



ANTONIO MENDES MAGALHÃES JÚNIOR

**UTILIZAÇÃO DE COMITÊS DE REDES NEURAIAS
ARTIFICIAIS NA CLASSIFICAÇÃO DE DANOS EM
SEMENTES DE GIRASSOL**

LAVRAS – MG

2020

ANTONIO MENDES MAGALHÃES JÚNIOR

**UTILIZAÇÃO DE COMITÊS DE REDES NEURAIAS ARTIFICIAIS NA
CLASSIFICAÇÃO DE DANOS EM SEMENTES DE GIRASSOL**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Profa. Dra. Thelma Sáfadi
Orientadora

LAVRAS – MG
2020

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da
Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Magalhães Júnior, Antonio Mendes
Utilização de comitês de Redes Neurais Artificiais na classificação
de danos em sementes de girassol / Antonio Mendes Magalhães
Júnior. - 2020.
150 p. : il.

Orientador(a): Thelma Sáfadi.

Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2020.
Bibliografia.

1. Redes Neurais Artificiais. 2. Análise de sementes. 3.
Aprendizado de máquinas. I. Sáfadi, Thelma. II. Título.

ANTONIO MENDES MAGALHÃES JÚNIOR

**UTILIZAÇÃO DE COMITÊS DE REDES NEURAIAS ARTIFICIAIS NA
CLASSIFICAÇÃO DE DANOS EM SEMENTES DE GIRASSOL**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 13 de Novembro de 2020.

Profa. Dra. Thelma Sáfadi	UFLA
Prof. Dr. Paulo Henrique Sales Guimarães	UFLA
Prof. Dr. Wilian Soares Lacerda	UFLA
Prof. Dr. Crysttian Arantes Paixão	UFSC

Profa. Dra. Thelma Sáfadi
Orientadora

**LAVRAS – MG
2020**

AGRADECIMENTOS

Aos meus familiares por todo incentivo dado, não somente durante a realização deste trabalho, mas durante toda a minha formação acadêmica. Obrigado à minha mãe Verli, ao meu pai Antonio, ao meu padrasto Ventura, à minha madrinha Ivany e a todos mais que contribuíram de alguma forma.

Aos meus amigos que estiveram presentes nesta etapa, contribuindo direta ou indiretamente, Ícaro Viterbre, Paula Santos, Rodrigo Afonso e Matheus Feres.

Aos amigos Daniel Amorim, Evanise Penido, Lara Sales, Layla Murad e Marília Mickaele pelos momentos de descontração e companhia.

À Livia Souza pela cooperação durante a execução deste trabalho e pelo enorme apoio nos momentos mais marcantes, sejam acadêmicos ou não.

À Renata Rabelo por estar presente em grande parte do tempo durante a realização desse trabalho, tornando essa tarefa menos árdua ao me fazer enxergar o mundo de maneira mais bela.

À Ariana Alves pela sincera e duradoura amizade, o que certamente aliviou o peso dessa jornada. *Ohana* quer dizer família.

Aos bons professores que encontrei até aqui em toda minha jornada de aprendizado, em especial àqueles que dispunham de poucos recursos para o exercício de sua profissão, mas ainda assim me encheram de esperança e admiração.

Ao Prof. Dr. Wilian Soares Lacerda pelo auxílio, esclarecimentos e diversas sugestões nas questões pertinentes ao trabalho.

À minha orientadora Profa. Dra. Thelma Sáfydi pelo incentivo, contribuições, sugestões, esclarecimentos e sobretudo pela compreensão demonstrada durante o desenvolvimento deste trabalho.

Ao Laboratório de Análise de Sementes do Departamento de Agricultura da Universidade Federal de Lavras pela cessão do banco de dados utilizado no estudo.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

“Bem, cientificamente, tradições são idiotas.”
(Rick Sanchez)

RESUMO

O Brasil detém o terceiro maior mercado mundial de sementes e movimenta bilhões de reais anualmente, tornando de extrema importância a análise de sementes. Por meio de técnicas de análise, é possível determinar o potencial de germinação e identificar danos nas sementes. O raio-x é uma das técnicas mais desejáveis, uma vez que fornece rápida análise e não causa a destruição das sementes. Entretanto, as imagens resultantes do processo raio-x muitas vezes necessitam de pós processamento, buscando a melhoria visual das imagens para a análise por parte dos avaliadores. O processo de avaliação pode ser realizado por um ou mais avaliadores, porém é carregado de subjetividade, tornando interessante a automatização dessa análise. As RNAs (Redes Neurais Artificiais) são reconhecidamente eficazes para a utilização em problemas de reconhecimento de padrões e classificação de dados, tornando-se candidatas para essa automatização. Neste trabalho, o objetivo foi realizar a classificação de imagens radiografadas de sementes de girassol, quanto ao seu nível de dano, utilizando diversas técnicas de processamento de imagens e extração de características para compor diferentes conjuntos de dados a fim de treinar as RNAs. Para tal, foi utilizado um conjunto de dados composto por imagens de radiografias de sementes de girassol, classificadas por avaliadores entre três categorias: sementes cheias, parcialmente cheias e deformadas. Utilizando-se essas imagens, foram compostos conjuntos de dados que, por sua vez, foram utilizados para treinar, validar e testar RNAs, que foram então utilizadas para compor comitês. Para cada caso foram compostos 10 comitês e obtidas as médias das métricas de acurácia, AUC e índice Kappa desses comitês. As médias das métricas de desempenho, aproximadamente 90% para acurácia, 0,97 para AUC e 0,84 para Kappa, no melhor caso, indicam a eficiência da metodologia utilizada neste trabalho e sugerem a possibilidade do uso desta em composição com as metodologias usuais no que tange à classificação e avaliação de sementes.

Palavras-chave: Análise de raio-x. Análise de sementes. Reconhecimento de padrões.

ABSTRACT

Brazil has the third largest world seed market and invoices billions of reais annually, making seed analysis extremely important. Through analysis techniques, it is possible to determine the germination potential and identify damage to the seeds. X-ray is one of the most desirable techniques, because it provides fast analysis and does not result in the destruction of seeds. However, the images resulting from the X-ray process often require post-processing, seeking visual improvement of the images for analysis by the evaluators. The evaluation process can be carried out by one or more evaluators, but it has a lot of subjectivity, making the automation of this analysis interesting. ANNs (Artificial Neural Networks) are known to be effective for use in pattern recognition and data classification problems, making them good candidates for this automation. In this work, the goal was to perform the classification of radiographic images of sunflower seeds, according to their level of damage, using multiple techniques of image processing and extraction of characteristics to compose different datasets in order to train the ANNs. For this, a dataset consisting of radiographic images of sunflower seeds was used, classified by evaluators into three categories: filled, partially filled or deformed seeds. Using these images, datasets were composed and used to train, validate and test ANNs, which were then used to compose committees. For each case, 10 committees were formed, and obtained averages of the metrics of accuracy, AUC and Kappa index of the committees. The averages of the performance metrics, approximately 90% for accuracy, 0.97 for AUC and 0.84 for Kappa, in the best case, indicate the efficiency of the methodology used in this work and suggest the possibility of using it in composition to usual methodologies for seed classification and evaluation.

Keywords: X-ray analysis. Seed analysis. Pattern Recognition.

LISTA DE FIGURAS

Figura 2.1 – Exemplo de sementes radiografadas de pinheiros.	28
Figura 2.2 – Exemplo de equalização de histograma: a) antes b) após a equalização	30
Figura 2.3 – Histograma redistribuído com a técnica de limitação de contraste . . .	31
Figura 2.4 – Exemplo de binarização de imagem.	32
Figura 2.5 – Direções para o cálculo das matrizes de coocorrência.	33
Figura 2.6 – Obtenção de matrizes de coocorrência de níveis de cinza.	34
Figura 2.7 – Representação de um neurônio biológico.	37
Figura 2.8 – Representação de um Perceptron.	38
Figura 2.9 – RNAs <i>feed-forward</i> : a) de camada única b) de múltiplas camadas. . .	40
Figura 2.10 – RNA recorrente.	40
Figura 2.11 – Diagrama do aprendizado supervisionado.	41
Figura 2.12 – Função de ativação limiar.	44
Figura 2.13 – Função de ativação linear.	45
Figura 2.14 – Função de ativação sigmoide.	45
Figura 2.15 – Função de ativação tangente hiperbólica.	46
Figura 2.16 – Função de ativação ReLu.	47
Figura 2.17 – Função de ativação Leaky ReLu.	48
Figura 2.18 – Funcionamento da função de ativação <i>softmax</i>	49
Figura 2.19 – Esquema de um comitê de máquinas.	50
Figura 2.20 – Esquema de uma matriz de confusão.	55
Figura 2.21 – Composição de uma matriz de confusão.	56
Figura 2.22 – Exemplos de classificadores no espaço ROC.	59
Figura 2.23 – Exemplos de curvas ROC e distribuição das classes.	61
Figura 3.1 – Diagrama esquemático da geração dos diversos conjuntos de dados. . .	69
Figura 3.2 – Máscara elíptica.	74
Figura 3.3 – Esquema de normalização dos dados.	79
Figura 3.4 – Diagrama esquemático da composição dos comitês.	82
Figura 4.1 – Exemplo de imagens de cada classe de sementes e suas variações para o aumento artificial do conjunto de dados: a) cheias b) parcialmente cheias c) deformadas.	84

Figura 4.2 – Exemplo de imagens de cada classe: a) antes do redimensionamento b) após o redimensionamento.	85
Figura 4.3 – Exemplo de imagens de cada classe: a) antes da equalização de histograma e redimensionamento b) após equalização de histograma c) após redimensionamento.	86
Figura 4.4 – Exemplo de imagens de cada classe: a) antes da segmentação e redimensionamento b) após segmentação c) após redimensionamento. . . .	86
Figura 4.5 – Exemplo de imagens de cada classe: a) antes da segmentação (maior área) e redimensionamento b) após segmentação (maior área) c) após redimensionamento.	87
Figura 4.6 – Exemplo de imagens de cada classe: a) antes da aplicação da máscara b) durante a aplicação da máscara (imagens já redimensionadas) c) pixels válidos de cada imagem após a aplicação da máscara.	87
Figura 4.7 – Exemplos de histogramas de imagens de cada uma das classes.	88
Figura 4.8 – Matriz de confusão (a) e curvas ROC (b) utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	92
Figura 4.9 – Matriz de confusão (a) e curvas ROC (b) utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	92
Figura 1 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	105
Figura 2 – Curvas ROC utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	105
Figura 3 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	106
Figura 4 – Curva ROC utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes cheias e parcialmente cheias.	106

Figura 5 –	Matriz de confusão utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes: (1) cheias e (2) deformadas.	107
Figura 6 –	Curva ROC utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes cheias e deformadas.	107
Figura 7 –	Matriz de confusão utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	108
Figura 8 –	Curva ROC utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes parcialmente cheias e deformadas.	108
Figura 9 –	Matriz de confusão utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	109
Figura 10 –	Curvas ROC utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	109
Figura 11 –	Matriz de confusão utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	110
Figura 12 –	Curva ROC utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes cheias e parcialmente cheias.	110
Figura 13 –	Matriz de confusão utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias e (2) deformadas.	111
Figura 14 –	Curva ROC utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes cheias e deformadas.	111
Figura 15 –	Matriz de confusão utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	112

Figura 16 – Curva ROC utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes parcialmente cheias e deformadas.	112
Figura 17 – Matriz de confusão utilizando o conjunto de dados “Histograma” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	113
Figura 18 – Curvas ROC utilizando o conjunto de dados “Histograma” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	113
Figura 19 – Matriz de confusão utilizando o conjunto de dados “Histograma” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	114
Figura 20 – Curva ROC utilizando o conjunto de dados “Histograma” e considerando as classes de sementes cheias e parcialmente cheias.	114
Figura 21 – Matriz de confusão utilizando o conjunto de dados “Histograma” e considerando as classes de sementes: (1) cheias e (2) deformadas. . . .	115
Figura 22 – Curva ROC utilizando o conjunto de dados “Histograma” e considerando as classes de sementes cheias e deformadas.	115
Figura 23 – Matriz de confusão utilizando o conjunto de dados “Histograma” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	116
Figura 24 – Curva ROC utilizando o conjunto de dados “Histograma” e considerando as classes de sementes parcialmente cheias e deformadas. . . .	116
Figura 25 – Matriz de confusão utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	117
Figura 26 – Curvas ROC utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	117
Figura 27 – Matriz de confusão utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	118
Figura 28 – Curva ROC utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes cheias e parcialmente cheias.	118
Figura 29 – Matriz de confusão utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) cheias e (2) deformadas. . . .	119

Figura 30 – Curva ROC utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes cheias e deformadas.	119
Figura 31 – Matriz de confusão utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	120
Figura 32 – Curva ROC utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes parcialmente cheias e deformadas.	120
Figura 33 – Matriz de confusão utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	121
Figura 34 – Curvas ROC utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	121
Figura 35 – Matriz de confusão utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	122
Figura 36 – Curva ROC utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes cheias e parcialmente cheias.	122
Figura 37 – Matriz de confusão utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes: (1) cheias e (2) deformadas.	123
Figura 38 – Curva ROC utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes cheias e deformadas.	123
Figura 39 – Matriz de confusão utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	124
Figura 40 – Curva ROC utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes parcialmente cheias e deformadas.	124
Figura 41 – Matriz de confusão utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	125

Figura 42 –	Curvas ROC utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	125
Figura 43 –	Matriz de confusão utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	126
Figura 44 –	Curva ROC utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes cheias e parcialmente cheias.	126
Figura 45 –	Matriz de confusão utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes: (1) cheias e (2) deformadas. . .	127
Figura 46 –	Curva ROC utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes cheias e deformadas.	127
Figura 47 –	Matriz de confusão utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	128
Figura 48 –	Curva ROC utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes parcialmente cheias e deformadas. . .	128
Figura 49 –	Curvas ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	129
Figura 50 –	Curvas ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	129
Figura 51 –	Matriz de confusão utilizando o conjunto de dados “Haralick” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias. . .	130
Figura 52 –	Curva ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes cheias e parcialmente cheias.	130
Figura 53 –	Matriz de confusão utilizando o conjunto de dados “Haralick” e considerando as classes de sementes: (1) cheias e (2) deformadas.	131
Figura 54 –	Curva ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes cheias e deformadas.	131
Figura 55 –	Matriz de confusão utilizando o conjunto de dados “Haralick” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	132

Figura 56 – Curva ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes parcialmente cheias e deformadas.	132
Figura 57 – Matriz de confusão utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	133
Figura 58 – Curvas ROC utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	133
Figura 59 – Matriz de confusão utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	134
Figura 60 – Curva ROC utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes cheias e parcialmente cheias.	134
Figura 61 – Matriz de confusão utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes: (1) cheias e (2) deformadas.	135
Figura 62 – Curva ROC utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes cheias e deformadas.	135
Figura 63 – Matriz de confusão utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	136
Figura 64 – Curva ROC utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes parcialmente cheias e deformadas.	136
Figura 65 – Matriz de confusão utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	137
Figura 66 – Curvas ROC utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	137
Figura 67 – Matriz de confusão utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	138

Figura 68 – Curva ROC utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes cheias e parcialmente cheias.	138
Figura 69 – Matriz de confusão utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) cheias e (2) deformadas.	139
Figura 70 – Curva ROC utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes cheias e deformadas. . . .	139
Figura 71 – Matriz de confusão utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	140
Figura 72 – Curva ROC utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes parcialmente cheias e deformadas.	140
Figura 73 – Matriz de confusão utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.	141
Figura 74 – Curvas ROC utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.	141
Figura 75 – Matriz de confusão utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.	142
Figura 76 – Curva ROC utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes cheias e parcialmente cheias.	142
Figura 77 – Matriz de confusão utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias e (2) deformadas.	143
Figura 78 – Curva ROC utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes cheias e deformadas.	143

Figura 79 – Matriz de confusão utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.	144
Figura 80 – Curva ROC utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes parcialmente cheias e deformadas.	144
Figura 81 – Intervalos de confiança para a acurácia média considerando as classes de sementes cheias, parcialmente cheias e deformadas em conjunto. . .	145
Figura 82 – Intervalos de confiança para a acurácia média considerando as classes de sementes cheias e parcialmente cheias.	145
Figura 83 – Intervalos de confiança para a acurácia média considerando as classes de sementes cheias e deformadas.	146
Figura 84 – Intervalos de confiança para a acurácia média considerando as classes de sementes parcialmente cheias e deformadas.	146
Figura 85 – Intervalos de confiança para a AUC média considerando as classes de sementes cheias, parcialmente cheias e deformadas em conjunto.	147
Figura 86 – Intervalos de confiança para a AUC média considerando as classes de sementes cheias e parcialmente cheias.	147
Figura 87 – Intervalos de confiança para a AUC média considerando as classes de sementes cheias e deformadas.	148
Figura 88 – Intervalos de confiança para a AUC média considerando as classes de sementes parcialmente cheias e deformadas.	148
Figura 89 – Intervalos de confiança para o índice Kappa médio considerando as classes de sementes cheias, parcialmente cheias e deformadas em conjunto.	149
Figura 90 – Intervalos de confiança para o índice Kappa médio considerando as classes de sementes cheias e parcialmente cheias.	149
Figura 91 – Intervalos de confiança para o índice Kappa médio considerando as classes de sementes cheias e deformadas.	150
Figura 92 – Intervalos de confiança para o índice Kappa médio considerando as classes de sementes parcialmente cheias e deformadas.	150

LISTA DE TABELAS

Tabela 2.1 – Descritores de Haralick e suas fórmulas.	35
Tabela 2.2 – Medidas de desempenho para os casos de classificação binária e multi-classes.	57
Tabela 2.3 – Proporções de classificações por dois classificadores com C classes. . .	64
Tabela 2.4 – Intensidade da concordância Kappa	65
Tabela 4.1 – Características dos conjuntos de dados formados.	89
Tabela 4.2 – Acurácia média dos comitês de RNAs para cada conjunto de dados. .	90
Tabela 4.3 – AUC média dos comitês de RNAs para cada conjunto de dados. . . .	90
Tabela 4.4 – Índice Kappa médio dos comitês de RNAs para cada conjunto de dados.	91

LISTA DE SIGLAS

AUC	<i>Area Under the ROC Curve</i>
CNNs	<i>Convolutional Neural Networks</i>
GLCMs	<i>Gray-Level Co-Occurrence Matrices</i>
MLP	<i>Multilayer Perceptron</i>
OvA	<i>One vs. All</i>
ReLU	<i>Rectified Linear Unit</i>
RNAs	Redes Neurais Artificiais
ROC	<i>Receiver Operating Characteristic</i>
Rprop	<i>Resilient Backpropagation</i>
SVM	<i>Support Vector Machine</i>

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Contextualização e motivação	21
1.2	Objetivos	22
1.3	Trabalhos correlatos	23
1.4	Estruturação do texto	25
2	REFERENCIAL TEÓRICO	27
2.1	Análise de sementes	27
2.1.1	Análise por raio-x	27
2.2	Processamento Digital de Imagens	28
2.2.1	Equalização de histograma	29
2.2.2	Segmentação	31
2.3	Extração de atributos de texturas	32
2.3.1	Abordagens estatísticas de descrição de padrões de textura . . .	32
2.3.1.1	Matrizes de Coocorrência de Níveis de Cinza	33
2.3.1.2	Descritores de textura de Haralick	34
2.4	Redes Neurais Artificiais	36
2.4.1	História	37
2.4.2	Arquitetura	39
2.4.2.1	Redes <i>feed-forward</i> de camada única	39
2.4.2.2	Redes <i>feed-forward</i> de múltiplas camadas	39
2.4.2.3	Redes recorrentes	40
2.4.3	Tipos de aprendizado	41
2.4.3.1	Aprendizado supervisionado	41
2.4.3.2	Aprendizado não-supervisionado	42
2.4.3.3	Aprendizado por reforço	42
2.4.4	Funções de ativação	43
2.4.4.1	Limiar (binária)	43
2.4.4.2	Linear	44
2.4.4.3	Sigmoide	45
2.4.4.4	Tangente hiperbólica	46
2.4.4.5	ReLU	47

2.4.4.6	Leaky ReLu	48
2.4.4.7	<i>Softmax</i>	49
2.5	Comitê de máquinas	50
2.5.1	Geração de componentes	51
2.5.2	Métodos de combinação	53
2.6	Métricas de avaliação de desempenho	54
2.6.1	Matrizes de confusão	54
2.6.2	Análise ROC	58
2.6.2.1	Espaço ROC	58
2.6.2.2	Curvas ROC	60
2.6.3	Índice Kappa de Cohen	63
3	MATERIAIS E MÉTODOS	67
3.1	Materiais	67
3.2	Métodos	68
3.2.1	Balanceamento de classes	68
3.2.2	Aumento artificial dos dados	70
3.2.3	Redimensionamento e transformação de cores	70
3.2.4	Redução de dimensionalidade dos dados de entrada	71
3.2.5	Conjuntos de dados por valores dos pixels	72
3.2.5.1	Valores dos pixels	72
3.2.5.2	Valores dos pixels com equalização de histograma	72
3.2.5.3	Segmentação	73
3.2.5.4	Segmentação (maior área)	73
3.2.5.5	Máscara elíptica	74
3.2.6	Conjuntos de dados por extração de características	75
3.2.6.1	Histograma	75
3.2.6.2	Descritores de textura de Haralick (14 descritores)	75
3.2.6.3	Descritores de textura de Haralick (4 descritores)	76
3.2.7	Conjuntos de dados por combinação de outros conjuntos	76
3.2.7.1	Descritores de textura de Haralick (duas distâncias)	77
3.2.7.2	Descritores de textura de Haralick e valores dos pixels com equalização	77

3.2.8	Normalização dos dados	78
3.2.9	Arquitetura e parâmetros de treinamento das RNAs	79
3.2.10	Treinamento das RNAs e formação dos comitês	81
3.2.11	Obtenção das métricas de avaliação	83
4	RESULTADOS E DISCUSSÃO	84
4.1	Aumento artificial dos dados	84
4.2	Conjuntos de dados por valores dos pixels	85
4.2.1	Valores dos pixels	85
4.2.2	Valores dos pixels com equalização de histograma	85
4.2.3	Segmentação	86
4.2.4	Segmentação (maior área)	86
4.2.5	Máscara elíptica	87
4.3	Conjuntos de dados por extração de características	87
4.3.1	Histograma	87
4.3.2	Descritores de textura de Haralick (14 descritores)	88
4.3.3	Descritores de textura de Haralick (4 descritores)	88
4.4	Conjuntos de dados por combinação de outros conjuntos	88
4.4.1	Descritores de textura de Haralick (duas distâncias)	89
4.4.2	Descritores de textura de Haralick e valores dos pixels com equalização	89
4.5	Resumo dos conjuntos de dados	89
4.6	Métricas de avaliação	90
4.7	Discussão	93
5	CONCLUSÃO	94
5.1	Trabalhos futuros	95
	REFERÊNCIAS BIBLIOGRÁFICAS	97
	APENDICE A – Matrizes de confusão e curvas ROC do conjunto de dados “Valores dos pixels sem equalização”	105
	APENDICE B – Matrizes de confusão e curvas ROC do conjunto de dados “Valores dos pixels com equalização”	109
	APENDICE C – Matrizes de confusão e curvas ROC do conjunto de dados “Histograma”	113

APENDICE D – Matrizes de confusão e ROC do conjunto de dados “Segmentação”	117
APENDICE E – Matrizes de confusão e curvas ROC do conjunto de dados “Segmentação (maior área)”	121
APENDICE F – Matrizes de confusão e curvas ROC do conjunto de dados “Máscara elíptica”	125
APENDICE G – Matrizes de confusão e curvas ROC do conjunto de dados “Haralick”	129
APENDICE H – Matrizes de confusão e curvas ROC do conjunto de dados “Haralick reduzido”	133
APENDICE I – Matrizes de confusão e curvas ROC do conjunto de dados “Haralick (duas distâncias)”	137
APENDICE J – Matrizes de confusão e curvas ROC do conjunto de dados “Haralick e valores dos pixels com equalização”	141
APENDICE K – Intervalos de confiança para a acurácia média	145
APENDICE L – Intervalos de confiança para a AUC média . .	147
APENDICE M – Intervalos de confiança para o índice Kappa médio	149

1 INTRODUÇÃO

Neste capítulo é realizada uma introdução ao assunto abordado neste trabalho, mostrando os principais aspectos relativos ao seu contexto, além de sua motivação e objetivos. Trabalhos correlatos e a estruturação do texto também estão presentes.

1.1 Contextualização e motivação

A busca por maior produção agrícola torna-se cada vez mais evidente com o crescimento populacional e o aumento da renda e consumo *per capita*, principalmente em países em desenvolvimento. Este fato não se traduz apenas na necessidade de expansão de áreas de cultivo, mas principalmente no aumento da produtividade das lavouras.

O Brasil é um dos principais produtores de alimentos no mundo, destacando-se principalmente na produção de carnes e grãos. Atualmente, soja e milho são os pilares da produção de grãos, tendo correspondido, em conjunto, a mais de 85% da produção de grãos na safra 2019/2020 (CONAB, 2020).

O girassol, embora pouco produzido em comparação ao milho e a soja, possui produção destacada quando comparada a outras oleaginosas, como canola e mamona (CONAB, 2020). De acordo com Grunvald et al. (2009), o girassol apresenta propriedades como alto teor de óleo nos grãos e ótima qualidade de subprodutos. Além disso, possui facilidade de adaptação a diferentes condições climáticas.

Nos últimos dez anos, entre as safras de 2009/2010 e 2019/2020, houve um salto na produção total de grãos de 149 milhões para 257 milhões toneladas. Neste mesmo período, a área plantada total passou de 47,4 milhões para 65,9 milhões de hectares, enquanto a produtividade média dos grãos saltou de 3148 para 3900 kg/ha (CONAB, 2020).

O aumento de produtividade está intimamente ligado ao desenvolvimento tecnológico, cada vez mais presente no setor agrícola. Este aumento de produtividade é, portanto, alcançado não apenas com mecanização das lavouras e utilização de técnicas modernas de irrigação, mas também com a seleção de sementes de boa qualidade para o plantio.

Diante deste cenário, surge a necessidade de se analisar as sementes, a fim de avaliar sua qualidade. Seja por doenças, ataques de pragas, danos mecânicos durante o beneficiamento ou questões sanitárias, as sementes podem apresentar danos que dificultem ou mesmo impeçam sua germinação e, conseqüentemente, a produção de uma plântula normal.

Na análise de sementes, a qualidade é avaliada a partir de atributos genéticos, físicos, fisiológicos e sanitários que possuam relação com sua capacidade de germinação e vigor. Dessa forma, os testes de avaliação consistem, em sua maioria, em examinar o poder de germinação e o vigor das sementes, além do nível de deterioração que possam ter sofrido.

Por se tratar de uma forma de avaliação rápida e não destrutiva, a análise de sementes por meio de raio-x tem se mostrado uma alternativa aos testes tradicionais. Ao se examinar as imagens radiográficas geradas no processo, torna-se possível a diferenciação entre sementes com boa formação e as que apresentem algum dano, seja mecânico, causado por ataque de pragas ou apenas má formação (BRASIL, 2009).

O processo de exame dessas imagens geralmente é realizado de forma visual por um analista de sementes, o que traz subjetividade ao processo por estar sujeito à interpretação do analista. Visando mitigar a subjetividade, métodos de processamento de imagens e aprendizagem de máquinas podem ser combinados para automatização deste processo.

Desta forma, conjuntos de dados podem ser formados por meio da utilização de metodologias de processamento de imagens digitais e extração de características de imagens. Por sua vez, estes conjuntos de dados podem ser utilizados por técnicas de aprendizagem de máquina para a classificação das sementes de forma automática.

A utilização dessas técnicas para a classificação automática das sementes permite a redução da subjetividade inerente ao usual processo de análise, realizado exclusivamente por analistas humanos. Além disso, resulta em economia de recursos temporais e financeiros.

1.2 Objetivos

O objetivo deste trabalho foi a proposição de uma metodologia para a avaliação e classificação de sementes por meio da utilização de comitês compostos por RNAs treinadas com conjuntos dados extraídos de imagens radiográficas dessas sementes. Para tal, destacaram-se como objetivos específicos:

- Analisar e aplicar correções às imagens radiográficas de sementes que constituem o conjunto de dados original.

- Aplicar técnicas de processamento de imagens e extração de características para compor diversos conjuntos de dados a serem utilizados no treinamento, validação e teste das RNAs.
- Verificar diferentes configurações de comitês e de suas RNAs a fim de identificar as que resultam em melhor desempenho.
- Avaliar cada abordagem utilizada para a geração dos conjuntos de dados com base em métricas de acurácia, curvas ROC (*Receiver Operating Characteristic*) e índice Kappa.

1.3 Trabalhos correlatos

Na literatura existem diversos exemplos de utilização de algoritmos de aprendizagem de máquinas para a classificação de sementes, seja por danos ou espécie. Na maioria desses exemplos, são utilizadas informações morfológicas e de textura para o treinamento destes algoritmos.

Wang, Ram e Dowell (2002) utilizaram espectroscopia NIR (*near-infrared*) para a classificação de danos em sementes de soja. Foram utilizados como classificadores modelos PLS (*Partial Least Squares*) e RNAs, tendo esse último obtido melhor desempenho, com acurácia de até 100% ao classificar sementes sadias e danificadas pelo clima.

Em Khatchaturian e Padilha (2008), após binarização de imagens de sementes de soja, foram utilizados os valores dos pixels (0 e 1) para o treinamento de RNAs do tipo MLP (*Multilayer Perceptron*). Considerando oito classes representando as variedades de sementes, a acurácia obtida na classificação foi de aproximadamente 57%.

No trabalho de Kiratiratanapruk e Sinthupinyo (2011) são utilizados padrões de textura baseados nas Matrizes de Coocorrência de Níveis de Cinza (*Gray Level Co-occurrence Matrices* - GLCMs) e Padrão Binário Local (*Local Binary Patterns* - LBP), além dos histogramas das imagens nos espaços de cores RGB (*red*, *green* e *blue*) e HSV (*hue*, *saturation* e *value*) para a classificação de diversos tipos de danos em sementes. A partir da extração desses recursos em um conjunto de mais de dez mil amostras, foram treinadas Máquinas de Vetores de Suporte (*Support Vector Machine* - SVM) e foi obtido nível de acurácia superior a 80% para sementes com danos (classificação entre os tipos de danos) e 95% para sementes normais.

Em Kurtulmuş, Alibas e Kavdir (2015) foram utilizadas informações de cores, formas e textura das sementes para compor um banco de dados. Por meio da utilização de seleção sequencial de recursos, foram selecionadas as características mais importantes para a classificação. Utilizando RNAs do tipo MLP, foi obtido nível de acurácia de aproximadamente 85% ao se classificar oito espécies de sementes de pimenta.

Redes Neurais Artificiais foram utilizadas para classificar sementes de trigo de três espécies em Nabeel e Quteishat (2015). Os autores utilizaram técnicas de processamento de imagens para a segmentação das mesmas. Posteriormente, a partir das imagens segmentadas, foram extraídas métricas como área, comprimento do maior eixo, excentricidade da elipse, diâmetro e perímetro. Essas informações foram utilizadas para o treinamento das RNAs e foi obtida acurácia de 95%.

Hussain e Ajaz (2015) utilizaram as informações de área, perímetro, compacidade, comprimento e largura do grão, coeficiente de assimetria e comprimento do sulco do grão para treinar diferentes classificadores. RNAs do tipo MLP se mostraram a melhor opção quando comparadas a classificadores como Naive Bayes, LWL (*Locally Weighted Regression*) e SMO (*Sequential Minimal Optimization*), resultando em acurácia de aproximadamente 95%.

No trabalho de Liu et al. (2015) também foram utilizadas algumas técnicas de processamento de imagens e foram extraídos recursos de forma, cor e textura para o treinamento de RNAs. Visando a classificação de diversos tipos de danos em sementes de soja, obtiveram acurácia superior a 97%.

Tujo (2019) utilizou algumas medidas como área, perímetro e comprimento do núcleo de sementes para o treinamento de RNAs. Neste trabalho foi obtida acurácia de cerca de 97% para a classificação de diferentes espécies de sementes.

Zhang, Dai e Cheng (2020) utilizaram imagens hiperespectrais com classificadores, dentre eles Redes Neurais Convolucionais (*Convolutional Neural Networks - CNNs*), para a classificação de danos causados por congelamento em sementes de milho. A utilização desse tipo de redes resultou em acurácia de 97,5%.

No trabalho de Ahmed et al. (2020) foram comparadas técnicas tradicionais de aprendizagem de máquina e *deep learning* com transferência de aprendizado para a classificação de sementes de melancia. Os autores consideraram duas classes de sementes: normais viáveis e anormais inviáveis. O treinamento dos classificadores tradicionais de

aprendizagem de máquinas se deu com a utilização de recursos (intensidade, contraste, FFT, LBP etc) extraídos de imagens de raio-x das sementes. Dentre as técnicas clássicas, o melhor desempenho foi utilizando LDA (*Linear Discriminant Analysis*), com 83,3% de acurácia. Na abordagem com *deep learning*, os autores utilizaram CNNs pré-treinadas (ConvNet, AlexNet, VGG-19, ResNet-50 e ResNet-101) para a classificação das sementes. As CNNs foram treinadas com as imagens de raio-x das sementes e o melhor desempenho foi obtido utilizando ResNet-50, resultando em 87,3% de acurácia.

Além destes trabalhos, é importante ressaltar os de Leite, Sáfadi e Carvalho (2013), Sáfadi et al. (2016) e Cassiano (2018). Estes autores utilizaram, total ou parcialmente, o mesmo conjunto de dados utilizado no desenvolvimento do presente trabalho.

Leite, Sáfadi e Carvalho (2013) optaram pelas abordagens PCA (*Principal Component Analysis*) e ICA (*Independent Component Analysis*) juntamente com análise de discriminante para a classificação das sementes. O índice de acurácia global obtido neste trabalho foi de 82%, considerando as três classes de sementes.

Em Sáfadi et al. (2016) foi utilizada a abordagem de espectro de ondaletas e SVM para a classificação das sementes. O índice médio de acurácia foi de 57% quando consideradas as três classes em conjunto, e de 82% quando consideradas apenas as classes de sementes cheias e deformadas.

Por fim, Cassiano (2018), com a utilização das abordagens de Expoente Direcional de Hurst e SVM, obteve índice de acurácia de 74,5% quando consideradas as três classes em conjunto, e de 99,76% quando consideradas apenas as classes de sementes cheias e deformadas.

1.4 Estruturação do texto

O presente trabalho está organizado em cinco capítulos. No primeiro capítulo há uma introdução ao tema em estudo por meio de sua contextualização, motivação e seus objetivos.

O capítulo dois traz uma revisão acerca das áreas de estudo que culminam nesse trabalho. Neste capítulo são apresentadas técnicas de análise de sementes e de processamento de imagens e extração de suas características, bem como a história das RNAs, suas arquiteturas, tipos de aprendizado e funcionamento, além de máquinas de comitês e métricas de avaliação de desempenho de sistemas de classificação.

No capítulo três são apresentados os materiais utilizados e a metodologia do trabalho. São abordadas as aplicações de técnicas para processamento e extração de características de imagens para a geração de diversos conjuntos de dados. Neste capítulo também são abordadas as configurações utilizadas nos treinamentos das RNAs, bem como a metodologia utilizada para a composição dos comitês formados por elas.

No capítulo quatro são apresentados os resultados obtidos com a aplicação das técnicas de processamento de imagens e extração de características. São comparadas as métricas de avaliação resultantes do treinamento dos comitês de RNAs com cada um dos conjuntos de dados utilizados. O capítulo é finalizado com uma discussão acerca desses resultados.

Por fim, no capítulo cinco é apresentada a conclusão sobre o trabalho, além de sugeridas possibilidades de uso para a metodologia desenvolvida. A seção de trabalhos futuros, em que são colocadas algumas possibilidades para a obtenção de resultados ainda melhores, finaliza este capítulo.

2 REFERENCIAL TEÓRICO

Neste capítulo é apresentada a base teórica para a realização deste estudo. Dessa forma, são conceituadas áreas de conhecimento utilizadas para o desenvolvimento técnico do presente trabalho. Este referencial teórico é estruturado em seis tópicos: análise de sementes, processamento digital de imagens, extração de atributos de textura em imagens, RNAs, comitês de máquinas e métricas de avaliação de desempenho.

2.1 Análise de sementes

A análise de sementes é de grande importância para se determinar a qualidade das mesmas. Testes para avaliar o potencial físico e fisiológico são realizados em laboratórios, a fim de determinar o potencial de germinação, identificar danos mecânicos ou causados por insetos, além de outras características das sementes.

A avaliação da qualidade de sementes visa auxiliar na tomada de decisão sobre aceitação ou descarte de lotes. Dessa forma, determinar a ocorrência de sementes defeituosas em um lote torna-se desejável, uma vez que essas sementes influenciam nos resultados de germinação (CRAVIOTTO et al., 2002).

2.1.1 Análise por raio-x

O teste de raio-x é um dos métodos mais utilizadas na análise de sementes. Este método propicia avaliação rápida e não destrutiva, permitindo a diferenciação entre sementes com boa formação e as que sofreram algum dano (ISTA, 2004; BRASIL, 2009). A utilização de raios-x não afeta a germinação das sementes (BINO; AARTSE; BURG, 1993; MENEZES; CÍCERO; VILLELA, 2005). Além disso, não é necessário um tratamento prévio das sementes para sua aplicação.

A premissa dessa técnica consiste na diferença de absorção dos raios-x nos diferentes tipos de tecidos das sementes, dependendo do comprimento de onda da radiação e da espessura, densidade e composição dos tecidos (BINO; AARTSE; BURG, 1993; ISTA, 2004; SIMAK, 1991). Nesta técnica, as sementes são colocadas entre uma fonte de raios-x e um filme radiográfico, quando os raios-x passam pelas sementes e atingem o filme, uma imagem latente é criada. Quando o filme é revelado, forma-se uma imagem permanente de sombras claras e escuras no filme radiográfico. Nesta imagem, as áreas mais escuras correspondem às partes em que os raios-x penetram de forma mais fácil, enquanto as

mais claras correspondem às partes mais densas das sementes (SIMAK, 1991). A Figura 2.1 mostra exemplos de radiografia de sementes de pinheiro. As sementes com pouco preenchimento são consideradas de pior qualidade.

Figura 2.1 – Exemplo de sementes radiografadas de pinheiros.



Fonte: bvcentre.ca/index.php/whitebark/restoration/seed_collection

A aplicação de raio-x para a análise de sementes tem ocorrido desde a década de 50, tendo como início o trabalho de Simak e Gustafsson (1953), em que avaliaram a qualidade de sementes de espécies florestais. Essa técnica tem sido utilizada com sucesso, relacionando a anatomia das sementes com a germinação e morfologia das plântulas (BURG et al., 1994; CICERO et al., 1998; CARVALHO et al., 1999).

2.2 Processamento Digital de Imagens

O PDI (Processamento Digital de Imagens) diz respeito ao processamento de imagens digitais por um computador. O processamento de imagens viabiliza aplicações em duas categorias: o aprimoramento de informações visuais das imagens para interpretação humana e a análise automática por computador de informações presentes nas imagens (MARQUES FILHO; VIEIRA NETO, 1999; GONZALEZ; WOODS, 2002).

As imagens digitais são divididas em duas categorias: vetoriais e matriciais. As imagens vetoriais são criadas com a utilização de formas geométricas como linhas, pontos e curvas aliadas às fórmulas matemáticas para a definição de início e fim dessas formas. Por sua vez, as imagens matriciais utilizam agrupamentos de pixels em sua formação, com cada pixel representando a intensidade da cor. Os pixels são os menores elementos de uma

imagem matricial. A localização de cada pixel é definida por um par de coordenadas (x , y) e seu valor representa a intensidade da cor, geralmente variando de 0 a 255.

As imagens podem apresentar imperfeições em razão de problemas durante suas aquisições. Imperfeições como pixels ruidosos e contraste e/ou brilho inadequado podem justificar a necessidade de se realizar um pré-processamento das imagens (MARQUES FILHO; VIEIRA NETO, 1999). O objetivo do pré-processamento é aprimorar a qualidade da imagem para sua utilização em outras etapas, como exibição, armazenamento e classificação. Existem dezenas de técnicas de processamento de imagens, mas apenas algumas delas são abordadas neste trabalho.

2.2.1 Equalização de histograma

O histograma de uma imagem indica a quantidade de pixels em cada um de seus níveis de cinza ou banda de cor, para imagens coloridas, e pode ser entendido como uma função de distribuição de frequência deles. Imagens diferentes podem apresentar o mesmo histograma, pois os histogramas não apresentam informações espaciais da distribuição dos pixels, apenas a quantidade de pixels com determinada intensidade.

Segundo Gonzalez e Woods (2002), o histograma é uma função discreta $h(r_k) = n_k$, em que r_k é o k -ésimo nível de intensidade e n_k é o número de pixels da imagem com intensidade r_k e cujos níveis de intensidade desta imagem estejam no intervalo $[0, L - 1]$, em que L é o maior valor possível da codificação da imagem.

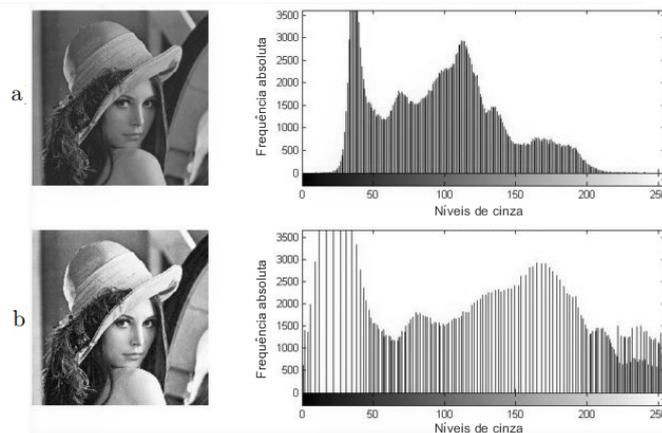
Ao se realizar a equalização do histograma, pretende-se torná-lo o mais uniforme possível, redistribuindo seus níveis e mapeando-os para novos níveis. A equalização de histograma resulta em uma imagem com maior contraste, como visto na Figura 2.2.

Ao se equalizar um histograma, uma função de transferência define o mapeamento das intensidades de cada pixel da imagem inicial para novas intensidades na imagem final (GONZALEZ; WOODS, 2002). A função de transferência deve ser monotonicamente crescente. Uma das formas de obter uma função de transferência é utilizar o histograma acumulado. Dessa forma, a equalização dá-se pela Equação 2.1.

$$s_k = T(r_k) = \frac{L-1}{MN} \times \sum_{j=0}^k n_j, \quad (2.1)$$

em que k é o número de níveis de intensidade da imagem ($k = 1, 2, \dots, L - 1$), s_k é o k -ésimo novo nível de intensidade, $T(r_k)$ é a função de transferência, L é o maior valor da

Figura 2.2 – Exemplo de equalização de histograma: a) antes b) após a equalização



Fonte: adaptado de digitalpadm.com/image-contrast-enhancement-cumulative-histogram-equalization

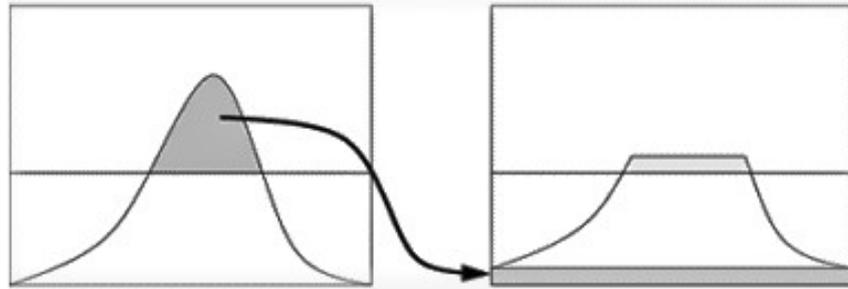
codificação da imagem ($L = 2^b$, onde b é o número de bits), $M \times N$ é a resolução (tamanho em linhas por colunas) da imagem e, por fim, n_j é o número de pixels da imagem com nível de intensidade j .

Em geral, a equalização global funciona de forma satisfatória se a distribuição das intensidades dos pixels é similar em toda a imagem. Porém, se isso não acontecer, essa técnica pode aumentar de forma demasiada o brilho das regiões mais luminosas e iluminar de forma insuficiente pequenas regiões muito escuras (CONCI, 2015). Buscando contornar este problema, surgiu a técnica de Equalização Adaptativa de Histograma (*Adaptive Histogram Equalization - AHE*).

Na Equalização Adaptativa de Histograma são criados histogramas regionais da imagem para vizinhanças definidas de cada pixel, como 15x15 pixels, por exemplo. Este método melhora o contraste local, realçando os detalhes da imagem, porém tende a ampliar o ruído em regiões muito homogêneas, uma vez que nessas regiões o histograma é altamente concentrado (PIZER et al., 1987). Dessa forma, surge a Equalização Adaptativa de Histograma com Limitação de Contraste (*Contrast Limited Adaptive Histogram Equalization - CLAHE*).

