



UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

EBER FABIANO PACANHELA

**ANÁLISE, CLASSIFICAÇÃO E DETECÇÃO AUTOMÁTICA DE NÓDULOS EM
RAÍZES DE CULTIVARES DE SOJA**

DISSERTAÇÃO DE MESTRADO

CORNÉLIO PROCÓPIO
2023

EBER FABIANO PACANHELA

ANÁLISE, CLASSIFICAÇÃO E DETECÇÃO AUTOMÁTICA DE NÓDULOS EM RAÍZES DE CULTIVARES DE SOJA

Analysis, Classification, and Automatic Detection of Nodules in Soybean Roots

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Tecnológica Federal do Paraná – UTFPR, como requisito parcial para obtenção do título de Mestre em Informática.

Orientador: Prof. Dr. Fabrício M. Lopes

Co-orientadora: Prof^ª. Dra. Priscila T. M. Saito

CORNÉLIO PROCÓPIO
2023



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



EBER FABIANO PACANHELA

ANÁLISE, CLASSIFICAÇÃO E DETECÇÃO AUTOMÁTICA DE NÓDULOS EM RAÍZES DE CULTIVARES DE SOJA

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Informática da Universidade Tecnológica Federal do Paraná (UTFPR).
Área de concentração: Computação Aplicada.

Data de aprovação: 09 de Agosto de 2023

Dr. Fabricio Martins Lopes, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Claiton De Oliveira, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Pedro Henrique Bugatti, Doutorado - Universidade Federal de São Carlos (Ufscar)

Dra. Priscila Tiemi Maeda Saito, Doutorado - Universidade Federal de São Carlos (Ufscar)

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 09/08/2023.

Dedico este trabalho aos meus pais, Eros e Rute, aos meus irmãos, Elizamar (*in memoriam*), Eder e Eliziane, ao meu marido Ronald e ao nosso filho Bento, aos meus amigos e a todos que contribuíram para esta caminhada.

AGRADECIMENTOS

Agradeço aos meus pais Eros e Rute, aos meus irmãos, Elizamar (*in memoriam*), Eder e Eliziane, ao meu marido Ronald e ao nosso filho Bento, pelo amor, apoio incondicional e compreensão durante os momentos de dedicação intensa. A vocês, minha gratidão eterna.

Agradeço aos meus orientadores Dr. Fabrício Lopes e Dra. Priscila Saito, que compartilharam seu conhecimento, paciência e sabedoria, guiando-me com seus ensinamentos valiosos e incentivando o meu crescimento acadêmico e pessoal. Muito obrigado!

Agradeço aos meus professores do PPGI, em especial ao professor Dr. Pedro Bugatti, cujas aulas sobre processamento de imagens foram essenciais no desenvolvimento deste estudo. Muito obrigado!

Agradeço aos meus amigos e a todos aqueles que, de alguma forma, estiveram presentes e contribuíram com essa trajetória e que tornaram possível a realização deste sonho acadêmico. Muito obrigado!

A inteligência é o que você usa quando não sabe o que fazer.
Jean Piaget

RESUMO

PACANHELA, Eber Fabiano. ANÁLISE, CLASSIFICAÇÃO E DETECÇÃO AUTOMÁTICA DE NÓDULOS EM RAÍZES DE CULTIVARES DE SOJA. 122 f. Dissertação – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2023.

A soja é a oleaginosa mais cultivada no mundo, possui grande número de variedades, alto valor econômico e elevado teor proteico. A produção de soja requer grandes quantidades de nitrogênio para a nutrição da planta. O nutriente pode ser fornecido por fertilizantes químicos, prejudiciais ao meio ambiente. No entanto, o processo de fixação biológica do nitrogênio (FBN) tem sido uma opção eficaz para aumentar a produtividade e minimizar os impactos ambientais. A FBN ocorre por meio da associação simbiótica de bactérias diazotróficas no sistema radicular de leguminosas e oleaginosas, onde nódulos radiculares são formados. Sabe-se que o avanço tecnológico tem impulsionado a adoção de tecnologias digitais na agricultura visando melhorar a eficiência e a produtividade do cultivo, reduzir impactos ambientais e promover a saúde humana. A sojicultura, devido à sua importância econômica, também pode ser beneficiada por essas tecnologias. Nesse contexto, este trabalho propõe o desenvolvimento de uma metodologia de inteligência computacional baseada em processamento de imagens e visão computacional para analisar, classificar e detectar automaticamente os nódulos radiculares resultantes da FBN na cultura da soja. No qual, em um estudo futuro a eficácia da nodulação em soja poderá ser avaliada, levando em consideração a quantidade de nódulos. Para isso, são adotados modelos clássicos e redes neurais convolucionais para a tarefa de classificação, bem como modelos de visão computacional para detecção de nódulos. Os resultados obtidos mostram que tanto os modelos tradicionais quanto as redes neurais convolucionais são promissores para a tarefa de classificação. Já na detecção, as redes neurais convolucionais demonstraram desempenho aprimorado. No entanto, ainda há espaço para melhorias visando aprimorar a precisão na detecção de nódulos em raízes de cultivares de soja.

Palavras-chave: Soja. Nodulação. Fixação Biológica do Nitrogênio. Visão Computacional. Aprendizado de Máquina. Aprendizado Profundo.

ABSTRACT

PACANELLA, Eber Fabiano. Analysis, Classification, and Automatic Detection of Nodules in Soybean Roots. 122 f. Dissertação – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2023.

Soybean is the most cultivated oilseed in the world, with many varieties, high economic value, and high protein content. Soybean production requires large amounts of nitrogen for the plant's nutrition. This nutrient can be supplied through chemical fertilizers, which harm the environment. However, the biological nitrogen fixation (BNF) process has effectively increased productivity and minimized environmental impacts. BNF occurs through the symbiotic association of diazotrophic bacteria in the root system of legumes and oilseeds, where root nodules are formed. It is known that technological advancements have driven the adoption of digital technologies in agriculture to improve efficiency and productivity, reduce environmental impacts, and promote human health. Due to its economic importance, soybean cultivation can benefit from these technologies. This study proposes developing a computational intelligence methodology based on image processing and computer vision to analyze, classify, and automatically detect the root nodules resulting from BNF in soybean crops. In which, a future study, the effectiveness of soybean nodulation can be evaluated, considering the quantity of nodules. For this purpose, classical models and convolutional neural networks are adopted for the classification task and computer vision models for nodule detection. The results show that traditional models and convolutional neural networks are promising for the classification task. In the detection task, convolutional neural networks demonstrated improved performance. However, there is still room for improvement to enhance nodule detection accuracy in soybean cultivar roots.

Keywords: Soybean. Nodulation. Biological Nitrogen Fixation. Computer Vision. Machine Learning. Deep Learning.

LISTA DE FIGURAS

FIGURA 1	– Convenção do sistema de coordenadas para representação de imagens digitais.	25
FIGURA 2	– Fases de um sistema de processamento de imagens digitais simplificado.	29
FIGURA 3	– Categorias de aprendizado de máquina.	34
FIGURA 4	– Esquema dos constituintes da célula neuronal.	37
FIGURA 5	– Modelo matemático de um neurônio artificial.	39
FIGURA 6	– Representação de uma rede neural multicamadas com três camadas ocultas.	40
FIGURA 7	– Representação de uma rede neural convolucional.	40
FIGURA 8	– Arcabouço da metodologia proposta. (a) Abordagem de classificação por meio de modelos tradicionais de aprendizado de máquina e arquiteturas profundas; (b) Abordagem de detecção por meio de modelos de visão computacional.	46
FIGURA 9	– Exemplo de imagens do conjunto de dados. (a) Raízes contendo nódulos; (b) Nódulos isolados das raízes.	47
FIGURA 10	– Exemplo de imagens com diferentes orientações do conjunto de dados.	48
FIGURA 11	– Exemplo de imagens com orientação padronizada.	49
FIGURA 12	– Exemplo de anotação das imagens utilizando a ferramenta LabelMe pela forma retângulo envolvente.	50
FIGURA 13	– Exemplo de anotação das regiões de interesse.	50
FIGURA 14	– Exemplo de regiões de interesse extraídas das imagens do conjunto de dados. (a) Nódulo; (b) Raiz; (c) Fundo.	52
FIGURA 15	– Exemplo de <i>data augmentation</i> nas regiões de interesse para cada uma das três classes (nódulo, raiz e fundo). (a), (b) e (c) Regiões de interesse originais; (d), (e) e (f) Regiões de interesse rotacionadas com o fator 0,20 aplicado; (g), (h) e (i) Regiões de interesse com ajuste no contraste aplicado no fator 0,20; (j), (k) e (l) Regiões de interesse com ajuste no brilho aplicado no fator 0,20.	54
FIGURA 16	– Tempo médio de treinamento e classificação obtidos no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas para os melhores resultados (i.e. maiores acurácias médias – Tabelas 3 e 4). Os modelos foram treinados no conjunto de dados original. (a) e (b) Treinamento e classificação no conjunto de dados original; (c) e (d) Treinamento e classificação no conjunto de dados aumentado.	63
FIGURA 17	– Média das métricas (acurácia, sensibilidade, especificidade, precisão, F1 e <i>kappa</i>) obtidas no conjunto de dados de teste pelos	

	classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas para os melhores resultados (i.e. maiores acurácias médias – Tabelas 3 e 4). Os modelos foram treinados no conjunto de dados original e aumentado. a) KNN (original); b) KNN (aumentado) c) RF (original); d) RF (aumentado); e) SVM (KL) (original); f) SVM (KL) (aumentado) g) SVM (KPG2) (original) e h) SVM (KPG2) (aumentado).	66
FIGURA 18 –	Média das matrizes de confusão obtidas no conjunto de dados de teste pelo classificador RF, utilizando as características obtidas pelo extrator tradicional GCH e arquitetura profunda ResNet101. O modelo foi treinado no conjunto de dados original. a) RF e GCH; b) RF e ResNet101.	66
FIGURA 19 –	Exemplos de confusões na classificação realizada pelo classificador RF em combinação com o extrator GCH. a) Raiz predita como nódulo e b) Fundo predito como nódulo.	67
FIGURA 20 –	Tempo médio de treinamento e classificação obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados nos conjuntos de dados original e aumentado. (a) e (b) Treinamento e classificação no conjunto de dados original; (c) e (d) Treinamento e classificação no conjunto de dados aumentado.	70
FIGURA 21 –	Média das métricas (acurácia, <i>loss</i> , sensibilidade, especificidade, precisão, F1 e <i>kappa</i>) obtidas no conjunto de dados de teste usando classificação por redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados nos conjuntos de dados original e aumentado. a) Conjunto de dados original; b) Conjunto de dados aumentado.	71
FIGURA 22 –	Progressão média das métricas acurácia x <i>loss</i> obtidas durante o treinamento das redes neurais convolucionais. Os modelos foram treinados no conjunto de dados original. a) Modelo VGG16; b) Modelo VGG19.	72
FIGURA 23 –	Média das matrizes de confusão obtidas no conjunto de dados de teste para as redes neurais convolucionais VGG16 e VGG19. Os modelos foram treinados no conjunto de dados original. (a) Modelo VGG16; (b) Modelo VGG19.	72
FIGURA 24–	Média das métricas <i>average precision</i> (AP@0,5), precisão e sensibilidade obtidas no conjunto de dados de teste pelo modelo de visão computacional YOLO v7.	75
FIGURA 25 –	Resultados da detecção de nódulos obtidos por cada modelo: (a) e (b) YOLO v7, (c) e (d) RetinaNet e (e) e (f) Faster R-CNN a partir de imagens de raízes de soja contendo nódulos (esquerda) e nódulos isolados das raízes (direita).	76
FIGURA 26 –	Média das matrizes de confusão obtidas no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e	

arquiteturas profundas para os melhores resultados (i.e. maiores acurácias médias – Tabela 3). Os modelos foram treinados no conjunto de dados original. a) KNN e ResNet101; b) KNN e ResNet152; c) RF e GCH; d) RF e ResNet101; e) RF e ResNet152; f) RF e ResNet50; g) SVM (KL) e DenseNet121; h) SVM (KL) e ResNet101; i) SVM (KL) e ResNet152; j) SVM (KL) e ResNet50; k) SVM (KPG2) e ResNet101; l) SVM (KPG2) e ResNet152; m) SVM (KPG2) e ResNet50. 99

FIGURA 27 – Média das matrizes de confusão obtidas no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas para os melhores resultados (i.e. maiores acurácias médias – Tabela 4). Os modelos foram treinados no conjunto de dados aumentado. a) KNN e ResNet101; b) KNN e ResNet152; c) KNN e ResNet50; d) RF e FOM (Color); e) RF e ResNet101; f) RF e ResNet152; g) RF e ResNet50; h) SVM (KL) e VGG16; i) SVM (KL) e VGG19; j) SVM (KL) e MobileNet; k) SVM (KL) e MobileNetV2; l) SVM (KL) e ResNet101; m) SVM (KL) e ResNet152; n) SVM (KL) e ResNet50; o) SVM (KPG2) e FOM (Color); p) SVM (KPG2) e VGG16; q) SVM (KPG2) e VGG19; r) SVM (KPG2) e MobileNet; s) SVM (KPG2) e MobileNetV2; t) SVM (KPG2) e ResNet101; u) SVM (KPG2) e ResNet152 e v) SVM (KPG2) e ResNet50. 103

FIGURA 28 – Progressão média das métricas acurácia x *loss* obtidas durante o treinamento das redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados no conjunto de dados original. a) VGG16; b) VGG19; c) DenseNet121; d) MobileNet; e) NASNetMobile; f) InceptionResNetV2; g) DenseNet169; h) DenseNet201; i) InceptionV3; j) ResNet101V2; k) ResNet152V2; l) ResNet50V2; m) Xception e n) NASNetLarge. 107

FIGURA 29 – Progressão média das métricas acurácia x *loss* obtidas durante o treinamento das redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados no conjunto de dados aumentado. a) VGG16; b) VGG19; c) DenseNet201 e d) Xception. 108

FIGURA 30 – Média das matrizes de confusão obtidas no conjunto de dados de teste usando classificação por redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados no conjunto de dados original. a) VGG16; b) VGG19; c) DenseNet121; d) MobileNet; e) NASNetMobile; f) InceptionResNetV2; g) DenseNet169; h) DenseNet201; i) InceptionV3; j) ResNet101V2; k) ResNet152V2; l) ResNet50V2; m) Xception e n) NASNetLarge. 111

FIGURA 31 – Média das matrizes de confusão obtidas no conjunto de dados de teste usando classificação por redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6).

Os modelos foram treinados no conjunto de dados aumentado. a) VGG16; b) VGG19; c) DenseNet201 e d) Xception. 112

LISTA DE TABELAS

TABELA 1	–	Descritores de imagens, tipos, quantidade de características e tempo de extração de características.	53
TABELA 2	–	Redes Neurais Convolucionais.	55
TABELA 3	–	Média de acurácias \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados original.	61
TABELA 4	–	Média de acurácias \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados aumentado.	62
TABELA 5	–	Redes Neurais Convolucionais.	67
TABELA 6	–	Média de acurácias \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados nos conjuntos de dados original e aumentado.	68
TABELA 7	–	Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados nos conjuntos de dados original e aumentado.	69
TABELA 8	–	Média da métrica <i>average precision</i> (AP@0,5) \pm desvio padrão obtidos no conjunto de dados de teste pelos modelos de visão computacional YOLO v7, RetinaNet e Faster R-CNN.	73
TABELA 9	–	Coefficiente de variação (%) da média da métrica <i>average precision</i> (AP@0,5) obtidos no conjunto de dados de teste pelos modelos de visão computacional YOLO v7, RetinaNet e Faster R-CNN.	74
TABELA 10	–	Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste pelo classificador KNN, utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados nos conjuntos de dados original e aumentado.	91
TABELA 11	–	Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste pelo classificador RF, utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados nos conjuntos de dados original e aumentado.	92
TABELA 12	–	Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste pelo classificador SVM (KL), utilizando as	

	características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados nos conjuntos de dados original e aumentado.	92
TABELA 13	– Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste pelo classificador SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados nos conjuntos de dados original e aumentado.	93
TABELA 14	– Tempo médio de treinamento e classificação \pm desvio padrão obtidos pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados original.	93
TABELA 15	– Tempo médio de treinamento e classificação \pm desvio padrão obtidos pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados aumentado.	94
TABELA 16	– Média das métricas (sensibilidade, especificidade, precisão, F1 e <i>kappa</i>) \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores KNN e RF, utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados original.	94
TABELA 17	– Média das métricas (sensibilidade, especificidade, precisão, F1 e <i>kappa</i>) \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados original.	95
TABELA 18	– Média das métricas (sensibilidade, especificidade, precisão, F1 e <i>kappa</i>) \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores KNN e RF, utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados aumentado.	95
TABELA 19	– Média das métricas (sensibilidade, especificidade, precisão, F1 e <i>kappa</i>) \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados aumentado.	96
TABELA 20	– Tempo médio de treinamento e classificação \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados nos conjuntos de dados original e aumentado.	96
TABELA 21	– Média das métricas (<i>loss</i> , sensibilidade, especificidade, precisão, F1 e <i>kappa</i>) \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos	

	foram treinados no conjunto de dados original.	99
TABELA 22 –	Média das métricas (<i>loss</i> , sensibilidade, especificidade, precisão, F1 e <i>kappa</i>) \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados no conjunto de dados aumentado.	104

LISTA DE SIGLAS

AA-DIR	Aprendizagem Ativa através da Diversidade, Incerteza e Representatividade
AP	<i>Average Precision</i> (Precisão Média)
BA	Bahia
BIC	<i>Border/Interior pixel Classification</i> (Classificação de Pixels de Borda/Interior)
CNN	<i>Convolutional Neural Network</i> (Rede Neural Convolucional)
CMY	<i>Cyan, Magenta, Yellow</i> (Ciano, Magenta, Amarelo)
CMYK	<i>Cyan, Magenta, Yellow, Black</i> (Ciano, Magenta, Amarelo, Preto)
Embrapa	Empresa Brasileira de Pesquisa Agropecuária
Fafen	Fábrica de Fertilizantes Nitrogenados
FOM	<i>First Order Measures</i> (Medidas de Primeira Ordem)
FBN	Fixação Biológica do Nitrogênio
GCH	<i>Global Color Histogram</i> (Histograma de Cor Global)
HD	<i>Hard Disk</i> (Disco Rígido)
HSI	<i>Hue, Saturation, Intensity</i> (Matiz, Saturação, Intensidade)
IBGE	Instituto Brasileiro de Geografia e Estatística
IA	Inteligência Artificial
IoT	Internet das Coisas
IoU	<i>Intersection over Union</i> (Interseção sobre União)
JSON	<i>JavaScript Object Notation</i> (Notação de Objeto JavaScript)
KNN	<i>K-Nearest Neighbors</i> (K-Vizinhos Mais Próximos)
KL	<i>Kernel Linear</i> (Núcleo Linear)
KPG2	<i>Kernel Polinomial de Grau 2</i> (Núcleo Polinomial de Grau 2)
LSPA	Levantamento Sistemático da Produção Agrícola
LDA	<i>Linear Discriminant Analysis</i> (Análise Discriminante Linear)
LBP	<i>Local Binary Patterns</i> (Padrões Binários Locais)
LCH	<i>Local Color Histogram</i> (Histograma de Cor Local)
Mapa	Ministério da Agricultura, Pecuária e Abastecimento
MLP	<i>Multi-layer Perceptron</i> (<i>Perceptron</i> de Múltiplas Camadas)
OPF	<i>Optimum-Path Forest</i> (Floresta de Caminho Ótimo)
PR	Paraná
RAM	<i>Random Access Memory</i> (Memória de Acesso Aleatório)
RF	<i>Random Forest</i> (Floresta Aleatória)
RNN	<i>Recurrent Neural Network</i> (Rede Neural Recorrente)
RGB	<i>Red, Green, Blue</i> (Vermelho, Verde, Azul)
RNA	Rede Neural Artificial
SP	São Paulo
SE	Sergipe
SSD	<i>Solid State Drives</i> (Unidades de Estado Sólido)
SVM	<i>Support Vector Machine</i> (Máquina de Vetores de Suporte)
USDA	<i>United States Department of Agriculture</i> (Departamento de Agricultura dos Estados Unidos)
WWF	<i>World Wide Fund for Nature</i> (Fundo Mundial para a Natureza)

LISTA DE SÍMBOLOS

NH ₃	Amônia Anidra
N ₂	Dinitrogênio
GB	Gigabyte
GHz	Giga-hertz
°	Grau
°C	Graus Celsius
H	Hidrogênio
N	Nitrogênio
%	Porcentagem
atm	Pressão Atmosférica
kg	Quilograma
s	Segundos
TB	Terabyte

SUMÁRIO

1	INTRODUÇÃO	9
1.1	Problemas e premissas	12
1.2	Objetivos	12
1.2.1	Objetivo geral	12
1.2.2	Objetivos específicos	13
1.3	Justificativa	13
1.4	Organização do texto	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	O cultivo de soja	16
2.1.1	Impactos ambientais da sojicultura	18
2.1.2	Fertilizantes nitrogenados e a fixação biológica do nitrogênio	19
2.2	Análise de imagens digitais	24
2.2.1	Processamento de imagens digitais e visão computacional	27
2.2.2	Extração de características	31
2.3	Aprendizado de máquina	33
2.4	Aprendizado profundo	36
2.5	Trabalhos relacionados	42
3	MATERIAIS E MÉTODOS	45
3.1	Aquisição do conjunto de dados	46
3.2	Pré-processamento e anotação do conjunto de dados	47
3.3	Segmentação das regiões de interesse e extração de características	51
3.4	Classificação	54
3.5	Detecção	57
4	RESULTADOS E DISCUSSÃO	59
4.1	Classificação por meio de modelos clássicos de aprendizado de máquina	60
4.2	Classificação por meio de redes neurais convolucionais	67
4.3	Detecção por meio de modelos de visão computacional	72
5	CONCLUSÃO E TRABALHOS FUTUROS	77
	REFERÊNCIAS	81
	APÊNDICE A – TABELAS E FIGURAS COMPLEMENTARES	91

1 INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill) é a oleaginosa mais cultivada no mundo. Ela possui grande número de variedades, alto valor econômico e elevado teor proteico em seus grãos, com cerca de 38%. Além disso, atualmente conta com cultivares que se adaptam facilmente aos diversos tipos de clima (DREONI et al., 2022; MORETTI et al., 2020; CONTE et al., 2020; BARRETO, 2004).

Existe uma crescente demanda mundial por alimentos proteicos e a produção de soja tem contribuído para suprir essa necessidade, pois ela é utilizada tanto na alimentação humana, por meio do tofu, óleo vegetal alimentício, leite de soja e proteínas texturizadas, quanto no preparo de rações para animais (CONTE et al., 2020; AGROLINK, 2007). Neste cenário, a soja tem se consolidado como um produto importante para a economia do Brasil. Na safra de 2022, o país conquistou a colocação de maior produtor mundial do grão, com uma produção total de aproximadamente 119,5 milhões de toneladas, o que representa cerca de 45,42% da produção brasileira de cereais, leguminosas e oleaginosas (USDA, 2023; IBGE, 2023).

No entanto, apesar dos benefícios econômicos que a sojicultura traz para os países que produzem e comercializam a soja, é importante destacar que sua produção apresenta impactos para o meio ambiente (BRUM et al., 2011). Um dos impactos ocorre por meio do uso intensivo de agroquímicos, especialmente os fertilizantes nitrogenados, pois a planta necessita de alta demanda de nitrogênio em seu processo produtivo (GOMES, 2019). De acordo com Hungria e Mendes (2015), são necessários 80 kg de nitrogênio para a produção de 1.000 kg de grãos de soja.

Como mencionado, a fertilização nitrogenada no cultivo da soja é prejudicial ao meio ambiente, pois contamina o solo e a água (YADAV et al., 2023; WWF, 2014).

Visando reduzir os impactos ambientais provocados pela utilização desnecessária ou incorreta de agroquímicos nitrogenados na sojicultura, o uso de biofertilizantes tem se apresentado como uma alternativa eficiente e de baixo custo para atingir uma produção agrícola sustentável (HUNGRIA et al., 2020). Visto que, a fixação biológica do nitrogênio (FBN) na cultura da soja tem se destacado pela capacidade de aumentar a produtividade sem aumentar excessivamente os custos, além de possuir impacto ambiental mínimo (HUNGRIA et al., 2001).

O processo de fixação biológica do nitrogênio ocorre por meio da associação simbiótica de bactérias diazotróficas no sistema radicular de leguminosas e oleaginosas, formando nódulos radiculares. As bactérias se abrigam nestes nódulos e recebem da planta proteção e alimento e, em compensação, elas capturam e transformam o nitrogênio atmosférico (N_2) em forma assimilável para a planta, fornecendo o nitrogênio necessário para o seu desenvolvimento (NOGUEIRA et al., 2018). No cultivo da soja, as bactérias capazes de nodular as raízes da soja pertencem ao gênero *Bradyrhizobium*, principalmente nas espécies *Bradyrhizobium japonicum* e *Bradyrhizobium elkanii*; e, comumente são chamadas de rizóbio ou *bradyrhizobium diazoefficiens* (HUNGRIA et al., 2001).

Além do processo de FBN no cultivo da soja, o avanço tecnológico tem impulsionado a adoção de tecnologias digitais na agricultura. A agricultura de precisão, como é conhecida, utiliza dispositivos eletrônicos, sensores e tecnologia de automação aliados a avançadas ferramentas como Internet das Coisas (IoT), mineração de dados, aprendizado de máquina, aprendizado profundo, visão computacional, ciência de dados, entre outras. Essas tecnologias atuam como um sistema agrícola integrado com foco em informação e produção, no qual possui o objetivo de minimizar o esforço de trabalho, auxiliar na tomada de decisões, maximizar a lucratividade, aumentar a eficiência e a produtividade da cultura melhorando a qualidade da produção, ao mesmo tempo em que minimiza os impactos ambientais e promove a saúde humana (DURAI; SHAMILI, 2022; COULIBALY et al., 2022; MESHARAM et al., 2021). De acordo com Lu e Young (2020), as tecnologias de visão computacional que utilizam imagens digitais são capazes de fornecer informações precisas e específicas sobre as culturas e seus ambientes. No entanto, nas áreas de

aprendizado de máquina, aprendizado profundo e visão computacional a identificação de nódulos em raízes de cultivares de soja tem sido alvo de poucos trabalhos na literatura.

Pereira et al. (2019c) propuseram uma abordagem para o processamento e a classificação automática do vigor de sementes de soja, por meio de uma perspectiva inteligente baseada em visão computacional. No qual, o trabalho consiste no processamento e análise de imagens de sementes de soja provenientes do teste de tetrazólio e aprendizado de máquina para a classificação automática do vigor das sementes. Pereira et al. (2021) também propuseram uma abordagem de aprendizado ativo para a classificação do vigor de sementes de soja e um estudo comparativo entre as abordagens de aprendizagem ativa e supervisionada.

Medeiros et al. (2020), por meio de métodos interativos e tradicionais de aprendizado de máquina, apresentaram uma metodologia para classificar sementes e mudas de soja de acordo com sua aparência e potencial fisiológico.

Chung et al. (2020) propuseram a caracterização do nódulo de raiz de soja exposta ao silício, por meio de aprendizado profundo baseado em imagens, no qual imagens 2D foram adotadas para a extração de características morfológicas das raízes, como área de superfície, comprimento e ângulo, bem como quantidade e tamanho dos nódulos radiculares.

Jubery et al. (2021) desenvolveram uma abordagem para detectar automaticamente nódulos em imagens de raízes de soja através de aprendizado profundo, fornecendo a quantidade de nódulos, tamanho dos nódulos e distribuições de nódulos ao longo da raiz principal.

Neste contexto, o objetivo deste trabalho consiste no desenvolvimento de uma metodologia de inteligência computacional, por meio de processamento de imagens digitais e visão computacional, para a análise, a classificação e a detecção automática de nódulos radiculares resultantes do processo de fixação biológica do nitrogênio na sojicultura, no qual em um trabalho futuro a eficácia da nodulação em soja poderá ser avaliada, levando em consideração a quantidade de nódulos.

1.1 Problemas e premissas

A fixação biológica do nitrogênio é de extrema importância para uma produção agrícola sustentável (HUNGRIA et al., 2020) e um forte indício de que o processo de fixação biológica do nitrogênio está ocorrendo no cultivo da soja é a presença de nódulos radiculares.

No entanto, o processo de análise e contagem destes nódulos radiculares, muitas vezes, ocorrem de maneira rudimentar, ou seja, a contagem da quantidade de nódulos presentes nas raízes de soja é realizada manualmente, o que pode resultar em erros de contagem. Além disso, como pode haver grande quantidade de nódulos em uma única raiz, o processo de contagem manual acaba se tornando difícil e moroso.

Esta pesquisa busca, através do processamento de imagens digitais e visão computacional, minimizar o esforço humano e a redução de erros na contagem dos nódulos radiculares presentes na raiz da soja. Logo, é proposto o desenvolvimento de uma metodologia de inteligência computacional para a análise, classificação e detecção automática destes nódulos radiculares.

1.2 Objetivos

1.2.1 Objetivo geral

Considerando a parceria com a Embrapa Soja Londrina - PR, em particular com a equipe liderada pela Dra. Mariangela Hungria, da qual houve a geração de um conjunto de imagens de raízes de soja.

O objetivo geral deste trabalho consiste no desenvolvimento de uma metodologia de inteligência computacional, por meio de processamento de imagens digitais e visão computacional, para a análise, a classificação e a detecção automática de nódulos radiculares resultantes do processo de fixação biológica do nitrogênio na sojicultura, no qual em um trabalho futuro a eficácia da nodulação em soja poderá ser avaliada, levando em consideração a quantidade de nódulos.

1.2.2 Objetivos específicos

Considerando o objetivo geral deste trabalho, são propostos alguns objetivos específicos, descritos a seguir e que compõem etapas importantes para a realização deste trabalho:

- Produzir e disponibilizar um conjunto de dados anotado de imagens de raízes de soja;
- Desenvolver metodologia de inteligência computacional, por meio de processamento de imagens digitais e visão computacional, para analisar, classificar e detectar nódulos radiculares em cultivares de soja;
- Validar a assertividade dos modelos clássicos de aprendizado de máquina na classificação de nódulos em raízes de soja;
- Identificar quais descritores de imagens, isto é, descritores tradicionais e arquiteturas profundas, em conjunto com os modelos de aprendizado de máquina são mais adequados para a classificação dos nódulos radiculares;
- Identificar e validar modelos de visão computacional para detectar automaticamente nódulos em raízes de cultivares de soja.

1.3 Justificativa

De acordo com Moreti et al. (2021), o uso de Inteligência Artificial (IA) tem sido realidade nos mais variados setores da economia, inclusive no agronegócio, e sua aplicação tem apresentado consequências positivas. Deste modo, a IA vem conquistando cada vez mais espaço no processo de produção do setor agro, visto que, a automatização de processos e otimização de tempo permitem facilitar os procedimentos e majorar a lucratividade. Logo, atender as demandas do setor por meio da transferência de tecnologia da computação para o agro, contribui portanto, com a competitividade no mercado.

No entanto, existem poucos trabalhos na literatura que visam a identificação

de nódulos em raízes de soja, e, menos ainda quando se trata do uso de abordagens tradicionais de aprendizado de máquina.

A identificação de nódulos em raízes de cultivares de soja é importante, pois permite avaliar a eficácia da nodulação em soja, visto que a presença de tais nódulos é um forte indício, porém não absoluto, de que o processo de fixação biológica do nitrogênio está ocorrendo. É importante destacar que somente a presença de nódulos radiculares em soja não garante que a fixação biológica do nitrogênio está ocorrendo, visto que, outros fatores são necessários, por exemplo, os nódulos apresentam em seu interior coloração rósea intensa quando estão em plena atividade (HUNGRIA et al., 2001).

Saber se a fixação biológica do nitrogênio está ocorrendo é importante, pois ela tem relevância como uma alternativa de baixo custo para aumentar a produtividade e atingir uma produção agrícola sustentável (HUNGRIA et al., 2020), isto é, apresenta a melhor relação custo/benefício para o produtor (HUNGRIA et al., 2001). Além disso, a fixação biológica do nitrogênio possui impacto ambiental mínimo (HUNGRIA et al., 2001), no qual permite dispensar o uso de agroquímicos nitrogenados – altamente prejudiciais ao meio ambiente. Pesquisadores como Nogueira et al. (2018), afirmam que a fixação biológica do nitrogênio representa um dos pilares de sustentabilidade no cultivo de soja no Brasil.

Entendendo a importância da fixação biológica do nitrogênio na sojicultura, esta pesquisa propõe o desenvolvimento de uma metodologia de inteligência computacional, por meio de processamento de imagens e visão computacional, para analisar, classificar e detectar nódulos radiculares em cultivares de soja de maneira automática, minimizando o trabalho manual. Com os resultados desta pesquisa será possível, em um trabalho futuro, avaliar a eficácia da nodulação em soja de forma automática, levando em consideração a quantidade de nódulos em raízes de cultivares de soja.

1.4 Organização do texto

O presente trabalho está organizado de forma que o Capítulo 2 contém o referencial teórico, no qual apresenta conceitos importantes para o desenvolvimento desta pesquisa. Além disso, este capítulo procura definir minimamente os elementos necessários para a total compreensão dos itens abordados neste trabalho.

No Capítulo 3, são apresentados os passos realizados neste trabalho para obter os descritores de imagens e os modelos tradicionais de aprendizado de máquina, assim como as redes neurais convolucionais mais adequadas e assertivas para a classificação de nódulos radiculares em cultivares de soja. Além disso, foi abordada a detecção de nódulos por meio dos modelos de visão computacional. O capítulo descreve de forma detalhada os procedimentos metodológicos adotados, oferecendo uma visão geral do processo de seleção e desenvolvimento dos métodos utilizados nesta pesquisa.

No Capítulo 4, são apresentados os resultados obtidos por meio da implementação da metodologia de inteligência computacional proposta neste trabalho. Nesse capítulo, são apresentados os resultados quantitativos, como acurácia, sensibilidade, especificidade, precisão, F1, *kappa*, *loss*, AP@0,5, bem como possíveis análises qualitativas dos resultados.

No Capítulo 5, é fornecida a conclusão do trabalho com base nos experimentos realizados e resultados obtidos. Nesse capítulo, são apresentadas as principais descobertas e contribuições deste estudo, bem como as limitações encontradas. Por fim, são propostas sugestões para trabalhos futuros com o objetivo de aprimorar e expandir os resultados obtidos, fornecendo uma base para pesquisas subsequentes.

2 FUNDAMENTAÇÃO TEÓRICA

Neste Capítulo é apresentado os principais conceitos utilizados neste trabalho. Na Seção 2.1 é esclarecido de forma sucinta a importância da sojicultura para o Brasil, permeando os impactos ambientais na produção de soja e a relevância da fixação biológica do nitrogênio como uma alternativa para atingir uma produção agrícola sustentável. Apresenta também os nódulos radiculares de soja, resultantes do processo de fixação biológica do nitrogênio, no qual é objeto de interesse neste trabalho. Na Seção 2.2, é abordado a importância da visão no desempenho de funções complexas, como, localização, detecção, reconhecimento e interpretação dos objetos no ambiente. Esclarece também conceitos pertinentes as áreas de processamento de imagens digitais e visão computacional. E por fim, nas Seções 2.3 e 2.4 é destacado a relevância do aprendizado de máquina e do aprendizado profundo nos dias atuais para resolver problemas complexos em diversas áreas, como visão computacional (imagem), processamento de linguagem natural (texto) e reconhecimento automático de fala (áudio) utilizando principalmente as redes neurais artificiais.

