocolo de procedimentos para padronizar a obtenção de imagens termográficas para o diagnóstico de câncer mamário. Com as imagens padronizadas, os autores puderam estimar os parâmetros termofísicos das mamas a partir dos termogramas.

Kandlikar et al., (2017) estudaram detalhadamente a termografia na detecção de mama, abordando as diferentes modalidades de termografia, os protocolos requeridos antes e durante o processo de obtenção das imagens e os esforços atuais para melhorar a detecção precoce do câncer de mama usando a termografia.

Em 2017, uma ferramenta complementar não invasiva para aquisição de termogramas foi introduzida (GARDUÑO-RAMÓN et al., 2017). Nesse método, o aparelho de termografia infravermelho é posicionado a 1,2 m de distância da paciente sentada. A essa distância, o tamanho mínimo do ponto mensurável é de 0,5 cm, permitindo a medição de tumores cancerígenos maiores que 0,5 cm. Dois tipos de protocolo podem ser empregados na captura da imagem: protocolos estáticos e dinâmicos. O protocolo estático envolve uma única imagem capturada após 10-15 minutos de estabilização térmica durante o repouso, enquanto o protocolo dinâmico envolve uma série de termogramas capturados a cada 15 segundos durante cinco minutos.

No mesmo ano, Gogoi et al. (2018) também propuseram um protocolo padrão de aquisição de termograma de mama. Fatores que influenciam a termografia, como informações pessoais e médicas do paciente, bem como a condição da sala, foram ajustados antes do exame. A aquisição dos termogramas mamários foi realizada em cubículo preto para adquirir um fundo homogêneo. Uma cama com mesa foi projetada para garantir a aclimação dos pacientes, que deveriam permanecer deitados por 15 minutos. Os pacientes foram posicionados a 1 metro de distância da câmera térmica, exceto aqueles com anatomia corporal maior do que a média e a distância de melhor ajuste foi determinada. As imagens foram capturadas nas vistas supina, frontal, lateral esquerda, lateral direita, oblíqua esquerda e oblíqua direita.

3.2 SEGMENTAÇÃO

O pré-processamento das imagens termográficas abrange a separação do background do paciente, removendo tudo que não será necessário para analisar o termograma. O principal

objetivo da segmentação da região de interesse é identificar a região da mama da qual as características serão extraídas e utilizadas como dados de entrada no classificador. A segmentação pode ser feita de forma manual, automática e semiautomática e, ao longo dos anos, variadas técnicas foram estudadas por diferentes pesquisadores (SINGH et al., 2020).

Qi et al (2008) aplicaram a transformada de Hough para segmentar as mamas e obter a curva de extração de características. A segmentação automática resultou na extração dos limites das mamas, que foram então processadas utilizando técnicas de extração de características e de classificação (SINGH et al., 2020).

Jin-Yu et al. (2009) propuseram um algoritmo genético para melhorar a relação sinal-ruído das imagens térmicas e para reduzir o tempo computacional envolvido na segmentação das imagens. Este método combina os benefícios do algoritmo genético com o método bidimensional de Otsu. De acordo com os autores, o tempo de redução média da segmentação das imagens térmicas com resolução de 198x173 pixels foi 4,83s (RAGHAVENDRA et al., 2019).

EtehadTavakol et al. (2010) aplicaram os algoritmos K-means e fuzzy C-means para segmentação de imagens infravermelhas de mamas. Seis imagens foram segmentadas e algumas características foram extraídas das regiões com as duas temperaturas mais elevadas.

Motta et al. (2010) apresentaram um método de segmentação automática contendo sete etapas principais:

- (i) identificar o limite inferior da região de interesse, ou seja, o sulco inframamário;
- (ii) eliminar o background através do método de Otsu;
- (iii) identificar o limite superior da região de interesse ao encontrar a axila;
- (iv) eliminar os braços e outras regiões;
- (v) determinar o eixo central e separar as mamas;
- (vi) detectar o sulco inframamário; e
- (vii) deslocar verticalmente.

Golestani et al. (2014) discutiram e compararam três técnicas de segmentação: K-means, fuzzy C-means e level set. Trinta imagens térmicas foram utilizadas nesse trabalho e resultados experimentais provaram que o algoritmo level set resultou em melhor eficiência, acurácia e robustez.

Mahmoudzadeh et al. (2015) apresentaram uma extensão única de um modelo oculto de Markov para melhorar a segmentação das imagens termográficas de mamas. Os resultados mostraram que houve uma redução no tempo de execução, comparado com outras técnicas de segmentação como fuzzy C-means.

De Oliveira et al. (2015) propuseram um método de segmentação automática utilizando limiar do histograma para balanceamento, agrupamento e detecção de cantos pelo método de Shi-Tomasi.

Pramanik et al. (2015) realizaram a segmentação utilizando o método de Otsu para remover o background, posteriormente aplicando uma técnica de reconstrução.

Kamath et al. (2015) apresentaram um método de segmentação automática utilizando Projection Profile Approach, método que pode ser generalizado para vários tipos de imagens termográficas com pequenas modificações, como padronização do background da imagem, altura e remoção do ruído.

Devido às limitações do baixo contraste, a etapa mais desafiadora da segmentação das mamas é identificar os limites inferiores e as dobras inframamárias das mamas. Suganthi et al. (2014) propuseram a integração do filtro de difusão anisotrópica com a estrutura do level set para evitar suavização através da fronteira e preservar fronteiras nítidas. A similaridade das regiões de interesse segmentadas e as imagens de base alcançam uma precisão média de 98%. O método level set foi aprimorado por Golestani et al. (2014) para aumentar a precisão com tempo mínimo de computação.

Sathish et al. (2017) destacaram as bordas da região aplicando um ajuste de curva polinomial. Este método foi capaz de detectar pequenas concavidades na borda superior e convexidades na parte inferior da mama. O algoritmo foi utilizado para funcionar mais rapidamente do que a transformação de Hough que é comumente usada para extração de curvas.

Pramanik et al. (2020) propuseram uma nova função para medição de intensidade local dos pixels da imagem para gerenciar a complexidade decorrente da intensidade da não-uniformidade e do ruído, chamada *multiscale spatially weighted pixel-contribution and shape-feature-embedded force* (MSPSF). O método de segmentação foi realizado em duas etapas: *energy functional* e baseado em level set, reduzindo o tempo de computação. Em outro método proposto, Pramanik et al. (2020) usaram a aproximação de arco para identificar o limite superior da mama em conjunto com um algoritmo desenvolvido para traçar a curva do limite inferior da região da mama. No entanto, em alguns casos, esse método falhou em segmentar a região da mama com precisão (ROSLIDAR et al., 2020).

Em vez do método level set, o método de segmentação recentemente proposto por Koshki et al. (2020) localiza o contorno inicial dentro do objeto desejado e converge para os pixels de contorno externos com intensidade semelhante à intensidade média. Este modelo

estendido de level set de contorno pode extrair objetos em imagens com múltiplas regiões, como termogramas de mama.

3.3 EXTRAÇÃO DE CARACTERÍSTICAS E CLASSIFICAÇÃO

A termografia apresenta algumas desvantagens tais como sensibilidade relativamente baixa para tumores profundos e pequenos, inabilidade de distinguir entre tumores e regiões com temperatura elevadas devido a inflamações, por exemplo, e divergências nas interpretações das imagens termográficas quando analisadas por diferentes especialistas. Uma maneira de contornar tais desvantagens é associar a imagem termográfica a ferramentas de aprendizagem de máquina e escolher e extrair, da melhor forma possível, as informações dos termogramas para compor o vetor de características a ser utilizado no classificador (BORCHARTT et al., 2013).

Ao final do processo de classificação tem-se quatro cenários possíveis: verdadeiro positivo (VP), onde a imagem termográfica com anormalidade é classificada corretamente como não-saudável; falso positivo (FP), onde a imagem termográfica é classificada incorretamente como não-saudável; verdadeiro negativo (VN), no qual a imagem termográfica é corretamente classificada como saudável e falso negativo (FN), no qual a imagem termográfica com anormalidade é incorretamente classificada como saudável. Os valores obtidos são combinados resultando nas medidas de acurácia, precisão, sensibilidade e especificidade, representadas pelas Equações 3.1-4 (BORCHARTT et al., 2013).

$$acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (3.1)$$

$$precisão = \frac{VP}{VP + FP} \quad (3.2)$$

$$sensibilidade = \frac{VP}{VP + FN} \quad (3.3)$$

$$especificidade = \frac{VN}{VN + FP} \quad (3.4)$$

NG et al. (2007) propuseram um diagnóstico combinando regressão linear, rede de função de base radial (RBFN - radial basis function network) e característica de operação do receptor (ROC - receiver characteristic curve) para analisar as temperaturas extraídas das

termografias. Um classificador do tipo ANN RBF obteve um aumento de 80,95% na precisão ao analisar 82 pacientes (30 pacientes normais, 48 com tumor benigno e 4 com câncer).

Qi et al (2008) desenvolveram um algoritmo de classificação assimétrica para extrair as características de termogramas: aprendizado não supervisionado com K-means e aprendizagem supervisionada com KNN. Os autores analisaram 6 termogramas normais e 18 termogramas com câncer; e foi observado que os momentos estatísticos de maior ordem (variância, distorção e curtose) indicam assimetria de forma mais eficiente comparado com os momentos de menor ordem (média e entropia) (SINGH et al., 2020).

Schaefer et al. (2009) extraíram 38 características de cada seio de 146 pacientes (29 casos malignos e 11 casos benignos) e treinaram um classificador fuzzy para classificar os tumores em malignos e benignos, atingindo uma taxa de acerto de 79,53%.

Serrano et al. (2010) extraíram 133 características de uma imagem térmica: 36 baseadas no coeficiente de Hurst e 97 baseadas na lacunaridade. Técnicas de aprendizagem de máquina como classificador Naïve-Bayes e classificador via agrupamento (clusters) apresentaram os melhores resultados (área de ROC = 0,958).

Prasad et al. (2010) extraíram características estatísticas de termogramas segmentados e utilizaram um classificador do tipo redes neurais perceptron multicamadas para classificar em normal, benigno e malignos. O classificador usou 50 termogramas como treino e 10 termogramas para teste e validação e obteve uma acurácia de 80%.

Acharya et al. (2012) extraíram características de textura das imagens termográficas e as utilizaram no classificador SVM (support vector machine), onde 36 pacientes (18 normais e 18 com tumores malignos) constituíram o grupo de treino e 14 pacientes, o grupo de teste. O sistema proposto aumentou a precisão em 88,10%, obtendo sensibilidade de 85,71% e especificidade de 90,48%.

Araújo et al. (2017) sugeriu o uso de dados intervalares para descrever uma técnica de extração de características que pudessem evidenciar a variabilidade dos dados de temperatura na região das mamas. Os autores utilizaram 50 pacientes divididos em três classes: malignos, benigno e cisto. Diferentes classificadores foram comparados na avaliação, atingindo 16% de taxa de erro global e 85% de sensibilidade para o tumor maligno.

Gaber et al. (2015) atingiram precisão de 88,41% ao classificar 63 pacientes (29 saudáveis e 24 malignos) utilizando um classificador support vector machine (SVM). A segmentação automática dos termogramas foi feita com com algoritmo Fast Fuzzy C-mean e conjuntos neutrosóficos.

Pramanik et al. (2016) propuseram um método de extração de características conhecido como Block Variance (BV), método que depende da textura local. Na fase de classificação, os autores usaram o método híbrido de uma regra de treinamento de descida de gradiente e uma rede neural de feed-forward. Eles avaliaram seu método usando 100 imagens (40 malignas e 60 benignas) de um banco de dados público. O sistema proposto é avaliado usando a abordagem de análise de assimetria. Os resultados mostraram que esse sistema é bom na precisão da classificação (menos de 0,1 taxa falso-positiva) em comparação com o trabalho relacionado (ALFAYEZ et al., 2020).

Gogoi et al. (2018) propuseram o uso de SVM para distinguir termogramas anormais dos normais. Em relação a precisão, especificidade e sensibilidade, os resultados experimentais mostraram que o método proposto alcançou 98%.

Sathish et al. (2018) realizaram um estudo comparativo entre o Ensemble Bagged Trees e o AdaBoost para a detecção de câncer de mama por termogramas. Os resultados da avaliação mostraram que o classificador Ensemble Bagged Trees é melhor que o AdaBoost. Essa avaliação foi realizada com precisão de 87%, sensibilidade de 83% e especificidade de 90,6%.

De Vasconcelos et al. (2018) calcularam 20 características baseadas nas temperaturas mínima e máxima das mamas direita e esquerda. No estudo, 233 termogramas foram utilizados na classificação binária (câncer e não-câncer) e na classificação multiclasse (maligno, benigno, cisto e normal). Além disso, o banco de dados existente foi balanceado pela construção de vetores para reduzir o viés do classificador. Um classificador baseado na Sequential Minimal Optimization (SMO) resultou nas sensibilidades de 94,73% e 80, 7% para as análises de classe binária e multiclasse, respectivamente.

Alfayes et al. (2020) compararam duas técnicas de aprendizado de máquina (MLP – perceptron multicamadas e ELM – aprendizagem de máquina extrema) para a detecção precoce do câncer de mama através de termogramas. Os resultados experimentais mostraram que a detecção de câncer de mama com base em ELM deu a melhor precisão (100%), enquanto os classificadores MLP deram apenas 82,20%. Além disso, verificou-se que o ELM é muito mais rápido que o MLP.

Brasileiro (2020) desenvolveu uma técnica de visão computacional com base em Rede Neural Convolucional para detectar o câncer de mama através de imagens termográficas. A autora utilizou uma base de dados contendo 97 amostras e obteve uma taxa de erro global de 7,5% e sensibilidade de 98,46% ao separar os indivíduos em três classes: benigno, maligno e cisto. A classificação binária da mesma base de dados encontrou uma taxa de erro global de

21,94% e 81,66% de sensibilidade ao classificar os indivíduos nas classes Câncer e Não Câncer.