A Equalização Adaptativa de Histograma com Limitação de Contraste limita a amplificação do contraste, reduzindo o problema com a amplificação do ruído. Neste método é feito um corte no histograma em um valor limite estabelecido e redistribui-se as intensidades que estão acima desse limite por todo o histograma, conforme Figura 2.3.

Figura 2.3 – Histograma redistribuído com a técnica de limitação de contraste



Fonte: adaptado de
en.wikipedia.org/wiki/Adaptive_histogram_equalization

2.2.2 Segmentação

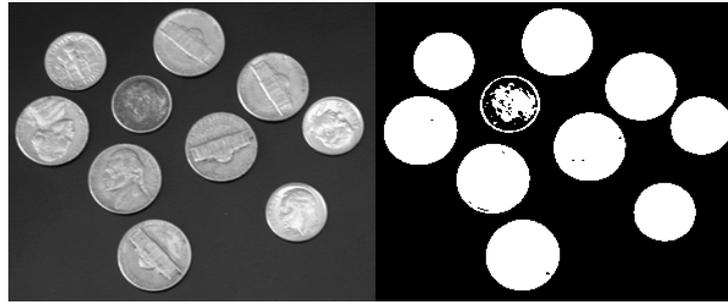
A segmentação tem como objetivo dividir a imagem nas regiões de interesse que a compõe, podendo ser considerada como um processo de classificação de pixels. Segundo Marques Filho e Vieira Neto (1999), a segmentação é uma tarefa de difícil implementação, sendo objeto de diversos estudos e possuindo uma variedade de métodos para sua realização.

O método de limiarização é um dos mais simples, porém é amplamente utilizado. Este método permite detectar regiões onde exista uma razoável uniformidade nos valores de intensidade do pixel, agrupando os níveis de cinza de uma imagem que satisfazem a um determinado critério de similaridade ou de um valor limiar pré-definido de intensidade dos pixels.

A limiarização consiste na subdivisão do histograma da imagem em K intervalos disjuntos I_k , $k = 1, 2, \dots, K$, utilizando $K - 1$ valores de limiar $T_1 < T_2 \dots < T_{K-1}$ e associar cada pixel da imagem com intensidade $f(x, y) \in I_k$ à região k . Para uma imagem com apenas um limiar, o processo é denominado binarização, sendo descrito pela Equação 2.2. A Figura 2.4 exemplifica um resultado do processo de binarização em uma imagem.

$$g(x, y) = \begin{cases} 1, & f(x, y) \geq \text{limiar} \\ 0, & f(x, y) < \text{limiar} \end{cases} \quad (2.2)$$

Figura 2.4 – Exemplo de binarização de imagem.



Fonte: mathworks.com/help/images/ref/imbinarize.html

2.3 Extração de atributos de texturas

A extração e seleção de atributos visam encontrar atributos relevantes para a classificação de imagens, tendo como principal objetivo a caracterização de objetos por meio de medidas (SILVA, 2013). A textura é uma das principais características utilizadas na identificação de objetos e regiões de interesse. Segundo Haralick, Shanmugam e Dinstein (1973), texturas definem parâmetros como uniformidade, densidade, regularidade e intensidade, entre outros. Eles ainda caracterizam textura como um conceito bidimensional, onde uma dimensão engloba as propriedades de tonalidade e a outra, as relações espaciais entre elas.

2.3.1 Abordagens estatísticas de descrição de padrões de textura

De acordo com Gonzalez e Woods (2002), existem três principais abordagens utilizadas para a descrição de padrões de textura:

1. **Abordagens estatísticas:** a textura é definida por um conjunto de estatísticas sobre a distribuição e relação entre os níveis de cinza de uma imagem.
2. **Abordagens estruturais:** a textura é representada como formada pela repetição de padrões que obedecem a alguma regra de posicionamento para a sua geração.
3. **Abordagens espectrais:** baseiam-se nas propriedades do espectro de Fourier, sendo utilizadas principalmente na detecção de periodicidade.

A extração de características utilizando a abordagem estatística proporciona a descrição de imagens por meio de regras estatísticas que governam a distribuição e relação

entre os diferentes níveis de pixels (ALVES; ARAÚJO, 2006). A abordagem estatística pode ser dividida em abordagem de primeira e segunda ordem.

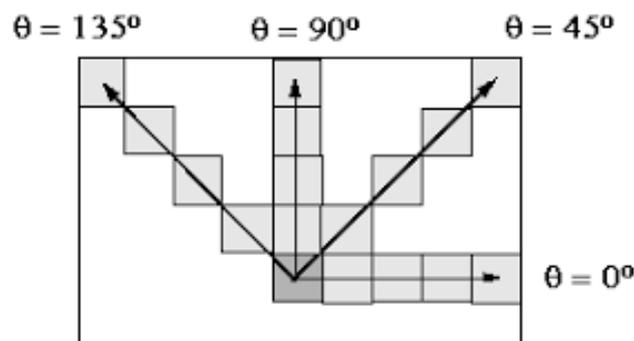
- Primeira ordem: relaciona-se à distribuição dos níveis de cinza em determinada imagem e usa-se o histograma de primeira ordem como base para extrair características como média, coeficiente de assimetria, desvio padrão, etc.
- Segunda ordem: considera o posicionamento espacial da ocorrência dos níveis de cinza nas imagens.

A diferença primordial entre essas abordagens é que uma imagem com mesma quantidade de diferentes níveis de cinza que outra, com diferença apenas em seus posicionamentos, irá apresentar mesmo histograma, conforme abordagem de primeira ordem. Estatísticas de segunda ordem, ou superiores, permitem diferenciar tais imagens.

2.3.1.1 Matrizes de Coocorrência de Níveis de Cinza

As Matrizes de Coocorrência de Níveis de Cinza mostram a organização espacial da ocorrência dos níveis de cinza em uma imagem. Essas matrizes são dispostas em organização bi-dimensional de níveis de cinza, onde pares de pixels são separados por uma relação espacial fixa que define a distância e a direção, d e θ , respectivamente, que um pixel referencial possui em relação ao pixel vizinho (MEDEIROS, 2001). Na obtenção dessas matrizes, a direção é definida em termos angulares, geralmente utilizando quatro direcionamentos: 0° , 45° , 90° e 135° , conforme se vê na Figura 2.5.

Figura 2.5 – Direções para o cálculo das matrizes de coocorrência.

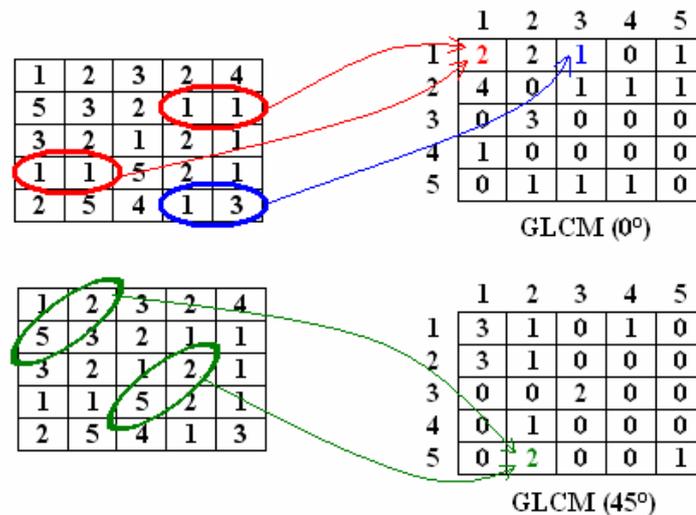


Fonte: adaptado de matlab.izmiran.ru/help/toolbox/images/enhanc15.html

A matriz de coocorrência de textura considera a relação entre dois pixels por vez, um chamado de pixel referência e o outro, de pixel vizinho. A matriz é formada a partir da contagem das combinações entre os tons de cinza dos referidos pixels. Portanto, cada elemento da matriz de coocorrência representa a contagem do número de ocorrência dos pares de níveis de cinza, ou seja, a quantidade de vezes que cada combinação foi observada, com uma distância d e direção θ .

A Figura 2.6 exemplifica a construção de duas matrizes de coocorrência. As matrizes do lado esquerdo são as imagens, enquanto as do lado direito são as GLCMs. Na primeira GLCM, considerou-se como vizinhos os pixels na direção 0° , enquanto na segunda considerou-se como vizinhos os pixels na direção 45° . Em ambas as foi considerada a distância $d = 1$.

Figura 2.6 – Obtenção de matrizes de coocorrência de níveis de cinza.



Fonte: Tou, Lau e Tay (2007)

2.3.1.2 Descritores de textura de Haralick

Em Haralick, Shanmugam e Dinstein (1973) são descritas diversas características que podem ser obtidas a partir de cálculos com as matrizes de coocorrência. Essas características são chamadas de Descritores de Haralick e servem como medidas para representação e diferenciação de texturas. A Tabela 2.1 mostra os descritores propostos e suas fórmulas.

Tabela 2.1 – Descritores de Haralick e suas fórmulas.

Descritor de textura	Fórmula do descritor
Segundo momento angular	$f_1 = \sum_i \sum_j \{p(i, j)\}^2$
Contraste	$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\}$, em que $n = i - j $.
Correlação	$f_3 = \frac{\sum_i \sum_j (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$, em que μ_x , μ_y , σ_x e σ_y são as médias e desvios-padrão de p_x e p_y .
Soma de quadrados: variância	$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j)$
Momento da diferença inversa	$f_5 = \sum_i \sum_j \frac{p(i, j)}{1 + (i - j)^2}$
Soma da média	$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i)$
Soma da variância	$f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 p_{x+y}(i)$
Soma da entropia	$f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log \{p_{x+y}(i)\}$
Entropia	$f_9 = - \sum_i \sum_j p(i, j) \log \{p(i, j)\}$
Diferença da variância	$f_{10} = - \sum_{i=0}^{N_g-1} (i - f_{11})^2 p_{x-y}(i)$
Diferença da entropia	$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log \{p_{x-y}(i)\}$
Medida de informação de correlação 1	$f_{12} = \frac{HXY - HXY1}{\max \{HX, HY\}}$ $HXY = - \sum_i \sum_j p(i, j) \log \{p(i, j)\}$ $HXY1 = - \sum_i \sum_j p(i, j) \log \{p_x(i)p_y(j)\}$ HX e HY são as entropias de p_x e p_y
Medida de informação de correlação 2	$f_{13} = \sqrt{1 - \exp \{-2(HXY2 - HXY)\}}$ $HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log \{p_x(i)p_y(j)\}$
Coeficiente de correlação máxima	$f_{14} = \sqrt{\text{segundo maior autovalor de } Q}$ $Q(i, j) = \sum_k \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)}$

Fonte: do autor.

Nas fórmulas presentes na Tabela 2.1, \sum_i e \sum_j são, respectivamente, $\sum_{i=1}^{N_g}$ e $\sum_{j=1}^{N_g}$, N_g é o número de níveis de cinza da imagem, $p(i, j)$ é a (i, j) -ésima entrada na matriz GLCM normalizada, $p(i, j) = P(i, j)/R$, em que R é a soma dos elementos da GLCM, $p_x(i)$ é a i -ésima entrada da matriz de probabilidade marginal, obtida pela soma das linhas de $p(i, j)$ e, por fim, $p_y(j)$ é a j -ésima entrada da matriz de probabilidade marginal, obtida pela soma das colunas de $p(i, j)$.

É importante ressaltar que alguns desses descritores estão altamente correlacionados entre si. Dessa forma, Haralick propõe utilizá-los como ponto de partida para a seleção ou combinação linear de alguns descritores. Na literatura encontra-se trabalhos justificando a utilização de apenas alguns descritores. Em Pacifici, Chini e Emery (2009), por exemplo, apenas seis descritores são considerados de maior relevância para a extração de características de textura (energia, contraste, variância, correlação, entropia e momento da diferença inversa). Outros exemplos de trabalhos que corroboram com esta ideia são (CONNERS; TRIVEDI; HARLOW, 1984) e (WESZKA; DYER; ROSENFELD, 1976).

2.4 Redes Neurais Artificiais

O cérebro humano é extremamente poderoso e complexo, sendo capaz de processar grande volume de informações em curto período de tempo. Os neurônios são as unidades mais importantes do cérebro, que conectados entre si, formam uma rede capaz de processar e transmitir informações (RAUBER, 1997).

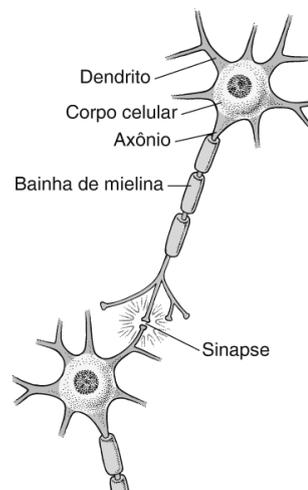
Ainda de acordo com Rauber (1997), as características de interesse que dão ao cérebro a capacidade de um comportamento inteligente e que são desejáveis em uma inteligência artificial são: robustez e tolerância a falhas, capacidade de aprendizagem, processamento de informações incertas e paralelismo.

A robustez garante que a eliminação de alguns neurônios não afete a funcionalidade global. A capacidade de aprendizagem garante ao cérebro a possibilidade de aprender tarefas nunca antes executadas. O processamento de informações incertas confere ao cérebro um raciocínio correto mesmo em situações em que tenha informações incompletas ou ruidosas. Por fim, com o paralelismo, não existe a restrição de execuções de instruções de forma sequencial, já que inúmeros neurônios podem estar ativos ao mesmo tempo.

Os neurônios são compostos pelos dendritos, corpo central e axônio. Os dendritos são considerados os terminais de entrada. O corpo do neurônio pode ser entendido como

uma unidade de soma, onde reações químicas e elétricas representam o processamento da informação. A saída das informações é realizada por impulsos elétricos propagados por meio do axônio e ao final do axônio existem ramificações responsáveis por distribuir a informação para outros neurônios, formando as sinapses. Os impulsos que são recebidos por determinado neurônio são processados no corpo e caso atinja um determinado linear, este neurônio dispara, enviando informações para outros neurônios, como ilustrado na Figura 2.7.

Figura 2.7 – Representação de um neurônio biológico.



Fonte: adaptado de msdmanuals.com/pt/casa/multimedia/figure/neu_nerve_cell_pt

Inspirados no modelo biológico, muitos pesquisadores buscaram simular o funcionamento do cérebro, principalmente o processo de aprendizagem por experiência, objetivando principalmente tarefas como classificação, processamento de imagens e reconhecimento de padrões (DATA SCIENCE ACADEMY, 2019). Essas pesquisas resultaram no modelo de neurônio artificial e, posteriormente, nas Redes Neurais Artificiais.

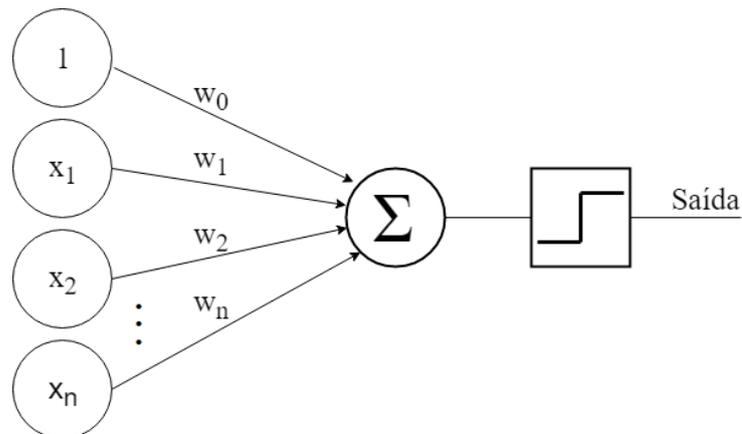
2.4.1 História

O pioneirismo na era das Redes Neurais é atribuído a McCulloch e Pitts (1943), um neurofisiologista e um matemático, respectivamente. Em seu trabalho, eles escreveram um artigo sobre como os neurônios poderiam funcionar e modelaram uma rede neural simples utilizando circuitos elétricos e lógica de limiar.

Em Hebb (1949) foi descrito um sistema de aprendizagem por correlação dos neurônios. Os caminhos neurais são fortalecidos sempre que são usados. Hebb argumentou que a ativação simultânea de células favorece um crescimento na força sináptica.

Inspirado no trabalho de McCulloch e Pitts (1943), Rosenblatt (1957) criou o Perceptron. O Perceptron opera somando os produtos das entradas e os pesos associados, adiciona um limite (bias) e resulta em um de dois valores possíveis (1 ou 0) (Figura 2.8).

Figura 2.8 – Representação de um Perceptron.



Fonte: do autor.

Na década de 1960, criou-se um otimismo exacerbado, pensava-se que as RNAs poderiam executar qualquer tarefa. Entretanto, Minsky e Papert (1969) demonstraram matematicamente a limitação do Perceptron em resolver problemas não separáveis linearmente. À época, Minsky e Papert não acreditavam ser possível construir um método de treinamento para redes com mais de uma camada, e concluíram que as redes neurais seriam sempre impactadas por essa limitação. Isso acabou por desmotivar os pesquisadores e os estudos nessa área reduziram drasticamente nos anos seguintes.

No trabalho de Rumelhart, Hinton e Williams (1986) surgiu o algoritmo de retropropagação (*backpropagation*). Eles mostraram ser possível treinar de forma eficiente redes com camadas intermediárias, resultando nas redes MLP. A presença de novas camadas de neurônios possibilitou o aumento da complexidade dos problemas que podem ser solucionados com Redes Neurais Artificiais. Este princípio é aplicado atualmente em aprendizagem profunda (*deep learning*), possibilitando a resolução de problemas cada vez mais complexos à medida que mais camadas são adicionadas (JONES, 2017).

2.4.2 Arquitetura

Segundo Haykin (2001), o modo de estruturação dos neurônios de uma Rede Neural Artificial tem forte relação com o algoritmo de aprendizagem a ser utilizado em seu treinamento. Ainda segundo Haykin (2001), em geral é possível identificar três classes fundamentais de arquiteturas: redes *feed-forward* de camada única, redes *feed-forward* de múltiplas camadas e redes recorrentes.

2.4.2.1 Redes *feed-forward* de camada única

Esta é a forma mais simples de uma rede em camadas. Nesta arquitetura, uma camada de entrada é diretamente associada aos neurônios que vão gerar a resposta de saída (PALMIERE, 2016).

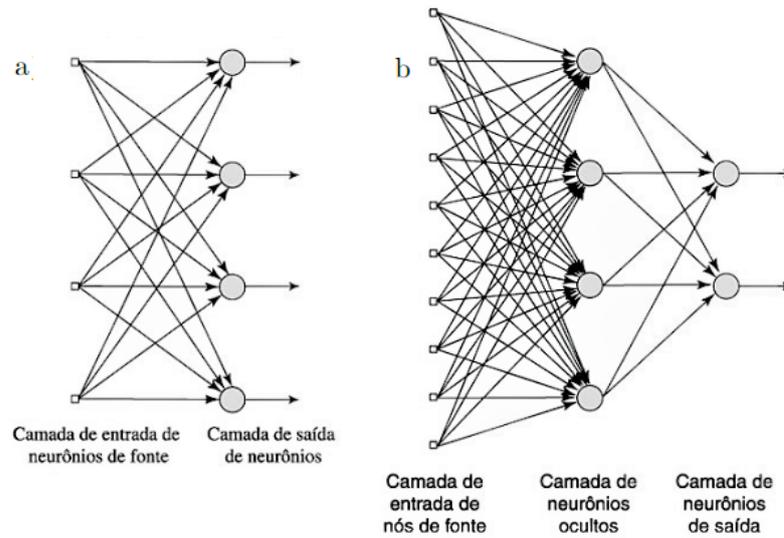
O fluxo de dados é estritamente em direção à camada de saída, por isso é chamada de rede *feed-forward*. Segundo Haykin (2001), esta arquitetura é chamada de camada única, pois a camada de entrada não é contada, uma vez que nela não é realizada qualquer computação. A Figura 2.9a mostra um exemplo de RNA do tipo *feed-forward* de camada única.

2.4.2.2 Redes *feed-forward* de múltiplas camadas

As redes de múltiplas camadas possuem uma ou mais camadas escondidas, também conhecidas como camadas intermediárias ou ocultas, o que resulta aumento de seu poder de processamento. A saída de cada camada intermediária é utilizada como entrada para a camada seguinte (HAYKIN, 2001).

As redes *feed-forward* de multicamadas geralmente são treinadas com o algoritmo de retropropagação do erro (*backpropagation*). Essas redes são empregadas em problemas de aproximação de funções, classificação de padrões, otimização, controle de processos, identificação de sistemas, entre outras aplicações. Exemplos desse tipo de rede são as redes MLP e Redes de Base Radial (*Radial Basis Function - RBF*). A Figura 2.9b exemplifica uma RNA *feed-forward* de múltiplas camadas para o caso com apenas uma camada escondida.

Figura 2.9 – RNAs *feed-forward*: a) de camada única b) de múltiplas camadas.

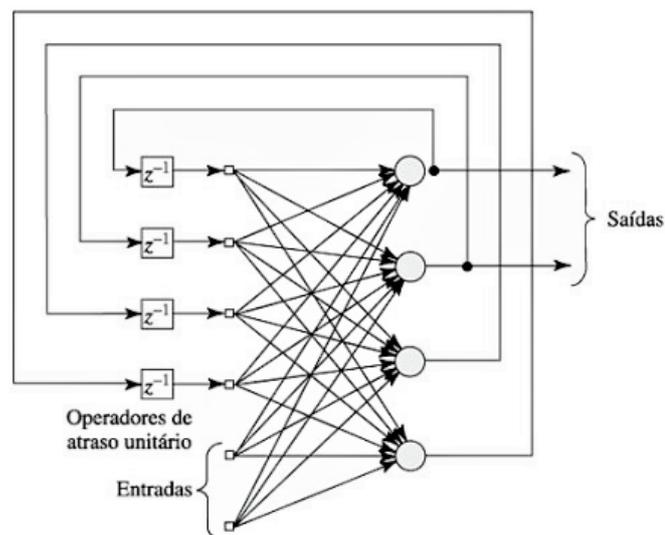


Fonte: adaptado de Haykin (2001).

2.4.2.3 Redes recorrentes

A distinção das redes recorrentes em relação às redes *feed-forward* dá-se pela presença de pelo menos um laço de recorrência (*feedback*). Segundo Haykin (2001), os laços de recorrência impactam profundamente no aprendizado e desempenho da rede. Os laços de recorrência envolvem ainda a utilização de ramos particulares, sendo compostos por unidades de atraso unitário (z^{-1}), resultando em um comportamento dinâmico não-linear, como mostra a Figura 2.10.

Figura 2.10 – RNA recorrente.



Fonte: Haykin (2001).

2.4.3 Tipos de aprendizado

A capacidade de aprendizagem é a habilidade mais importante das RNAs. O aprendizado dá-se, basicamente, por meio de um processo iterativo de ajustes dos pesos e limiares (bias), denominado treinamento.

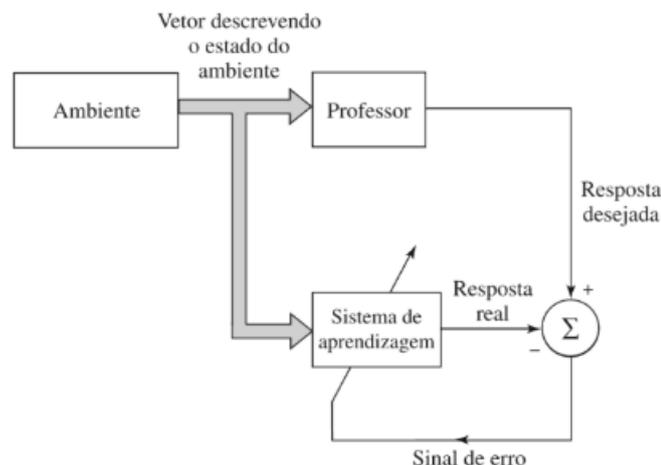
De acordo com Haykin (2001), o treinamento é o processo no qual os parâmetros livres (pesos e limiares) de uma RNA são ajustados por meio de um mecanismo de apresentação de estímulos fornecidos. Os principais modelos de aprendizagem são: supervisionada, não-supervisionada e por reforço.

2.4.3.1 Aprendizado supervisionado

No aprendizado supervisionado, um agente, ou professor, indica para a RNA a resposta desejada para determinado padrão de entrada, ou seja, a rede consegue medir a diferença entre seu comportamento atual e o comportamento desejado (OSÓRIO; VEIRA, 1999). Dessa forma, os parâmetros são ajustados sob a influência dos dados de treinamento e do sinal de erro.

O ajuste é realizado de forma iterativa, compartilhando o conhecimento do agente com a RNA. O término do treinamento dá-se quando a rede é capaz de emular o agente, tornando possível dispensá-lo, uma vez que a partir desse momento ela consegue lidar de forma autônoma com os problemas para os quais foi treinada (HAYKIN, 2001). A Figura 2.11 mostra o diagrama do aprendizado supervisionado.

Figura 2.11 – Diagrama do aprendizado supervisionado.



Fonte: Haykin (2001).

A grande desvantagem deste tipo de aprendizagem se deve ao fato de que a rede não aprenderá novas estratégias para lidar com situações que não estavam presentes no conjunto de treinamento. Por exemplo, uma rede treinada para o reconhecimento de imagens de carros e barcos não seria capaz de reconhecer imagens de aviões, uma vez que não havia imagens de aviões em seu conjunto de treinamento.

Existem diversos algoritmos de otimização para a realização de treinamento supervisionado, dentre eles (MATHWORKS, c2020):

- *Backpropagation* padrão (BP);
- Método do gradiente (GRAD);
- Gradiente conjugado (GC);
- Gradiente conjugado escalonado (SCG);
- Levenberg-Marquardt (LM);
- Broyden-Fletcher-Goldfarb-Shanno (BFGS);
- *Resilient backpropagation* (Rprop);
- *One-step Secant* (OSS).

2.4.3.2 Aprendizado não-supervisionado

O aprendizado não-supervisionado (auto-organização) acontece quando não é apresentada para a RNA a resposta desejada para determinado padrão de entrada, ou seja, os dados não são rotulados. Neste caso, somente os padrões de entrada estão disponíveis. A rede deve se adaptar às regularidades estatísticas dos dados de entrada, identificando e categorizando (*cluster*) esses dados (BECKER, 1991).

2.4.3.3 Aprendizado por reforço

Por fim, no aprendizado por reforço, um avaliador externo avalia a resposta fornecida pela rede. Durante este tipo de aprendizado, ao gerar um sinal de saída, a RNA recebe um sinal de reforço, indicando a qualidade de sua resposta.

2.4.4 Funções de ativação

Os neurônios são caracterizados pelos pesos, bias e função de ativação. A função de ativação processa a combinação linear resultante da entrada pelos pesos das sinapses, gerando então o sinal de saída do neurônio. As funções de ativação são de extrema importância, pois são elas que decidem se um neurônio deverá ser ativado ou não, isto é, se a informação processada é relevante ou não (DATA SCIENCE ACADEMY, 2019). A Equação 2.3 descreve o funcionamento das funções de ativação:

$$y = \text{ativação} \left(\sum (\text{pesos} \times \text{entrada}) + \text{bias} \right). \quad (2.3)$$

As funções de ativação são responsáveis pelas transformações não lineares realizadas ao longo do sinal de entrada. Na ausência de uma função de ativação, os pesos fariam somente uma simples transformação linear, uma rede neural seria apenas um modelo de regressão linear.

Existem várias funções de ativação, cada uma recomendada para determinadas situações. As mais importantes são descritas a seguir.

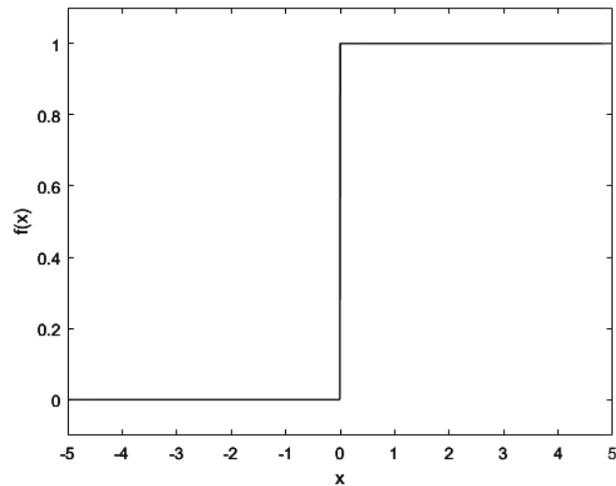
2.4.4.1 Limiar (binária)

Com esta função de ativação, a saída do neurônio é igual a 0 quando seu valor for negativo, e igual a 1 quando seu valor for positivo. A função limiar é descrita pela Equação 2.4.

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}. \quad (2.4)$$

Esta função de ativação é extremamente simples. Seu uso é justificado quando se quer um classificador binário, ou seja, quando é preciso apenas dizer “sim” ou “não” sobre uma única classe. Entretanto, isso não acontece na maioria dos casos, pois geralmente é desejável classificar os dados em várias classes, o que faz com que essa função de ativação raramente seja utilizada (SHARMA V, 2017). Além disso, o gradiente dessa função é zero, o que a torna ineficiente para treinamentos com algoritmos que dependam do gradiente. A função de ativação limiar pode ser vista na Figura 2.12.

Figura 2.12 – Função de ativação limiar.



Fonte: do autor.

2.4.4.2 Linear

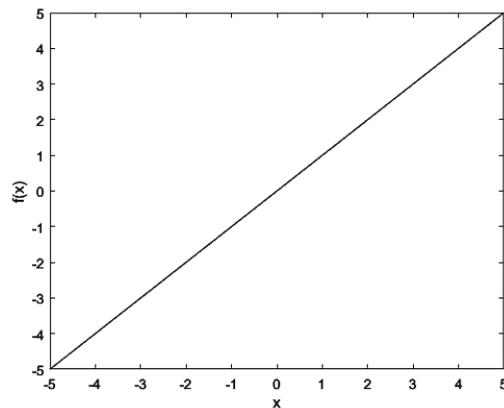
A função de ativação linear cria um sinal de saída proporcional ao sinal de entrada. De certo modo, essa função é melhor que uma função de ativação limiar, uma vez que permite várias saídas e não apenas 0 ou 1. A função de ativação linear é dada pela Equação 2.5.

$$f(x) = cx, \quad (2.5)$$

em que c é uma constante ($c \in \mathbb{R}$).

Existem alguns problemas com a função de ativação linear. Não importa quantas camadas componham a rede neural, se todas elas possuírem função de ativação linear, a função de ativação da última camada ainda será uma função linear da entrada da primeira camada. Além disso, a derivada dessa função é constante, não dependendo da entrada. Isso significa que ocorre uma descida constante do gradiente, o que faz com que a rede não aprenda com a utilização de algoritmos de treinamento que dependam do gradiente (MAHMOOD, 2018). A função de ativação linear é mostrada na Figura 2.13.

Figura 2.13 – Função de ativação linear.



Fonte: do autor.

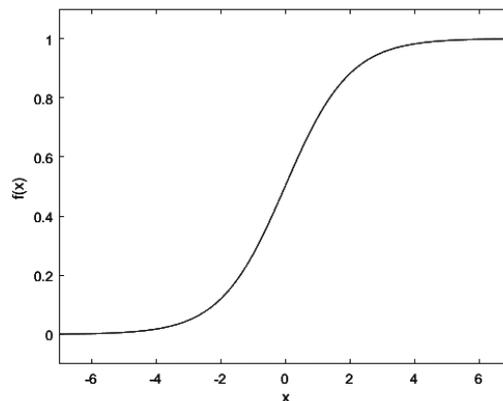
2.4.4.3 Sigmoide

A função de ativação sigmoide é amplamente utilizada. Essa função é não-linear, diferenciável continuamente e gera valores de saída entre 0 e 1. A função sigmoide é descrita pela Equação 2.6.

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2.6)$$

Os problemas com a função sigmoide ocorrem quando os gradientes tornam-se muito pequenos, o que atrapalha seu aprendizado. Outro problema com essa função é que os valores variam apenas entre 0 e 1, e nem sempre é desejável que os valores propagados para outros neurônios sejam apenas positivos (DATA SCIENCE ACADEMY, 2019). A função sigmoide é mostrada na Figura 2.14.

Figura 2.14 – Função de ativação sigmoide.



Fonte: do autor.

Nota-se que na função sigmoide, para os valores de x entre -2 e 2, há grande inclinação em y , o que indica que pequenas alterações nos valores de x nessa região refletem em mudanças significativas nos valores y . Entretanto, em suas extremidades, os valores de y tendem a responder menos às mudanças em x , o que indica que o gradiente nessa região é pequeno e, conseqüentemente, a rede terá aprendido muito lento ou deixará de aprender (MAHMOOD, 2018). Este problema é denominado gradiente de fuga.

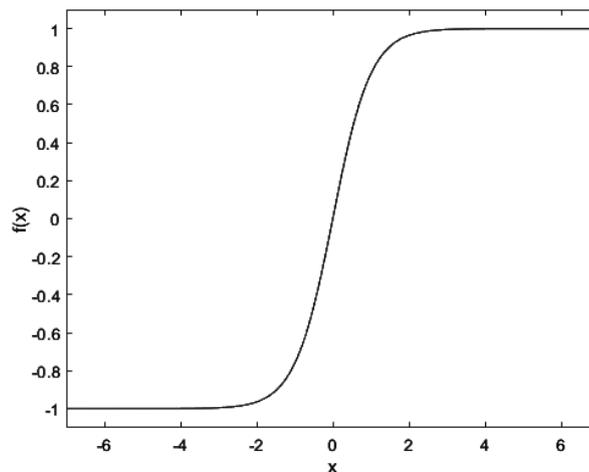
2.4.4.4 Tangente hiperbólica

A função tangente hiperbólica é uma versão ampliada da função sigmoide. Enquanto na função sigmoide os valores são limitados entre 0 e 1, na função tangente hiperbólica os valores são limitados entre -1 e 1 (MAHMOOD, 2018). A função tangente hiperbólica é descrita pela Equação 2.7.

$$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (2.7)$$

Esta função de ativação possui maior gradiente que a sigmoide quando x está entre -2 e 2, ou seja, sua derivada tem maior inclinação nessa região. Entretanto, tal como a função de ativação sigmoide, a tangente hiperbólica também apresenta o problema de gradiente de fuga (SHARMA V, 2017). A Figura 2.15 mostra a função de ativação tangente hiperbólica.

Figura 2.15 – Função de ativação tangente hiperbólica.



Fonte: do autor.

2.4.4.5 ReLu

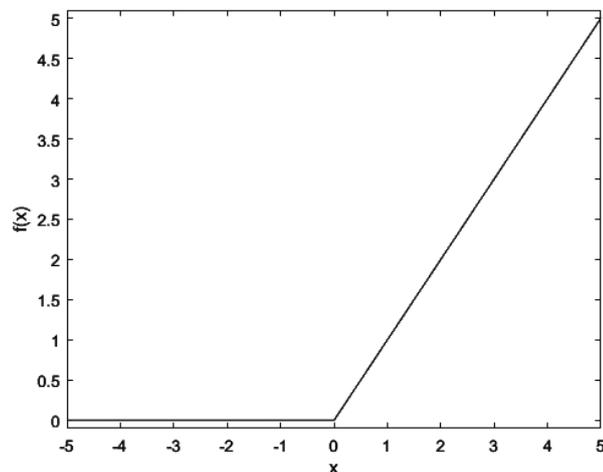
A função de ativação Unidade Linear Retificada (*Rectified Linear Unit - ReLu*) é uma das funções de ativação mais utilizadas atualmente, principalmente em modelos com muitas camadas escondidas (*deep learning*). A ReLu é definida pela Equação 2.8.

$$f(x) = \max(0, x). \quad (2.8)$$

Essa função retorna o valor 0 para qualquer entrada negativa, mas para valores positivos, ela funciona como uma função linear. Entretanto, a ReLu é uma função não-linear, o que torna possível sua utilização com algoritmos que dependam do gradiente (MAHMOOD, 2018).

Ao contrário das funções sigmoide e tangente hiperbólica, na ReLu não existe o problema de gradiente de fuga, já que sua derivada é 0 para valores negativos e 1 para valores positivos. Além disso, a ReLu possui menor custo computacional se comparada a essas funções (MAHMOOD, 2018). A Figura 2.16 mostra a função ReLu.

Figura 2.16 – Função de ativação ReLu.



Fonte: do autor.

A principal vantagem da ReLu é que ela não ativa todos os neurônios ao mesmo tempo. A utilização das funções sigmoide ou tangente hiperbólica em grandes RNAs faz com que a ativação seja densa, uma vez que quase todos os neurônios poderiam ser disparados para descrever a saída da rede. Porém, devido à característica de saída 0 para valores negativos de entrada, a utilização da ReLu faz com que menos neurônios disparem ao mesmo tempo, tornando a ativação esparsa e, conseqüentemente, demandando menor

custo computacional (DATA SCIENCE ACADEMY, 2019). Em contrapartida, esta propriedade da função faz com que seu gradiente seja 0 para entradas negativas. Isto significa que para ativações nessa região da ReLu, não é possível executar a retropropagação dos erros, pois o gradiente é nulo, e os pesos não são ajustados (SHARMA V, 2017).

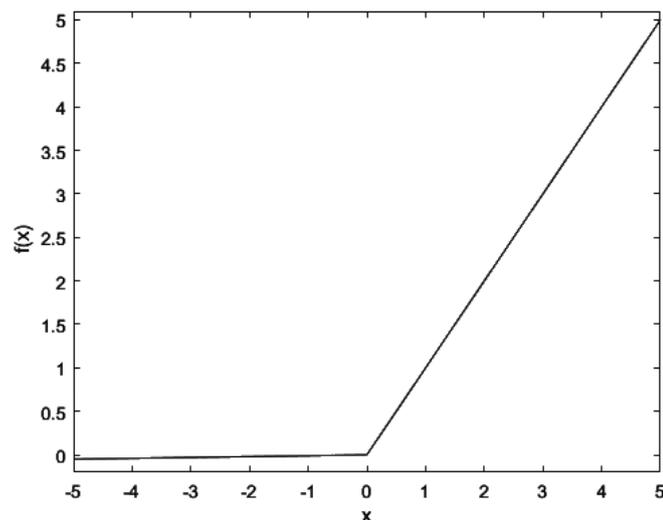
2.4.4.6 Leaky ReLu

A função Leaky ReLu é apenas uma variante da ReLu. Na função de ativação ReLu, o gradiente é 0 para entradas negativas. A Leaky ReLu corrige esse problema ao se definir um pequeno componente linear de x na região de entradas negativas, conforme Equação 2.9.

$$f(x) = \begin{cases} \alpha x, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (2.9)$$

A Leaky ReLU tem uma pequena inclinação positiva, definida de forma arbitrária (geralmente 0,01), na região negativa de entrada ($x < 0$), possibilitando a retropropagação do erro mesmo nessa região (DATA SCIENCE ACADEMY, 2019), como mostra a Figura 2.17.

Figura 2.17 – Função de ativação Leaky ReLu.



Fonte: do autor.

2.4.4.7 Softmax

A *Softmax*, idealizada por Bridle (1990), é uma função de ativação utilizada quando é desejável construir um classificador de várias classes. De acordo com MIKULSKI (2019), essa função é utilizada na camada de saída das RNAs, transformando o processamento efetuado pela rede em valores interpretáveis pelos humanos. A função *Softmax* é dada pela Equação 2.10.

$$\sigma(z_j) = softmax(j) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \text{ para } j = 1, \dots, K, \quad (2.10)$$

em que z_j é o j -ésimo *logit* e K é a quantidade de neurônios presentes na camada de saída.

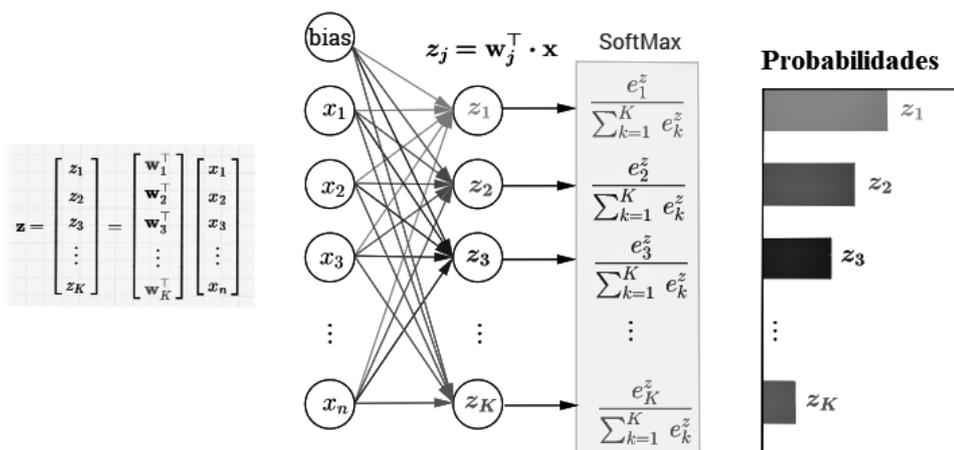
Os *logits* são os valores brutos resultantes da multiplicação dos pesos pelos valores de entrada da última camada, antes da aplicação de uma função de ativação, e são definidos pela Equação 2.11.

$$z_j = w_j^T \cdot x, \quad (2.11)$$

em que w_j^T é o vetor de pesos associados ao j -ésimo neurônio e x o vetor de entradas.

A função *Softmax* transforma as saídas para cada classe em valores entre 0 e 1, com a soma desses valores totalizando 1, o que pode-se interpretar como a probabilidade da entrada pertencer a determinada classe. A Figura 2.18 representa este processo.

Figura 2.18 – Funcionamento da função de ativação *softmax*.

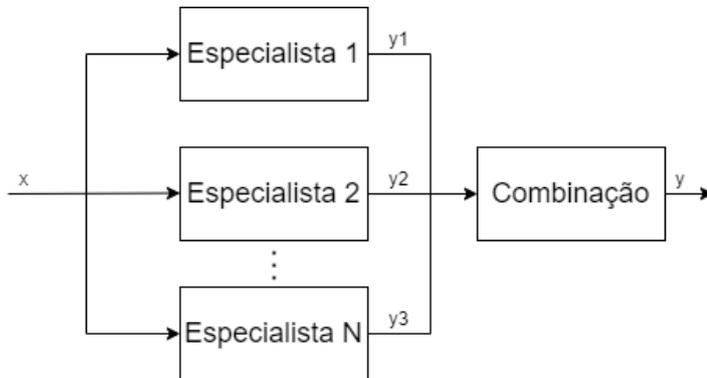


Fonte: adaptado de rinterested.github.io/statistics/softmax.html

2.5 Comitê de máquinas

De acordo com Haykin (2001), comitês de máquinas utilizam a abordagem de dividir e conquistar, ou seja, tarefas computacionais de grande complexidade são resolvidas dividindo-as em tarefas de menor complexidade, e então combinando as soluções destas últimas. Tratando-se da aprendizagem supervisionada, a simplicidade computacional é atingida ao se distribuir a aprendizagem entre um número de especialistas. À essa combinação de especialistas dá-se o nome de comitês de máquinas. Elas conciliam o conhecimento assimilado pelos especialistas para obter uma decisão global, supostamente melhor que a decisão alcançável por qualquer especialista individualmente (Figura 2.19).

Figura 2.19 – Esquema de um comitê de máquinas.



Fonte: do autor.

Ainda segundo Haykin (2001), os comitês de máquinas são aproximadores universais, podendo ser classificadas em duas grandes categorias:

1. Estruturas estáticas, onde as respostas de vários especialistas são combinadas por meio de um mecanismo, sem que se envolva o sinal de entrada. Os seguintes métodos estão inclusos nesta categoria:
 - Média de *ensemble*: as saídas são combinadas linearmente, produzindo uma saída global.
 - Reforço: um algoritmo fraco de aprendizagem¹ é transformado em um que alcança alta precisão.

¹ Algoritmo de acurácia pouco superior a de um sistema aleatório, isto é, pouco superior a 50%.

2. Estruturas dinâmicas, onde o sinal de entrada é envolvido na integração das saídas dos especialistas individuais em uma saída global. Fazem parte desta categoria os métodos:

- Mistura de especialistas: as respostas dos especialistas são combinadas, de forma não-linear, por meio de uma única rede de passagem.
- Mistura hierárquica de especialistas: as respostas dos especialistas são combinadas, também de forma não-linear, por meio de redes de passagens dispostas de forma hierárquica.

2.5.1 Geração de componentes

Segundo Lima (2004), a principal razão para a utilização de comitês de máquinas é propiciar um incremento de desempenho. Dessa forma, a geração de componentes para o comitê possui grande relevância. Os requisitos principais referentes aos componentes são:

1. Os componentes devem apresentar bom desempenho isoladamente.
2. O comportamento dos componentes deve ser o mais descorrelacionado possível, a fim de conferir diversidade ao comitê.

Ainda de acordo com Lima (2004), os resultados de maior expressão encontrados na literatura, no que diz respeito a geração de componentes, empregam RNAs como componentes do comitê. A geração de RNAs com diversidade pode se dar das seguintes formas (LAI et al., 2006):

1. Diferentes inicializações de pesos para cada uma das RNAs que compõem o comitê.
2. Variação da arquitetura das RNA, variando a quantidade de neurônios ou mesmo a quantidade de camadas.
3. Utilização de diferentes algoritmos de treinamento.
4. Treinamento das RNAs com diferentes subconjuntos de treinamento.