2.1 O cultivo de soja

A soja (*Glycine max* (L.) Merril) é uma oleaginosa com grãos que apresentam um alto teor proteico, com cerca de 38%. Devido à sua adaptabilidade a diferentes climas, variedades diversas e valor econômico significativo, a soja é a leguminosa mais cultivada no mundo (DREONI et al., 2022; MORETTI et al., 2020; CONTE et al., 2020; BARRETO, 2004).

Conforme Bonato e Vello (1999) a cultura da soja é classificada como uma

planta de dias curtos e sensível às variações entre latitudes ou datas de semeadura, graças a suas respostas às mudanças no fotoperíodo. Sendo a época de semeadura um dos principais fatores que influenciam o porte, o ciclo e o rendimento da cultura (EGEWARTH, 2019). Segundo Egewarth (2019), a temperatura do ar e do solo são essenciais para o desenvolvimento da soja. Para a pesquisadora a planta se adapta melhor a temperaturas do ar e do solo variando entre 20 °C e 30 °C, sendo que a temperatura do ar ideal para o desenvolvimento da soja é em torno de 30 °C e a temperatura do solo em torno de 25 °C, nestas condições a planta possui um crescimento rápido e uniforme.

Em virtude da alta demanda global por alimentos proteicos, a cultura da soja está se tornando cada vez mais importante. Aproximadamente 85% da produção mundial de soja é processada para produzir óleo, torta e farelo, sendo este dois últimos amplamente utilizados na preparação de rações para animais. No entanto, cerca de apenas 10% são utilizados diretamente para alimentação humana, por exemplo, tofu, óleo vegetal alimentício, leite de soja, proteínas texturizadas, entre outros. E os 5% restantes são usados como sementes ou para alimentação dos animais nas próprias fazendas produtoras (CONTE et al., 2020; AGROLINK, 2007). Além disso, a soja tem sido utilizada no combate ao alcoolismo, tratamento de câncer de mama e próstata, na produção de tintas, colas, protetores solares e na fabricação de biocombustíveis (PINHO, 2004). No cenário brasileiro, a sojicultura também possui vasta relevância socioeconômica, pois se trata de um dos principais produtos exportados (FLORENCIO, 2017).

De acordo com o Levantamento Sistemático da Produção Agrícola (LSPA) do Instituto Brasileiro de Geografia e Estatística (IBGE), em dezembro de 2022 o Brasil atingiu aproximadamente 263,2 milhões de toneladas na produção de cereais, leguminosas e oleaginosas. A produção foi ligeiramente maior que a produção do mesmo período no ano de 2021, cerca de 9,9 milhões de toneladas. No entanto, a produção de soja teve uma redução 11,42%, atingindo cerca de 119,5 milhões de toneladas em dezembro de 2022, representando em torno de 45,42% da produção de cereais, leguminosas e oleaginosas (IBGE, 2023). Ainda em conformidade com o IBGE (2023), neste período, os estados brasileiros com maior produção de soja, em

milhões de toneladas, foram: Mato Grosso (38), Goiás (15) e Paraná (12,2).

Em dezembro de 2022 a área plantada de soja foi próxima de 41,1 milhões de hectares, tendo uma diferença de 223,2 mil hectares a menos no total da área colhida. O rendimento médio da soja, neste período, foi de 2.923 quilogramas por hectare (IBGE, 2023).

Com base nos dados supracitados, é evidente que o Brasil tem se consolidado como um importante produtor de soja, sendo confirmado pelo relatório *Soybeans: World Supply and Distribution* divulgado pelo *United States Department of Agriculture* (USDA), no qual o país alcançou o patamar de maior produtor mundial do grão na safra de 2021/22, ficando à frente dos Estados Unidos, Argentina e China. Essa tendência deve se manter na safra de 2022/23, visto que, o relatório mantém o Brasil na mesma posição. O relatório classifica também o Brasil como maior exportador de soja do mundo, seguido pelos Estados Unidos, Paraguai, Argentina e Canadá (USDA, 2023). Além disso, é importante destacar que a China, a União Europeia, o México e a Argentina são os maiores importadores de soja do mundo, conforme evidencia o referido relatório da USDA.

2.1.1 Impactos ambientais da sojicultura

O cultivo de soja teve um crescimento acentuado nas últimas décadas e essa tendência de crescimento deve se manter nos próximos anos. Visto que, a soja é uma rica fonte de proteína e energia, considerada como produto fundamental para o abastecimento mundial de alimentos, pois a população está aumentando e estima-se que chegue a 10 bilhões até 2050 (SHAHWAR et al., 2023). O consumo da soja pela população pode ser direto ou indireto, ou seja, através do consumo da soja em grão (tofu, molho de soja, entre outros) ou do consumo da carne animal, no qual a soja é moída e transformada em farelo para ração (CONTE et al., 2020; WWF, 2014).

É importante ressaltar que, apesar da sojicultura trazer benefícios econômicos para os países que produzem e comercializam a soja, a produção desta oleaginosa também apresenta impactos para o meio ambiente. A alta demanda nacional ou internacional deste produto, faz com que milhões de hectares de florestas, pastagens

e savanas sejam convertidas, de forma direta ou indiretamente, em terra agrícola, colocando em xeque a biodiversidade (BRUM et al., 2011).

Além dos impactos ambientais, o uso intensivo de fertilizantes, pesticidas e herbicidas na produção de soja, independentemente do tamanho da unidade agrícola, ameaçam demasiadamente o meio ambiente, provocando a contaminação do solo e da água, bem como a sua biodiversidade em geral, além de representar uma séria ameaça para a saúde das pessoas (YADAV et al., 2023; WWF, 2014). Arvor et al. (2010) reforçam essa ideia, pois para os autores a poluição dos cursos d'água, em consequência do uso de agroquímicos e à erosão do solo, também afetam negativamente os ecossistemas naturais.

2.1.2 Fertilizantes nitrogenados e a fixação biológica do nitrogênio

Segundo Faquin (2005) o nitrogênio (N) desempenha um papel fundamental no crescimento e produção das culturas, participando de diversos processos fisiológicos essenciais para o desenvolvimento das plantas. Destacando-se por ser um nutriente constituinte dos aminoácidos, clorofila e muitas enzimas essenciais, que contribuem para o crescimento e desenvolvimento da área foliar e do sistema radicular das plantas (CABEZAS et al., 2000). Castro et al. (1999) afirmam que o nitrogênio é considerado um dos principais nutrientes para alcançar produtividade elevada em culturas anuais, inclusive para a produção de soja, conforme destaca Malavolta (1987).

A deficiência de nitrogênio em cultivares de soja pode promover baixa produção, com plantas menores e de crescimento lento (FAQUIN, 2005), ou seja, compromete o crescimento e o desenvolvimento da planta, pois a falta de nitrogênio provoca a degradação da molécula de clorofila, interferindo negativamente na absorção e utilização da luz solar no processo de fotossíntese (SILVA et al., 2012). De acordo com Taiz et al. (2017) a carência de nitrogênio faz com que a maioria das espécies apresente clorose (amarelecimento das folhas), principalmente nas folhas mais velhas e próximas à base da planta, podendo tornar-se completamente amarelas ou castanhas, além de se desprender da planta quando o nitrogênio é escasso.

Gomes (2019) ressalta que a cultura de soja necessita de alta demanda de

nitrogênio em seu processo produtivo, por causa do alto teor proteico dos grãos desta leguminosa. Hungria et al. (2001) salienta que os grãos de soja apresentam um teor médio de 6,5% de nitrogênio. Posto isto, são necessários 65 kg de nitrogênio para produzir 1.000 kg de grãos de soja, soma-se a isso, pelo menos mais 15 kg de nitrogênio para as folhas, caule e raízes, resultando em uma demanda de 80 kg de nitrogênio. Desse modo, são necessários 240 kg de nitrogênio para a obter o rendimento médio nacional de 3.000 kg de grãos/ha (HUNGRIA; MENDES, 2015).

Apesar de aproximadamente 78% da atmosfera ser composta por nitrogênio (MAPA, 2011), ele praticamente é inexistente nas rochas que dão origem aos solos, conforme afirma Bigaton (2017). Já o reservatório de nitrogênio existente na matéria orgânica do solo é restrito, esgotando-se rapidamente após alguns cultivos. Além do mais, os processos de decomposição da matéria orgânica e de perdas de nitrogênio se aceleram no território brasileiro, por efeito das condições de temperatura e umidade predominantes. Ocasionalmente em solos com baixos teores de nitrogênio, capazes de fornecer, em média, apenas 15 a 30 kg de nitrogênio por lavoura (HUNGRIA; MENDES, 2015).

Uma das maneiras de suprir a deficiência de nutrientes na cultura de soja é por meio do uso de agroquímicos, especialmente os fertilizantes nitrogenados para atender as necessidades da planta em relação à carência de nitrogênio. Conforme Dias e Fernandes (2006) os agroquímicos nitrogenados possuem o nitrogênio como nutriente principal em sua composição, sendo este, originado da fabricação da amônia anidra (NH_3), que é a matéria-prima elementar de todos os fertilizantes nitrogenados sintéticos. Por sua vez, a amônia anidra é um gás resultante da reação do gás de síntese, uma mistura na relação 1:3 de nitrogênio (N) proveniente do ar com o hidrogênio (H) de fontes diversas – do gás natural, da nafta, do óleo combustível ou de outros derivados de petróleo, sendo o gás natural o mais utilizado e também a melhor fonte de hidrogênio para a produção deste agroquímico.

Visto que a amônia é o insumo fundamental para obtenção dos fertilizantes nitrogenados, Dias e Fernandes (2006) apontam que existem poucas fábricas que produzem matérias-primas para adubos nitrogenados no Brasil, sendo elas a Fábrica de Fertilizantes Nitrogenados (Fafen), em Laranjeiras (SE) e Camaçari (BA), e a

Ultrafertil, em Cubatão (SP) e Araucária (PR). Para cobrir o déficit nacional de NH_3 , em torno de 55%, as importações procedem, principalmente, da Rússia e da Ucrânia.

Hungria et al. (2001) revela que o processo industrial que transforma o dinitrogênio (N_2) em amônia (NH_3) necessita: a) hidrogênio (derivado de gás de petróleo); b) catalisador contendo ferro; c) altas temperaturas (300° a 600°C); e d) altas pressões (200 a 800 atm). Dessa forma, Hungria et al. (2001, p. 12) enfatizam que “o gasto de fontes energéticas não-renováveis é estimado em seis barris de petróleo por tonelada de NH_3 sintetizada”.

É importante destacar que o fertilizante nitrogenado incorporado ao solo pode sofrer ação de diversos processos, tais como, processos microbiológicos (nitrificação, desnitrificação, imobilização), químicos (trocas, fixação, precipitação, hidrólise) e físicos (lixiviação, volatilização). Todos esses processos afetam negativamente a disponibilidade do nutriente para o cultivo, podendo resultar em baixa produtividade da cultura e, principalmente, o risco de contaminação ambiental, havendo implicações na sustentabilidade dos agrossistemas (VILLALBA et al., 2014). Sabe-se que os agroquímicos nitrogenados estão entre os mais utilizados e são os que causam maior impacto ambiental (BIGATON, 2017), por exemplo, contaminação das águas dos rios, lagos e lençóis subterrâneos (CABEZAS et al., 2000).

Por causa dos impactos ambientais causados pela utilização desnecessária ou incorreta de adubos nitrogenados na produção da soja, cada vez mais a sociedade global – por intermédio do consumo racional, tratados internacionais e implementações de políticas públicas – tem se conscientizado da importância da produção sustentável, na qual seja cada vez menos agressivo ao meio ambiente (CABEZAS et al., 2000).

Além disso, para Hungria et al. (2001, p. 3), “a cultura da soja no Brasil seria inviabilizada economicamente se os produtores tivessem que aplicar todo o nitrogênio necessário para suprir as demandas da planta”, pois a adubação nitrogenada possui custo elevado, cerca de US\$ 15 bilhões para atender a produção anual de soja no país (HUNGRIA; MENDES, 2015).

A alta competitividade mundial, em virtude da globalização da economia, entre

os setores produtivos, principalmente no setor primário, faz com que os agricultores enfrentem desafios para maximizar a eficiência produtiva, seja por meio do aumento da produtividade ou da redução dos custos da produção, para não serem excluídos do cenário global (HUNGRIA et al., 2001).

Neste cenário, o uso de biofertilizantes tem se demonstrado uma alternativa importante e de baixo custo para atingir uma produção agrícola sustentável (HUNGRIA et al., 2020). Além disso, de acordo com Shahwar et al. (2023), Telles et al. (2023), uma agricultura mais sustentável pode ser atendida em parte pelos microrganismos do solo, que também desempenham um papel fundamental nos serviços econômicos e ecossistêmicos, contribuindo com biodiversidade do solo e no manejo coordenado de nutrientes. Sendo, provavelmente, a contribuição mais significativa proveniente da fixação biológica do nitrogênio (FBN) (HUNGRIA; MENDES, 2015).

Para Hungria et al. (2001), a fixação biológica do nitrogênio na cultura da soja tem se destacado pela capacidade de aumentar a produtividade sem aumentar excessivamente os custos, ou seja, representa a melhor relação custo/benefício para o produtor, além de ter impacto ambiental mínimo. Hungria et al. (2020) e Nogueira et al. (2018), salientam que a fixação biológica do nitrogênio representa um dos pilares de sustentabilidade no cultivo de soja no Brasil, apresentando um grande benefício para o produtor e para o meio ambiente, visto que, permite dispensar o uso de agroquímicos nitrogenados na cultura, e ainda, aumenta a competitividade do produto no mercado internacional.

O processo de fixação biológica do nitrogênio ocorre através da simbiose¹ entre bactérias diazotróficas e a planta, sendo que essas bactérias se associam ao sistema radicular de leguminosas e oleaginosas, formando nódulos radiculares, onde elas se abrigam e recebem da planta hospedeira proteção e alimento. Em compensação, as bactérias capturam o nitrogênio atmosférico (N_2) e, por intermédio da enzima bacteriana nitrogenase, o N_2 é reduzido a amônia (NH_3), que é transformada em compostos nitrogenados e exportados para a planta hospedeira, ou seja, elas fornecem o nitrogênio, em forma assimilável, necessário para o

¹Simbiose é um tipo de associação entre dois organismos em que ambos se beneficiam (ALCÂNTARA et al., 2017, p. 96).

desenvolvimento da planta (NOGUEIRA et al., 2018).

No cultivo da soja, as bactérias capazes de nodular as raízes da planta pertencem ao gênero *Bradyrhizobium*, principalmente nas espécies *Bradyrhizobium japonicum* e *Bradyrhizobium elkanii*; e, geralmente são denominadas pelos termos genéricos: rizóbio ou *bradyrhizobium diazoefficiens* (HUNGRIA et al., 2001).

Alcântara et al. (2017) afirmam que a fixação biológica do nitrogênio está sujeita a fatores bióticos (ligados aos organismos vivos) e abióticos (fatores de solo e clima). Em relação aos fatores abióticos, ela é afetada principalmente pela acidez do solo, temperatura, fertilidade do solo e umidade. A acidez do solo afeta principalmente os aspectos nutricionais, como menores teores de fósforo, cálcio e magnésio, e excessivos teores de alumínio e manganês. Temperaturas elevadas do solo limitam a fixação biológica do nitrogênio, dado que, elas afetam praticamente todas as etapas de crescimento do rizóbio e das plantas hospedeiras. Já a escassez hídrica, além de prejudicar o desenvolvimento das plantas, influencia a atividade fisiológica dos rizóbios e sua sobrevivência. E por fim, o excesso de umidade também impede a nodulação e a fixação biológica do nitrogênio, pois afeta a atividade das enzimas encarregadas pela redução do nitrato e a assimilação de amônia.

É imprescindível destacar que o processo de fixação biológica do nitrogênio se dá no interior dos nódulos radiculares da planta. Esses nódulos são estruturas especializadas que alojam os rizóbios e são formados mediante um complexo processo, que provocam o crescimento de células específicas na planta hospedeira, além de englobar diversas outras etapas e mudanças (fisiológicas e morfológicas), tanto na planta quanto na bactéria (ALCÂNTARA et al., 2017). E, ainda de acordo com Hungria et al. (2001) os nódulos apresentam, em seu interior, coloração rósea intensa quando estão em plena atividade, isto deve-se ao fato de atividade da leghemoglobina, cuja função é a mesma da hemoglobina do sangue humano, isto é, o transporte do oxigênio, no qual é primordial às funções vitais desses microrganismos aeróbios.

Conforme Hungria et al. (2001) os solos brasileiros, primitivamente, são desprovidos de bactérias do gênero *Bradyrhizobium* capazes de nodular de modo eficaz a soja. No entanto, grande parte dos solos cultivados já foram inoculados²

²Inoculação é uma prática segundo a qual bactérias fixadoras de nitrogênio selecionadas pela

anteriormente e apresentam uma população elevada de estirpes, estimada em torno de 10^3 a 10^5 células/g de solo. Mendes et al. (2014) recomendam as estirpes bacterianas SEMIA 5079 e SEMIA 5080, da espécie *Bradyrhizobium japonicum*, e SEMIA 587 e SEMIA 5019, pertencentes à espécie *Bradyrhizobium elkanii* para inoculação da soja no Brasil.

Alcântara et al. (2017) citam algumas vantagens da inoculação para as culturas e que estas vantagens estão intimamente relacionadas aos benefícios da fixação biológica do nitrogênio, sendo elas: a) utilização biológica do nitrogênio, substituindo gastos com a aquisição de adubos nitrogenados; b) promoção do crescimento da planta, que origina maior produção das culturas; e c) melhoria das condições do solo, pelo aumento da incorporação da massa verde, oriunda de uma maior produção da cultura. Em geral, a inoculação contribui com a diminuição dos custos de produção, com o aumento do rendimento da cultura e com os ganhos econômicos para a agropecuária brasileira e para o setor produtivo, além de auxiliar na preservação sustentável do meio ambiente, uma vez que o nitrogênio mineral, em quantidade acentuada, é um poluente ambiental.

Hungria et al. (2001) ressaltam que os fertilizantes nitrogenados são assimilados com maior rapidez pelas plantas, visto que, está em uma forma prontamente disponível enquanto na fixação biológica do nitrogênio a planta necessita investir um capital energético ínfimo na formação dos nódulos. Porém, esse gasto é recuperado brevemente no início do ciclo vegetativo, resultando em grande retorno nutricional às plantas. Ademais, Hungria et al. (2001) advertem que a utilização de fertilizantes nitrogenados atrapalham a fixação biológica do N_2 , mesmo em uma dose inicial, não resultando em benefícios ao rendimento de grãos.

2.2 Análise de imagens digitais

Os humanos são dotados de capacidades sensoriais que propiciam adequada percepção do ambiente que os cerca, dentre esses sentidos, um dos mais importantes

pesquisa são adicionadas às sementes das plantas no momento da semeadura. A inoculação é feita com um produto chamado inoculante, que não polui o solo, fornece nitrogênio para as plantas e é muito mais barato do que o adubo químico nitrogenado (ALCÂNTARA et al., 2017, p. 99).

e avançados é a visão. A visão abrange várias funções complexas, por exemplo, localização, detecção, reconhecimento e interpretação de objetos no ambiente (GONZALEZ; WOODS, 2010; PEDRINI; SCHWARTZ, 2008). Costa e Cesar (2009) ressaltam que a aquisição e análise da informação visual produzida pela interação entre a luz e os objetos do mundo têm representado meios poderosos pelos quais humanos e animais podem aprender de forma rápida e eficiente sobre seus ambientes circundantes. Logo, não é de se surpreender que a capacidade visual exerçam papel fundamental na percepção humana (GONZALEZ; WOODS, 2010).

Conforme Gonzalez e Woods (2010), uma imagem pode ser definida como uma função bidimensional de intensidade luminosa, denotada por $f(x,y)$, no qual x e y são coordenadas espaciais. O valor ou amplitude de f nas coordenadas espaciais (x,y) fornece a intensidade ou nível de cinza da imagem naquele ponto. Uma imagem é dita como digital, quando x,y e os valores de intensidade de f são quantidades finitas e discretas. Logo, uma imagem digital é composta por um número finito de elementos, denominados pixels, em que cada um possui localização e valor específico. A Figura 1 apresenta um exemplo de imagem digital, bem como a convenção do sistema de coordenadas adotada para sua representação, no qual a origem da imagem, ou seja, o ponto $(0,0)$, está localizado no canto superior esquerdo da imagem.

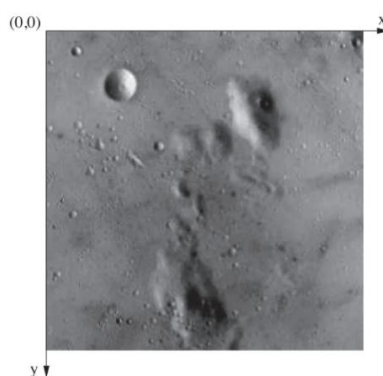


Figura 1: Convenção do sistema de coordenadas para representação de imagens digitais.

Fonte: Pedrini e Schwartz (2008)

Uma imagem digital pode ser obtida por um processo conhecido como digitalização, no qual envolve duas etapas, a amostragem e a quantização (PEDRINI; SCHWARTZ, 2008). Para Gonzalez e Woods (2010), a ideia primordial por trás da

amostragem e da quantização consiste em converter uma imagem contínua f em um formato digital, ou seja, a amostragem é responsável em discretizar o domínio de definição da imagem nas direções x e y e a quantização consiste na escolha de um número inteiro L de níveis de cinza possíveis para cada ponto da imagem. A imagem digital pode ser representada por uma matriz numérica $M \times N$, como definida na equação 1.

$$f(x,y) \approx \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix} \quad (1)$$

Cada elemento $f(x,y)$ dessa matriz é nomeado de pixel, com $0 \leq x \leq M-1$ e $0 \leq y \leq N-1$. Além disso, cada pixel possui um valor L associado, no qual $L_{min} \leq f(x,y) \leq L_{max}$, tal que o intervalo $[L_{min}, L_{max}]$ é intitulado como escala de cinza (PEDRINI; SCHWARTZ, 2008).

O valor de cada pixel representa a intensidade/profundidade ou a cor da imagem naquele ponto e dependendo do valor as imagens podem ser categorizadas em três tipos: binárias, tons de cinza e coloridas. Uma imagem binária é uma imagem digital, na qual os valores de cada pixel são representados apenas com dígitos binários, ou seja, 0 ou 1, representando, respectivamente, a cor preta e a cor branca (LOPES, 2003). A binarização de imagens é um processo no qual uma imagem em tons de cinza ou colorida é convertida em uma imagem binária, composta apenas por pixels pretos e brancos. Esse processo é baseado na definição de um limite de intensidade de pixel, chamado de limiar, acima do qual os pixels são considerados brancos e abaixo dos quais são considerados pretos (LOPES; CONSULARO, 2005; LOPES; CONSULAR, 2002). A segunda categoria é denominada de imagens em tons de cinza, em que o valor de cada pixel pode ser representado por meio de vários tons de cinza. O número padrão usado é 256 de tons ou níveis de cinza, pois podem ser representados usando números binários de 8 bits, e, definidos por potência de 2, ou seja, $L = 2^b$, no qual L é o número de níveis de cinza da imagem e b é a profundidade da imagem. Isso significa que cada pixel pode ter associado um valor

de cinza entre 0 e 255, representando, respectivamente, a cor preta e a cor branca. A terceira categoria é chamada de imagens coloridas, na qual o valor de cada pixel pode ser representado por meio de vários tons de cor. Todas as cores são representadas como uma combinação de três cores primárias, vermelha (R, *red*), verde (G, *green*) e azul (B, *blue*), que são chamadas de canais de cores em uma imagem. Portanto, uma imagem colorida possui três canais, enquanto uma imagem em tons de cinza possui um único canal (PAREKH, 2021). É importante evidenciar que existem outros modelos de cores além do RGB, por exemplo, modelo CMY (*cyan, magenta, yellow* – ciano, magenta, amarelo), modelo CMYK (*cyan, magenta, yellow, black* – ciano, magenta, amarelo, preto), modelo HSI (*hue, saturation, intensity* – matiz, saturação, intensidade), entre outros (GONZALEZ; WOODS, 2010).

A quantidade de pixels presentes em uma imagem define a sua resolução, quanto maior o número de pixels, melhor é a sua qualidade. As resoluções de imagens são descritas, por exemplo, como 320×240 , 640×480 , 800×600 , 1024×768 e assim por diante. Isto significa que, por exemplo, uma imagem com resolução de 1024×768 possui 1024 pixels no eixo x (colunas), 768 pixels no eixo y (linhas) e 786.432 pixels no total (SINGH, 2019).

Costa e Cesar (2009) afirmam que algumas dificuldades no processo de análise de informações visuais derivam de ruídos, oclusões e distorções, bem como do fato de que a formação da imagem envolve o mapeamento de um espaço tridimensional para um bidimensional, o que implica na perda de informações. Portanto, alguns cuidados são requeridos na digitalização de imagens para que seja minimizada a perda de informação no processo de amostragem. Um desses cuidados consiste na escolha correta do espaçamento entre as amostras tomadas da imagem contínua, isto é, o objetivo é encontrar condições de amostragem, no qual a partir de um conjunto de valores amostrados a imagem contínua possa ser completamente recuperada.

2.2.1 Processamento de imagens digitais e visão computacional

De acordo com Gonzalez e Woods (2010) não existe um consenso na literatura de onde termina a área de processamento de imagens digitais e inicia a

área de visão computacional.

Frequentemente na literatura o processamento de imagens é dito como uma disciplina na qual tanto a entrada quanto a saída de um processo são imagens (BHUYAN, 2019), porém Gonzalez e Woods (2010), afirmam que esta definição é muito restritiva, pois uma tarefa corriqueira como calcular a intensidade média de uma imagem (que resulta em um único número) não seria considerada, neste contexto, como uma operação de processamento de imagens. Já para Pedrini e Schwartz (2008, p. 1), o processamento de imagens corresponde a “um conjunto de técnicas para capturar, representar e transformar imagens com o auxílio de computador”, no qual a aplicação dessas técnicas permite extrair e identificar informações das imagens, bem como melhorar a qualidade visual de certos aspectos estruturais, para facilitar a percepção humana e a interpretação automática por meio de máquinas.

O campo da visão computacional pertence a um ramo da inteligência artificial, no qual o objetivo é emular a visão humana através de computadores, para aprender, ser capaz de fazer inferências e agir com base em informações visuais (GONZALEZ; WOODS, 2010). Bhuyan (2019) discute que a visão computacional visa permitir que os computadores processem e identifiquem imagens e vídeos da mesma forma que a visão humana, em que o objetivo principal é descrever uma cena do mundo real em uma ou mais imagens e identificar e reconstruir suas propriedades, como características de cor, informações de forma, características de textura, iluminação da cena, entre outros.

Segundo Gonzalez e Woods (2010), um sistema de processamento de imagens digitais é composto por um conjunto de fases: aquisição, pré-processamento, segmentação, extração de características, e, reconhecimento e interpretação, que são capazes de produzir um resultado a partir do domínio do problema. A Figura 2 apresenta a relação simplificada entre as fases de um sistema de processamento de imagens digitais.

A fase de aquisição consiste na captura da imagem por meio de um dispositivo ou sensor (câmeras de vídeo, tomógrafos médicos, satélites, scanners, etc.) e a converte em uma representação digital. A fase de pré-processamento propõe-se melhorar a qualidade da imagem – por meio de técnicas (atenuação de ruído,

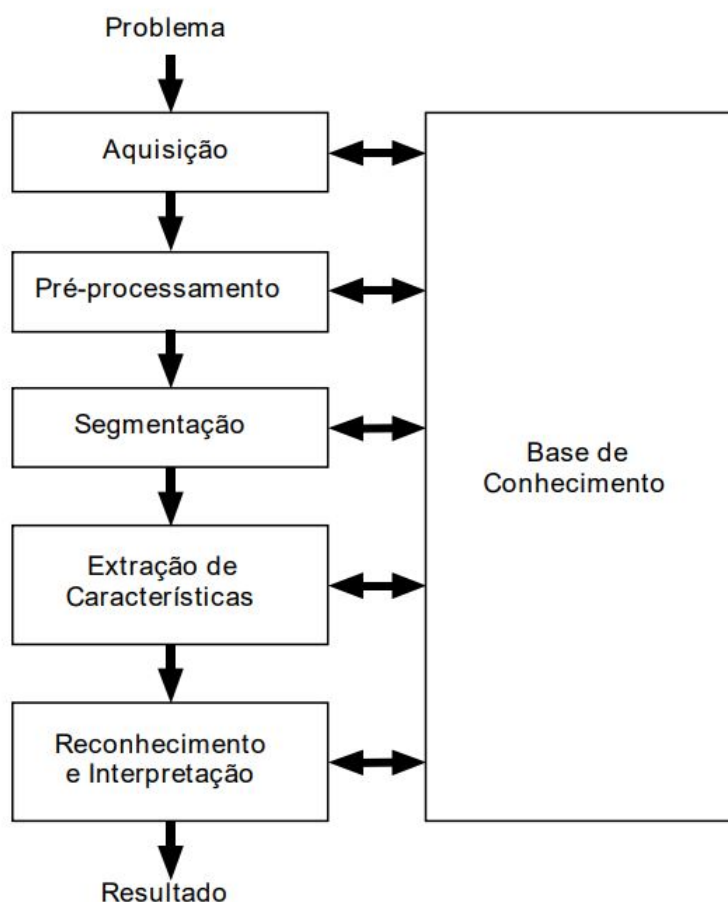


Figura 2: Fases de um sistema de processamento de imagens digitais simplificado.

Fonte: Adaptado de Gonzalez e Woods (2010)

correção de contraste, brilho, suavização, entre outras) – resultante do processo de aquisição, pois a mesma pode apresentar imperfeições tais como: presença de pixels ruidosos, contraste e/ou brilho inadequado decorrentes das condições de iluminação ou características do próprio dispositivo de captura. A fase de segmentação visa a extração e identificação de unidades significativas da imagem, ou seja, dos objetos de interesse que a compõem. Geralmente, ela se baseia na detecção de descontinuidades (bordas) ou de similaridades (regiões) na imagem. Na fase de extração de características, as características das imagens, resultantes da fase de segmentação, são extraídas por meio de descritores que permitem caracterizar, geralmente, por atributos numéricos, propriedades dessas imagens de modo que possam ser utilizados para diferenciar classes de objetos. A última fase, reconhecimento e interpretação, consiste na atribuição de um rótulo a um objeto com base em suas características (reconhecimento) e a atribuição de um significado a

um conjunto de objetos reconhecidos (interpretação). Todas as atividades das fases supracitadas pressupõem a existência de um conhecimento sobre o problema a ser resolvido, armazenado em uma base de conhecimento que pode ser utilizada para orientar a comunicação entre as fases (FILHO; NETO, 1999).

De acordo com Pedrini e Schwartz (2008), uma das tarefas mais significativas em aplicações envolvendo um sistema de visão computacional é a interpretação ou a análise de imagens, no qual o objetivo principal é conseguir uma descrição que contenha informações consideráveis para diferenciar objetos de interesse, de maneira confiável e com a mínima intervenção humana. Entretanto, é necessário salientar que as etapas que envolvem tanto o processamento de imagens, quanto visão computacional, muitas vezes é realizada por um operador humano, no qual detém o conhecimento ou a experiência sobre o domínio da aplicação.

A crescente evolução da tecnologia digital, associado ao desenvolvimento de novos algoritmos, tem surgido um número cada vez maior de aplicações. Muitas áreas, como medicina, biologia, automação industrial, agronomia, astronomia, microscopia, artes, área militar, segurança, vigilância, entre outros, tem se beneficiado de técnicas de processamento de imagens e visão computacional para solucionar problemas. Por exemplo, no campo da medicina, o uso de imagens capturadas por raios X, tomografia computadorizada, ressonância magnética e ultra-sonografia, aliadas à análise e interpretação dessas imagens – promovidas pelas áreas de processamento de imagens e visão computacional – tem auxiliado com maior precisão e rapidez diagnósticos médicos, no qual permite a identificação de lesões ou regiões atingidas por câncer, permitindo assim melhor planejamento dos tratamentos e das cirurgias (PEDRINI; SCHWARTZ, 2008).

O avanço tecnológico também tem impulsionado a adoção de tecnologias digitais na agricultura. A agricultura de precisão, como é conhecida, utiliza dispositivos eletrônicos, sensores e tecnologia de automação aliados a avançadas ferramentas como Internet das Coisas (IoT), mineração de dados, aprendizado de máquina, aprendizado profundo, visão computacional, ciência de dados, entre outras. Essas tecnologias atuam como um sistema agrícola integrado com foco em informação e produção, no qual possui o objetivo de minimizar o esforço de trabalho, auxiliar

na tomada de decisões, maximizar a lucratividade, aumentar a eficiência e a produtividade da cultura melhorando a qualidade da produção ao mesmo tempo em que minimiza os impactos ambientais e promove a saúde humana (DURAI; SHAMILI, 2022; COULIBALY et al., 2022; MESHRAM et al., 2021). De acordo com Lu e Young (2020), as tecnologias de visão computacional que utilizam imagens digitais para interpretar e entender o mundo são capazes de fornecer informações precisas e específicas sobre as culturas e seus ambientes. Além disso, o grupo de pesquisas em Bioinformática e Reconhecimento de Padrões têm desenvolvido diversos estudos e metodologias relevantes (PACANHELA et al., 2022; PEREIRA et al., 2021; de Lima et al., 2019; PEREIRA et al., 2019a; PEREIRA et al., 2019b; PEREIRA et al., 2019c; de Lima et al., 2015; BRILHADOR et al., 2015; BRILHADOR et al., 2013) que contribuem com o avanço nas pesquisas e aplicações na agricultura.

2.2.2 Extração de características

Um dos problemas clássicos no desenvolvimento de um sistema de visão computacional consiste na seleção de um conjunto de características extraídas do objeto de interesse para o propósito de classificação (PEDRINI; SCHWARTZ, 2008). Para Barelli (2018), a extração de características de um objeto ou região de interesse de uma imagem representa um processo fundamental de qualquer sistema de visão computacional, no qual consiste em obter informações que tornam possível classificar ou identificar um objeto.

De acordo com Choras (2007), uma característica pode ser definida como uma função de uma ou mais medidas, na qual quantifica alguma propriedade de um objeto. Segundo Kunaver e Tasic (2005), as características de um objeto ou região de interesse podem ser classificadas em categorias, sendo as principais: características espectrais, geométricas e de textura.

Conforme Kunaver e Tasic (2005), nas características espectrais a imagem é processada como um matriz de pixels, em que cada pixel representa o brilho e o nível de cor, e, a conexão entre pixels é irrelevante, pois nessa categoria não há preocupação com formas e contornos na imagem. O tipo mais básico de características espectrais é o espaço de cores que baseia-se na distribuição de cores

na imagem. Patel e Gamit (2016), em conformidade com Kunaver e Tasic (2005), afirmam que a cor é um recurso visual muito básico usado em recuperação de imagens com base na semelhança de cores e isto é comparativamente robusto, pois não depende do tamanho da imagem, orientação e mudanças de escala. Para esta categoria, podemos destacar o descritor de imagem *Global Color Histogram* (GCH) proposto por Stricker e Orengo (1995).

Dentre as características geométricas, a forma é uma importante característica visual e é uma das características primitivas para descrição do conteúdo da imagem. A recuperação de imagem baseada na forma consiste na medição da similaridade entre as formas representadas por suas características (CHORAS, 2007). Os descritores de forma podem ser classificados em duas categorias principais: técnicas baseadas em fronteiras (contornos) e técnicas baseadas em regiões (COSTA; CESAR, 2009). Uso de técnicas baseadas em região considera toda a área de um objeto para descrição da forma, enquanto as técnicas baseadas em contorno usam apenas as informações presentes no contorno de um objeto (CHORAS, 2007). De acordo com Costa e Cesar (2009), o contorno de um objeto pode transmitir informações suficientes para o reconhecimento do objeto original e ainda destacam que uma das maneiras mais comuns de descrever formas é medindo características específicas como área, perímetro, número de cantos, medidas baseadas em curvatura, entre outros. Além disso, os descritores de forma também podem ser divididos em dois grupos: no primeiro estão características que são invariáveis à translação, rotação e dimensionamento, enquanto no outro grupo permanecem características que são sensíveis à essas propriedades (KUNAVÉR; TASIC, 2005). Momentos Invariantes de Hu (HU, 1962) e Zernike (TEAGUE, 1980) são alguns exemplos de descritores de forma.