3.4 ANÁLISE DE DADOS SIMBÓLICOS

Segundo Billard (2011), a análise de dados simbólicos baseia-se na ideia de que as inferências estatísticas são comumente exigidas em nível de grupo, e não individual. Portanto, os modelos de análise são construídos com base nos grupos de dados, isto é, os símbolos (BERANGER et al., 2018).

O dado intervalar é uma das maneiras mais comuns de representar os dados simbólicos, pois refletem a variabilidade e incerteza das medidas observadas. Sendo o conjunto de observações individuais $X_1, X_2, \dots, X_n \in \mathbb{R}$, o intervalo é construído como $S = (\min_i X_i, \max_i X_i) \subset \mathbb{R}$. O conjunto de dados agrupados ou intervalos $S_1, S_2, \dots, S_m \in S$ são considerados os novos pontos do processo. As inferências estatísticas serão então feitas no conjunto de dados simbólicos sem qualquer referência aos dados no âmbito individual (BERANGER et al., 2018).

Ciampi et al. (2000) introduziram uma generalização da binary decision trees para prever a associação de classe de dados simbólicos e em 2002, Rossi e Conan-Guez generalizaram multiperceptrons para trabalhar com dados intervalares.

Em 2004, de Souza et al. propuseram dois métodos de agrupamento baseados na distância de Mahalanobis para dados intervalares. No primeiro método, a distância utilizada é adaptativa e comum para todas as classes, e os protótipos são vetores de intervalos. No segundo método, cada classe tem sua própria distância e o protótipo de cada classe é composto por um vetor de intervalos e uma distância adaptativa.

Em 2006, Appice et al. introduziram uma abordagem lazy-learning que estende um algoritmo de classificação KNN (k-nearest neighbor) de distância tradicional para dados intervalares e modal, e Silva e Brito propuseram três abordagens para a análise multivariada de dados intervalares com foco na análise discriminante linear. E em 2011, de Souza et al. apresentaram quatro classificadores baseados na regressão logística, nos quais utilizaram variáveis intervalares.

Araújo et al. (2017) propuseram um classificador ponderado por classes para dados intervalares baseado numa adaptação do classificador de distância mínima sugerido por Souza et al. (2004). O classificador é baseado na distância de Mahalanobis e foi ajustado para classificar dados intervalares de temperatura, obtidos a partir de termogramas de mama, em

uma das três classes: maligno, benigno ou cisto. Os autores obtiveram taxa de erro global de 16% e sensibilidade ao tumor maligno de 93%.

Queiroz et al. (2018) generalizaram a análise discriminante para dados intervalares ao mudar o produto interno utilizado na matriz central para dados de intervalos. A abordagem proposta foi aplicada em um problema de classificação de anormalidades de temperatura da mama em relação às classes maligno e não-maligno, resultou em uma taxa de sensibilidade de 78%.

3.5 CONTRIBUIÇÃO CIENTÍFICA DO TRABALHO

O presente trabalho expande o classificador baseado em variáveis intervalares desenvolvido por Araújo (2017) de forma que a combinação de diferentes tipos de variáveis possa ser utilizada em um mesmo processo de classificação.

4 METODOLOGIA

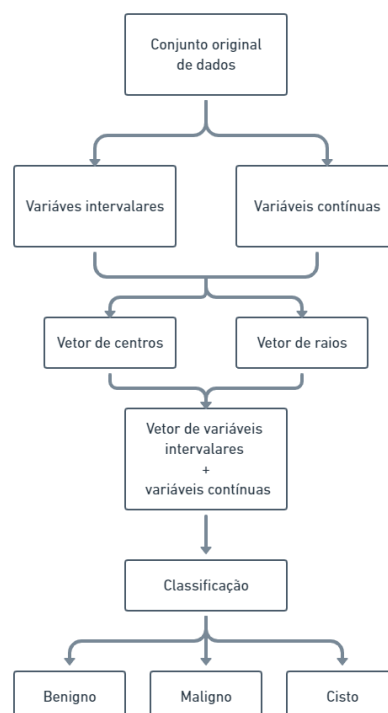
Este trabalho propõe uma modificação da proposta de extração de características sugerida por Araújo (2014), acrescentando variáveis contínuas junto ao conjunto de variáveis intervalares no processo de classificação. Os testes serão realizados a partir de uma base de dados estendida, contendo 97 pacientes que serão submetidos ao classificador original e ao classificador proposto.

A metodologia deste trabalho está dividida nas seguintes etapas: processamento do conjunto de dados original, extração de características/variáveis contínuas representativas do problema físico e classificação.

A extração de características utiliza dados intervalares contidos na base original, que foram obtidos das matrizes de temperatura de cada mama e das suas respectivas matrizes morfológicas que são geradas por meio do processamento morfológico prévio (ARAÚJO, 2014). O processo de classificação se baseará no uso de dados intervalares e dados contínuos, utilizados como vetor de entrada em um classificador de distância mínima, que usa uma adaptação da distância de Mahalanobis baseada em intervalos como função de proximidade.

Uma representação geral da metodologia pode ser vista no fluxograma apresentado na Figura 5.

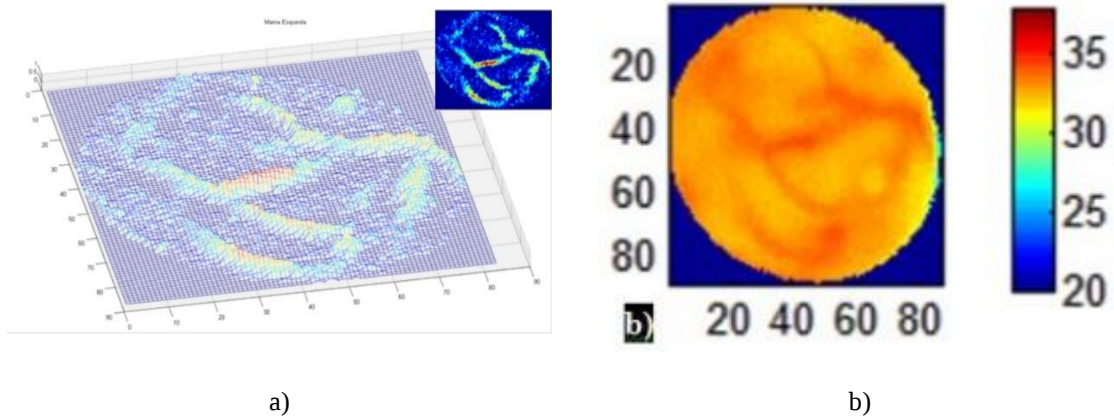
Figura 5 - Fluxograma simplificado da metodologia



4.1 BASE DE DADOS

A base utilizada neste trabalho contém dados de imagens termográficas de pacientes do Hospital das Clínicas da Universidade Federal de Pernambuco, obtidas durante o período de 2006 a 2014, utilizando-se uma câmera de infravermelho FLIS S45 (ARAÚJO, 2014; BEZERRA et al., 2013a, 2013b; OLIVEIRA, 2012; QUEIROZ et al., 2018). As pacientes das quais foram tomadas as imagens concordaram em participar da pesquisa a partir da assinatura do Termo de Consentimento Livre e Esclarecido (TCLE), documento obrigatório do projeto cadastrado no Ministério da Saúde sob o Registro CEP/CCS/UFPE N°279/05 e aprovado pelo Comitê de Ética da Universidade Federal de Pernambuco. As imagens contidas na base de dados foram previamente processadas seguindo a metodologia de Araújo (2014), gerando um conjunto inicial de dados de temperatura e dados morfológicos de cada paciente (Figura 6).

Figura 6 - a) Processamento morfológico sobre as temperaturas na mama; b) Imagem termográfica segmentada da mama



O conjunto inicial de dados corresponde às variáveis intervalares dos dados de temperatura e dados morfológicos das imagens termográficas de mama das pacientes voluntárias. O conjunto é composto por 97 indivíduos divididos em três classes – benigno, maligno e cisto – contendo 43 elementos na classe benigno, 28 elementos na classe maligno e 26 elementos na classe cisto.

Cada indivíduo na base de dados é representado pelas seguintes variáveis intervalares:

$$MD = [Tmin_{MD}, Tmax_{MD}]$$

$$ME = [Tmin_{ME}, Tmax_{ME}]$$

$$T = [\min(Tmin_{MD}, Tmin_{ME}), \max(Tmax_{MD}, Tmax_{ME})]$$

$$Morf = [\min(morf_{MD}, morf_{ME}), \max(morf_{MD}, morf_{ME})]$$

onde MD e ME representam os intervalos de temperatura obtidos para a mama esquerda e mama direita, respectivamente; T corresponde ao intervalo de temperatura envolvendo as duas mamas em conjunto; e Morf representa o intervalo correspondente aos dados morfológicos dos pacientes.

4.2 EXTRAÇÃO DE CARACTERÍSTICAS

A abordagem desenvolvida neste trabalho utiliza os intervalos de temperatura e intervalos morfológicos das imagens termográficas, que correspondem às quatro variáveis intervalares contidas no conjunto inicial de dados: MD, ME, T e Morf; em adição a variáveis contínuas que são extraídas das matrizes de temperaturas das mamas.

4.2.1 Obtenção de dados contínuos

As características contínuas usadas neste trabalho se baseiam nas medidas de dissimilaridade de Gowda-Diday para intervalos (BILLARD et al., 2006). Billard e Diday (2016) definiram três medidas de dissimilaridade para intervalos: Componente de Posição (D_1), Componente de Conteúdo (D_2) e Componente de Posição Relativa (D_3). Aqui, usaremos as medidas de posição D_1 e de conteúdo D_2 , definidas como mostrado a seguir.

Sejam dois intervalos $z = [a_z, b_z]$ e $w = [a_w, b_w]$, onde a_z e b_z , representam, respectivamente, os valores de infimum e maximum de z , e a_w e b_w correspondem aos valores de infimum e maximum de w , respectivamente. As medidas de dissimilaridade de Gowda-Diday de posição e de conteúdo entre as duas variáveis intervalares z e w são dadas por:

$$D_1(z, w) = \frac{||b_z - a_z| - |b_w - a_w||}{K} \quad (4.1)$$

onde D_1 representa o componente de posição entre os dois intervalos e K corresponde ao comprimento total compreendido por z e w , dado por:

$$K = |\max(b_z, b_w) - \min(a_z, a_w)| \quad (4.2)$$

O componente D_2 representa o componente de conteúdo entre os dois intervalos e é dado por:

$$D_2(z, w) = \frac{|b_z - a_z| + |b_w - a_w| - 2\gamma}{K} \quad (4.3)$$

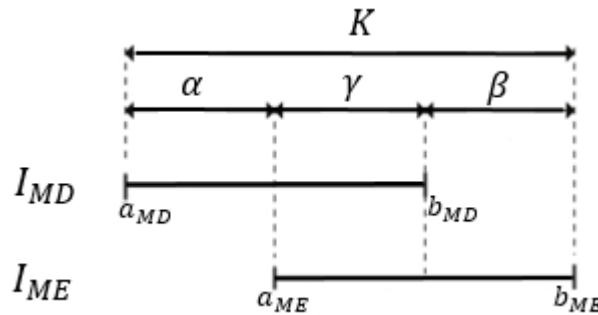
onde γ corresponde ao comprimento da interseção entre os intervalos z e w , dado por:

$$\gamma = |\max(a_z, a_w) - \min(b_z, b_w)| \quad (4.4)$$

Associando os intervalos de temperatura das mamas às medidas de dissimilaridade de Diday:

Sejam $I_{MD} = [a_{MD}, b_{MD}]$ e $I_{ME} = [a_{ME}, b_{ME}]$ os intervalos de temperaturas das mamas direita e esquerda obtidos de uma imagem termográfica de mama, respectivamente. Podemos representar graficamente as variáveis intervalares I_{MD} e I_{ME} ao longo de um plano, como mostrado na Figura 7.

Figura 7 - Representação gráfica dos intervalos de temperatura para a mama direita e mama esquerda



Fonte: ARAÚJO, 2014

Da Figura 7, pode-se definir as distâncias α , γ e β como:

$$\alpha = |a_{ME} - a_{MD}|$$

$$\gamma = |b_{MD} - a_{ME}|$$

$$\beta = |b_{ME} - b_{MD}|$$

Dessa forma, os intervalos I_{MD} e I_{ME} , referentes aos dados de temperaturas para a mama direita e mama esquerda, respectivamente, podem ser representados por

$$I_{MD} = |b_{MD} - a_{MD}| = \alpha + \gamma \quad (4.5)$$

$$I_{ME} = |b_{ME} - a_{ME}| = \beta + \gamma \quad (4.6)$$

onde γ representa o comprimento da interseção entre os dois intervalos.

Assim, os componentes D_1 e D_2 da medida de dissimilaridade de Gowda-Diday, que servirão como base para as características contínuas, podem ser obtidas das medidas α , β e K :

$$D_1(I_{MD}, I_{ME}) = \frac{||b_{MD} - a_{MD}| - |b_{ME} - a_{ME}||}{K} = \frac{(\alpha + \gamma) - (\beta + \gamma)}{K}$$

$$D_1(I_{MD}, I_{ME}) = \frac{\alpha - \beta}{K} \quad (4.7)$$

$$D_2(I_{MD}, I_{ME}) = \frac{||b_{MD} - a_{MD}| + |b_{ME} - a_{ME}| - 2\gamma|}{K} = \frac{(\alpha + \gamma) + (\beta + \gamma) - 2\gamma}{K}$$

$$D_2(I_{MD}, I_{ME}) = \frac{\alpha + \beta}{K} \quad (4.8)$$

Dessa forma, para cada paciente, calcula-se os valores de α , β , K e β_M , onde β_M corresponde à variável β medida sobre os dados morfológicos (ARAÚJO, 2014), obtendo-se um vetor de características contínuas $w = (\alpha, \beta, K, \beta_M)$. Esses dados contínuos, junto com outras variáveis - como temperatura máxima de cada mama, média de temperaturas das mamas e média entre as temperaturas máximas de ambas as mamas - resultará em um conjunto de dados contínuos que será adicionado ao vetor de intervalos original $y = (MD, ME, T, Morf)$, explanados na seção 4.1. O conjunto final de variáveis compõe os dados de entrada do processo de classificação.