No que diz respeito ao treinamento com diferentes subconjuntos, o objetivo é que haja conjuntos distintos, capazes de conduzir as RNAs à diversidade de generalização,

mesmo que se mantenham fixos os parâmetros das mesmas (LIMA, 2004). As principais formas de obtenção desses conjuntos são:

- **Conjuntos disjuntos:** os conjuntos de treinamento devem ser disjuntos, ou seja, amostras sem repetição de um conjunto original (SHARKEY; SHARKEY; CHANDROTH, 1995). Dessa forma, não há qualquer sobreposição de dados utilizados para o treinamento de diferentes componentes. O problema é que, se o conjunto original não for suficientemente grande, o tamanho dos conjuntos de treinamento será reduzido, o que pode resultar em baixo desempenho de generalização (TUMER; GHOSH, 1996).
- **Reamostragem:** consiste em obter vários subconjuntos de treinamento distintos a partir de um único conjunto de treinamento, sendo a abordagem *bagging*, introduzida por Breiman (1996), uma das mais utilizadas. Nesta abordagem, diversos conjuntos de treinamento são gerados a partir da amostragem uniforme, com reposição, do conjunto de treinamento original. Dessa forma, se obtém uma solução a partir de cada um desses conjuntos de treinamento. Como a amostragem é com reposição, apesar de os conjuntos de treinamento possuírem a mesma quantidade de amostras do conjunto original, algumas amostras do conjunto original podem aparecer mais de uma vez em cada conjunto de treinamento, fazendo com que algumas amostras não sejam selecionadas. Essa diferença entre cada conjunto de treinamento concede diversidade aos modelos obtidos por meio desses conjuntos de treinamento.
- **Boosting adaptativo (AdaBoost):** abordagem proposta por Freund e Schapire (1996), é uma variação da abordagem *boosting* de Schapire (1990), mesclando ideias dessa última e da abordagem *bagging*. Os conjuntos de treinamento são reamostrados de forma adaptativa, fazendo com que as amostras que mais contribuíram para o erro de treinamento dos componentes já treinados possuam maior probabilidade de comporem os conjuntos de treinamento dos próximos componentes a serem treinados. Dessa forma, apenas o primeiro componente é treinado a partir de uma amostragem uniforme do conjunto de dados original e os demais componentes são treinados sequencialmente, a fim de redefinir a probabilidade de seleção das amostras na geração dos próximos conjuntos de treinamento.

2.5.2 Métodos de combinação

Existem diversos modos de combinar as saídas dos componentes dos comitês. Os mais comuns são listados a seguir.

- **Combinações lineares:** médias simples e ponderada são os métodos mais populares para a combinação das saídas dos componentes. A saída global (resposta do comitê) pode ser criada a partir de um conjunto de saídas fornecidas pelos componentes do comitê (PERRONE; COOPER, 1993), como descrito na Equação 2.12.

$$y_c = \sum_{i=1}^n w_i \times y_i, \quad (2.12)$$

onde y_c é a resposta do comitê, n é o número de membros do comitê, w_i é o peso associado a cada componente do comitê e y_i é a resposta individual de cada componente do comitê. Se todos os membros do comitê possuírem mesmo peso, então a combinação torna-se uma média simples.

- **Métodos de combinação não-linear:** as saídas dos componentes do comitê são relacionadas de forma não-linear (Equação 2.13). Essas relações podem ser realizadas por meio da implementação de diversas técnicas de inteligência artificial, tais como redes neurais, algoritmos genéticos e lógica *fuzzy* (PALIT; POPOVIC, 2000).

$$y_c = \psi(y_1, y_2, y_3, \dots, y_n), \quad (2.13)$$

onde y_c é a resposta do comitê, $y_1, y_2, y_3, \dots, y_n$ são as saídas individuais de cada componente e $\psi(y_1, y_2, y_3, \dots, y_n)$ é uma função não-linear.

- **Votação:** o resultado que obtiver apoio da maioria dos componentes é definido como a saída da máquina de comitê. Como cada rede possui mesmo peso de votação, um componente com desempenho ruim na máquina pode acabar por diminuir significativamente o desempenho do comitê. Existem algumas variações desse método (LIMA, 2004), listadas a seguir:

- **Unanimidade,** onde o comitê decide que determinada observação só pertence a determinada classe somente se todos os seus componentes tomarem a mesma decisão.

- **Unanimidade modificada**, onde o comitê decide que determinada observação só pertence a determinada classe se um certo percentual de componentes tomarem a mesma decisão.
- **Majoritário**, onde o comitê decide que determinada observação só pertence a determinada classe se mais da metade de seus componentes tomarem a mesma decisão.
- **Pluralidade limitada**, onde o comitê decide que determinada observação só pertence a determinada classe se o número de componentes que a classificam nessa classe é maior que o número de componentes que a classificam em qualquer outra classe.

2.6 Métricas de avaliação de desempenho

A partir do desenvolvimento dos sistemas de classificação, torna-se de extrema importância avaliar seus desempenhos. Tal avaliação visa a comparação entre esses sistemas, uma vez que geralmente são construídos com diferentes metodologias e configurações para a solução de determinado problema. Alguns dos critérios de avaliação mais utilizados são descritos a seguir.

2.6.1 Matrizes de confusão

As matrizes de confusão resumem o desempenho da classificação realizada por um sistema classificador, representando visualmente as classificações reais e as previstas pelo sistema (ALI et al., 2017). Essas matrizes mostram as frequências com que um sistema classificou instâncias como pertencentes a determinada classe e as frequências reais dessas instâncias em cada classe.

Nas matrizes de confusão, as linhas geralmente representam as classificações realizadas por um sistema, e as colunas os dados reais (referência). O resultado é uma matriz quadrada $k \times k$, em que k é a quantidade de classes. Na diagonal principal ficam as instâncias que foram corretamente classificados pelo sistema. A coluna marginal (n_{i+}) representa a quantidade total de instâncias classificadas como pertencentes à classe i e a linha marginal (n_{+j}) representa a quantidade total de instâncias verdadeiramente pertencentes à classe i (PASSO, 2013). A Figura 2.20 representa uma matriz de confusão.

Figura 2.20 – Esquema de uma matriz de confusão.

		j = Colunas (referência)			Total linhas
		1	2	k	n_{i+}
i = Linhas (classificação)	1	n_{11}	n_{12}	n_{1k}	n_{1+}
	2	n_{21}	n_{22}	n_{2k}	n_{2+}
	k	n_{k1}	n_{k2}	n_{kk}	n_{k+}
Total colunas n_{+j}		n_{+1}	n_{+2}	n_{+k}	n

Fonte: adaptado de Congalton e Green (2009)

As matrizes de confusão relacionam os acertos e erros de um sistema de classificação, mostrando o número de classificações corretas e as classificações preditas para cada classe. Os valores Verdadeiro Positivo, Falso Negativo, Falso Positivo e Verdadeiro Negativo compõe a estrutura das matrizes de confusão (MATSUNAGA, 2007; SOKOLOVA; LAPALME, 2009). A Figura 2.21 mostra a disposição desses valores em uma matriz de confusão, enquanto suas descrições são dadas a seguir:

- Verdadeiro Positivo (*True Positive* - **TP**): indica a quantidade de instâncias classificadas corretamente como pertencentes a uma determinada classe.
- Falso Negativo (*False Negative* - **FN**): indica a quantidade de instâncias classificadas erroneamente como não pertencentes a determinada classe.
- Falso Positivo (*False Positive* - **FP**): indica a quantidade de instâncias classificadas erroneamente como pertencentes a determinada classe.
- Verdadeiro Negativo (*True Negative* - **TN**): indica a quantidade de instâncias classificadas corretamente como não pertencentes a determinada classe.

Figura 2.21 – Composição de uma matriz de confusão.

		Classe esperada (real)	
		Classe A	Classe B
Classe prevista	Classe A	Verdadeiro positivo (TP)	Falso positivo (FP)
	Classe B	Falso negativo (FN)	Verdadeiro negativo (TN)

Fonte: do autor

Por meio das matrizes de confusão é possível extrair diversas métricas que permitem avaliar o sistema de classificação utilizado (MATSUNAGA, 2007; OLSON; DELEN, 2008; SWALIN, 2018). Essas métricas são descritas a seguir:

- Acurácia: indica a proporção de instâncias classificadas corretamente.
- Precisão: indica a proporção de verdadeiros positivos dentre todas as instâncias classificadas como positivas.
- Sensibilidade, *Recall* ou TPR (*True Positive Rate*): indica qual a proporção de verdadeiros positivos corretamente classificados.
- Especificidade ou TNR (*True Negative Rate*): indica qual a proporção de verdadeiros negativos corretamente classificados.
- *F-score*: é uma maneira de combinar precisão e *recall* em um único número. O escore F_1 é calculado usando a média harmônica entre precisão e *recall*.

Os problemas multi-classes, isto é, quando há mais de duas classes, resultam em matrizes de confusão maiores e mais complicadas (OLSON; DELEN, 2008). Nestes casos, as métricas utilizadas são as mesmas de uma classificação binária, entretanto, as métricas são calculadas para cada classe. Dessa forma, o problema é tratado como classificação binária do tipo OvA (*One vs All*), comparando uma classe contra o agrupamento de todas as outras (AMAZON, 2020). Tschorn (2019) e Shmueli (2019) descrevem como as métricas referentes a cada classe podem ser combinadas por meio de médias:

- Média macro: dá-se peso igual para todas as classes presentes. Na média macro, somam-se as métricas por classe e então divide-se a soma pela quantidade de classes.

- Média macro ponderada: semelhante à media macro, porém pondera-se pela quantidade de instâncias em cada classe. É a soma da métrica por classe multiplicada pela frequência relativa da classe. Classes com maior quantidade de instâncias tornam-se mais significativas.
- Média micro: dá-se peso igual para cada amostra. Classes com maior quantidade de instâncias tendem a dominar o cálculo desta média. Na média micro, a precisão é sempre igual ao *recall*, já que um falso positivo (FP) para uma classe é sempre um falso negativo (FN) para outra classe. Além disso, precisão e *recall* também são iguais à acurácia, pois resumem-se à soma da diagonal principal dividida pelo total de instâncias.

A Tabela 2.2 resume os cálculos das métricas acurácia, precisão, *recall*, especificidade e F_1 -score para as classificações binárias e multi-classes utilizando a média micro e a média macro. A média macro ponderada, por se tratar apenas de uma variação da média macro, não foi adicionada à tabela.

Tabela 2.2 – Medidas de desempenho para os casos de classificação binária e multi-classes.

Medida	Classificação binária	Multi-classe média micro (μ)	Multi-classe média macro (M)
Acurácia	$\frac{TP+TN}{TP+FP+TN+FN}$	$\frac{C}{\sum_{i=1}^C TP_i/N}$	$\frac{C}{\sum_{i=1}^C TP_i/N}$
Precisão (Pr)	$\frac{TP}{TP+FP}$	$\frac{\sum_{i=1}^C TP_i}{\sum_{i=1}^C (TP_i+FP_i)}$	$\frac{\sum_{i=1}^C \frac{TP_i}{TP_i+FP_i}}{C}$
<i>Recall</i> (Re)	$\frac{TP}{TP+FN}$	$\frac{\sum_{i=1}^C TP_i}{\sum_{i=1}^C (TP_i+FN_i)}$	$\frac{\sum_{i=1}^C \frac{TP_i}{TP_i+FN_i}}{C}$
Especificidade	$\frac{TN}{FP+TN}$	$\frac{\sum_{i=1}^C TN_i}{\sum_{i=1}^C (FP_i+TN_i)}$	$\frac{\sum_{i=1}^C \frac{TN_i}{FP_i+TN_i}}{C}$
F_1 -score	$\frac{2 \times Pr \times Re}{Pr + Re}$	$\frac{2 \times Pr_\mu \times Re_\mu}{Pr_\mu + Re_\mu}$	$\frac{2 \times Pr_M \times Re_M}{Pr_M + Re_M}$

TP - *True Positive*; TN - *True Negative*; FP - *False Positive*; FN - *False Negative*; C - quantidade de classes; N - quantidade total de instâncias.

Fonte: do autor.

Em sua forma geral, o F -score é dado pela Equação 2.14.

$$F_\beta = \frac{(1 + \beta^2) \times \text{Precisão} \times \text{Recall}}{\beta^2 \times \text{Precisão} + \text{Recall}}, \quad (2.14)$$

em que β define o grau de importância da precisão em relação ao *recall* (SOKOLOVA; LAPALME, 2009). Quando $\beta = 1$, precisão e *recall* possuem mesma importância; quando

$\beta = 2$, *recall* tem o dobro de importância que precisão; quando $\beta = 0,5$, precisão tem o dobro de importância que *recall* (MATSUNAGA, 2007). Isso é útil, pois em alguns casos é preferível dar mais importância à precisão ou ao *recall*, a depender do trabalho conduzido.

2.6.2 Análise ROC

A análise ROC é uma técnica visual que permite avaliar e selecionar classificadores com base em seu desempenho (FAWCETT, 2006). Acredita-se que a análise ROC tenha surgido durante a Segunda Guerra Mundial como uma forma de auxiliar a análise de sinais provenientes de radares. A análise ROC tem sido utilizada em diversas áreas como medicina, radiologia, ciências sociais e aprendizado de máquina (MATSUBARA, 2008).

Considerando um sistema classificador e um conjunto de instâncias, cada uma das instâncias é classificada como positiva (P) ou negativa (N), e se encaixa em uma das quatro situações possíveis: verdadeiro positivo (TP), verdadeiro negativo (TN), falso positivo (FP) ou falso negativo (FN), compondo uma matriz de confusão de classificação binária, tamanho 2x2.

A partir dessa matriz de confusão, obtêm-se a Taxa de Verdadeiros Positivos (*True Positive Rate* - **TPR**)² e a Taxa de Falsos Positivos (*False Positive Rate* - **FPR**)³, que servem como base para a construção do espaço ROC (OLSON; DELEN, 2008).

2.6.2.1 Espaço ROC

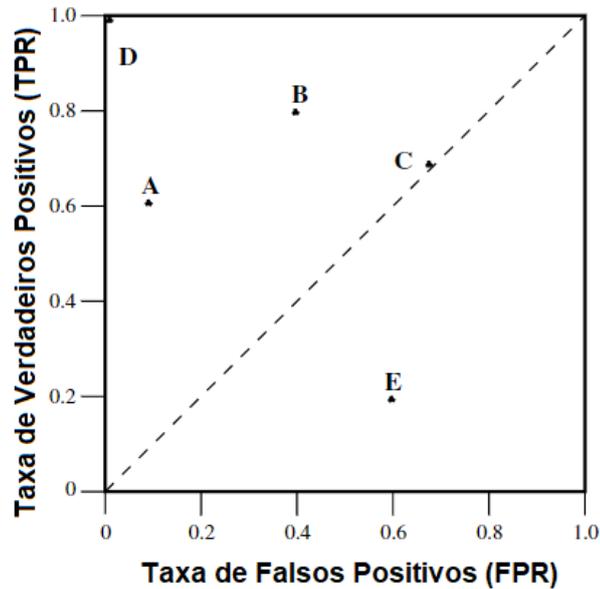
O espaço ROC é um quadrado de lados iguais a 1. Os gráficos ROC têm duas dimensões, em que nos eixos X e Y são colocadas as taxas de falsos e verdadeiros positivos, respectivamente. Essencialmente, um gráfico ROC representa o *trade-off* relativo entre os benefícios, representados pelos verdadeiros positivos, e os custos, representados pelos falsos positivos (OLSON; DELEN, 2008).

A Figura 2.22 mostra um gráfico ROC com exemplos de classificadores discretos, como árvores de decisão e conjunto de regras. Este tipo de classificadores tem como saída somente uma classe para cada instância, fornecendo um par (FPR, TPR) que corresponde a um ponto no espaço ROC.

² **TPR** também é chamada de Sensibilidade ou *Recall*.

³ **FPR** é equivalente ao complemento da Especificidade ($FPR = 1 - \text{Especificidade}$).

Figura 2.22 – Exemplos de classificadores no espaço ROC.



Fonte: adaptado de Fawcett (2006)

De acordo com Fawcett (2006), existem alguns pontos importantes a serem observados no espaço ROC:

- Ponto inferior esquerdo (0,0): representa a estratégia do classificador nunca emitir uma classificação positiva, o que faz com que ele não cometa erros do tipo falsos positivos, mas também não consegue obter verdadeiros positivos.
- O ponto superior direito (1,1): representa o oposto do anterior. Neste caso, o classificador emite classificações positivas de forma incondicional, o que faz com que ele obtenha todos os verdadeiros positivos, mas cometa muitos erros do tipo falsos positivos.
- O ponto superior esquerdo (0,1): representa uma classificação perfeita. Neste caso, o classificador obterá todos os verdadeiros positivos e não cometerá erros do tipo falsos positivos.

Os classificadores dispostos do lado esquerdo do gráfico ROC são considerados conservadores, uma vez que apenas classifica uma instância como positiva quando há fortes evidências, o que resulta em poucos erros do tipo falsos positivos. Por outro lado, os classificadores dispostos à direita no gráfico ROC são considerados liberais, pois fazem

classificações positivas com pouca evidência, o que resulta em maior quantidade de erros do tipo falsos positivos (OLSON; DELEN, 2008).

A linha diagonal entre os pontos (0,0) e (1,1) representa um classificador aleatório. Dessa forma, qualquer classificador abaixo dessa linha é considerado pior que um classificador aleatório. Entretanto, esses classificadores são capazes de distinguir entre classes, apenas não o fazem corretamente. A negação desse classificador produzirá pontos acima da linha diagonal (FAWCETT, 2006). Na Figura 2.22, por exemplo, o classificador B é equivalente à negação do classificador E.

2.6.2.2 Curvas ROC

Os classificadores contínuos, como as RNAs e *Naive Bayes*, ao contrário dos classificadores discretos, produzem como saída para cada instância uma probabilidade ou *score*, que representa a pertinência da instância à classe. A partir desses classificadores, com a aplicação de um valor limiar, pode-se gerar classificadores discretos (FAWCETT, 2006).

A obtenção da curva ROC dá-se a partir da variação de um limiar de decisão (*threshold*) sobre toda a faixa de *scores* produzida pelo classificador (CASTRO; BRAGA, 2011). O classificador atribui um *score* a cada exemplo e os ordena por seus *scores*. Os exemplos acima dos limiares são classificados como positivos, enquanto os abaixo são classificados como negativos (ISHIDA; POZO, 2007).

Por meio da variação do limiar, determina-se valores para as taxas de verdadeiros positivos e falsos positivos para cada valor do limiar. Dessa forma, tem-se vários pontos no espaço ROC e gera-se uma curva que mostra, de forma gráfica, o *trade-off* entre TPR e FPR (ISHIDA; POZO, 2007; CASTRO; BRAGA, 2011). A construção detalhada de curvas ROC pode ser encontrada em Fawcett (2006) e Matsubara (2008).

As curvas ROC são, basicamente, representações bidimensionais do desempenho de classificadores. Portanto, para comparar classificadores ou julgar o quão adequado é um classificador, reduz-se as curvas ROC a um valor escalar que também represente o desempenho dos classificadores (OLSON; DELEN, 2008).

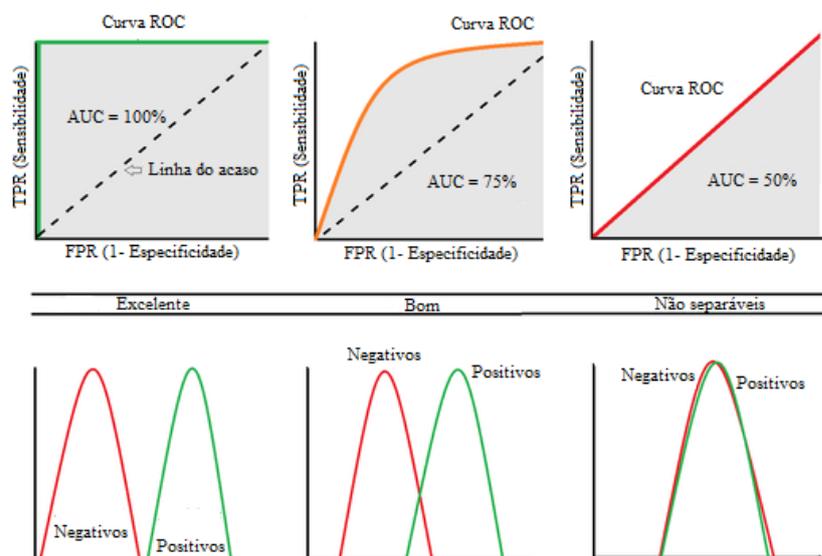
A área sob a curva ROC (*Area Under the ROC Curve* - AUC) é o método mais comum para realizar essa redução (HANLEY; MCNEIL, 1982; BRADLEY, 1997). A AUC é parte do espaço ROC, um quadrado unitário, portanto seu valor sempre estará entre 0 e 1. O classificador perfeito resultaria em AUC de valor 1. A linha diagonal representa

um classificador aleatório e possui AUC de valor 0,5. Dessa forma, nenhum classificador realista deve possuir AUC menor 0,5 (FAWCETT, 2006).

A métrica AUC mede a capacidade do sistema classificador prever um *score* maior para instâncias positivas em comparação com as negativas (AMAZON, 2020). Escolhidas uma instância positiva e uma negativa, a probabilidade de que a instância positiva supere a negativa, isto é, seja ordenada primeiramente em relação ao *score*, é dada pela AUC (SWALIN, 2018). Hanley e McNeil (1982) demonstraram que a AUC é equivalente à estatística de Wilcoxon. Além disso, a AUC também se relaciona com o coeficiente Gini (HAND; TILL, 2001).

A Figura 2.23 mostra três curvas ROC, as distribuições das classes de onde vieram e seus valores de AUC. Nota-se que a primeira curva ROC possui AUC igual a 1, ou seja, o classificador conseguiu classificar corretamente todas as instâncias. Isso é possível pois não há interseção entre as classes. A segunda curva ROC possui AUC de 0,75 e as classes possuem uma pequena interseção, demonstrando o que normalmente acontece em problemas práticos. Por fim, a terceira curva ROC possui AUC de 0,5, desempenho igual a de um classificador aleatório. A distribuição das classes mostra que há uma interseção quase completa entre elas, impossibilitando que o classificador consiga distingui-las.

Figura 2.23 – Exemplos de curvas ROC e distribuição das classes.



Fonte: adaptado de datasciencecentral.com/profiles/blogs/roc-curve-explained-in-one-picture

Os problemas multi-classes resultam em uma matriz de confusão de tamanho $n \times n$, para n classes. Essa matriz contém uma diagonal principal com as n classificações corretas

e $n^2 - n$ possíveis erros. Neste caso não se compara o *trade-off* em relação aos valores de TPR e FPR, mas dos n benefícios e $n^2 - n$ erros. Em um problema com apenas três classes, por exemplo, a superfície torna-se um polítopo de $3^2 - 3 = 6$ dimensões (FAWCETT, 2006).

Nos problemas multi-classes, uma forma de lidar com as n classes é produzir n curvas ROC, utilizando a estratégia *OvA*. Dado um conjunto C com todas as classes, o i -ésimo gráfico ROC refere-se ao desempenho do classificador utilizando a classe c_i como classe positiva (P), como mostra a Equação 2.15, e todas as outras $C - 1$ classes como classes negativas (N), como mostra a Equação 2.16 (FAWCETT, 2006).

$$P_i = c_i \quad (2.15)$$

$$N_i = \bigcup_{j \neq i} c_j \in C \quad (2.16)$$

Nos problemas multi-classes há diferenças também na métrica AUC. Em Provost e Domingos (2000) é apresentada uma alternativa para a generalização multi-classes de AUC. Nesta generalização, calcula-se as AUCs gerando uma curva para cada classe, medindo a área sob a curva e fazendo uma somatório ponderado das AUCs, de acordo com a prevalência de cada classe no conjunto de dados. A Equação 2.17 descreve este cálculo.

$$AUC_{total} = \sum_{c_i \in C} AUC(c_i) \times p(c_i), \quad (2.17)$$

em que $AUC(c_i)$ é a área sob a curva considerando a i -ésima classe como positiva e $p(c_i)$ é a ponderação da prevalência da classe no conjunto de dados. A vantagem dessa alternativa é que a AUC_{total} é gerada diretamente a partir das curvas ROC de cada classe referência. Entretanto, esta alternativa apresenta como desvantagem ser sensível às distribuições de classe e custos de erro (FAWCETT, 2006).

Hand e Till (2001) adotaram uma abordagem diferente para a generalização multi-classes de AUC. Baseada no fato de que a AUC é equivalente à probabilidade do classificador ordenar uma instância positiva escolhida aleatoriamente de forma mais alta que uma instância negativa também escolhida aleatoriamente, eles encontraram uma formulação que mede a discriminação não ponderada de pares de classes. A Equação 2.18 descreve essa formulação.

$$AUC_{total} = \frac{2}{|C|(|C|-1)} \times \sum_{(c_i, c_j) \in C} AUC(c_i, c_j), \quad (2.18)$$

em que $AUC(c_i, c_j)$ é a área abaixo da curva ROC de duas classes, c_i e c_j . A soma refere-se a todos os pares de classes distintas, independentemente da ordem. Existem $|C|(|C|-1)/2$ pares de classes. A formulação de Hand e Till (2001) é insensível à distribuição de classes e aos custos de erro, entretanto, não há maneira fácil para se visualizar a superfície cuja área está sendo calculada (FAWCETT, 2006).

As abordagens de Provost e Domingos (2000) e Hand e Till (2001) são extensões da métrica AUC para problemas multi-classes. Essas extensões tornaram-se populares devido à complexidade de se calcular o Volume Sob a Superfície ROC (*Volume of the Under Surface - VUS*). Entretanto, o VUS foi estudado como métrica de desempenho de problemas multi-classes, como nos trabalhos de Ferri, Hernández-Orallo e Salido (2003) e Landgrebe e Duin (2006).

2.6.3 Índice Kappa de Cohen

O índice Kappa de Cohen (κ), desenvolvido por Cohen (1960), descreve a intensidade de concordância entre dois avaliadores independentes que realizam uma avaliação de uma mesma amostra em que as categorias sejam mutuamente exclusivas.

Os dados referentes às avaliações realizadas são dispostos em tabelas $n \times n$, em que as linhas representam as avaliações de um dos avaliadores e as colunas representam as avaliações do outro avaliador. O índice Kappa de Cohen é descrito pela Equação 2.19 (COHEN, 1960; CONGALTON; GREEN, 2009).

$$\kappa = \frac{p_o - p_e}{1 - p_e} = \frac{\sum_{i=1}^c p_{ii} - \sum_{i=1}^c p_{i+} \times p_{+i}}{1 - \sum_{i=1}^c p_{i+} \times p_{+i}} \quad (2.19)$$

em que κ representa o índice Kappa, p_o e p_e representam, respectivamente, as proporções de concordância observada (acurácia) e esperada ao acaso entre os avaliadores. A proporção p_o é calculada pelo somatório de p_{ii} , que são as proporções presentes na i -ésima linha e i -ésima coluna da tabela, enquanto p_e é calculada pelo somatório da multiplicação entre p_{i+} e p_{+i} , que representam as somas das proporções presentes na i -ésima linha e i -ésima coluna, respectivamente. Finalmente, c é o número de classes, equivalente ao número de linhas e colunas da tabela.

A Tabela 2.3 mostra a disposição das proporções de classificações para o caso com dois avaliadores, A e B, e c classes. A diagonal principal representa as proporções de concordância entre os avaliadores para cada uma das classes.

Tabela 2.3 – Proporções de classificações por dois classificadores com C classes.

Avaliador A	Avaliador B				Total
	1	2	...	c	
1	p_{11}	p_{12}	...	p_{1c}	p_{1+}
2	p_{21}	p_{22}	...	p_{2c}	p_{2+}
⋮	⋮	⋮		⋮	⋮
c	p_{c1}	p_{c2}	...	p_{cc}	p_{c+}
Total	p_{+1}	p_{+2}	...	p_{+c}	1

Fonte: adaptado de Fleiss, Levin e Paik (2003).

Na Equação 2.20 é demonstrado o cálculo de p_o e na Equação 2.21 o cálculo de p_e , ambas para o caso mostrado na Tabela 2.3, com dois avaliadores e c classes.

$$p_o = \sum_{i=1}^c p_{ii} = p_{11} + p_{22} + \dots + p_{cc}. \quad (2.20)$$

$$p_e = \sum_{i=1}^c p_{i+} \times p_{+i} = p_{1+} \times p_{+1} + p_{2+} \times p_{+2} + \dots + p_{c+} \times p_{+c}. \quad (2.21)$$

O índice Kappa pode ser definido como a proporção de concordância entre os avaliadores, excluindo-se a proporção de concordância esperada ao acaso, e varia entre -1 e 1. Valor Kappa próximo a 1 indica concordância perfeita entre os avaliadores, próximo a 0 indica concordância próxima à esperada pelo acaso e abaixo de 0 indica discordância entre os avaliadores (FONSECA; SILVA; SILVA, 2013).

De acordo com Landis e Koch (1977), a concordância entre os avaliadores, utilizando o índice Kappa, pode ser classificado conforme a Tabela 2.4.

Em Cohen (1960) é apresentada uma aproximação para o erro-padrão do índice Kappa, como mostra a Equação 2.22.

$$EP(\kappa) = \sqrt{\frac{p_o(1-p_o)}{N(1-p_e)^2}}. \quad (2.22)$$

Segundo Cohen (1960), a fórmula do erro padrão é uma aproximação, pois trata p_e como uma constante e p_o como um valor populacional. Como p_o não varia muito em

Tabela 2.4 – Intensidade da concordância Kappa

Índice Kappa	Concordância
<0,00	Pobre
0,00-0,20	Desprezível
0,21-0,40	Razoável
0,41-0,60	Moderada
0,61-0,80	Substancial
0,81-1,00	Quase perfeita

Fonte: Landis e Koch (1977).

relação ao valor de Kappa, especialmente para N grande (>100), as considerações são adequadas. Além disso, com N suficientemente grande, a distribuição amostral de κ se aproxima da normalidade, tornando possível o cálculo dos limites de confiança da forma usual.

Entretanto, Fleiss, Cohen e Everitt (1969) ponderam que o cálculo do erro padrão proposto em Cohen (1960) deriva de premissas contraditórias sobre os totais marginais fixos e variação binomial das frequências nas células da matriz de confusão. O cálculo proposto em Cohen (1960) superestima o erro padrão. Fleiss, Cohen e Everitt (1969) propuseram então uma formulação para a variância assintótica de κ , resultando na expressão para o erro padrão descrita na Equação 2.23 (FLEISS; LEVIN; PAIK, 2003).

$$EP(\kappa) = \frac{1}{\sqrt{N} \times (1 - p_e)} \times \sqrt{(A + B - C)}, \quad (2.23)$$

em que:

$$A = \sum_{i=1}^C p_{ii} \times [1 - (p_{i+} + p_{+i})(1 - \kappa)]^2,$$

$$B = (1 - \kappa)^2 \times \sum_{i=1}^C \sum_{\substack{j=1 \\ j \neq i}}^C p_{ij} (p_{+i} + p_{j+})^2 \text{ e finalmente,}$$

$$C = [\kappa - p_e(1 - \kappa)]^2.$$

Devido à simplicidade da formulação de Cohen (1960), ela é frequentemente utilizada para fins de planejamento. Além disso, em NCSS (2020) é ressaltado ainda que a

aproximação de Cohen (1960) geralmente resulta em valores próximos aos resultados da formulação mais precisa de Fleiss, Cohen e Everitt (1969).

Dessa forma, o IC (Intervalo de Confiança) para Kappa é dado pela Equação 2.24.

$$IC = \kappa \pm z_{\frac{\alpha}{2}} \times EP(\kappa), \quad (2.24)$$

em que $z_{\frac{\alpha}{2}}$ é o percentil superior da distribuição normal padrão para $\frac{\alpha}{2}$ e α é o nível de significância,

O índice Kappa mais comum avalia a concordância entre dois avaliadores, porém existem algumas extensões (SILVA; PAES, 2012). Cohen (1968) propôs o Kappa ponderado, visando distinguir os níveis de concordâncias e discordâncias (entre leves, moderadas e graves, por exemplo), atribuindo pesos diferentes para cada tipo delas. O Kappa original trata as concordâncias e discordâncias como igualmente graves, enquanto o ponderado permite especificar diferentes níveis gravidade. Além disso, Fleiss (1971) propôs o Kappa generalizado, uma extensão para os casos onde existam mais de dois avaliadores.

3 MATERIAIS E MÉTODOS

O presente capítulo apresenta uma descrição dos materiais que foram necessários para a realização do trabalho, bem como a metodologia utilizada para seu desenvolvimento. Este capítulo é estruturado nos seguintes tópicos:

- Materiais, onde há uma descrição dos dados, *software* e *hardware* utilizados.
- Métodos, onde é descrita toda a metodologia utilizada para o processamento dos dados originais, estruturação e treinamento das RNAs e composição dos comitês de RNAs.

3.1 Materiais

Neste trabalho, os dados utilizados foram imagens radiografadas de sementes de girassol. Essas imagens foram submetidas a métodos de processamento de imagens e extração de características a fim de servir como dados de treinamento para RNAs.

As imagens radiografadas são de sementes de girassol (*Helianthus annuus* L.), cultivar Hélio-250, e foram obtidas pelo Laboratório de Análise de Sementes do Departamento de Agricultura da Universidade Federal de Lavras. As sementes passaram pelo processo de radiografia com intensidade de 22kV e exposição média de 11 segundos em um aparelho de raios-x Faxitron MX20, resultando em imagens digitalizadas que foram arquivadas em formato jpeg.

Após análise visual por parte de profissionais da área de análise de sementes, as mesmas foram classificadas como cheias, parcialmente cheias ou danificadas. Foram classificadas como sementes cheias aquelas com embrião opaco na radiografia; parcialmente cheias as com danos, porém com preservação do eixo embrionário; por fim, como deformadas as com danos que afetam o eixo embrionário. Ao término da classificação, o grupo de sementes cheias contou com 175 imagens, o de sementes parcialmente cheias com 130 imagens, e o de sementes deformadas com 140 imagens, totalizando 445 imagens.

Foi utilizado o *software* Matlab® versão 2014a em um computador com sistema operacional Windows® em sua versão 10, processador Intel® i5-8265U *quad-core* com *clock* de 1,6 GHz (com tecnologia Turbo *Boost* até 3,9 GHz) e 6 MB de memória *cache*, memória RAM de 16 GB DDR4 de 2400 MHz, além de placa gráfica NVIDIA® GeForce® MX130 com 2 GB de memória VRAM DDR5.

3.2 Métodos

A metodologia desenvolvida neste trabalho consiste em realizar uma série de processamentos nos dados originais (radiografias de sementes) a fim de se extrair características e criar diferentes conjuntos de dados. Por sua vez, os conjuntos de dados são utilizados para o treinamento, validação e teste de RNAs de diferentes configurações, compondo comitês de RNAs. Por fim, foram calculadas diversas métricas de desempenho para identificar a melhor metodologia de geração desses conjuntos de dados. As próximas seções tratam detalhadamente da metodologia, enquanto a Figura 3.1 apresenta um diagrama com o passo a passo partindo dos dados originais até a geração dos diversos conjuntos de dados para treinamento, validação e teste das RNAs.

3.2.1 Balanceamento de classes

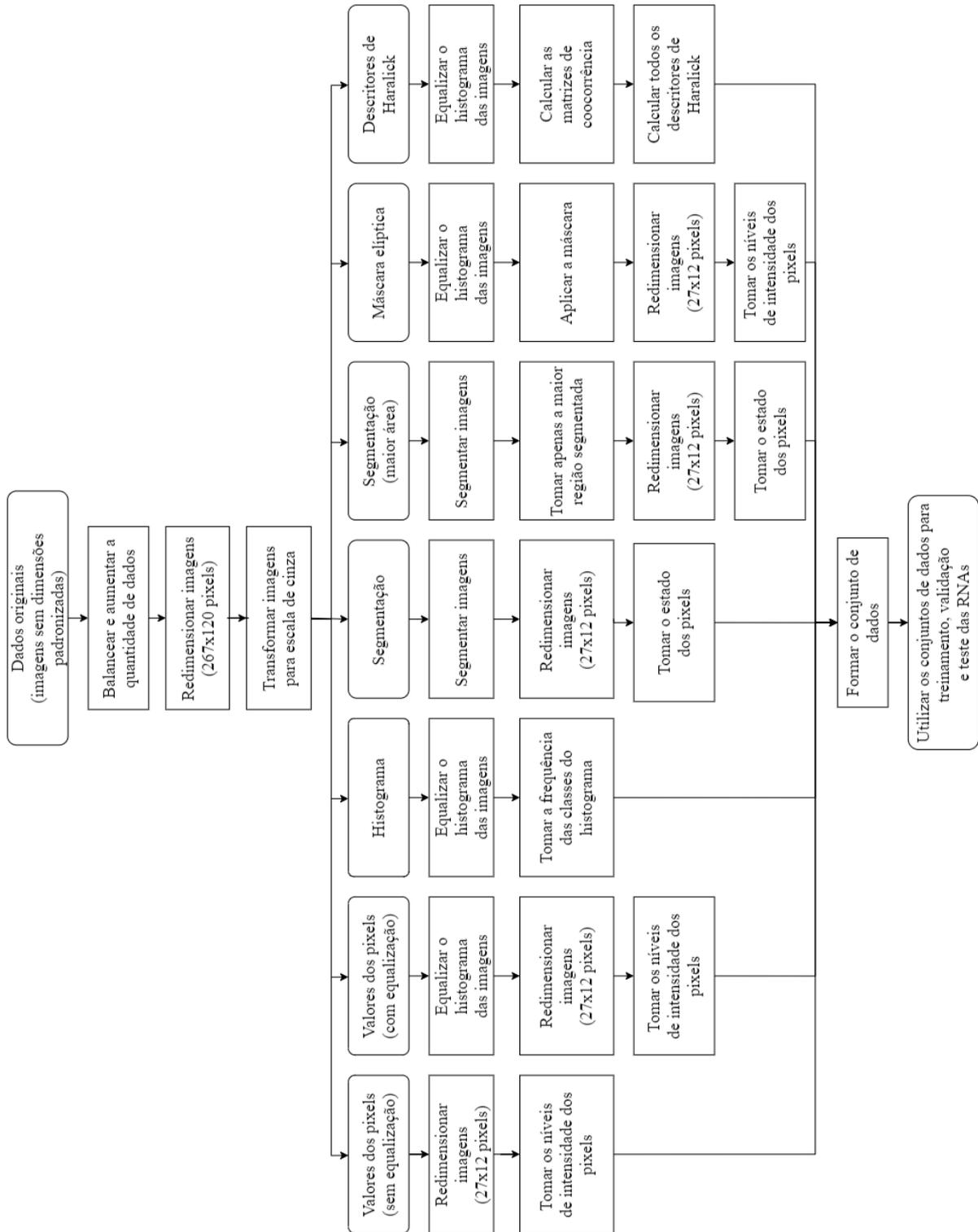
O treinamento com classes desbalanceadas pode resultar em RNAs que beneficiem as classes com maior número de elementos, diminuindo a acurácia. Geralmente, isso faz com que as RNAs tenham bom desempenho em relação às classes com maior quantidade de elementos, mas desempenho muito ruim em relação às classes com menor quantidade de elementos.

Existem diversos métodos para minimizar o problema do desbalanceamento. A reamostragem do tipo *oversampling* é uma das mais comuns e foi utilizada neste trabalho. Este tipo de re-amostragem consiste em balancear as classes por meio da geração de novos elementos para as classes com menor quantidade de amostras (HE; MA, 2013). Essa geração de novos elementos dá-se até que todas as classes possuam a mesma quantidade de elementos.

Neste trabalho, os dados utilizados possuíam quantidade distinta de elementos em cada classe. O grupo de sementes cheias continha 175 imagens, o de sementes parcialmente cheias continha 130 imagens e o de sementes deformadas continha 140 imagens.

Visando o balanceamento das classes, foram selecionadas, aleatoriamente, 45 imagens de sementes parcialmente cheias e 35 imagens de sementes deformadas. As imagens selecionadas foram então duplicadas e adicionadas às suas respectivas classes, resultando em 175 imagens em todas as três classes e, conseqüentemente, 525 imagens no total.

Figura 3.1 – Diagrama esquemático da geração dos diversos conjuntos de dados.



Fonte: do autor.

3.2.2 Aumento artificial dos dados

A Maldição da Dimensionalidade, termo cunhado por Richard Bellman, diz que a quantidade de amostras necessárias para se estimar uma função cresce exponencialmente em relação a quantidade de variáveis de entrada dessa função (CHEN, 2009). Isto indica a necessidade de grande quantidade de amostras quando se tem muitas variáveis de entrada, como ocorre neste trabalho.

A quantidade de amostras no conjunto de dados original, mesmo com o pequeno aumento ocasionado pelo balanceamento das classes, ainda pode ser considerada baixa. Isso é um fator complicador principalmente nos casos em que são usados os valores dos pixels como variáveis de entrada.

Tal problema geralmente resulta em *overfitting*¹ das RNAs, fazendo com que elas percam poder de generalização e apenas “memorizem” os dados de treinamento. Como consequência, apesar de terem bom desempenho nos dados de treinamento, o desempenho nos dados de teste costuma ser baixo.

Idealmente deve-se coletar mais dados até que se obtenha uma quantidade adequada de amostras, mas em alguns casos essa coleta não é possível por diversos fatores. A fim de contornar esse problema, o conjunto de dados original pode ser aumentado de forma artificial.

Neste trabalho, foram tomadas cópias das imagens do conjunto de dados original, espelhadas em relação aos eixos x e y . Desta forma, o conjunto de dados que antes contava com 175 amostras em cada classe passou a contar com 700 amostras. Esse aumento, apesar de artificial, confere maior diversidade ao conjunto de dados e pode atenuar um possível problema de *overfitting*.

3.2.3 Redimensionamento e transformação de cores

Originalmente, as imagens possuíam dimensões distintas entre si. Por essa razão, visando padronizar as dimensões, as imagens foram redimensionadas para 267×120 pixels (média das dimensões originalmente encontradas), utilizando a função *imresize* do Matlab®. Essa função implementa interpolação bicúbica, onde o valor do pixel de saída é uma média ponderada dos valores dos pixels na região 4×4 mais próxima.

¹ *Overfitting* ou sobreajuste é um termo usado para descrever um modelo que se ajusta bem aos dados já observados, mas tem um desempenho ruim para novos dados. Isso indica que o modelo decorou os dados, incluindo ruídos, e não aprendeu a generalizar para novos dados.

Além disso, originalmente as imagens foram salvas no formato de cores RGB, razão pela qual todas passaram também pela transformação RGB para escala de cinzas. Essa transformação consiste na eliminação das informações de matiz e saturação, mantendo a informação de luminosidade (TYAGI, 2018). Este processo ocorre com aplicação da Equação 3.1 e deu-se com a utilização da função *rgb2gray* do Matlab®.

$$Pixel = 0,2989 \times R + 0,5870 \times G + 0,1140 \times B, \quad (3.1)$$

onde R , G e B correspondem aos valores de intensidade dos canais vermelho, verde e azul, respectivamente.

3.2.4 Redução de dimensionalidade dos dados de entrada

A quantidade de variáveis de entrada impacta diretamente no treinamento e desempenho das RNAs. A maldição da dimensionalidade, termo cunhado por Bellman (1961), diz que a quantidade de atributos dos dados impacta de forma exponencial na quantidade de dados necessários para se obter o conhecimento desejado.

Dada uma quantidade fixa de elementos, a implicação prática da maldição da dimensionalidade é a de que existe um número máximo de atributos a partir do qual o desempenho do classificador irá sofrer uma piora ao invés de melhorar. Dessa forma, caso não seja possível aumentar a quantidade de elementos, deve-se então reduzir a dimensionalidade, diminuir a quantidade de atributos.

O conjunto de dados utilizado neste trabalho é relativamente pequeno, contando com apenas 700 amostras em cada classe, mesmo após o balanceamento das classes e aumento artificial dos dados. Além disso, após o redimensionamento inicial, as imagens possuíam 267x120 pixels, o que resultaria em vetores de entrada com 32040 valores. Somando-se estes fatos, o treinamento das RNAs seria dificultoso, demandando grande custo computacional para o processamento.

Neste trabalho não foram utilizados métodos de seleção de atributos. A solução encontrada, apesar da perda de informações inerente ao processo, foi realizar um novo redimensionamento das imagens nos casos em que os valores dos pixels foram utilizados diretamente como entrada para as RNAs. O novo redimensionamento foi de ordem dez, reduzindo a imagem de 267x120 pixels para 27x12 pixels e, conseqüentemente, reduzindo

os vetores de entrada de 32040 para 324 valores. Os casos em que esse novo redimensionamento foi necessário são informados nas seções seguintes.

3.2.5 Conjuntos de dados por valores dos pixels

Os conjuntos de dados formados por valores dos pixels foram formados após o emprego de técnicas de processamento de imagens aos dados originais. Cada conjunto de dados tem suas particularidades acerca de quais técnicas foram utilizadas e são descritos a seguir, mostrando também a ordem das técnicas aplicadas.

3.2.5.1 Valores dos pixels

A composição desse conjunto de dados foi realizada tomando diretamente os valores dos pixels das imagens. Este é o conjunto mais simples, tendo sido formado apenas pelo emprego das etapas utilizadas de forma comum em todos os conjuntos de dados formados. As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;
5. Redução de dimensionalidade (redimensionamento para 27x12 pixels).