As características de textura referem-se a padrões visuais que propriedades de homogeneidade não resultam da presença de uma única cor ou intensidade. Essas características contêm informações importantes sobre a estrutura de disposição das superfícies e sua relação com o meio ambiente (KUNAVÉR; TASIC, 2005). Para Gonzalez e Woods (2010) e Patel e Gamit (2016), os descritores de textura fornecem medidas de propriedades como suavidade, direcionalidade, rugosidade, regularidade

e aspereza. Haralick (HARALICK et al., 1973) e *Local Binary Patterns* (LBP) (OJALA et al., 2000) são alguns exemplos de descritores de textura.

Com o surgimento das Redes Neurais Convolucionais (*Convolutional Neural Networks* - CNNs) no final dos anos 80 (LECUN et al., 1989) e a sua popularização em 2012, devido ao surgimento de grandes bancos de dados públicos, por exemplo, *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) (KRIZHEVSKY et al., 2012). As CNNs passaram a atuar também como extratores de características profundas (NOBREGA et al., 2018). VGG16 e VGG19 (SIMONYAN; ZISSERMAN, 2015), DenseNet121 (HUANG et al., 2017), InceptionV3 (SZEGEDY et al., 2016), ResNet101 (HE et al., 2016a) e Xception (CHOLLET, 2017) são alguns exemplos de arquiteturas profundas.

Conforme Pedrini e Schwartz (2008), a descrição de um objeto depende do descritor adotado, e, este deve permitir a caracterização sem ambiguidades e com a quantidade reduzida de características ou medidas extraídas do objeto.

2.3 Aprendizado de máquina

A inteligência artificial é compreendida por diversas habilidades desejadas aos sistemas, e o aprendizado de máquina é uma das mais utilizadas e de extrema importância nos dias atuais, devido a sua capacidade de resolução de problemas complexos ou de grande volume de dados (LENZ et al., 2020; COPPIN, 2013). Geron (2019, p. 4) define aprendizado de máquina como “a ciência (e a arte) da programação de computadores para que eles possam aprender com os dados”.

Para Lenz et al. (2020, p. 13), “o processo de aprendizado consiste no treinamento de um algoritmo ou modelo para que possa criar regras que relacionam os dados de entrada (atributos previsores) com os dados de saída (atributo alvo)”, no qual permite realizar tarefas como classificação, previsão e agrupamento de dados. De acordo com Faceli et al. (2021), as aplicações baseadas em aprendizado de máquina usam heurísticas que procuram por modelos aptos de representar o conhecimento presente em um conjunto de dados. De modo geral, os conjuntos de dados possuem estruturas em formato tabular, uma matriz atributo-valor, em que cada linha representa

um objeto (instância) e cada coluna representa um atributo (característica). Sendo assim, o aprendizado de máquina busca por semelhanças entre as características de instâncias diferentes de determinado conjunto de dados (LENZ et al., 2020).

Conforme Faceli et al. (2021), as tarefas de aprendizado de máquina podem ser classificadas em categorias, sendo elas: aprendizado supervisionado, semissupervisionado, por reforço, ativo e não supervisionado, conforme pode ser observado na Figura 3.

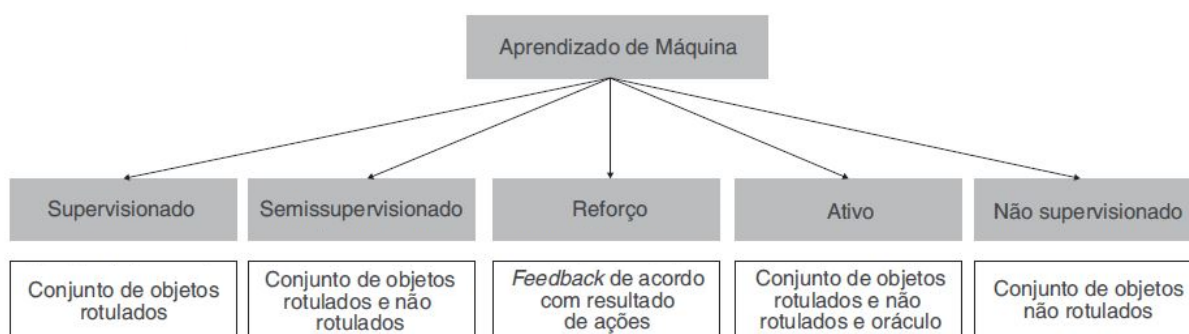


Figura 3: Categorias de aprendizado de máquina.

Fonte: Adaptado de Faceli et al. (2021)

De acordo com Theodoridis (2020), o aprendizado supervisionado refere-se ao tipo de aprendizado de máquina em que todos os dados disponíveis foram rotulados, isto é, os dados de treinamento fornecidos ao algoritmo incluem as soluções desejadas, denominadas de rótulos (GERON, 2019). Segundo Hurbans (2020), o aprendizado supervisionado pode ser dividido em duas categorias: classificação e regressão. A classificação visa prever categorias de exemplos com base em suas características, isto é, consiste em atribuir um rótulo para cada instância de teste dentre um conjunto de classes possíveis (HURBANS, 2020; THEODORIDIS, 2020). A regressão procura determinar uma relação entre variáveis dependentes e independentes através de métodos estatísticos, no qual o resultado, isto é, a variável dependente é um valor numérico (IZBICKI; SANTOS, 2020). *K-Nearest Neighbors* (KNN), *Random Forest* (RF) e *Support Vector Machine* (SVM) são alguns exemplos de modelos de aprendizado supervisionado.

O aprendizado semissupervisionado está entre o aprendizado supervisionado e o não supervisionado, no qual existem dados rotulados, mas não o suficiente para

obter uma boa estimativa (THEODORIDIS, 2020). Neste contexto, o aprendizado semissupervisionado busca aumentar o número de objetos rotulados, no qual o modelo considera os dados rotulados disponíveis para rotular dados não rotulados. Logo, esses dados podem ser adicionados ao conjunto de treinamento e o processo se repete até que os dados que satisfaçam uma dada condição (FACELI et al., 2021).

O aprendizado por reforço é inspirado na psicologia comportamental e opera recompensando ou punindo um algoritmo baseado em suas ações em um ambiente (HURBANS, 2020). Nesse contexto, o agente observa o ambiente, seleciona, executa ações e obtém recompensas em troca — ou penalidades na forma de recompensas negativas. Logo, o modelo deve aprender por si só a melhor estratégia para obter o maior número de recompensas ao longo do tempo (GERON, 2019).

Segundo Faceli et al. (2021), uma estratégia de aprendizado ativo consiste em selecionar iterativamente os dados a serem rotulados e o rótulo a ser atribuído a cada um deles, podendo consultar um oráculo. Deste modo, apenas os dados que apresentam uma característica faltante nos dados rotulados, com a probabilidade de melhorar a qualidade do modelo induzido, são selecionados. Já os dados que não foram selecionados são descartados.

No aprendizado não supervisionado os dados de treinamento não são rotulados, isto é, o modelo tenta aprender sem um “professor” (GERON, 2019). Para Theodoridis (2020), o aprendizado não supervisionado apenas as variáveis de entrada são fornecidas, em que nenhuma informação de saída ou rótulo está disponível. Sendo assim, o seu objetivo é desvendar a estrutura subjacente a um determinado conjunto de dados. Um dos tipos mais importantes de aprendizado não supervisionado é o agrupamento (*clustering*).

De acordo com Lenz et al. (2020), o aprendizado de máquina pressupõe o acúmulo de conhecimento, e, conseqüentemente, pode-se extrair significado a partir de um determinado conjunto de dados. Para isto, este processo pode ser decomposto em etapas: coleta de dados, preparação dos dados, seleção do modelo, treinamento, avaliação, ajuste de parâmetros e aplicação.

Na primeira etapa os dados devem ser selecionados de modo que o

quantidade e a qualidade das amostras sejam suficientes para extrair o conhecimento necessário. Os dados devem ser preparados, isto é, na etapa de preparação dos dados podem sofrer transformações de unidades, conversão de escala, discretização, normalização e até mesmo mudança de representação dos dados. Nesta fase é importante examinar o balanceamento dos dados, ou seja, se os dados coletados para classes distintas são equivalentes. Dados incompletos, inconsistentes, redundantes ou com ruídos devem ser tratados ou eliminados de modo que não influenciem o processo de treinamento. Além disso, os dados podem ser divididos em grupos distintos, de maneira que esses grupos possam ser utilizados na etapa de treinamento e de teste separadamente. Na etapa de seleção do modelo, de acordo com o objetivo, os dados e suas complexidades, devem ser observadas as vantagens oferecidas por cada um dos modelos, sendo alguns exemplos, regressões lineares, regressões logísticas, classificação, agrupamento, aprendizado profundo, entre outros. Com a escolha do modelo, na etapa de treinamento o intuito é aperfeiçoar o modelo a cada nova amostra de treinamento avaliada, sendo que o treinamento corresponde na inferência de hipóteses que mais se aproximam do resultado apropriado para a saída a partir de determinado conjunto de dados de entrada. A etapa de avaliação procura determinar se o modelo obtido após o treinamento possui precisão satisfatória na predição do alvo a partir do conjunto de dados de testes, inicialmente separado. A etapa de ajuste de parâmetros consiste na alteração dos hiperparâmetros do modelo, parâmetros esses que controlam o processo de aprendizado em si, de modo a obter um modelo mais adequado (FACELI et al., 2021; LENZ et al., 2020). No entanto, é necessário ter cautela ao alterar os hiperparâmetros do modelo, pois corre-se o risco de produzir um modelo de bom desempenho com os dados de treinamento, porém o modelo não lida muito bem com novos dados, este problema é conhecido como sobreajuste (*overfitting*) (GRUS, 2016). E por fim, na etapa de aplicação o modelo é utilizado para realizar previsões a partir de seu treinamento (LENZ et al., 2020).

2.4 Aprendizado profundo

De acordo com Ugail (2022) o aprendizado profundo é um subcampo do aprendizado de máquina, no qual permite que modelos computacionais aprendam

e representem dados de forma que imite como o cérebro humano percebe e compreende as informações. Neste sentido, o aprendizado profundo trata-se de um mecanismo pelo qual um algoritmo de máquina pode aprender por exemplos, obtendo uma configuração ótima para um modelo de modo que a saída desejada possa ser obtida a partir de um conjunto de dados de entrada.

O aprendizado profundo é utilizado para resolver tarefas práticas em diversas áreas, como visão computacional (imagem), processamento de linguagem natural (texto) e reconhecimento automático de fala (áudio) usando principalmente as redes neurais artificiais (RNAs) (TRASK, 2019).

De acordo com Carvalho (2023), RNAs são abordagens computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes. Essas técnicas têm a capacidade de adquirir conhecimento por meio da experiência. O sistema nervoso é composto por uma rede altamente complexa de células conhecidas como neurônios. Esses neurônios desempenham um papel fundamental na regulação das funções e comportamentos do corpo humano, bem como no processo de pensamento. Eles consistem em dendritos, que são terminais de entrada, um corpo celular central e axônios, que são extensões longas responsáveis pela transmissão de sinais para outras células, conforme pode ser observado na Figura 4.

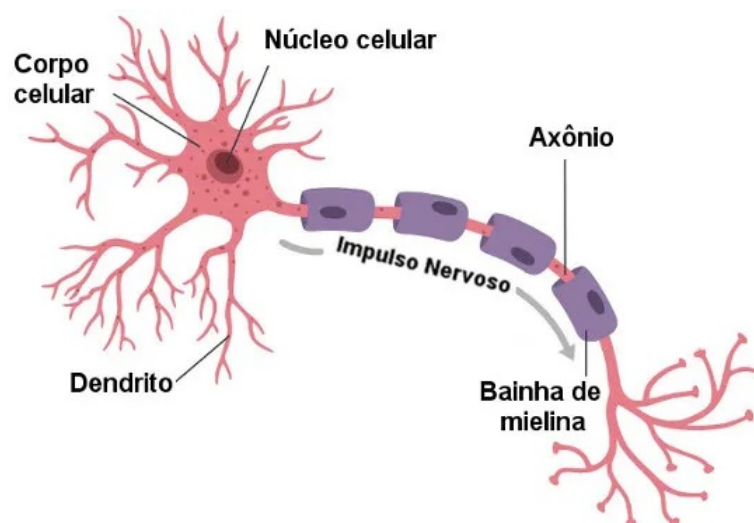


Figura 4: Esquema dos constituintes da célula neuronal.

Fonte: Santos (2023)

A comunicação entre os neurônios ocorre por meio das sinapses, regiões onde dois neurônios se conectam e permitem a transmissão de impulsos nervosos entre eles. Quando um neurônio A recebe impulsos em um determinado momento, esses impulsos são processados. Se o limiar de ação é alcançado, o neurônio A é ativado e libera substâncias neurotransmissoras que se propagam do corpo celular para o axônio. O axônio pode estar conectado aos dendritos de outros neurônios, como o neurônio B. O neurotransmissor pode modular a polaridade da membrana pós-sináptica, tanto diminuindo quanto aumentando, o que inibe ou excita a geração de pulsos no neurônio B (CARVALHO, 2023).

Na década de 1940, Warren McCulloch e Walter Pitts, pesquisadores do campo da neurociência computacional, propuseram o primeiro modelo computacional de aprendizado de máquina, inspirado nos neurônios cerebrais (MCCULLOCH; PITTS, 1943). Eles buscavam compreender como o cérebro humano é capaz de produzir padrões complexos por meio da interconexão de neurônios. Essa abordagem serviu como base para o modelo *perceptron* proposto posteriormente por Rosenblatt (1958).

Semelhantemente ao funcionamento de um neurônio humano, o modelo neural *perceptron* recebe uma série de sinais de entrada (x_1, x_2, \dots, x_m) , que podem ser excitatórios ou inibitórios. Esses sinais podem ser multiplicados pelos pesos correspondentes (w_1, w_2, \dots, w_m) , permitindo que o efeito de cada sinal seja ajustado conforme necessário. Em seguida, é realizada a soma ponderada dos sinais de entrada. Se essa soma ponderada atingir um limite escolhido – neste caso, a função de ativação binária – o modelo emitirá uma saída de 1; caso contrário, a saída será 0. Dessa forma, o modelo toma sua decisão de saída com base nos sinais de entrada (x_1, x_2, \dots, x_m) , realizando uma soma ponderada que resulta em uma saída binária, ou seja, 0 ou 1 (UGAIL, 2022).

A Equação 2 apresenta como a saída do neurônio pode ser descrita, bem como a Figura 5 ilustra o modelo matemático de um neurônio artificial:

$$y = f(u) = f\left(\sum_{j=1}^m w_j x_j\right) \quad (2)$$

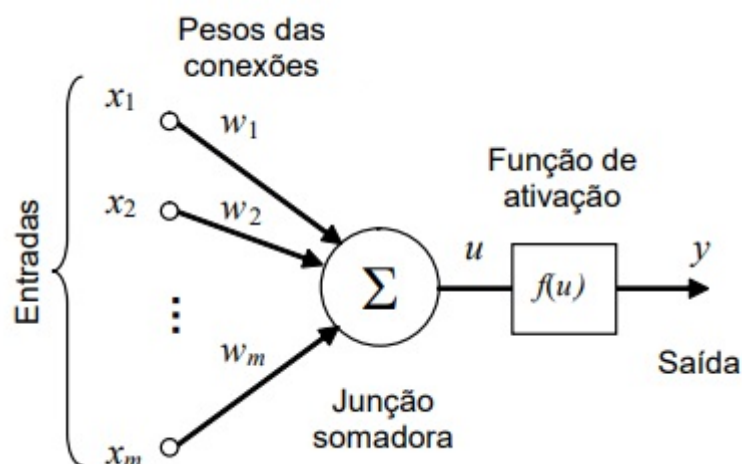


Figura 5: Modelo matemático de um neurônio artificial.

Fonte: Adaptado de Boccato (2023)

No entanto, um único neurônio é bastante limitado em sua capacidade, pois muitos problemas do mundo real exigem soluções mais complexas do que simples classificadores binários. Para lidar com essa limitação, a ideia de *perceptron* pode ser estendida para redes neurais artificiais, na qual vários neurônios podem ser interconectados em uma rede com múltiplas camadas, conforme pode ser observado na Figura 6. Neste sentido, uma RNA é, essencialmente, composta por uma coleção de *perceptrons* individuais interconectados, formando uma rede de neurônios, onde cada unidade *perceptron* na rede contém pesos, que são utilizados para controlar a influência entre si. Esses pesos são ajustados durante o processo de treinamento da rede, permitindo que ela aprenda e adapte seu comportamento para realizar tarefas específicas (UGAIL, 2022).

Além disso, com o avanço da tecnologia e o aumento da capacidade de processamento, as redes neurais artificiais têm se tornado cada vez mais sofisticadas, possibilitando a criação de modelos mais complexos, como as redes neurais convolucionais (CNNs) e as redes neurais recorrentes (RNNs).

De acordo com Anjos (2019), as redes neurais convolucionais também são redes inspiradas no funcionamento do cérebro humano, que utilizam primitivas básicas, como linhas, curvas e formas, para processar informações visuais. Essas redes ganharam o nome de “convolucionais” devido à sua camada mais importante, chamada camada de convolução. Ao realizar operações de convolução nas imagens

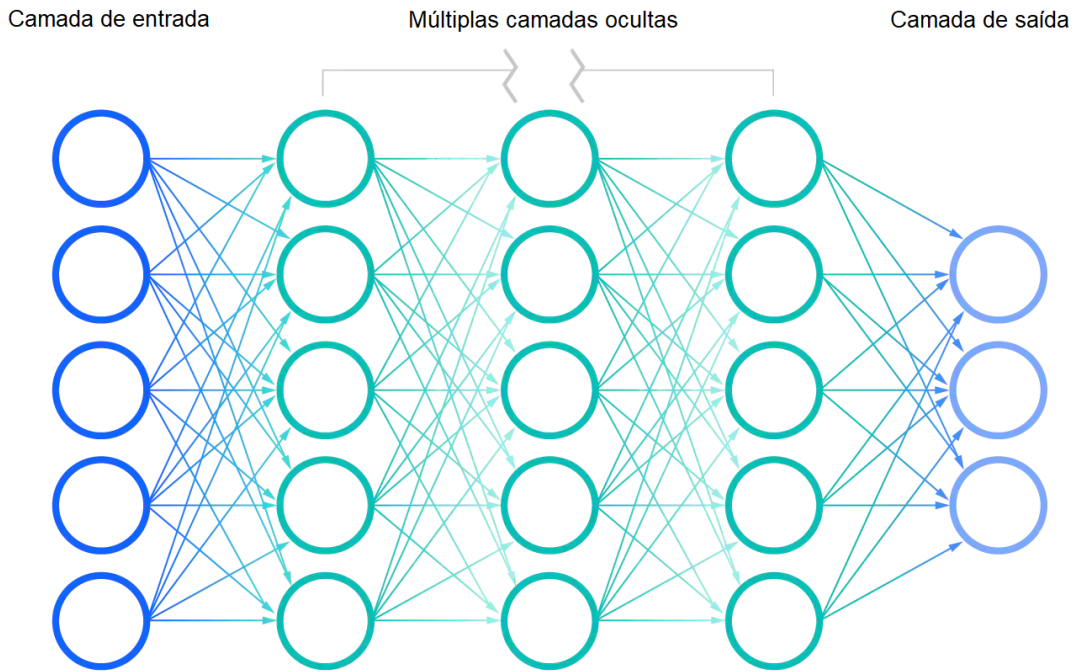


Figura 6: Representação de uma rede neural multicamadas com três camadas ocultas.
 Fonte: Adaptado de IBM (2023)

de entrada, as CNNs são capazes de extrair características e padrões relevantes para a tarefa de visão computacional, conforme pode ser observado na Figura 7. No entanto, as CNNs possuem outras camadas importantes além da camada de convolução, por exemplo, a camada de *pooling* ou *downsampling*, a camada de *flattening* e a camada totalmente conectada ou camada densa.

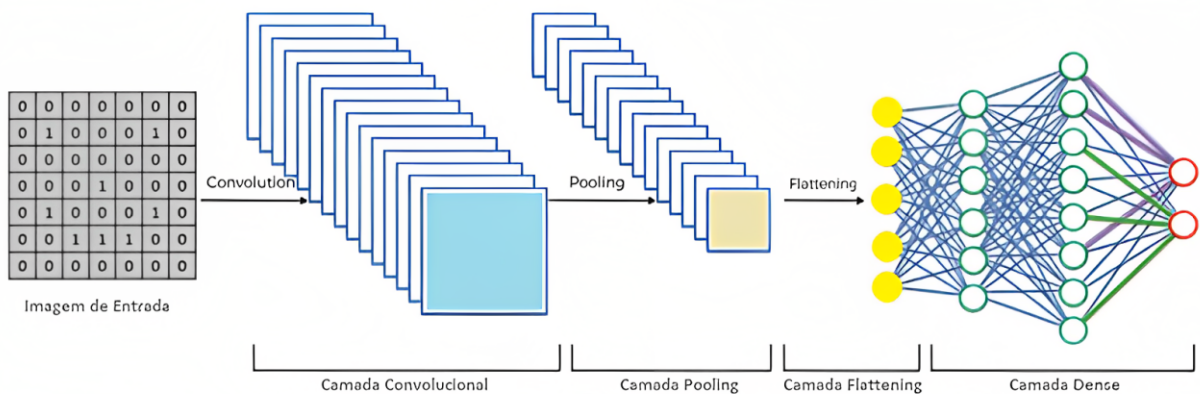


Figura 7: Representação de uma rede neural convolucional.
 Fonte: Viceri (2023)

Uma camada de convolução é composta por múltiplos filtros ou *kernels* que

realizam operações de convolução. Durante a convolução, o produto escalar entre a matriz de entrada e o filtro é calculado. O filtro itera sobre o vetor de entrada, realizando operações de multiplicação e soma para obter um valor resultante. A camada de *pooling* ou *downsampling* reduz a quantidade de informação espacial nas representações intermediárias. Isso é alcançado pela redução do tamanho espacial dos mapas de características, mantendo as características mais importantes e descartando o ruído ou detalhes menos relevantes. Após as características serem extraídas na camada de convolução e a dimensão reduzida pela camada de *pooling*, os dados multidimensionais são reduzidos a uma única dimensão, na camada de *flattening*, para serem processados pelas camadas totalmente conectadas. Diferentemente das camadas convolucionais e de *pooling*, a camada totalmente conectada tem conexões entre todos os neurônios da camada anterior e da camada seguinte. Na qual, a camada densa é responsável por combinar as características extraídas nas camadas anteriores e gerar as saídas desejadas, permitindo que a rede neural tome decisões finais e realize tarefas específicas utilizando pesos ajustáveis para aprender e adaptar seu comportamento durante o treinamento (KUMAR; RENUKA, 2023).

Nos últimos anos, houve avanços significativos na detecção de objetos por meio das redes neurais convolucionais. Uma das principais razões para isso é a capacidade dessas redes de serem treinadas de ponta a ponta, eliminando a necessidade de extração de características explícita (LIU et al., 2020; O'SHEA; NASH, 2015).

A detecção de objetos é uma tarefa fundamental no campo da visão computacional e processamento de imagens. Embora os seres humanos possam perceber e reconhecer objetos em uma imagem com facilidade, essa continua sendo uma tarefa complexa e desafiadora para os computadores. No entanto, as redes neurais convolucionais têm mostrado grande potencial nesse aspecto, permitindo o desenvolvimento de sistemas capazes de identificar e localizar objetos com precisão e eficiência (ARULPRAKASH; ARULDOSS, 2022). Faster R-CNN (REN et al., 2016), YOLO (REDMON et al., 2016), SSD (LIU et al., 2016) e RetinaNet (LIN et al., 2018) são alguns exemplos de redes neurais convolucionais para detecção de objetos em

imagens.

2.5 Trabalhos relacionados

A identificação de nódulos em raízes de soja tem sido alvo de poucos trabalhos na literatura. Uma abordagem foi proposta em 2019 para o processamento e a classificação automática do vigor de sementes de soja, por meio de uma perspectiva inteligente baseada em visão computacional (PEREIRA et al., 2019c). A metodologia utilizada no trabalho consiste no processamento e análise de imagens de sementes de soja provenientes do teste de tetrazólio, bem como o aprendizado e classificação automática do vigor das sementes. Para os experimentos, foram considerados seis conjuntos de dados (D1-D6), totalizando 3.851 amostras. Para o conjunto de dados D1, foram consideradas as características já extraídas e disponibilizadas publicamente (MICHALSKI; CHILAUSSKY, 1980), enquanto os conjuntos de dados D2-D6 foram consideradas as características extraídas pelos extratores *Border/Interior pixel Classification* (BIC) (STEHLING et al., 2002), GCH (STRICKER; ORENGO, 1995), *Local Color Histogram* (LCH) (SWAIN; BALLARD, 1991) e LBP (OJALA et al., 1996). Considerando as características extraídas, os modelos de aprendizado de máquina J48, KNN, *Multi-layer Perceptron* (MLP), RF, SVM e *Optimum-Path Forest* (OPF) foram utilizados no trabalho para classificação das amostras.

Em 2020, uma abordagem foi proposta por meio de métodos interativos e tradicionais de aprendizado de máquina para classificar sementes e mudas de soja de acordo com sua aparência e potencial fisiológico (MEDEIROS et al., 2020). Foram selecionadas 700 sementes de soja da safra 2019/2020 e a partir das sementes as mudas foram produzidas. As características das sementes e das mudas foram baseadas em cor e intensidade de pixel e descritores de borda e textura. Os modelos de classificação selecionados pelos autores foram: *Linear Discriminant Analysis* (LDA), RF e SVM, no qual utilizaram a técnica conhecida como validação cruzada (REFAEILZADEH et al., 2009) para avaliação dos modelos, sendo nove partes do conjunto de dados para treinamento e uma parte para teste com iteração de dez vezes.

Outro trabalho também desenvolvido em 2020 procurou caracterizar nódulo

de raiz de soja exposta ao silício, através de aprendizado de máquina baseado em imagens (CHUNG et al., 2020). Imagens 2D foram adotadas para a extração de características morfológicas das raízes, como área de superfície, comprimento e ângulo, bem como a quantidade e tamanho dos nódulos radiculares. A quantidade e o tamanho dos nódulos foram determinados pela detecção e segmentação baseada em aprendizado profundo. A arquitetura U-Net (RONNEBERGER et al., 2015) foi utilizada para a segmentação dos nódulos. Foram anotadas 35 imagens manualmente, através de polígonos desenhados no contorno de cada nódulo e realizado o treinamento da rede. Porém, mais dados de treinamento foram adicionados por meio de uma abordagem de anotação semiautomática, resultando em um total de 135 imagens para treinamento da rede profunda.

Pereira et al. (2021) propuseram uma abordagem de aprendizado ativo para a classificação do vigor de sementes de soja e um estudo comparativo entre as abordagens de aprendizagem ativa e supervisionada. A estratégia denominada Aprendizagem Ativa através da Diversidade, Incerteza e Representatividade (AA-DIR) consiste em selecionar (efetivamente e eficientemente) as amostras mais informativas para o treinamento do classificador, pois ela considera os critérios de diversidade, incerteza e representatividade. Este último, segundo os autores, permite agregar mais informações ao processo de aprendizagem, visto que, além do conhecimento do classificador ao longo das iterações, o conhecimento do especialista é explorado, no qual as anotações do especialista em iterações anteriores são consideradas para obter amostras mais representativas de cada classe. As amostras ativamente selecionadas pelos critérios (diversidade, incerteza e representatividade) juntamente com os rótulos verificados (confirmados ou corrigidos) pelo especialista são adicionados ao conjunto de treinamento. O processo de aprendizagem é iterativo de tal maneira que após cada iteração de feedback do especialista, o classificador é retreinado. Para realizar os experimentos, foram considerados quatro conjuntos de dados compostos por imagens de sementes de soja obtidas a partir do teste de tetrazólio, totalizando 4.708 amostras. Para cada conjunto de dados, as características foram obtidas por meio de diferentes extratores, os baseados em cor, BIC (STEHLING et al., 2002), GCH (STRICKER; ORENGO, 1995), LCH (SWAIN; BALLARD, 1991) e baseado em textura, LBP (OJALA et al., 1996). No trabalho foi adotado o

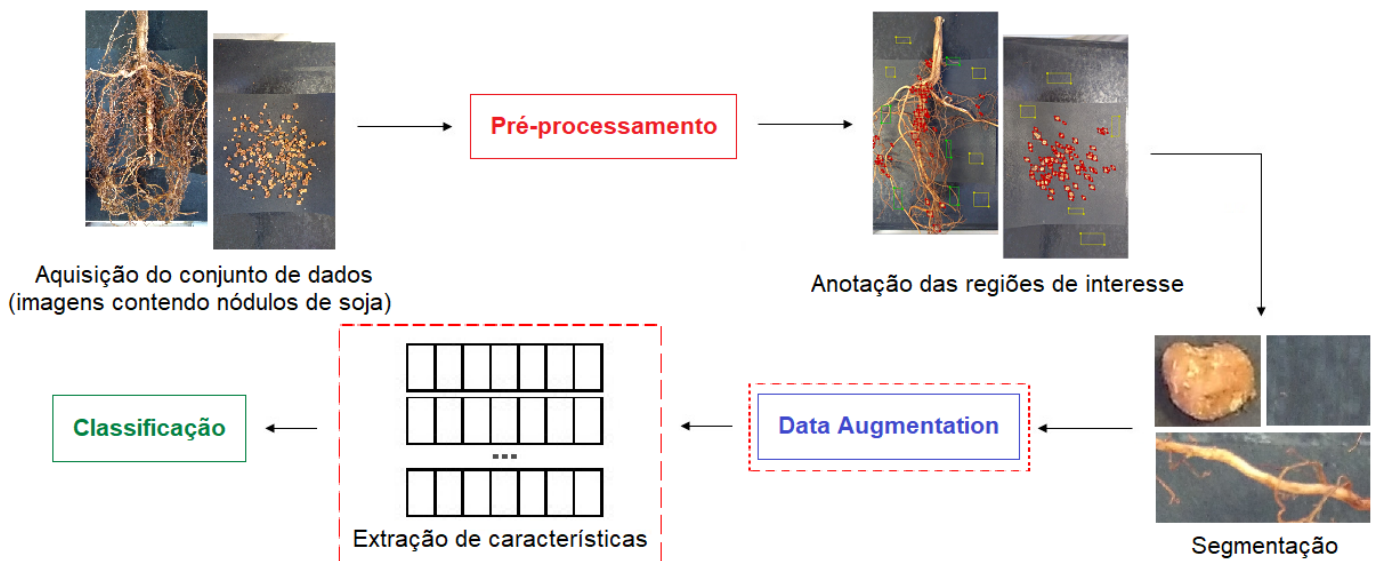
classificador OPF. Os experimentos comprovaram que as abordagens de aprendizagem ativa mostraram um forte desempenho sobre as de aprendizagem supervisionada simples.

Em 2021 também foi desenvolvido uma abordagem para detectar automaticamente nódulos em imagens de raízes de soja através de aprendizado de máquina, fornecendo a quantidade de nódulos, tamanho dos nódulos e distribuições de nódulos ao longo da raiz principal (JUBERY et al., 2021). Para isto, foram adotadas 691 imagens e aprendizado profundo. Para detecção dos nódulos, foi utilizada a arquitetura RetinaNet (LIN et al., 2017) e para a detecção da raiz principal foi utilizada a arquitetura U-Net (RONNEBERGER et al., 2015), adotando 80% das amostras para o treinamento e 20% para teste. Além disso, foi adotado o aumento de dados com técnicas de transformações geométricas (*flipping*), zoom de 120%, translação nas direções horizontal e vertical em 5% e rotação das imagens até 15 graus (SHORTEN; KHOSHGOFTAAR, 2019).

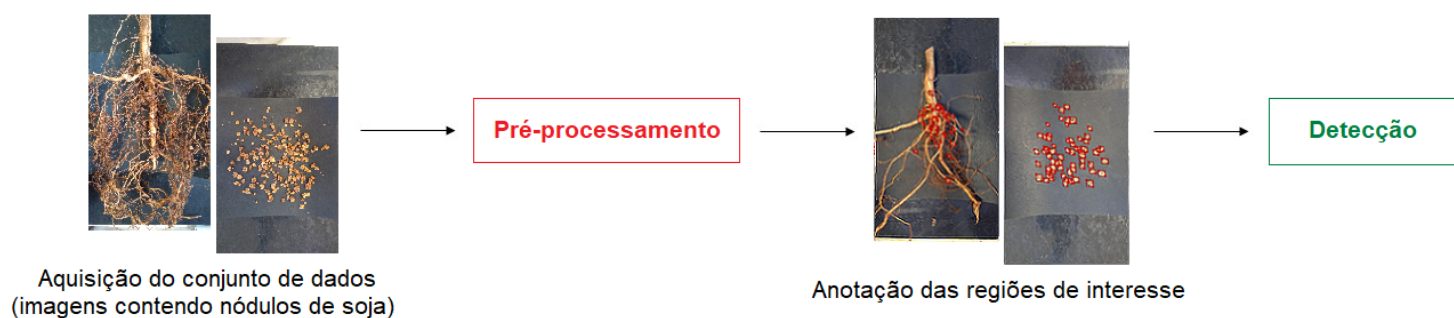
Diferente dos trabalhos anteriores, este trabalho propõe uma abordagem com modelos clássicos de aprendizado de máquina, tais como KNN, RF e SVM, utilizando descritores tradicionais e arquiteturas profundas para extrair as características, além de redes neurais convolucionais para classificação de nódulos em raízes de cultivares de soja. E, além disso, identificar e validar modelos de visão computacional, tais como YOLO v7 (WANG et al., 2022), RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016) para detectar automaticamente nódulos provenientes da fixação biológica do nitrogênio em raízes de cultivares de soja.

3 MATERIAIS E MÉTODOS

Para descrever a metodologia deste trabalho, é apresentado na Seção 3.1 o conjunto de dados utilizado. Na Seção 3.2 é descrito como foi realizado o pré-processamento e a anotação das imagens do conjunto de dados. A Seção 3.3 descreve o processo de segmentação das regiões de interesse das imagens anotadas, a técnica aplicada para aumentar a quantidade de amostras e quais descritores de imagens foram utilizados para realizar a extração de características. Por fim, nas Seções 3.4 e 3.5 são apresentados os modelos de aprendizado de máquina e aprendizado profundo utilizados na classificação das amostras do conjunto de dados, bem como os modelos de visão computacional usados para detectar automaticamente os nódulos em raízes de cultivares de soja. A Figura 8 ilustra o arcabouço da metodologia proposta com as abordagens de classificação por meio de modelos tradicionais de aprendizado de máquina e arquiteturas profundas e detecção por meio de modelos de visão computacional.



(a)



(b)

Figura 8: Arcabouço da metodologia proposta. (a) Abordagem de classificação por meio de modelos tradicionais de aprendizado de máquina e arquiteturas profundas; (b) Abordagem de deteção por meio de modelos de visão computacional.