Para agregar os dados contínuos aos dados de intervalo é necessário representar as variáveis contínuas como variáveis intervalares. Segundo Jaulin et al. (2001), qualquer número real pode ser representado com um intervalo degenerado e vice-versa. Para tal, o número real a pode ser representando como o intervalo degenerado $Y = [a, a]$.

Por fim, obtém-se um vetor de doze dimensões definido como:

$$\dot{y} = (MD, ME, T, Morf, e, f, g, h, i, j, k, l) \quad (4.11)$$

onde:

$$\begin{aligned} MD &= [Tmin_{MD}, Tmax_{MD}]; & g &= [média_{Tmáx}, média_{Tmáx}]; \\ ME &= [Tmin_{ME}, Tmax_{ME}]; & h &= [média_{média}, média_{média}]; \\ T &= [\min(Tmin_{MD}, Tmin_{ME}), \max(Tmax_{MD}, Tmax_{ME})]; & i &= [\alpha, \alpha]; \\ d &= [\min\{\max(LTD_{MD}, LTD_{ME})\}, \max\{\max(LTD_{MD}, LTD_{ME})\}]; & j &= [\beta, \beta]; \\ e &= [Tmáx_{MD}, Tmáx_{MD}]; & k &= [K, K]; \\ f &= [Tmáx_{ME}, Tmáx_{ME}]; & l &= [\beta_M, \beta_M]. \end{aligned}$$

4.2.2 Etapa de classificação

O processo de classificação utilizado nessa metodologia segue o modelo definido por Araújo (2017), baseado em um classificador de distância mínima utilizando a distância de Mahalanobis parametrizada para intervalos.

Para a etapa de classificação, cada um dos intervalos é separado em suas componentes de centro e raio, valores escalares denotados, respectivamente, por:

$$centro = \frac{a + b}{2} \quad \quad \quad raio = \frac{|a - b|}{2}$$

onde a e b representam os limites inferior e superior, respectivamente, de uma observação intervalar $\dot{y} = [a, b]$.

Essa etapa vai gerar dois novos vetores contínuos: o vetor de centros, w_c , contendo os centros dos intervalos e o vetor de raios, w_r , contendo os raios dos intervalos.

Em seguida, esses dois vetores contínuos são transformados para um novo espaço de características a partir do critério de Fisher, que tem a função de reduzir a dimensionalidade do vetor de entrada (ARAÚJO et al., 2017). Após a transformação de Fisher, os dados são classificados por meio da adaptação do classificador intervalar de Mahalanobis utilizado em Araújo et al. (2017).

4.3 CLASSIFICAÇÃO

Um classificador de distância mínima para dados intervalares foi utilizado para a classificação. Seja $\Gamma = \dot{y}_1, \dots, \dot{y}_m$ um grupo de m objetos simbólicos descritos por p variáveis intervalares, onde cada objeto $\dot{y}_i (i = 1, \dots, m)$ é representando por um vetor de intervalos $\dot{y}_i = ([a_{i1}, b_{i1}], \dots, [a_{ip}, b_{ip}])^T$. O grupo Γ é dividido em P partições que contêm K classes ($C_1, \dots, C_K$), sendo que cada classe $C_k (k = 1, \dots, K)$ possui um protótipo L_k que também é representando por um vetor de intervalos $\dot{g}_k = ([\alpha_{k1}, \beta_{k1}], \dots, [\alpha_{kp}, \beta_{kp}])^T$, onde g_k corresponde ao vetor de médias dos intervalos de C_k . Logo, para $j = 1, \dots, p$:

$$\left[\alpha_{kj} = \frac{\sum_{i \in C_k} a_{ij}}{n_k}, \beta_{kj} = \frac{\sum_{i \in C_k} b_{ij}}{n_k} \right] \quad (4.10)$$

onde n_k representa o número de elementos na classe C_k .

O classificador de distância mínima aloca um objeto $\dot{v}_i$ para a classe C_k se

$$d^\lambda(\dot{v}_i, \dot{g}_k) \leq d^\lambda(\dot{v}_i, \dot{g}_r) \quad \forall r, r = 1, \dots, K \quad (4.11)$$

Neste trabalho, a função $d^\lambda(\dot{v}_i, \dot{g}_k)$ corresponde a uma distância de Mahalanobis parametrizada para intervalos, descrita na seção a seguir.

4.3.1 Distância de Mahalanobis parametrizada para intervalos

Seja $v_{iL} = (a_{i1}, \dots, a_{ip})^T$ e $v_{iS} = (b_{i1}, \dots, b_{ip})^T$ dois vetores contendo os valores inferiores e superiores de v_i , respectivamente; e $g_{kL} = (\alpha_{k1}, \dots, \alpha_{kp})^T$ e $g_{kS} = (\beta_{k1}, \dots, \beta_{kp})^T$ dois vetores contendo os valores inferiores e superiores de g_k , respectivamente.

A distância de Mahalanobis $\delta(\dot{v}_i, \dot{g}_k)$ entre duas observações intervalares $\dot{v}_i$ e $\dot{g}_k$ foi definida por Souza et al (2004) como:

$$\delta(\dot{v}_i, \dot{g}_k) = d(v_{iL}, g_{kL}) + d(v_{iS}, g_{kS}) \quad (4.12)$$

onde

$$d(v_{il}, g_{kl}) = (v_{il} - g_{kl})^T S_I (v_{il} - g_{kl}) \quad (4.13)$$

representa a distância de Mahalanobis entre os vetores dos valores inferiores, v_{il} e g_{kl} , e S_I corresponde à matriz de covariância para os valores inferiores; e

$$d(v_{is}, g_{ks}) = (v_{is} - g_{ks})^T S_S (v_{is} - g_{ks}) \quad (4.14)$$

representa a distância de Mahalanobis entre os vetores dos valores superiores, v_{is} e g_{ks} , e S_S corresponde à matriz de covariância para os valores superiores.

Segundo Araújo (2017), a distância de Mahalanobis parametrizada pode ser definida através de uma matriz de covariâncias comum entre os valores inferiores e os valores superiores, parametrizada para cada classe. O parâmetro de controle para a classe C_k $\lambda_k \in [0,1]$, $(k = 1, \dots, K)$ mede o grau de relevância entre os limites dos intervalos, isto é, entre os valores inferiores e os valores superiores, no cálculo da matriz de covariâncias parametrizada da classe.

A distância de Mahalanobis parametrizada entre os vetores intervalares $\dot{v}_i$ e $\dot{g}_k$ é definida da seguinte forma:

$$d^\lambda(\dot{v}_i, \dot{g}_k) = (v_{il} - g_{kl})^T S_k(\lambda_k)(v_{il} - g_{kl}) + (v_{is} - g_{ks})^T S_k(\lambda_k)(v_{is} - g_{ks}) \quad (4.15)$$

Onde $S_k(\lambda_k)$, $(k = 1, \dots, K)$ representa a matriz de covariâncias parametrizada a classe C_k , definida com base em um parâmetro de controle $\lambda_k \in [0,1]$.

Neste trabalho, são utilizadas e comparadas três métricas para o cálculo das matrizes de covariâncias parametrizadas da classe C_k . São elas:

$$S_k^1(\lambda_k) = \frac{(1 - \lambda_k)S_{kl} + \lambda_k S_{ks}}{(1 - \lambda_k)n_k + \lambda_k n} \quad (4.16a)$$

$$S_k^2(\lambda_k) = \frac{(1 - \lambda_k)S_{kl} + \lambda_k S_{ks}}{\lambda_k n} \quad (4.16b)$$

$$S_k^3(\lambda_k) = (1 - \lambda_k)S_{kI} + \lambda_k S_{kS} \quad (4.16c)$$

onde λ_k representa o parâmetro de controle da classe C_k ; S_{kI} corresponde à matriz de covariância para os valores de infimum dos intervalos da classe C_k ; e S_{kS} corresponde à matriz de covariância para os valores de maximum dos intervalos da classe C_k . Cada métrica se difere entre si no que diz respeito ao peso aplicado em relação ao número de indivíduos da amostra e/ou número de indivíduos de cada classe. Assim, a Equação 4.16a é ponderada em função do número de indivíduos da base e do número de indivíduos da classe. A Equação 4.16b é ponderada apenas em função do número de indivíduos da base e a Equação 4.16c não sofre a influência do número de indivíduos.

Pode-se observar que para as variáveis contínuas, da forma $Y = [a, a]$, onde o valor de infimum é igual ao valor do maximum, teremos $S_{kI} = S_{kS} = S_k$ e:

$$S_k^1(\lambda_k) = \frac{(1 - \lambda_k)S_k + \lambda_k S_k}{(1 - \lambda_k)n_k + \lambda_k n} = \frac{S_k}{(1 - \lambda_k)n_k + \lambda_k n} \quad (4.17a)$$

$$S_k^2(\lambda_k) = \frac{(1 - \lambda_k)S_k + \lambda_k S_k}{\lambda_k n} = \frac{S_k}{\lambda_k n} \quad (4.17b)$$

$$S_k^3(\lambda_k) = (1 - \lambda_k)S_k + \lambda_k S_k = S_k \quad (4.17c)$$

Assim, para as variáveis contínuas os numeradores das Equações 4.16a, 4.16b e 4.16c se resumem a S_k , que corresponde à própria matriz de covariância da variável contínua e o mesmo classificador pode ser aplicado sem perda de generalidade entre dados intervalares e dados contínuos.

O parâmetro λ_k representa a influência das matrizes de covariância dos valores inferiores e dos valores superiores no cálculo da matriz de covariâncias comum. Então, para a classe C_k , se:

- $\lambda_k = 0$, a matriz de covariâncias comum será representada apenas pela matriz de covariâncias dos valores inferiores;
- $\lambda_k = 1$, a matriz de covariâncias comum será representada apenas pela matriz dos valores superiores;

- $0 < \lambda_k < 1$, a matriz de covariâncias comum será uma representada por uma combinação das matrizes de covariância dos valores inferiores e dos valores superiores.

O valor de cada parâmetro λ_k será variado no intervalo de 0 a 1, com passo de 0,1, para testar cada combinação possível dos três parâmetros. As iterações sucessivas tiveram como objetivo encontrar os valores ótimos dos parâmetros de λ_k , ou seja, determinar as condições que resultam na menor taxa de erro global de classificação combinado com os maiores índices de sensibilidade e de especificidade para as três classes.

A taxa de erro global de classificação leva em consideração as amostras do tipo falso negativo (FN), ou seja, os indivíduos doentes erroneamente classificados como saudáveis, e é calculada da seguinte forma:

$$taxa\ de\ erro = \frac{1}{n} \sum_{i=1}^C FN_i \quad (4.18)$$

onde:

n = número total de amostras

C = número total de classes

FN_i = falso negativo para a classe Ω_i

O índice de sensibilidade refere-se à capacidade do classificador em identificar corretamente um indivíduo verdadeiro positivo dentre aqueles que possuem a doença, avaliando o quão sensível o classificador é. A sensibilidade é calculada da forma:

$$Sens(i) = \frac{VP_i}{VP_i + FN_i} \quad (4.19)$$

onde:

VP_i = verdadeiro positivo para a classe Ω_i

FN_i = falso negativo para a classe Ω_i

O índice de especificidade refere-se à capacidade do classificador em identificar corretamente um indivíduo verdadeiro negativo dentre aqueles que são saudáveis, ou seja, avalia o quão específico é o classificador. A especificidade de cada classe é calculada da seguinte forma:

$$Espec(i) = \frac{VN_i}{VN_i + FP_i} \quad (4.20)$$

onde:

VN_i = verdadeiro negativo para a classe Ω_i

FP_i = falso positivo para a classe Ω_i

5 RESULTADOS

Este capítulo apresenta os resultados obtidos pela metodologia proposta no capítulo anterior. A avaliação é feita em relação à taxa de erro global de classificação e aos índices de sensibilidade (Sens) e de especificidade (Espec) para cada classe. São computadas duas análises, a primeira considerando a classificação dos indivíduos nas três classes (maligno, benigno e cisto); e a segunda análise classifica o conjunto de dados considerando duas classes: câncer e não-câncer. Para tal, as classes benigno e cisto são agrupadas como uma única classe.

Durante o processo de classificação, cada indivíduo foi utilizado como grupo de teste, enquanto o restante do conjunto foi utilizado como grupo de treino. Esse procedimento caracteriza o critério de validação cruzada *Leave-One-Out*, técnica aplicada quando a amostra analisada é pequena, dificultando a divisão do conjunto em grupo de teste e grupo de treino. A validação repete-se de forma que todos os indivíduos atuem como amostra de teste.

5.1 RESULTADOS DA METODOLOGIA PROPOSTA

Os resultados serão comparados com a metodologia de Araújo (2017), considerando apenas variáveis intervalares e com os resultados obtidos por Brasileiro (2020) usando alguns classificadores da literatura para dados intervalares.

5.1.1 Comparação com Araújo (2014) na base original

Como forma inicial de avaliação, a metodologia proposta foi aplicada à mesma base usada por Araújo (2014) contendo 50 amostras, das quais 14 contêm anomalias malignas, 19 são casos benignos e 17 são casos de cistos. Os resultados desta metodologia obtiveram os mesmos valores encontrados por Araújo (2017), não indicando ganho ou perda ao acrescentar variáveis contínuas ao processo de classificação intervalar. O classificador proposto, assim como o original, encontrou uma taxa de erro global de classificação de 16%, ou seja, classificou corretamente 84% do número total de amostras. Em relação à classe maligno, a metodologia proposta conseguiu classificar corretamente 93% das amostras que contêm anormalidades malignas, valor representado pelo índice de sensibilidade.