3.2.5.2 Valores dos pixels com equalização de histograma

Tal como no conjunto de valores dos pixels, este conjunto foi composto pelos valores dos pixels das imagens e passou pelas mesmas etapas. Entretanto, os valores dos pixels foram tomados após as imagens terem sido submetidas ao processo de equalização de seus histogramas. As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);

4. Transformação para escala de cinzas;
5. Equalização de histograma;
6. Redução de dimensionalidade (redimensionamento para 27x12 pixels).

3.2.5.3 Segmentação

Este conjunto de dados foi composto também pelos valores dos pixels, porém neste caso as imagens passaram pelo processo de limiarização binária. Desta forma, cada pixel assumiu valor 0 ou 1 (preto ou branco). As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;
5. Limiarização binária;
6. Redução de dimensionalidade (redimensionamento para 27x12 pixels).

3.2.5.4 Segmentação (maior área)

O processo de segmentação original resultou em pequenos fragmentos formados por pixels brancos fora da região de interesse (parte central das sementes). Visando eliminar esse problema, foi utilizada apenas a maior região segmentada da imagem, que também é a região central. Este conjunto foi, portanto, formado pelos valores dos pixels após a etapa de limiarização binária e redimensionamento da região de interesse. As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;

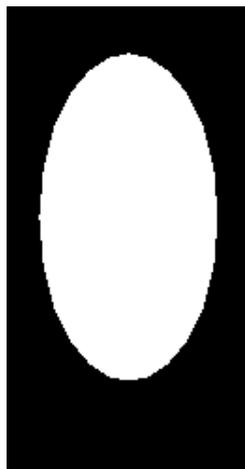
5. Limiarização binária;
6. Seleção da maior área segmentada;
7. Redução de dimensionalidade (redimensionamento para 27x12 pixels).

3.2.5.5 Máscara elíptica

As sementes possuem formato elipsoidal, portanto, considerando como área de interesse uma região elíptica a partir do centro das imagens, foi criada uma máscara lógica (pixels 0 e 1) com as mesmas dimensões das imagens padronizadas (267x120 pixels).

Em um processo denominado aritmética pontual, a máscara foi aplicada. Este processo nada mais é do que multiplicação, elemento por elemento, das matrizes correspondentes à imagem e à máscara. A máscara criada é mostrada na Figura 3.2.

Figura 3.2 – Máscara elíptica.



Fonte: do autor.

Este conjunto de dados foi formado pelos valores dos pixels da região elíptica (parte branca da máscara). As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;
5. Equalização de histogramas;

6. Redução de dimensionalidade (redimensionamento para 27x12 pixels);
7. Aplicação da máscara.

3.2.6 Conjuntos de dados por extração de características

Os conjuntos de dados formados por extração de características das imagens foram formados por meio do emprego de técnicas de extração de características. Os conjuntos formados são descritos a seguir.

3.2.6.1 Histograma

A geração deste conjunto de dados se deu por meio da construção de histogramas das imagens, já com dimensões padronizadas. Os histogramas foram construídos utilizando 64 classes (níveis de cinza) e o conjunto de dados foi formado pelas frequências, ou quantidade de pixels, presentes em cada uma das classes. As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;
5. Construção e equalização de histogramas (64 classes).

3.2.6.2 Descritores de textura de Haralick (14 descritores)

Este conjunto de dados foi gerado por meio do cálculo dos descritores de textura propostos por Haralick. As matrizes de coocorrência das imagens foram calculadas para as direções 0° , 45° , 90° e 135° , considerando a distância de 20 pixels e 64 níveis de cinza.

A partir das matrizes de coocorrência foram calculados todos os 14 descritores de textura propostos por Haralick para cada uma das direções. Para o cálculo dos descritores, foi utilizada a função *haralickTextureFeatures*, desenvolvida por Monzel (2018) para utilização no Matlab[®]. As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;
5. Equalização dos histogramas;
6. Cálculo das matrizes de coocorrência para as quatro direções;
7. Cálculo dos quatorze descritores em cada uma das direções.

3.2.6.3 Descritores de textura de Haralick (4 descritores)

Este conjunto de dados também foi composto por descritores de Haralick, entretanto foram considerados apenas quatro dos principais: contraste, correlação, entropia e variância. As matrizes de coocorrência das imagens também foram calculadas para as direções 0° , 45° , 90° e 135° , com distância de 20 pixels e considerando 64 níveis de cinza.

A partir das matrizes de coocorrência foram calculados quatro dos principais descritores de textura para cada uma das direções. As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;
5. Equalização dos histogramas;
6. Cálculo das matrizes de coocorrência para as quatro direções;
7. Cálculo de quatro dos principais descritores em cada uma das direções.

3.2.7 Conjuntos de dados por combinação de outros conjuntos

Os conjuntos de dados desta seção foram formados por meio da combinação de conjuntos de dados das seções anteriores. A formação desses conjuntos é descrita a seguir.

3.2.7.1 Descritores de textura de Haralick (duas distâncias)

Este conjunto de dados foi gerado por meio do cálculo dos descritores de textura propostos por Haralick considerando duas diferentes distâncias entre os pixels. As matrizes de coocorrência das imagens foram calculadas para as direções 0° , 45° , 90° e 135° , com distâncias de 20 e 25 pixels e considerando 64 níveis de cinza.

Os 14 descritores de textura propostos por Haralick foram calculados para cada uma das quatro direções e cada uma das duas distâncias escolhidas. Para o cálculo dos descritores, foi utilizada a função *haralickTextureFeatures*, desenvolvida por Monzel (2018).

As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;
5. Equalização dos histogramas;
6. Cálculo das matrizes de coocorrência para as quatro direções considerando distância de 20 pixels;
7. Cálculo das matrizes de coocorrência para as quatro direções considerando distância de 25 pixels;
8. Cálculo dos quatorze descritores em cada uma das direções e cada uma das distâncias.

3.2.7.2 Descritores de textura de Haralick e valores dos pixels com equalização

Este conjunto de dados foi composto pela junção dos conjuntos dos descritores de Haralick e dos valores do pixel com equalização de histograma.

As matrizes de coocorrência das imagens foram calculadas para as direções 0° , 45° , 90° e 135° , com distâncias de 20 pixels e considerando 64 níveis de cinza. A partir

dessas matrizes, foram calculados os 14 descritores de textura de Haralick, que compõe a primeira parte deste conjunto.

A segunda parte desse conjunto foi composta pelos valores dos pixels das imagens tomados após equalização de histograma das imagens. As etapas percorridas para formação deste conjunto de dados foram:

1. Balanceamento de classes;
2. Aumento artificial dos dados;
3. Redimensionamento de imagens (padronização, 267x120 pixels);
4. Transformação para escala de cinzas;
5. Equalização de histogramas;
6. Cálculo das matrizes de coocorrência para as quatro direções considerando distância de 20 pixels;
7. Cálculo dos quatorze descritores em cada uma das direções;
8. Redução de dimensionalidade (redimensionamento para 27x12 pixels).

3.2.8 Normalização dos dados

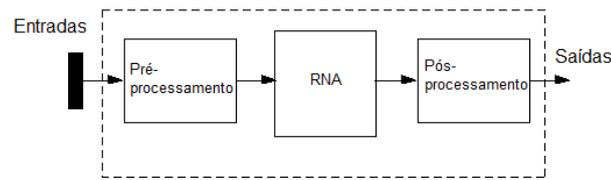
O treinamento das RNAs geralmente é mais eficiente quando se normaliza os dados, pois a função de custo resultante de dados normalizados tende a ter maior simetria, reduzindo o tempo gasto para o treinamento.

A normalização é essencial quando há variáveis com ordens de grandeza muito distintas, mas por se tratar de um processo pouco custoso em termos computacionais, a normalização dos dados é considerada uma boa prática em todas as situações.

Neste trabalho, todos os conjuntos de dados formados descritos nas seções anteriores passaram pelo processo de normalização. A Figura 3.3 representa o processo de normalização dos dados.

A função de normalização escolhida foi a *mapminmax*, que transforma os dados a fim de deixá-los com valores mínimo e máximo de -1 e 1, respectivamente. A Equação 3.2 descreve essa normalização.

Figura 3.3 – Esquema de normalização dos dados.



Fonte: adaptado de mathworks.com/help/deeplearning/ug/choose-neural-network-input-output-processing-functions.html.

$$y = (y_{max} - y_{min}) \times \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right) + y_{min}, \quad (3.2)$$

onde y é o valor normalizado; x_{min} e x_{max} são o menor e o maior valor encontrados nos dados, respectivamente; y_{min} e y_{max} são, respectivamente, o menor e maior valores desejados após a normalização, neste caso -1 e 1.

A normalização dos dados é aplicada para cada variável separadamente. Dessa forma, cada variável tem seus valores normalizados considerando m elementos que compõem o conjunto de dados.

Em um conjunto de dados representado por uma matriz $n \times m$, onde n é a quantidade de variáveis e m a quantidade de elementos, a normalização age em cada uma das linhas individualmente, normalizando os valores de cada variável para todas as amostras.

3.2.9 Arquitetura e parâmetros de treinamento das RNAs

As RNAs utilizadas foram de topologia *feed-forward*, com uma camada escondida e função de ativação do tipo tangente hiperbólica e *softmax* nas camadas escondida e de saída, respectivamente. Tal escolha se deu por parcimônia, uma vez que testes com maior quantidade de camadas escondidas não resultou diretamente em melhor desempenho.

Visando diversidade na composição do comitê de RNAs, a quantidade de neurônios na camada escondida de cada rede foi definido aleatoriamente, em intervalo compreendido entre 1/4 e 1/2 do número de variáveis de entrada, que varia de acordo com o conjunto de dados utilizado no treinamento. A camada de saída das RNAs contou com dois ou três neurônios, a depender de quantas classes foram testadas, uma vez que a classificação foi realizada utilizando as três classes de sementes em conjunto e também comparando de duas em duas classes.

O treinamento das RNAs foi do tipo supervisionado, tendo sido fornecida a classificação esperada de parte dos conjunto de dados. Foram compostos 10 comitês, com 10 RNAs em cada um deles, onde se visou a obtenção da média e intervalo de confiança de cada uma das métricas de desempenho desses comitês: acurácia, índice Kappa e AUC.

Os conjuntos de dados foram divididos, de maneira estratificada, em 10 subconjuntos. A cada composição de comitê, um desses conjuntos foi utilizado para teste, enquanto os nove restantes foram utilizados para treinamento e validação da RNAs que compuseram o referido comitê.

Buscando maior diversidade entre os componentes dos comitês, o conjunto de treinamento de cada rede variou conforme a técnica de reamostragem *bagging*, descrita na Seção 2.5.1. O algoritmo Rprop (*Resilient Backpropagation*), proposto em Riedmiller e Braun (1993), foi utilizado para a atualização dos pesos nas RNAs.

O algoritmo Rprop tem como objetivo eliminar o efeito de gradiente de fuga, presente quando se utiliza funções de ativação sigmoidais. Neste algoritmo apenas o sinal da derivada é utilizado para determinar a direção da atualização do peso, ou seja, a magnitude da derivada não tem efeito sobre a atualização do peso. Dessa forma, o tamanho da mudança no peso é dado por um valor de atualização separado.

A atualização do valor de cada peso é dada por fatores de incremento e decremento, δ_i e δ_d , respectivamente. Sempre que a derivada da função de desempenho em relação a determinado peso possui mesmo sinal em duas iterações sucessivas, o peso é incrementado pelo fator δ_i . Se o sinal da derivada for diferente, o peso sofre decremento com fator δ_d . Por fim, se a derivada for zero, o valor do peso permanece.

Neste trabalho foi utilizado δ_i de 1,2 e δ_d de 0,5. Além disso, foi utilizada taxa de aprendizagem de 0,01, máximo de épocas de 200, função de desempenho entropia cruzada e parada antecipada considerando 10 épocas subsequentes. A escolha desses valores se deu de forma empírica.

A parada antecipada tem por objetivo prevenir o *overfitting*, interrompendo o treinamento precocemente quando se detecta perda do poder de generalização da rede por determinado número de épocas subsequentes. Essa detecção se dá por meio do teste de desempenho da rede sobre os dados de validação.

Desta forma, cada RNA nos comitês pôde diferir das outras em relação aos valores dos pesos iniciais, quantidade de neurônios na camada escondida e elementos do con-

junto de treinamento efetivamente utilizados para o treinamento (técnica *bagging*). Essa estratégia confere diversidade ao comitê.

3.2.10 Treinamento das RNAs e formação dos comitês

Os conjuntos de dados gerados foram subdivididos, de maneira estratificada, em 10 subconjuntos. Na formação de cada um dos 10 comitês, um dos subconjuntos foi utilizado como teste, enquanto os demais como dados de treinamento e validação.

Essa abordagem é denominada validação cruzada *k-fold*, em que k é a quantidade de subdivisões realizadas. A validação *k-fold* visa garantir que cada um desses conjuntos seja utilizado como teste em uma iteração. A cada iteração, $k-1$ conjuntos são utilizados para treinamento e um conjunto para teste.

A partir dessa abordagem, foi possível obter a média e o intervalo de confiança para as métricas de avaliação. Isso faz com que o resultado obtido seja mais robusto, diminuindo a chance de o desempenho obtido ter sido fruto do acaso ao ter se escolhido elementos mais fáceis ou difíceis de se classificar para compor o conjunto de testes.

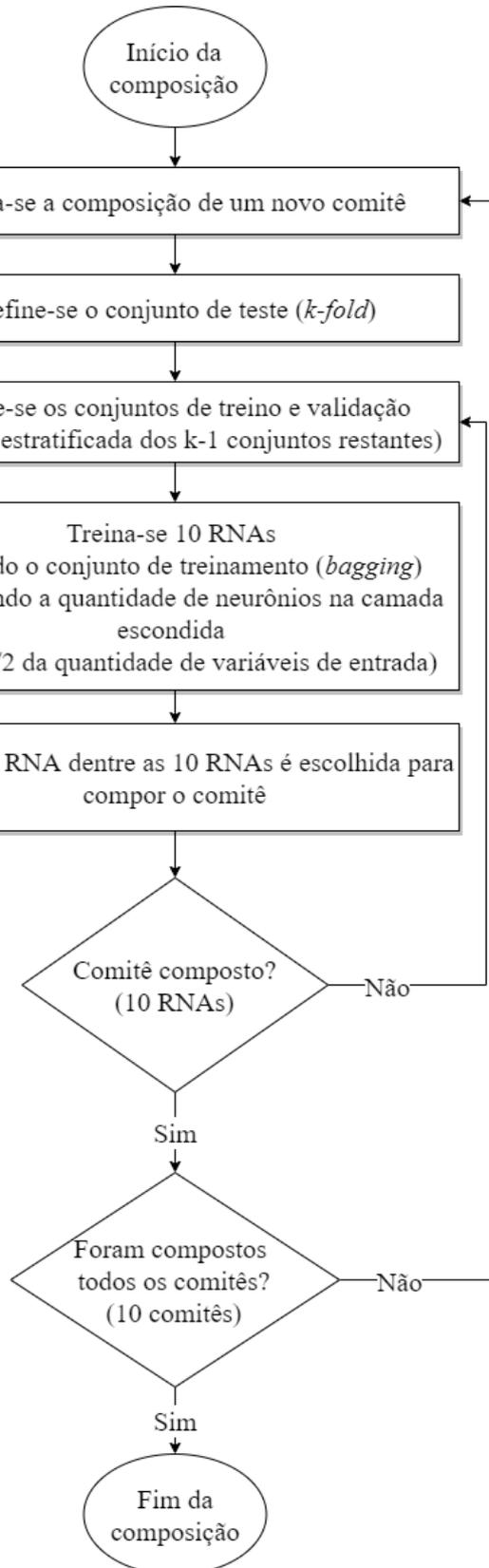
Foram compostos 10 comitês com 10 RNAs em cada um deles. Os comitês foram testados com os dados de teste, onde a cada comitê formado os dados de teste foram alternados (*k-fold*). Ao final deste processo, foram calculadas as médias das métricas de avaliação e seus respectivos intervalos de confiança.

A cada seleção de um novo membro para o comitê houve alterações nos conjuntos de treinamento e validação, permanecendo apenas o conjunto de testes como fixo. O conjunto de treinamento variou conforme técnica de amostragem *bagging* a cada RNA treinada. A quantidade de neurônios na camada escondida também foi variada a cada rede treinada, sendo escolhida aleatoriamente entre $1/4$ e $1/2$ da quantidade de variáveis de entrada. Por meio dessas variações, buscou-se conferir diversidade ao comitê de RNAs. O fluxograma da Figura 3.4 mostra o processo de composição dos comitês.

A resposta de cada comitê foi dada pela combinação das respostas das RNAs que os integram. O método de combinação escolhido foi o de média ponderada, dando maior peso para as RNAs com maior acurácia em relação aos dados de validação.

Por fim, as respostas de cada um dos 10 comitês foram utilizadas para calcular-se a média e o intervalo de confiança de cada uma das métricas de avaliação.

Figura 3.4 – Diagrama esquemático da composição dos comitês.



Fonte: do autor.

3.2.11 Obtenção das métricas de avaliação

As matrizes de confusão para as respostas de cada comitê foram geradas a partir de suas respostas. Por meio das matrizes de confusão calculou-se o índice Kappa e a acurácia de cada comitê. Dessa forma, obteve-se a média dessas métricas e os respectivos intervalos de confiança.

As curvas ROC referentes a cada resposta foram combinadas fixando os valores de seus limiares durante suas construções, garantindo que fossem tomados sempre os mesmos limiares. Foram então calculadas as médias dos valores em seus eixos (x e y).

Os casos em que são consideradas apenas duas classes de sementes, por se tratar de uma classificação binária, resultam em apenas uma curva ROC por comitê. Nestes casos, a combinação das curvas ROC resultou em apenas uma curva, a partir da qual calculou-se a métrica AUC.

Por fim, os casos em que são consideradas as três classes de sementes resultam em três curvas ROC, conforme abordagem *OvA*, considerando sempre uma classe contra as outras duas. A combinação das curvas ROC resulta em três curvas ROC. A métrica AUC nestes casos é a média das AUCs das três curvas ROC resultantes.

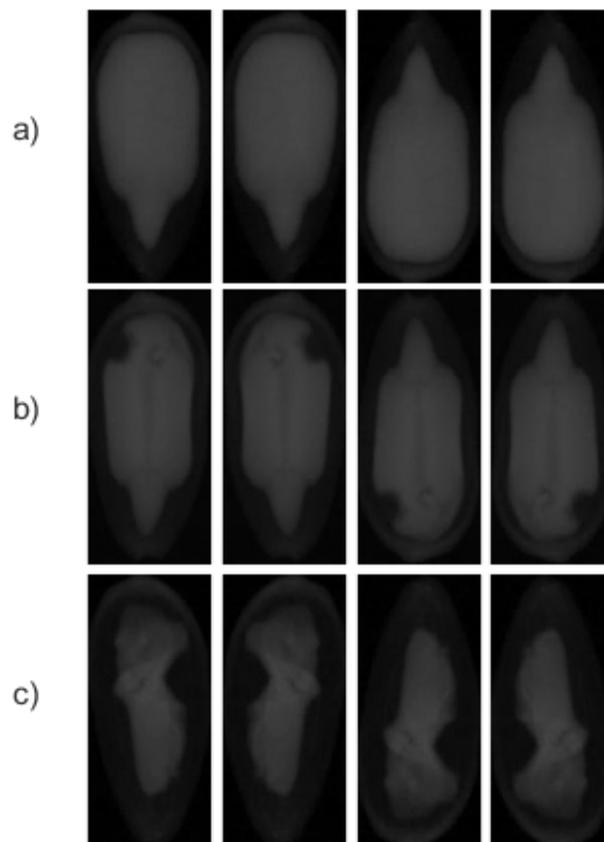
4 RESULTADOS E DISCUSSÃO

O presente capítulo apresenta os resultados obtidos por meio do processamento do conjunto de dados original para a formação dos diversos conjuntos utilizados. São apresentados também os resultados dos treinamentos dos comitês de RNAs para cada conjunto de dados formado, bem como suas métricas de avaliação e discussão acerca dos resultados encontrados.

4.1 Aumento artificial dos dados

Por meio da aplicação das técnicas de aumento de dados foi possível quintuplicar o número de amostras do conjunto de dados original. A Figura 4.1 exemplifica esse processo em imagens de cada uma das classes de sementes em estudo, em que a primeira imagem de cada letra é a original e as subsequentes são as imagens espelhadas de forma vertical e horizontal.

Figura 4.1 – Exemplo de imagens de cada classe de sementes e suas variações para o aumento artificial do conjunto de dados: a) cheias b) parcialmente cheias c) deformadas.



Fonte: do autor.

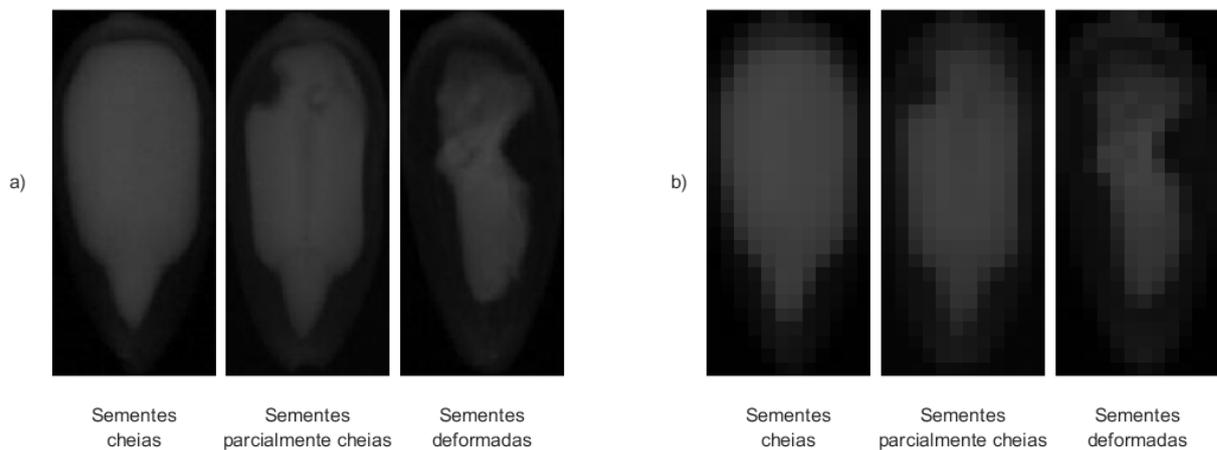
4.2 Conjuntos de dados por valores dos pixels

Os conjuntos de dados formados por valores de pixels têm suas características descritas a seguir. Tais conjuntos foram formados por meio da aplicação de técnicas de processamento de imagens.

4.2.1 Valores dos pixels

O conjunto de dados composto pelos valores dos pixels contou com 324 variáveis de entrada (quantidade de pixels - 27x12) e 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente. A Figura 4.2 mostra exemplos de imagens de cada classe antes e após o redimensionamento das imagens.

Figura 4.2 – Exemplo de imagens de cada classe: a) antes do redimensionamento b) após o redimensionamento.

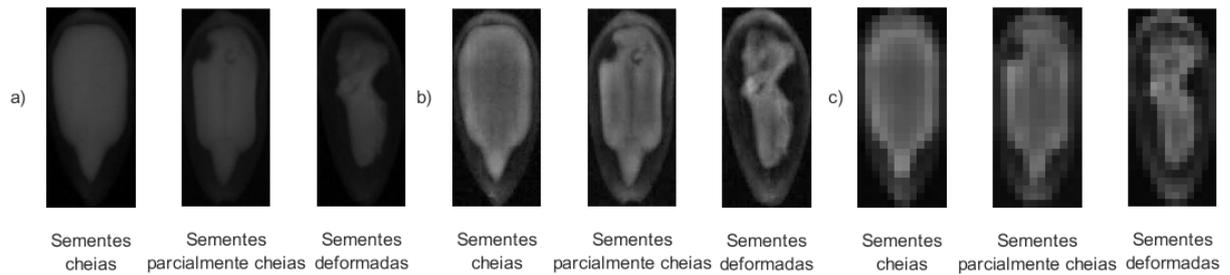


Fonte: do autor.

4.2.2 Valores dos pixels com equalização de histograma

O conjunto de dados composto pelos valores dos pixels com equalização de histograma contou com 324 variáveis de entrada (quantidade de pixels - 27x12) e 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente. A Figura 4.3 mostra exemplos de imagens de cada classe antes e após a equalização de histograma e redimensionamento das imagens.

Figura 4.3 – Exemplo de imagens de cada classe: a) antes da equalização de histograma e redimensionamento b) após equalização de histograma c) após redimensionamento.

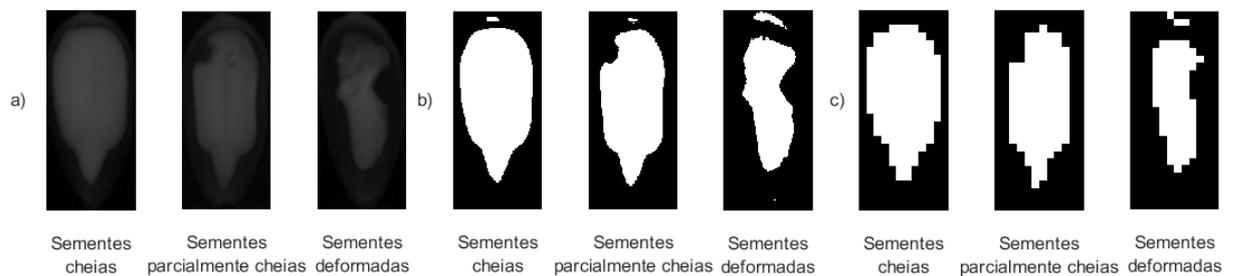


Fonte: do autor.

4.2.3 Segmentação

A segmentação das imagens resultou em conjunto de dados de 324 variáveis de entrada (quantidade de pixels - 27x12) e 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente. A Figura 4.4 exemplifica imagens de cada uma das classes com a aplicação da segmentação e redimensionamento.

Figura 4.4 – Exemplo de imagens de cada classe: a) antes da segmentação e redimensionamento b) após segmentação c) após redimensionamento.

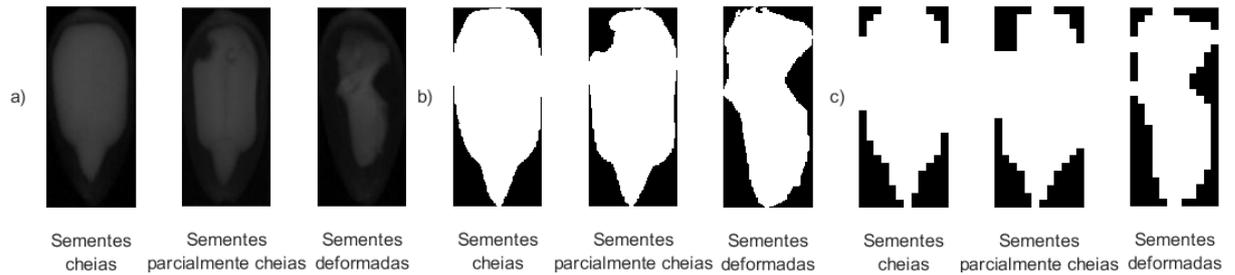


Fonte: do autor.

4.2.4 Segmentação (maior área)

O conjunto de dados gerado por meio da segmentação utilizando a maior área segmentada das imagens contou com 324 variáveis de entrada (quantidade de pixels - 27x12) e 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente. Exemplos de imagens de cada classe com a aplicação da segmentação de maior área e redimensionamento estão presentes na Figura 4.5.

Figura 4.5 – Exemplo de imagens de cada classe: a) antes da segmentação (maior área) e redimensionamento b) após segmentação (maior área) c) após redimensionamento.

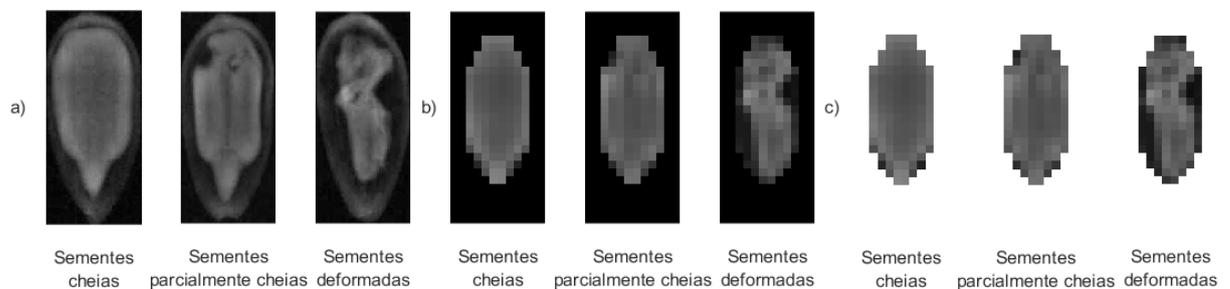


Fonte: do autor.

4.2.5 Máscara elíptica

A aplicação da máscara resultou em conjunto de dados formado por 126 variáveis de entrada (quantidade de pixels dentro da região elíptica) e 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente. A Figura 4.6 exemplifica a aplicação da máscara em imagens de cada uma das classes.

Figura 4.6 – Exemplo de imagens de cada classe: a) antes da aplicação da máscara b) durante a aplicação da máscara (imagens já redimensionadas) c) pixels válidos de cada imagem após a aplicação da máscara.



Fonte: do autor.

4.3 Conjuntos de dados por extração de características

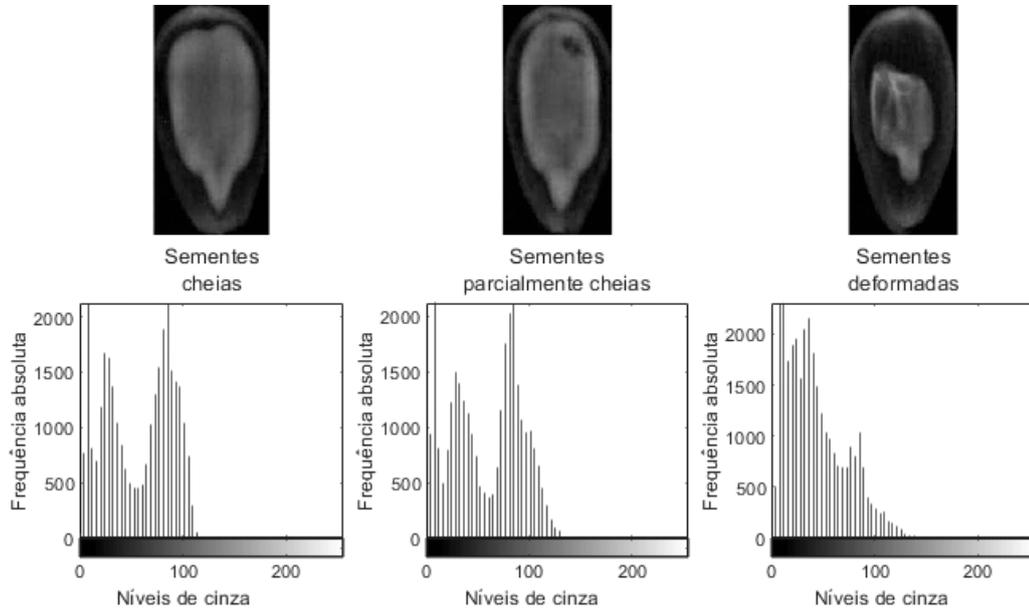
Nesta seção, são apresentados os conjuntos de dados formados por meio do emprego de técnicas de extração de características de imagens. Os conjuntos são descritos a seguir.

4.3.1 Histograma

O conjunto de dados formado pelos histogramas das imagens é composto por 64 variáveis de entrada (níveis de cinza ou classes do histograma) e 1400 ou 2100 amostras,

quando consideradas duas ou três classes, respectivamente. A Figura 4.7 exemplifica os resultados obtidos com este método para imagens de cada uma das classes.

Figura 4.7 – Exemplos de histogramas de imagens de cada uma das classes.



Fonte: do autor.

4.3.2 Descritores de textura de Haralick (14 descritores)

O conjunto de dados formado por meio dos 14 descritores de Haralick calculados nas quatro direções contou com 56 variáveis de entrada (14 descritores calculados em cada uma das 4 direções) e 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente.

4.3.3 Descritores de textura de Haralick (4 descritores)

O conjunto de dados formado por quatro dos principais descritores de Haralick contou com 16 variáveis de entrada (4 descritores calculados em cada uma das 4 direções) e 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente.

4.4 Conjuntos de dados por combinação de outros conjuntos

Os conjuntos de dados apresentados nessa seção foram formados por meio da combinação de outros conjuntos de dados que resultaram em melhor performance das RNAs. Esses conjuntos são descritos a seguir.

4.4.1 Descritores de textura de Haralick (duas distâncias)

O conjunto de dados formado por meio dos 14 descritores de Haralick calculados nas quatro direções e considerando duas distâncias distintas contou com 112 variáveis de entrada, sendo elas 14 descritores em cada uma das 4 direções e em cada uma das duas distâncias. Este conjunto contou com 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente.

4.4.2 Descritores de textura de Haralick e valores dos pixels com equalização

Por fim, o conjunto de dados formado pelos descritores de textura de Haralick e valores dos pixels com equalização foi composto pelos 14 descritores de Haralick calculados em cada uma das quatro direções e pelo nível de intensidade dos pixels com equalização de histograma. Esse conjunto contou com 380 variáveis de entrada, sendo 56 resultantes dos descritores de Haralick e 324 dos valores dos pixels. Além disso, contou com 1400 ou 2100 amostras, quando consideradas duas ou três classes, respectivamente.

4.5 Resumo dos conjuntos de dados

A fim de sintetizar as informações sobre os conjuntos de dados formados em cada etapa e visando facilitar a leitura e o entendimento do presente trabalho, foram resumidas na Tabela 4.1 as características de cada um dos conjuntos de dados formados.

Tabela 4.1 – Características dos conjuntos de dados formados.

Método	N. de variáveis	N. de amostras	Descrição
Valores dos pixels s/ equalização	324	1400 ou 2100	Nível de intensidade de cada pixel
Valores dos pixels c/ equalização	324	1400 ou 2100	Nível de intensidade de cada pixel
Segmentação	324	1400 ou 2100	Estado do pixel (0 ou 1)
Segmentação (maior região)	324	1400 ou 2100	Estado do pixel (0 ou 1)
Máscara elíptica	126	1400 ou 2100	Nível de intensidade de cada pixel
Histograma	64	1400 ou 2100	Quant. de pixels em cada nível de cinza
Haralick	56	1400 ou 2100	Valores dos descritores de textura
Haralick reduzido	14	1400 ou 2100	Valores dos descritores de textura
Haralick duplo	112	1400 ou 2100	Valores dos descritores de textura
Haralick e valores dos pixels com equalização	380	1400 ou 2100	Valores dos descritores de textura e nível de intensidade de cada pixel

Fonte: do autor.

4.6 Métricas de avaliação

Os comitês de RNAs foram utilizados para classificar as sementes de girassol quanto aos seus níveis de danos, utilizando cada um dos conjuntos de dados.

Os resultados obtidos dizem respeito às médias de acurácia, AUC e índice Kappa. Foram consideradas as classificações com as três classes ao mesmo tempo, e também de duas em duas classes.

Os resultados dispostos na Tabela 4.2 dizem respeito à média de acurácia para cada um dos conjuntos gerados e combinações de classes de sementes.

Tabela 4.2 – Acurácia média dos comitês de RNAs para cada conjunto de dados.

Método	Nível de acerto C x PC x D	Nível de acerto C x PC	Nível de acerto C x D	Nível de acerto PC x D
Pixels s/ equalização	75,7143% ± 1,1217%	75,4286% ± 2,3905%	93,8571% ± 1,0682%	88,5714% ± 2,6399%
Pixels c/ equalização	77,5714% ± 2,1213%	80,2857% ± 1,6460%	94,8571% ± 1,1952%	89,2857% ± 1,9909%
Histograma	90,3333% ± 1,2756%	93,6429% ± 1,2600%	100,0000% ± 0,0000%	98,8571% ± 0,7290%
Segmentação	70,8095% ± 1,7436%	66,5000% ± 2,6104%	93,3571% ± 0,9800%	86,5000% ± 1,9516%
Segmentação (maior região)	74,3333% ± 1,8690%	73,7143% ± 1,9105%	94,5714% ± 1,4489%	88,5714% ± 1,3199%
Máscara elíptica	75,6667% ± 1,6133%	75,5714% ± 1,7311%	94,3571% ± 1,0320%	89,1429% ± 1,3812%
Haralick	83,7143% ± 1,2351%	86,4286% ± 2,3146%	99,1429% ± 0,7467%	96,8571% ± 1,1657%
Haralick reduzido	72,9048% ± 1,7896%	73,6429% ± 3,4950%	95,6429% ± 0,7364%	88,6429% ± 2,0175%
Haralick duplo	89,5714% ± 1,0172%	92,0714% ± 1,4970%	99,2857% ± 0,5112%	98,4286% ± 0,9287%
Haralick e pixels com equalização	82,7619% ± 0,8208%	81,8571% ± 2,6000%	96,6429% ± 1,0020%	91,8571% ± 1,7361%

C - sementes cheias, PC - sementes parcialmente cheias e D - sementes deformadas.

Fonte: do autor.

Os resultados referentes às médias da métrica AUC para cada um dos conjuntos de dados utilizados no trabalho e para cada combinação de classes de sementes encontram-se na Tabela 4.3.

Tabela 4.3 – AUC média dos comitês de RNAs para cada conjunto de dados.

Método	Nível de acerto C x PC x D	Nível de acerto C x PC	Nível de acerto C x D	Nível de acerto PC x D
Pixels s/ equalização	0,9034 ± 0,0078	0,8241 ± 0,0170	0,9789 ± 0,0054	0,9425 ± 0,0177
Pixels c/ equalização	0,9130 ± 0,0108	0,8627 ± 0,0162	0,9875 ± 0,0032	0,9552 ± 0,0116
Histograma	0,9810 ± 0,0045	0,9843 ± 0,0072	1,0000 ± 0,0000	0,9990 ± 0,0014
Segmentação	0,8753 ± 0,0116	0,7340 ± 0,0201	0,9791 ± 0,0056	0,9387 ± 0,0116
Segmentação (maior região)	0,8942 ± 0,0098	0,8119 ± 0,0209	0,9798 ± 0,0070	0,9454 ± 0,0085
Máscara elíptica	0,9028 ± 0,0106	0,8352 ± 0,0159	0,9826 ± 0,0042	0,9555 ± 0,0090
Haralick	0,9495 ± 0,0070	0,9400 ± 0,0153	0,9996 ± 0,0005	0,9949 ± 0,0030
Haralick reduzido	0,8839 ± 0,0152	0,8150 ± 0,0301	0,9935 ± 0,0019	0,9530 ± 0,0110
Haralick duplo	0,9754 ± 0,0041	0,9757 ± 0,0085	0,9998 ± 0,0002	0,9990 ± 0,0013
Haralick e pixels com equalização	0,9364 ± 0,0098	0,8906 ± 0,0163	0,9923 ± 0,0069	0,9706 ± 0,0091

C - sementes cheias, PC - sementes parcialmente cheias e D - sementes deformadas.

Fonte: do autor.

Na Tabela 4.4 são apresentados os resultados referentes à média do índice Kappa para cada um dos conjuntos de dados utilizados e considerando cada uma das combinações de classes de sementes.

Tabela 4.4 – Índice Kappa médio dos comitês de RNAs para cada conjunto de dados.

Método	Nível de acerto C x PC x D	Nível de acerto C x PC	Nível de acerto C x D	Nível de acerto PC x D
Pixels s/ equalização	0,6357 ± 0,0168	0,5086 ± 0,0478	0,8771 ± 0,0214	0,7714 ± 0,0528
Pixels c/ equalização	0,6636 ± 0,0318	0,6057 ± 0,0329	0,8971 ± 0,0239	0,7857 ± 0,0398
Histograma	0,8550 ± 0,0191	0,8729 ± 0,0252	1,0000 ± 0,0000	0,9771 ± 0,0146
Segmentação	0,5621 ± 0,0262	0,3300 ± 0,0522	0,8671 ± 0,0196	0,7300 ± 0,0390
Segmentação (maior região)	0,6150 ± 0,0280	0,4743 ± 0,0382	0,8914 ± 0,0290	0,7714 ± 0,0264
Máscara elíptica	0,6350 ± 0,0242	0,5114 ± 0,0346	0,8871 ± 0,0206	0,7829 ± 0,0276
Haralick	0,7557 ± 0,0185	0,7286 ± 0,0463	0,9829 ± 0,0149	0,9371 ± 0,0233
Haralick reduzido	0,5936 ± 0,0268	0,4729 ± 0,0699	0,9129 ± 0,0147	0,7729 ± 0,0403
Haralick duplo	0,8436 ± 0,0153	0,8414 ± 0,0299	0,9857 ± 0,0102	0,9686 ± 0,0186
Haralick e pixels com equalização	0,7414 ± 0,0123	0,6371 ± 0,0520	0,9329 ± 0,0200	0,8371 ± 0,0347

C - sementes cheias, PC - sementes parcialmente cheias e D - sementes deformadas.

Fonte: do autor.

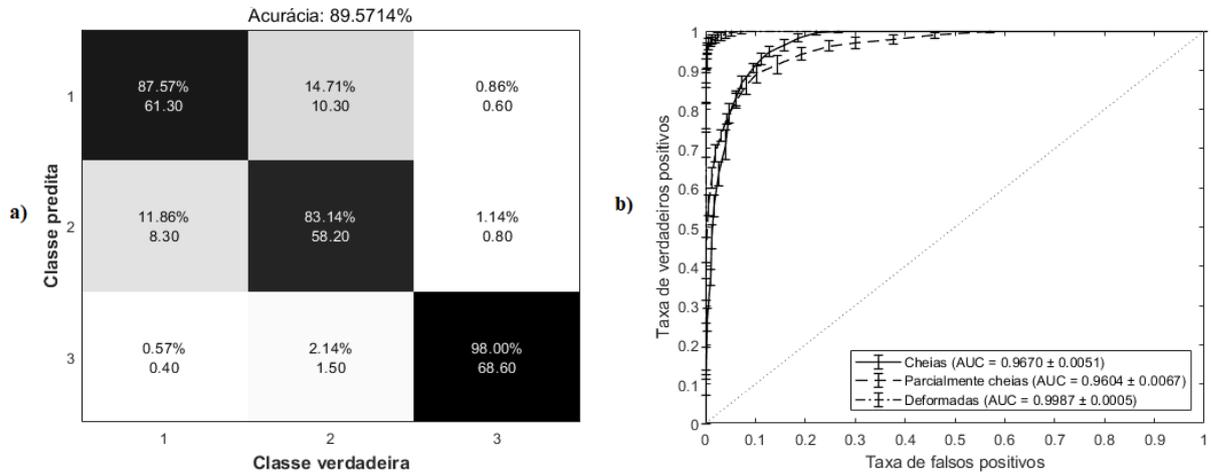
Apesar de o conjunto de dados formado pelo histograma das imagens ter resultado em melhores níveis de acurácia, índice Kappa e AUC, é importante ressaltar que os histogramas são gerados apenas pela contagem de pixels em cada uma das classes do histograma. Não há, portanto, diferença prática quando a imagem está na posição original, rotacionada ou espelhada, pois seus histogramas serão iguais.

Considerando essa questão sobre os histogramas, não haveria qualquer diferença para a geração deste conjunto o aumento artificial de dados ter sido realizado por meio do espelhamento, rotação ou simples cópias idênticas das imagens. Este fato possivelmente resultou em benefícios para este conjunto de dados.

Dessa forma, conclui-se que o melhor conjunto de dados é o formado pelos descritores de Haralick calculados para duas distâncias distintas, “Haralick duplo”. Desconsiderando o conjunto formado pelos histogramas, este foi o com melhor performance considerando todas as métricas de avaliação utilizadas e, portanto, seus resultados foram destacados nas tabelas.

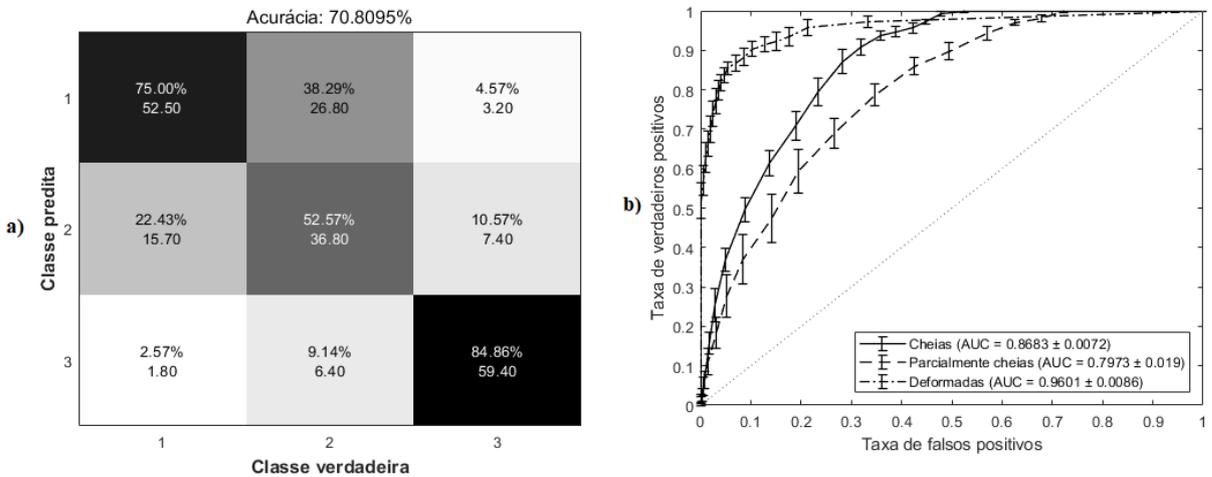
As matrizes de confusão médias e as curvas ROC resultantes dos conjuntos “Haralick duplo” (melhor caso) e “Segmentação” (pior caso), quando consideradas as três classes de sementes em conjunto, são mostradas na Figura 4.8 e Figura 4.9, respectivamente. Como o espaço ROC é bidimensional, cada uma das curvas é feita considerando uma das classes contra todas as demais.