3.1 Aquisição do conjunto de dados

As imagens de raízes de soja contendo nódulos foram obtidas no Laboratório de Biotecnologia do Solo da Embrapa Soja, Londrina - PR, por meio do Instituto Nacional de Ciência e Tecnologia (INCT) Microrganismos Promotores do Crescimento de Plantas Visando à Sustentabilidade Agrícola e à Responsabilidade Ambiental, coordenado pela Dra. Mariangela Hungria (EMBRAPA, 2022).

O conjunto de dados contempla um total de 190 imagens com resolução de 2.608×4.632 pixels de largura e altura, respectivamente. Deste total de 190 imagens, 162 imagens são amostras de raízes de soja contendo nódulos e 28 imagens são amostras de nódulos isolados das raízes.

As imagens foram padronizadas em um fundo de cor escura, no qual as raízes e os nódulos ficaram sobrepostos a esse fundo. Além disso, foi removido o excesso de terra fixado nas raízes, maximizando a visibilidade dos nódulos, conforme pode ser observado na Figura 9.

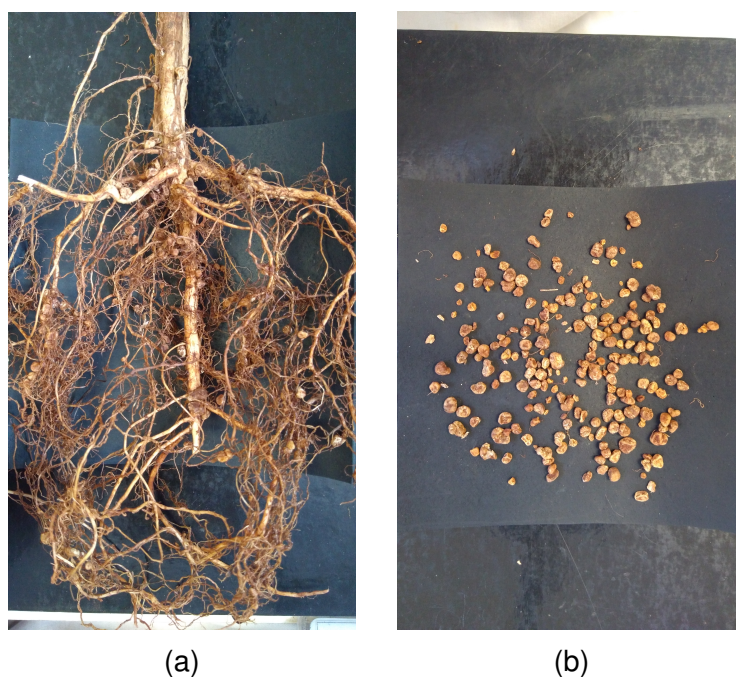


Figura 9: Exemplo de imagens do conjunto de dados. (a) Raízes contendo nódulos; (b) Nódulos isolados das raízes.

Fonte: Embrapa (2022)

3.2 Pré-processamento e anotação do conjunto de dados

A princípio optou-se em padronizar a orientação das imagens antes de iniciar o processo de anotação, pois conforme pode ser observado na Figura 10, o conjunto de dados contém imagens com diferentes orientações.





Figura 10: Exemplo de imagens com diferentes orientações do conjunto de dados.

Fonte: Embrapa (2022)

Adotou-se os seguintes critérios para a padronização de orientação das imagens do conjunto de dados: a) as imagens contendo raízes foram rotacionadas para modo retrato e com o caule da planta voltado para a parte superior da imagem; b) as imagens contendo nódulos isolados das raízes foram rotacionadas para modo retrato. É possível observar na Figura 11 como foram padronizadas as imagens.





Figura 11: Exemplo de imagens com orientação padronizada.

Após a padronização de orientação das imagens do conjunto de dados, seguiu-se com a anotação das mesmas. Para isto, foi adotada a ferramenta LabelMe (WADA, 2022). Esta ferramenta é uma aplicação desktop com interface gráfica, desenvolvida em Python, para anotação de regiões de interesse nas imagens e está disponível para os sistemas operacionais Windows, Ubuntu e macOS.

A ferramenta LabelMe disponibiliza algumas formas para anotar as regiões de interesse na imagem, por exemplo, polígono, retângulo envolvente, círculo, linha, ponto, dentre outras. Como o processo de anotação de regiões de interesse nas imagens é uma atividade morosa e que demanda muita atenção, visando otimizar o tempo elegeu-se a forma retângulo envolvente, pois com apenas dois pontos na diagonal é possível definir o retângulo que envolve a região de interesse na imagem, conforme pode ser observado na Figura 12. Além disso, o discente de iniciação científica Igor Gabrig contribuiu com o processo de anotação do conjunto de dados.

Para o processo de anotação do conjunto de dados, foram definidas três classes para anotação, sendo elas, nódulo, raiz e fundo. Adotou-se o seguinte padrão para anotação das imagens: anotação de todos os nódulos existentes (anotação em vermelho), anotação de cinco amostras de raízes (anotação em verde) e anotação de cinco amostras de fundo (anotação em amarelo), e, como existem imagens presentes no conjunto de dados em que os nódulos estão isolados das raízes da soja, optou-se

em manter a anotação de todos os nódulos e a anotação de cinco amostras de fundo, conforme mostra a Figura 13.

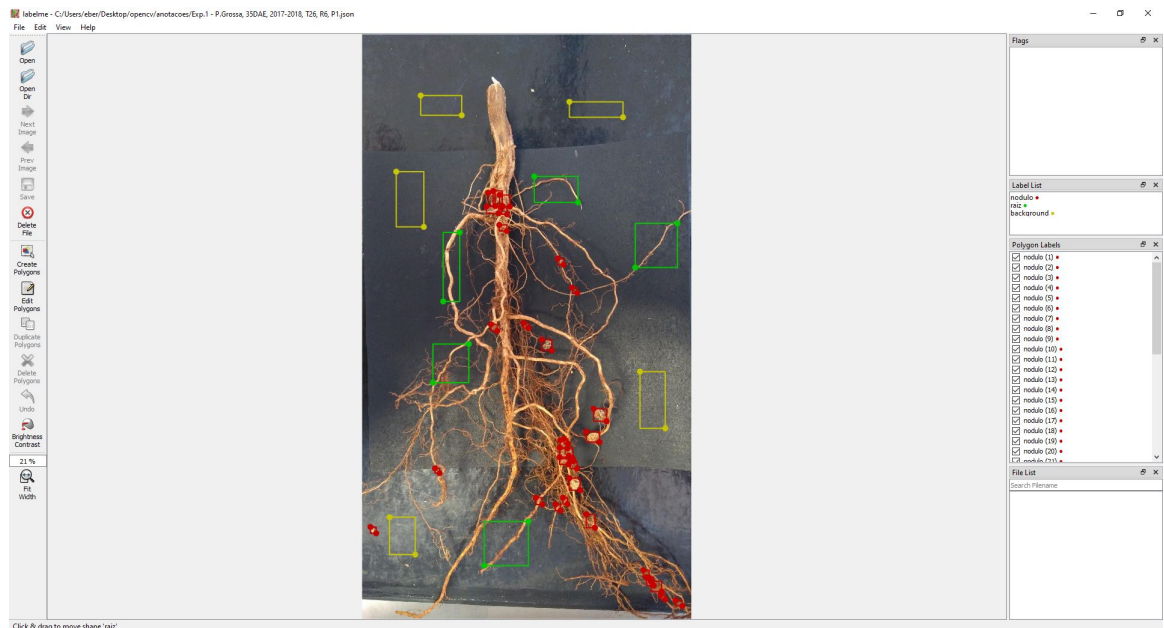


Figura 12: Exemplo de anotação das imagens utilizando a ferramenta LabelMe pela forma retângulo envolvente.

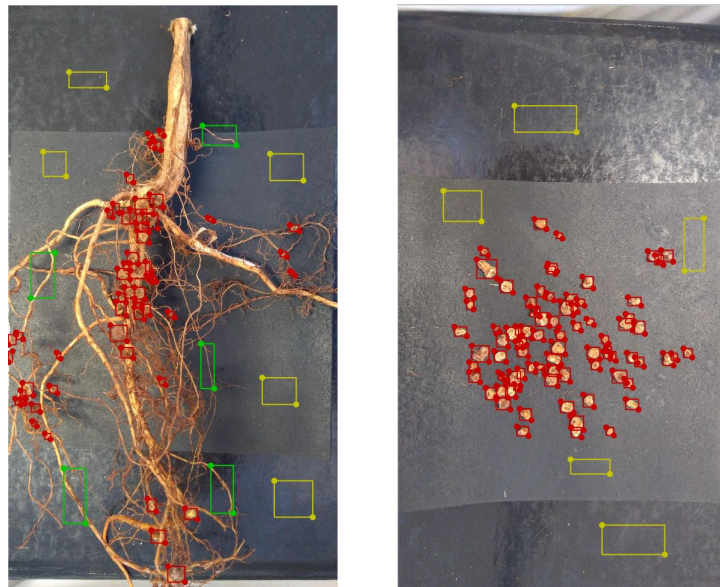


Figura 13: Exemplo de anotação das regiões de interesse.

A ferramenta LabelMe além de possibilitar atribuir o rótulo para cada uma das classes anotadas, permite também atribuir um identificador para cada uma das regiões de interesse na imagem. A atribuição do rótulo para cada uma das regiões de interesse ocorreu da seguinte forma: poderia ser definido os seguintes rótulos n

ou nódulo para representar a classe nódulo, r ou raiz para representar a classe raiz e f ou fundo para representar a classe fundo, as abreviações n, r e f foram definidas para facilitar o processo de anotação, da mesma forma que não foram definidos neste momento os identificadores para as regiões de interesse. Foi implementado um algoritmo, em Python, para interpretar a saída da ferramenta LabelMe convertendo, respectivamente, os rótulos n, r e f para os rótulos nódulo, raiz e fundo. Além disso, o algoritmo atribuiu o identificador, por rótulo, das regiões de interesse, sendo este, um número inteiro que inicia em 1 e é incrementado a cada nova região de interesse.

A saída gerada pela ferramenta LabelMe é um arquivo JSON (*JavaScript Object Notation*). Este arquivo contém uma lista de objetos da forma que foi anotada cada uma das regiões de interesse, e, que neste caso foi retângulo envolvente. Esta lista contém os rótulos das classes anotadas, uma lista de dois pontos com suas respectivas coordenadas cartesianas e o identificador da região de interesse. Além dessa lista de objetos, o arquivo JSON contém também informações da imagem anotada, como o caminho da imagem salva no sistema operacional, informações de altura e largura da imagem, bem como a imagem convertida em base 64.

3.3 Segmentação das regiões de interesse e extração de características

Após padronizar todos arquivos JSON com os nomes das classes e atribuição do identificador das regiões de interesse, iniciou-se o processo de extração destas regiões de interesse. Para isto, foi desenvolvido um algoritmo em Python no qual processou os pontos das coordenadas de cada uma das regiões de interesse dos 190 arquivos JSON. Estes arquivos foram associados às imagens originais do conjunto de dados, possibilitando extrair cada região de interesse, gerando novas subimagens.

As subimagens geradas foram salvas com o seguinte padrão de nome: identificador, classe que a região de interesse representa e nome da imagem original.

Como resultado deste processamento gerou-se um total de 11.302 amostras, deste total de amostras, 9.542 são nódulos, 810 são raízes e 950 são fundos. Pode-se observar na Figura 14 um exemplo de cada uma das classes. Este processamento durou cerca de 1 minuto e 55 segundos para ser concluído, considerando as seguintes

configurações da máquina utilizada: sistema operacional Windows 10 (64 bits) com processador Intel(R) Core(TM) i7-9700 com 3.00 GHz, 16,00 GB de memória RAM, 512 GB de armazenamento, sendo este um SSD, 1 TB de armazenamento, sendo este um HD e uma placa de vídeo da NVIDIA GeForce GT 1030.

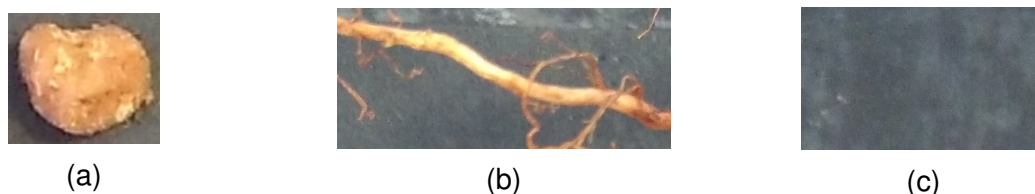


Figura 14: Exemplo de regiões de interesse extraídas das imagens do conjunto de dados. (a) Nódulo; (b) Raiz; (c) Fundo.

Após a segmentação das regiões de interesse das imagens originais, o conjunto de dados contendo as subimagens foi dividido físico e aleatoriamente em uma proporção de 80/20, sendo que 80% dos dados utilizados para treinamento e 20% dos dados utilizados para teste. O conjunto de dados de treinamento (80%) compreende um total de 9.041 amostras, sendo deste total, 7.650 amostras da classe nódulo, 634 amostras da classe raiz e 757 amostras da classe fundo. Já o conjunto de dados de teste (20%) compreende um total de 2.261 amostras, sendo deste total, 1.892 amostras da classe nódulo, 176 amostras da classe raiz e 193 amostras da classe fundo.

Com a divisão do conjunto de dados na proporção 80/20, ficou evidente a desproporção entre a quantidade de amostras de cada uma das classes (nódulo, raiz e fundo). Para resolver essa questão, optou-se em aplicar a técnica de *data augmentation* (SHORTEN; KHOSHGOFTAAR, 2019) no conjunto de dados de treinamento (80%), na qual consiste em aumentar a quantidade de amostras para as classes nódulo, raiz e fundo. Neste caso, optou-se em aumentar a quantidade de amostras de todas as classes até atingir 10.000 amostras para cada uma.

As estratégias utilizadas para aumentar a quantidade de amostras das classes nódulo, raiz e fundo foram: rotacionar, ajustar o contraste e o brilho das subimagens, iniciando com o fator de rotação, contraste e brilho em 0,20 sendo incrementado 0,15 ao fator a cada nova iteração. É importante destacar que nem todas as subimagens foram submetidas às três estratégias da técnica de *data augmentation*. Como

mencionado anteriormente, só foram aplicadas as estratégias de *data augmentation* até atingir a quantidade de 10.000 amostras para cada uma das classes, partindo do conjunto de dados de treinamento. Com isso, após o processo de *data augmentation* a classe nódulo teve um total de 2.350 amostras adicionadas, a classe raiz teve um total de 9.366 amostras adicionadas e a classe fundo teve um total de 9.243 amostras adicionadas. Além disso, as imagens resultantes do processo de *data augmentation* foram salvas fisicamente no disco rígido do computador. O processamento desta etapa durou cerca de 7 horas, 8 minutos e 16 segundos, gerando um total de 20.959 novas amostras. A Figura 15 apresenta um exemplo deste processamento para as três classes (nódulo, raiz e fundo).

Após o balanceamento entre a quantidade de amostras de cada uma das classes (nódulo, raiz e fundo), aplicou-se em cada uma das regiões de interesse 25 descritores de características (BRESSAN et al., 2018). A Tabela 1 apresenta as informações dos descritores (tradicionais e arquiteturas profundas) considerados, tais como, os tipos, as quantidades de características e o tempo de extração de características (em segundos) para cada descritor.

Tabela 1: Descritores de imagens, tipos, quantidade de características e tempo de extração de características.

	Descritor	Tipo	Características	Tempo (s)	Referência
Tradicionais	Momentos Invariantes de Hu	Forma	7	44	Hu (1962)
	First Order Measures (FOM) Gray	Textura	8	13.910	Irons e Petersen (1981)
	Haralick	Textura	13	614	Haralick et al. (1973)
	First Order Measures (FOM) Color	Textura	24	39.217	Irons e Petersen (1981)
	Global Color Histogram (GCH)	Cor	30	8.707	Stricker e Orengo (1995)
	Zernike	Forma	72	197	Teague (1980)
	Local Binary Patterns (LBP)	Textura	352	1.163	Ojala et al. (2000)
Arquiteturas Profundas	VGG16	Genérico	512	11.759	Simonyan e Zisserman (2015)
	VGG19	Genérico	512	13.063	Simonyan e Zisserman (2015)
	DenseNet121	Genérico	1.024	7.405	Huang et al. (2017)
	MobileNet	Genérico	1.024	4.308	Howard et al. (2017)
	NASNetMobile	Genérico	1.056	3.951	Zoph et al. (2018)
	MobileNetV2	Genérico	1.280	3.099	Sandler et al. (2018)
	InceptionResNetV2	Genérico	1.536	9.290	Szegedy et al. (2017)
	DenseNet169	Genérico	1.664	7.802	Huang et al. (2017)
	DenseNet201	Genérico	1.920	8.975	Huang et al. (2017)
	InceptionV3	Genérico	2.048	6.789	Szegedy et al. (2016)
	ResNet101	Genérico	2.048	10.403	He et al. (2016a)
	ResNet101V2	Genérico	2.048	7.605	He et al. (2016b)
	ResNet152	Genérico	2.048	12.954	He et al. (2016a)
	ResNet152V2	Genérico	2.048	8.343	He et al. (2016b)
	ResNet50	Genérico	2.048	8.389	He et al. (2016a)
	ResNet50V2	Genérico	2.048	7.951	He et al. (2016b)
Xception	Genérico	2.048	6.654	Chollet (2017)	
NASNetLarge	Genérico	4.032	13.435	Zoph et al. (2018)	

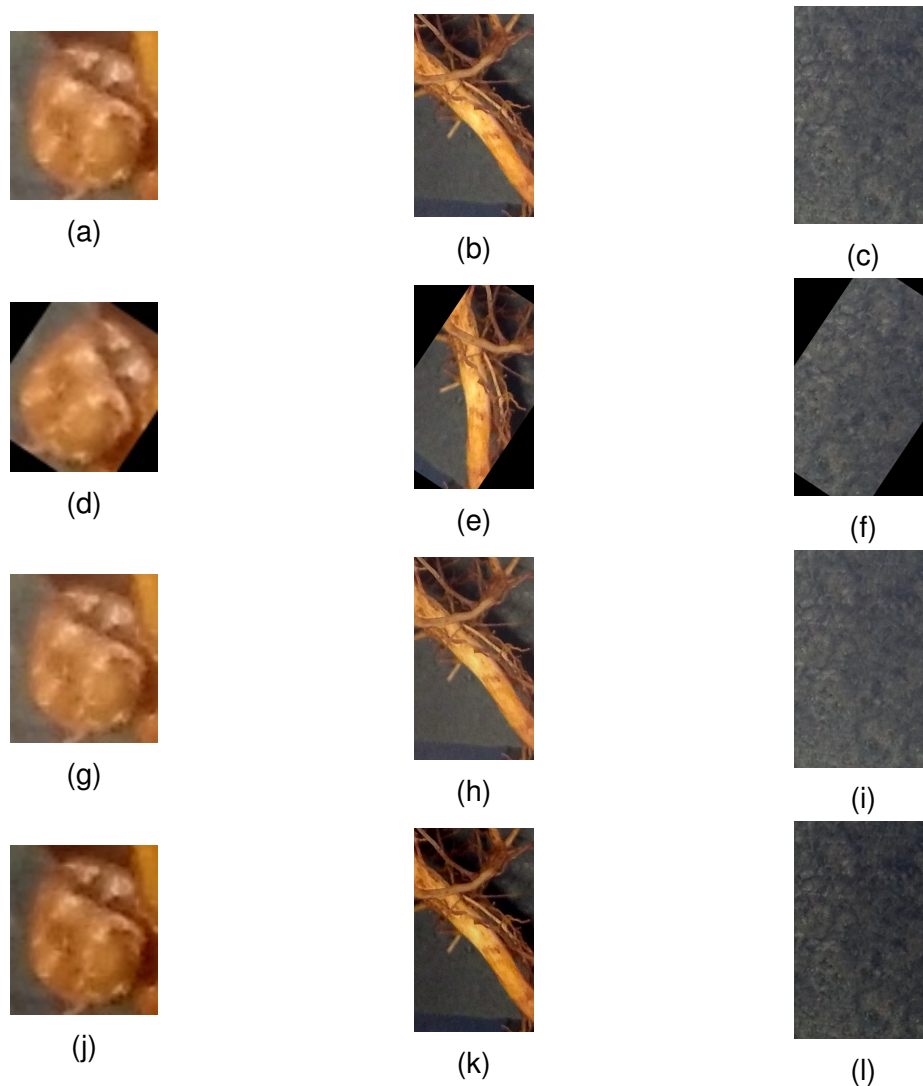


Figura 15: Exemplo de *data augmentation* nas regiões de interesse para cada uma das três classes (nódulo, raiz e fundo). (a), (b) e (c) Regiões de interesse originais; (d), (e) e (f) Regiões de interesse rotacionadas com o fator 0,20 aplicado; (g), (h) e (i) Regiões de interesse com ajuste no contraste aplicado no fator 0,20; (j), (k) e (l) Regiões de interesse com ajuste no brilho aplicado no fator 0,20.

3.4 Classificação

Após a extração das características foi realizado o treinamento dos modelos de aprendizado de máquina e aprendizado profundo. É necessário evidenciar que as características foram reescaladas no intervalo $[0,1]$ utilizando a técnica conhecida como MinMax (WITTEN et al., 2011), de forma a serem comparáveis entre si antes de serem submetidas aos classificadores tradicionais de aprendizado de máquina.

Foram selecionados três classificadores clássicos amplamente utilizados na literatura: KNN, RF e SVM. No caso do SVM, foram considerados dois tipos de

kernel: Linear (KL) e Polinomial de Grau 2 (KPG2). Além disso, para a abordagem de aprendizado profundo, foram selecionadas dezoito redes neurais convolucionais para realizar a classificação das amostras do conjunto de dados. Os detalhes sobre essas redes neurais estão apresentados na Tabela 2.

Tabela 2: Redes Neurais Convolucionais.

Nome	Referência
VGG16	Simonyan e Zisserman (2015)
VGG19	Simonyan e Zisserman (2015)
DenseNet121	Huang et al. (2017)
MobileNet	Howard et al. (2017)
NASNetMobile	Zoph et al. (2018)
MobileNetV2	Sandler et al. (2018)
InceptionResNetV2	Szegedy et al. (2017)
DenseNet169	Huang et al. (2017)
DenseNet201	Huang et al. (2017)
InceptionV3	Szegedy et al. (2016)
ResNet101	He et al. (2016a)
ResNet101V2	He et al. (2016b)
ResNet152	He et al. (2016a)
ResNet152V2	He et al. (2016b)
ResNet50	He et al. (2016a)
ResNet50V2	He et al. (2016b)
Xception	Chollet (2017)
NASNetLarge	Zoph et al. (2018)

É fundamental ressaltar que os hiperparâmetros dos modelos tradicionais de classificação e das redes neurais convolucionais não foram alterados. Isso ocorreu porque o objetivo deste trabalho não é determinar os melhores parâmetros para esses classificadores específicos neste conjunto de dados. As bibliotecas e os *frameworks* utilizados foram: NumPy, Pandas, Scikit-learn, TensorFlow e Keras.

As arquiteturas profundas e os classificadores tradicionais foram avaliados usando tanto o conjunto de dados original quanto o conjunto de dados aumentado por meio da técnica de *data augmentation*. Como mencionado anteriormente, o conjunto de dados original foi dividido na proporção de 80/20, em que 80% dos dados foram utilizados para treinamento e 20% dos dados foram usados para teste. A técnica de *data augmentation* foi aplicada exclusivamente ao conjunto de dados de treinamento, ou seja, aos 80% do conjunto de dados.

Para o processo de treinamento e avaliação os modelos, também foi empregada a técnica de validação cruzada (REFAEILZADEH et al., 2009) com 10-partes. Esse procedimento consistiu em realizar dez iterações no conjunto de dados de treinamento (80%), no qual nove partes foram utilizados para treinamento dos modelos e uma parte foi reservado para avaliação dos mesmos. Além disso, o conjunto de teste (20%), que foi inicialmente separado na proporção de 80/20, foi utilizado para verificar a capacidade de generalização dos modelos, ou seja, verificar o desempenho dos modelos ao lidar com dados completamente desconhecidos do processo de treinamento e validação.

A técnica de transferência de aprendizado foi empregada no treinamento das redes neurais convolucionais neste estudo. Essa técnica aproveita o conhecimento adquirido por um modelo pré-treinado em um conjunto de dados para melhorar o aprendizado em um novo conjunto de dados específico (MARQUES et al., 2022). No caso deste estudo, os pesos iniciais das camadas de convoluções foram obtidos a partir do pré-treinamento com o conjunto de dados ImageNet (DENG et al., 2009). Isso permitiu que os modelos se beneficiassem do conhecimento prévio na tarefa de classificação. As camadas de convoluções tiveram seus pesos travados, enquanto apenas os pesos da camada totalmente conectada foram atualizados. A atualização dos pesos ocorreram a cada trinta e duas imagens, valor padrão de implementação das arquiteturas.

Inicialmente, o treinamento das redes neurais convolucionais foi definido para 6.000 épocas. No entanto, a técnica de *early stopping* (PRECHELT, 2012) também foi aplicada para evitar o sobreajuste (*overfitting*) do modelo aos dados de treinamento. O *early stopping* interrompe o treinamento assim que um critério de parada pré-definido é atingido, ajudando a obter um modelo mais generalizável. Neste estudo, o critério utilizado foi o valor mínimo da função de perda (*loss*), com uma tolerância de três épocas consecutivas sem melhora.

Durante o processo de treinamento, os pesos das redes neurais convolucionais foram avaliados e salvos com base na melhor acurácia obtida.

Neste estudo, foram utilizadas diversas métricas para avaliar tanto os modelos tradicionais de aprendizado de máquina quanto as arquiteturas profundas. As métricas

incluíram acurácia, sensibilidade, especificidade, precisão, F1, *kappa* e *loss*. A acurácia mede a taxa geral de classificações corretas do modelo. A sensibilidade avalia a capacidade do modelo em identificar corretamente os casos positivos. A especificidade mede a capacidade do modelo em identificar corretamente os casos negativos. A precisão quantifica a proporção de resultados corretamente identificados como positivos entre todos os casos identificados como positivos pelo modelo. A métrica F1 é uma média harmônica entre precisão e sensibilidade, fornecendo uma medida equilibrada de desempenho do modelo. O *kappa* é uma medida que leva em consideração o desempenho do modelo além do acaso, corrigindo a acurácia observada. A métrica *loss* é específica para as redes neurais convolucionais (CNNs) e mede a discrepância entre as classificações previstas pelo modelo e os rótulos verdadeiros. Todas essas métricas são importantes para avaliar diferentes aspectos do desempenho dos modelos e fornecer uma visão abrangente de sua eficácia (SAMMUT; WEBB, 2017).

3.5 Detecção

Neste estudo, foram utilizados três modelos de visão computacional para a detecção automática de nódulos em raízes de cultivares de soja. Os modelos adotados foram o YOLO v7 (WANG et al., 2022), o RetinaNet (LIN et al., 2018) e o Faster R-CNN (REN et al., 2016).

Para realizar essa tarefa, o conjunto de dados adquirido, pré-processado e anotado nas Seções 3.1 e 3.2 foi utilizado. Esse conjunto de dados foi dividido fisicamente e de forma aleatória em dois subconjuntos. O primeiro subconjunto, que representou 80% das imagens, foi utilizado para o treinamento das redes neurais convolucionais. O segundo subconjunto, correspondendo a 20% das imagens, foi reservado para teste das redes neurais convolucionais.

No subconjunto de dados de treinamento, que correspondia a 80% do conjunto total, foi aplicada a técnica de validação cruzada (REFAEILZADEH et al., 2009) com 10-partes para treinar e avaliar os modelos mencionados anteriormente. Essa técnica consiste em dividir o conjunto de treinamento em 10 partes, realizando 10 iterações.

Em cada iteração, nove partes foram utilizados para treinar o modelo e uma parte foi utilizado para validação. Isso permitiu uma avaliação mais robusta e precisa do desempenho dos modelos, levando em consideração diferentes combinações de dados de treinamento e validação.

Além disso, o subconjunto de teste, que foi separado inicialmente e correspondia a 20% do conjunto total, foi utilizado para verificar a capacidade de generalização dos modelos. Essa etapa permitiu avaliar como os modelos se comportam ao lidar com dados completamente desconhecidos, ou seja, dados que não foram usados no processo de treinamento e validação. Isso é importante para verificar se os modelos são capazes de generalizar e realizar previsões precisas em novos dados.

Para o treinamento dos modelos foi empregado o método de transferência de aprendizado. Essa técnica consiste em aproveitar o conhecimento de um modelo pré-treinado para aprender outro conjunto de dados, com o objetivo de melhorar o aprendizado no domínio destino específico (MARQUES et al., 2022). No caso deste estudo, os pesos iniciais para treinar as redes neurais convolucionais YOLO v7, RetinaNet e Faster R-CNN foram obtidas a partir do pré-treinamento realizado no conjunto de dados Microsoft COCO (LIN et al., 2015). Essa abordagem permitiu que os modelos se beneficiassem do conhecimento prévio adquirido na tarefa de detecção de objetos em imagens.

As três redes neurais convolucionais foram treinadas com diferentes épocas, especificamente 200, 400, 600, 800 e 1.000, com atualização dos pesos a cada 4 imagens.

Durante a validação cruzada, foram realizadas várias execuções para avaliar o desempenho e a robustez dos modelos de visão computacional. A métrica selecionada para essa avaliação foi a *average precision* (AP) com um limiar de 0,5 (AP@0,5). Essa métrica considera a sobreposição (IoU - *Intersection over Union*) entre os objetos previstos pelos modelos e os *ground truth*, que são as anotações de referência. A escolha dessa métrica se baseia na sua ampla utilização na literatura (SU et al., 2023; REN et al., 2023; DANSO et al., 2023; MORABIA et al., 2020).

4 RESULTADOS E DISCUSSÃO

Com o intuito de avaliar a eficácia da metodologia de inteligência computacional proposta neste estudo para a análise, a classificação e a detecção automática de nódulos em raízes de cultivares de soja, foram realizadas cinco etapas. Primeiramente, utilizou-se o conjunto de dados de imagens descrito na Seção 3.1, que incluiu as regiões de interesse e a extração de características (conforme explicado na Seção 3.3). Em seguida, foram aplicados processos de classificação utilizando modelos clássicos de aprendizado de máquina e redes neurais convolucionais (conforme descrito na Seção 3.4), bem como o processo de detecção de nódulos por meio de modelos de visão computacional (apresentado na Seção 3.5).

Neste estudo, foram conduzidos diversos experimentos envolvendo classificação utilizando modelos tradicionais de aprendizado de máquina, bem como classificação utilizando modelos de aprendizado profundo, juntamente com a detecção por meio de modelos de visão computacional. Para facilitar a compreensão dos resultados obtidos, neste capítulo são apresentados os resultados em seções separadas.

Na Seção 4.1, são apresentados os resultados dos experimentos de classificação utilizando modelos clássicos com o conjunto de dados original e também com o conjunto de dados aumentado. Em seguida, na Seção 4.2, são descritos os resultados dos experimentos de classificação utilizando redes neurais convolucionais também com o conjunto de dados original e aumentado. Por fim, na Seção 4.3, são apresentados os resultados dos experimentos de detecção utilizando modelos de visão computacional.

4.1 Classificação por meio de modelos clássicos de aprendizado de máquina

O experimento de classificação foi realizado utilizando diferentes modelos: KNN, RF, SVM (KL) e SVM (KPG2). Nesse experimento, os hiperparâmetros dos modelos não foram modificados sendo adotados os parâmetros padrões dos métodos. As características obtidas por extratores tradicionais e arquiteturas profundas foram utilizadas como entrada para os modelos.

As Tabelas 3 e 4 exibem a média das acurácias nos conjuntos de dados de teste, em que os modelos foram treinados utilizando os conjuntos de dados original e aumentado, respectivamente. Essas tabelas incluem a quantidade de características de cada descritor, juntamente com as médias de acurácia \pm desvio padrão obtidas para cada modelo, considerando separadamente os descritores tradicionais e as arquiteturas profundas. Os resultados mais destacados (ou seja, os melhores classificadores) para cada extrator estão em negrito, os resultados mais destacados (ou seja, os melhores extratores) para cada classificador estão sublinhados e os melhores resultados (ou seja, as maiores médias de acurácia) obtidos estão marcados com um asterisco.

Ao analisar a Tabela 3, na qual os modelos foram treinados utilizando o conjunto de dados original, pode-se observar que o descritor tradicional GCH em combinação com o classificador RF obteve o melhor resultado entre os extratores tradicionais. Por outro lado, os descritores profundos DenseNet121, ResNet101, ResNet152 e ResNet50, combinados com os classificadores KNN, RF, SVM (KL) e SVM (KPG2), apresentaram resultados superiores em comparação com as outras combinações de descritores profundos e classificadores. É importante ressaltar que as combinações de classificadores e extratores profundos resultaram em acurácias maiores em relação aos descritores tradicionais.

Além disso, conforme mostrado na Tabela 3, pode-se observar que a quantidade de características dos descritores tradicionais é menor em comparação aos descritores profundos. Segundo Geron (2019), quanto maior a quantidade de características para cada instância, mais lento se torna o treinamento do modelo.

Ao analisar a Tabela 4, na qual os modelos foram treinados utilizando

Tabela 3: Média de acurácias \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados original.

	Descritor	Características	Conjunto de dados original			
			KNN	RF	SVM (KL)	SVM (KPG2)
Tradicionais	Momentos Invariantes de Hu	7	0,930 \pm 0,002	0,953 \pm 0,001	0,911 \pm 0,000	0,934 \pm 0,001
	FOM (Gray)	8	0,980 \pm 0,001	0,992 \pm 0,001	0,976 \pm 0,001	0,986 \pm 0,001
	Haralick	13	0,987 \pm 0,001	0,991 \pm 0,001	0,987 \pm 0,000	0,992 \pm 0,000
	FOM (Color)	24	0,990 \pm 0,000	0,995 \pm 0,000	0,992 \pm 0,001	0,993 \pm 0,000
	GCH	30	0,987 \pm 0,001	0,996 \pm 0,000*	0,975 \pm 0,001	0,988 \pm 0,001
	Zernike	72	0,859 \pm 0,003	0,899 \pm 0,001	0,891 \pm 0,000	0,891 \pm 0,000
	LBP	352	0,969 \pm 0,001	0,969 \pm 0,001	0,976 \pm 0,000	0,982 \pm 0,000
Arquiteturas Profundas	VGG16	512	0,995 \pm 0,000	0,997 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000
	VGG19	512	0,996 \pm 0,000	0,997 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000
	DenseNet121	1.024	0,995 \pm 0,000	0,995 \pm 0,000	0,999 \pm 0,000*	0,998 \pm 0,000
	MobileNet	1.024	0,994 \pm 0,000	0,994 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000
	NASNetMobile	1.056	0,956 \pm 0,001	0,959 \pm 0,001	0,971 \pm 0,001	0,962 \pm 0,000
	MobileNetV2	1.280	0,993 \pm 0,000	0,993 \pm 0,001	0,998 \pm 0,000	0,998 \pm 0,000
	InceptionResNetV2	1.536	0,957 \pm 0,001	0,962 \pm 0,001	0,947 \pm 0,001	0,929 \pm 0,000
	DenseNet169	1.664	0,995 \pm 0,000	0,994 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000
	DenseNet201	1.920	0,995 \pm 0,000	0,996 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000
	InceptionV3	2.048	0,967 \pm 0,001	0,968 \pm 0,001	0,985 \pm 0,001	0,969 \pm 0,001
	ResNet101	2.048	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*
	ResNet101V2	2.048	0,971 \pm 0,001	0,974 \pm 0,001	0,990 \pm 0,001	0,965 \pm 0,000
	ResNet152	2.048	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*
	ResNet152V2	2.048	0,978 \pm 0,001	0,980 \pm 0,001	0,992 \pm 0,000	0,980 \pm 0,001
	ResNet50	2.048	0,998 \pm 0,000	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*
	ResNet50V2	2.048	0,976 \pm 0,001	0,975 \pm 0,001	0,992 \pm 0,001	0,978 \pm 0,000
	Xception	2.048	0,982 \pm 0,001	0,981 \pm 0,001	0,994 \pm 0,000	0,987 \pm 0,000
NASNetLarge	4.032	0,961 \pm 0,001	0,966 \pm 0,001	0,978 \pm 0,001	0,973 \pm 0,001	

o conjunto de dados aumentado, pode-se observar que o descritor tradicional FOM (Color), combinado com os classificadores RF e SVM (KPG2), obteve o melhor resultado entre os extratores tradicionais. Por outro lado, os descritores profundos VGG16, VGG19, MobileNet, MobileNetV2, ResNet101, ResNet152 e ResNet50, combinados com os classificadores KNN, RF, SVM (KL) e SVM (KPG2), apresentaram resultados superiores em comparação com as outras combinações de descritores profundos e classificadores. É importante ressaltar que as combinações de classificadores e extratores profundos resultaram em acurácias maiores em relação aos descritores tradicionais.