O valor de cada parâmetro λ_k foi variado no intervalo de 0 a 1, com passo de 0,1, para testar cada combinação possível dos três parâmetros, resultando em 1331 iterações no total,

com o objetivo de encontrar a combinação que apresente o melhor conjunto de resultados que possua a menor taxa de erro global de classificação e os maiores índices de sensibilidade para as três classes.

Os resultados completos encontram-se sintetizados na Tabela 1.

Tabela 1 - Resumo dos resultados de classificação com base de dados contendo 50 amostras

Classificador	Taxa de erro	Sens. Maligno	Sens. Benigno	Sens. Cisto	λ_1	λ_2	λ_3
Araújo (2014)	0,16	0,93	0,84	0,76	0,3	0,4	0
Proposto (4.16a)	0,16	0,93	0,84	0,76	0	0,5	0,3
Proposto (4.16b)	0,16	0,93	0,84	0,76	0,4	0,6	0,4
Proposto (4.16c)	0,16	0,93	0,84	0,76	0,7	1	0,7

De acordo com a Tabela 1, observa-se que não houve divergências entre os resultados alcançados pelo classificador proposto e os resultados alcançados pelo classificador de Araújo (2014) na base utilizada por Araújo (2014).

5.1.2 Aplicação da metodologia à base estendida

A avaliação seguinte representa a aplicação da metodologia proposta na base estendida. A base de dados estendida contém 97 indivíduos agrupados em três classes:

- Maligno, composta por 28 indivíduos;
- Benigno, composta por 43 indivíduos;
- Cisto, composta por 26 indivíduos.

A Figura 8 exibe os histogramas referentes aos erros de classificação geral obtidos dos resultados encontrados a partir das Equações 4.16a, 4.16b e 4.16c, respectivamente. Observando os histogramas apresentados, verifica-se que, nos três casos, as taxas de erro mínimas atingidas encontram-se em torno de 45%, ou seja, o classificador proposto classificou corretamente 55% das amostras. Nas Tabelas 2, 3 e 4 encontram-se todas as observações que atingiram o menor erro, das quais observa-se que a menor taxa, 45,36%, combinada com os maiores índices de sensibilidade para as classes maligno (64,29%), benigno (53,49%) e cisto (46,15%), foi atingida na classificação utilizando a matriz de covariância parametrizada descrita pela Equação 4.16a. Esses valores foram alcançados com a combinação de parâmetros $\lambda_1 = 0,2$, $\lambda_2 = 1,0$ e $\lambda_3 = 0,9$.

Figura 8 - Histogramas do erro de classificação geral para a base estendida. a) considerando $\mathbf{S}_k^1(\lambda_k)$ como matriz de covariâncias; b) considerando $\mathbf{S}_k^2(\lambda_k)$ como matriz de covariâncias; c) considerando $\mathbf{S}_k^3(\lambda_k)$ como matriz de covariâncias.

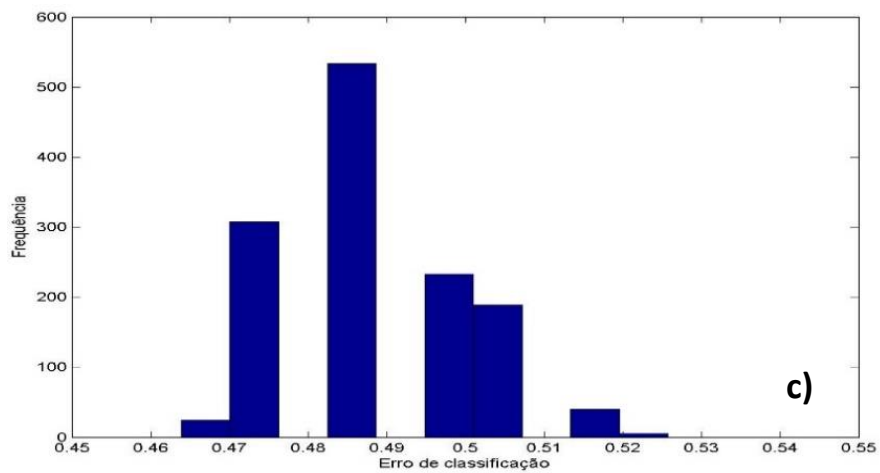
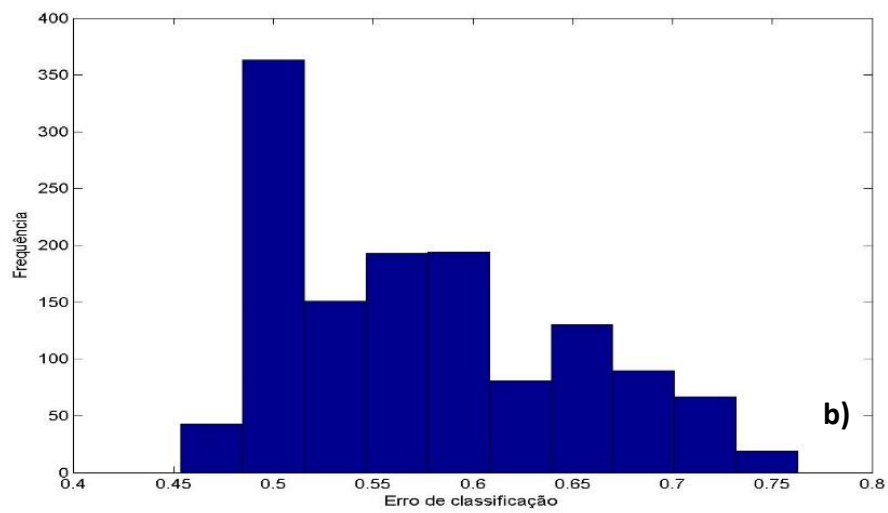
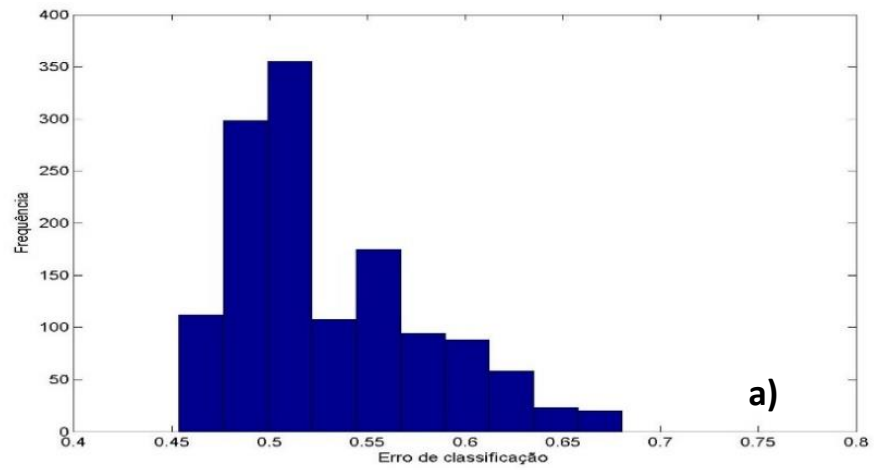


Tabela 2 - Observações que atingiram a menor taxa de erro geral – Equação 4.16a

Erro	Número de acertos	Sens. Maligno	Espec. Maligno	Sens. Benigno	Espec. Benigno	Sens. Cisto	Espec. Cisto	λ_1	λ_2	λ_3
0,45361	53	0,64286	0,84	0,44186	0,79167	0,61538	0,69014	0	0,6	0,3
0,45361	53	0,64286	0,85333	0,44186	0,77083	0,61538	0,69014	0	0,7	0,3
0,45361	53	0,64286	0,84932	0,51163	0,68	0,5	0,76056	0	0,7	0,6
0,45361	53	0,60714	0,86301	0,53488	0,66	0,5	0,76056	0	0,9	0,6
0,45361	53	0,64286	0,84	0,44186	0,79167	0,61538	0,69014	0,1	0,7	0,4
0,45361	53	0,64286	0,85333	0,44186	0,77083	0,61538	0,69014	0,1	0,8	0,4
0,45361	53	0,64286	0,84932	0,51163	0,68	0,5	0,76056	0,1	0,8	0,7
0,45361	53	0,64286	0,85333	0,4186	0,79167	0,65385	0,67606	0,2	0,9	0,4
0,45361	53	0,64286	0,85333	0,44186	0,77083	0,61538	0,69014	0,2	0,9	0,5
0,45361	53	0,64286	0,84932	0,51163	0,68	0,5	0,76056	0,2	0,9	0,8
0,45361	53	0,64286	0,85333	0,4186	0,79167	0,65385	0,67606	0,2	1	0,4
0,45361	53	0,64286	0,85333	0,44186	0,77083	0,61538	0,69014	0,2	1	0,5
0,45361	53	0,64286	0,84932	0,51163	0,68	0,5	0,76056	0,2	1	0,8
0,45361	53	0,64286	0,84932	0,53488	0,66	0,46154	0,77465	0,2	1	0,9
0,45361	53	0,64286	0,83784	0,44186	0,77551	0,61538	0,70423	0,3	1	0,6
0,45361	53	0,64286	0,83562	0,51163	0,7	0,5	0,76056	0,3	1	0,9

Tabela 3 - Observações que atingiram a menor taxa de erro geral – Equação 4.16b

Erro	Número de acertos	Sens. Maligno	Espec. Maligno	Sens. Benigno	Espec. Benigno	Sens. Cisto	Espec. Cisto	λ_1	λ_2	λ_3
0,45361	53	0,64286	0,84932	0,51163	0,68	0,5	0,76056	0,4	0,7	0,6
0,45361	53	0,64286	0,85135	0,44186	0,7551	0,61538	0,70423	0,6	1	0,7
0,45361	53	0,64286	0,84932	0,51163	0,68	0,5	0,76056	0,6	1	0,9

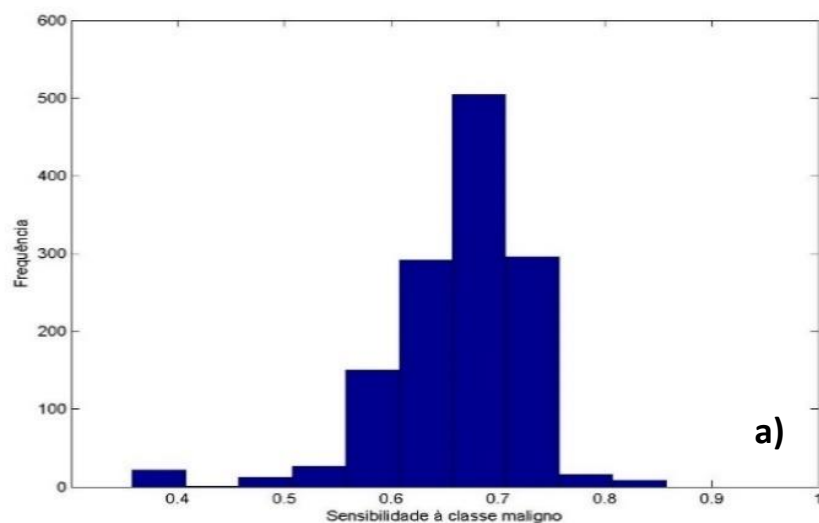
Tabela 4 - Observações que atingiram a menor taxa de erro geral – Equação 4.16c

Erro	Número de acertos	Sens. Maligno	Espec. Maligno	Sens. Benigno	Espec. Benigno	Sens. Cisto	Espec. Cisto	λ_1	λ_2	λ_3
0,46392	52	0,64286	0,81333	0,39535	0,8125	0,65385	0,69014	0	0	0,6
0,46392	52	0,64286	0,82667	0,39535	0,8125	0,65385	0,67606	0,1	0,1	0
0,46392	52	0,64286	0,82667	0,39535	0,8125	0,65385	0,67606	0,1	0,2	0
0,46392	52	0,67857	0,78667	0,37209	0,85417	0,65385	0,69014	0,5	0	0
0,46392	52	0,67857	0,78667	0,37209	0,85417	0,65385	0,69014	0,5	0,1	0
0,46392	52	0,67857	0,78667	0,37209	0,85417	0,65385	0,69014	0,5	0,2	0
0,46392	52	0,67857	0,78667	0,37209	0,85417	0,65385	0,69014	0,5	0,3	0
0,46392	52	0,67857	0,8	0,39535	0,83333	0,61538	0,69014	0,5	0,6	0
0,46392	52	0,67857	0,81333	0,37209	0,83333	0,65385	0,67606	0,5	0,6	0,2
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,8	0	0
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,8	0,1	0
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,8	0,2	0

0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,8	0,3	0
0,46392	52	0,67857	0,78378	0,44186	0,79592	0,53846	0,73239	0,8	0,9	0
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,9	0	0
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,9	0,1	0
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,9	0,2	0
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,9	0,3	0
0,46392	52	0,67857	0,78378	0,44186	0,79592	0,53846	0,73239	0,9	1	0
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	1	0,1	0,1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	1	0,2	0,1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	1	0,3	0,1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	1	0,4	0,1
0,46392	52	0,67857	0,78378	0,44186	0,79592	0,53846	0,73239	1	1	0,1

Analisando os histogramas da Figura 9, observa-se que um conjunto de resultados da classificação utilizando a matriz de covariância 4.16b obteve 100% de sensibilidade à classe maligno. Esses resultados estão listados na Tabela 5 e nela pode-se observar que, apesar do classificador ter identificado corretamente todas as amostras da classe maligno, não identificou nenhuma amostra das classes benigno e cisto. Alocar todos os indivíduos numa única classe não é o resultado esperado para um classificador, resultando numa elevada taxa do erro de classificação geral (71,13% de erro) e na aparente boa taxa de sensibilidade da classe maligno.