Figura 4.8 – Matriz de confusão (a) e curvas ROC (b) utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



Fonte: do autor.

Figura 4.9 – Matriz de confusão (a) e curvas ROC (b) utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



Fonte: do autor.

As médias das matrizes de confusão e as curvas ROC com suas respectivas AUCs para todos os casos estudados estão disponíveis nos apêndices deste trabalho. A fim de facilitar a diferenciação entre os desempenhos resultantes da utilização de cada conjunto de dados, encontram-se também nos apêndices os intervalos de confiança, dispostos graficamente, das métricas de avaliação utilizadas.

4.7 Discussão

Novamente desconsiderando o conjunto de dados “Histograma”, nota-se que para todas as métricas de avaliação utilizadas a ordem de desempenho é a mesma. O conjunto de dados denominado “Haralick duplo” é o melhor colocado, seguido pelos conjuntos “Haralick” e “Haralick e pixels com equalização de histograma”.

As classificações com a combinação de apenas duas classes, de modo geral, também seguem a mesma ordem de desempenho descrita anteriormente. Entretanto, nota-se um ganho de desempenho desproporcional, se comparados aos outros, dos conjuntos “Segmentação”, “Segmentação (maior área)” e “Máscara elíptica” quando envolvida a classe de sementes deformadas.

É possível observar que a diferenciação entre as sementes cheias e deformadas é a que resulta em melhor desempenho com a utilização de qualquer dos conjuntos de dados. Isso acontece em razão de haver uma maior diferenciação visual entre essas duas classes de sementes. É possível observar que todos os conjuntos de dados formados são suficientemente bons para que os comitês de RNAs consigam diferenciar satisfatoriamente entre essas duas classes de sementes, já que todos resultam em acurácia maior que 90%.

Como esperado em razão da pouca diferença visual, a grande dificuldade foi a diferenciação entre as sementes cheias e parcialmente cheias. De forma geral, essa combinação resultou em menor desempenho para todos os conjuntos de dados e métricas de avaliação utilizadas. Essa baixa diferença visual fica muito evidente ao se comparar os histogramas de sementes dessas duas classes, como mostrado anteriormente na Figura 4.7.

Essa alegação referente à dificuldade de classificação entre sementes cheias e parcialmente cheias é comprovada ao se observar não só os níveis de acurácia, mas as matrizes de confusão e curvas ROC. Nota-se, para todos os conjuntos, menor AUC e maior confundimento entre essas classes. Além disso, o índice Kappa para este caso é sempre menor, indicando baixa concordância entre a classificação real e a realizada pelas RNAs.

Por fim, por ter sido utilizado o mesmo conjunto de dados, os resultados obtidos são comparáveis aos dos trabalhos de Leite, Sáfadi e Carvalho (2013), Sáfadi et al. (2016) e Cassiano (2018), que obtiveram níveis de acurácia de 82%, 57% e 74,5%, respectivamente, quando utilizadas as três classes em conjunto. Dessa forma, os resultados obtidos no presente trabalho representam significativo ganho de desempenho na classificação das sementes de girassol quanto ao seu nível de dano.

5 CONCLUSÃO

Neste capítulo são apresentadas as conclusões sobre o trabalho e sugerida possível utilização para a metodologia desenvolvida. Além disso, como trabalhos futuros, são colocadas opções que possivelmente resultem em ganho de desempenho.

O conjunto de dados formado pela combinação dos descritores de Haralick calculados considerando duas diferentes distâncias mostrou-se como a melhor opção para o treinamento das RNAs. A utilização desse conjunto resultou em comitês com níveis de acurácia de 89,57% com a utilização das três classes em conjunto, 92,07% quando consideradas as classes de sementes cheias e parcialmente cheias, 99,28% quando consideradas as classes de sementes cheias e deformadas e, por fim, 98,43% quando consideradas as classes de sementes parcialmente cheias e deformadas.

Ressalta-se que tais resultados foram obtidos mesmo com a utilização de um conjunto de dados composto por uma quantidade relativamente baixa de amostras, apesar do aumento artificial executado. A quantidade de amostras mostrou-se problemática principalmente ao considerar os conjuntos de dados que utilizam os valores dos pixels como variáveis de entrada, uma vez que isso resulta em uma grande quantidade dessas variáveis para um conjunto limitado de amostras.

O aumento artificial do conjunto de dados contribuiu para amenizar o problema da quantidade de amostras, resultando em aumento significativo dos dados disponíveis e também ocasionando aumento de diversidade no conjunto de dados. Entretanto, é importante ressaltar que a coleta de novos dados seria uma alternativa mais interessante, embora geralmente seja algo dificultoso na prática.

Ressalta-se ainda o ganho de desempenho propiciado pela metodologia utilizada neste trabalho quando comparada aos trabalhos que utilizaram o mesmo conjunto de dados, já citados anteriormente. Considerando a métrica acurácia, o ganho foi de aproximadamente 7,5% quando consideradas as três classes de sementes em conjunto em comparação com o trabalho de melhor desempenho até então.

A metodologia aqui desenvolvida poderia ser utilizada, por exemplo, para compor um comitê de avaliação misto entre humanos e um comitê de RNAs. Portanto, a avaliação das sementes poderia ser realizada pelo comitê de RNAs e por um avaliador humano. Por meio dessa composição, seria necessário um segundo avaliador humano apenas em caso de divergência entre o primeiro avaliador humano e o comitê de RNAs, a fim de dirimir a

dúvida. Tal composição poderia resultar em economia de recursos humanos, temporais e financeiros.

Desta forma, este trabalho sugere a possibilidade da utilização da metodologia nele desenvolvida como composição aos meios de avaliação de sementes usuais atualmente, realizada apenas por avaliadores humanos. A utilização de comitês de RNAs para tal avaliação poderá, portanto, reduzir a subjetividade inerente ao processo atual, garantindo maior precisão e rapidez nas avaliações e auxiliando os avaliadores humanos nas tomadas de decisão, além de potencialmente reduzir o gasto de recursos.

Por fim, ressalta-se a importância de que o emprego da metodologia aqui desenvolvida seja acompanhado da utilização de um conjunto de dados substancialmente maior do que o disponível para a realização deste trabalho. A disponibilidade de uma quantidade maior de amostras para o treinamento das RNAs seguramente resultará em maior diversidade do conjunto de dados e trará melhorias ao aprendizado das RNAs, fazendo com que possuam maior capacidade de generalização e, conseqüentemente, melhor desempenho na classificação das sementes.

5.1 Trabalhos futuros

Propõe-se como trabalhos futuros a criação de um classificador de sementes capaz não apenas de classificar sementes quanto ao seu nível de dano, mas também diferenciar diferentes espécies de sementes em determinadas imagens.

Os métodos utilizados para composição dos conjuntos de dados utilizados neste estudo, com exceção dos descritores de Haralick, não levam em consideração a correlação espacial entre os pixels das imagens. Em razão disso e do melhor desempenho ter sido encontrado utilizando um dos conjuntos de dados composto por descritores de Haralick, é razoável supor que a utilização de outros métodos que considerem essa correlação espacial possam também performar bem neste tipo problema.

Redes Neurais Convolucionais são reconhecidamente eficazes para problemas classificação de imagens, embora notavelmente mais complexas que as RNAs utilizadas neste trabalho. Constituem, assim, uma excelente alternativa para ser utilizada futuramente.

Há ainda a possibilidade da utilização de transferência de aprendizado, utilizando-se de redes convolucionais pré-treinadas e reconhecidamente eficazes para encurtar as

etapas de treinamento, dado que o treinamento deste tipo de rede costuma ser mais custoso em termos temporais e computacionais.

A partir da utilização de CNNs pré treinadas, poderia torna-se viável a criação de um *software* para classificação de sementes de diversas espécies, seja em relação ao nível de dano, qualidade das sementes ou diferenciação entre espécies.

Além da criação de um *software*, há ainda a possibilidade da criação de uma máquina com o sistema classificador embarcado, a fim de realizar a classificação e separação automática das sementes. Isso poderia ser interessante do ponto de vista comercial, já que se tornaria possível precificar de maneira diferente as sementes separadas.

A utilização da técnica BoVW (*Bag of Visual Words*) também é vista como uma alternativa plausível, embora bem mais simples. Essa técnica visa extrair características das imagens para servirem como conjunto de treinamento para algoritmos de classificação, em especial SVM. Semelhante ao histograma, com esta técnica são criadas “palavras visuais”(pequenas partes das imagens) que depois têm suas frequências contadas e podem formar um conjunto de dados para o treinamento de algoritmos classificadores.

REFERÊNCIAS BIBLIOGRÁFICAS

- AHMED, M. R. et al. Classification of watermelon seeds using morphological patterns of x-ray imaging: A comparison of conventional machine learning and deep learning. **Sensors**, MDPI AG, v. 20, n. 23, p. 6753, nov 2020.
- ALI, M. et al. An accurate CT saturation classification using a deep learning approach based on unsupervised feature extraction and supervised fine-tuning strategy. **Energies**, MDPI AG, v. 10, n. 11, p. 1830, 2017.
- ALVES, W. A. L.; ARAÚJO, S. A. de. Avaliação da utilização de dois diferentes modelos de redes neurais na classificação de imagens com base em características texturais. **Exacta**, Universidade Nove de Julho, v. 4, n. 1, p. 77–86, 2006.
- AMAZON WEB SERVICE. **Amazon Machine Learning: guia do desenvolvedor**. Seattle (Estados Unidos da América), 2020. Disponível em: <https://docs.aws.amazon.com/pt_br/machine-learning/latest/dg/machinelearning-dg.pdf>. Acesso em: 12 abr. 2020.
- BECKER, S. Unsupervised learning procedures for neural networks. v. 02, p. 17–33, 1991. ISSN 0129-0657.
- BELLMAN, R. **Adaptive control processes: a guided tour**. New Jersey, EUA: Princeton University Press, 1961.
- BINO, R. J.; AARTSE, J. W.; BURG, W. J. van der. Non-destructive x-ray analysis of arabidopsis embryo mutants. **Seed Science Research**, Cambridge University Press, v. 3, n. 3, p. 167–170, 1993.
- BRADLEY, A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. **Pattern Recognition**, Elsevier BV, v. 30, n. 7, p. 1145–1159, jul 1997.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. **Regras para análise de sementes**. Brasília, 2009. 399 p.
- BREIMAN, L. Bagging predictors. **Machine learning**, Springer, v. 24, n. 2, p. 123–140, 1996.
- BRIDLE, J. S. Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In: **Neurocomputing**. Heidelberg (Alemanha): Springer, 1990. v. 68, p. 227–236.
- BURG, W. J. van der et al. Predicting tomato seedling morphology by x-ray analysis of seeds. **Journal of the American Society for Horticultural Science**, American Society for Horticultural Science, v. 119, n. 2, p. 258–263, 1994.
- CARVALHO, M. L. de et al. Pre-harvest stress cracks in maize (*Zea mays* L.) kernels as characterized by visual, x-ray and low temperature scanning electron microscopical analysis: effect on kernel quality. **Seed Science Research**, Cambridge University Press, v. 9, n. 3, p. 227–236, 1999.

CASSIANO, F. R. **Expoente direcional de Hurst na análise de similaridade de imagens de sementes**. Tese (Doutorado em Estatística e Experimentação Agropecuária) — Universidade Federal de Lavras, 2018.

CASTRO, C. L. de; BRAGA, A. P. Aprendizado supervisionado com conjuntos de dados desbalanceados. **Sba Controle & Automação**, v. 22, n. 5, p. 441–466, 2011.

CHEN, L. Curse of dimensionality. In: LIU, L.; ÖZSU, M. T. (Ed.). **Encyclopedia of Database Systems**. Boston (Estados Unidos): Springer, 2009. p. 545–546.

CICERO, S. M. et al. Evaluation of mechanical damage in seeds of maize (zea mays l.) by x-ray and digital imaging. **Seed Science and Technology**, v. 26, p. 603–612, 1998.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, Sage Publications Sage CA: Thousand Oaks, CA, v. 20, n. 1, p. 37–46, 1960.

COHEN, J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. **Psychological Bulletin**, American Psychological Association (APA), v. 70, n. 4, p. 213–220, 1968.

CONAB – Companhia Nacional de Abastecimento. **Série histórica das safras – Grãos por produtos**. Brasília, 2020. Disponível em: <<https://www.conab.gov.br/info-agro/safras/serie-historica-das-safras>>. Acesso em: 16 dez. 2020.

CONCI, A. **Análise de Imagens**. 2015. Universidade Federal Fluminense. Notas de aula. Disponível em: <<http://www.ic.uff.br/~aconci/aula-2-2015-AI.pdf>>. Acesso em: 10 abr. 2020.

CONGALTON, R.; GREEN, K. **Assessing the Accuracy of Remotely Sensed Data: Principles and Practices**. 2. ed. Boca Rato (Estados Unidos): CRC Press, 2009.

CONNERS, R. W.; TRIVEDI, M. M.; HARLOW, C. A. Segmentation of a high-resolution urban scene using texture operators. **Computer vision, graphics, and image processing**, Elsevier, v. 25, n. 3, p. 273–310, 1984.

CRAVIOTTO, R. M. et al. Description of a pure seed fraction of oat through usual evaluations and radiographic images. **Pesquisa Agropecuária Brasileira**, SciELO Brasil, v. 37, n. 8, p. 1183–1188, 2002.

DATA SCIENCE ACADEMY. **Deep Learning Book**. Brasília, Distrito Federal, Brasil.: DAS, 2019. Disponível em: <<http://www.deeplearningbook.com.br>>. Acesso em: 26 ago. 2019.

FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, Elsevier BV, v. 27, n. 8, p. 861–874, 2006.

FERRI, C.; HERNÁNDEZ-ORALLO, J.; SALIDO, M. A. Volume under the ROC surface for multi-class problems. In: **Machine Learning: ECML 2003**. Heidelberg (Alemanha): Springer Berlin Heidelberg, 2003. p. 108–120.

- FLEISS, J. L. Measuring nominal scale agreement among many raters. **Psychological Bulletin**, American Psychological Association (APA), v. 76, n. 5, p. 378–382, 1971.
- FLEISS, J. L.; COHEN, J.; EVERITT, B. S. Large sample standard errors of kappa and weighted kappa. **Psychological Bulletin**, American Psychological Association (APA), v. 72, n. 5, p. 323–327, 1969.
- FLEISS, J. L.; LEVIN, B.; PAIK, M. C. **Statistical Methods for Rates and Proportions**. 3. ed. Hoboken (Estados Unidos): John Wiley & Sons, 2003.
- FONSECA, R.; SILVA, P.; SILVA, R. Acordo inter-juízes: O caso do coeficiente kappa. **Laboratório de Psicologia**, ISPA - Instituto Universitario, v. 5, n. 1, nov 2013.
- FREUND, Y.; SCHAPIRE, R. E. Experiments with a new boosting algorithm. In: **Proceedings of the Thirteenth International Conference on International Conference on Machine Learning**. São Francisco (Estados Unidos): Morgan Kaufmann Publishers, 1996. (ICML 96), p. 148–156.
- GONZALEZ, R. C.; WOODS, R. E. **Digital image processing**. 2. ed. Upper Saddle River (Estados Unidos): Prentice Hall, 2002. 793 p.
- GRUNVALD, A. K. et al. Adaptabilidade e estabilidade de genótipos de girassol nos estados do rio grande do sul e paraná. **Ciência e Agrotecnologia**, v. 33, n. 9, p. 1195–1204, 2009.
- HAND, D. J.; TILL, R. J. **Machine Learning**, Springer Science and Business Media LLC, v. 45, n. 2, p. 171–186, 2001.
- HANLEY, J. A.; MCNEIL, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. **Radiology**, v. 143, p. 29–36, 1982.
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. H. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, n. 6, p. 610–621, 1973.
- HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001.
- HE, H.; MA, Y. **Imbalanced Learning: Foundations, Algorithms, and Applications**. Hoboken (Estados Unidos): John Wiley & Sons, 2013.
- HEBB, D. O. **The organization of behavior: A Neuropsychological Theory**. Nova Iorque (Estados Unidos): John Wiley & Sons, 1949.
- HUSSAIN, L.; AJAZ, R. Seed classification using machine learning techniques. **Journal of Multidisciplinary Engineering Science and Technology (JMEST)**, v. 02, p. 1098–1102, 05 2015.
- ISHIDA, C. Y.; POZO, A. T. Pareto front elite. In: **XXVII Congresso SBC / VI ENIA - Encontro Nacional de Inteligência Artificial**. Rio de Janeiro: [s.n.], 2007.
- ISTA. **International rules for seed testing**. Zurique (Suíça), 2004. Manual Técnico.

- JONES, M. T. A neural networks deep dive. 2017. Disponível em: <<https://developer.ibm.com/articles/cc-cognitive-neural-networks-deep-dive/>>. Acesso em: 12 ago. 2019.
- KHATCHATOURIAN, O.; PADILHA, F. R. R. Reconhecimento de variedades de soja por meio do processamento de imagens digitais usando redes neurais artificiais. **Engenharia Agrícola**, FapUNIFESP (SciELO), v. 28, n. 4, p. 759–769, 2008.
- KIRATIRATANAPRUK, K.; SINTHUPINYO, W. Color and texture for corn seed classification by machine vision. In: **International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS)**. Chiang Mai (Tailândia): IEEE, 2011.
- KURTULMUŞ, F.; ALIBAS, I.; KAVDIR, I. Classification of pepper seeds using machine vision based on neural network. **International Journal of Agricultural and Biological Engineering**, v. 9, p. 51–62, 12 2015.
- LAI, K. K. et al. A novel nonlinear neural network ensemble model for financial time series forecasting. In: SPRINGER. **International Conference on Computational Science**. Heidelberg (Alemanha), 2006. p. 790–793.
- LANDGREBE, T.; DUIN, R. A simplified volume under the roc hypersurface. **SAIEE Africa Research Journal**, v. 98, 11 2006.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, p. 159–174, 1977.
- LEITE, I.; SÁFADI, T.; CARVALHO, M. Evaluation of seed radiographic images by independent component analysis and discriminant analysis. **Seed Science and Technology**, International Seed Testing Association, v. 41, n. 2, p. 235–244, aug 2013.
- LIMA, C. A. d. M. **Comitê de Máquinas: uma abordagem unificada empregando máquinas de vetores-suporte**. Tese (Doutorado em Engenharia Elétrica) — UNICAMP - Universidade Estadual de Campinas, 2004.
- LIU, D. et al. Discriminating and elimination of damaged soybean seeds based on image characteristics. **Journal of Stored Products Research**, Elsevier BV, v. 60, p. 67–74, jan 2015.
- MAHMOOD, H. Activation functions in neural networks. 2018. Disponível em: <<https://towardsdatascience.com/activation-functions-in-neural-networks-83ff7f46a6bd/>>. Acesso em: 13 ago. 2019.
- MARQUES FILHO, O.; VIEIRA NETO, H. **Processamento digital de imagens**. Rio de Janeiro: Brasport, 1999.
- MATHWORKS INC. **Choose a Multilayer Neural Network Training Function**. Natick (Estados Unidos da América), c2020. Disponível em: <<https://www.mathworks.com/help/deeplearning/ug/choose-a-multilayer-neural-network-training-function.html>>. Acesso em: 17 nov. 2020.

- MATSUBARA, E. T. **Relações entre Ranking, Análise ROC e Calibração em Aprendizado de Máquina**. Tese (Doutorado em Ciências de Computação e Matemática Computacional) — Instituto de Ciências Matemáticas e de Computação - ICMC-USP, 2008.
- MATSUNAGA, L. A. **Uma Metodologia de Categorização Automática de Textos para a Distribuição dos Projetos de Lei às Comissões Permanentes da Câmara Legislativa do Distrito Federal**. Tese (Doutorado em Sistemas Computacionais) — Universidade Federal do Rio de Janeiro, 2007.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.
- MEDEIROS, G. C. F. **Uso de texturas para o acompanhamento da evolução do tratamento de úlceras dermatológicas**. Dissertação (Mestrado em Engenharia Elétrica) — Universidade de São Paulo, São Carlos, 2001.
- MENEZES, N. L. de; CÍCERO, S. M.; VILLELA, F. A. Identificação de fissuras em sementes de arroz após a secagem artificial, por meio de raios-x. **Ciência Rural**, Universidade Federal de Santa Maria, v. 35, n. 5, p. 1194–1196, 2005.
- MIKULSKI, B. Understanding the softmax activation function. 2019. Disponível em: <www.mikulskibartosz.name/understanding-the-softmax-activation-function>. Acesso em: 13 ago. 2019.
- MINSKY, M.; PAPER, S. **Perceptron: an introduction to computational geometry**. MIT Press, Cambridge (Estados Unidos), 1969.
- MONZEL, R. **Haralick Texture Features v. 1.3.1.0**. MATLAB Central File Exchange, 2018. Disponível em: <<https://www.mathworks.com/matlabcentral/fileexchange/58769-haralicktexturefeatures>>. Acesso em: 17 ago. 2019.
- NABEEL, A.; QUTEISHAT, A. Wheat seeds classification using multi-layer perceptron artificial neural network. **International Journal of Electronics Communication and Computer Engineering**, v. 6, p. 2278–4209, 06 2015.
- NCSS - STATISTICAL SOFTWARE. **Power Analysis & Sample Size Documentation**. Kaysville (Estados Unidos), 2020. Disponível em: <https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/Confidence_Intervals_for_Kappa.pdf>.
- OLSON, D. L.; DELEN, D. Performance evaluation for predictive modeling. In: **Advanced Data Mining Techniques**. Heidelberg (Alemanha): Springer-Verlag, 2008. p. 137–147.
- OSÓRIO, F. S.; VIEIRA, R. Sistemas híbridos inteligentes. In: **ENIA-Encontro Nacional de Inteligência Artificial/Congresso da SBC**. Rio de Janeiro: [s.n.], 1999. p. 1–59.
- PACIFICI, F.; CHINI, M.; EMERY, W. J. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use

classification. **Remote Sensing of Environment**, Elsevier, v. 113, n. 6, p. 1276–1292, 2009.

PALIT, A. K.; POPOVIC, D. Nonlinear combination of forecasts using artificial neural network, fuzzy logic and neuro-fuzzy approaches. In: IEEE. **Ninth IEEE International Conference on Fuzzy Systems**. San Antonio (Estados Unidos), 2000. v. 2, p. 566–571.

PALMIERE, S. E. Inteligência artificial - arquiteturas e topologias de redes neurais artificiais. 2016. Disponível em: <www.embarcados.com.br/redes-neurais-artificiais>. Acesso em: 10 ago. 2019.

PASSO, D. P. **Análise da qualidade de classificadores para identificação de alvos urbanos em imagens de alta resolução espacial - Uma aplicação com as imagens do satélite WORLDVIEW II**. Dissertação (Mestrado em Geociências) — Universidade de Brasília, Brasília, 2013.

PERRONE, M.; COOPER, L. When networks disagree: Ensemble methods for hybrid neural networks. In: **Neural networks for speech and image processing**. Londres (Inglaterra): Chapman and Hall, 1993.

PIZER, S. M. et al. Adaptive histogram equalization and its variations. **Computer vision, graphics, and image processing**, Elsevier, v. 39, n. 3, p. 355–368, 1987.

PROVOST, F.; DOMINGOS, P. **Well-trained PETs: Improving probability estimation trees**. Nova Iorque (Estados Unidos): Stern School of Busines, 2000. Relatório Técnico CeDER Working Paper IS-00-04.

RAUBER, T. W. **Redes neurais artificiais**. Vitória, 1997. Apostila. Disponível em: <http://www.inf.ufes.br/~thomas/html/disc_rn.html>. Acesso em: 13 ago. 2019.

RIEDMILLER, M.; BRAUN, H. A direct adaptive method for faster backpropagation learning: the RPROP algorithm. In: **IEEE International Conference on Neural Networks**. San Francisco (Estados Unidos): IEE, 1993. p. 586–591.

ROSENBLATT, F. **The perceptron: a perceiving and recognizing automaton**. Ithaca (Estados Unidos): Cornell Aeronautical Laboratory, 1957. Relatório Técnico 85-460-1.

RUMELHART, D.; HINTON, G.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986.

SÁFADI, T. et al. Wavelet-based spectral descriptors for detection of damage in sunflower seeds. **International Journal of Wavelets, Multiresolution and Information Processing**, World Scientific Publishing Company, v. 14, n. 04, p. 1–13, 2016.

SCHAPIRE, R. E. The strength of weak learnability. **Machine learning**, Springer, v. 5, n. 2, p. 197–227, 1990.

SHARKEY, A. J.; SHARKEY, N. E.; CHANDROTH, G. O. Neural nets and diversity. In: RABE, G. (Ed.). **Safe Comp 95**. Londres (Inglaterra): Springer London, 1995. p. 375–389.

- SHARMA V, A. Understanding activation functions in neural networks. 2017. Disponível em: <<https://medium.com/the-theory-of-everything/understanding-activationfunctions-in-neural-networks-9491262884e0>>. Acesso em: 15 ago. 2019.
- SHMUELI, B. Multi-class metrics made simple, part I: Precision and recall. 2019. Disponível em: <<https://towardsdatascience.com/multi-class-metrics-made-simple-part-i-precision-and-recall-9250280bddc2>>. Acesso em: 08 mai. 2020.
- SILVA, A. C. O. **Classificação de áreas cafeeiras em imagens multiespectrais através de redes neurais artificiais e extratores de textura**. 121 p. Monografia (Bacharelado em Ciência de Computação) — Universidade Federal de Lavras, Lavras, 2013.
- SILVA, R. de Souza e; PAES Ângela T. Teste de concordância Kappa. **Educação Continuada em Saúde Einstein**, v. 10, n. 4, p. 165–166, 2012. Disponível em: <<http://apps.einstein.br/revista/ecs/index.asp?SPREEdic=36>>. Acesso em: 10 mai. 2020.
- SIMAK, M. Testing of forest tree and shrub seeds by x-radiography. In: GORDON P. GOSLING, B. W. A. (Ed.). **Tree and shrub seed handbook**. Madison (Estados Unidos): International Seed Testing Association, 1991. p. 1–28.
- SIMAK, M.; GUSTAFSSON, Å. X-ray photography and sensitivity in forest tree species. **Hereditas**, Wiley Online Library, v. 39, n. 3-4, p. 458–468, 1953.
- SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **Information Processing & Management**, Elsevier BV, v. 45, n. 4, p. 427–437, 2009.
- SWALIN, A. Choosing the right metric for evaluating machine learning models — part 2. 2018. Disponível em: <<https://medium.com/usf-msds/choosing-the-right-metric-for-evaluating-machine-learning-models-part-2-86d5649a5428>>. Acesso em: 11 maio. 2020.
- TOU, J. Y.; LAU, P. Y.; TAY, Y. H. Computer vision-based wood recognition system. In: **International workshop on advanced image technology**. Bangkok (Tailândia): Thai Embedded Systems Association (TESA), 2007. p. 197–202.
- TSCHORN, P. Accuracy, precision and recall: Multi-class performance metrics for supervised learning—elixir. 2019. Disponível em: <<https://tech.labs.oliverwyman.com/blog/2019/10/17/accuracy-precision-recal-elixir/>>. Acesso em: 11 maio. 2020.
- TUJO, T. A predictive model to predict seed classes using machine learning. **International Journal of Engineering and Technical Research**, v. 6, p. 334–344, 12 2019.
- TUMER, K.; GHOSH, J. Error correlation and error reduction in ensemble classifiers. **Connection science**, Taylor & Francis, v. 8, n. 3-4, p. 385–404, 1996.
- TYAGI, V. **Understanding Digital Image Processing**. 1. ed. Boca Raton (Estados Unidos): CRC Press, 2018. 374 p.

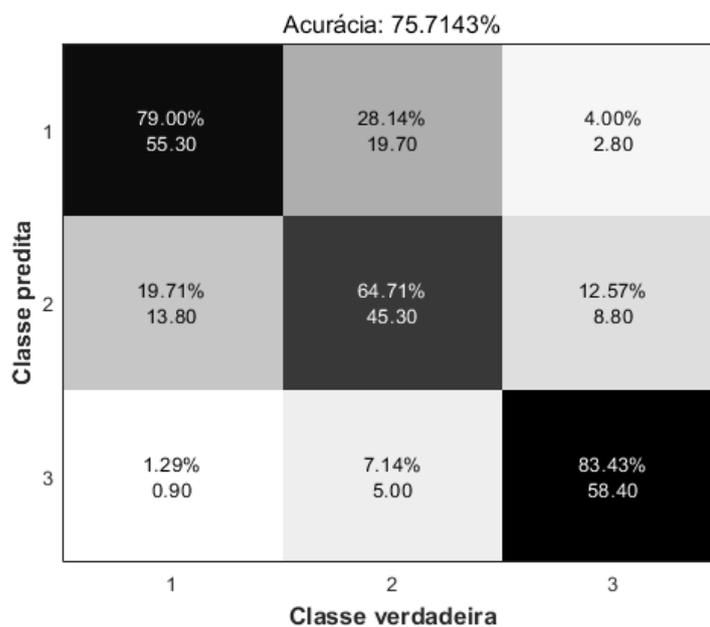
WANG, D.; RAM, M. S.; DOWELL, F. E. Classification of damaged soybean seeds using near-infrared spectroscopy. **Transactions of the ASAE**, American Society of Agricultural and Biological Engineers (ASABE), v. 45, n. 6, 2002.

WESZKA, J. S.; DYER, C. R.; ROSENFELD, A. A comparative study of texture measures for terrain classification. **IEEE Transactions on Systems, Man, and Cybernetics**, IEEE, n. 4, p. 269–285, 1976.

ZHANG, J.; DAI, L.; CHENG, F. Identification of corn seeds with different freezing damage degree based on hyperspectral reflectance imaging and deep learning method. **Food Analytical Methods**, Springer Science and Business Media LLC, 2020.

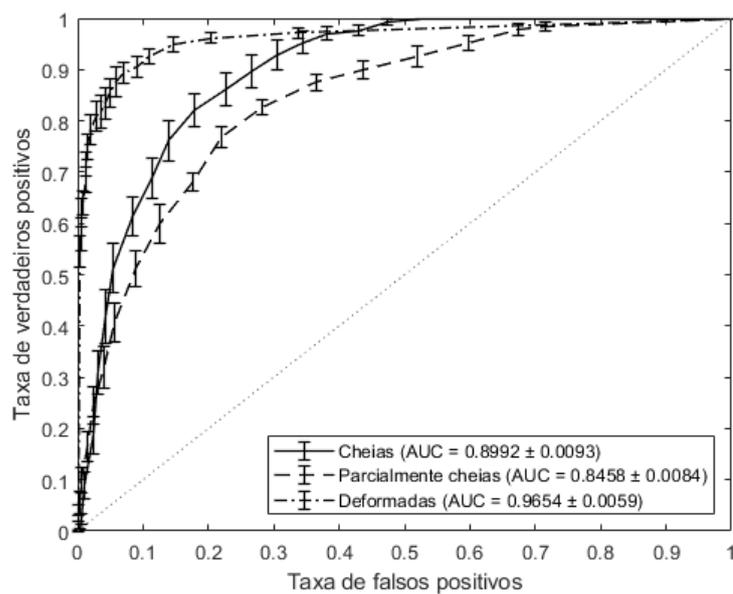
APÊNDICE A – Matrizes de confusão e curvas ROC do conjunto de dados
“Valores dos pixels sem equalização”

Figura 1 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



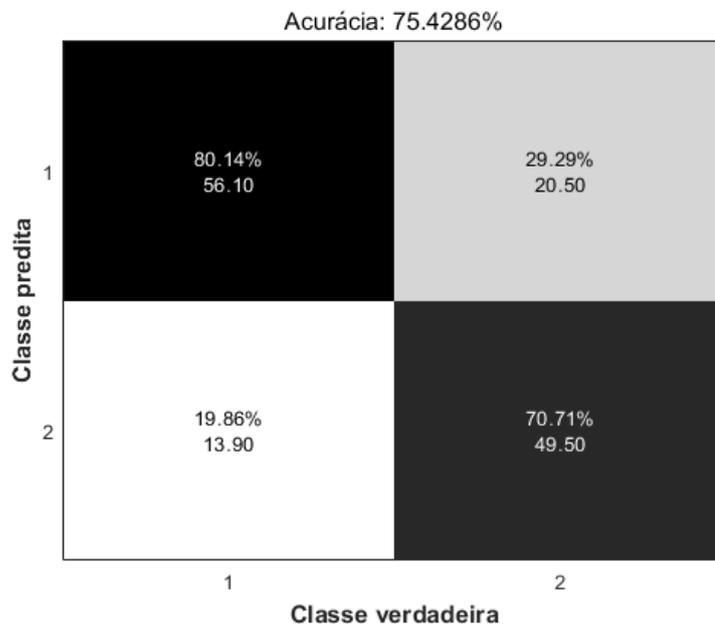
Fonte: do autor.

Figura 2 – Curvas ROC utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



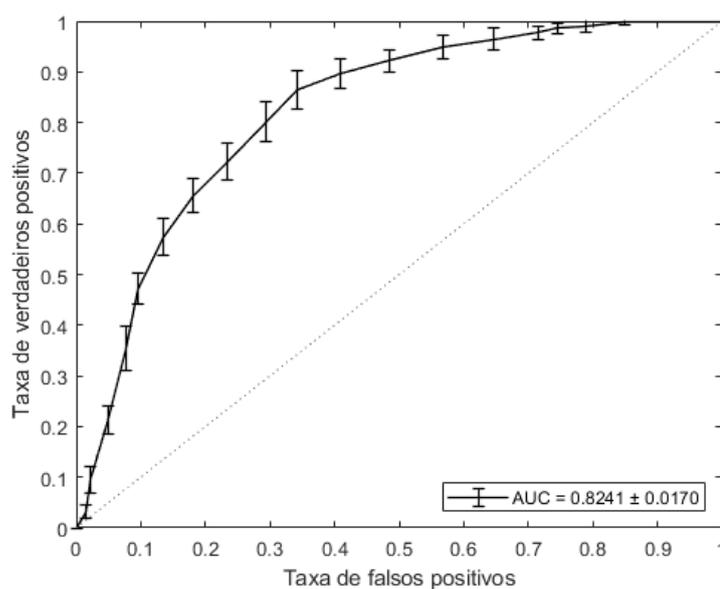
Fonte: do autor.

Figura 3 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



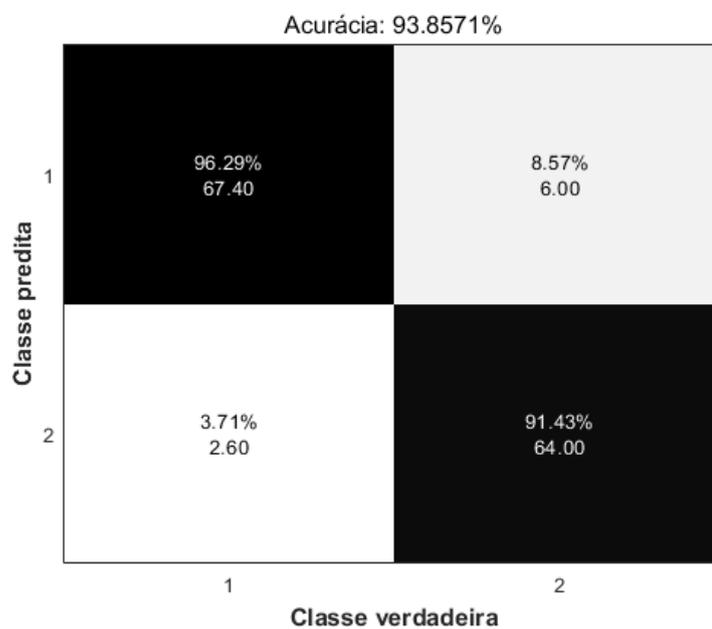
Fonte: do autor.

Figura 4 – Curva ROC utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes cheias e parcialmente cheias.



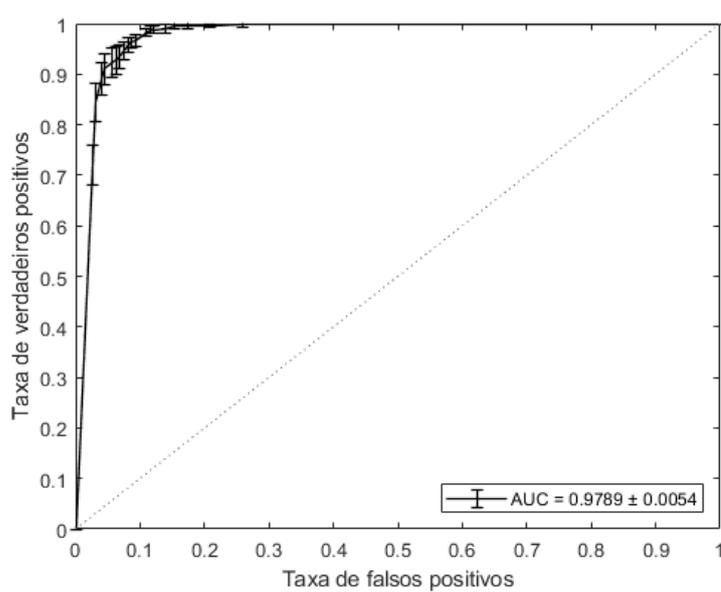
Fonte: do autor.

Figura 5 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes: (1) cheias e (2) deformadas.



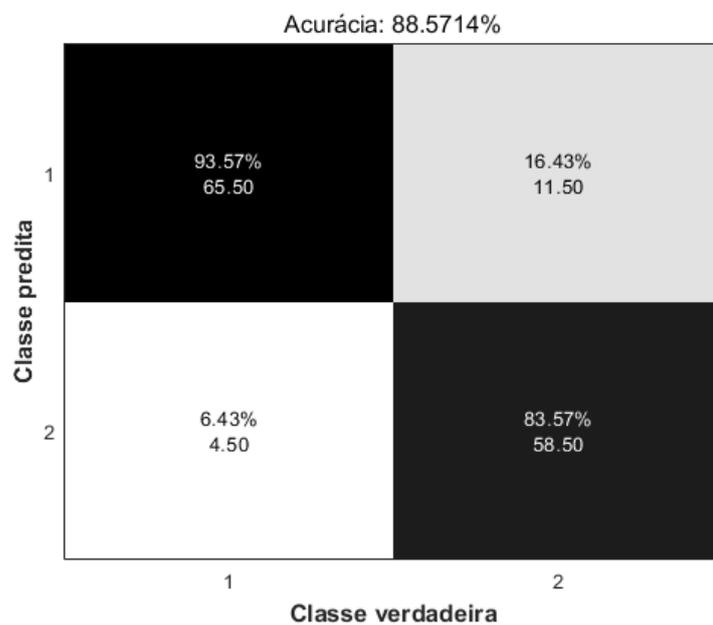
Fonte: do autor.

Figura 6 – Curva ROC utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes cheias e deformadas.



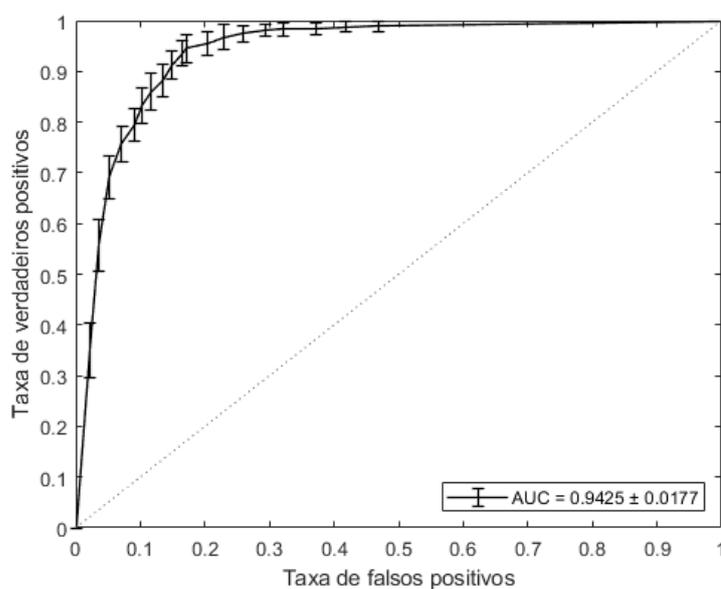
Fonte: do autor.

Figura 7 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.



Fonte: do autor.

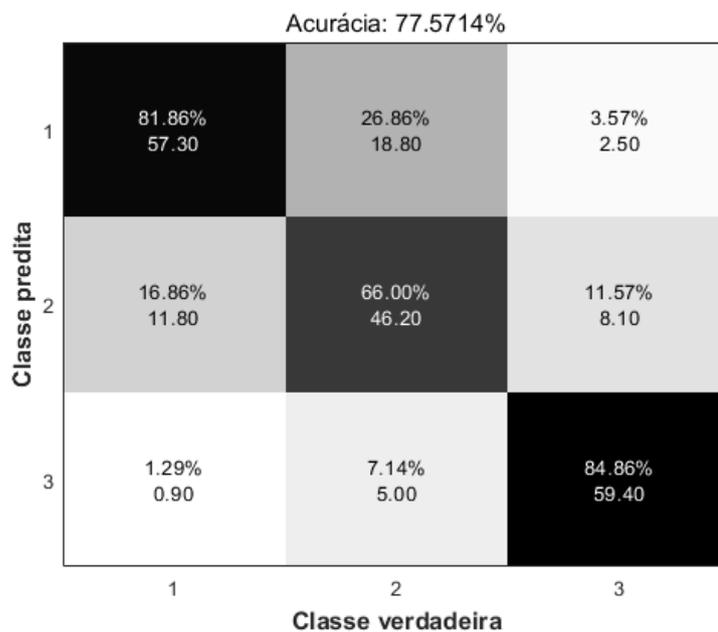
Figura 8 – Curva ROC utilizando o conjunto de dados “Valores dos pixels sem equalização” e considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

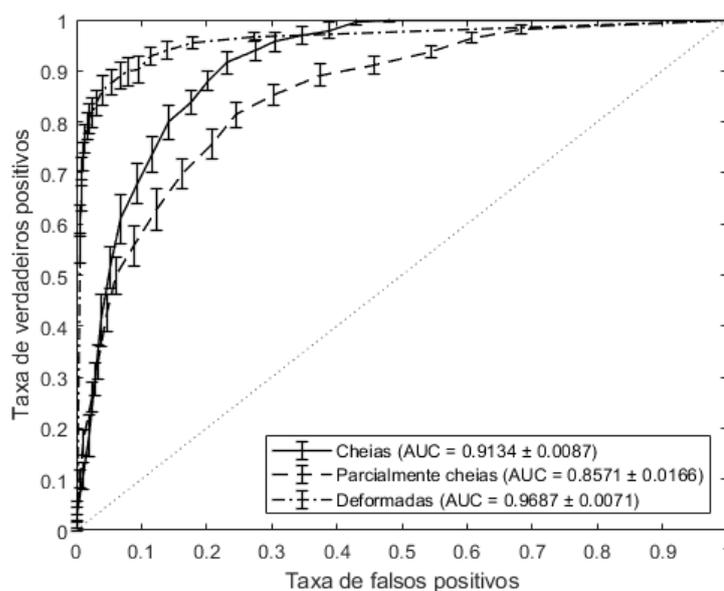
APÊNDICE B – Matrizes de confusão e curvas ROC do conjunto de dados
“Valores dos pixels com equalização”

Figura 9 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



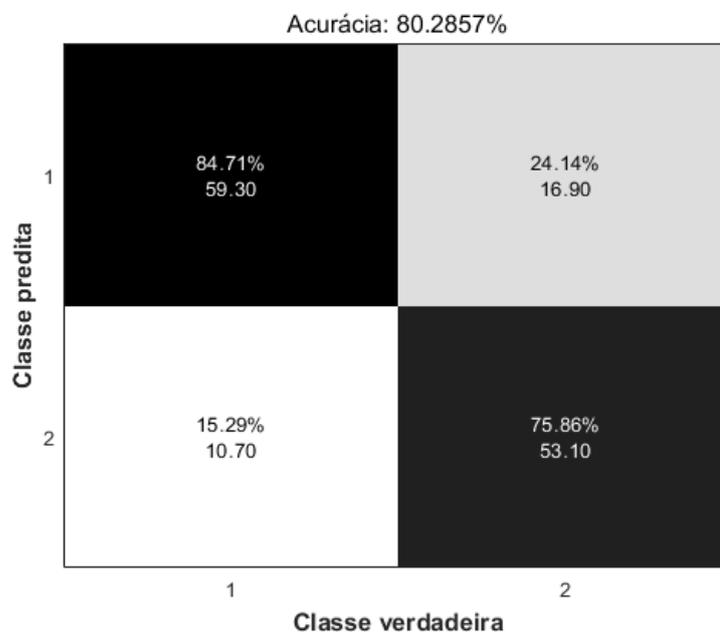
Fonte: do autor.