Além disso, conforme mostrado na Tabela 4, pode-se observar que o treinamento utilizando o conjunto de dados aumentado resultou em uma maior variedade de combinações de descritores e classificadores com melhores resultados, porém com resultados semelhantes ao treinamento dos modelos no conjunto de dados original. É interessante notar que os descritores profundos ResNet101, ResNet152 e

Tabela 4: Média de acurácias \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados aumentado.

	Descriptor	Características	Conjunto de dados aumentado			
			KNN	RF	SVM (KL)	SVM (KPG2)
Tradicionais	Momentos Invariantes de Hu	7	0,848 \pm 0,003	0,891 \pm 0,000	0,640 \pm 0,254	0,639 \pm 0,252
	FOM (Gray)	8	0,982 \pm 0,001	0,992 \pm 0,001	0,983 \pm 0,000	0,990 \pm 0,000
	Haralick	13	0,988 \pm 0,000	0,991 \pm 0,000	0,981 \pm 0,000	0,987 \pm 0,000
	FOM (Color)	24	0,990 \pm 0,000	0,996 \pm 0,000*	0,995 \pm 0,000	0,996 \pm 0,000*
	GCH	30	0,990 \pm 0,000	0,995 \pm 0,000	0,988 \pm 0,000	0,990 \pm 0,000
	Zernike	72	0,707 \pm 0,004	0,794 \pm 0,003	0,670 \pm 0,001	0,709 \pm 0,002
	LBP	352	0,973 \pm 0,001	0,978 \pm 0,001	0,987 \pm 0,000	0,985 \pm 0,000
Arquiteturas Profundas	VGG16	512	0,998 \pm 0,000	0,998 \pm 0,000	0,999 \pm 0,000*	0,999 \pm 0,000*
	VGG19	512	0,998 \pm 0,000	0,998 \pm 0,000	0,999 \pm 0,000*	0,999 \pm 0,000*
	DenseNet121	1.024	0,995 \pm 0,000	0,997 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000
	MobileNet	1.024	0,998 \pm 0,000	0,998 \pm 0,000	0,999 \pm 0,000*	0,999 \pm 0,000*
	NASNetMobile	1.056	0,943 \pm 0,002	0,957 \pm 0,002	0,957 \pm 0,001	0,955 \pm 0,000
	MobileNetV2	1.280	0,997 \pm 0,000	0,998 \pm 0,000	0,999 \pm 0,000*	0,999 \pm 0,000*
	InceptionResNetV2	1.536	0,925 \pm 0,001	0,936 \pm 0,002	0,883 \pm 0,001	0,769 \pm 0,001
	DenseNet169	1.664	0,995 \pm 0,000	0,994 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000
	DenseNet201	1.920	0,997 \pm 0,000	0,996 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000
	InceptionV3	2.048	0,970 \pm 0,001	0,965 \pm 0,001	0,982 \pm 0,001	0,968 \pm 0,000
	ResNet101	2.048	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*
	ResNet101V2	2.048	0,955 \pm 0,001	0,969 \pm 0,002	0,985 \pm 0,000	0,949 \pm 0,000
	ResNet152	2.048	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*
	ResNet152V2	2.048	0,971 \pm 0,001	0,980 \pm 0,000	0,990 \pm 0,001	0,976 \pm 0,000
	ResNet50	2.048	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*	0,999 \pm 0,000*
	ResNet50V2	2.048	0,969 \pm 0,001	0,974 \pm 0,001	0,990 \pm 0,001	0,973 \pm 0,001
	Xception	2.048	0,983 \pm 0,001	0,982 \pm 0,001	0,992 \pm 0,001	0,989 \pm 0,000
NASNetLarge	4.032	0,931 \pm 0,002	0,960 \pm 0,001	0,967 \pm 0,002	0,967 \pm 0,001	

ResNet50 mostraram-se eficientes para a classificação de nódulos em raízes de soja tanto no treinamento com o conjunto de dados original quanto no treinamento com o conjunto de dados aumentado.

É relevante destacar que as Tabelas 3 e 4, além de apresentarem as médias gerais de acurácia juntamente com seus respectivos desvios padrão, também demonstraram consistência nos resultados. Isso indica que o treinamento e a validação dos modelos foram executados de forma consistente, sem grandes variações na dispersão dos resultados em relação à média.

Também é possível observar, no Apêndice A, por meio das Tabelas Complementares 10-13, que todas as classes consideradas no experimento (nódulo, raiz e fundo) obtiveram valores uniformes tanto no treinamento com o conjunto de dados original quanto no treinamento com o conjunto de dados aumentado. Isso evidencia que os modelos foram capazes de realizar a classificação de forma equilibrada e consistente para todas as classes, sem apresentar resultados

discrepantes entre elas.

Em relação ao tempo de treinamento e classificação, em segundos, das amostras utilizando os modelos em combinação com os descritores tradicionais e arquiteturas profundas, é possível observar na Figura 16 que o classificador RF demandou mais tempo de treinamento tanto no conjunto de dados original quanto no conjunto de dados aumentado. No entanto, em termos de tempo de classificação, o classificador RF apresentou um desempenho superior em relação aos demais modelos. Por outro lado, o modelo KNN mostrou o comportamento oposto, com um tempo de treinamento mais rápido, mas um tempo de classificação mais lento em comparação com o RF.

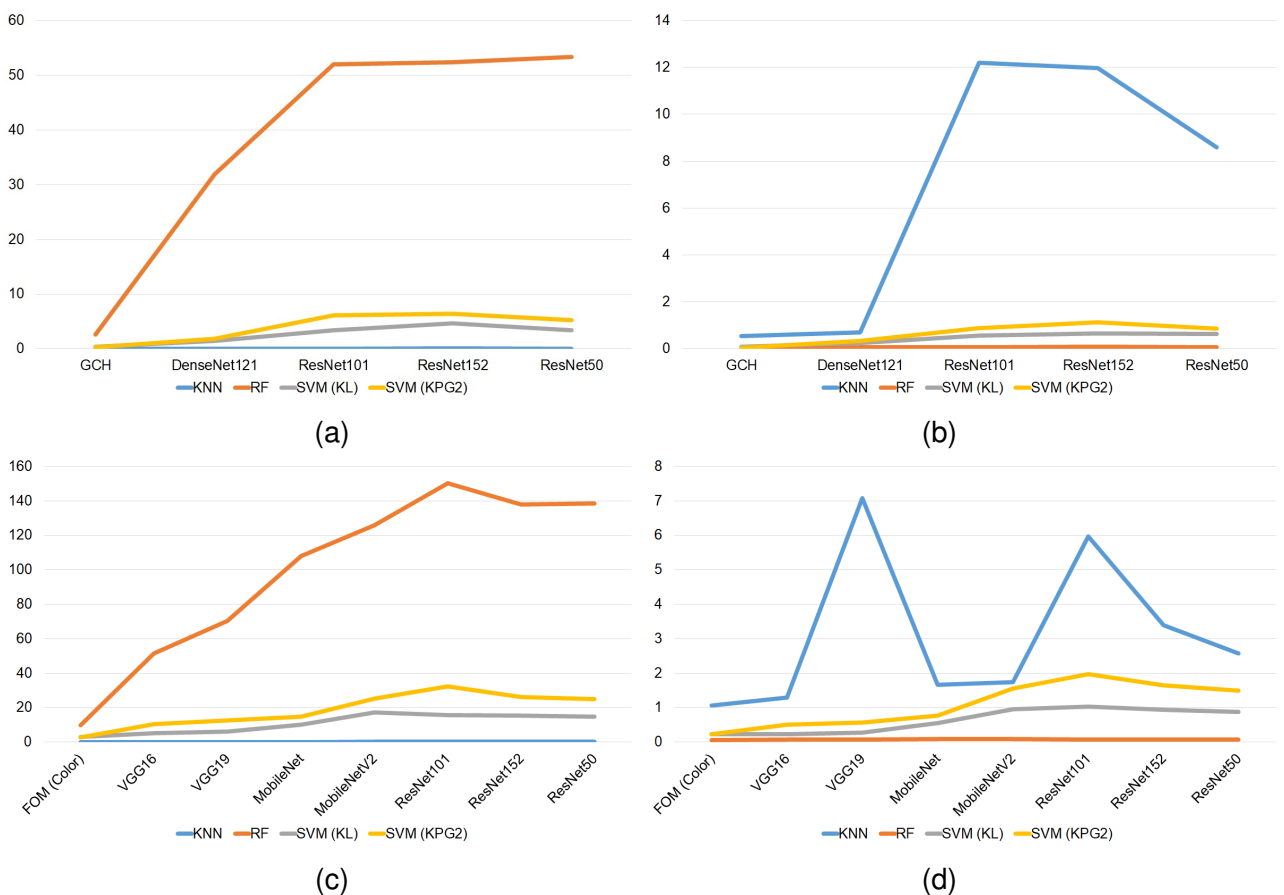
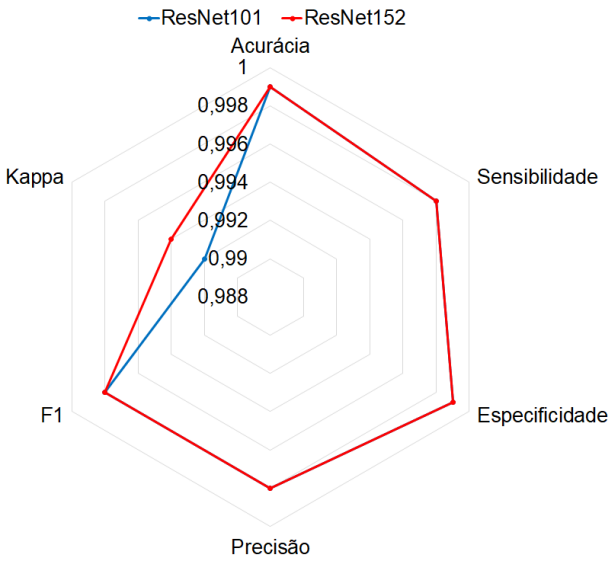


Figura 16: Tempo médio de treinamento e classificação obtidos no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas para os melhores resultados (i.e. maiores acurácias médias – Tabelas 3 e 4). Os modelos foram treinados no conjunto de dados original. (a) e (b) Treinamento e classificação no conjunto de dados original; (c) e (d) Treinamento e classificação no conjunto de dados aumentado.

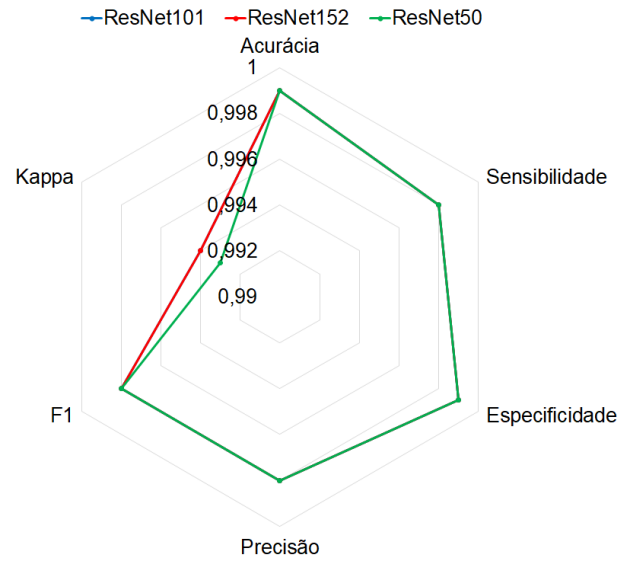
Na Figura 17, é apresentada a média das demais métricas (sensibilidade, especificidade, precisão, F1 e *kappa*) obtidas no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Essas métricas correspondem aos melhores resultados (ou seja, as maiores acurácias médias) encontrados nas Tabelas 3 e 4, onde os modelos foram treinados tanto no conjunto de dados original quanto no conjunto de dados aumentado, respectivamente.

Nos gráficos de radar apresentados na Figura 17, é possível realizar uma comparação visual das múltiplas métricas simultaneamente. Esses gráficos contemplam métricas como acurácia, sensibilidade, especificidade, precisão, F1 e *kappa*, proporcionando uma visão abrangente do desempenho do modelo em diferentes aspectos. Essa representação visual facilita a identificação rápida dos pontos fortes e fracos do modelo em cada métrica. Por exemplo, se um segmento estiver mais afastado do centro em uma métrica específica, indica um desempenho superior nessa métrica. Dessa forma, os gráficos de radar auxiliam na compreensão e análise das métricas, permitindo uma avaliação abrangente do desempenho do modelo.

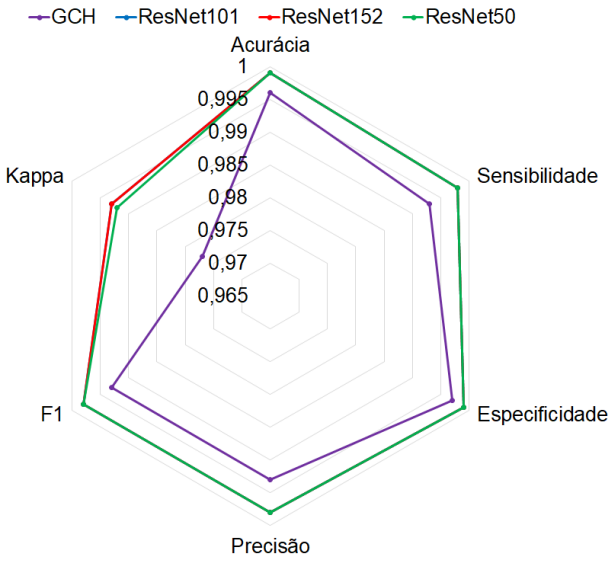
Apesar dos modelos terem obtidos resultados significativos na tarefa de classificação de nódulos radiculares de soja, pequenas confusões ocorreram, como evidenciados pelas duas matrizes de confusão apresentadas na Figura 18. Em particular, o modelo RF apresentou predições incorretas, classificando erroneamente amostras de raízes e fundo como pertencentes à classe de nódulo, conforme ilustrado na Figura 19. Esses exemplos destacam a ocorrência de falsos positivos, indicando uma dificuldade do modelo em distinguir corretamente entre diferentes classes, resultando em uma classificação equivocada nessas situações específicas. No entanto, no geral, as combinações de descritores e classificadores, destacados com asterisco nas Tabelas 3 e 4, apresentaram resultados promissores para a classificação de nódulos em raízes de cultivares de soja.



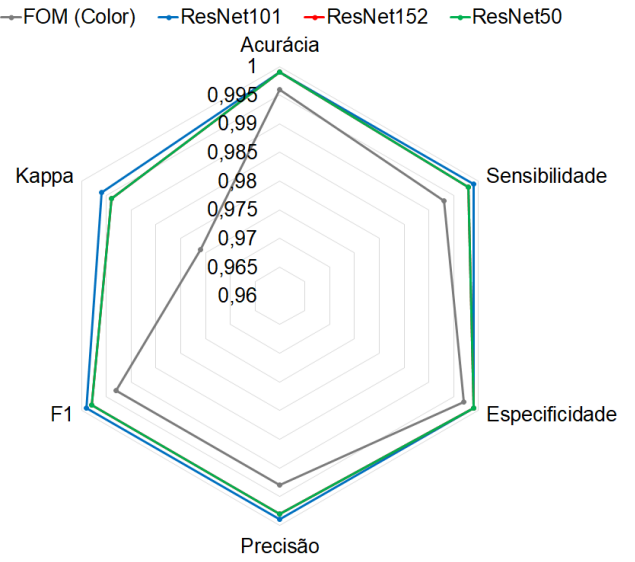
(a)



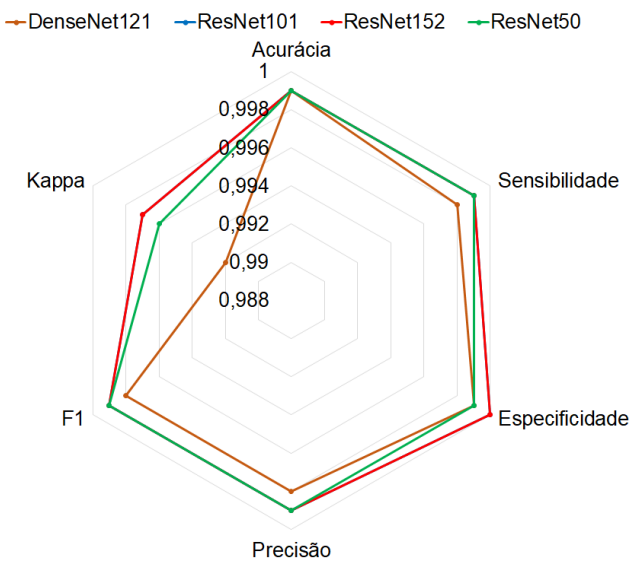
(b)



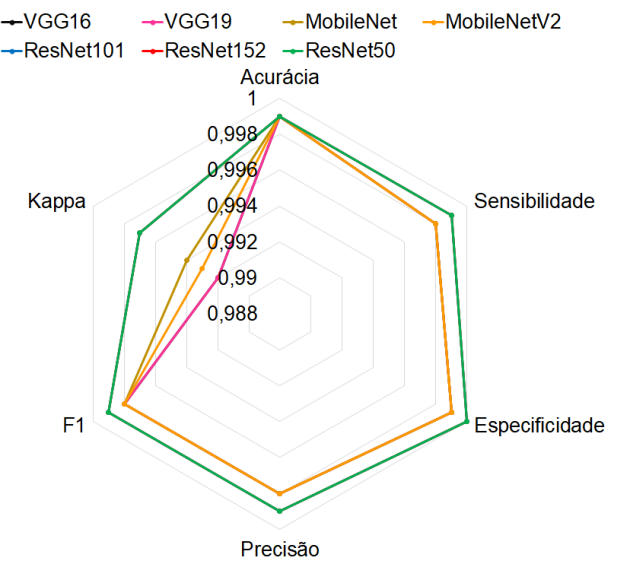
(c)



(d)



(e)



(f)

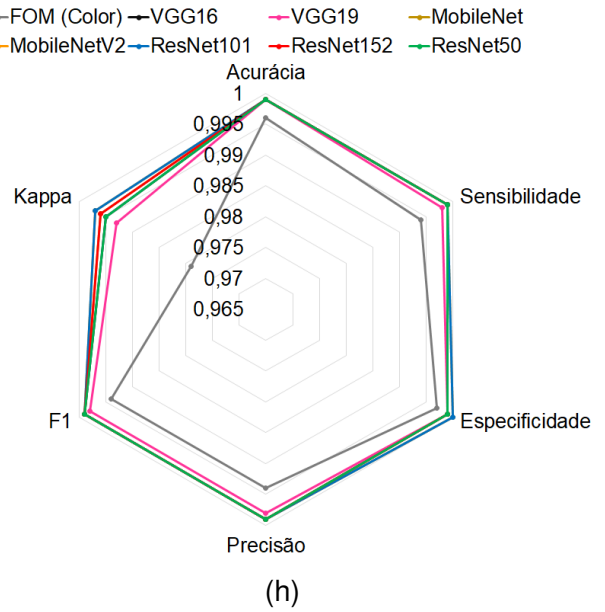
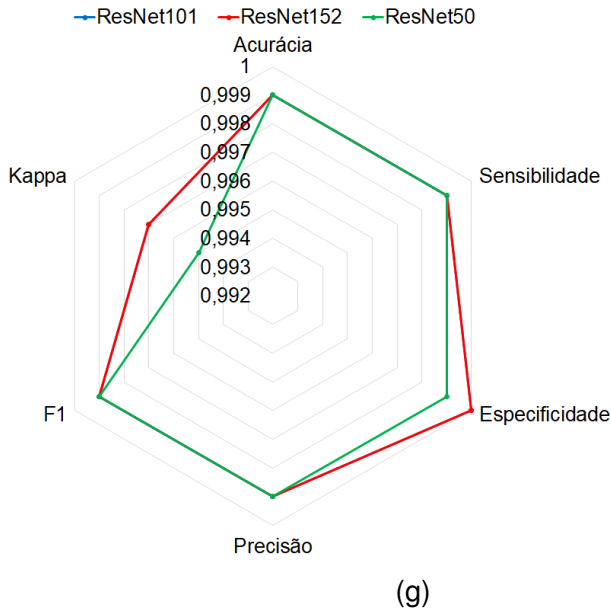


Figura 17: Média das métricas (acurácia, sensibilidade, especificidade, precisão, F1 e kappa) obtidas no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas para os melhores resultados (i.e. maiores acurácias médias – Tabelas 3 e 4). Os modelos foram treinados no conjunto de dados original e aumentado. a) KNN (original); b) KNN (aumentado) c) RF (original); d) RF (aumentado); e) SVM (KL) (original); f) SVM (KL) (aumentado) g) SVM (KPG2) (original) e h) SVM (KPG2) (aumentado).

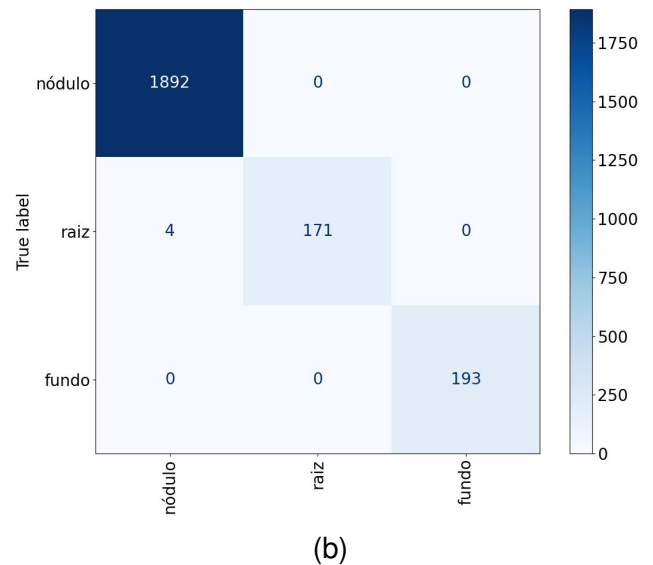
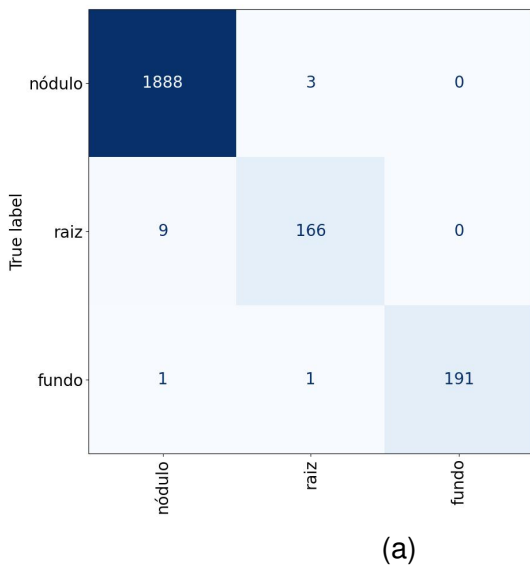


Figura 18: Média das matrizes de confusão obtidas no conjunto de dados de teste pelo classificador RF, utilizando as características obtidas pelo extrator tradicional GCH e arquitetura profunda ResNet101. O modelo foi treinado no conjunto de dados original. a) RF e GCH; b) RF e ResNet101.



Figura 19: Exemplos de confusões na classificação realizada pelo classificador RF em combinação com o extrator GCH. a) Raiz predita como nódulo e b) Fundo predito como nódulo.

4.2 Classificação por meio de redes neurais convolucionais

Foi conduzido também um experimento de classificação utilizando diversas redes neurais convolucionais, conforme apresentado na Tabela 5. Nesse experimento, os hiperparâmetros padrões dos modelos não foram alterados.

Tabela 5: Redes Neurais Convolucionais.

Nome	Referência
VGG16	Simonyan e Zisserman (2015)
VGG19	Simonyan e Zisserman (2015)
DenseNet121	Huang et al. (2017)
MobileNet	Howard et al. (2017)
NASNetMobile	Zoph et al. (2018)
MobileNetV2	Sandler et al. (2018)
InceptionResNetV2	Szegedy et al. (2017)
DenseNet169	Huang et al. (2017)
DenseNet201	Huang et al. (2017)
InceptionV3	Szegedy et al. (2016)
ResNet101	He et al. (2016a)
ResNet101V2	He et al. (2016b)
ResNet152	He et al. (2016a)
ResNet152V2	He et al. (2016b)
ResNet50	He et al. (2016a)
ResNet50V2	He et al. (2016b)
Xception	Chollet (2017)
NASNetLarge	Zoph et al. (2018)

Na Tabela 6 é apresentada a média das acurácias, acompanhada do desvio padrão, para cada modelo nos conjuntos de dados de teste. Esses modelos foram treinados tanto com o conjunto de dados original quanto com o conjunto de dados aumentado. Os melhores resultados, indicados com um asterisco, correspondem às maiores médias de acurácia obtidas entre todos os modelos adotados. Essa

tabela permite uma análise comparativa dos modelos, identificando quais deles apresentaram um desempenho superior na tarefa de classificação.

Tabela 6: Média de acurácias \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados nos conjuntos de dados original e aumentado.

CNN	Conjunto de dados original	Conjunto de dados aumentado
VGG16	$0,993 \pm 0,012^*$	$0,997 \pm 0,003^*$
VGG19	$0,984 \pm 0,036^*$	$0,982 \pm 0,028^*$
DenseNet121	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
MobileNet	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
NASNetMobile	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
MobileNetV2	$0,998 \pm 0,000$	$0,997 \pm 0,000$
InceptionResNetV2	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
DenseNet169	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
DenseNet201	$0,999 \pm 0,000^*$	$1,000 \pm 0,000^*$
InceptionV3	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
ResNet101	$0,958 \pm 0,008$	$0,901 \pm 0,046$
ResNet101V2	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
ResNet152	$0,933 \pm 0,012$	$0,830 \pm 0,054$
ResNet152V2	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
ResNet50	$0,960 \pm 0,014$	$0,932 \pm 0,021$
ResNet50V2	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$
Xception	$0,999 \pm 0,000^*$	$1,000 \pm 0,000^*$
NASNetLarge	$0,999 \pm 0,000^*$	$0,999 \pm 0,000$

No conjunto de dados original, a maioria das redes neurais convolucionais obteve resultados satisfatórios na classificação das amostras. No entanto, ao considerar o conjunto de dados aumentado, as CNNs VGG16, VGG19, DenseNet201 e Xception se destacaram, apresentando as maiores médias de acurácia. Esses resultados evidenciam a eficácia dessas redes neurais convolucionais na tarefa de classificação quando o conjunto de dados é expandido por meio da técnica de aumento de dados. No entanto, é importante ressaltar que as demais redes neurais não alcançaram o mesmo desempenho nessas condições específicas, o que resultou em uma seleção reduzida de redes neurais convolucionais em comparação com os modelos treinados no conjunto de dados original.

A Tabela 7 exibe a média de acurácias por classe, acompanhada pelo desvio padrão, obtida nos conjuntos de dados de teste utilizando redes neurais convolucionais. Os modelos foram treinados nos conjuntos de dados original e

aumentado. Ao analisar essa tabela, é possível observar que as médias de acurácia entre as classes foram consistentes, não havendo um afastamento significativo da média. Isso indica que os modelos foram capazes de realizar a classificação de forma equilibrada e consistente para todas as classes, sem apresentar resultados discrepantes entre elas.

Tabela 7: Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados nos conjuntos de dados original e aumentado.

CNN	Conjunto de dados original			Conjunto de dados aumentado		
	Nódulo	Raiz	Fundo	Nódulo	Raiz	Fundo
VGG16	0,990 \pm 0,018	0,990 \pm 0,018	1,000 \pm 0,000	0,995 \pm 0,004	0,995 \pm 0,004	0,999 \pm 0,001
VGG19	0,976 \pm 0,054	0,977 \pm 0,054	0,999 \pm 0,001	0,974 \pm 0,043	0,978 \pm 0,039	0,995 \pm 0,006
DenseNet121	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
MobileNet	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
NASNetMobile	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
MobileNetV2	0,996 \pm 0,001	0,997 \pm 0,001	1,000 \pm 0,000	0,996 \pm 0,000	0,996 \pm 0,000	1,000 \pm 0,000
InceptionResNetV2	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
DenseNet169	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
DenseNet201	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000
InceptionV3	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
ResNet101	0,936 \pm 0,011	0,967 \pm 0,003	0,969 \pm 0,012	0,855 \pm 0,068	0,956 \pm 0,035	0,891 \pm 0,036
ResNet101V2	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
ResNet152	0,899 \pm 0,018	0,948 \pm 0,003	0,952 \pm 0,016	0,749 \pm 0,079	0,918 \pm 0,053	0,823 \pm 0,035
ResNet152V2	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
ResNet50	0,941 \pm 0,022	0,960 \pm 0,004	0,980 \pm 0,018	0,903 \pm 0,031	0,959 \pm 0,011	0,935 \pm 0,025
ResNet50V2	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
Xception	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000
NASNetLarge	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000

Na Figura 20, é possível observar o tempo de treinamento e classificação, em segundos, das amostras utilizando redes neurais convolucionais. No conjunto de dados original, a CNN NASNetLarge demandou mais tempo tanto no treinamento quanto na classificação. Já no conjunto de dados aumentado, a CNN Xception apresentou o maior tempo de treinamento e classificação. Por outro lado, os modelos VGG16, VGG19, DenseNet121, MobileNet e NASNetMobile demonstraram tempos menores no treinamento no conjunto de dados original. Na classificação, os modelos com os tempos mais curtos foram DenseNet121, MobileNet, NASNetMobile e ResNet50V2. No caso do conjunto de dados aumentado, a CNN VGG16 obteve o melhor desempenho em termos de tempo de treinamento e classificação. Essas informações destacam a variação nos tempos de treinamento e classificação entre as diferentes redes neurais convolucionais, tanto no conjunto de dados original quanto no conjunto de dados aumentado.

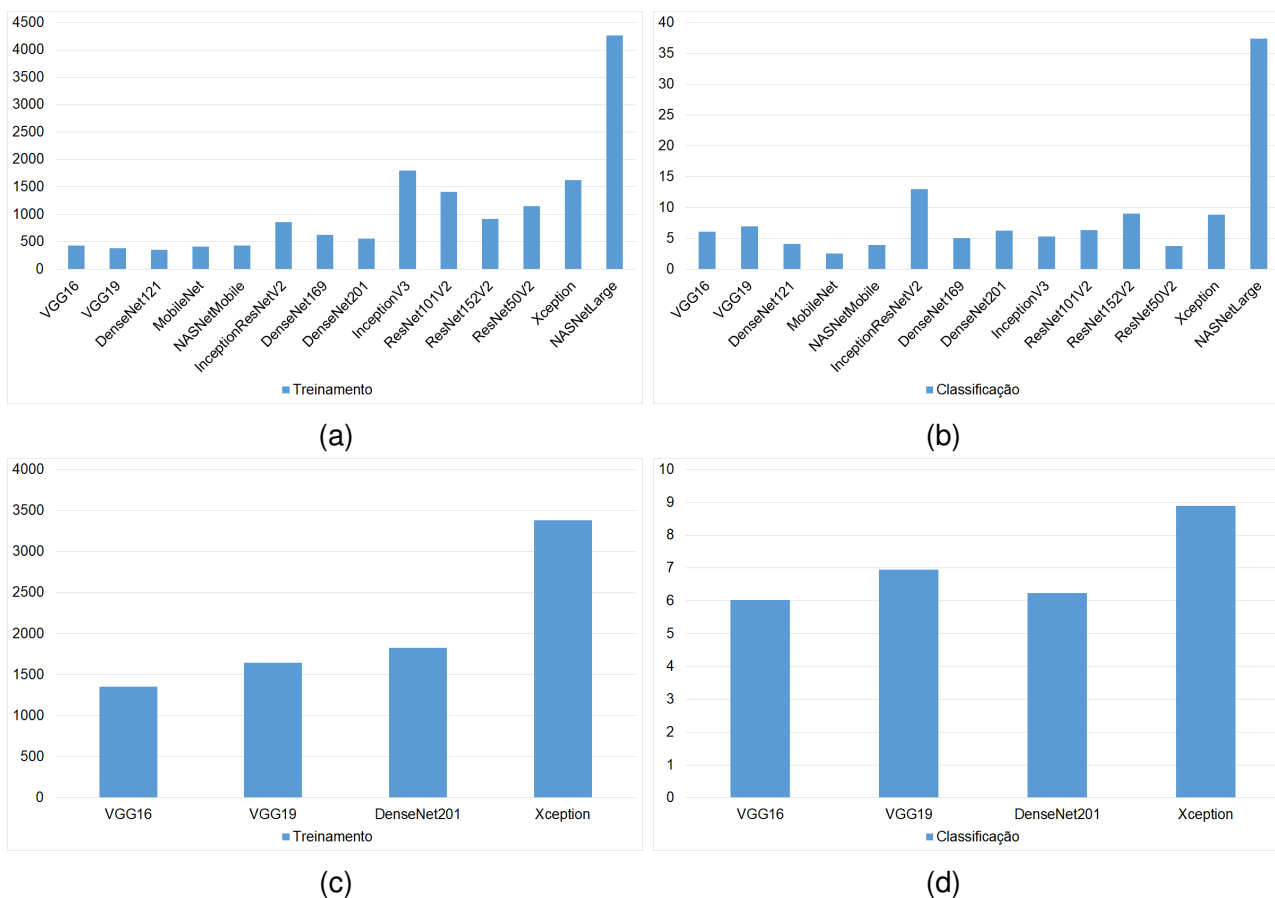


Figura 20: Tempo médio de treinamento e classificação obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados nos conjuntos de dados original e aumentado. (a) e (b) Treinamento e classificação no conjunto de dados original; (c) e (d) Treinamento e classificação no conjunto de dados aumentado.

Na Figura 21 são exibidas as médias das demais métricas (*loss*, sensibilidade, especificidade, precisão, F1 e *kappa*) obtidas no conjunto de dados de teste para a classificação por redes neurais convolucionais (CNN). Essas métricas correspondem aos melhores resultados encontrados, ou seja, as maiores médias de acurácia, conforme destacado na Tabela 6, onde os modelos foram treinados tanto no conjunto de dados original quanto no conjunto de dados aumentado. Essas métricas fornecem uma visão abrangente do desempenho dos modelos em termos de diferentes aspectos de classificação, permitindo avaliar a eficácia das CNNs em relação a essas métricas específicas.