Figura 9 - Histogramas da sensibilidade para a classe maligno para a base estendida. a) considerando $S_k^1(\lambda_k)$ como matriz de covariâncias; b) considerando $S_k^2(\lambda_k)$ como matriz de covariâncias; c) considerando $S_k^3(\lambda_k)$ como matriz de covariâncias.



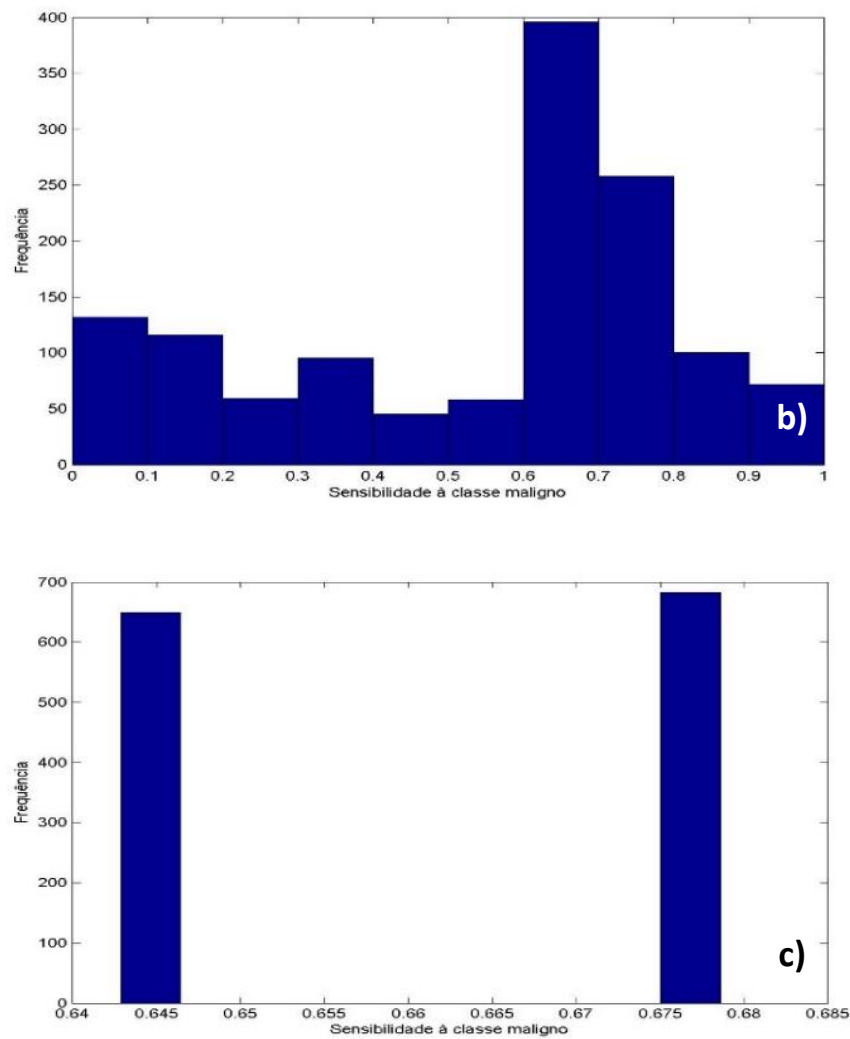


Tabela 5 - Observações que atingiram o maior índice de sensibilidade para classe Maligno – Equação 4.16b

Erro	Número de acertos	Sens. Maligno	Espec. Maligno	Sens. Benigno	Espec. Benigno	Sens. Cisto	Espec. Cisto	λ_1	λ_2	λ_3
0,71134	28	1	0	0	1	0	1	0	0,1	0
0,71134	28	1	0	0	1	0	1	0	0,2	0
0,71134	28	1	0	0	1	0	1	0	0,3	0
0,71134	28	1	0	0	1	0	1	0	0,4	0
0,71134	28	1	0	0	1	0	1	0	0,5	0
0,71134	28	1	0	0	1	0	1	0	0,6	0
0,71134	28	1	0	0	1	0	1	0	0,7	0
0,71134	28	1	0	0	1	0	1	0	0,8	0
0,71134	28	1	0	0	1	0	1	0	0,9	0
0,71134	28	1	0	0	1	0	1	0	1	0

Nos processos de classificação utilizando as matrizes de covariância 4.16a e 4.16c, não houveram resultados que atingiram 100% de sensibilidade à classe maligno; os índices

máximos atingidos foram 85,71% e 67,86%, respectivamente, e os resultados completos para cada caso estão listados nos Apêndices A e B. Na Tabela 6 encontra-se o resumo das melhores combinações de resultados para os três casos, considerando o maior índice de sensibilidade à classe maligno.

Tabela 6 - Resumo de resultados para a metodologia proposta

Classificador	Taxa de erro	Sens. Maligno	Sens. Benigno	Sens. Cisto	λ_1	λ_2	λ_3
$S_k^1(\lambda_k)$	0,6495	0,8571	0,0930	0,2301	1	0	1
$S_k^2(\lambda_k)$	0,7113	1	0	0	0	0,1	0
$S_k^3(\lambda_k)$	0,4639	0,6786	0,4419	0,5385	0,8	0,9	0

Na Tabela 7 estão exibidos os resultados obtidos a partir da metodologia original de Araújo (2014).

Tabela 7 - Resultados de classificação de Araújo (2017) com a base de dados estendida.

Classificador	Taxa de erro	Sens. Maligno	Sens. Benigno	Sens. Cisto	λ_1	λ_2	λ_3
Araújo (2017)	0,5260	0,8880	0,2560	0,3850	0,5	0,2	0,9

Os resultados relacionados à base estendida apontam para um grau de dificuldade maior para os ambos os classificadores (proposto e Araújo (2017)), uma vez que se obteve na base original de 50 indivíduos um desempenho melhor e idêntico cada um dos classificadores. Comparando-se os resultados descritos nas Tabelas 6 e 7 sobre as duas abordagens, pode-se observar que a metodologia de Araújo (2017) apresentou menor taxa de erros global com relação à metodologia proposta para $S_k^1(\lambda_k)$ e $S_k^2(\lambda_k)$. A metodologia proposta sob $S_k^3(\lambda_k)$ apresentou resultados mais equilibrados para as três classes, com menor taxa de erro global mas, no entanto, menor sensibilidade à classe maligno com relação à metodologia adotada por Araújo (2014). Pode-se destacar ainda a maior sensibilidade às classes benigno e cisto a partir da metodologia proposta sob $S_k^3(\lambda_k)$.

Outros classificadores intervalares clássicos aplicados a esta mesma base estendida foram investigados por Brasileiro (2020). Em seu trabalho, Brasileiro (2020) fez uma comparação entre o uso de técnicas de Deep Learning e classificadores tradicionais como o Mapa Multilayer Perceptron (MLP), *Support Vector Classification* (SVC) e modelos de árvore de decisão Random Forest. Os resultados obtidos para esta base foram avaliados em

função da taxa de erro global e da sensibilidade à classe maligno, avaliado para as três classes (benigno, maligno e cisto) e também para duas classes (classificação binária: câncer × não câncer). Os resultados descritos para os classificadores avaliados sob a ótica de três classes estão na Tabela 8. Pode-se observar que o classificador Random Forest foi o que obteve menor taxa de erro global, porém, ainda acima de 50% de erro que aponta para uma baixa sensibilidade às outras classes avaliadas (benigno e cisto). O classificador SVC acertou 100% dos indivíduos da classe maligno com alta taxa de erros global, indicando que pode ter alocado majoritariamente quase a totalidade dos indivíduos na mesma classe, como aconteceu com o modelo proposto sob a matriz de covariâncias $S_k^2(\lambda_k)$. O classificador MLP, por sua vez, não foi capaz de classificar corretamente nenhum indivíduo da classe maligno.

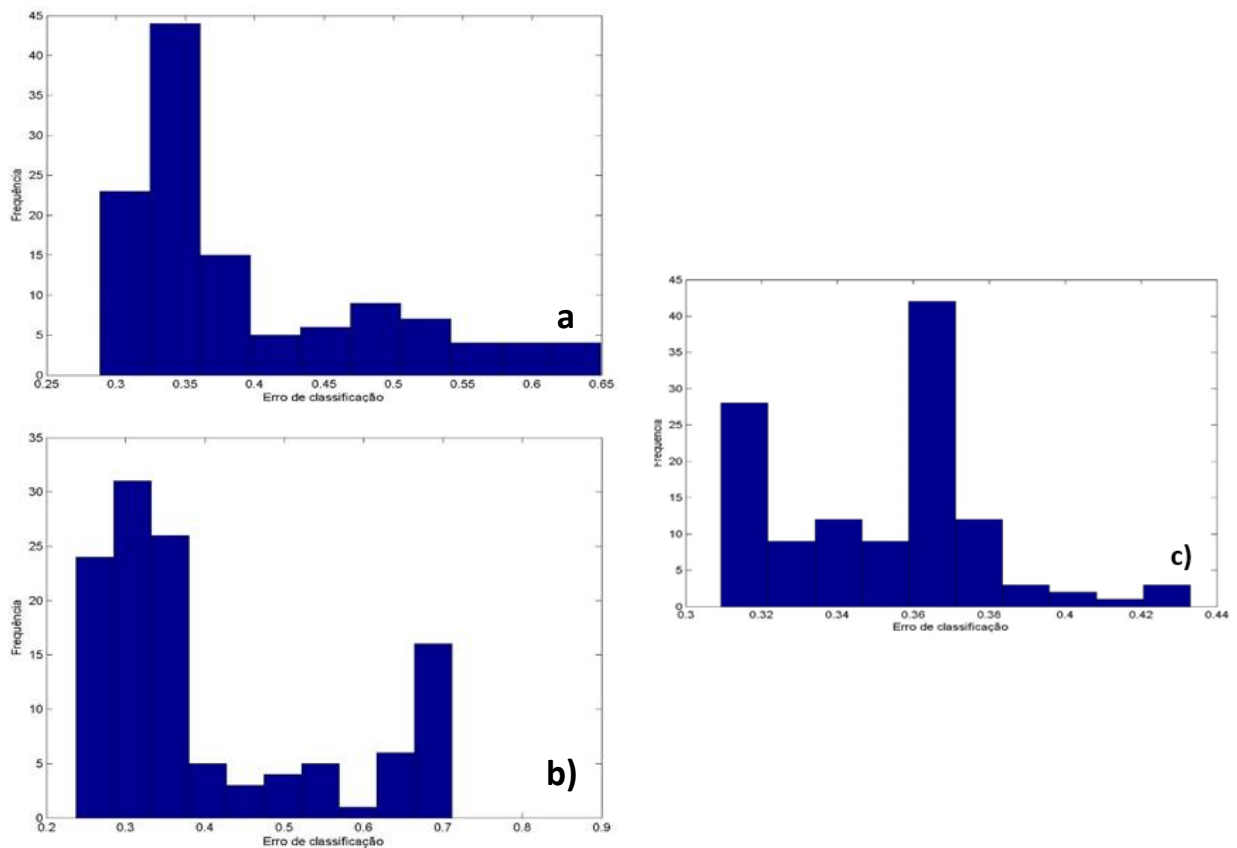
Tabela 8 - Taxa de erro global e Sensibilidade da Classe Maligno para modelos clássicos da literatura para classificação com a base estendida para 3 classes (Brasileiro, 2020).

Classificador	Taxa de erro	Sens. Maligno
MLP	0,6	0
SVC	0,7667	1
Random Forest	0,5667	0,7143

5.2 CLASSIFICAÇÃO BINÁRIA

Um estudo complementar foi realizado ao ajustar o classificador proposto para classificar os indivíduos em duas classes: câncer, contendo as amostras da classe maligno; e não-câncer, contendo as amostras das classes benigno e cisto. Os histogramas da taxa de erro geral de classificação podem ser vistos na Figura 10 para a metodologia proposta. Nela observa-se que, para os três casos, a menor taxa de erro encontra-se abaixo de 30%, ou seja, o classificador proposto classificou de forma correta em torno de 70% das amostras. Na Tabelas 9-11 encontram-se os resultados para os casos utilizando as três matrizes de covariância parametrizadas descritas pelas Equações 4.16, incluindo os índices de sensibilidade para as classes câncer e não-câncer.

Figura 10 - Histogramas do erro de classificação geral. a) considerando $S_k^1(\lambda_k)$ como matriz de covariâncias; b) considerando $S_k^2(\lambda_k)$ como matriz de covariâncias; c) considerando $S_k^3(\lambda_k)$ como matriz de covariâncias.



A menor taxa de erro geral alcançada – 23,7% –, entre os três casos, foi atingida no classificador que utilizou a matriz de covariância parametrizada $S_k^2(\lambda_k)$, como visto na Tabela 10 e representa uma redução de aproximadamente 21,6 pontos percentuais em relação ao problema com três classes e 36,8% menor que o erro encontrado por Araújo (2017), cujos resultados encontram-se dispostos na Tabela 12. Entretanto, o índice de sensibilidade para a classe câncer, para o mesmo caso, foi de 28,7%; ou seja, menos de um terço dos indivíduos com anormalidades malignas foi classificado corretamente. Por fim, 95,6% dos indivíduos que apresentam anormalidades benignas ou cistos foram classificados corretamente na classe não-câncer.

Por sua vez, o classificador com matriz de covariâncias $S_k^3(\lambda_k)$ foi o que obteve maior taxa de erro global (30,93%) mas foi também, dentre os três, o que apresentou melhor sensibilidade à classe maligno, com 64,28% de sensibilidade, e boa sensibilidade à classe não câncer, com 71% de sensibilidade para esta classe.