Figura 10 – Curvas ROC utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



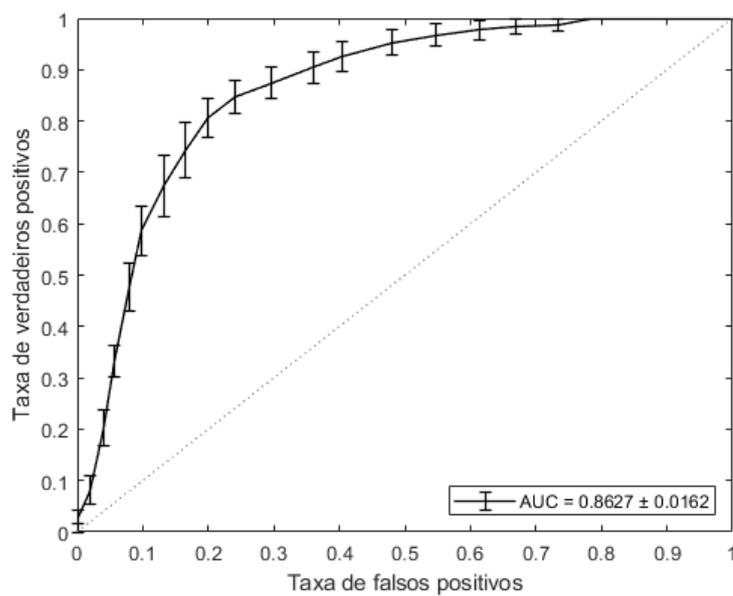
Fonte: do autor.

Figura 11 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



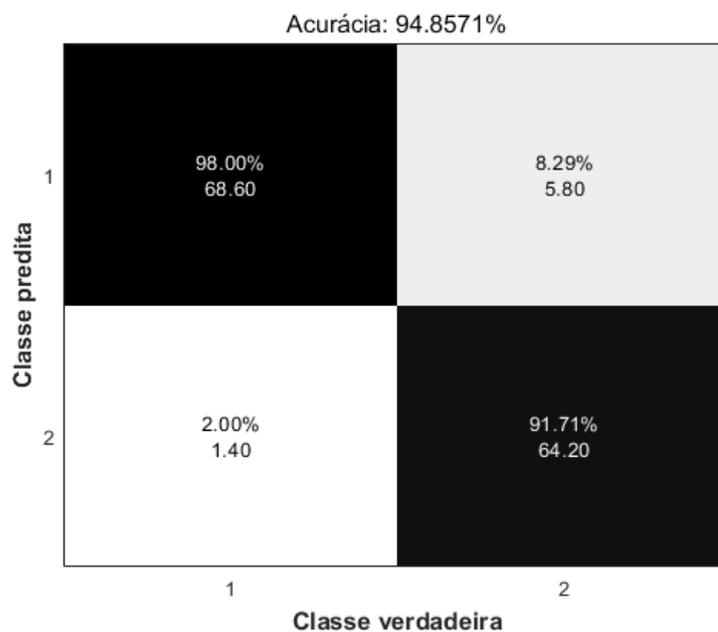
Fonte: do autor.

Figura 12 – Curva ROC utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes cheias e parcialmente cheias.



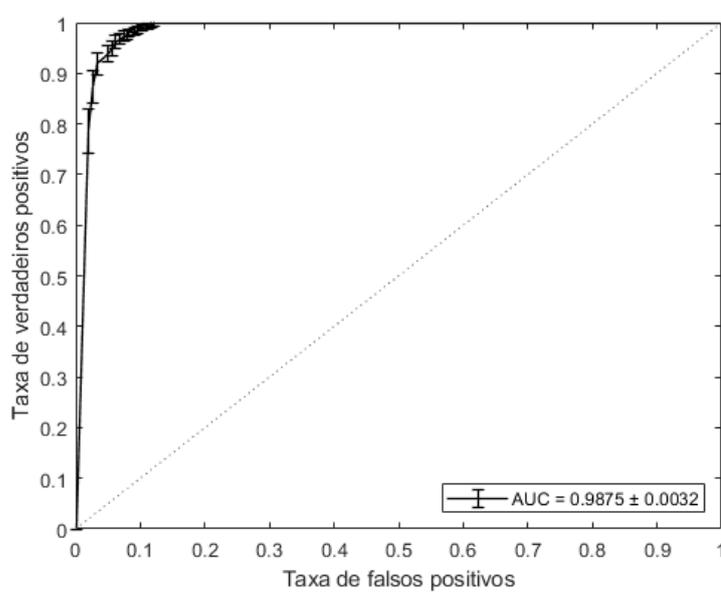
Fonte: do autor.

Figura 13 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias e (2) deformadas.



Fonte: do autor.

Figura 14 – Curva ROC utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes cheias e deformadas.



Fonte: do autor.

Figura 15 – Matriz de confusão utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.

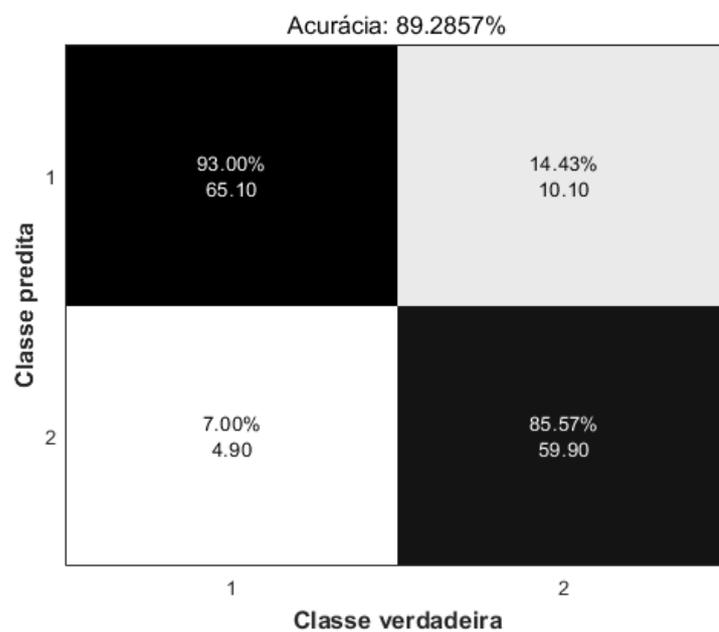
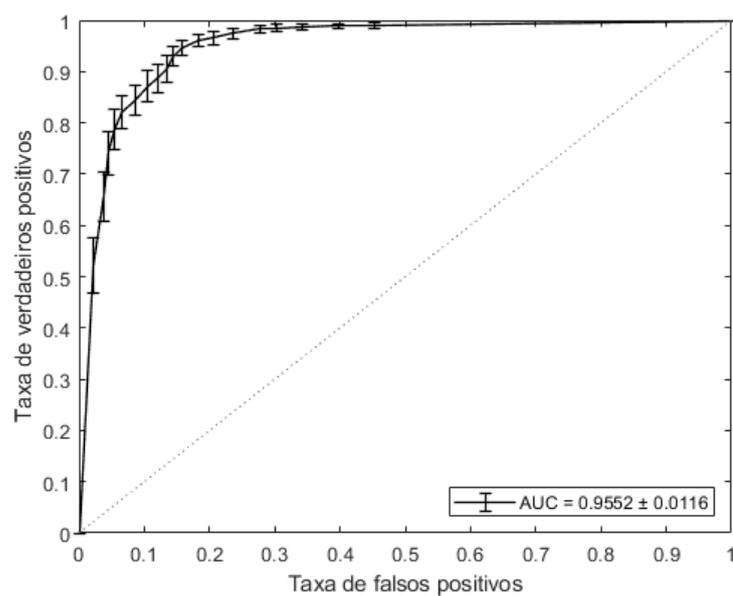
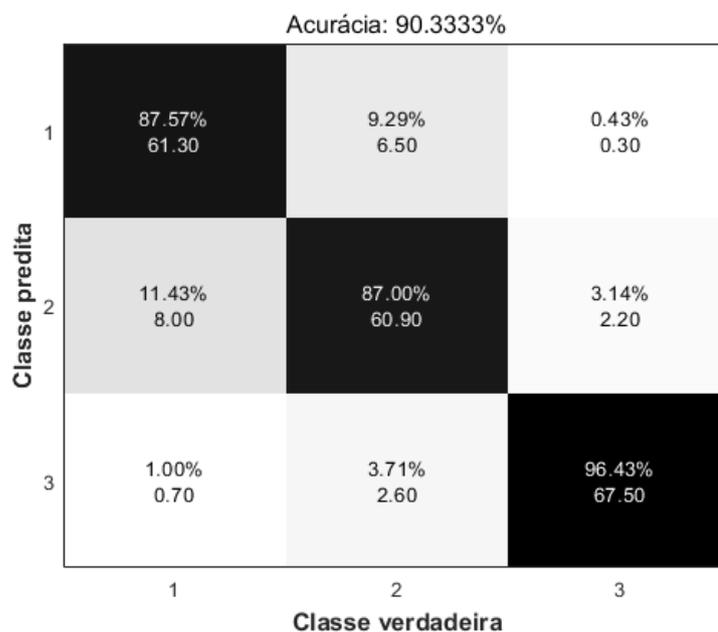


Figura 16 – Curva ROC utilizando o conjunto de dados “Valores dos pixels com equalização” e considerando as classes de sementes parcialmente cheias e deformadas.



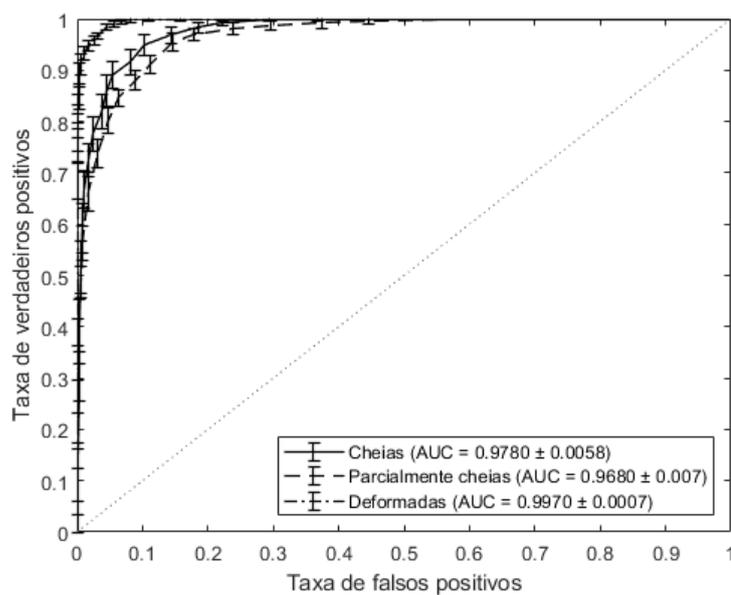
APÊNDICE C – Matrizes de confusão e curvas ROC do conjunto de dados
“Histograma”

Figura 17 – Matriz de confusão utilizando o conjunto de dados “Histograma” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



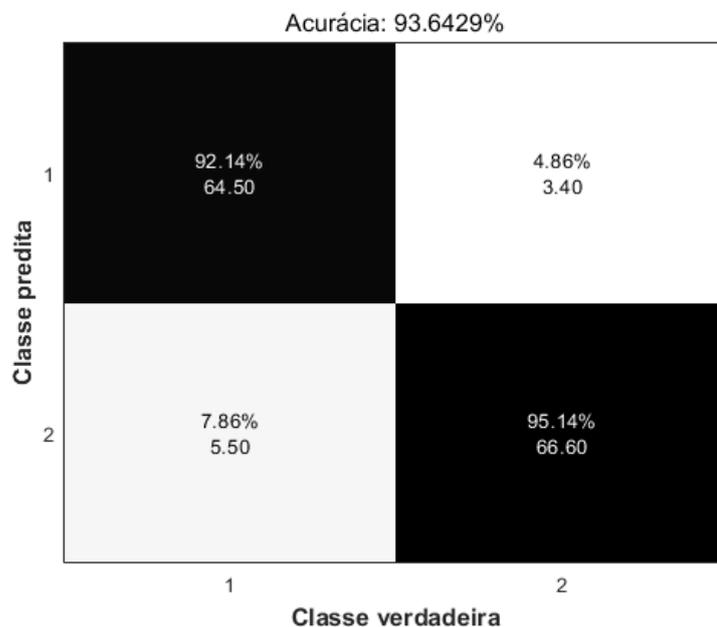
Fonte: do autor.

Figura 18 – Curvas ROC utilizando o conjunto de dados “Histograma” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



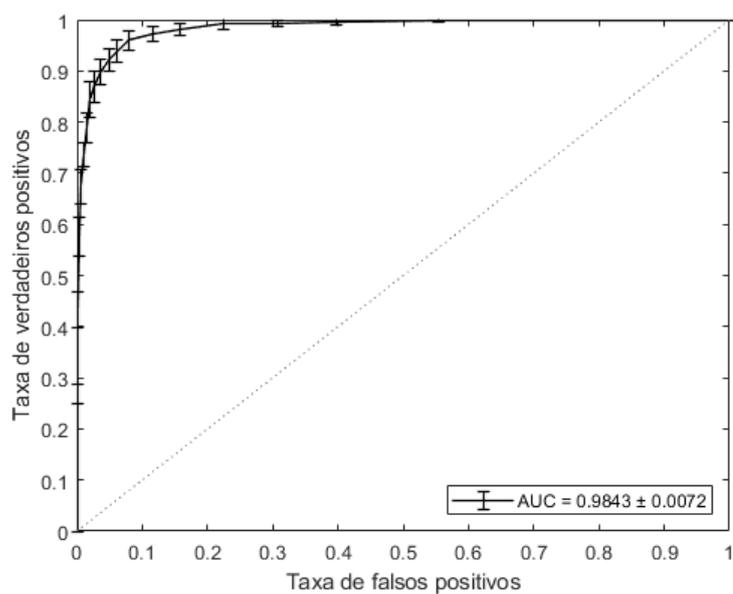
Fonte: do autor.

Figura 19 – Matriz de confusão utilizando o conjunto de dados “Histograma” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



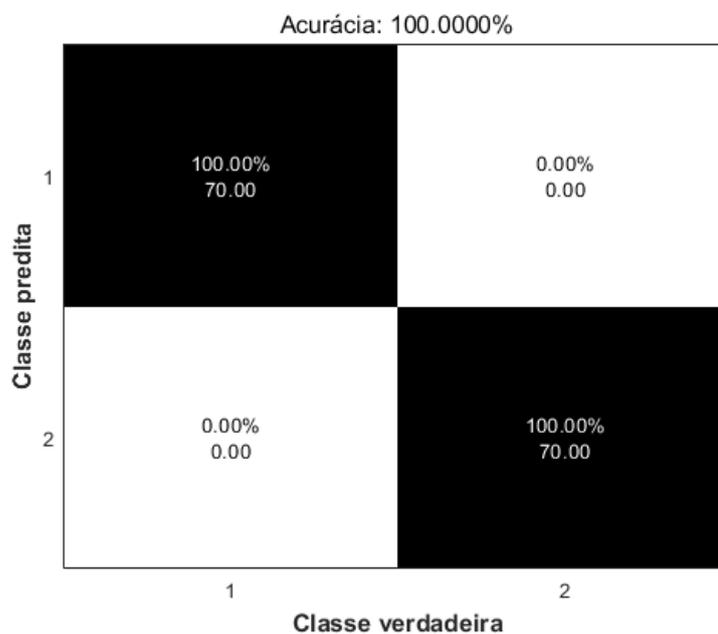
Fonte: do autor.

Figura 20 – Curva ROC utilizando o conjunto de dados “Histograma” e considerando as classes de sementes cheias e parcialmente cheias.



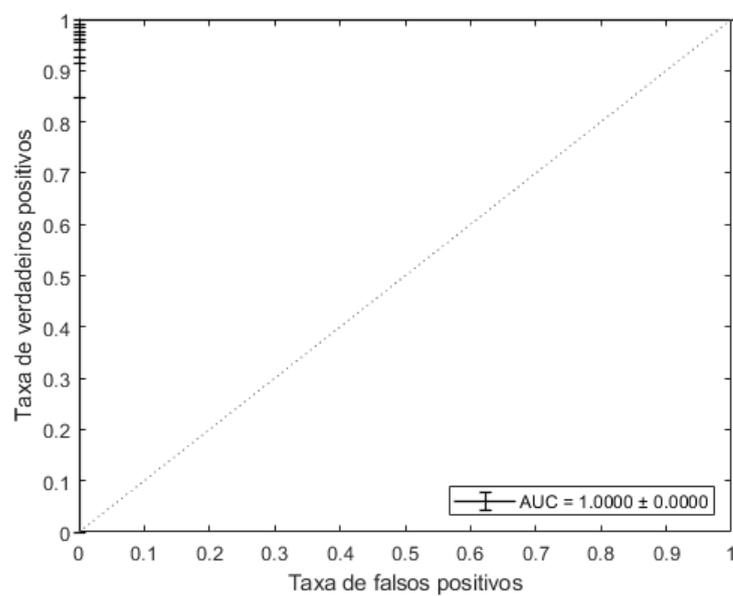
Fonte: do autor.

Figura 21 – Matriz de confusão utilizando o conjunto de dados “Histograma” e considerando as classes de sementes: (1) cheias e (2) deformadas.



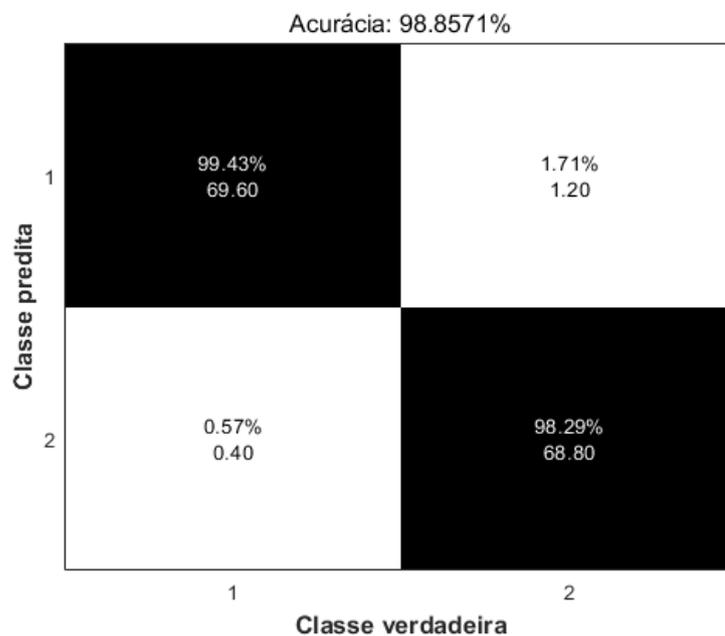
Fonte: do autor.

Figura 22 – Curva ROC utilizando o conjunto de dados “Histograma” e considerando as classes de sementes cheias e deformadas.



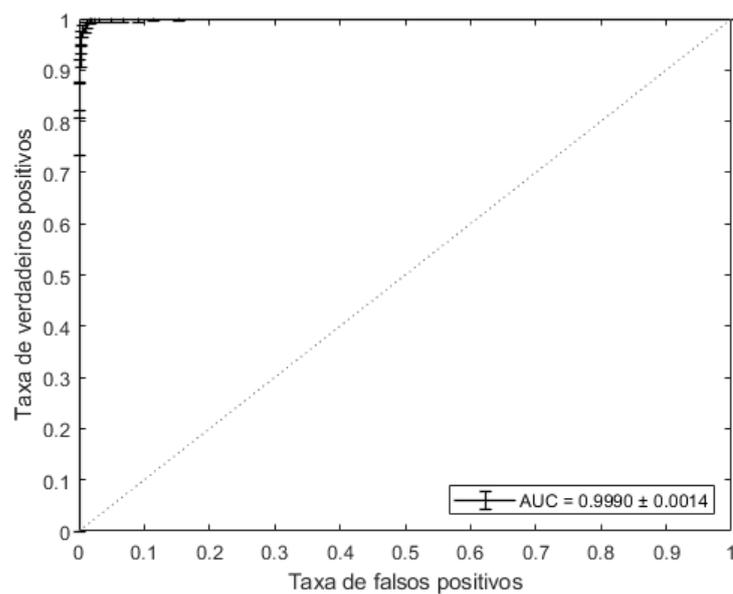
Fonte: do autor.

Figura 23 – Matriz de confusão utilizando o conjunto de dados “Histograma” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.



Fonte: do autor.

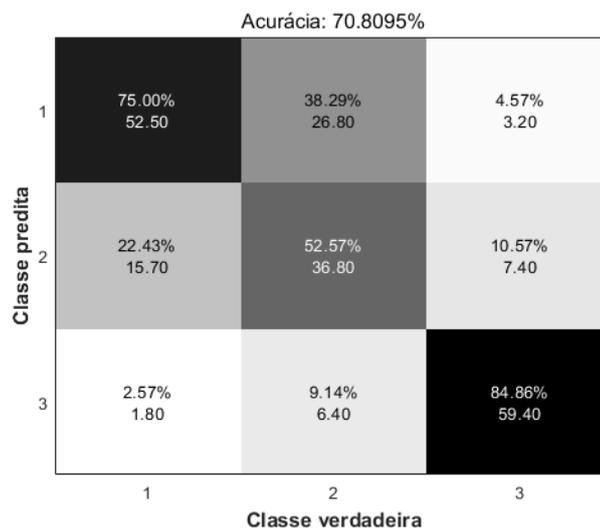
Figura 24 – Curva ROC utilizando o conjunto de dados “Histograma” e considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

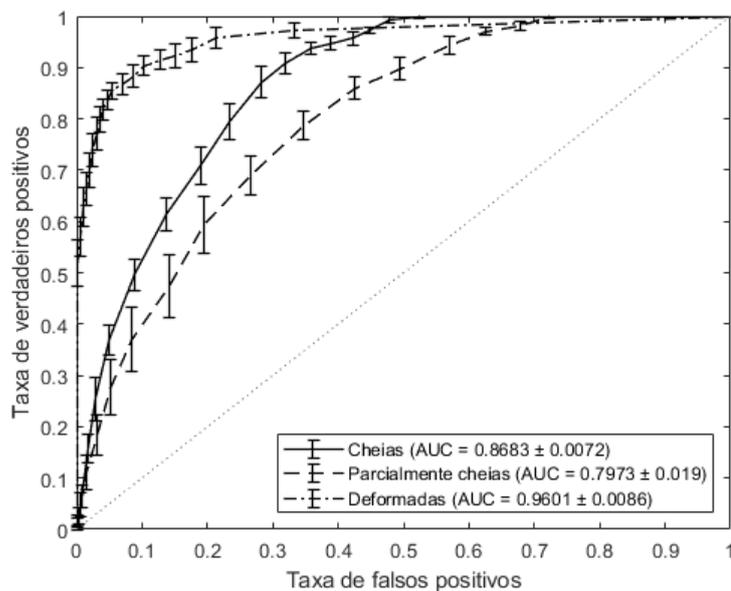
APÊNDICE D – Matrizes de confusão e ROC do conjunto de dados
“Segmentação”

Figura 25 – Matriz de confusão utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



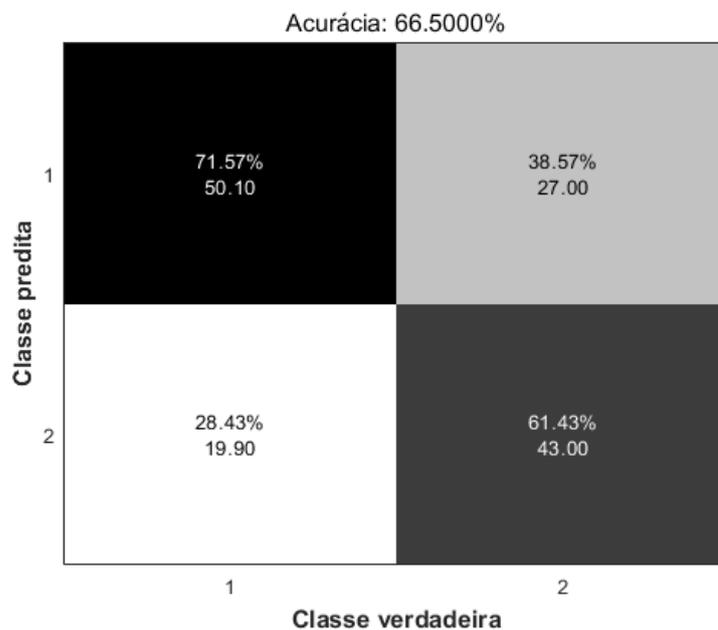
Fonte: do autor.

Figura 26 – Curvas ROC utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



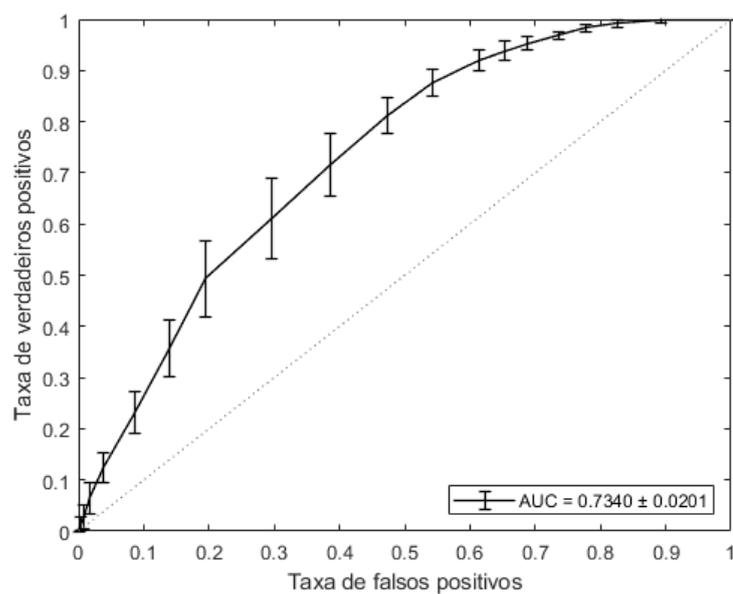
Fonte: do autor.

Figura 27 – Matriz de confusão utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



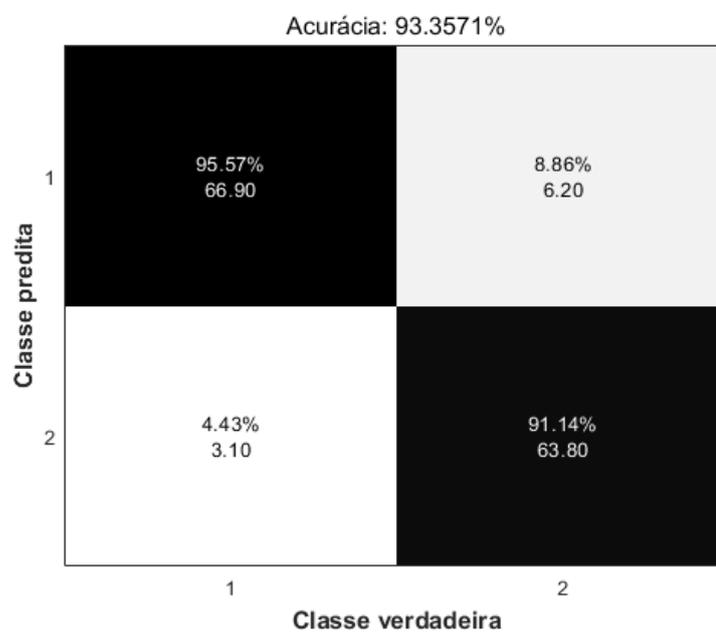
Fonte: do autor.

Figura 28 – Curva ROC utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes cheias e parcialmente cheias.



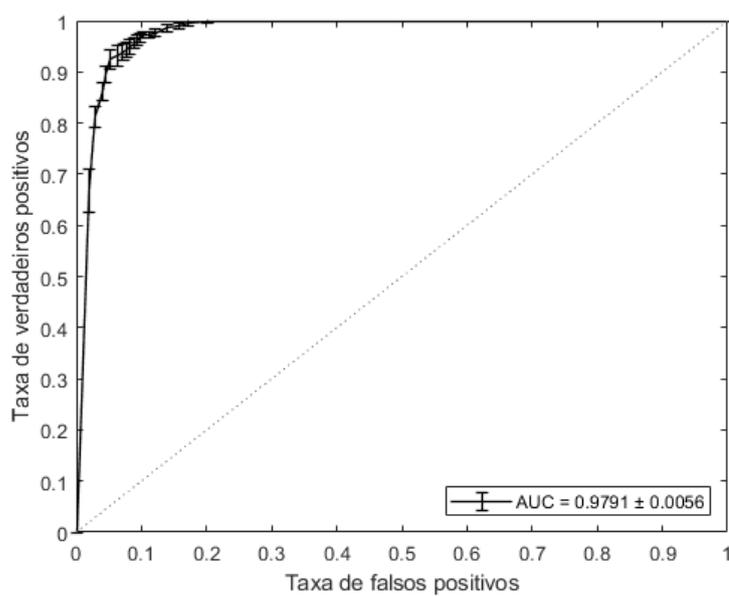
Fonte: do autor.

Figura 29 – Matriz de confusão utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) cheias e (2) deformadas.



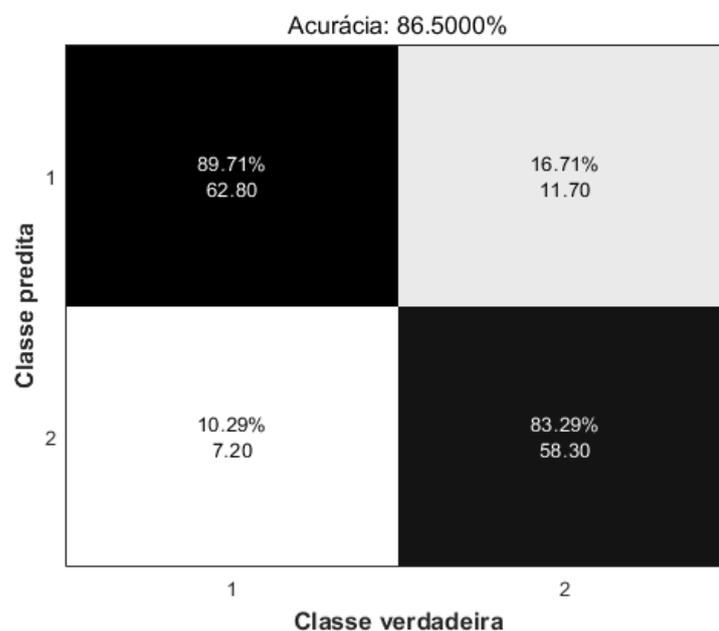
Fonte: do autor.

Figura 30 – Curva ROC utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes cheias e deformadas.



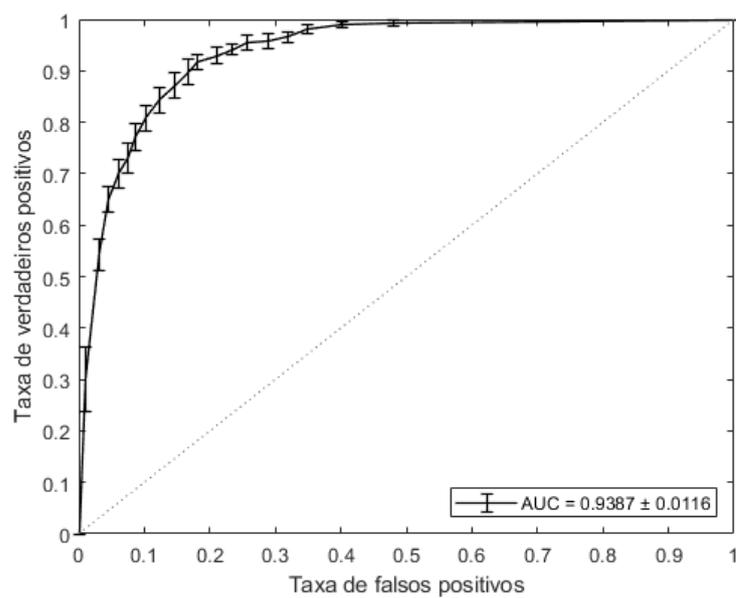
Fonte: do autor.

Figura 31 – Matriz de confusão utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.



Fonte: do autor.

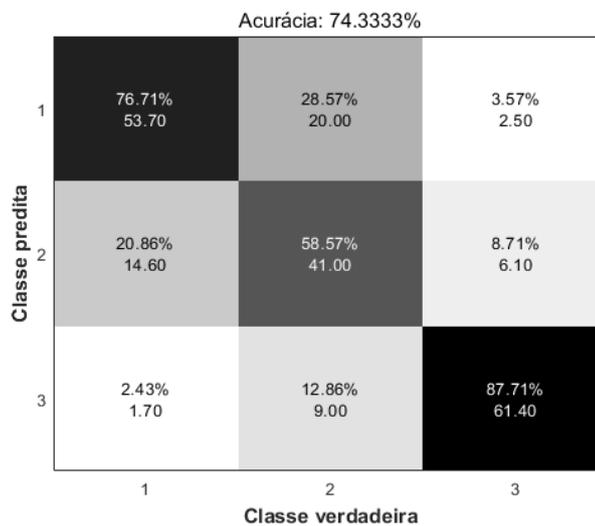
Figura 32 – Curva ROC utilizando o conjunto de dados “Segmentação” e considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

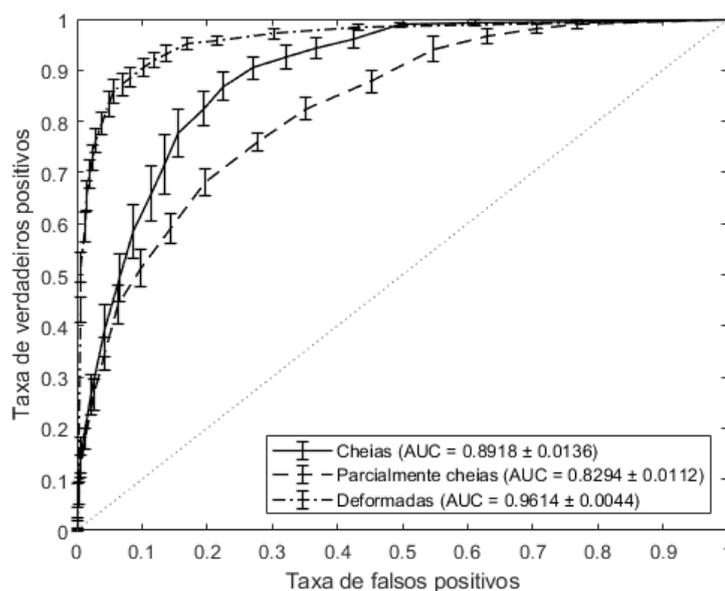
APÊNDICE E – Matrizes de confusão e curvas ROC do conjunto de dados
“Segmentação (maior área)”

Figura 33 – Matriz de confusão utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



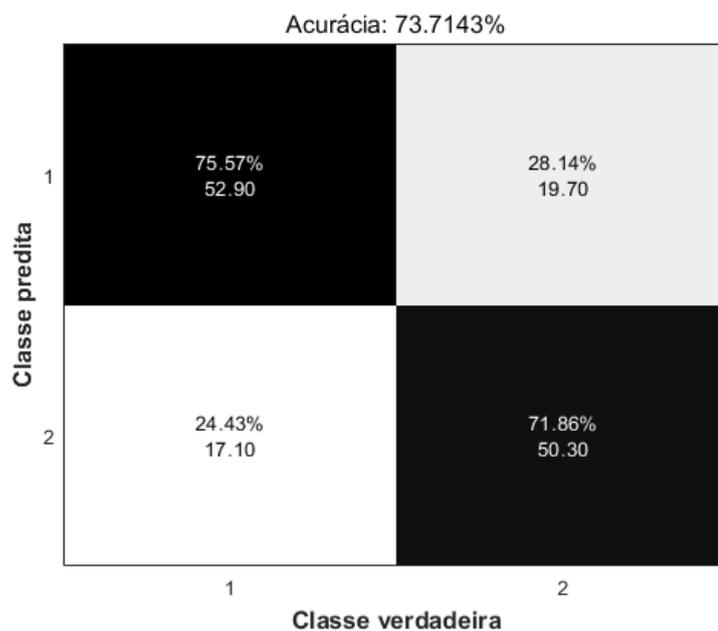
Fonte: do autor.

Figura 34 – Curvas ROC utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



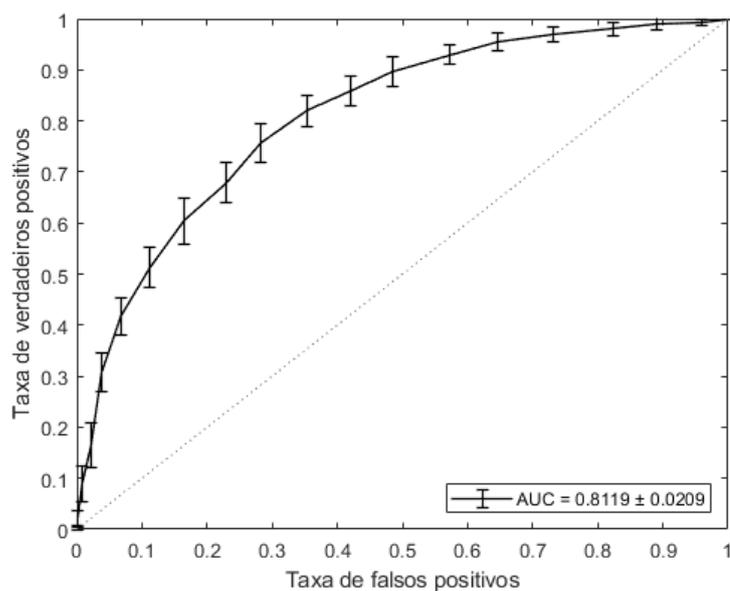
Fonte: do autor.

Figura 35 – Matriz de confusão utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



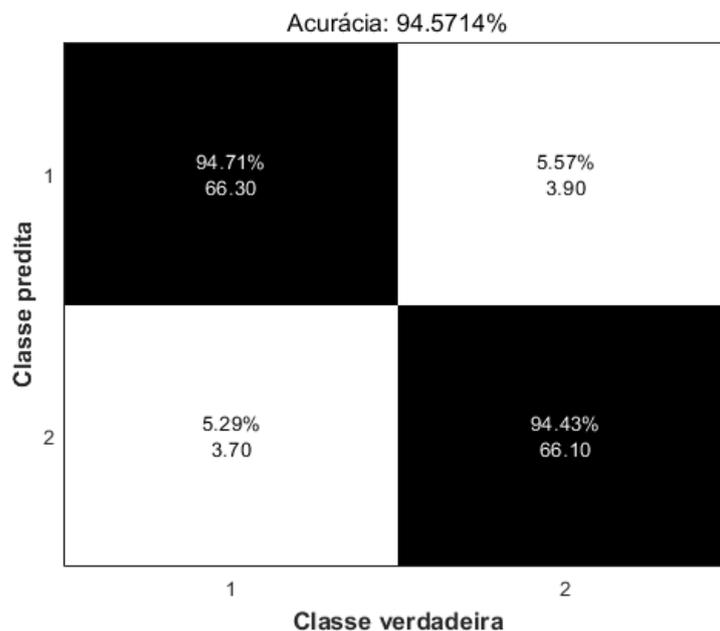
Fonte: do autor.

Figura 36 – Curva ROC utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes cheias e parcialmente cheias.



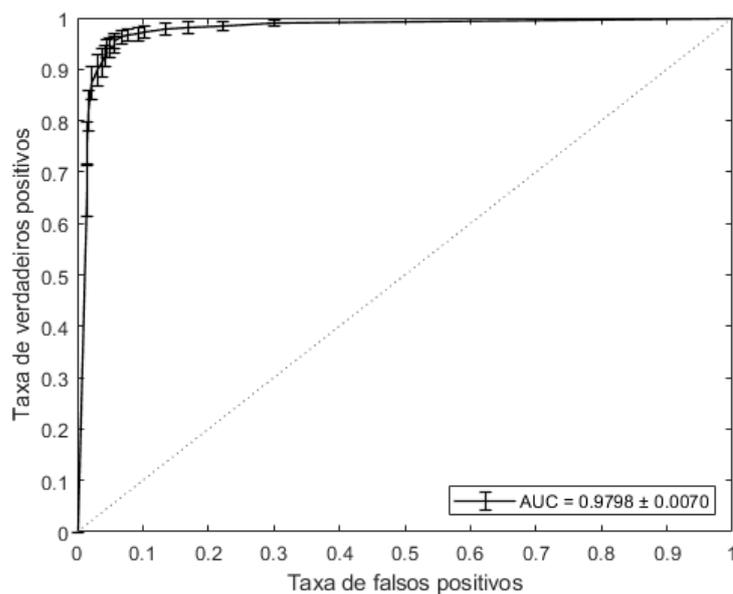
Fonte: do autor.

Figura 37 – Matriz de confusão utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes: (1) cheias e (2) deformadas.



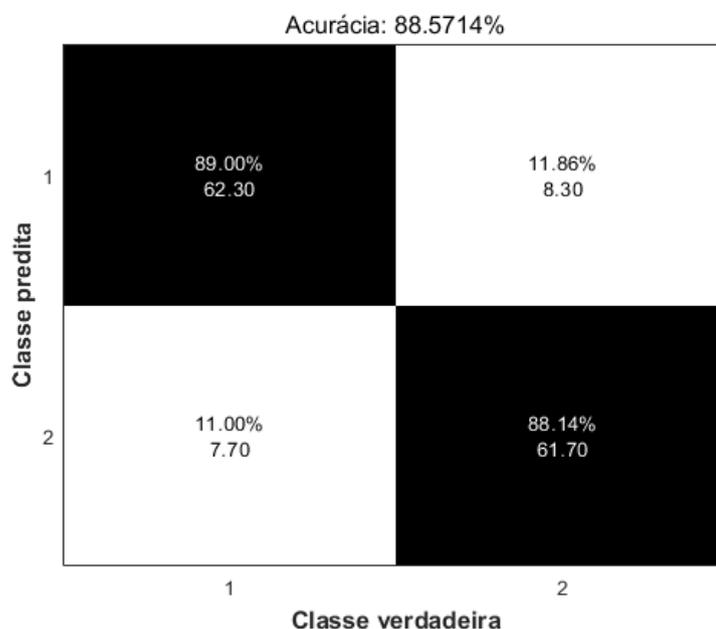
Fonte: do autor.

Figura 38 – Curva ROC utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes cheias e deformadas.



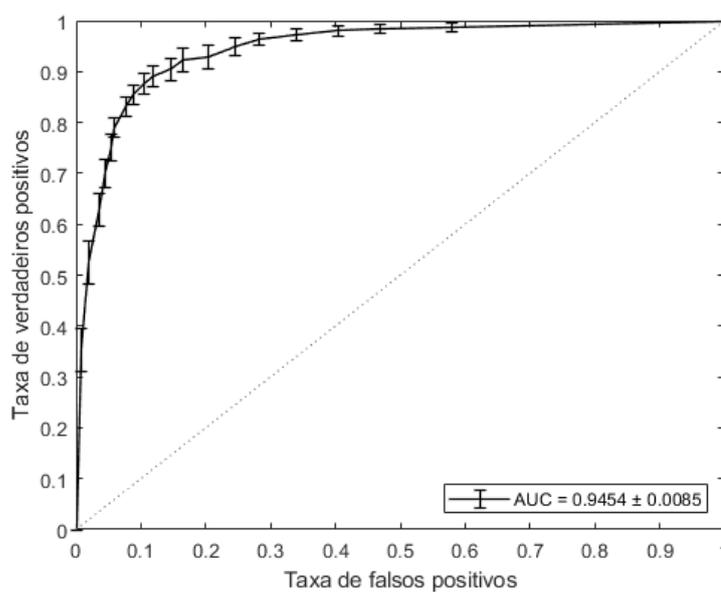
Fonte: do autor.

Figura 39 – Matriz de confusão utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.



Fonte: do autor.