Na Figura 22, é exibida a progressão média das métricas acurácia x *loss* durante o treinamento das redes neurais convolucionais VGG16 e VGG19 no conjunto

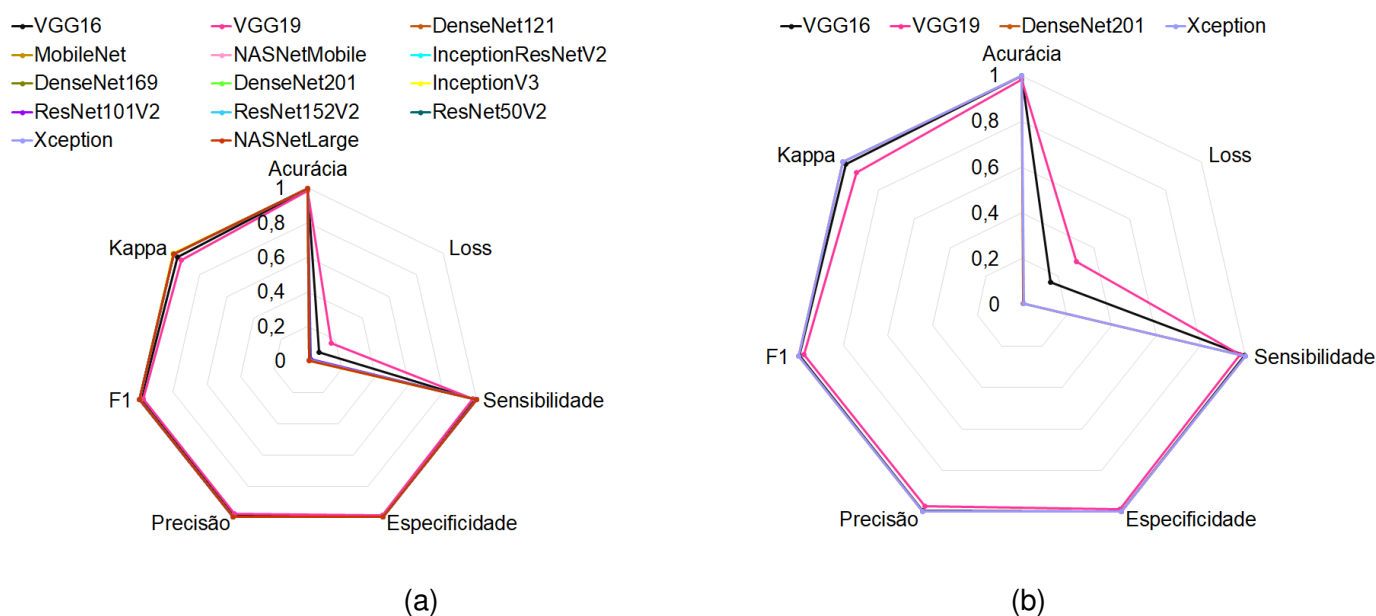


Figura 21: Média das métricas (acurácia, *loss*, sensibilidade, especificidade, precisão, F1 e *kappa*) obtidas no conjunto de dados de teste usando classificação por redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados nos conjuntos de dados original e aumentado. a) Conjunto de dados original; b) Conjunto de dados aumentado.

de dados original. Essa figura representa a evolução dessas métricas ao longo das épocas de treinamento, fornecendo *insights* sobre o desempenho e o progresso das redes neurais durante o processo de aprendizado. Através da análise dessa progressão, é possível avaliar a convergência das redes, identificar tendências e realizar comparações entre as duas arquiteturas, auxiliando na compreensão do desempenho dessas redes na tarefa de classificação.

Apesar dos modelos terem alcançado resultados significantes na classificação de nódulos radiculares de soja, algumas confusões ocorreram, conforme evidenciados pelas duas matrizes de confusão na Figura 23. Especificamente, o modelo VGG16 apresentou predições incorretas, classificando erroneamente amostras de raízes como nódulos e nódulos como raízes. Essas ocorrências destacam a presença de falsos positivos, indicando uma dificuldade do modelo em distinguir corretamente entre as diferentes classes e resultando em classificações equivocadas em situações específicas. No entanto, de forma geral, as redes neurais convolucionais, destacadas com asterisco na Tabela 6, apresentaram resultados promissores para a classificação de nódulos em raízes de culturas de soja.

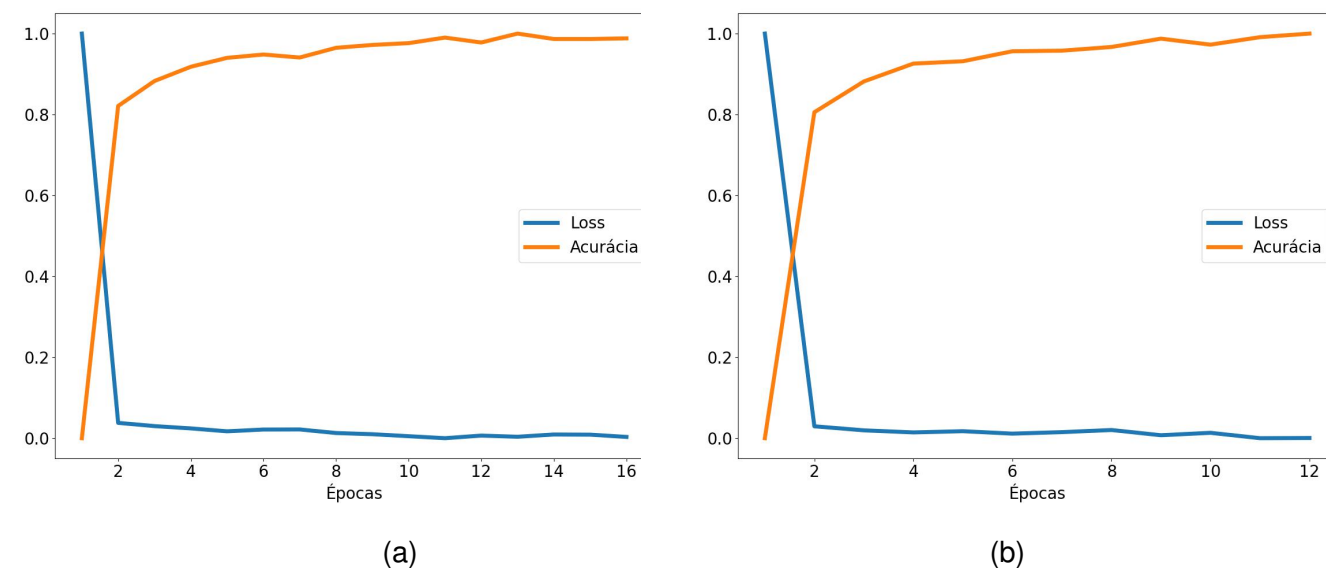


Figura 22: Progressão média das métricas acurácia x *loss* obtidas durante o treinamento das redes neurais convolucionais. Os modelos foram treinados no conjunto de dados original. a) Modelo VGG16; b) Modelo VGG19.

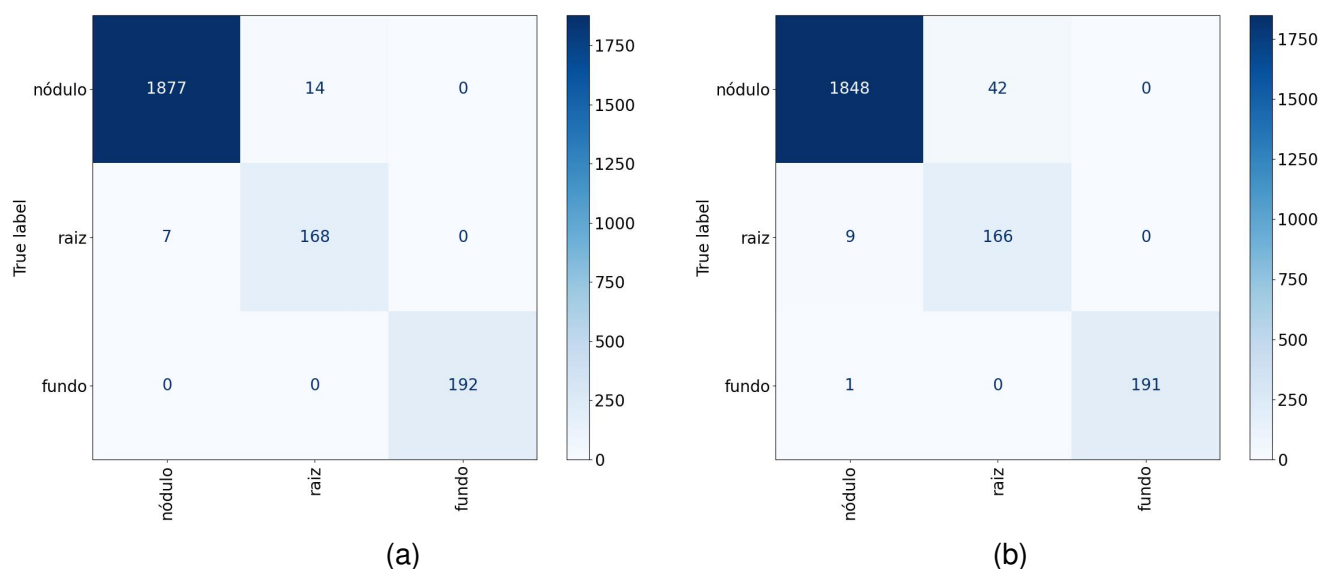


Figura 23: Média das matrizes de confusão obtidas no conjunto de dados de teste para as redes neurais convolucionais VGG16 e VGG19. Os modelos foram treinados no conjunto de dados original. (a) Modelo VGG16; (b) Modelo VGG19.

4.3 Detecção por meio de modelos de visão computacional

Foi realizado também um experimento para detectar nódulos em raízes de cultivares de soja usando modelos de visão computacional, como o YOLO v7 (WANG

et al., 2022), o RetinaNet (LIN et al., 2018) e o Faster R-CNN (REN et al., 2016).

A Tabela 8 apresenta a média da métrica *average precision* com um limiar de 0,5 (AP@0,5), acompanhada do desvio padrão, para cada um dos modelos nos conjuntos de dados de teste em diferentes épocas de treinamento. A tabela revela que as três redes neurais convolucionais tiveram um desempenho melhor em termos da métrica AP@0,5 a medida que o número de épocas de treinamento aumentou, com exceção da CNN YOLO v7 (WANG et al., 2022), que apresentou um desempenho inferior com 1.000 épocas de treinamento em comparação com 800 épocas.

Além disso, ao observar a Tabela 8, nota-se que a arquitetura YOLO v7 (WANG et al., 2022) apresentou um desvio padrão maior em comparação com as arquiteturas RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016). Isso indica uma maior variabilidade em relação à média da métrica AP@0,5 nas execuções da validação cruzada para o modelo YOLO v7 (WANG et al., 2022).

Tabela 8: Média da métrica *average precision* (AP@0,5) \pm desvio padrão obtidos no conjunto de dados de teste pelos modelos de visão computacional YOLO v7, RetinaNet e Faster R-CNN.

Épocas	YOLO v7	RetinaNet	Faster R-CNN
200	0,503 \pm 0,151	0,514 \pm 0,004	0,528 \pm 0,007
400	0,600 \pm 0,120	0,547 \pm 0,004	0,568 \pm 0,009
600	0,720 \pm 0,099	0,566 \pm 0,003	0,588 \pm 0,006
800	0,730 \pm 0,091	0,582 \pm 0,003	0,600 \pm 0,006
1.000	0,700 \pm 0,083	0,590 \pm 0,004	0,617 \pm 0,008

Com o objetivo de avaliar o grau de dispersão em relação à média aritmética da métrica AP@0,5 para a arquitetura YOLO v7 (WANG et al., 2022) em comparação com as arquiteturas RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016), a Tabela 9 apresenta o coeficiente de variação (CV) expresso em percentual. O coeficiente de variação é calculado utilizando a fórmula mostrada em 3:

$$CV = \frac{\text{Desvio Padrão}}{\text{Média Aritmética}} \cdot 100. \quad (3)$$

Tabela 9: Coeficiente de variação (%) da média da métrica *average precision* (AP@0,5) obtidos no conjunto de dados de teste pelos modelos de visão computacional YOLO v7, RetinaNet e Faster R-CNN.

Épocas	YOLO v7	RetinaNet	Faster R-CNN
200	30,020	0,778	1,326
400	20,000	0,731	1,585
600	13,750	0,530	1,020
800	12,466	0,515	1,000
1.000	11,857	0,678	1,297

Ao analisar o coeficiente de variação da métrica AP@0,5, pode-se constatar que houve um maior afastamento da média aritmética para o modelo YOLO v7 (WANG et al., 2022) durante as execuções da validação cruzada. Isso é evidenciado pelos valores mais altos de coeficiente de variação no conjunto de dados de teste para esse modelo. Por exemplo, o menor coeficiente de variação registrado foi de 11,857% para 1.000 épocas de treinamento. Em contrapartida, os modelos RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016), considerando o mesmo número de épocas de treinamento, obtiveram coeficientes de variação de 0,678% e 1,297%, respectivamente.

No entanto, ao observar a Tabela 9, é possível notar que o coeficiente de variação diminuiu a medida que o número de épocas de treinamento aumentou para a CNN YOLO v7 (WANG et al., 2022). Por outro lado, para as CNNs RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016), não houve uma variação significativa no CV da métrica AP@0,5 com o aumento das épocas de treinamento, mantendo-se abaixo de 2%.

Considerando que o modelo YOLO v7 (WANG et al., 2022) teve um desempenho superior aos modelos RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016) em termos da maior métrica AP@0,5. A Figura 24 também apresenta a média aritmética das métricas de *average precision* (AP@0,5), precisão e sensibilidade para o conjunto de dados de teste para o modelo YOLO v7 (WANG et al., 2022). Pode-se observar que o treinamento com 800 épocas resultou melhores resultados para as métricas mencionadas acima.

Para ilustrar a detecção de nódulos nas raízes de soja, a Figura 25 apresenta imagens do conjunto de dados de teste, nas quais o modelo YOLO v7 (WANG et al.,

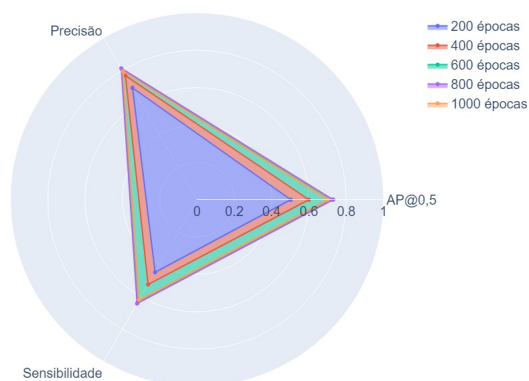


Figura 24: Média das métricas *average precision* (AP@0,5), precisão e sensibilidade obtidas no conjunto de dados de teste pelo modelo de visão computacional YOLO v7.

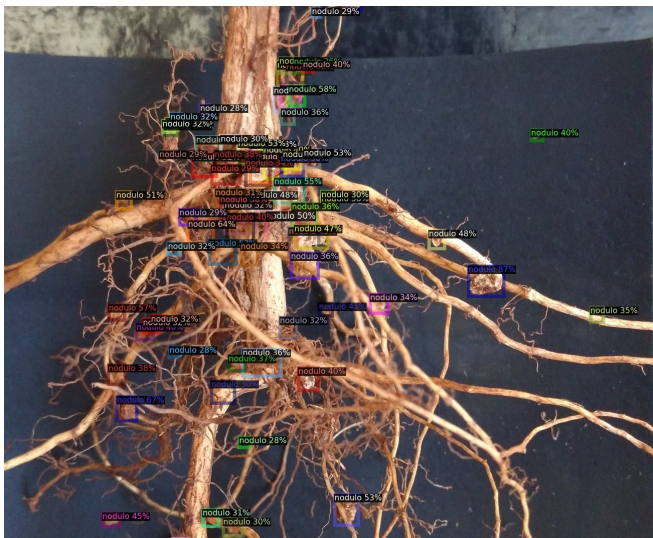
2022), bem como as arquiteturas RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016), detectaram os nódulos. No entanto, ainda é necessário realizar alguns experimentos nos modelos de visão computacional considerados nos experimentos de modo a melhorar a assertividade na detecção.



(a)



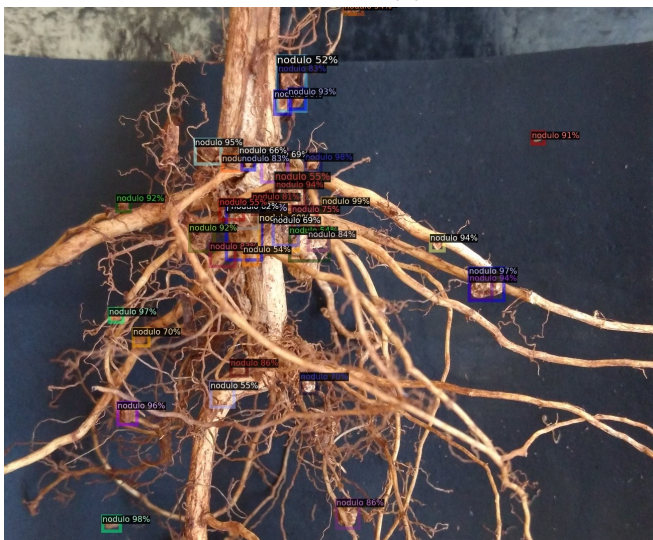
(b)



(c)



(d)



(e)



(f)

Figura 25: Resultados da detecção de nódulos obtidos por cada modelo: (a) e (b) YOLO v7, (c) e (d) RetinaNet e (e) e (f) Faster R-CNN a partir de imagens de raízes de soja contendo nódulos (esquerda) e nódulos isolados das raízes (direita).

5 CONCLUSÃO E TRABALHOS FUTUROS

O objetivo deste estudo foi propor uma metodologia baseada em inteligência computacional, utilizando processamento de imagens digitais e visão computacional, para realizar a análise, classificação e detecção de nódulos radiculares decorrentes da fixação biológica do nitrogênio no cultivo da soja, no qual em um trabalho futuro a eficácia da nodulação em soja poderá ser avaliada, considerando a quantidade de nódulos radiculares.

No presente estudo, foram conduzidos diversos experimentos utilizando técnicas de aprendizado de máquina e aprendizado profundo. Para a tarefa de classificação, foram selecionados modelos clássicos como KNN, RF, SVM (KL) e SVM (KPG2), combinados com extratores de características tradicionais e arquiteturas profundas. Além disso, foram exploradas dezoito redes neurais convolucionais, conforme detalhado na Tabela 2 (Seção 3.4 do Capítulo 3), para realizar a classificação das amostras nos conjuntos de dados.

No contexto da detecção, adotou-se uma abordagem baseada em visão computacional, utilizando os modelos YOLO v7 (WANG et al., 2022), RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016) para detectar automaticamente os nódulos em raízes de cultivares de soja.

Com o objetivo de aumentar o número de amostras nos conjuntos de dados e avaliar a eficácia e robustez dos modelos, foram empregadas técnicas de *data augmentation* e validação cruzada. Essas abordagens contribuíram para gerar uma maior diversidade e representatividade nos dados utilizados tanto para o treinamento quanto para a avaliação dos modelos de classificação.

No caso da abordagem de detecção, foi adotada exclusivamente a estratégia

de validação cruzada, a fim de avaliar o desempenho e robustez dos modelos em detectar os nódulos em raízes de cultivares de soja.

Ao analisar os resultados da classificação utilizando modelos tradicionais treinados com o conjunto de dados original, observou-se que a combinação do descritor tradicional GCH com o classificador RF apresentou o melhor desempenho entre os extratores tradicionais. Por outro lado, os descritores profundos DenseNet121, ResNet101, ResNet152 e ResNet50, em conjunto com os classificadores KNN, RF, SVM (KL) e SVM (KPG2), alcançaram resultados superiores em comparação com outras combinações de descritores profundos e classificadores. Vale ressaltar que as combinações de classificadores e extratores profundos obtiveram acurácias mais altas em comparação aos descritores tradicionais.

No conjunto de dados aumentado, observou-se que o descritor tradicional FOM (Color), combinado com os classificadores RF e SVM (KPG2), obteve o melhor desempenho entre os extratores tradicionais. Da mesma forma, os descritores profundos VGG16, VGG19, MobileNet, MobileNetV2, ResNet101, ResNet152 e ResNet50, combinados com os classificadores KNN, RF, SVM (KL) e SVM (KPG2), apresentaram resultados superiores em relação a outras combinações de descritores profundos e classificadores. Novamente, as combinações de classificadores e extratores profundos resultaram em acurácias mais altas em comparação aos descritores tradicionais.

Além disso, observou-se que o treinamento utilizando o conjunto de dados aumentado resultou em uma maior variedade de combinações de descritores e classificadores com melhores resultados, embora esses resultados fossem semelhantes aos obtidos no treinamento com o conjunto de dados original. É interessante notar que os descritores profundos ResNet101, ResNet152 e ResNet50 mostraram-se eficientes para a classificação de nódulos em raízes de soja, tanto no treinamento com o conjunto de dados original quanto no treinamento com o conjunto de dados aumentado.

Ao analisar a abordagem de classificação utilizando redes neurais convolucionais, observou-se que a maioria delas obteve resultados satisfatórios na classificação das amostras no conjunto de dados original. No entanto, ao considerar

o conjunto de dados aumentado, as redes neurais convolucionais VGG16, VGG19, DenseNet201 e Xception se destacaram, apresentando as maiores médias de acurácia. Isso evidencia a eficácia dessas redes neurais convolucionais na tarefa de classificação quando o conjunto de dados é expandido por meio da técnica de aumento de dados.

É relevante destacar que as demais redes neurais convolucionais obtiveram desempenho ligeiramente inferior nessas condições específicas, o que levou a uma seleção limitada de redes neurais convolucionais em comparação com os modelos treinados no conjunto de dados original.

No contexto da abordagem de detecção, as três redes neurais convolucionais, YOLO v7 (WANG et al., 2022), RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016), demonstraram um desempenho aprimorado em termos da métrica AP@0,5 a medida que o número de épocas de treinamento aumentou. Uma exceção foi a CNN YOLO v7 (WANG et al., 2022), que apresentou um desempenho inferior com 1.000 épocas de treinamento em comparação com 800 épocas. Além disso, observa-se que a arquitetura YOLO v7 (WANG et al., 2022) exibiu um desvio padrão mais elevado em comparação com as arquiteturas RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016). Isso indica uma maior variabilidade em relação à média da métrica AP@0,5 nas execuções da validação cruzada para o modelo YOLO v7 (WANG et al., 2022). No entanto, é possível notar que o coeficiente de variação diminuiu a medida que o número de épocas de treinamento aumentou para a CNN YOLO v7 (WANG et al., 2022).

Em conclusão, os resultados obtidos nas abordagens de classificação de nódulos em raízes de soja, tanto por modelos tradicionais de aprendizado de máquina quanto por redes neurais convolucionais, foram promissores.

No entanto, na abordagem de detecção, ainda há espaço para melhorias nos modelos de visão computacional YOLO v7 (WANG et al., 2022), RetinaNet (LIN et al., 2018) e Faster R-CNN (REN et al., 2016), a fim de aprimorar a precisão na detecção dos nódulos radiculares. Para trabalhos futuros, recomenda-se dedicar mais tempo a esses modelos, explorando diferentes estratégias. Uma possibilidade seria aumentar o número de imagens anotadas disponíveis no conjunto de dados,

possibilitando um treinamento mais abrangente e uma melhor adaptação dos modelos aos padrões dos nódulos. Além disso, é importante realizar experimentos adicionais para explorar e ajustar os hiperparâmetros dos modelos, visando obter uma maior precisão na detecção. A aplicação de um limiar de confiança nas detecções também pode ser considerada, a fim de filtrar possíveis falsos positivos e melhorar a acurácia geral do sistema. Além disso, é recomendado prospectar outros modelos de visão computacional adequados para a detecção de nódulos, explorando diferentes arquiteturas e técnicas avançadas disponíveis na literatura.

No geral, essas descobertas destacam o potencial da inteligência computacional, do processamento de imagens e da visão computacional para auxiliar na análise, classificação e detecção de nódulos radiculares em cultivos de soja. A continuidade da pesquisa e aprimoramento das abordagens utilizadas podem contribuir para o desenvolvimento de métodos mais precisos e confiáveis na avaliação da nodulação e eficácia da fixação biológica de nitrogênio na soja.

REFERÊNCIAS

- AGROLINK. A Soja: História, Tendências e Virtudes. **Revista Funcionais e Nutracêuticos**, São Paulo, p. 28–40, 2007.
- ALCÂNTARA, R. M. C. M. de; XAVIER, G. R.; ZILLI, J. É.; MARTINS, L. M. V.; RUMJANEK, N. G.; JÚNIOR, P. I. F.; LEITE, J.; JÚNIOR, E. B. d. S.; SANTOS, C. E. d. R. e. S. Fixação Biológica de Nitrogênio. In: **Feijão-Caupi: o produtor pergunta, a Embrapa responde**. 1. ed. Brasília: Embrapa, 2017. cap. 8, p. 95–105.
- ANJOS, J. R. L. dos. **InferenCNN: Uma biblioteca para Inferência de Redes Neurais Convolucionais Multiplataforma em OpenCL**. 87 p. Dissertação (Mestrado em Ciência da Computação) — Universidade Federal de Pernambuco, 2019.
- ARULPRAKASH, E.; ARULDOSS, M. A study on generic object detection with emphasis on future research directions. **Journal of King Saud University - Computer and Information Sciences**, v. 34, n. 9, p. 7347–7365, 2022.
- ARVOR, D.; MEIRELLES, M. S. P.; VARGAS, R.; SKORUPA, L. A.; FIDALGO, E. C. C.; DUBREUIL, V.; HERLIN, I.; BERROIR, J.-P. Monitoring land use changes around the indigenous lands of the xingu basin in mato grosso, brazil. In: **2010 IEEE International Geoscience and Remote Sensing Symposium**. Honolulu: IEEE, 2010. p. 3190–3193.
- BARELLI, F. **Introdução à Visão Computacional: Uma abordagem prática com Python e OpenCV**. São Paulo: Casa do Código, 2018.
- BARRETO, C. Os impactos socioambientais da soja no Brasil. In: **II Encontro da Associação Nacional de Pós Graduação e Pesquisa em Ambiente e Sociedade**. Indaiatuba: II, 2004.
- BHUYAN, M. K. **Computer Vision and Image Processing: Fundamentals and Applications**. Boca Raton: CRC Press, 2019.
- BIGATON, A. **Contribuição dos Fertilizantes para o Desempenho Econômico e Ambiental do Etanol de Cana-de-Açúcar**. 98 p. Dissertação (Mestrado em Planejamento de Sistemas Energéticos) — Faculdade de Engenharia Mecânica - Universidade Estadual de Campinas, 2017.
- BOCCATO, L. **Rede MLP: Perceptron de Múltiplas Camadas**. 2023. Disponível em: <https://www.dca.fee.unicamp.br/~lboccatto/topico_4.1_redes_MLP.pdf>.
- BONATO, E. R.; VELLO, N. A. Aspectos genéticos do tempo para o florescimento em variantes naturais de soja. **Pesquisa Agropecuária Brasileira**, Brasília, p. 988–993, 1999.

BRESSAN, R. S.; ALVES, D. H. A.; VALERIO, L. M.; BUGATTI, P. H.; SAITO, P. T. M. Doctor: The role of deep features in content-based mammographic image retrieval. In: **2018 IEEE 31st Intl Symposium on Computer-Based Medical Systems**. [S.l.: s.n.], 2018. p. 158–163.

BRILHADOR, A.; COLONHEZI, T. P.; BUGATTI, P. H.; LOPES, F. M. Combining texture and shape descriptors for bioimages classification: A case of study in imageclef dataset. In: RUIZ-SHULCLOPER, J.; BAJA, G. Sanniti di (Ed.). **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. p. 431–438.

BRILHADOR, A.; SERRARENS, D. A.; LOPES, F. M. A computer vision approach for automatic measurement of the inter-plant spacing. In: PARDO, A.; KITTLER, J. (Ed.). **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications**. Cham: Springer International Publishing, 2015. p. 219–227.

BRUM, A. L.; DALFOVO, W. C. T.; AZUAGA, F. L. Alguns impactos da expansão da produção de soja no município de sorriso-mt. **Desenvolvimento em Questão**, v. 7, n. 14, p. 173–200, 2011.

CABEZAS, W. A. R. L.; TRIVELIN, P. C. O.; KONDÖRFER, G. H.; PEREIRA, S. Balanço da adubação nitrogenada sólida e fluida de cobertura na cultura de milho, em sistema plantio direto no Triângulo Mineiro (MG). **Revista Brasileira de Ciência do Solo**, v. 24, n. 2, p. 363–376, 2000.

CARVALHO, A. P. de Leon F. de. **Redes Neurais Artificiais**. 2023. Disponível em: <<https://sites.icmc.usp.br/andre/research/neural/>>.

CASTRO, C. d.; BALLA, A.; CASTIGLIONI, V. B. R.; SFREDO, G. J. Doses e métodos de aplicação de nitrogênio em girassol. **Scientia Agricola**, Piracicaba, v. 56, n. 4, p. 827–833, 1999.

CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. **2017 IEEE Conf. on Computer Vision and Pattern Recognition**, p. 1800–1807, 2017.

CHORAS, R. S. Image feature extraction techniques and their applications for cbir and biometrics systems. **International Journal of Biology and Biomedical Engineering**, v. 1, p. 6–16, 2007.

CHUNG, Y. S.; LEE, U.; HEO, S.; SILVA, R. R.; NA, C.-I.; KIM, Y. Image-based machine learning characterizes root nodule in soybean exposed to silicon. **Frontiers in Plant Science**, v. 11, 2020. Disponível em: <<https://www.frontiersin.org/article/10.3389/fpls.2020.520161>>.

CONTE, O.; POSSAMAI, E. J.; FILHO, P. C. **Resultados do monitoramento integrado da colheita da soja na safra 2019/2020 no Paraná**. Londrina, 2020.

COPPIN, B. **Inteligência Artificial**. Rio de Janeiro: LTC, 2013. 628 p.

COSTA, L. da F.; CESAR, J. R. M. **Shape Classification and Analysis: Theory and Practice**. 2. ed. Boca Raton: CRC Press, 2009.

COULIBALY, S.; KAMSU-FOGUEM, B.; KAMISSOKO, D.; TRAORE, D. Deep learning for precision agriculture: A bibliometric analysis. **Intelligent Systems with Applications**, v. 16, 2022.

DANSO, S. A.; LIPING, S.; HU, D.; AFOAKWA, S.; BADZONGOLY, E. L.; ODOOM, J.; MUHAMMAD, O.; MUSHTAQ, M. U.; QAYOOM, A.; ZHOU, W. An optimal defect recognition security-based terahertz low resolution image system using deep learning network. **Egyptian Informatics Journal**, v. 24, n. 3, p. 100384, 2023.

de Lima, G. V.; SAITO, P. T.; LOPES, F. M.; BUGATTI, P. H. Classification of texture based on bag-of-visual-words through complex networks. **Expert Systems with Applications**, v. 133, p. 215–224, 2019. ISSN 0957-4174.

de Lima, G. V. L.; CASTILHO, T. R.; BUGATTI, P. H.; SAITO, P. T. M.; LOPES, F. M. A complex network-based approach to the analysis and classification of images. In: PARDO, A.; KITTLER, J. (Ed.). **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications**. Cham: Springer International Publishing, 2015. p. 322–330.

DENG, J.; DONG, W.; SOCHER, R.; LI, L.-J.; LI, K.; FEI-FEI, L. Imagenet: A large-scale hierarchical image database. In: **2009 IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2009. p. 248–255.

DIAS, V. P.; FERNANDES, E. Fertilizantes: uma visão global sintética. **BNDES Setorial**, n. 24, p. 137–138, 2006.

DREONI, I.; MATTHEWS, Z.; SCHAAFSMA, M. The impacts of soy production on multi-dimensional well-being and ecosystem services: A systematic review. **Journal of Cleaner Production**, v. 335, 2022.

DURAI, S. K. S.; SHAMILI, M. D. Smart farming using Machine Learning and Deep Learning techniques. **Decision Analytics Journal**, v. 3, 2022.

EGEWARTH, V. A. **Características agrônômicas, adaptabilidade e estabilidade produtiva de genótipos de soja**. 117 p. Tese (Doutorado em Agronomia) — Universidade Estadual do Oeste do Paraná, 2019.

EMBRAPA. **INCT MPCP-AGRO**. 2022. Disponível em: <<https://www.microagro.com.br>>.

FACELI, K.; LORENA, A. C.; GAMA, J.; ALMEIDA, T. A. de; CARVALHO, A. C. P. L. F. de. **Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina**. 2. ed. Rio de Janeiro: LTC, 2021.

FAQUIN, V. **Nutrição Mineral de Plantas**. Lavras: UFLA / FAEPE, 2005.

FILHO, O. M.; NETO, H. V. **Processamento Digital de Imagens**. Rio de Janeiro: Brasport, 1999.

FLORENCIO, V. H. R. **Caracterização de variedades cultivadas de soja de diferentes grupos de maturação em função dos atributos morfológicos, fenológicos e a produtividade**. 68 p. Dissertação (Mestrado em Fitotecnia) — Escola Superior de Agricultura Luiz de Queiroz - Universidade de São Paulo, 2017.

GERON, A. **Mãos à Obra Aprendizado de Máquina com Scikit-Learn TensorFlow: Conceitos, Ferramentas e Técnicas Para a Construção de Sistemas Inteligentes**. Rio de Janeiro: Alta Books, 2019.

GOMES, M. S. **Fertilização Nitrogenada: produção, produtividade e trocas gasosas da soja cultivada em latossolo amarelo em casa de vegetação**. 70 p. Dissertação (Mestrado em Agronomia) — Universidade Federal Rural da Amazônia, 2019.

GONZALEZ, R. C.; WOODS, R. C. **Processamento Digital de Imagens**. 3. ed. São Paulo: Pearson Prentice Hall, 2010.

GRUS, J. **Data Science do Zero: Primeiras Regras com o Python**. Rio de Janeiro: Alta Books, 2016.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, n. 6, p. 610–621, 1973.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. **2016 IEEE Conf. on Computer Vision and Pattern Recognition**, p. 770–778, 2016.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Identity mappings in deep residual networks. **ArXiv**, abs/1603.05027, 2016.

HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. **ArXiv**, abs/1704.04861, 2017.

HU, M.-K. Visual pattern recognition by moment invariants. **IRE Transactions on Information Theory**, v. 8, n. 2, p. 179–187, 1962.

HUANG, G.; LIU, Z.; WEINBERGER, K. Q. Densely connected convolutional networks. **2017 IEEE Conf. on Computer Vision and Pattern Recognition**, p. 2261–2269, 2017.

HUNGRIA, M.; CAMPO, R. J.; MENDES, I. C. **Fixação Biológica do Nitrogênio na Cultura da Soja**. Londrina, 2001. v. 35, n. 13, 48 p.

HUNGRIA, M.; MENDES, I. C. Nitrogen fixation with soybean: The perfect symbiosis? In: _____. **Biological Nitrogen Fixation**. John Wiley Sons, Ltd, 2015. cap. 99, p. 1009–1024. ISBN 9781119053095. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119053095.ch99>>.

HUNGRIA, M.; MORETTI, L. G.; CRUSCIOL, C. A. C.; KURAMAE, E. E.; BOSSOLANI, J. W.; MOREIRA, A.; COSTA, N. R.; ALVES, C. J.; PASCOALOTO, I. M.; RONDINA, A. B. L. Effects of growth-promoting bacteria on soybean root activity, plant development, and yield. **Agronomy Journal**, v. 112, n. 1, p. 418–428, 2020. Disponível em: <<https://access.onlinelibrary.wiley.com/doi/abs/10.1002/agj2.20010>>.

HURBANS, R. **Grokking Artificial Intelligence Algorithms**. Shelter Island: Manning, 2020.

IBGE. **Levantamento Sistemático da Produção Agrícola: Tabelas**. 2023. Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9201-levantamento-sistemico-da-producao-agricola.html?=&t=resultados>>.