Tabela 9 - Observações que atingiram a menor taxa de erro geral – Equação 4.16a

Erro	Número de acertos	Sens. Câncer	Espec. Câncer	Sens. Não-Câncer	Espec. Não-Câncer	λ_1	λ_2
0,28866	69	0,5	0,7971	0,7971	0,5	0	1

Tabela 10 - Observações que atingiram a menor taxa de erro geral – Equação 4.16b

Erro	Número de acertos	Sens. Câncer	Espec. Câncer	Sens. Não-Câncer	Espec. Não-Câncer	λ_1	λ_2
0,23711	74	0,28571	0,95652	0,95652	0,28571	0,1	0,4

Tabela 11 - Observações que atingiram a menor taxa de erro geral – Equação 4.16c

Erro	Número de acertos	Sens. Câncer	Espec. Câncer	Sens. Não-Câncer	Espec. Não-Câncer	λ_1	λ_2
0,30928	67	0,64286	0,71014	0,71014	0,64286	0	0,9
0,30928	67	0,64286	0,71014	0,71014	0,64286	0,1	0,9
0,30928	67	0,64286	0,71014	0,71014	0,64286	0,2	0,9
0,30928	67	0,64286	0,71014	0,71014	0,64286	0,3	0,9

Na Tabela 12, encontram-se os resultados que atingiram as menores taxas de erro no classificador proposto e também os resultados obtidos da metodologia de Araújo (2017) aplicada à base estendida para duas classes.

Tabela 12 - Resumo dos resultados de classificação binária com a base de dados estendida – Menor taxa de erro apresentada para a metodologia proposta e a metodologia de Araújo (2017).

Classificador	Taxa de erro	Sens. Câncer	Sens. Não-Câncer	λ_1	λ_2
Araújo (2017)	0,375	0,592	0,638	0,1	0,8
Proposto ($S_k^1(\lambda_k)$)	0,289	0,50	0,797	0	1
Proposto ($S_k^2(\lambda_k)$)	0,237	0,286	0,956	0,1	0,4
Proposto ($S_k^3(\lambda_k)$)	0,309	0,643	0,710	0,3	0,9

Na Tabela 12 pode-se observar que a metodologia proposta obteve menor taxa de erro global em relação à metodologia adotada por Araújo (2017) nas três abordagens sugeridas. Dessas três abordagens, a matriz de covariâncias $S_k^3(\lambda_k)$ se mostrou mais consistente, obtendo uma maior taxa de sensibilidade para ambas as classes na comparação com a metodologia original usada por Araújo (2017).

A Tabela 13 mostra os resultados investigados por Brasileiro (2020) para os classificadores clássicos em uma análise binária: câncer $\times$ não câncer. Na Tabela, observa-se

que houve uma redução das taxas de erro com relação ao problema envolvendo três variáveis e uma maior dificuldade em obter uma boa classificação da classe maligno.

Tabela 13 - Taxa de erro global e Sensibilidade da Classe Maligno para os modelos clássicos da literatura usados para classificação binária: câncer x não câncer

Classificador	Taxa de erro	Sens. Câncer
MLP	0,2333	0
SVC	0,2333	0
Random Forest	0,3333	0,2857

O baixo índice de sensibilidade à classe câncer pode ser decorrente da diferença dos números de indivíduos entre as classes câncer e não câncer, uma vez que ao se agrupar as classes benigno e cisto, o número total da classe não câncer tenha passado a ser 69 contra 28 indivíduos da classe câncer (maligno). Essa diferença pode ter resultado em um viés no processo de classificação, pois não há um balanceamento das classes.

6 CONCLUSÕES

Neste trabalho foram apresentadas duas modificações no classificador proposto por Araújo (2017): o acréscimo de variáveis contínuas ao conjunto de variáveis intervalares utilizado pelo autor e a utilização de três métricas diferentes para calcular as matrizes de covariância parametrizadas de cada classe.

A metodologia aplicada na base de dados original de Araújo (2014) com 50 amostras resultou nos mesmos valores encontrados no classificador do autor, não havendo divergências nos resultados ao acrescentar variáveis contínuas ao processo de classificação intervalar.

Na base de dados estendida, o classificador proposto, assim como o classificador de Araújo (2017), apresentou um grau de dificuldade maior no processo de classificação. Em relação aos resultados encontrados por Araújo (2017), a abordagem $S_k^3(\lambda_k)$ foi a que apresentou valores mais equilibrados para as três classes, alcançando menor taxa de erro global e maiores índices de sensibilidade às classes benigno e cisto; porém, o índice de sensibilidade à classe maligno foi menor quando comparado ao classificador original.

Na classificação binária aplicada à base de dados estendida, o acréscimo das variáveis contínuas contribuiu para uma queda da taxa de erro global e da sensibilidade à classe não-câncer em relação à metodologia anterior de Araújo (2017). A metodologia sugerida obteve menor taxa de erro global nas três abordagens sugeridas e, dentre elas, a matriz de covariâncias $S_k^3(\lambda_k)$ se mostrou a mais consistente, obtendo a maior taxa de sensibilidade para ambas as classes: 64,3% para a classe câncer e 71,0% para a classe não-câncer.

Em relação aos resultados encontrados por Brasileiro (2020), o classificador proposto aplicado às três classes alcançou taxas de erros menores do que aquelas encontradas nos modelos clássicos da literatura utilizados pelo autor. Dos modelos utilizados, o Random Forest obteve a menor taxa de erro - 56,67% -, porém ainda superior à taxa de erro encontrada pela abordagem $S_k^3(\lambda_k)$. Na classificação binária, os classificadores MLP e SVC não conseguiram identificar nenhuma das amostras com anormalidades malignas, e o classificador Random Forest obteve sensibilidade para a classe câncer de 28,57%. Considerando esses valores, a metodologia proposta - representada pela abordagem $S_k^3(\lambda_k)$ - apresenta resultados mais satisfatórios.

Neste estudo, observa-se que a abordagem que obteve os melhores resultados, tanto na classificação com três classes como na classificação binária, foi aquela cuja matriz de covariância parametrizada não sofre a influência do número de indivíduos.

Como trabalhos futuros, sugere-se:

- Investigação de outros grupos de dados contínuos como, por exemplo, variáveis de textura;
- Utilização de deep learning como próximo passo na classificação de anormalidades de mamas.

REFERÊNCIAS

- ACHARYA, U. R.; NG, E. Y. K.; TAN, J.-H.; SREE, S. V. Thermography based breast cancer detection using texture features and support vector machine. **Journal of Medical Systems**, v. 36, n. 3, p. 1503–1510, 2012.
- 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.; SAYED, G. I.; GABER, T.; HASSANIEN, A. E.; SNASEL, V.; SILVA, L. F. Detection of breast abnormalities of thermograms based on a new segmentation method. **Proceedings of the 2015 Federated Conference on Computer Science and Information Systems**, FedCSIS 2015, v. 5, p. 255–261, 2015.
- ANDERSON, J. P.; PARIKH, J. R.; SHENFELD, D. K.; CHURCH, B. W.; LARAMIE, J. M.; MARDEKIAN, J.; PIPER, B. A.; WILLKE, R. J.; RUBLEE, D. A. Reverse Engineering and Evaluation of Prediction Models for Progression to Type 2 Diabetes: An Application of Machine Learning Using Electronic Health Records. **Journal of Diabetes Science and Technology**, v. 10, n. 1, p. 6–18, 2016.
- APPICE, A. et al. Classification of symbolic objects: A lazy learning approach. **Intelligent Data Analysis**, v. 10, n. 4, p. 301–324, 2006.
- ARAKERI, M. P.; REDDY, G. R. M. Computer-aided diagnosis system for tissue characterization of brain tumor on magnetic resonance images. **Signal, Image and Video Processing**, v. 9, n. 2, p. 409–425, 2015.
- ARAÚJO, M. C. **Uso de imagens termográficas para classificação de anormalidades de mama baseado em variáveis simbólicas intervalares**. Tese de doutorado, Universidade Federal de Perna.
- ARAÚJO, M. C. et al. An interval prototype classifier based on a parameterized distance applied to breast thermographic images. **Medical and Biological Engineering and Computing**, v. 55, n. 6, p. 873–884, 2017.
- ARAÚJO, M. C.; LIMA, R. C. F.; DE SOUZA, R. M. C. R. Interval symbolic feature extraction for thermography breast cancer detection. **Expert Systems with Applications**, v. 41, n. 15, p. 6728–6737, 2014.
- AYDINER, A. **Nuclear Medicine Imaging in Breast Cancer**. [s.l.: s.n.].
- BERANGER, B.; LIN, H.; SISSON, S. A. New models for symbolic data analysis. p. 1–33, 2018.
- BEZERRA, L. A. et al. Estimation of breast tumor thermal properties using infrared images. **Signal Processing**, v. 93, n. 10, p. 2851–2863, 2013a.
- BEZERRA, L. A. et al. Infrared imaging for breast cancer detection with proper selection of

properties: From acquisition protocol to numerical simulation. **Multimodality Breast Imaging: Diagnosis and Treatment**, p. 285–332, 2013b.

BILLARD, L.; DIDAY, E. **Symbolic Data Analysis: Conceptual Statistics and Data Mining**. [s.l: s.n.], 2006.

BILLARD, L.; DIDAY, E. **Clustering Methodology for Symbolic Data**. West Sussex, John Wiley & Sons Ltd, 2019.

BOCK, H.H.; DIDAY, E. **Analysis of Symbolic Data**. [s.l: s.n.], 2000.

BORCHARTT, T. B. et al. Breast thermography from an image processing viewpoint: A survey. **Signal Processing**, v. 93, n. 10, p. 2785–2803, 2013.

BRASILEIRO, F. R. S., **Análise de Imagens Termográficas para Classificação de Câncer de Mama Utilizando Rede Neural Convolucional**, Dissertação de mestrado, [s.l.] Universidade Federal de Pernambuco, 2020.

CIAMPI, A. et al. Growing a tree classifier with imprecise data. **Pattern Recognition Letters**, v. 21, n. 9, p. 787–803, 2000.

DAYAKSHINI, D. et al. Segmentation of Breast Thermogram Images for the Detection of Breast Cancer – A Projection Profile Approach. **Journal of Image and Graphics**, v. 3, n. 1, p. 47–51, 2015.

DE MEIRA, L. F. et al. Termografia na Área Biomédica. **Pan American Journal of Medical Thermology**, v. 1, n. 1, p. 31–41, 2014.

DE OLIVEIRA, J. P. S. et al. Segmentation of infrared images: A new technology for early detection of breast diseases. **Proceedings of the IEEE International Conference on Industrial Technology**, v. 2015-June, n. June, p. 1765–1771, 2015.

DE SOUZA, R. M. C. R.; QUEIROZ, D. C. F.; CYSNEIROS, F. J. A. Logistic regression-based pattern classifiers for symbolic interval data. **Pattern Analysis and Applications**, v. 14, n. 3, p. 273–282, 2011.

DE VASCONCELOS, J. H.; DOS SANTOS, W. P.; DE LIMA, R. D. C. F. Investigations on statistical classification methods for use in breast thermography. **IEEE Latin America Transactions**, v. 16, n. 6, p. 1631–1637, 2018.

DIDAY, E; NOIRHOMME-FRAITURE, M. **Symbolic Data Analysis and the SODAS Software**. [s.l: s.n.].

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification.pdf**. [s.l: s.n.], 2009.

EKICI, S.; JAWZAL, H. Breast cancer diagnosis using thermography and convolutional neural networks. **Medical Hypotheses**, v. 137, n. December 2019, p. 109542, 2020.

ESTEVA, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. **Nature**, v. 542, n. 7639, p. 115–118, 2017.

- ETEHADTAVAKOL, M. et al. Analysis of breast thermography using fractal dimension to establish possible difference between malignant and benign patterns. **Journal of Healthcare Engineering**, v. 1, n. 1, p. 27–43, 2010.
- ETEHADTAVAKOL, M. et al. Breast cancer detection from thermal images using bispectral invariant features. **International Journal of Thermal Sciences**, v. 69, p. 21–36, 2013.
- 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. 2, p. 1–20, 2013.
- FERREIRA, M. R. . Análise Discriminante Clássica E De Núcleo: Avaliações E Algumas Contribuições Relativas Aos Métodos Boosting E. p. 110, 2007.
- GABER, T. et al. Thermogram breast cancer prediction approach based on Neutrosophic sets and fuzzy c-means algorithm. **Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS**, v. 2015- Novem, p. 4254–4257, 2015.
- GARDUÑO-RAMÓN, M. A. et al. Supportive noninvasive tool for the diagnosis of breast cancer using a thermographic camera as sensor. **Sensors (Switzerland)**, v. 17, n. 3, 2017.
- GOGOI, U. R. et al. Singular value based characterization and analysis of thermal patches for early breast abnormality detection. **Australasian Physical and Engineering Sciences in Medicine**, v. 41, n. 4, p. 861–879, 2018.
- GOLESTANI, N.; ETEHADTAVAKOL, M.; NG, E. Y. K. Level set method for segmentation of infrared breast thermograms. **EXCLI Journal**, v. 13, p. 241–251, 2014.
- GONZALES, R.C.; WOODS, R.E.; EDDINS, S. L. **Digital image processing using MATLAB**. [s.l: s.n.].
- GONZALEZ-HERNANDEZ, J. L. et al. Technology, application and potential of dynamic breast thermography for the detection of breast cancer. **International Journal of Heat and Mass Transfer**, v. 131, p. 558–573, 2019.
- HASTIE, T; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**. [s.l: s.n.], 2009.
- JAULIN, LUC; KIEFFER, MICHEL; DIDRIT, OLIVIER; WALTER, É. **Applied Interval Analysis**. [s.l: s.n.], 2001.
- KANDLIKAR, S. 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.; PRASAD, S. V. A. V. Image processing for early diagnosis of breast cancer using infrared images. **2010 The 2nd International Conference on Computer and Automation Engineering, ICCAE 2010**, v. 3, p. 564–566, 2010.