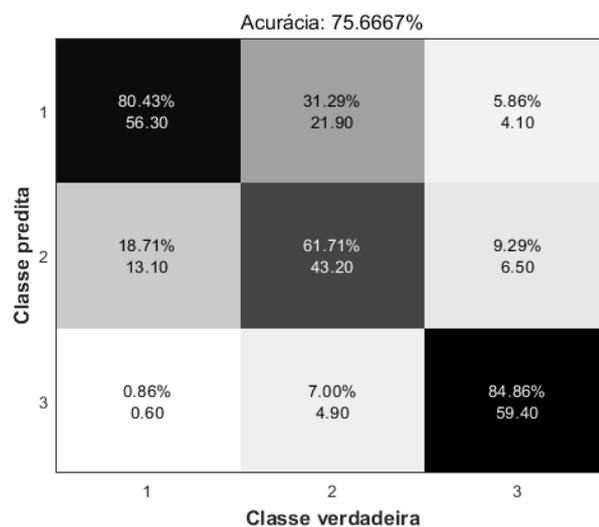
Figura 40 – Curva ROC utilizando o conjunto de dados “Segmentação (maior área)” e considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

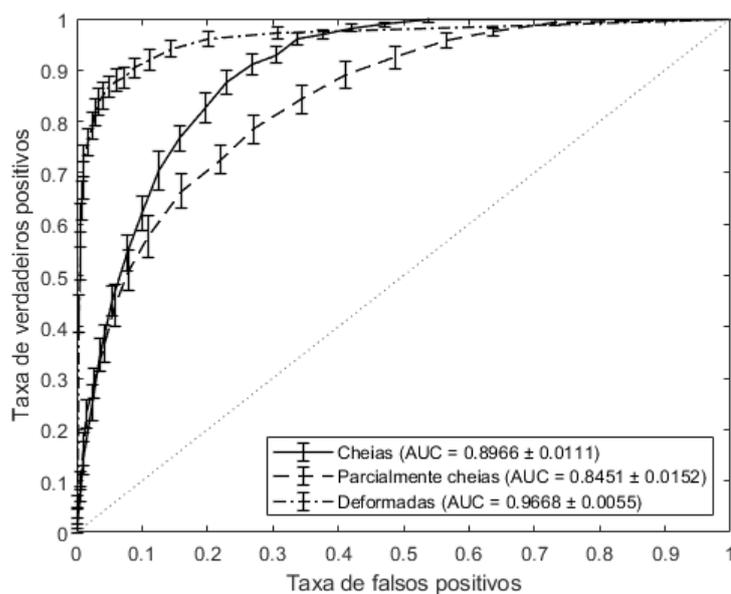
**APÊNDICE F – Matrizes de confusão e curvas ROC do conjunto de dados
“Máscara elíptica”**

Figura 41 – Matriz de confusão utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



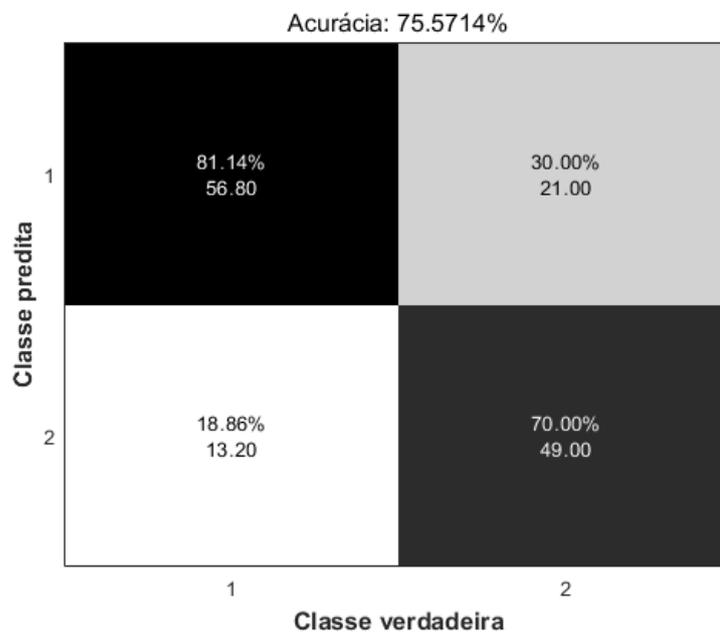
Fonte: do autor.

Figura 42 – Curvas ROC utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



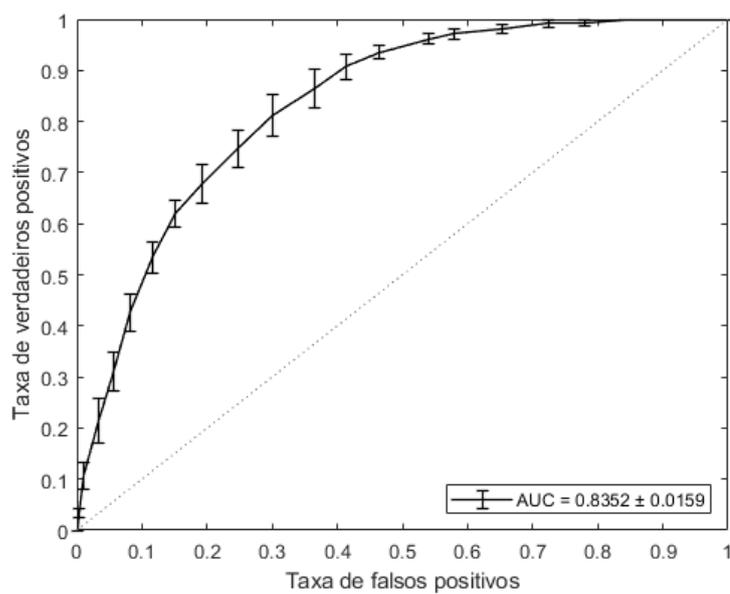
Fonte: do autor.

Figura 43 – Matriz de confusão utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



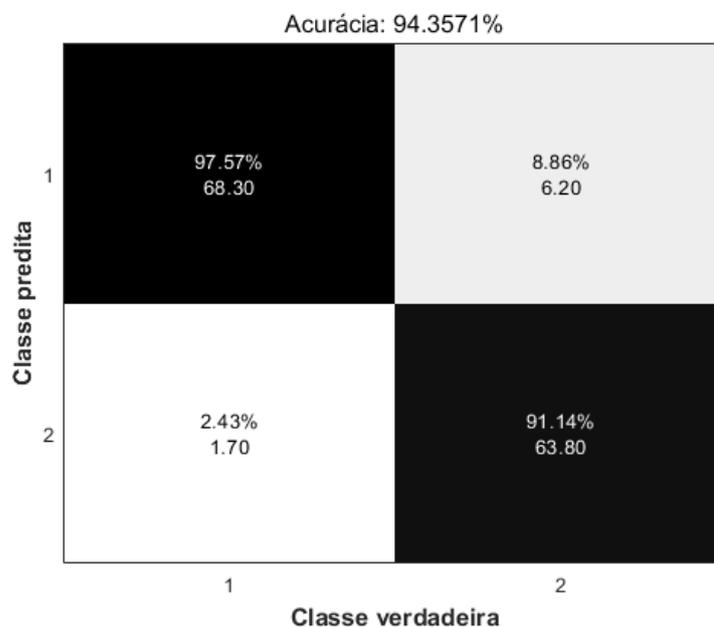
Fonte: do autor.

Figura 44 – Curva ROC utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes cheias e parcialmente cheias.



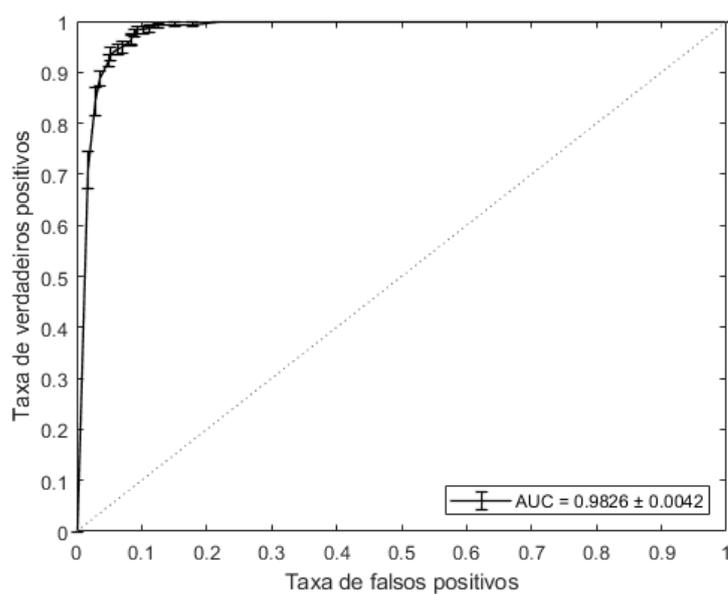
Fonte: do autor.

Figura 45 – Matriz de confusão utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes: (1) cheias e (2) deformadas.



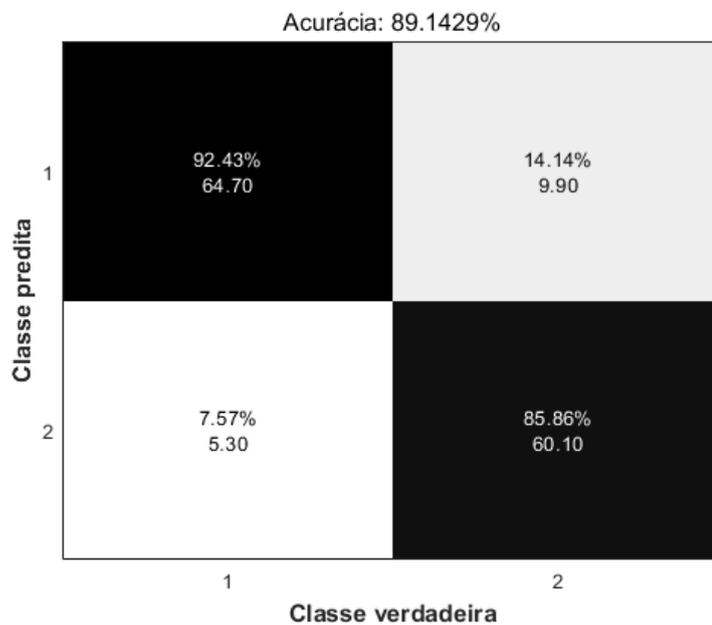
Fonte: do autor.

Figura 46 – Curva ROC utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes cheias e deformadas.



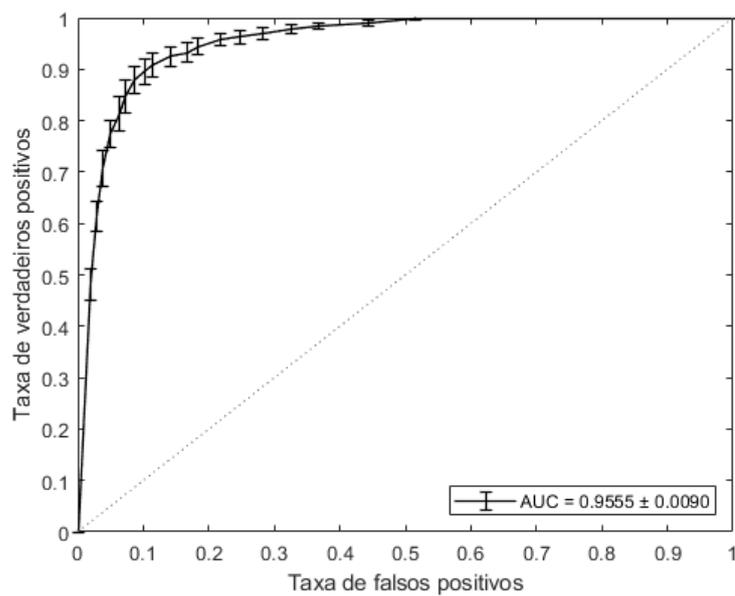
Fonte: do autor.

Figura 47 – Matriz de confusão utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.



Fonte: do autor.

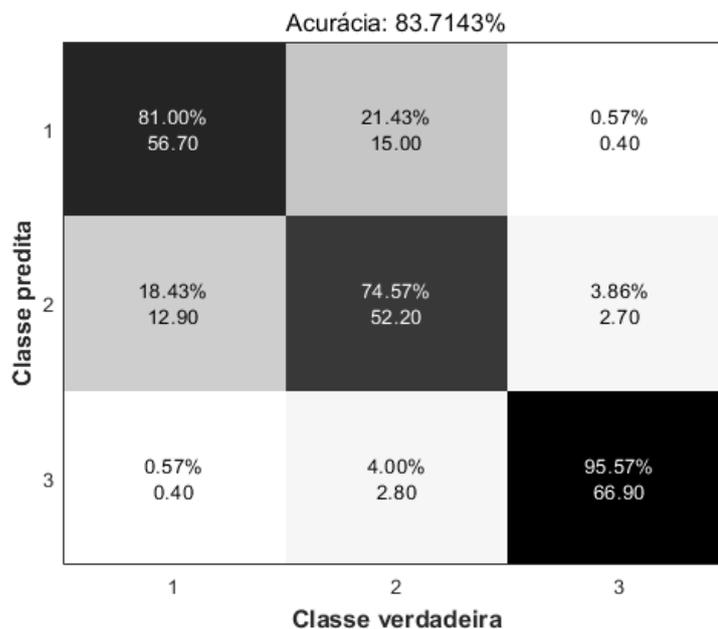
Figura 48 – Curva ROC utilizando o conjunto de dados “Máscara elíptica” e considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

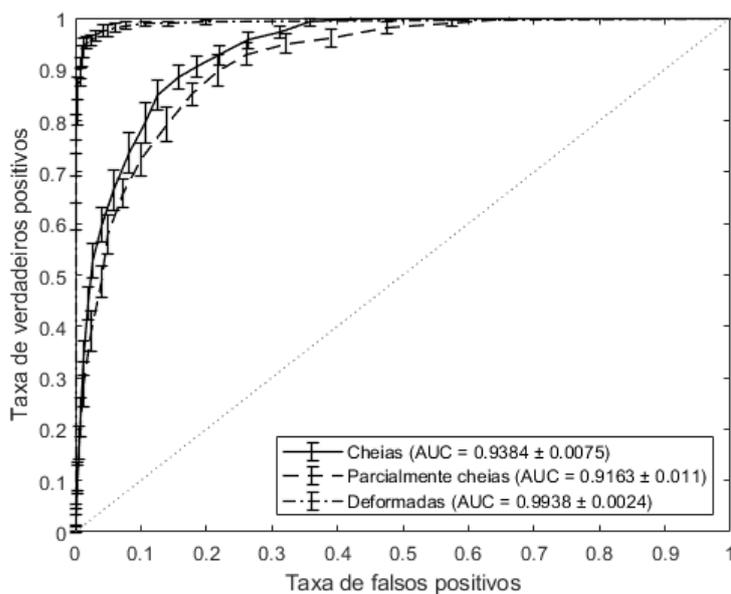
APÊNDICE G – Matrizes de confusão e curvas ROC do conjunto de dados
“Haralick”

Figura 49 – Curvas ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



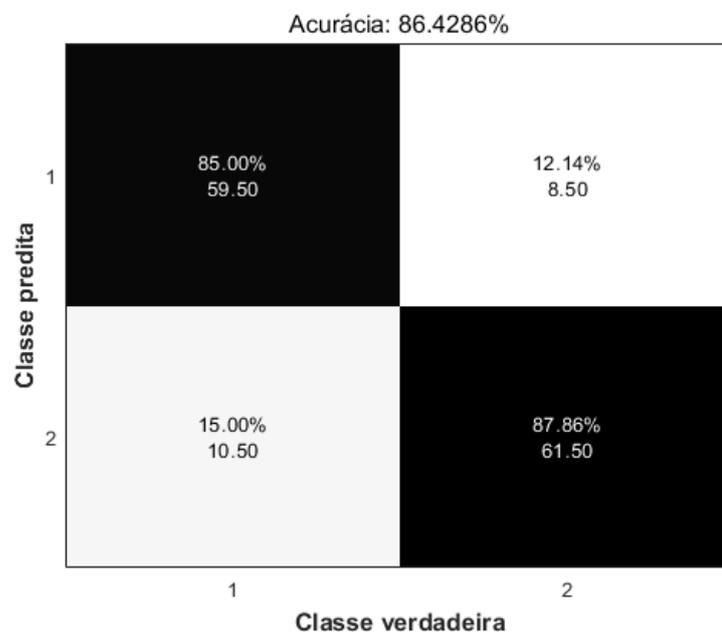
Fonte: do autor.

Figura 50 – Curvas ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



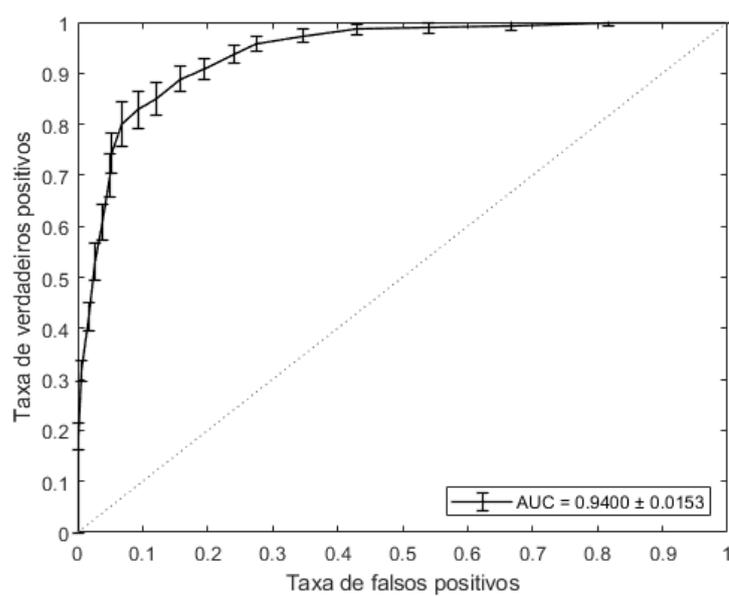
Fonte: do autor.

Figura 51 – Matriz de confusão utilizando o conjunto de dados “Haralick” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



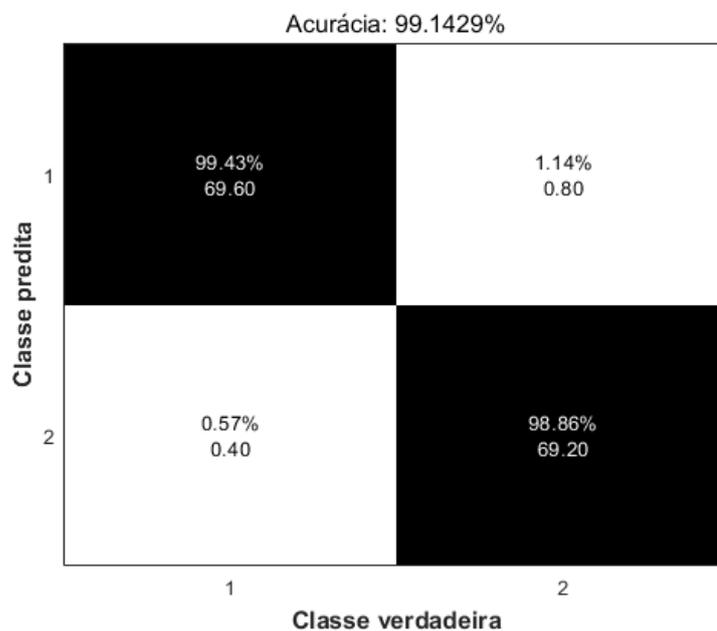
Fonte: do autor.

Figura 52 – Curva ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes cheias e parcialmente cheias.



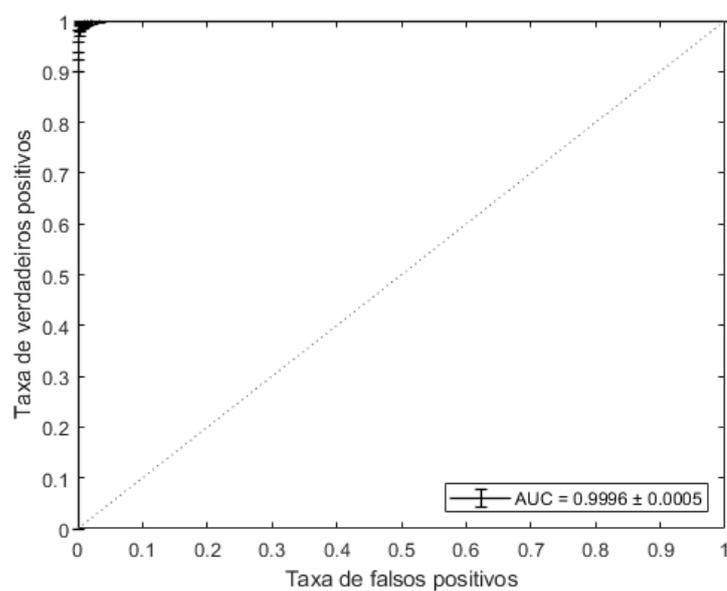
Fonte: do autor.

Figura 53 – Matriz de confusão utilizando o conjunto de dados “Haralick” e considerando as classes de sementes: (1) cheias e (2) deformadas.



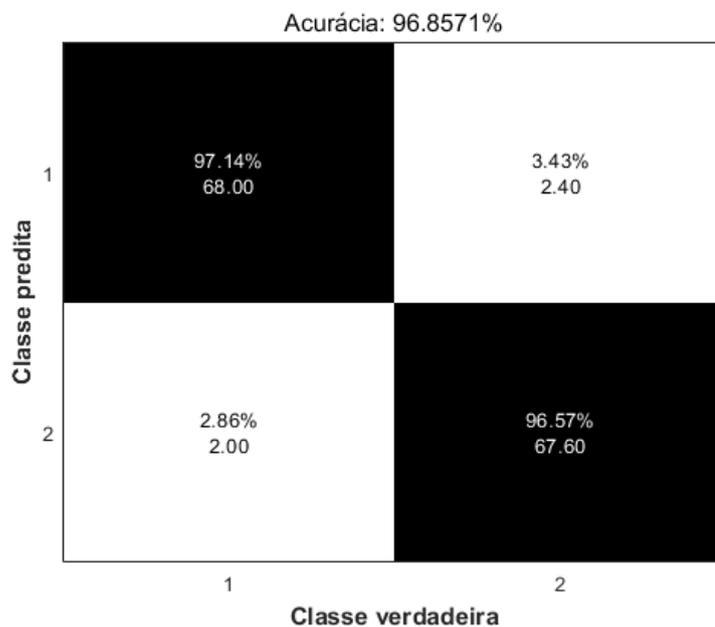
Fonte: do autor.

Figura 54 – Curva ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes cheias e deformadas.



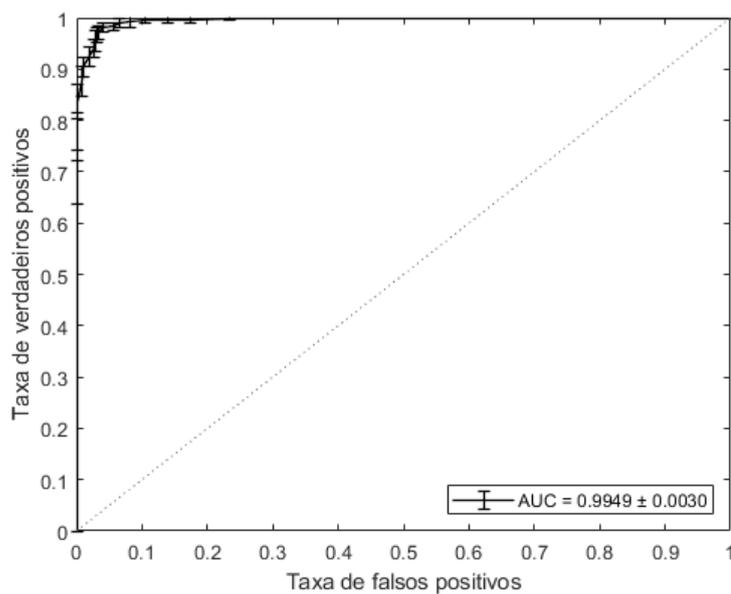
Fonte: do autor.

Figura 55 – Matriz de confusão utilizando o conjunto de dados “Haralick” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.



Fonte: do autor.

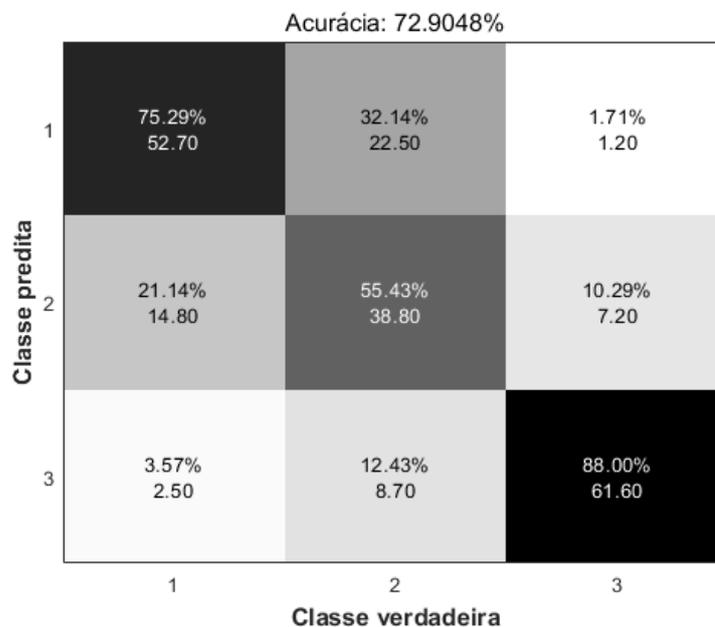
Figura 56 – Curva ROC utilizando o conjunto de dados “Haralick” e considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

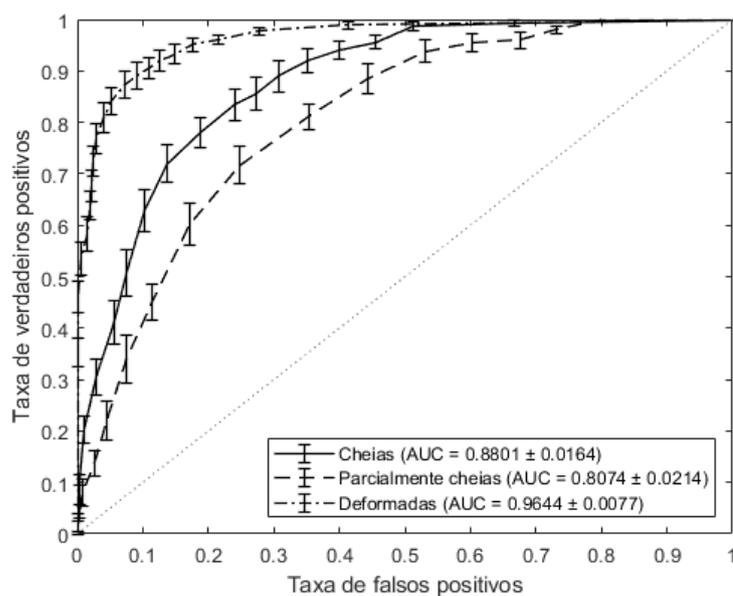
APÊNDICE H – Matrizes de confusão e curvas ROC do conjunto de dados
“Haralick reduzido”

Figura 57 – Matriz de confusão utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



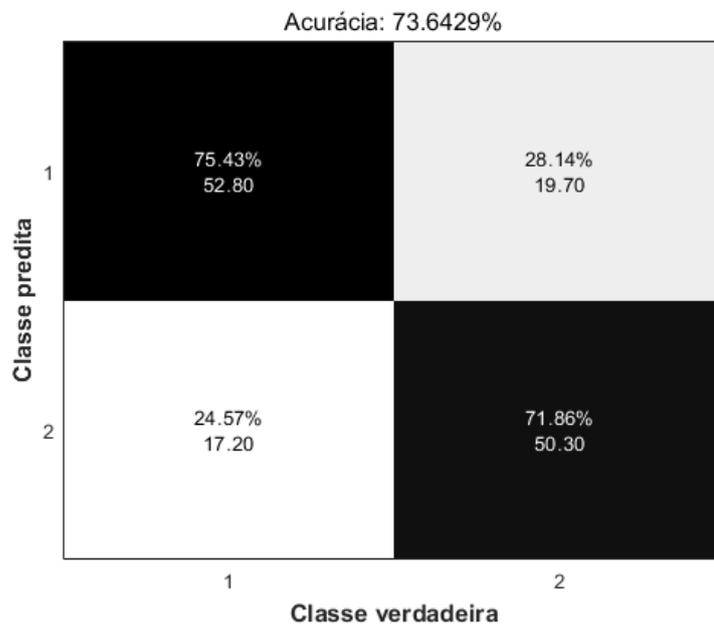
Fonte: do autor.

Figura 58 – Curvas ROC utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



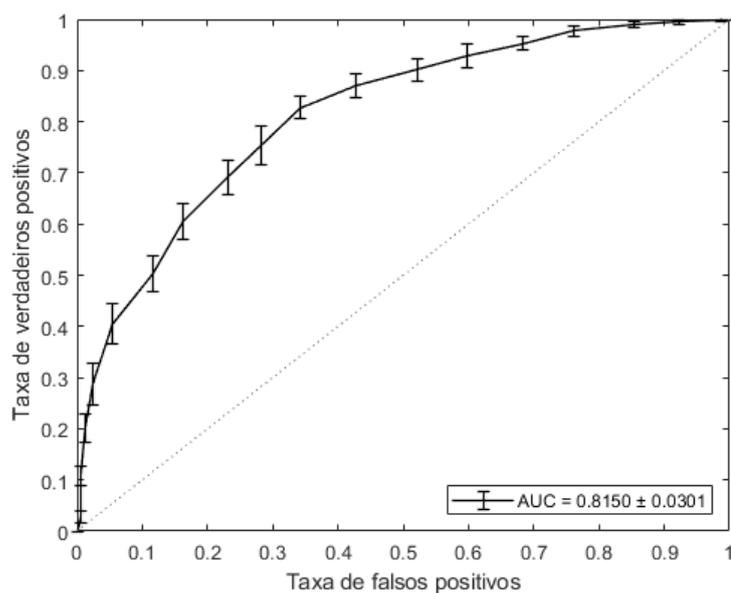
Fonte: do autor.

Figura 59 – Matriz de confusão utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



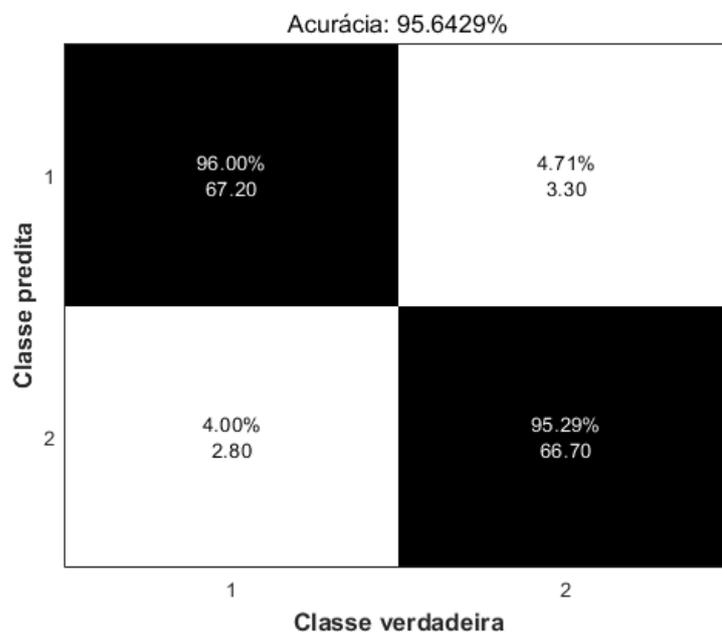
Fonte: do autor.

Figura 60 – Curva ROC utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes cheias e parcialmente cheias.



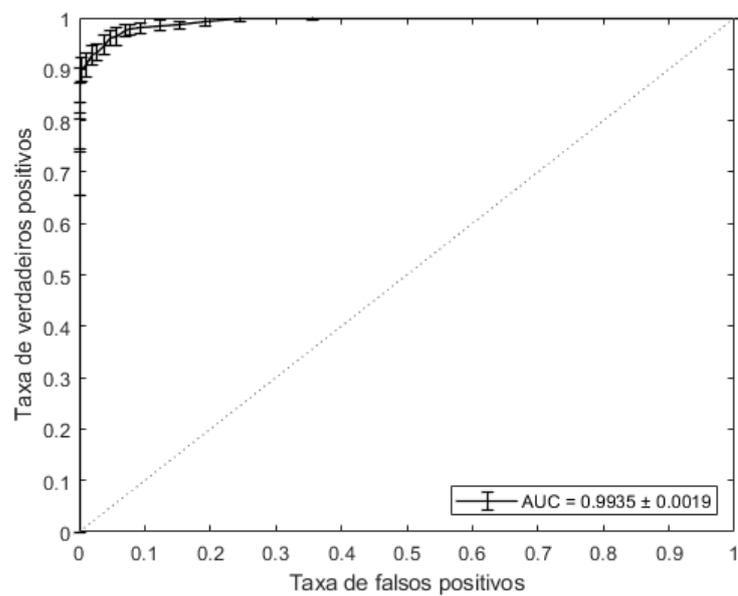
Fonte: do autor.

Figura 61 – Matriz de confusão utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes: (1) cheias e (2) deformadas.



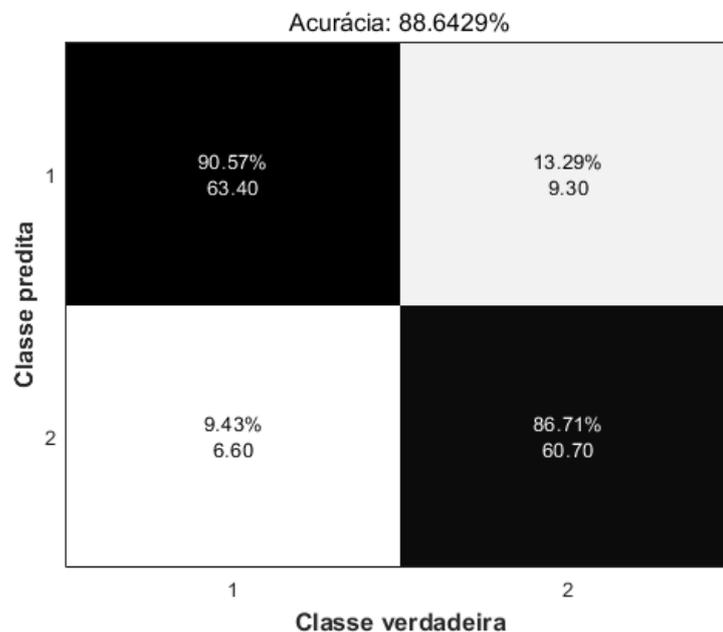
Fonte: do autor.

Figura 62 – Curva ROC utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes cheias e deformadas.



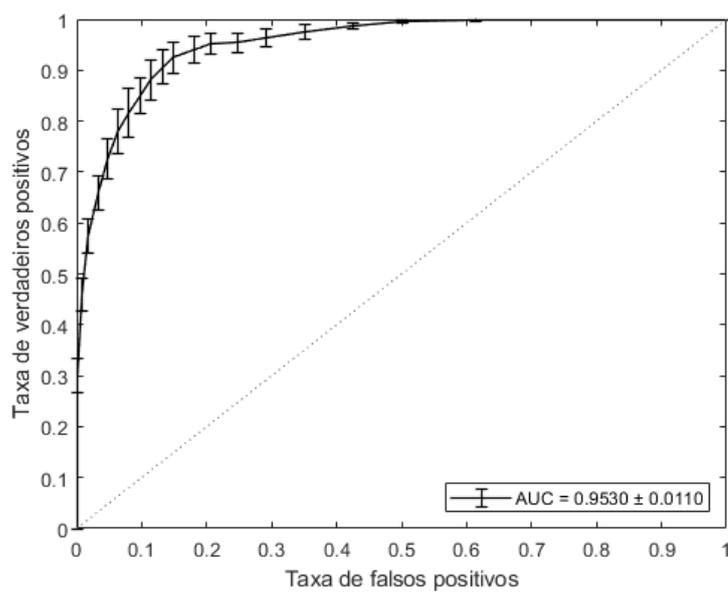
Fonte: do autor.

Figura 63 – Matriz de confusão utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.



Fonte: do autor.

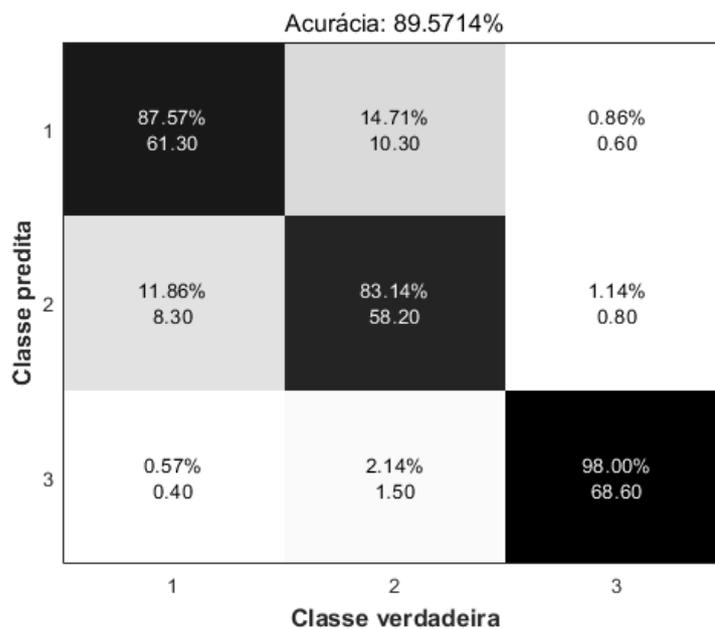
Figura 64 – Curva ROC utilizando o conjunto de dados “Haralick reduzido” e considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

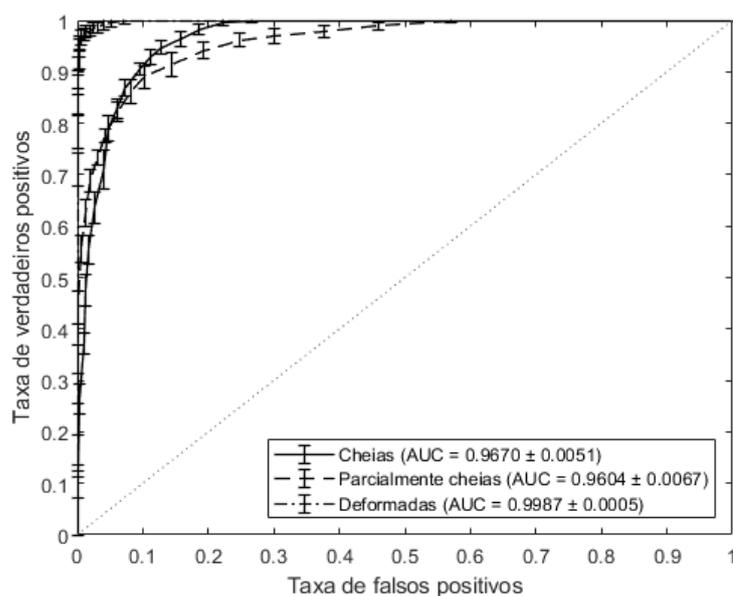
APÊNDICE I – Matrizes de confusão e curvas ROC do conjunto de dados
“Haralick (duas distâncias)”

Figura 65 – Matriz de confusão utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



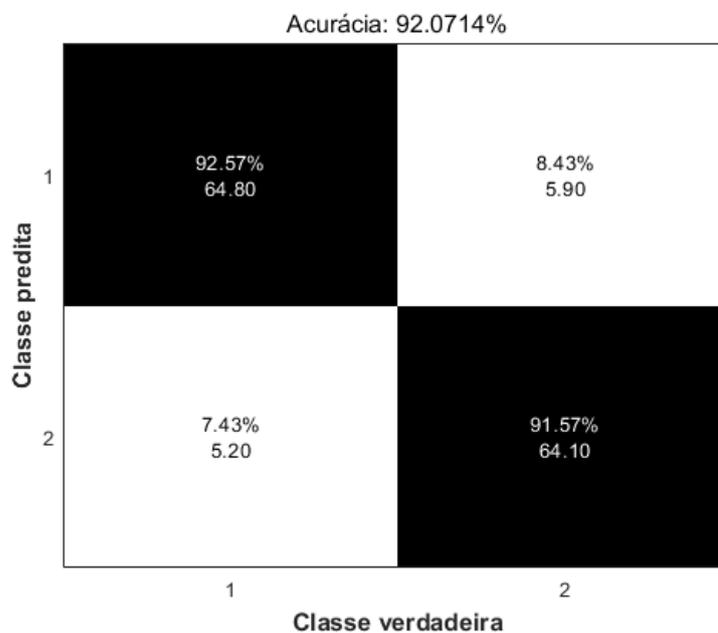
Fonte: do autor.

Figura 66 – Curvas ROC utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



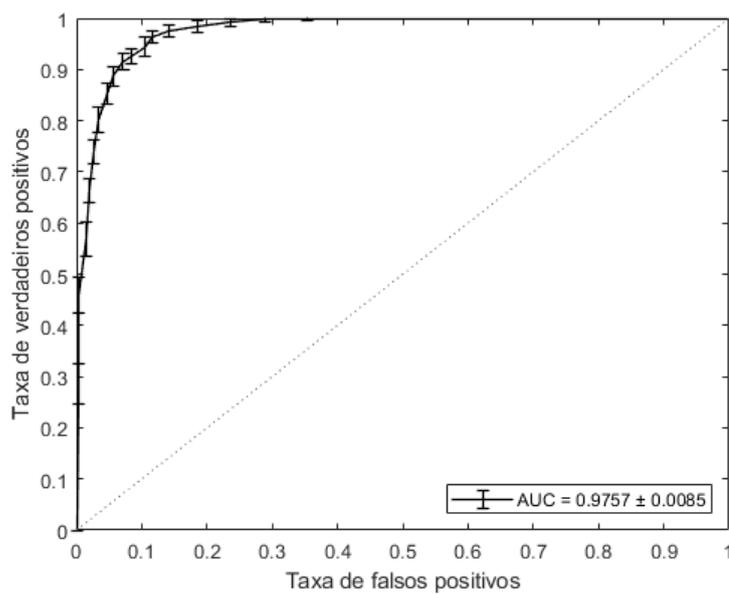
Fonte: do autor.

Figura 67 – Matriz de confusão utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



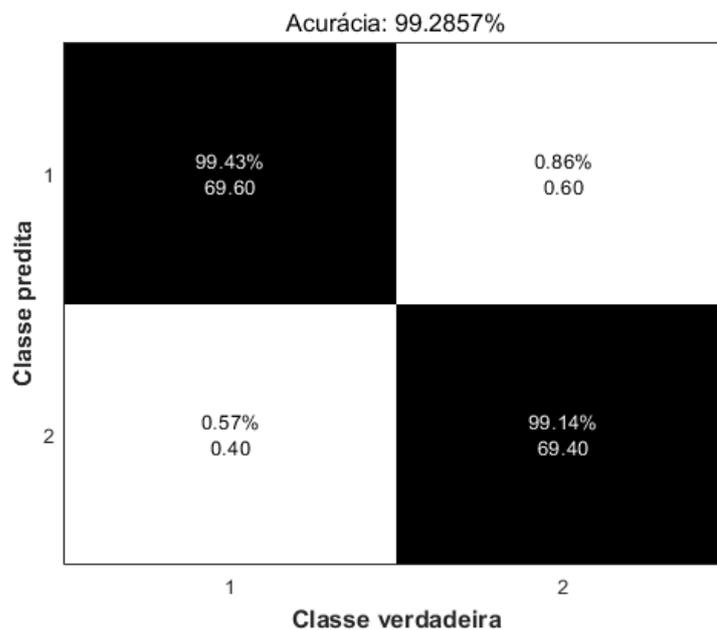
Fonte: do autor.

Figura 68 – Curva ROC utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes cheias e parcialmente cheias.



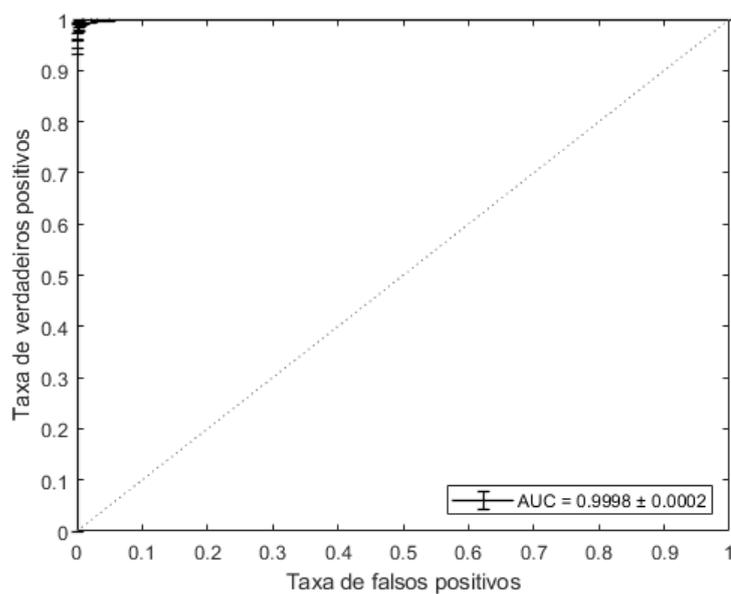
Fonte: do autor.

Figura 69 – Matriz de confusão utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) cheias e (2) deformadas.



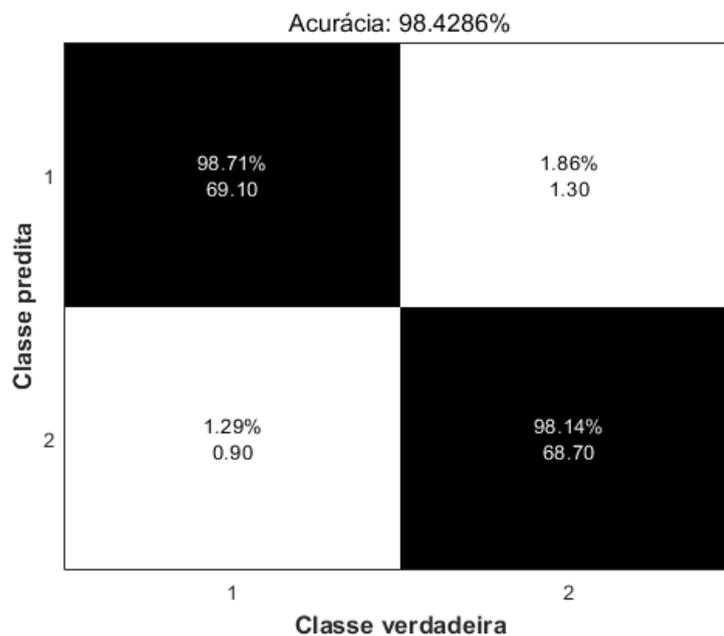
Fonte: do autor.