IBM. **O que são Redes Neurais?** 2023. Disponível em: <<https://www.ibm.com/br-pt/topics/neural-networks>>.

IRONS, J. R.; PETERSEN, G. W. Texture transforms of remote sensing data. **Remote Sensing of Environment**, v. 11, p. 359–370, 1981. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/003442578190033X>>.

IZBICKI, R.; SANTOS, T. M. dos. **Aprendizado de máquina: uma abordagem estatística**. São Carlos: [s.n.], 2020.

JUBERY, T. Z.; CARLEY, C. N.; SINGH, A.; SARKAR, S.; GANAPATHYSUBRAMANIAN, B.; SINGH, A. K. Using machine learning to develop a fully automated soybean nodule acquisition pipeline (snap). **Plant Phenomics**, v. 2021, 2021. Disponível em: <<https://doi.org/10.34133/2021/9834746>>.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. Imagenet classification with deep convolutional neural networks. **Neural Information Processing Systems**, v. 25, 01 2012.

KUMAR, L. A.; RENUKA, D. K. **Deep Learning Approach for Natural Language Processing, Speech, and Computer Vision: Techniques and Use Cases**. Boca Raton: CRC Press, 2023.

KUNAVAR, M.; TASIC, J. Image feature extraction - an overview. In: **EUROCON 2005 - The International Conference on "Computer as a Tool"**. [S.l.: s.n.], 2005. v. 1, p. 183–186.

LECUN, Y.; BOSER, B.; DENKER, J.; HENDERSON, D.; HOWARD, R.; HUBBARD, W.; JACKEL, L. Handwritten digit recognition with a back-propagation network. In: **Advances in Neural Information Processing Systems**. [S.l.]: Morgan-Kaufmann, 1989. v. 2.

LENZ, M. L.; NEUMANN, F. B.; SANTARELLI, R.; SALVADOR, D. **Fundamentos de Aprendizagem de Máquina**. Porto Alegre: SAGAH, 2020. 311 p.

LIN, T.-Y.; GOYAL, P.; GIRSHICK, R.; HE, K.; DOLLÁR, P. **Focal Loss for Dense Object Detection**. 2018.

LIN, T.-Y.; GOYAL, P.; GIRSHICK, R. B.; HE, K.; DOLLÁR, P. Focal loss for dense object detection. **2017 IEEE Intl Conf. on Computer Vision**, p. 2999–3007, 2017.

LIN, T.-Y.; MAIRE, M.; BELONGIE, S.; BOURDEV, L.; GIRSHICK, R.; HAYS, J.; PERONA, P.; RAMANAN, D.; ZITNICK, C. L.; DOLLÁR, P. **Microsoft COCO: Common Objects in Context**. 2015.

LIU, L.; OUYANG, W.; WANG, X.; FIEGUTH, P.; CHEN, J.; LIU, X.; PIETIKÄINEN, M. Deep Learning for Generic Object Detection: A Survey. **International Journal of Computer Vision**, v. 128, 2020.

LIU, W.; ANGUELOV, D.; ERHAN, D.; SZEGEDY, C.; REED, S.; FU, C.-Y.; BERG, A. C. SSD: Single shot MultiBox detector. In: **Computer Vision – ECCV 2016**. [S.l.]: Springer International Publishing, 2016. p. 21–37.

LOPES, F. M. **Um modelo perceptivo de limiarização de imagens digitais**. 129 p. Dissertação (Mestrado em Informática) — Universidade Federal do Paraná, 2003.

LOPES, F. M.; CONSULAR, L. A. Automatic gray level thresholding based on relevant features for subjective decisions. In: **Proceedings. XV Brazilian Symposium on Computer Graphics and Image Processing**. [S.l.: s.n.], 2002. p. 411.

LOPES, F. M.; CONSULARO, L. A. A rbfm perceptivo model for image thresholding. In: **XVIII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAP'05)**. [S.l.: s.n.], 2005. p. 225–232.

LU, Y.; YOUNG, S. A survey of public datasets for computer vision tasks in precision agriculture. **Computers and Electronics in Agriculture**, v. 178, 2020.

MALAVOLTA, E. **Manual de calagem e adubação das principais culturas**. São Paulo: Ceres, 1987.

MAPA. **Plano ABC - Agricultura de Baixa Emissão de Carbono**. 2011. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/sustentabilidade/plano-abc/plano-abc-agricultura-de-baixa-emissao-de-carbono>>.

MARQUES, J. A. L.; GOIS, F. N. B.; MADEIRO, J. P. do V.; LI, T.; FONG, S. J. Chapter 4 - Artificial neural network-based approaches for computer-aided disease diagnosis and treatment. In: BHOI, A. K.; de Albuquerque, V. H. C.; SRINIVASU, P. N.; MARQUES, G. (Ed.). **Cognitive and Soft Computing Techniques for the Analysis of Healthcare Data**. [S.l.]: Academic Press, 2022, (Intelligent Data-Centric Systems). p. 79–99.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, 1943.

MEDEIROS, A. D. d.; CAPOBIANGO, N. P.; SILVA, J. M. d.; SILVA, L. J. d.; SILVA, C. B. d.; DIAS, D. C. F. d. S. Interactive machine learning for soybean seed and seedling quality classification. **Scientific Reports**, v. 10, 2020. Disponível em: <<https://doi.org/10.1038/s41598-020-68273-y>>.

MENDES, I. de C.; JUNIOR, F. B. dos R.; PERES, J. R. R.; VARGAS, M. A. T.; SUHET, A. R. Embrapa Cerrados: 37 anos de contribuições para o avanço da FBN no Brasil. In: **Anais da XVI RELARE - Reunião da Rede de Laboratórios para Recomendação, Padronização e Difusão de Tecnologia de Inoculantes Microbianos de Interesse Agrícola**. Londrina: 16, 2014. p. 60–61.

MESHARAM, V.; PATIL, K.; MESHARAM, V.; HANCHATE, D.; RAMKTEKE, S. Machine learning in agriculture domain: A state-of-art survey. **Artificial Intelligence in the Life Sciences**, v. 1, p. 100010, 2021. ISSN 2667-3185. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2667318521000106>>.

MICHALSKI, R.; CHILAUSSKY, R. "learning by being told and learning from example: An experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis". **International Journal of Policy Analysis and Information Systems**, v. 4, 01 1980.

MORABIA, K.; ARORA, J.; VIJAYKUMAR, T. **Attention-based Joint Detection of Object and Semantic Part**. 2020.

MORETI, M. P.; OLIVEIRA, T.; SARTORI, R.; CAETANO, W. Inteligência artificial no agronegócio e os desafios para a proteção da propriedade intelectual. **Cadernos de Prospecção**, v. 14, n. 1, p. 60, jan. 2021. Disponível em: <<https://periodicos.ufba.br/index.php/nit/article/view/33098>>.

MORETTI, L. G.; CRUSCIOL, C. A. C.; BOSSOLANI, J. W.; MOMESSO, L.; GARCIA, A.; KURAMAE, E. E.; HUNGRIA, M. Bacterial Consortium and Microbial Metabolites Increase Grain Quality and Soybean Yield. **Journal of Soil Science and Plant Nutrition**, v. 20, 2020.

NOBREGA, R. V. M. d.; PEIXOTO, S. A.; SILVA, S. P. P. da; FILHO, P. P. R. Lung nodule classification via deep transfer learning in ct lung images. In: **2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.: s.n.], 2018. p. 244–249.

NOGUEIRA, M. A.; PRANDO, A. M.; OLIVEIRA, A. B. de; LIMA, D. de; CONTE, O.; HARGER, N.; OLIVEIRA, F. T. de; HUNGRIA, M. **Ações de transferência de tecnologia em inoculação/coinoculação com Bradyrhizobium e Azospirillum na cultura da soja na safra 2017/18 no estado do Paraná**. Londrina, 2018. 15 p.

OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Gray scale and rotation invariant texture classification with local binary patterns. In: **Computer Vision**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000. p. 404–420. ISBN 978-3-540-45054-2.

OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, v. 29, n. 1, p. 51–59, 1996. ISSN 0031-3203. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0031320395000674>>.

O'SHEA, K.; NASH, R. **An Introduction to Convolutional Neural Networks**. 2015.

PACANHELA, E.; RONDINA, A.; NOGUEIRA, M.; HUNGRIA, M.; SAITO, P.; LOPES, F. Análise e classificação automática de nódulos em raízes de cultivares de soja. In: **Anais do XVI Brazilian e-Science Workshop**. Porto Alegre, RS, Brasil: SBC, 2022. p. 41–48. ISSN 2763-8774.

PAREKH, R. **Fundamentals of image, audio, and video processing using MATLAB with applications to pattern recognition**. Boca Raton: CRC Press, 2021.

PATEL, J. M.; GAMIT, N. C. A review on feature extraction techniques in content based image retrieval. In: **2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)**. [S.l.: s.n.], 2016. p. 2259–2263.

PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais : princípios, algoritmos e aplicações**. São Paulo: Cengage Learning, 2008.

PEREIRA, D.; BUGATTI, P.; LOPES, F.; SOUZA, A. de; SAITO, P. Isoy - sistema inteligente para melhorias na produtividade e sustentabilidade em empresas beneficiadoras de sementes de soja. In: **Anais do XV Simpósio Brasileiro de Sistemas de Informação**. Porto Alegre, RS, Brasil: SBC, 2019. p. 439–446. ISSN 0000-0000.

PEREIRA, D. F.; BUGATTI, P. H.; LOPES, F. M.; SOUZA, A. L.; SAITO, P. T. Contributing to agriculture by using soybean seed data from the tetrazolium test. **Data in Brief**, v. 23, p. 103652, 2019. ISSN 2352-3409.

PEREIRA, D. F.; BUGATTI, P. H.; LOPES, F. M.; SOUZA, A. L. S. M. de; SAITO, P. T. M. An intelligent system to enhance the productivity and sustainability in soybean crop enterprises. In: **Proceedings of the XV Brazilian Symposium on Information Systems**. New York, NY, USA: Association for Computing Machinery, 2019. (SBSI'19). ISBN 9781450372374. Disponível em: <<https://doi.org/10.1145/3330204.3330266>>.

PEREIRA, D. F.; BUGATTI, P. H.; LOPES, F. M.; SOUZA, A. L. S. M. de; SAITO, P. T. M. Assessing active learning strategies to improve the quality control of the soybean seed vigor. **IEEE Transactions on Industrial Electronics**, v. 68, n. 2, p. 1675–1683, 2021.

PINHO, C. A soja é pop. **Istoé**, São Paulo, 1796, p. 80–81, 2004.

PRECHELT, L. Early stopping — but when? In: _____. **Neural Networks: Tricks of the Trade: Second Edition**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 53–67.

REDMON, J.; DIVVALA, S.; GIRSHICK, R.; FARHADI, A. **You Only Look Once: Unified, Real-Time Object Detection**. 2016.

REFAEILZADEH, P.; TANG, L.; LIU, H. Cross-validation. In: _____. **Encyclopedia of Database Systems**. Boston, MA: Springer US, 2009. p. 532–538. ISBN 978-0-387-39940-9. Disponível em: <https://doi.org/10.1007/978-0-387-39940-9_565>.

REN, M.; ZHANG, X.; CHEN, X.; ZHOU, B.; FENG, Z. YOLOv5s-M: A deep learning network model for road pavement damage detection from urban street-view imagery. **International Journal of Applied Earth Observation and Geoinformation**, v. 120, p. 103335, 2023.

REN, S.; HE, K.; GIRSHICK, R.; SUN, J. **Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks**. 2016.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: **MICCAI**. [S.l.: s.n.], 2015.

ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. **Psychological Review**, 1958.

SAMMUT, C.; WEBB, G. I. **Encyclopedia of Machine Learning and Data Mining**. 2. ed. Boston: Springer, 2017.

SANDLER, M.; HOWARD, A. G.; ZHU, M.; ZHMOGINOV, A.; CHEN, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. **2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition**, p. 4510–4520, 2018.

SANTOS, V. S. dos. **Neurônio**. 2023. Disponível em: <<https://mundoeducacao.uol.com.br/biologia/neuronios.htm>>.

SHAHWAR, D.; MUSHTAQ, Z.; MUSHTAQ, H.; ALQARAWI, A. A.; PARK, Y.; ALSHAHRANI, T. S.; FAIZAN, S. Role of microbial inoculants as bio fertilizers for improving crop productivity: A review. **Heliyon**, v. 9, n. 6, 2023.

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. **Journal of Big Data**, p. 2196–1115, 2019.

SILVA, T. R. B. da; REIS, A. C. de S.; MACIEL, C. D. de G. Relationship between chlorophyll meter readings and total n in crambe leaves as affected by nitrogen topdressing. **Industrial Crops and Products**, v. 39, p. 135–138, 2012.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **CoRR**, abs/1409.1556, 2015.

SINGH, H. **Practical Machine Learning and Image Processing: For Facial Recognition, Object Detection, and Pattern Recognition Using Python**. Berkeley: Apress, 2019.

STEHLING, R. O.; NASCIMENTO, M. A.; aO, A. X. F. A compact and efficient image retrieval approach based on border/interior pixel classification. In: **Proceedings of the Eleventh International Conference on Information and Knowledge Management**. New York, NY, USA: Association for Computing Machinery, 2002. (CIKM '02), p. 102–109. ISBN 1581134924. Disponível em: <<https://doi.org/10.1145/584792.584812>>.

STRICKER, M. A.; ORENGO, M. Similarity of color images. In: NIBLACK, W.; JAIN, R. C. (Ed.). **Storage and Retrieval for Image and Video Databases III**. SPIE, 1995. v. 2420, p. 381 – 392. Disponível em: <<https://doi.org/10.1117/12.205308>>.

SU, R.; YUE, J.; LI, Z.; JIA, S.; SHENG, G. Detection and counting method of juvenile abalones based on improved SSD network. **Information Processing in Agriculture**, 2023.

SWAIN, M. J.; BALLARD, D. H. Color indexing. **International Journal of Computer Vision**, n. 7, p. 1573–1405, 1991.

SZEGEDY, C.; IOFFE, S.; VANHOUCKE, V.; ALEMI, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. In: **AAAI**. [S.l.: s.n.], 2017.

SZEGEDY, C.; VANHOUCKE, V.; IOFFE, S.; SHLENS, J.; WOJNA, Z. Rethinking the inception architecture for computer vision. **2016 IEEE Conf. on Computer Vision and Pattern Recognition**, p. 2818–2826, 2016.

TAIZ, L.; ZEIGER, E.; MØLLER, I. M.; MURPHY, A. **Fisiologia e desenvolvimento vegetal**. 6. ed. Porto Alegre: Artmed, 2017.

TEAGUE, M. R. Image analysis via the general theory of moments*. **J. Opt. Soc. Am.**, OSA, v. 70, n. 8, p. 920–930, Aug 1980. Disponível em: <<http://opg.optica.org/abstract.cfm?URI=josa-70-8-920>>.

TELLES, T. S.; NOGUEIRA, M. A.; HUNGRIA, M. Economic value of biological nitrogen fixation in soybean crops in Brazil. **Environmental Technology Innovation**, v. 31, 2023.

THEODORIDIS, S. **Machine Learning: A Bayesian and Optimization Perspective**. 2. ed. Cambridge: Academic Press, 2020.

TRASK, A. W. **Grokking Deep Learning**. Shelter Island: MANNING, 2019.

UGAIL, H. **Deep Learning in Visual Computing: Explanations and Examples**. Boca Raton: CRC Press, 2022.

USDA. **United States Departamento of Agriculture: PSD Reports**. 2023. Disponível em: <<https://apps.fas.usda.gov/psdonline/app/index.html/app/downloads>>.

VICERI. **Arquiteturas de Redes Neurais Convolucionais para reconhecimento de imagens**. 2023. Disponível em: <<https://viceri.com.br/insights/arquiteturas-de-redes-neurais-convolucionais-para-reconhecimento-de-imagens/>>.

VILLALBA, H. A. G.; LEITE, J. M.; OTTO, R.; TRIVELIN, P. C. O. Fertilizantes nitrogenados: Novas tecnologias. **Informações Agronômicas**, n. 148, p. 12–20, 2014.

WADA, K. **LabelMe**. 2022. Disponível em: <<https://github.com/wkentaro/labelme>>.

WANG, C.-Y.; BOCHKOVSKIY, A.; LIAO, H.-Y. M. **YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors**. 2022.

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data mining: practical machine learning tools and techniques**. 3. ed. Burlington: Morgan Kaufmann, 2011.

WWF. **O crescimento da soja: impactos e soluções**. Gland, Suíça: WWF, 2014. 96 p.

YADAV, N.; GARG, V. K.; CHHILLAR, A. K.; RANA, J. S. Recent advances in nanotechnology for the improvement of conventional agricultural systems: A Review. **Plant Nano Biology**, 2023.

ZOPH, B.; VASUDEVAN, V.; SHLENS, J.; LE, Q. V. Learning transferable architectures for scalable image recognition. **2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition**, p. 8697–8710, 2018.

APÊNDICE A – TABELAS E FIGURAS COMPLEMENTARES

Tabela 10: Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste pelo classificador KNN, utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados nos conjuntos de dados original e aumentado.

	Descritor	KNN					
		Conjunto de dados original			Conjunto de dados aumentado		
		Nódulo	Raiz	Fundo	Nódulo	Raiz	Fundo
Tradicionais	Momentos Invariantes de Hu	0,941 \pm 0,002	0,924 \pm 0,002	0,925 \pm 0,002	0,841 \pm 0,003	0,841 \pm 0,004	0,861 \pm 0,003
	FOM (Gray)	0,975 \pm 0,002	0,972 \pm 0,002	0,994 \pm 0,001	0,980 \pm 0,001	0,974 \pm 0,002	0,990 \pm 0,001
	Haralick	0,985 \pm 0,001	0,981 \pm 0,001	0,995 \pm 0,000	0,989 \pm 0,000	0,983 \pm 0,001	0,994 \pm 0,000
	FOM (Color)	0,986 \pm 0,000	0,985 \pm 0,000	0,999 \pm 0,000	0,985 \pm 0,001	0,985 \pm 0,001	1,000 \pm 0,000
	GCH	0,989 \pm 0,001	0,981 \pm 0,001	0,990 \pm 0,001	0,992 \pm 0,001	0,984 \pm 0,001	0,993 \pm 0,000
	Zernike	0,801 \pm 0,004	0,917 \pm 0,001	0,860 \pm 0,004	0,597 \pm 0,006	0,740 \pm 0,003	0,784 \pm 0,004
	LBP	0,991 \pm 0,001	0,955 \pm 0,002	0,962 \pm 0,002	0,990 \pm 0,000	0,962 \pm 0,001	0,967 \pm 0,001
Arquiteturas Profundas	VGG16	0,995 \pm 0,000	0,992 \pm 0,001	0,997 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	0,999 \pm 0,000
	VGG19	0,995 \pm 0,000	0,995 \pm 0,000	0,998 \pm 0,000	0,997 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	DenseNet121	0,992 \pm 0,000	0,992 \pm 0,000	0,999 \pm 0,000	0,993 \pm 0,001	0,993 \pm 0,001	1,000 \pm 0,000
	MobileNet	0,992 \pm 0,000	0,991 \pm 0,000	0,999 \pm 0,000	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000
	NASNetMobile	0,946 \pm 0,001	0,951 \pm 0,001	0,972 \pm 0,002	0,923 \pm 0,003	0,930 \pm 0,002	0,974 \pm 0,001
	MobileNetV2	0,990 \pm 0,001	0,990 \pm 0,001	0,998 \pm 0,000	0,996 \pm 0,000	0,995 \pm 0,000	0,999 \pm 0,000
	InceptionResNetV2	0,942 \pm 0,002	0,944 \pm 0,001	0,986 \pm 0,001	0,896 \pm 0,002	0,900 \pm 0,002	0,978 \pm 0,001
	DenseNet169	0,993 \pm 0,001	0,992 \pm 0,001	1,000 \pm 0,000	0,992 \pm 0,000	0,992 \pm 0,000	1,000 \pm 0,000
	DenseNet201	0,994 \pm 0,000	0,993 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,000	0,995 \pm 0,000	1,000 \pm 0,000
	InceptionV3	0,958 \pm 0,001	0,955 \pm 0,001	0,989 \pm 0,001	0,960 \pm 0,001	0,959 \pm 0,001	0,992 \pm 0,001
	ResNet101	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	ResNet101V2	0,972 \pm 0,001	0,959 \pm 0,002	0,981 \pm 0,001	0,950 \pm 0,002	0,934 \pm 0,002	0,980 \pm 0,001
	ResNet152	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	ResNet152V2	0,982 \pm 0,001	0,968 \pm 0,001	0,985 \pm 0,001	0,966 \pm 0,002	0,957 \pm 0,002	0,991 \pm 0,001
	ResNet50	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	ResNet50V2	0,975 \pm 0,001	0,965 \pm 0,001	0,987 \pm 0,001	0,965 \pm 0,001	0,954 \pm 0,001	0,988 \pm 0,000
	Xception	0,980 \pm 0,001	0,973 \pm 0,001	0,992 \pm 0,000	0,980 \pm 0,001	0,975 \pm 0,001	0,995 \pm 0,000
NASNetLarge	0,957 \pm 0,001	0,951 \pm 0,001	0,975 \pm 0,001	0,913 \pm 0,002	0,905 \pm 0,003	0,975 \pm 0,001	

Tabela 11: Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste pelo classificador RF, utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados nos conjuntos de dados original e aumentado.

	Descritor	RF					
		Conjunto de dados original			Conjunto de dados aumentado		
		Nódulo	Raiz	Fundo	Nódulo	Raiz	Fundo
Tradicionais	Momentos Invariantes de Hu	0,957 \pm 0,001	0,941 \pm 0,002	0,961 \pm 0,002	0,837 \pm 0,000	0,922 \pm 0,000	0,915 \pm 0,000
	FOM (Gray)	0,992 \pm 0,000	0,988 \pm 0,001	0,995 \pm 0,001	0,994 \pm 0,000	0,988 \pm 0,001	0,993 \pm 0,001
	Haralick	0,989 \pm 0,001	0,987 \pm 0,001	0,997 \pm 0,001	0,989 \pm 0,001	0,986 \pm 0,001	0,998 \pm 0,000
	FOM (Color)	0,994 \pm 0,000	0,993 \pm 0,000	1,000 \pm 0,000	0,994 \pm 0,000	0,993 \pm 0,000	1,000 \pm 0,000
	GCH	0,994 \pm 0,000	0,994 \pm 0,000	0,999 \pm 0,000	0,994 \pm 0,000	0,993 \pm 0,000	0,999 \pm 0,000
	Zernike	0,852 \pm 0,002	0,922 \pm 0,000	0,922 \pm 0,002	0,721 \pm 0,005	0,812 \pm 0,003	0,850 \pm 0,002
LBP	0,991 \pm 0,000	0,955 \pm 0,001	0,962 \pm 0,001	0,993 \pm 0,000	0,968 \pm 0,001	0,974 \pm 0,001	
Arquiteturas Profundas	VGG16	0,995 \pm 0,000	0,995 \pm 0,000	1,000 \pm 0,000	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000
	VGG19	0,995 \pm 0,000	0,995 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	DenseNet121	0,993 \pm 0,000	0,993 \pm 0,001	1,000 \pm 0,000	0,995 \pm 0,001	0,995 \pm 0,001	1,000 \pm 0,000
	MobileNet	0,992 \pm 0,001	0,992 \pm 0,001	1,000 \pm 0,000	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000
	NASNetMobile	0,945 \pm 0,002	0,949 \pm 0,002	0,984 \pm 0,001	0,940 \pm 0,002	0,943 \pm 0,002	0,986 \pm 0,001
	MobileNetV2	0,990 \pm 0,001	0,990 \pm 0,001	0,999 \pm 0,000	0,997 \pm 0,000	0,996 \pm 0,001	1,000 \pm 0,000
	InceptionResNetV2	0,948 \pm 0,002	0,952 \pm 0,001	0,988 \pm 0,000	0,912 \pm 0,002	0,906 \pm 0,002	0,988 \pm 0,001
	DenseNet169	0,991 \pm 0,001	0,991 \pm 0,001	1,000 \pm 0,000	0,991 \pm 0,001	0,991 \pm 0,001	1,000 \pm 0,000
	DenseNet201	0,993 \pm 0,000	0,993 \pm 0,000	1,000 \pm 0,000	0,994 \pm 0,000	0,994 \pm 0,000	1,000 \pm 0,000
	InceptionV3	0,958 \pm 0,001	0,956 \pm 0,002	0,989 \pm 0,001	0,953 \pm 0,001	0,952 \pm 0,001	0,991 \pm 0,001
	ResNet101	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	ResNet101V2	0,972 \pm 0,001	0,962 \pm 0,001	0,987 \pm 0,001	0,965 \pm 0,002	0,954 \pm 0,002	0,987 \pm 0,001
	ResNet152	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	ResNet152V2	0,981 \pm 0,001	0,970 \pm 0,001	0,989 \pm 0,001	0,977 \pm 0,001	0,970 \pm 0,000	0,993 \pm 0,001
	ResNet50	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	ResNet50V2	0,971 \pm 0,001	0,963 \pm 0,001	0,990 \pm 0,001	0,969 \pm 0,001	0,962 \pm 0,001	0,992 \pm 0,000
Xception	0,977 \pm 0,001	0,973 \pm 0,001	0,993 \pm 0,001	0,976 \pm 0,001	0,973 \pm 0,001	0,996 \pm 0,001	
NASNetLarge	0,957 \pm 0,001	0,956 \pm 0,001	0,986 \pm 0,001	0,949 \pm 0,001	0,944 \pm 0,001	0,987 \pm 0,001	

Tabela 12: Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste pelo classificador SVM (KL), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados nos conjuntos de dados original e aumentado.

	Descritor	SVM (KL)					
		Conjunto de dados original			Conjunto de dados aumentado		
		Nódulo	Raiz	Fundo	Nódulo	Raiz	Fundo
Tradicionais	Momentos Invariantes de Hu	0,894 \pm 0,000	0,922 \pm 0,000	0,919 \pm 0,000	0,506 \pm 0,343	0,583 \pm 0,412	0,832 \pm 0,249
	FOM (Gray)	0,969 \pm 0,002	0,966 \pm 0,002	0,994 \pm 0,000	0,983 \pm 0,000	0,979 \pm 0,000	0,988 \pm 0,000
	Haralick	0,986 \pm 0,000	0,982 \pm 0,000	0,995 \pm 0,000	0,984 \pm 0,000	0,972 \pm 0,000	0,986 \pm 0,000
	FOM (Color)	0,990 \pm 0,001	0,988 \pm 0,001	0,998 \pm 0,000	0,992 \pm 0,000	0,992 \pm 0,000	1,000 \pm 0,000
	GCH	0,974 \pm 0,001	0,972 \pm 0,001	0,981 \pm 0,001	0,989 \pm 0,000	0,984 \pm 0,000	0,992 \pm 0,000
	Zernike	0,837 \pm 0,000	0,922 \pm 0,000	0,915 \pm 0,000	0,546 \pm 0,002	0,875 \pm 0,001	0,589 \pm 0,001
	LBP	0,989 \pm 0,000	0,966 \pm 0,001	0,972 \pm 0,001	0,993 \pm 0,000	0,981 \pm 0,001	0,987 \pm 0,000
Arquiteturas Profundas	VGG16	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	VGG19	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	DenseNet121	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000
	MobileNet	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	NASNetMobile	0,962 \pm 0,001	0,964 \pm 0,001	0,986 \pm 0,001	0,946 \pm 0,001	0,948 \pm 0,002	0,978 \pm 0,001
	MobileNetV2	0,998 \pm 0,001	0,998 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	InceptionResNetV2	0,924 \pm 0,001	0,931 \pm 0,001	0,984 \pm 0,001	0,834 \pm 0,002	0,844 \pm 0,002	0,971 \pm 0,000
	DenseNet169	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000
	DenseNet201	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000
	InceptionV3	0,980 \pm 0,001	0,978 \pm 0,001	0,995 \pm 0,001	0,976 \pm 0,001	0,976 \pm 0,001	0,996 \pm 0,001
	ResNet101	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	ResNet101V2	0,992 \pm 0,001	0,985 \pm 0,001	0,992 \pm 0,001	0,986 \pm 0,001	0,977 \pm 0,001	0,991 \pm 0,000
	ResNet152	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	ResNet152V2	0,992 \pm 0,000	0,988 \pm 0,001	0,996 \pm 0,001	0,987 \pm 0,001	0,984 \pm 0,001	0,997 \pm 0,000
	ResNet50	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	ResNet50V2	0,992 \pm 0,001	0,989 \pm 0,001	0,995 \pm 0,001	0,990 \pm 0,001	0,986 \pm 0,001	0,995 \pm 0,000
Xception	0,993 \pm 0,001	0,991 \pm 0,000	0,998 \pm 0,001	0,991 \pm 0,001	0,988 \pm 0,001	0,997 \pm 0,000	
NASNetLarge	0,974 \pm 0,001	0,969 \pm 0,001	0,990 \pm 0,001	0,962 \pm 0,002	0,955 \pm 0,003	0,985 \pm 0,001	

Tabela 13: Média de acurácias por classe \pm desvio padrão obtidos no conjunto de dados de teste pelo classificador SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados nos conjuntos de dados original e aumentado.

	Descritor	SVM (KPG2)					
		Conjunto de dados original			Conjunto de dados aumentado		
		Nódulo	Raiz	Fundo	Nódulo	Raiz	Fundo
Tradicionais	Momentos Invariantes de Hu	0,927 \pm 0,001	0,938 \pm 0,001	0,936 \pm 0,001	0,500 \pm 0,337	0,584 \pm 0,414	0,832 \pm 0,249
	FOM (Gray)	0,984 \pm 0,001	0,980 \pm 0,001	0,994 \pm 0,001	0,991 \pm 0,000	0,986 \pm 0,001	0,994 \pm 0,001
	Haralick	0,992 \pm 0,000	0,989 \pm 0,000	0,996 \pm 0,000	0,989 \pm 0,000	0,982 \pm 0,000	0,991 \pm 0,000
	FOM (Color)	0,992 \pm 0,000	0,989 \pm 0,000	0,998 \pm 0,000	0,994 \pm 0,000	0,994 \pm 0,000	1,000 \pm 0,000
	GCH	0,986 \pm 0,000	0,987 \pm 0,001	0,992 \pm 0,001	0,987 \pm 0,000	0,988 \pm 0,000	0,995 \pm 0,000
	Zernike	0,837 \pm 0,000	0,922 \pm 0,000	0,915 \pm 0,000	0,603 \pm 0,002	0,877 \pm 0,001	0,648 \pm 0,002
	LBP	0,989 \pm 0,000	0,976 \pm 0,001	0,982 \pm 0,001	0,987 \pm 0,000	0,980 \pm 0,001	0,989 \pm 0,000
Arquiteturas Profundas	VGG16	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	VGG19	0,998 \pm 0,000	0,997 \pm 0,000	0,999 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	DenseNet121	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	MobileNet	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	NASNetMobile	0,949 \pm 0,000	0,956 \pm 0,001	0,981 \pm 0,000	0,939 \pm 0,001	0,949 \pm 0,001	0,979 \pm 0,001
	MobileNetV2	0,998 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	InceptionResNetV2	0,910 \pm 0,000	0,922 \pm 0,000	0,956 \pm 0,000	0,682 \pm 0,002	0,726 \pm 0,001	0,900 \pm 0,001
	DenseNet169	0,996 \pm 0,000	0,996 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	1,000 \pm 0,000
	DenseNet201	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000	0,997 \pm 0,000	0,997 \pm 0,000	1,000 \pm 0,000
	InceptionV3	0,958 \pm 0,003	0,965 \pm 0,001	0,985 \pm 0,002	0,960 \pm 0,001	0,961 \pm 0,000	0,984 \pm 0,001
	ResNet101	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	ResNet101V2	0,956 \pm 0,001	0,957 \pm 0,001	0,981 \pm 0,000	0,936 \pm 0,000	0,938 \pm 0,000	0,974 \pm 0,000
	ResNet152	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	ResNet152V2	0,983 \pm 0,000	0,971 \pm 0,001	0,988 \pm 0,001	0,975 \pm 0,000	0,965 \pm 0,000	0,988 \pm 0,000
	ResNet50	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	1,000 \pm 0,000
	ResNet50V2	0,978 \pm 0,001	0,969 \pm 0,001	0,988 \pm 0,001	0,967 \pm 0,001	0,966 \pm 0,001	0,987 \pm 0,001
	Xception	0,987 \pm 0,000	0,980 \pm 0,001	0,993 \pm 0,000	0,987 \pm 0,001	0,984 \pm 0,001	0,997 \pm 0,000
NASNetLarge	0,968 \pm 0,001	0,964 \pm 0,001	0,987 \pm 0,000	0,961 \pm 0,001	0,954 \pm 0,001	0,986 \pm 0,000	

Tabela 14: Tempo médio de treinamento e classificação \pm desvio padrão obtidos pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados original.