KENNEDY, D. A.; LEE, T.; SEELY, D. A comparative review of thermography as a breast cancer screening technique. **Integrative Cancer Therapies**, v. 8, n. 1, p. 9–16, 2009.

LAWSON, R.N; CHUTGAI, M. S. Breast Cancer and Body Temperature. **JAMA: The Journal of the American Medical Association**, v. 183, n. 7, p. 221, 1963.

LAWSON, R. Implications of surface temperatures in the diagnosis of breast cancer. **Canadian Medical Association journal**, v. 75, n. 4, p. 309–311, 1956.

LEE, H.; CHEN, Y. P. P. Image based computer aided diagnosis system for cancer detection. **Expert Systems with Applications**, v. 42, n. 12, p. 5356–5365, 2015.

LOZANO, A.; HASSANIPOUR, F. Infrared imaging for breast cancer detection: An objective review of foundational studies and its proper role in breast cancer screening. **Infrared Physics and Technology**, v. 97, n. December 2018, p. 244–257, 2019.

MACHADO, D. A.; GIRALDI, G.; NOVOTNY, A. A. Topological Derivative Applied to Automatic Segmentation of Frontal Breast Thermograms. 2013.

MAHMOUDZADEH, E. et al. Extended hidden Markov model for optimized segmentation of breast thermography images. **Infrared Physics and Technology**, v. 72, p. 19–28, 2015.

MOTTA, L., CONCI, A., LIMA, R., DINIZ, E., LUÍS, S. Automatic segmentation on thermograms in order to aid diagnosis and 2D modeling. **Proceedings of 10th Workshop em Informática Médica**, n. January 2015, p. 1610–1619, 2010.

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. Integrative Computer-Aided Diagnostic With Breast Thermogram. **Journal of Mechanics in Medicine and Biology**, v. 07, n. 01, p. 1–10, 2007.

NOVA, R. DE L. V. USO DE IMAGENS TERMOGRÁFICAS DE MAMA PARA ANÁLISE DE PATOLOGIAS ATRAVÉS DA COMPARAÇÃO ENTRE DIVERSOS CLASSIFICADORES ESTATÍSTICOS. v. 91, p. 399–404, 2017.

NOVER, A. B. et al. Modern breast cancer detection: A technological review. **International Journal of Biomedical Imaging**, v. 2009, 2009.

OLIVEIRA, J. P. S. DE. **Extração automática da região de interesse em imagens térmicas laterais da mama**. [s.l: s.n.].

OLIVEIRA, L. F. M. avaliação de técnicas de processamento de imagem e extração de características em imagens de pulmão e de coluna vertebral. p. 39, 2015.

PRAMANIK, S. et al. Segmentation of Breast-Region in Breast Thermogram Using Arc-Approximation and Triangular-Space Search. **IEEE Transactions on Instrumentation and Measurement**, v. 69, n. 7, p. 4785–4795, 2020.

PRAMANIK, S.; BHATTACHARJEE, D.; NASIPURI, M. Wavelet based thermogram

analysis for breast cancer detection. **2015 International Symposium on Advanced Computing and Communication, ISACC 2015**, p. 205–212, 2015.

QUEIROZ, D. C. F. et al. Kernelized inner product-based discriminant analysis for interval data. **Pattern Analysis and Applications**, v. 21, n. 3, p. 731–740, 2018.

QUEIROZ, K. F. F. C. **Desenvolvimento e implementação de uma ferramenta computacional de uso médico para análise de imagens termográficas**. [s.l.] 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 and Technology**, v. 102, n. June, 2019.

ROSLIDAR, R. et al. A Review on Recent Progress in Thermal Imaging and Deep Learning Approaches for Breast Cancer Detection. **IEEE Access**, v. 8, p. 116176–116194, 2020.

ROSSI, F; CONAN-GUEZ, B. **Multi-layer Perceptron on Interval Data in: Classification, clustering and data analysis**. [s.l.: s.n.].

ROY, A.; KLEIN, D. Testing of mean interval for interval-valued data. **Communications in Statistics - Theory and Methods**, v. 0, n. 0, p. 1–17, 2019.

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.

SATHISH, D.; KAMATH, S. Detection of breast thermograms using ensemble classifiers. **Journal of Telecommunication, Electronic and Computer Engineering**, v. 10, n. 3–2, p. 35–39, 2018.

SATHYA, P.; DEEPA, V. B. Analysis of Supervised Image Classification Method for Satellite Images. v. 5, n. 2, p. 16–19, 2017.

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.

SERRANO, R. C. et al. Using Hurst Coefficient and Lacunarity to diagnosis early breast diseases. **17th International Conference on Systems, Signals and Image Processing**, p. 10–13, 2010.

SHAMSI KOSHKI, A. et al. Extending contour level set model for multi-class image segmentation with Application to Breast Thermography Images. **Infrared Physics and Technology**, v. 105, n. September 2019, p. 103174, 2020.

SIDEY-GIBBONS, J. A. M.; SIDEY-GIBBONS, C. J. Machine learning in medicine: a practical introduction. **BMC Medical Research Methodology**, v. 19, n. 1, p. 1–18, 2019.

SILVA, A. P. D.; BRITO, P. Linear discriminant analysis for interval data. **Computational**

Statistics, v. 21, n. 2, p. 289–308, 2006.

SIVA TEJA KAKILETI, G. M. et al. Advances in Breast Thermography. In: **Trends in Helicobacter pylori Infection**. [s.l.: s.n.]. v. ip. 13.

SOUZA, R. M. C. R. de. Classificação de imagens SAR baseada em uma abordagem simbólica. Dissertação (Mestrado) — Universidade Federal de Pernambuco, Recife, 1999.

SOUZA, R. M. C. et al. Dynamic Cluster Methods for Interval Data Based on Mahalanobis Distances. **Classification, Clustering, and Data Mining Applications**, p. 351–360, 2004.

SUGANTHI, S. S.; RAMAKRISHNAN, S. Anisotropic diffusion filter based edge enhancement for segmentation of breast thermogram using level sets. **Biomedical Signal Processing and Control**, v. 10, n. 1, p. 128–136, 2014.

TRAVIS, R. C. et al. Night shift work and breast cancer incidence: Three prospective studies and meta-analysis of published studies. **Journal of the National Cancer Institute**, v. 108, n. 12, p. 1–9, 2016.

WEBB, A. R. **Statistical Pattern Recognition**. [s.l.: s.n.]. v. 9

ZHOU, Q.; LI, Z.; AGGARWAL, J. K. Boundary extraction in thermal images by edge map. **Proceedings of the ACM Symposium on Applied Computing**, v. 1, p. 254–258, 2004.

**APÊNDICE A - OBSERVAÇÕES QUE ATINGIRAM O MAIOR ÍNDICE DE
SENSIBILIDADE PARA CLASSE MALIGNO – EQUAÇÃO 4.24^a**

Erro	Número de acertos	Sens. Maligno	Espec. Maligno	Sens. Benigno	Espec. Benigno	Sens. Cisto	Espec. Cisto	λ_1	λ_2	λ_3
0,65979	33	0,85714	0,33803	0,093023	0,88462	0,19231	0,84507	0,9	0	1
0,64948	34	0,85714	0,32394	0,093023	0,92308	0,23077	0,84507	1	0	1

**APÊNDICE B - OBSERVAÇÕES QUE ATINGIRAM O MAIOR ÍNDICE DE
SENSIBILIDADE PARA CLASSE MALIGNO – EQUAÇÃO 4.24c**

Erro	Número de acertos	Sens. Maligno	Espec. Maligno	Sens. Benigno	Espec. Benigno	Sens. Cisto	Espec. Cisto	λ_1	λ_2	λ_3
0,46392	52	0,67857	0,78667	0,37209	0,85417	0,65385	0,69014	0,5	0	0
0,47423	51	0,67857	0,77333	0,34884	0,85417	0,65385	0,69014	0,5	0	0,1
0,47423	51	0,67857	0,77333	0,34884	0,85417	0,65385	0,69014	0,5	0	0,2
0,48454	50	0,67857	0,77333	0,34884	0,83333	0,61538	0,69014	0,5	0	0,3
0,48454	50	0,67857	0,77333	0,34884	0,83333	0,61538	0,69014	0,5	0	0,4
0,48454	50	0,67857	0,77333	0,34884	0,83333	0,61538	0,69014	0,5	0	0,5
0,48454	50	0,67857	0,77333	0,34884	0,83333	0,61538	0,69014	0,5	0	0,6
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0	0,7
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0	0,8
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0	0,9
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,5	0	1
0,46392	52	0,67857	0,78667	0,37209	0,85417	0,65385	0,69014	0,5	0,1	0
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,1	0,1
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,1	0,2
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,1	0,3
0,48454	50	0,67857	0,77333	0,34884	0,83333	0,61538	0,69014	0,5	0,1	0,4
0,48454	50	0,67857	0,77333	0,34884	0,83333	0,61538	0,69014	0,5	0,1	0,5
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,1	0,6
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0,1	0,7
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0,1	0,8
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,1	0,9
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,1	1
0,46392	52	0,67857	0,78667	0,37209	0,85417	0,65385	0,69014	0,5	0,2	0
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,2	0,1
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,2	0,2
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,2	0,3
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,2	0,4
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,2	0,5
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,2	0,6
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0,2	0,7
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0,2	0,8
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,2	0,9
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,2	1
0,46392	52	0,67857	0,78667	0,37209	0,85417	0,65385	0,69014	0,5	0,3	0
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,3	0,1
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,3	0,2
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,3	0,3
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,3	0,4
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,5	0,3	0,5
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,5	0,3	0,6
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0,3	0,7

0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0,3	0,8
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,3	0,9
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,3	1
0,47423	51	0,67857	0,77333	0,37209	0,85417	0,61538	0,69014	0,5	0,4	0
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,4	0,1
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,4	0,2
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,4	0,3
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,4	0,4
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,5	0,4	0,5
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,5	0,4	0,6
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,5	0,4	0,7
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,5	0,4	0,8
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,4	0,9
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,4	1
0,47423	51	0,67857	0,77333	0,37209	0,85417	0,61538	0,69014	0,5	0,5	0
0,48454	50	0,67857	0,77333	0,34884	0,85417	0,61538	0,67606	0,5	0,5	0,1
0,47423	51	0,67857	0,78667	0,34884	0,85417	0,65385	0,67606	0,5	0,5	0,2
0,48454	50	0,67857	0,78667	0,34884	0,83333	0,61538	0,67606	0,5	0,5	0,3
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,5	0,5	0,4
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,5	0,5	0,5
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,5	0,5	0,6
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,5	0,5	0,7
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,5	0,5	0,8
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,5	0,9
0,49485	49	0,67857	0,76	0,34884	0,83333	0,57692	0,69014	0,5	0,5	1
0,46392	52	0,67857	0,8	0,39535	0,83333	0,61538	0,69014	0,5	0,6	0
0,47423	51	0,67857	0,8	0,37209	0,83333	0,61538	0,67606	0,5	0,6	0,1
0,46392	52	0,67857	0,81333	0,37209	0,83333	0,65385	0,67606	0,5	0,6	0,2
0,47423	51	0,67857	0,81333	0,37209	0,8125	0,61538	0,67606	0,5	0,6	0,3
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,6	0,4
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,6	0,5
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,6	0,6
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,5	0,6	0,7
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,5	0,6	0,8
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,6	0,9
0,48454	50	0,67857	0,78667	0,37209	0,8125	0,57692	0,69014	0,5	0,6	1
0,47423	51	0,67857	0,78667	0,37209	0,83333	0,61538	0,69014	0,5	0,7	0
0,47423	51	0,67857	0,8	0,37209	0,83333	0,61538	0,67606	0,5	0,7	0,1
0,47423	51	0,67857	0,8	0,37209	0,83333	0,61538	0,67606	0,5	0,7	0,2
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,7	0,3
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,7	0,4
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,7	0,5
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,7	0,6
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,5	0,7	0,7
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,5	0,7	0,8
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,7	0,9
0,48454	50	0,67857	0,8	0,37209	0,8125	0,57692	0,67606	0,5	0,7	1
0,48454	50	0,67857	0,78378	0,37209	0,79592	0,57692	0,70423	0,6	0	0

0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,4	0,5
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,4	0,6
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,4	0,7
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,6	0,4	0,8
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,6	0,4	0,9
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,6	0,4	1
0,49485	49	0,67857	0,77027	0,37209	0,79592	0,53846	0,70423	0,6	0,5	0
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,6	0,5	0,1
0,47423	51	0,67857	0,78667	0,37209	0,83333	0,61538	0,69014	0,6	0,5	0,2
0,47423	51	0,67857	0,78667	0,37209	0,83333	0,61538	0,69014	0,6	0,5	0,3
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,5	0,4
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,5	0,5
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,5	0,6
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,5	0,7
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,5	0,8
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,6	0,5	0,9
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,6	0,5	1
0,49485	49	0,67857	0,77027	0,37209	0,79592	0,53846	0,70423	0,6	0,6	0
0,49485	49	0,67857	0,77333	0,34884	0,83333	0,57692	0,67606	0,6	0,6	0,1
0,47423	51	0,67857	0,78667	0,37209	0,83333	0,61538	0,69014	0,6	0,6	0,2
0,47423	51	0,67857	0,78667	0,37209	0,83333	0,61538	0,69014	0,6	0,6	0,3
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,6	0,4
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,6	0,5
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,6	0,6
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,6	0,7
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,6	0,8
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,6	0,6	0,9
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,6	0,6	1
0,48454	50	0,67857	0,78378	0,39535	0,79592	0,53846	0,70423	0,6	0,7	0
0,48454	50	0,67857	0,78667	0,37209	0,83333	0,57692	0,67606	0,6	0,7	0,1
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	0,2
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	0,3
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	0,4
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	0,5
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	0,6
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	0,7
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	0,8
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	0,9
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,7	1
0,49485	49	0,67857	0,77027	0,37209	0,79592	0,53846	0,70423	0,6	0,8	0
0,48454	50	0,67857	0,78667	0,37209	0,83333	0,57692	0,67606	0,6	0,8	0,1
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,8	0,2
0,48454	50	0,67857	0,77333	0,39535	0,83333	0,53846	0,69014	0,6	0,8	0,3
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,8	0,4
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,8	0,5
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,8	0,6
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,8	0,7
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,8	0,8