Figura 70 – Curva ROC utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes cheias e deformadas.



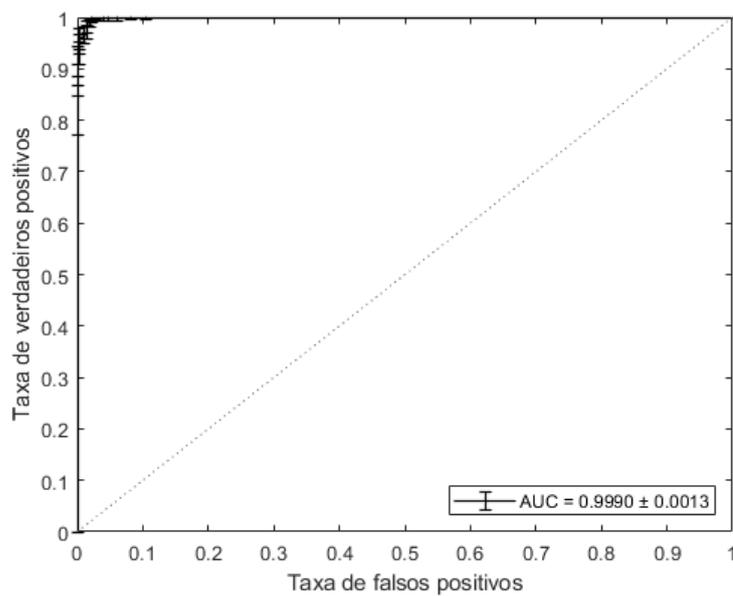
Fonte: do autor.

Figura 71 – Matriz de confusão utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.



Fonte: do autor.

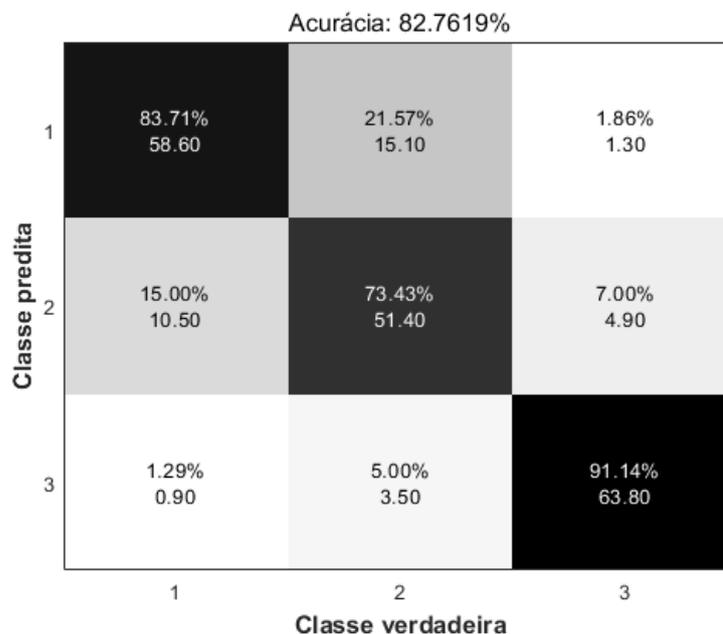
Figura 72 – Curva ROC utilizando o conjunto de dados “Haralick (duas distâncias)” e considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

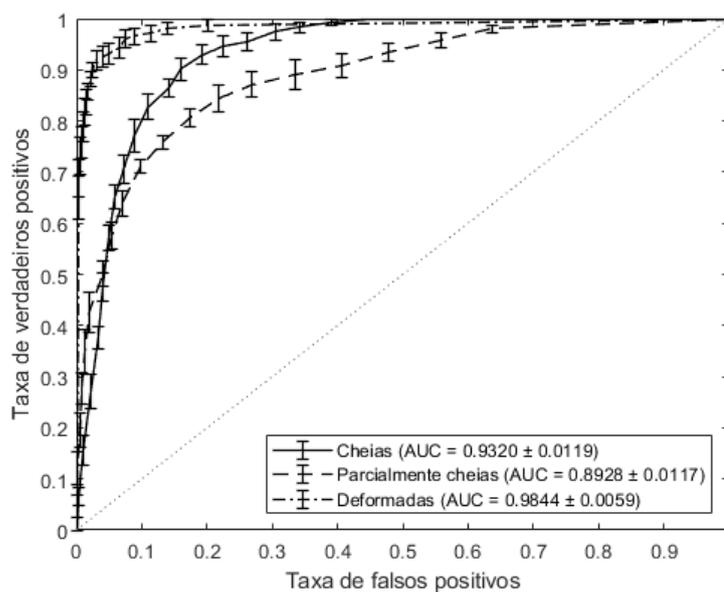
**APÊNDICE J – Matrizes de confusão e curvas ROC do conjunto de dados
“Haralick e valores dos pixels com equalização”**

Figura 73 – Matriz de confusão utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias, (2) parcialmente cheias e (3) deformadas.



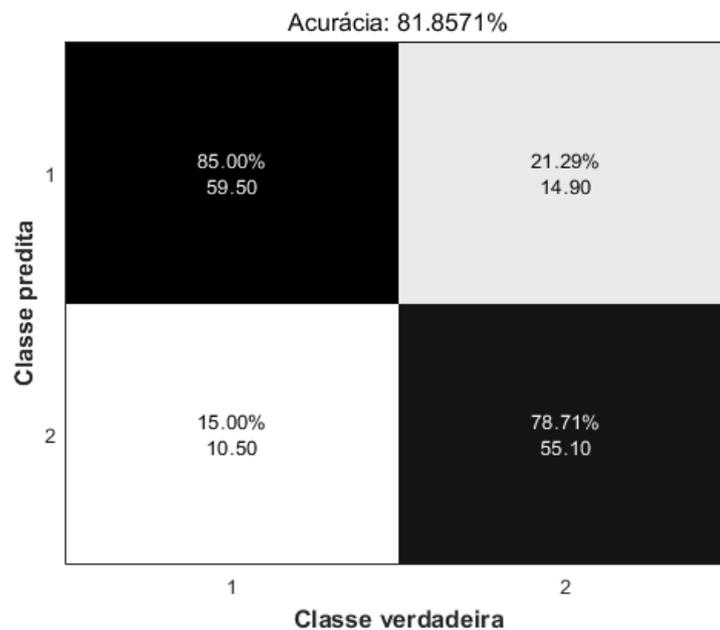
Fonte: do autor.

Figura 74 – Curvas ROC utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes cheias, parcialmente cheias e deformadas.



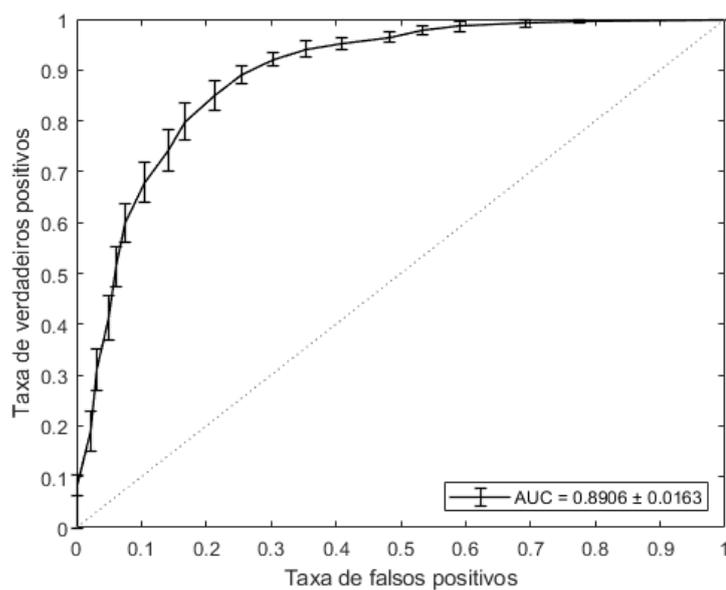
Fonte: do autor.

Figura 75 – Matriz de confusão utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias e (2) parcialmente cheias.



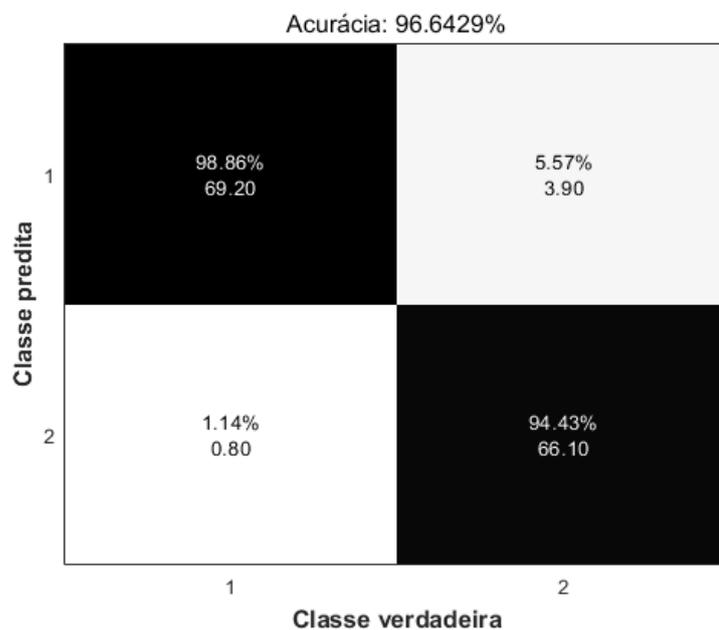
Fonte: do autor.

Figura 76 – Curva ROC utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes cheias e parcialmente cheias.



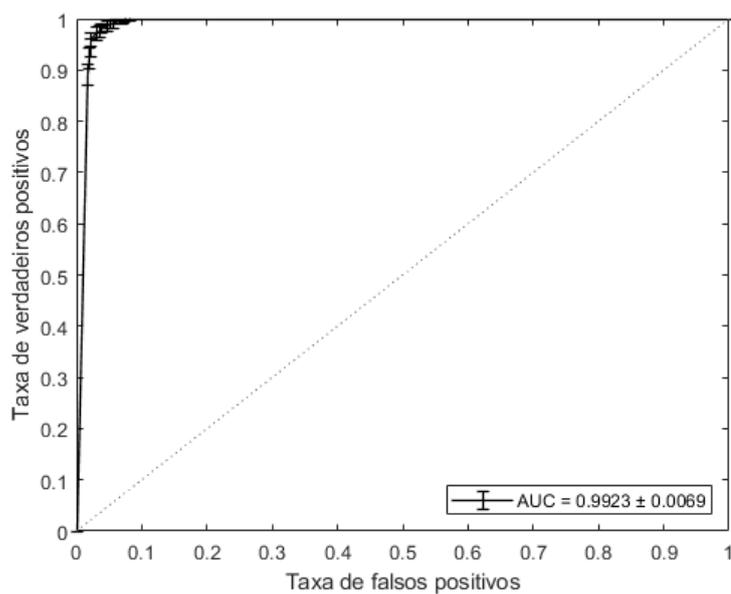
Fonte: do autor.

Figura 77 – Matriz de confusão utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes: (1) cheias e (2) deformadas.



Fonte: do autor.

Figura 78 – Curva ROC utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes cheias e deformadas.



Fonte: do autor.

Figura 79 – Matriz de confusão utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes: (1) parcialmente cheias e (2) deformadas.

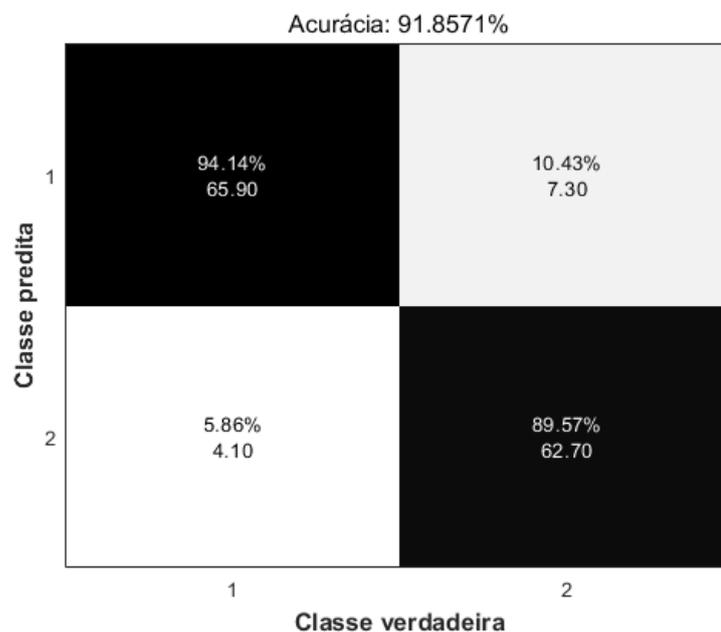
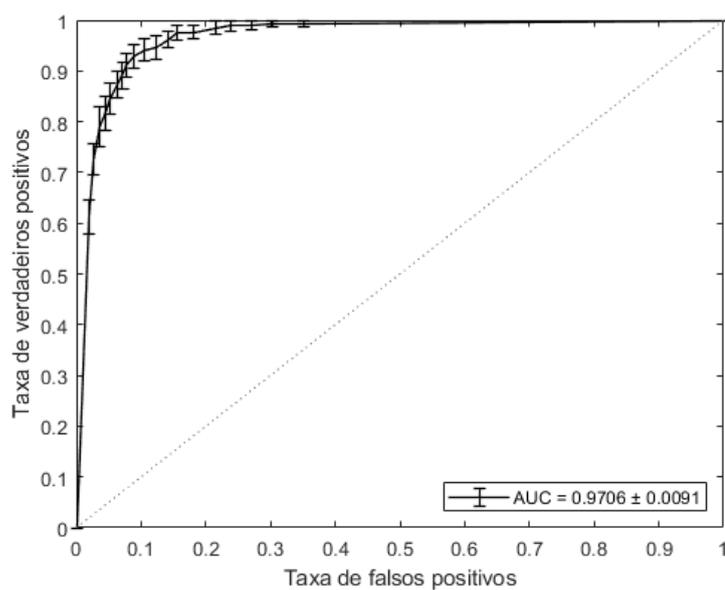
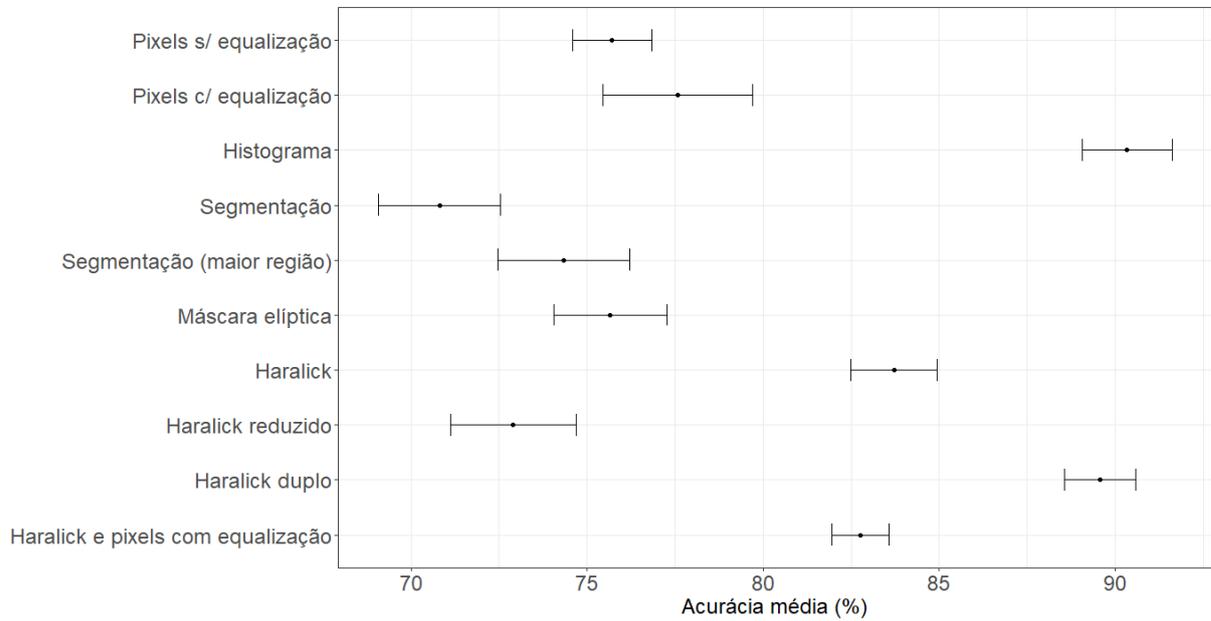


Figura 80 – Curva ROC utilizando o conjunto de dados “Haralick e valores dos pixels com equalização” e considerando as classes de sementes parcialmente cheias e deformadas.



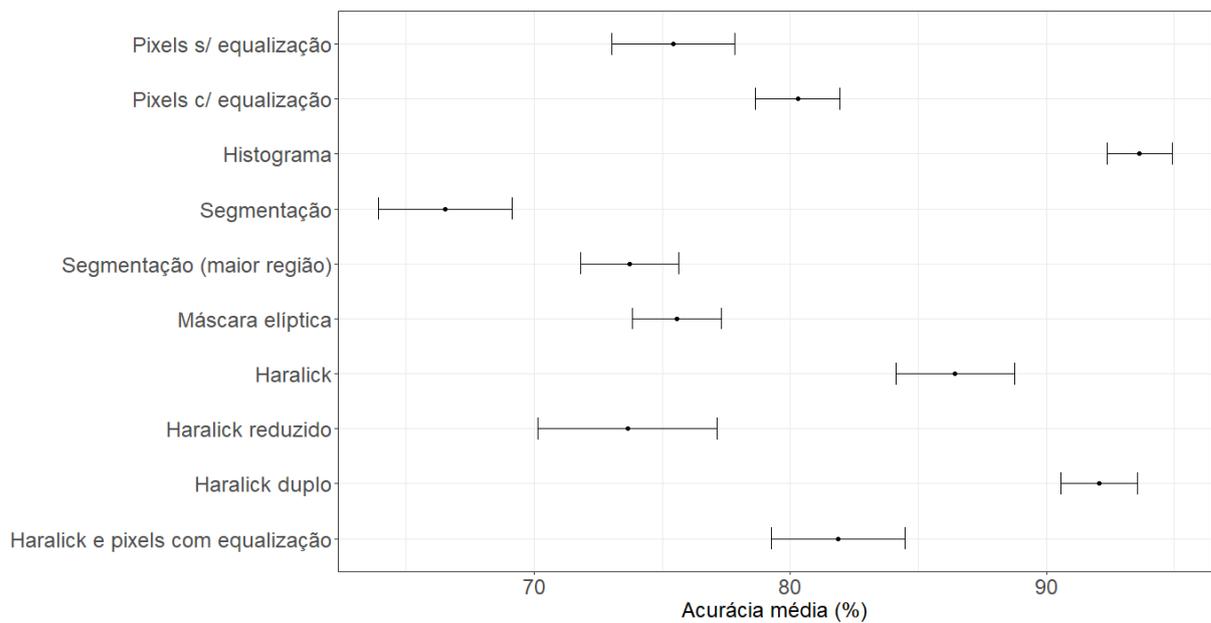
APÊNDICE K – Intervalos de confiança para a acurácia média

Figura 81 – Intervalos de confiança para a acurácia média considerando as classes de sementes cheias, parcialmente cheias e deformadas em conjunto.



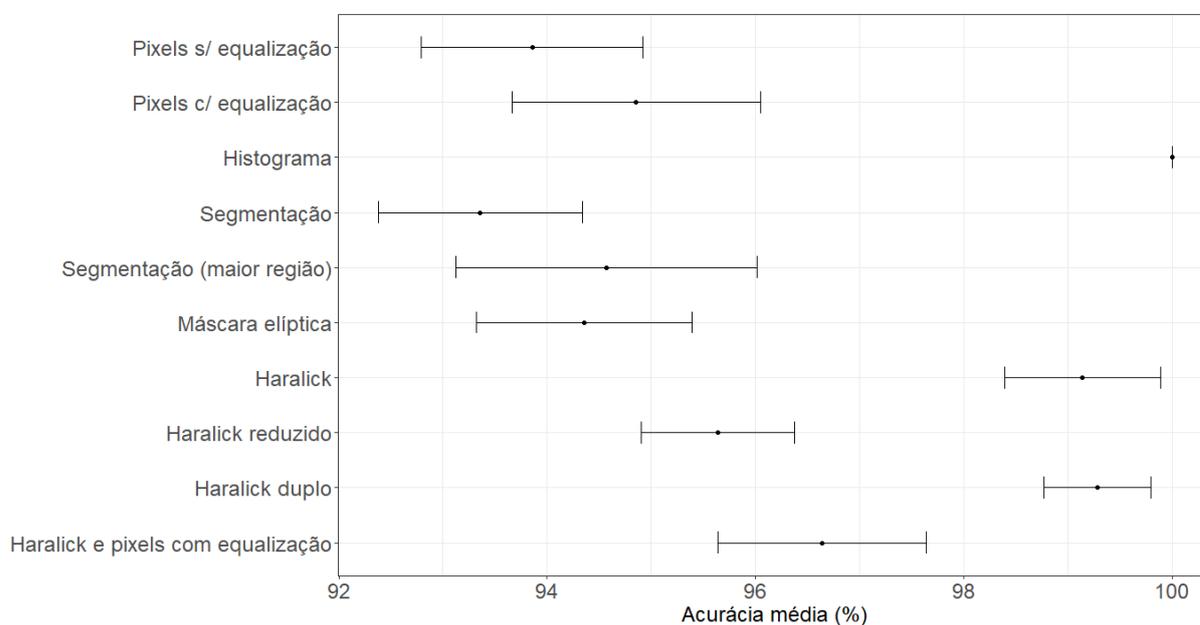
Fonte: do autor.

Figura 82 – Intervalos de confiança para a acurácia média considerando as classes de sementes cheias e parcialmente cheias.



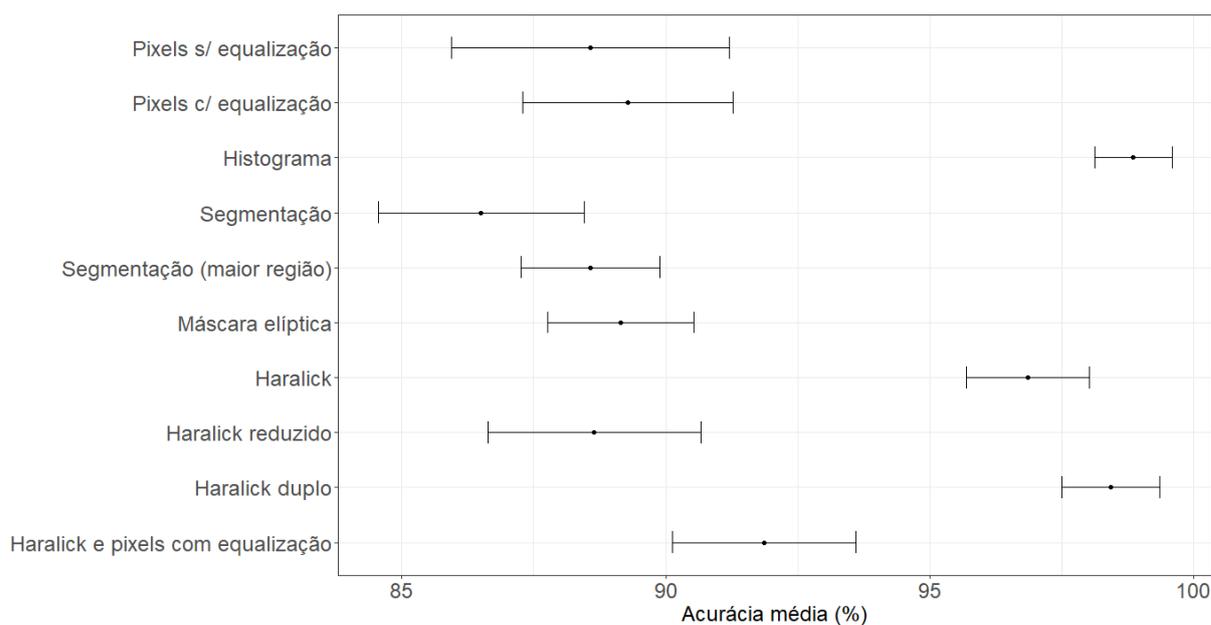
Fonte: do autor.

Figura 83 – Intervalos de confiança para a acurácia média considerando as classes de sementes cheias e deformadas.



Fonte: do autor.

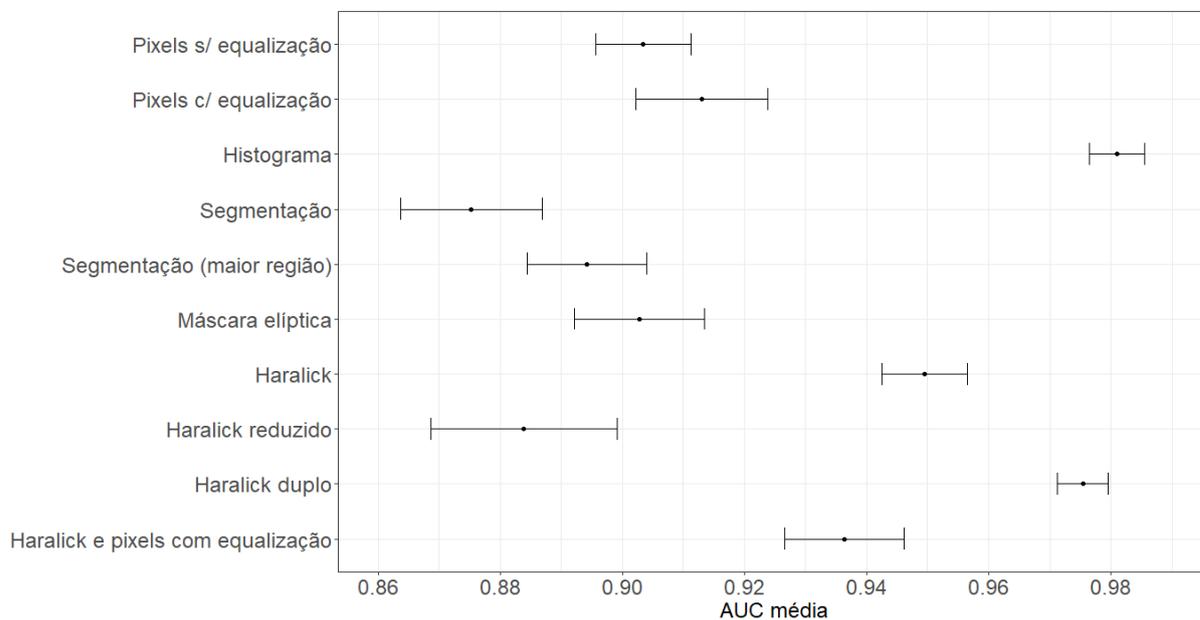
Figura 84 – Intervalos de confiança para a acurácia média considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

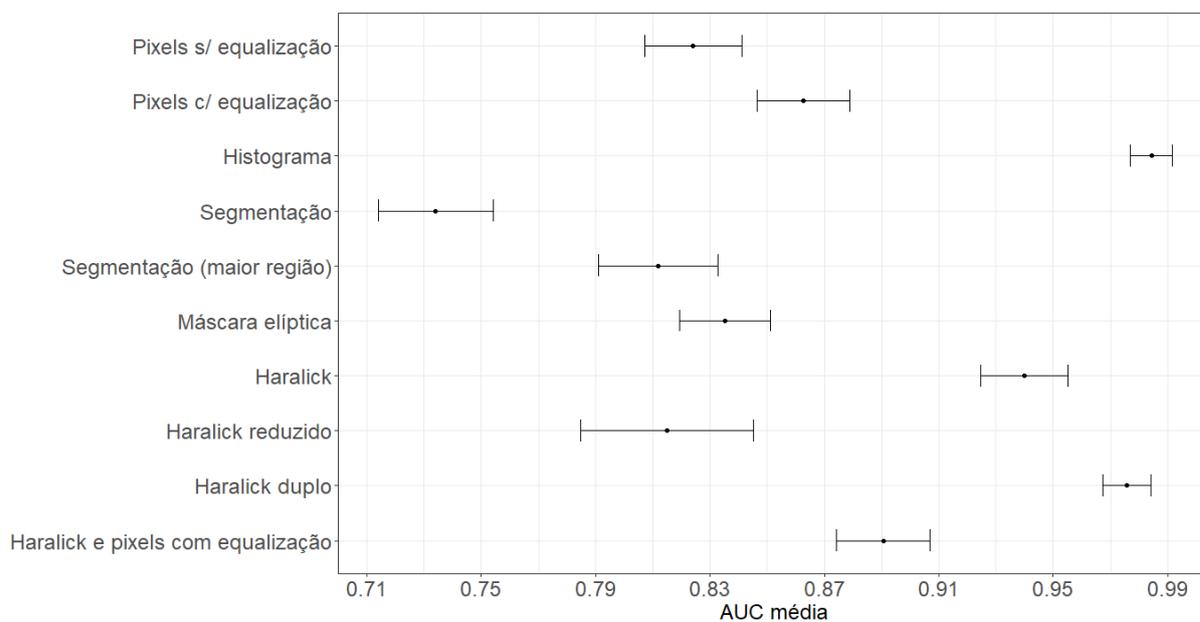
APÊNDICE L – Intervalos de confiança para a AUC média

Figura 85 – Intervalos de confiança para a AUC média considerando as classes de sementes cheias, parcialmente cheias e deformadas em conjunto.



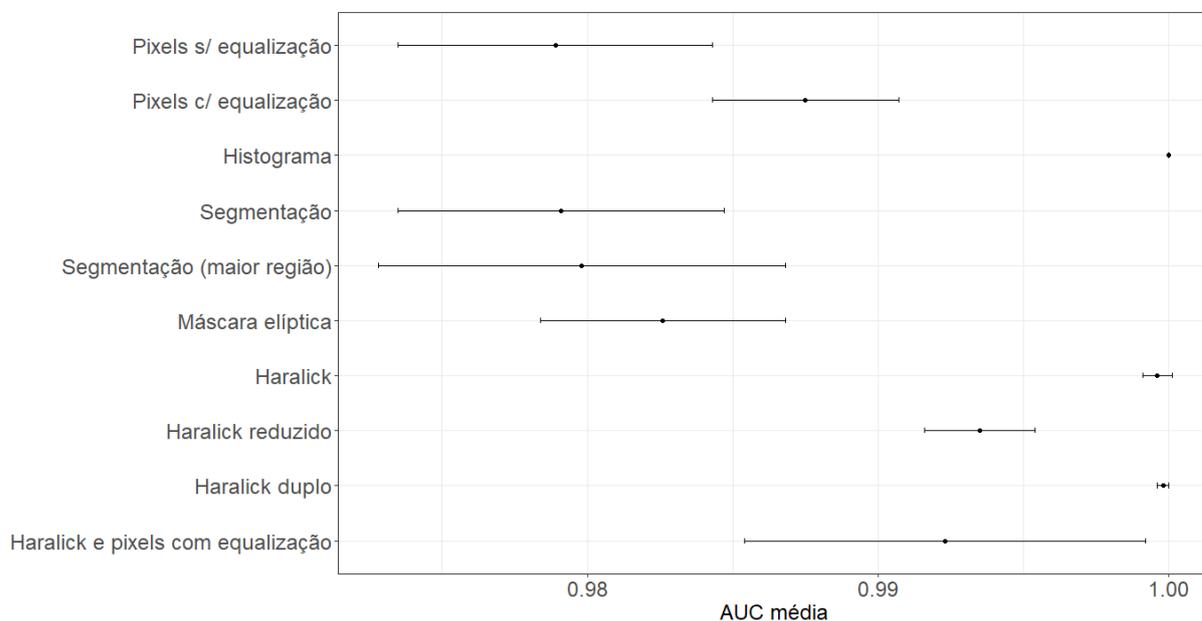
Fonte: do autor.

Figura 86 – Intervalos de confiança para a AUC média considerando as classes de sementes cheias e parcialmente cheias.



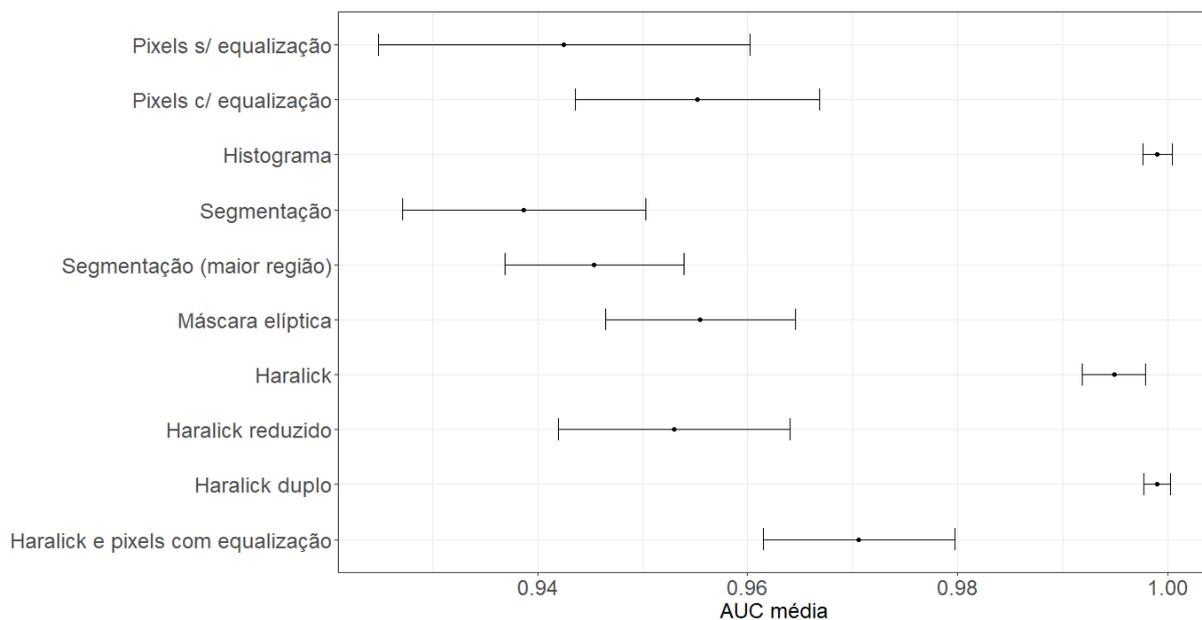
Fonte: do autor.

Figura 87 – Intervalos de confiança para a AUC média considerando as classes de sementes cheias e deformadas.



Fonte: do autor.

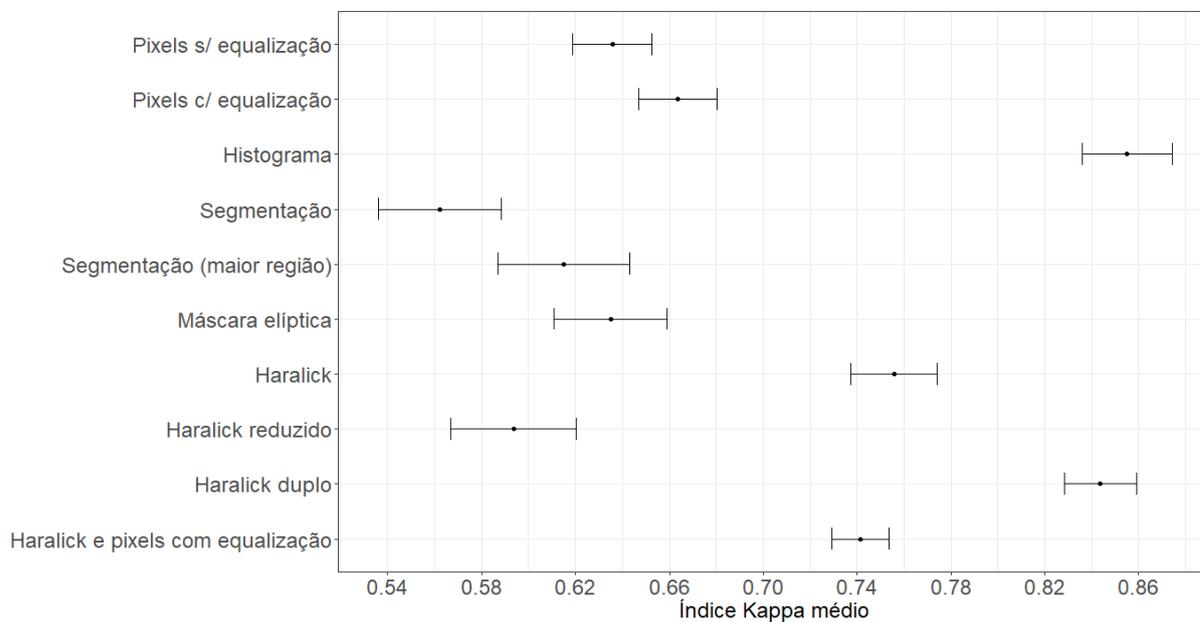
Figura 88 – Intervalos de confiança para a AUC média considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.

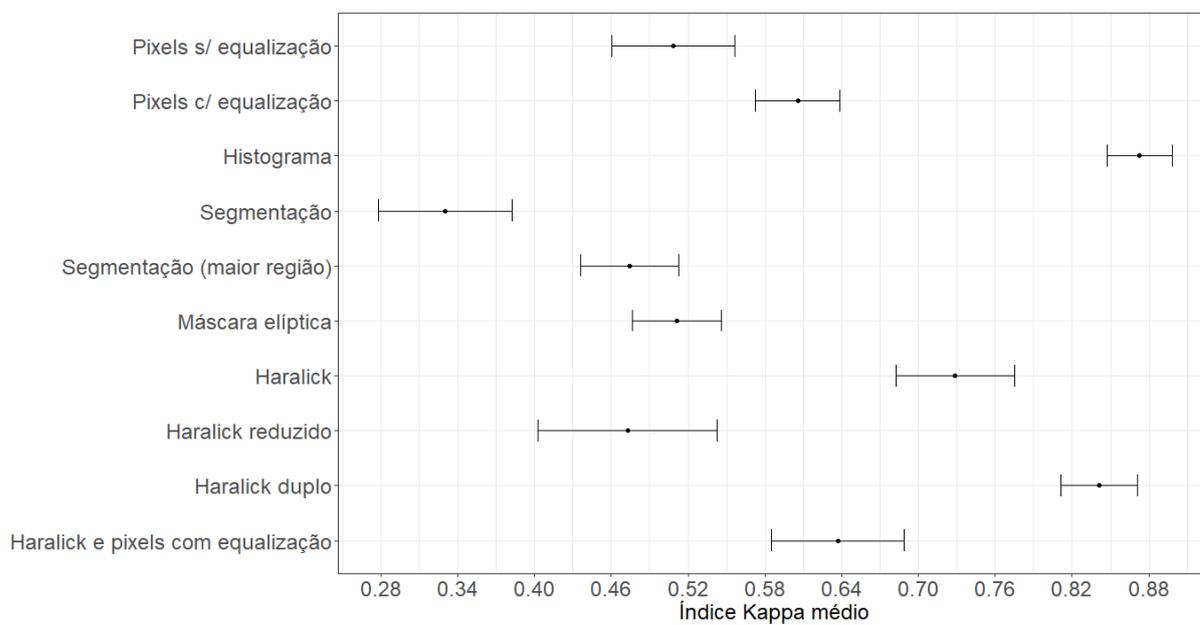
APÊNDICE M – Intervalos de confiança para o índice Kappa médio

Figura 89 – Intervalos de confiança para o índice Kappa médio considerando as classes de sementes cheias, parcialmente cheias e deformadas em conjunto.



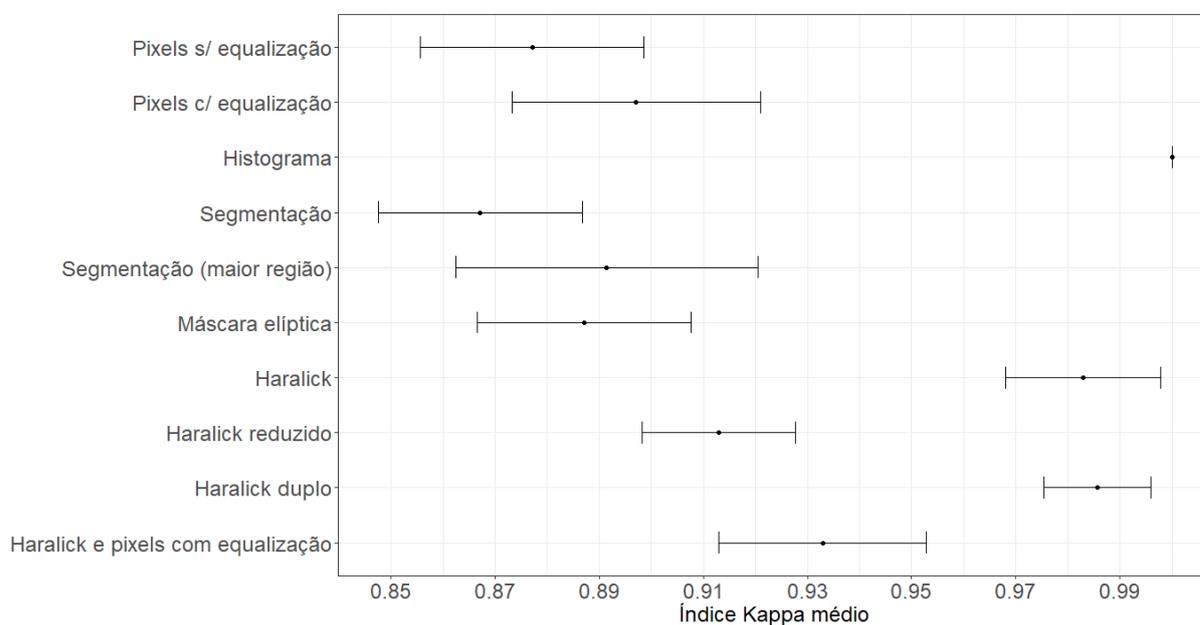
Fonte: do autor.

Figura 90 – Intervalos de confiança para o índice Kappa médio considerando as classes de sementes cheias e parcialmente cheias.



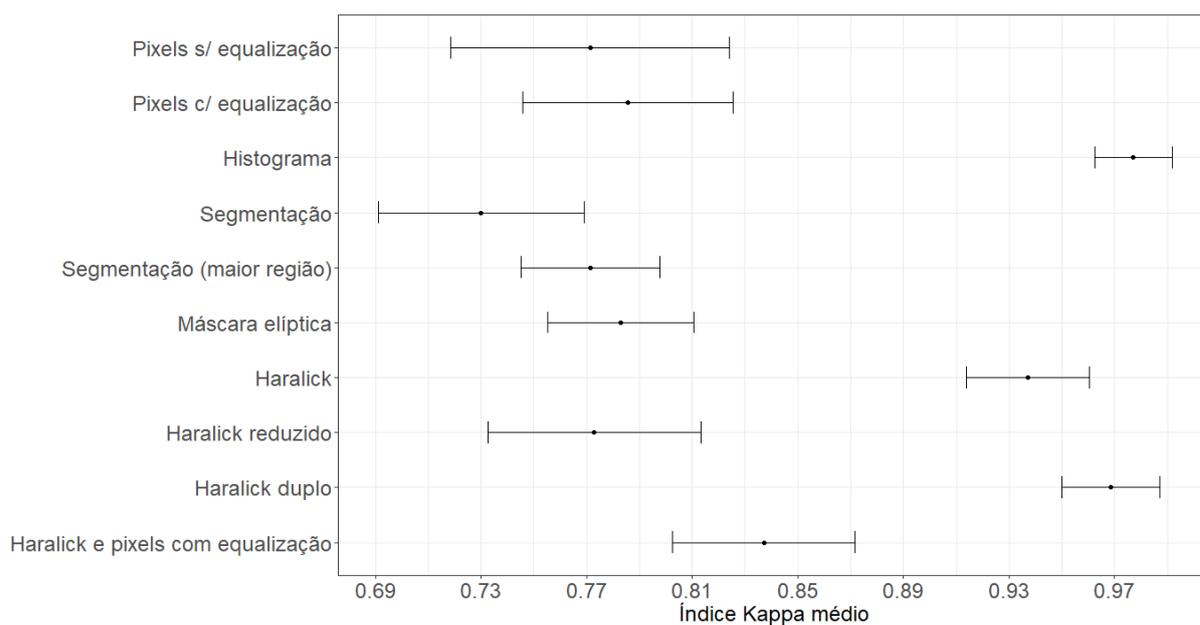
Fonte: do autor.

Figura 91 – Intervalos de confiança para o índice Kappa médio considerando as classes de sementes cheias e deformadas.



Fonte: do autor.

Figura 92 – Intervalos de confiança para o índice Kappa médio considerando as classes de sementes parcialmente cheias e deformadas.



Fonte: do autor.