0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,8	0,9
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,6	0,8	1
0,49485	49	0,67857	0,78378	0,37209	0,77551	0,53846	0,70423	0,6	0,9	0
0,48454	50	0,67857	0,78667	0,37209	0,8125	0,57692	0,69014	0,6	0,9	0,1
0,48454	50	0,67857	0,78667	0,39535	0,8125	0,53846	0,69014	0,6	0,9	0,2
0,48454	50	0,67857	0,78667	0,39535	0,8125	0,53846	0,69014	0,6	0,9	0,3
0,48454	50	0,67857	0,78667	0,39535	0,8125	0,53846	0,69014	0,6	0,9	0,4
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	0,9	0,5
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	0,9	0,6
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	0,9	0,7
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	0,9	0,8
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	0,9	0,9
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	0,9	1
0,49485	49	0,67857	0,78378	0,37209	0,77551	0,53846	0,70423	0,6	1	0
0,48454	50	0,67857	0,78667	0,37209	0,8125	0,57692	0,69014	0,6	1	0,1
0,48454	50	0,67857	0,78667	0,39535	0,8125	0,53846	0,69014	0,6	1	0,2
0,48454	50	0,67857	0,78667	0,39535	0,8125	0,53846	0,69014	0,6	1	0,3
0,48454	50	0,67857	0,78667	0,39535	0,8125	0,53846	0,69014	0,6	1	0,4
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	1	0,5
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	1	0,6
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	1	0,7
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	1	0,8
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	1	0,9
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,6	1	1
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,7	0	0
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,7	0	0,1
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,7	0	0,2
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,7	0	0,3
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,7	0	0,4
0,47423	51	0,67857	0,77333	0,37209	0,83333	0,61538	0,70423	0,7	0	0,5
0,47423	51	0,67857	0,77333	0,37209	0,83333	0,61538	0,70423	0,7	0	0,6
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,7	0	0,7
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,7	0	0,8
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,7	0	0,9
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,7	0	1
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,7	0,1	0
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,7	0,1	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,7	0,1	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,7	0,1	0,3
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,7	0,1	0,4
0,47423	51	0,67857	0,77333	0,37209	0,83333	0,61538	0,70423	0,7	0,1	0,5
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,7	0,1	0,6
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,7	0,1	0,7
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,7	0,1	0,8
0,48454	50	0,67857	0,76	0,37209	0,83333	0,57692	0,70423	0,7	0,1	0,9
0,48454	50	0,67857	0,75676	0,37209	0,81633	0,57692	0,71831	0,7	0,1	1
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,7	0,2	0
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,7	0,2	0,1

0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,6	0,6
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,6	0,7
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,6	0,8
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,6	0,9
0,48454	50	0,67857	0,75676	0,37209	0,81633	0,57692	0,71831	0,7	0,6	1
0,48454	50	0,67857	0,77027	0,39535	0,79592	0,53846	0,71831	0,7	0,7	0
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,7	0,7	0,1
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,7	0,7	0,2
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,7	0,7	0,3
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,7	0,7	0,4
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,7	0,5
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,7	0,6
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,7	0,7
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,7	0,8
0,48454	50	0,67857	0,77333	0,37209	0,83333	0,57692	0,69014	0,7	0,7	0,9
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,7	0,7	1
0,47423	51	0,67857	0,78378	0,4186	0,79592	0,53846	0,71831	0,7	0,8	0
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,7	0,8	0,1
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,7	0,8	0,2
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,7	0,8	0,3
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,7	0,8	0,4
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,8	0,5
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,8	0,6
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,8	0,7
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,8	0,8
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,8	0,9
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,7	0,8	1
0,48454	50	0,67857	0,77027	0,39535	0,79592	0,53846	0,71831	0,7	0,9	0
0,47423	51	0,67857	0,77027	0,39535	0,81633	0,57692	0,71831	0,7	0,9	0,1
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,7	0,9	0,2
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,7	0,9	0,3
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,7	0,9	0,4
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,9	0,5
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,9	0,6
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,9	0,7
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,9	0,8
0,47423	51	0,67857	0,78667	0,39535	0,83333	0,57692	0,69014	0,7	0,9	0,9
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,7	0,9	1
0,48454	50	0,67857	0,78378	0,39535	0,77551	0,53846	0,71831	0,7	1	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,7	1	0,1
0,48454	50	0,67857	0,78378	0,39535	0,79592	0,53846	0,70423	0,7	1	0,2
0,48454	50	0,67857	0,78378	0,39535	0,79592	0,53846	0,70423	0,7	1	0,3
0,48454	50	0,67857	0,78378	0,39535	0,79592	0,53846	0,70423	0,7	1	0,4
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,7	1	0,5
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,7	1	0,6
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,7	1	0,7
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,7	1	0,8
0,47423	51	0,67857	0,8	0,39535	0,8125	0,57692	0,69014	0,7	1	0,9

0,47423	51	0,67857	0,7973	0,39535	0,79592	0,57692	0,70423	0,7	1	1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,8	0	0
0,47423	51	0,67857	0,77027	0,39535	0,79592	0,57692	0,73239	0,8	0	0,1
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,8	0	0,2
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,8	0	0,3
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,8	0	0,4
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,8	0	0,5
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,8	0	0,6
0,48454	50	0,67857	0,75676	0,37209	0,81633	0,57692	0,71831	0,8	0	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0	0,9
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,8	0	1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,8	0,1	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,8	0,1	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,1	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,1	0,3
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,8	0,1	0,4
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,8	0,1	0,5
0,48454	50	0,67857	0,75676	0,37209	0,81633	0,57692	0,71831	0,8	0,1	0,6
0,48454	50	0,67857	0,75676	0,37209	0,81633	0,57692	0,71831	0,8	0,1	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,1	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,1	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,1	1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,8	0,2	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,8	0,2	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,2	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,2	0,3
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,2	0,4
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,2	0,5
0,48454	50	0,67857	0,75676	0,37209	0,81633	0,57692	0,71831	0,8	0,2	0,6
0,48454	50	0,67857	0,75676	0,37209	0,81633	0,57692	0,71831	0,8	0,2	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,2	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,2	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,2	1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,8	0,3	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,8	0,3	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,3	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,3	0,3
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,3	0,4
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,8	0,3	0,5
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,8	0,3	0,6
0,48454	50	0,67857	0,75676	0,37209	0,81633	0,57692	0,71831	0,8	0,3	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,3	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,3	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,8	0,3	1
0,47423	51	0,67857	0,77027	0,4186	0,79592	0,53846	0,73239	0,8	0,4	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,8	0,4	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,8	0,4	0,2

0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,8	0,8	0,7
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,8	0,8	0,8
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,8	0,8	0,9
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,8	0,8	1
0,46392	52	0,67857	0,78378	0,44186	0,79592	0,53846	0,73239	0,8	0,9	0
0,47423	51	0,67857	0,78378	0,4186	0,79592	0,53846	0,71831	0,8	0,9	0,1
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,8	0,9	0,2
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,8	0,9	0,3
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,8	0,9	0,4
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,8	0,9	0,5
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,8	0,9	0,6
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,8	0,9	0,7
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	0,8	0,9	0,8
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	0,8	0,9	0,9
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	0,8	0,9	1
0,47423	51	0,67857	0,77027	0,4186	0,79592	0,53846	0,73239	0,8	1	0
0,48454	50	0,67857	0,77027	0,39535	0,79592	0,53846	0,71831	0,8	1	0,1
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,8	1	0,2
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,8	1	0,3
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,8	1	0,4
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,8	1	0,5
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,8	1	0,6
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,8	1	0,7
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	0,8	1	0,8
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	0,8	1	0,9
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	0,8	1	1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,9	0	0
0,47423	51	0,67857	0,77027	0,39535	0,79592	0,57692	0,73239	0,9	0	0,1
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,9	0	0,2
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,9	0	0,3
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,9	0	0,4
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,9	0	0,5
0,47423	51	0,67857	0,76712	0,37209	0,8	0,61538	0,73239	0,9	0	0,6
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0	0,9
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0	1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,9	0,1	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,9	0,1	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,1	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,1	0,3
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,9	0,1	0,4
0,47423	51	0,67857	0,77027	0,37209	0,81633	0,61538	0,71831	0,9	0,1	0,5
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,1	0,6
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,1	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,1	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,1	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,1	1

0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,9	0,2	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,9	0,2	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,2	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,2	0,3
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,2	0,4
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,2	0,5
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,2	0,6
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,2	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,2	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,2	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,2	1
0,46392	52	0,67857	0,78378	0,4186	0,79592	0,57692	0,73239	0,9	0,3	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,9	0,3	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,3	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,3	0,3
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,3	0,4
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,3	0,5
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,3	0,6
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,3	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,3	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,3	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,3	1
0,47423	51	0,67857	0,77027	0,4186	0,79592	0,53846	0,73239	0,9	0,4	0
0,47423	51	0,67857	0,78378	0,39535	0,79592	0,57692	0,71831	0,9	0,4	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,4	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,4	0,3
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,4	0,4
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,4	0,5
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,4	0,6
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,4	0,7
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,4	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,4	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,4	1
0,47423	51	0,67857	0,77027	0,4186	0,79592	0,53846	0,73239	0,9	0,5	0
0,48454	50	0,67857	0,77027	0,39535	0,79592	0,53846	0,71831	0,9	0,5	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,5	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,5	0,3
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,5	0,4
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,5	0,5
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,5	0,6
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,5	0,7
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,5	0,8
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,5	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,5	1
0,47423	51	0,67857	0,77027	0,4186	0,79592	0,53846	0,73239	0,9	0,6	0
0,48454	50	0,67857	0,77027	0,39535	0,79592	0,53846	0,71831	0,9	0,6	0,1
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,6	0,2
0,47423	51	0,67857	0,78378	0,37209	0,81633	0,61538	0,70423	0,9	0,6	0,3

0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,6	0,4
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,6	0,5
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,6	0,6
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,6	0,7
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,6	0,8
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,6	0,9
0,48454	50	0,67857	0,75342	0,37209	0,8	0,57692	0,73239	0,9	0,6	1
0,47423	51	0,67857	0,77027	0,4186	0,79592	0,53846	0,73239	0,9	0,7	0
0,48454	50	0,67857	0,77027	0,39535	0,79592	0,53846	0,71831	0,9	0,7	0,1
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,7	0,2
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,7	0,3
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,7	0,4
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,7	0,5
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,7	0,6
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,7	0,7
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,7	0,8
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,7	0,9
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,7	1
0,47423	51	0,67857	0,77027	0,4186	0,79592	0,53846	0,73239	0,9	0,8	0
0,48454	50	0,67857	0,77027	0,39535	0,79592	0,53846	0,71831	0,9	0,8	0,1
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,8	0,2
0,49485	49	0,67857	0,75676	0,37209	0,81633	0,53846	0,70423	0,9	0,8	0,3
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,8	0,4
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,8	0,5
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,8	0,6
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,8	0,7
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,8	0,8
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,8	0,9
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,8	1
0,47423	51	0,67857	0,77027	0,4186	0,79592	0,53846	0,73239	0,9	0,9	0
0,48454	50	0,67857	0,77027	0,39535	0,79592	0,53846	0,71831	0,9	0,9	0,1
0,49485	49	0,67857	0,75676	0,37209	0,81633	0,53846	0,70423	0,9	0,9	0,2
0,49485	49	0,67857	0,75676	0,37209	0,81633	0,53846	0,70423	0,9	0,9	0,3
0,49485	49	0,67857	0,75676	0,37209	0,81633	0,53846	0,70423	0,9	0,9	0,4
0,48454	50	0,67857	0,77027	0,37209	0,81633	0,57692	0,70423	0,9	0,9	0,5
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,9	0,6
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,9	0,7
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,9	0,8
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,9	0,9
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	0,9	0,9	1
0,46392	52	0,67857	0,78378	0,44186	0,79592	0,53846	0,73239	0,9	1	0
0,47423	51	0,67857	0,78378	0,4186	0,79592	0,53846	0,71831	0,9	1	0,1
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,9	1	0,2
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,9	1	0,3
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	0,9	1	0,4
0,47423	51	0,67857	0,78378	0,39535	0,81633	0,57692	0,70423	0,9	1	0,5
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	0,9	1	0,6
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	0,9	1	0,7

0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	1	0,9	0,5
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	1	0,9	0,6
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	1	0,9	0,7
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	1	0,9	0,8
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	1	0,9	0,9
0,48454	50	0,67857	0,76712	0,37209	0,8	0,57692	0,71831	1	0,9	1
0,47423	51	0,67857	0,78378	0,4186	0,79592	0,53846	0,71831	1	1	0
0,46392	52	0,67857	0,78378	0,44186	0,79592	0,53846	0,73239	1	1	0,1
0,48454	50	0,67857	0,77027	0,4186	0,79592	0,5	0,71831	1	1	0,2
0,48454	50	0,67857	0,77027	0,39535	0,81633	0,53846	0,70423	1	1	0,3
0,48454	50	0,67857	0,76712	0,39535	0,8	0,53846	0,71831	1	1	0,4
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	1	1	0,5
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	1	1	0,6
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	1	1	0,7
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	1	1	0,8
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	1	1	0,9
0,47423	51	0,67857	0,78082	0,39535	0,8	0,57692	0,71831	1	1	1