	Descritor	Características	Conjunto de dados original							
			KNN		RF		SVM (KL)		SVM (KPG2)	
			Treinamento	Classificação	Treinamento	Classificação	Treinamento	Classificação	Treinamento	Classificação
Tradicionais	Momentos Invariantes de Hu	7	0,014 \pm 0,000	0,092 \pm 0,001	1,555 \pm 0,016	0,063 \pm 0,001	0,471 \pm 0,004	0,090 \pm 0,001	0,582 \pm 0,022	0,092 \pm 0,001
	FOM (Gray)	8	0,015 \pm 0,000	0,090 \pm 0,001	1,192 \pm 0,027	0,049 \pm 0,000	0,229 \pm 0,006	0,057 \pm 0,002	0,173 \pm 0,005	0,037 \pm 0,001
	Haralick	13	0,020 \pm 0,000	0,147 \pm 0,004	2,166 \pm 0,021	0,053 \pm 0,000	0,164 \pm 0,002	0,043 \pm 0,001	0,140 \pm 0,006	0,030 \pm 0,000
	FOM (Color)	24	0,007 \pm 0,000	0,540 \pm 0,010	1,892 \pm 0,049	0,046 \pm 0,001	0,171 \pm 0,007	0,033 \pm 0,001	0,166 \pm 0,005	0,032 \pm 0,001
	GCH	30	0,007 \pm 0,000	0,547 \pm 0,011	2,649 \pm 0,036	0,050 \pm 0,000	0,428 \pm 0,008	0,091 \pm 0,002	0,243 \pm 0,011	0,042 \pm 0,001
	Zernike	72	0,007 \pm 0,000	0,558 \pm 0,009	11,006 \pm 0,149	0,092 \pm 0,000	1,765 \pm 0,034	0,307 \pm 0,003	3,931 \pm 0,067	0,378 \pm 0,005
	LBP	352	0,011 \pm 0,000	0,610 \pm 0,018	3,946 \pm 0,080	0,058 \pm 0,000	0,970 \pm 0,057	0,308 \pm 0,003	1,091 \pm 0,074	0,296 \pm 0,009
Arquiteturas Profundas	VGG16	512	0,011 \pm 0,000	0,622 \pm 0,009	12,842 \pm 0,197	0,060 \pm 0,002	0,588 \pm 0,026	0,078 \pm 0,004	1,230 \pm 0,064	0,174 \pm 0,027
	VGG19	512	0,011 \pm 0,000	0,629 \pm 0,011	14,518 \pm 0,153	0,059 \pm 0,001	0,594 \pm 0,009	0,078 \pm 0,002	1,187 \pm 0,021	0,183 \pm 0,005
	DenseNet121	1.024	0,014 \pm 0,000	0,695 \pm 0,011	31,864 \pm 0,595	0,065 \pm 0,001	1,417 \pm 0,057	0,243 \pm 0,011	1,814 \pm 0,052	0,337 \pm 0,009
	MobileNet	1.024	0,014 \pm 0,000	0,689 \pm 0,013	34,623 \pm 0,275	0,066 \pm 0,001	1,490 \pm 0,081	0,250 \pm 0,010	2,137 \pm 0,095	0,359 \pm 0,008
	NASNetMobile	1.056	0,039 \pm 0,024	10,429 \pm 0,398	33,397 \pm 6,303	0,263 \pm 0,164	22,375 \pm 3,505	3,382 \pm 1,429	22,673 \pm 2,861	5,627 \pm 1,742
	MobileNetV2	1.280	0,015 \pm 0,000	0,740 \pm 0,016	39,699 \pm 0,750	0,070 \pm 0,001	2,196 \pm 0,042	0,408 \pm 0,008	3,103 \pm 0,053	0,533 \pm 0,010
	InceptionResNetV2	1.536	0,017 \pm 0,001	0,787 \pm 0,015	39,889 \pm 0,979	0,082 \pm 0,003	19,639 \pm 1,449	3,042 \pm 0,046	25,568 \pm 2,106	4,024 \pm 0,098
	DenseNet169	1.664	0,020 \pm 0,003	1,045 \pm 0,001	44,429 \pm 5,763	0,072 \pm 0,005	2,659 \pm 0,186	0,462 \pm 0,040	3,616 \pm 0,312	0,638 \pm 0,059
	DenseNet201	1.920	0,035 \pm 0,021	5,647 \pm 5,148	71,225 \pm 34,881	0,075 \pm 0,005	4,006 \pm 2,360	0,610 \pm 0,164	5,074 \pm 1,968	1,063 \pm 0,602
	InceptionV3	2.048	0,020 \pm 0,000	0,951 \pm 0,076	89,939 \pm 1,958	0,095 \pm 0,002	12,089 \pm 0,556	1,981 \pm 0,062	19,427 \pm 0,891	3,671 \pm 0,134
	ResNet101	2.048	0,032 \pm 0,013	12,208 \pm 0,516	52,016 \pm 3,598	0,076 \pm 0,012	3,362 \pm 0,849	0,566 \pm 0,069	6,091 \pm 1,545	0,887 \pm 0,265
	ResNet101V2	2.048	0,030 \pm 0,003	5,672 \pm 3,386	29,899 \pm 4,413	0,098 \pm 0,010	11,732 \pm 1,421	2,279 \pm 0,206	24,381 \pm 2,337	4,961 \pm 0,803
	ResNet152	2.048	0,056 \pm 0,025	11,976 \pm 0,658	52,412 \pm 2,675	0,094 \pm 0,048	4,682 \pm 0,918	0,660 \pm 0,300	6,364 \pm 1,365	1,142 \pm 0,554
	ResNet152V2	2.048	0,028 \pm 0,004	1,996 \pm 0,555	46,882 \pm 4,885	0,090 \pm 0,010	7,187 \pm 0,797	1,211 \pm 0,096	14,186 \pm 1,016	2,848 \pm 0,252
	ResNet50	2.048	0,025 \pm 0,004	8,604 \pm 4,948	53,340 \pm 10,862	0,074 \pm 0,008	3,381 \pm 0,855	0,624 \pm 0,350	5,260 \pm 1,127	0,853 \pm 0,203
	ResNet50V2	2.048	0,053 \pm 0,022	12,025 \pm 0,343	47,943 \pm 4,756	0,099 \pm 0,006	12,806 \pm 1,858	1,890 \pm 0,647	29,700 \pm 4,286	5,838 \pm 1,230
	Xception	2.048	0,020 \pm 0,000	0,862 \pm 0,018	40,274 \pm 0,768	0,090 \pm 0,003	7,415 \pm 0,541	1,295 \pm 0,027	15,461 \pm 1,137	3,106 \pm 0,083
NASNetLarge	4.032	0,051 \pm 0,033	10,872 \pm 6,926	105,967 \pm 28,166	0,177 \pm 0,125	38,030 \pm 8,332	10,489 \pm 3,892	50,775 \pm 12,672	14,648 \pm 5,831	

Tabela 15: Tempo médio de treinamento e classificação \pm desvio padrão obtidos pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados aumentado.

Descritor	Características	Conjunto de dados aumentado								
		KNN		RF		SVM (KL)		SVM (KPG2)		
		Treinamento	Classificação	Treinamento	Classificação	Treinamento	Classificação	Treinamento	Classificação	
Tradicionais	Momentos Invariantes de Hu	7	0,074 ± 0,008	0,146 ± 0,065	2,685 ± 0,583	0,049 ± 0,026	42,470 ± 3,954	2,400 ± 0,657	49,107 ± 6,835	2,482 ± 0,600
	FOM (Gray)	8	0,070 ± 0,014	0,126 ± 0,023	6,803 ± 1,024	0,058 ± 0,003	10,793 ± 2,175	0,627 ± 0,177	6,547 ± 1,439	0,542 ± 0,149
	Haralick	13	0,108 ± 0,011	0,205 ± 0,025	11,660 ± 1,107	0,058 ± 0,004	7,701 ± 1,257	0,705 ± 0,023	7,698 ± 0,164	0,647 ± 0,015
	FOM (Color)	24	0,031 ± 0,007	1,059 ± 0,021	9,764 ± 1,450	0,052 ± 0,001	2,896 ± 0,702	0,220 ± 0,066	2,598 ± 0,564	0,222 ± 0,064
	GCH	30	0,029 ± 0,004	0,858 ± 0,171	9,687 ± 1,296	0,060 ± 0,001	14,688 ± 3,472	0,958 ± 0,277	12,188 ± 2,909	0,744 ± 0,161
	Zernike	72	0,038 ± 0,017	2,963 ± 3,504	45,662 ± 13,512	0,126 ± 0,004	98,367 ± 38,961	4,804 ± 1,480	110,667 ± 41,419	5,622 ± 2,570
Arquiteturas Profundas	LBP	352	0,039 ± 0,007	0,950 ± 0,361	25,183 ± 0,435	0,081 ± 0,002	57,073 ± 1,078	3,935 ± 0,090	88,338 ± 1,507	6,114 ± 0,532
	VGG16	512	0,045 ± 0,009	1,288 ± 0,243	51,628 ± 7,543	0,069 ± 0,005	5,048 ± 0,768	0,231 ± 0,023	10,461 ± 1,705	0,501 ± 0,040
	VGG19	512	0,055 ± 0,008	7,090 ± 4,502	70,204 ± 12,442	0,074 ± 0,004	6,214 ± 1,221	0,274 ± 0,038	12,685 ± 3,166	0,569 ± 0,025
	DenseNet121	1.024	0,066 ± 0,010	3,823 ± 3,705	132,349 ± 24,323	0,099 ± 0,020	15,102 ± 2,029	0,788 ± 0,247	31,733 ± 5,736	2,018 ± 0,385
	MobileNet	1.024	0,062 ± 0,014	1,667 ± 0,302	108,068 ± 7,930	0,084 ± 0,002	9,975 ± 0,984	0,553 ± 0,022	14,756 ± 1,316	0,769 ± 0,077
	NASNetMobile	1.056	0,074 ± 0,025	6,279 ± 5,639	50,999 ± 13,189	0,146 ± 0,050	252,206 ± 55,194	17,055 ± 2,643	341,036 ± 70,323	29,995 ± 8,089
	MobileNetV2	1.280	0,073 ± 0,014	1,744 ± 0,408	125,908 ± 12,513	0,094 ± 0,004	17,336 ± 1,438	0,951 ± 0,049	25,315 ± 1,615	1,550 ± 0,072
	InceptionResNetV2	1.536	0,074 ± 0,011	3,946 ± 4,600	195,593 ± 59,645	0,118 ± 0,006	616,106 ± 139,539	48,894 ± 12,589	798,946 ± 82,435	60,340 ± 16,644
	DenseNet169	1.664	0,083 ± 0,011	6,043 ± 6,125	194,734 ± 37,309	0,105 ± 0,029	23,815 ± 3,730	1,236 ± 0,389	46,103 ± 6,456	4,596 ± 1,236
	DenseNet201	1.920	0,099 ± 0,036	6,295 ± 6,634	186,033 ± 32,516	0,097 ± 0,010	25,718 ± 4,988	1,539 ± 0,254	49,002 ± 5,363	4,709 ± 0,974
	InceptionV3	2.048	0,085 ± 0,012	2,918 ± 1,414	270,682 ± 64,351	0,148 ± 0,042	225,704 ± 55,144	18,345 ± 5,598	485,093 ± 119,304	35,104 ± 4,626
	ResNet101	2.048	0,081 ± 0,012	5,974 ± 7,707	150,392 ± 46,674	0,076 ± 0,009	15,727 ± 1,255	1,035 ± 0,263	32,467 ± 11,576	1,978 ± 1,000
	ResNet101V2	2.048	0,086 ± 0,019	6,109 ± 7,687	98,848 ± 22,087	0,117 ± 0,015	252,100 ± 51,122	19,676 ± 1,161	609,597 ± 30,360	53,491 ± 12,067
	ResNet152	2.048	0,084 ± 0,019	3,397 ± 2,840	137,957 ± 20,103	0,078 ± 0,011	15,424 ± 1,221	0,942 ± 0,063	26,053 ± 2,122	1,643 ± 0,166
	ResNet152V2	2.048	0,085 ± 0,016	8,840 ± 8,688	212,917 ± 56,874	0,119 ± 0,013	134,606 ± 20,271	11,004 ± 1,206	479,124 ± 72,754	41,578 ± 13,094
	ResNet50	2.048	0,089 ± 0,013	2,567 ± 0,757	138,618 ± 15,576	0,076 ± 0,007	14,758 ± 0,934	0,876 ± 0,043	24,885 ± 1,596	1,491 ± 0,076
	ResNet50V2	2.048	0,086 ± 0,017	9,744 ± 9,268	123,988 ± 27,202	0,120 ± 0,020	157,701 ± 30,445	14,056 ± 4,714	491,741 ± 90,730	47,906 ± 11,111
	Xception	2.048	0,080 ± 0,015	2,459 ± 0,723	93,575 ± 12,400	0,110 ± 0,010	109,205 ± 6,989	10,820 ± 1,889	218,366 ± 11,273	21,700 ± 1,593
	NASNetLarge	4.032	0,143 ± 0,030	14,164 ± 11,477	234,305 ± 68,480	0,145 ± 0,026	665,762 ± 148,504	56,379 ± 16,083	991,289 ± 220,499	94,993 ± 33,258

Tabela 16: Média das métricas (sensibilidade, especificidade, precisão, F1 e kappa) \pm desvio padrão obtidos no conjunto de dados de teste pelos classificadores KNN e RF, utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados original.

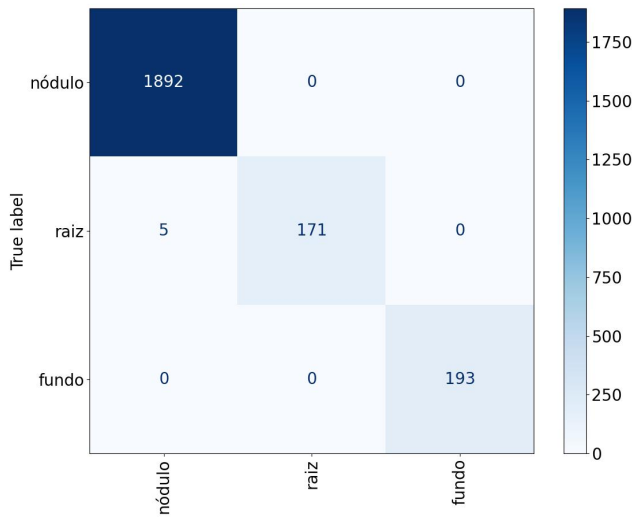
Descritor	Conjunto de dados original										
	KNN					RF					
	Sensibilidade	Especificidade	Precisão	F1	Kappa	Sensibilidade	Especificidade	Precisão	F1	Kappa	
Tradicionais	Momentos Invariantes de Hu	0,895 ± 0,003	0,947 ± 0,001	0,895 ± 0,003	0,895 ± 0,003	0,609 ± 0,011	0,930 ± 0,002	0,965 ± 0,001	0,930 ± 0,002	0,930 ± 0,002	0,733 ± 0,007
	FOM (Gray)	0,971 ± 0,002	0,985 ± 0,001	0,971 ± 0,002	0,971 ± 0,002	0,891 ± 0,008	0,988 ± 0,001	0,994 ± 0,000	0,988 ± 0,001	0,988 ± 0,001	0,957 ± 0,003
	Haralick	0,981 ± 0,001	0,990 ± 0,000	0,981 ± 0,001	0,981 ± 0,001	0,931 ± 0,003	0,987 ± 0,001	0,993 ± 0,001	0,987 ± 0,001	0,987 ± 0,001	0,953 ± 0,004
	FOM (Color)	0,985 ± 0,000	0,993 ± 0,000	0,985 ± 0,000	0,985 ± 0,000	0,947 ± 0,002	0,993 ± 0,000	0,997 ± 0,000	0,993 ± 0,000	0,993 ± 0,000	0,976 ± 0,001
	GCH	0,980 ± 0,001	0,990 ± 0,001	0,980 ± 0,001	0,980 ± 0,001	0,929 ± 0,004	0,993 ± 0,000	0,997 ± 0,000	0,993 ± 0,000	0,993 ± 0,000	0,977 ± 0,002
	Zernike	0,789 ± 0,004	0,894 ± 0,002	0,789 ± 0,004	0,789 ± 0,004	0,134 ± 0,012	0,848 ± 0,002	0,924 ± 0,001	0,848 ± 0,002	0,848 ± 0,002	0,184 ± 0,017
Arquiteturas Profundas	LBP	0,954 ± 0,002	0,977 ± 0,001	0,954 ± 0,002	0,954 ± 0,002	0,836 ± 0,006	0,954 ± 0,001	0,977 ± 0,000	0,954 ± 0,001	0,954 ± 0,001	0,838 ± 0,004
	VGG16	0,992 ± 0,001	0,996 ± 0,000	0,992 ± 0,001	0,992 ± 0,001	0,972 ± 0,002	0,995 ± 0,000	0,998 ± 0,000	0,995 ± 0,000	0,995 ± 0,000	0,983 ± 0,002
	VGG19	0,994 ± 0,000	0,997 ± 0,000	0,994 ± 0,000	0,994 ± 0,000	0,979 ± 0,001	0,995 ± 0,000	0,998 ± 0,000	0,995 ± 0,000	0,995 ± 0,000	0,983 ± 0,002
	DenseNet121	0,992 ± 0,000	0,996 ± 0,000	0,992 ± 0,000	0,992 ± 0,000	0,971 ± 0,001	0,993 ± 0,001	0,996 ± 0,000	0,993 ± 0,001	0,993 ± 0,001	0,974 ± 0,002
	MobileNet	0,991 ± 0,000	0,995 ± 0,000	0,991 ± 0,000	0,991 ± 0,000	0,967 ± 0,001	0,992 ± 0,001	0,996 ± 0,000	0,992 ± 0,001	0,992 ± 0,001	0,970 ± 0,002
	NASNetMobile	0,935 ± 0,002	0,967 ± 0,001	0,935 ± 0,002	0,935 ± 0,002	0,754 ± 0,006	0,939 ± 0,002	0,970 ± 0,001	0,939 ± 0,002	0,939 ± 0,002	0,755 ± 0,008
	MobileNetV2	0,989 ± 0,001	0,994 ± 0,000	0,989 ± 0,001	0,989 ± 0,001	0,960 ± 0,002	0,990 ± 0,001	0,995 ± 0,000	0,990 ± 0,001	0,990 ± 0,001	0,964 ± 0,004
	InceptionResNetV2	0,936 ± 0,002	0,968 ± 0,001	0,936 ± 0,002	0,936 ± 0,002	0,755 ± 0,007	0,944 ± 0,002	0,972 ± 0,001	0,944 ± 0,002	0,944 ± 0,002	0,778 ± 0,008
	DenseNet169	0,992 ± 0,001	0,996 ± 0,000	0,992 ± 0,001	0,992 ± 0,001	0,972 ± 0,002	0,991 ± 0,001	0,995 ± 0,000	0,991 ± 0,001	0,991 ± 0,001	0,968 ± 0,002
	DenseNet201	0,993 ± 0,000	0,996 ± 0,000	0,993 ± 0,000	0,993 ± 0,000	0,974 ± 0,001	0,993 ± 0,000	0,997 ± 0,000	0,993 ± 0,000	0,993 ± 0,000	0,976 ± 0,001
	InceptionV3	0,951 ± 0,001	0,976 ± 0,000	0,951 ± 0,001	0,951 ± 0,001	0,812 ± 0,003	0,952 ± 0,002	0,976 ± 0,001	0,952 ± 0,002	0,952 ± 0,002	0,814 ± 0,008
	ResNet101	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,992 ± 0,000	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,993 ± 0,001
	ResNet101V2	0,956 ± 0,002	0,978 ± 0,001	0,956 ± 0,002	0,956 ± 0,002	0,846 ± 0,007	0,961 ± 0,001	0,980 ± 0,000	0,961 ± 0,001	0,961 ± 0,001	0,854 ± 0,003
	ResNet152	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,994 ± 0,001	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,993 ± 0,001
	ResNet152V2	0,967 ± 0,001	0,984 ± 0,001	0,967 ± 0,001	0,967 ± 0,001	0,885 ± 0,004	0,970 ± 0,001	0,985 ± 0,000	0,970 ± 0,001	0,970 ± 0,001	0,893 ± 0,003
	ResNet50	0,997 ± 0,000	0,998 ± 0,000	0,997 ± 0,000	0,997 ± 0,000	0,989 ± 0,001	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,992 ± 0,001
	ResNet50V2	0,964 ± 0,001	0,982 ± 0,001	0,964 ± 0,001	0,964 ± 0,001	0,870 ± 0,005	0,962 ± 0,001	0,981 ± 0,000	0,962 ± 0,001	0,962 ± 0,001	0,860 ± 0,003
	Xception	0,972 ± 0,001	0,986 ± 0,001	0,972 ± 0,001	0,972 ± 0,001	0,899 ± 0,004	0,971 ± 0,001	0,986 ± 0,001	0,971 ± 0,001	0,971 ± 0,001	0,894 ± 0,005
	NASNetLarge	0,942 ± 0,001	0,971 ± 0,001	0,942 ± 0,001	0,942 ± 0,001	0,791 ± 0,004	0,949 ± 0,001	0,975 ± 0,001	0,949 ± 0,001	0,949 ± 0,001	0,807 ± 0,005

Tabela 19: Média das métricas (sensibilidade, especificidade, precisão, F1 e kappa) ± desvio padrão obtidos no conjunto de dados de teste pelos classificadores SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas. Os modelos foram treinados no conjunto de dados aumentado.

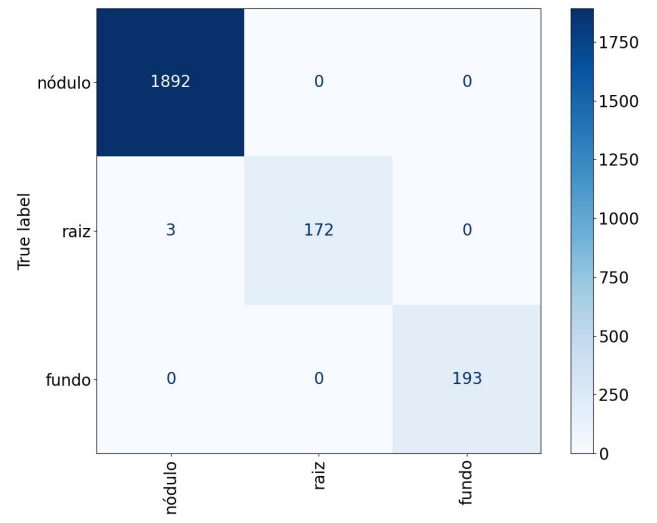
Descritor	Conjunto de dados aumentado										
	SVM (KL)					SVM (KPG2)					
	Sensibilidade	Especificidade	Precisão	F1	Kappa	Sensibilidade	Especificidade	Precisão	F1	Kappa	
Tradicionais	Momentos Invariantes de Hu	0,460 ± 0,381	0,730 ± 0,190	0,460 ± 0,381	0,460 ± 0,381	0,038 ± 0,077	0,458 ± 0,379	0,729 ± 0,189	0,458 ± 0,379	0,458 ± 0,379	0,000 ± 0,000
	FOM (Gray)	0,975 ± 0,000	0,987 ± 0,000	0,975 ± 0,000	0,975 ± 0,000	0,910 ± 0,000	0,986 ± 0,001	0,993 ± 0,000	0,986 ± 0,001	0,986 ± 0,001	0,949 ± 0,002
	Haralick	0,972 ± 0,000	0,986 ± 0,000	0,972 ± 0,000	0,972 ± 0,000	0,901 ± 0,001	0,981 ± 0,000	0,990 ± 0,000	0,981 ± 0,000	0,981 ± 0,000	0,933 ± 0,001
	FOM (Color)	0,992 ± 0,000	0,996 ± 0,000	0,992 ± 0,000	0,992 ± 0,000	0,971 ± 0,001	0,994 ± 0,000	0,997 ± 0,000	0,994 ± 0,000	0,994 ± 0,000	0,979 ± 0,001
	GCH	0,983 ± 0,000	0,991 ± 0,000	0,983 ± 0,000	0,983 ± 0,000	0,938 ± 0,001	0,985 ± 0,000	0,993 ± 0,000	0,985 ± 0,000	0,985 ± 0,000	0,947 ± 0,001
	Zernike	0,505 ± 0,002	0,752 ± 0,001	0,505 ± 0,002	0,505 ± 0,002	0,161 ± 0,001	0,564 ± 0,002	0,782 ± 0,001	0,564 ± 0,002	0,564 ± 0,002	0,198 ± 0,001
	LBP	0,980 ± 0,001	0,990 ± 0,000	0,980 ± 0,001	0,980 ± 0,001	0,931 ± 0,002	0,978 ± 0,001	0,989 ± 0,000	0,978 ± 0,001	0,978 ± 0,001	0,924 ± 0,002
Arquiteturas Profundas	VGG16	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,992 ± 0,001	0,999 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,995 ± 0,001
	VGG19	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,992 ± 0,000	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,993 ± 0,001
	DenseNet121	0,997 ± 0,000	0,999 ± 0,000	0,997 ± 0,000	0,997 ± 0,000	0,990 ± 0,001	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,992 ± 0,001
	MobileNet	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,994 ± 0,001	0,999 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,996 ± 0,001
	NASNetMobile	0,936 ± 0,002	0,968 ± 0,001	0,936 ± 0,002	0,936 ± 0,002	0,783 ± 0,005	0,933 ± 0,001	0,966 ± 0,000	0,933 ± 0,001	0,933 ± 0,001	0,751 ± 0,003
	MobileNetV2	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,993 ± 0,001	0,999 ± 0,000	1,000 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,997 ± 0,000
	InceptionResNetV2	0,824 ± 0,002	0,912 ± 0,001	0,824 ± 0,002	0,824 ± 0,002	0,557 ± 0,003	0,654 ± 0,002	0,827 ± 0,001	0,654 ± 0,002	0,654 ± 0,002	0,332 ± 0,002
	DenseNet169	0,997 ± 0,000	0,999 ± 0,000	0,997 ± 0,000	0,997 ± 0,000	0,991 ± 0,001	0,998 ± 0,000	0,999 ± 0,000	0,998 ± 0,000	0,998 ± 0,000	0,992 ± 0,001
	DenseNet201	0,997 ± 0,000	0,999 ± 0,000	0,997 ± 0,000	0,997 ± 0,000	0,991 ± 0,001	0,997 ± 0,000	0,998 ± 0,000	0,997 ± 0,000	0,997 ± 0,000	0,989 ± 0,001
	InceptionV3	0,974 ± 0,001	0,987 ± 0,001	0,974 ± 0,001	0,974 ± 0,001	0,910 ± 0,004	0,952 ± 0,001	0,976 ± 0,000	0,952 ± 0,001	0,952 ± 0,001	0,834 ± 0,002
	ResNet101	0,999 ± 0,000	1,000 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,997 ± 0,000	0,999 ± 0,000	1,000 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,997 ± 0,000
	ResNet101V2	0,977 ± 0,001	0,988 ± 0,000	0,977 ± 0,001	0,977 ± 0,001	0,920 ± 0,003	0,924 ± 0,000	0,962 ± 0,000	0,924 ± 0,000	0,924 ± 0,000	0,707 ± 0,001
	ResNet152	0,999 ± 0,000	1,000 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,997 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,996 ± 0,001
	ResNet152V2	0,984 ± 0,001	0,992 ± 0,000	0,984 ± 0,001	0,984 ± 0,001	0,946 ± 0,003	0,964 ± 0,000	0,982 ± 0,000	0,964 ± 0,000	0,964 ± 0,000	0,873 ± 0,001
	ResNet50	0,999 ± 0,000	1,000 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,997 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,999 ± 0,000	0,995 ± 0,000
	ResNet50V2	0,985 ± 0,001	0,993 ± 0,000	0,985 ± 0,001	0,985 ± 0,001	0,949 ± 0,003	0,960 ± 0,001	0,980 ± 0,000	0,960 ± 0,001	0,960 ± 0,001	0,866 ± 0,003
	Xception	0,988 ± 0,001	0,994 ± 0,000	0,988 ± 0,001	0,988 ± 0,001	0,958 ± 0,003	0,984 ± 0,001	0,992 ± 0,000	0,984 ± 0,001	0,984 ± 0,001	0,945 ± 0,002
	NASNetLarge	0,951 ± 0,003	0,976 ± 0,001	0,951 ± 0,003	0,951 ± 0,003	0,833 ± 0,010	0,950 ± 0,001	0,975 ± 0,001	0,950 ± 0,001	0,950 ± 0,001	0,830 ± 0,004

Tabela 20: Tempo médio de treinamento e classificação ± desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados nos conjuntos de dados original e aumentado.

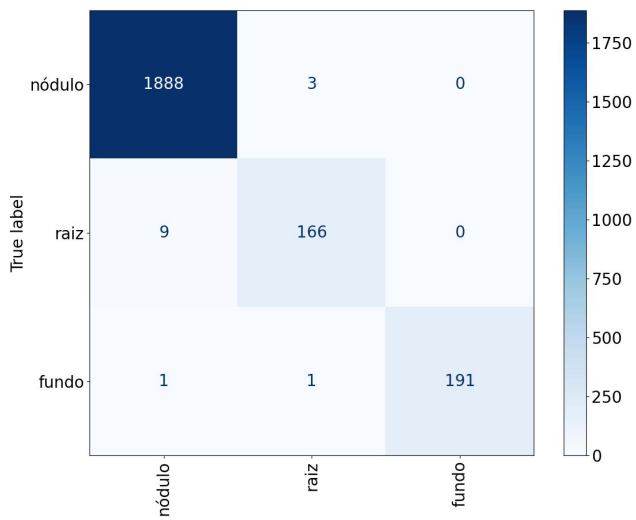
CNN	Conjunto de dados original		Conjunto de dados aumentado	
	Treinamento	Classificação	Treinamento	Classificação
VGG16	431,638 ± 134,143	6,057 ± 0,158	1356,099 ± 306,509	6,029 ± 0,164
VGG19	377,881 ± 68,534	6,959 ± 0,030	1643,364 ± 498,137	6,942 ± 0,020
DenseNet121	349,723 ± 79,239	4,055 ± 0,117	1165,057 ± 334,089	4,033 ± 0,078
MobileNet	410,655 ± 259,046	2,517 ± 0,116	1634,076 ± 1181,465	2,525 ± 0,083
NASNetMobile	428,061 ± 178,119	3,940 ± 0,050	1367,892 ± 484,238	3,907 ± 0,041
MobileNetV2	187,283 ± 49,687	2,568 ± 0,055	910,177 ± 172,460	2,582 ± 0,220
InceptionResNetV2	857,911 ± 229,625	12,948 ± 0,113	4001,244 ± 1151,672	12,940 ± 0,133
DenseNet169	625,671 ± 234,820	4,992 ± 0,069	1645,204 ± 483,998	4,972 ± 0,054
DenseNet201	550,568 ± 144,882	6,227 ± 0,094	1827,538 ± 487,855	6,234 ± 0,077
InceptionV3	1801,433 ± 767,942	5,257 ± 0,162	2096,590 ± 2627,578	5,310 ± 0,171
ResNet101	3616,321 ± 951,816	7,311 ± 0,017	10367,669 ± 3313,341	7,280 ± 0,062
ResNet101V2	1407,381 ± 776,482	6,331 ± 0,035	2729,093 ± 3046,806	6,293 ± 0,027
ResNet152	4380,810 ± 892,898	10,312 ± 0,033	11874,626 ± 2568,869	10,251 ± 0,039
ResNet152V2	916,577 ± 860,357	9,023 ± 0,044	3014,832 ± 3097,171	9,059 ± 0,020
ResNet50	2417,574 ± 622,885	4,486 ± 0,073	7442,891 ± 1574,707	4,490 ± 0,062
ResNet50V2	1145,790 ± 337,474	3,734 ± 0,053	1445,191 ± 1180,603	3,719 ± 0,072
Xception	1625,925 ± 1801,254	8,826 ± 0,063	3379,700 ± 1619,901	8,887 ± 0,070
NASNetLarge	4260,906 ± 4099,762	37,344 ± 0,071	26264,936 ± 28095,703	37,228 ± 0,125



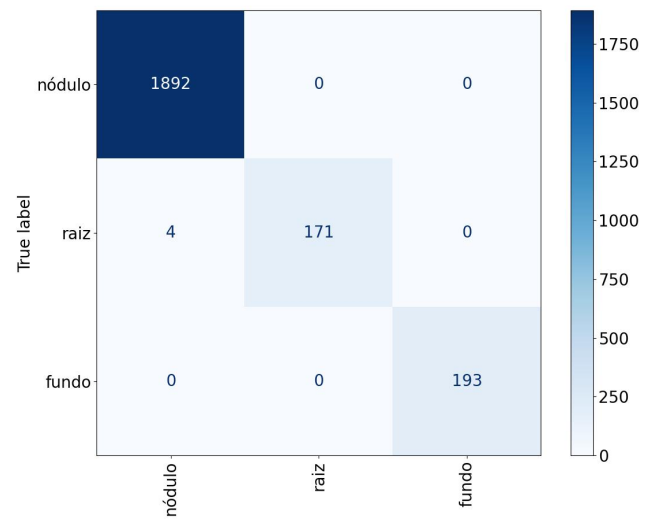
(a)



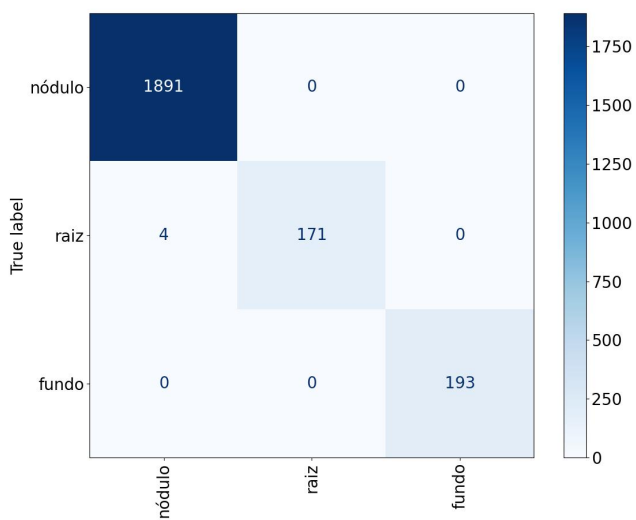
(b)



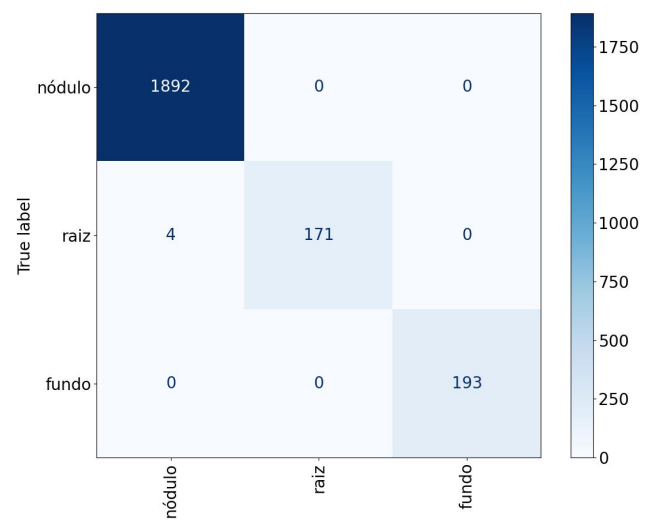
(c)



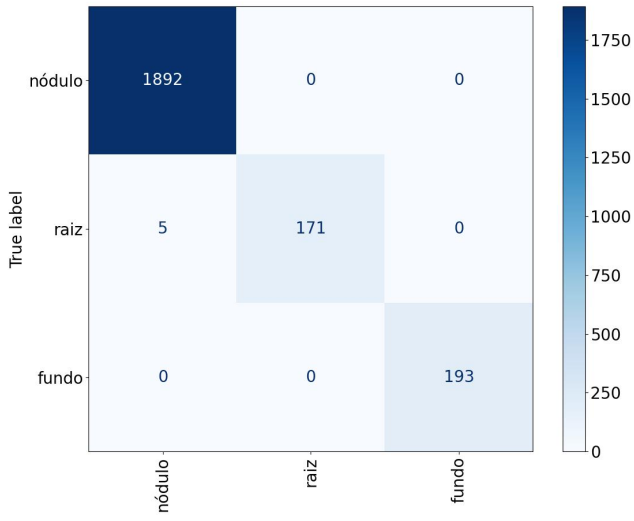
(d)



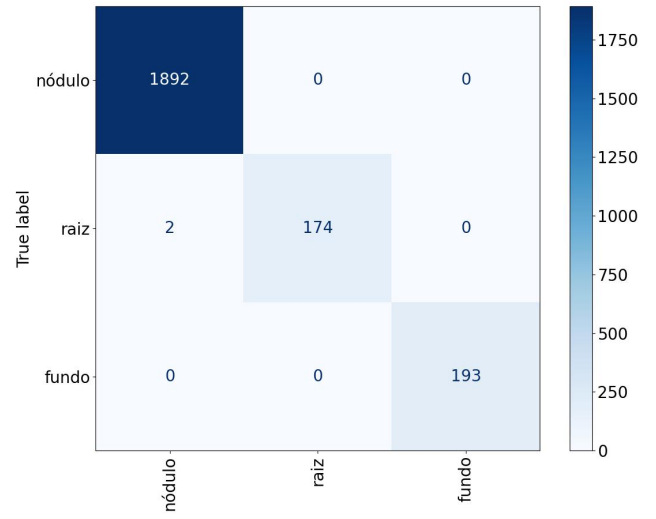
(e)



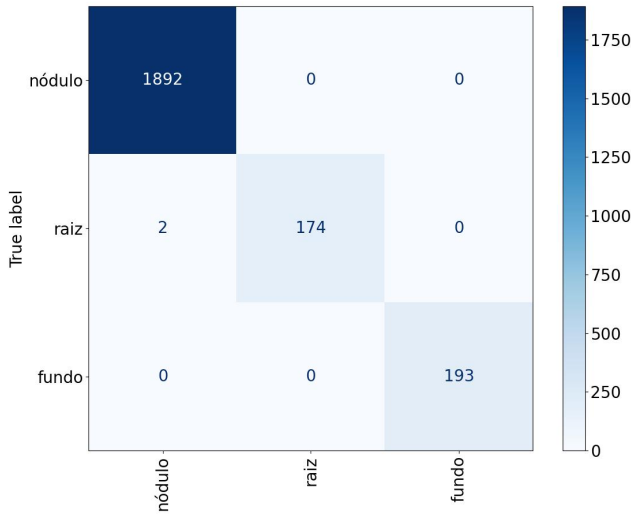
(f)



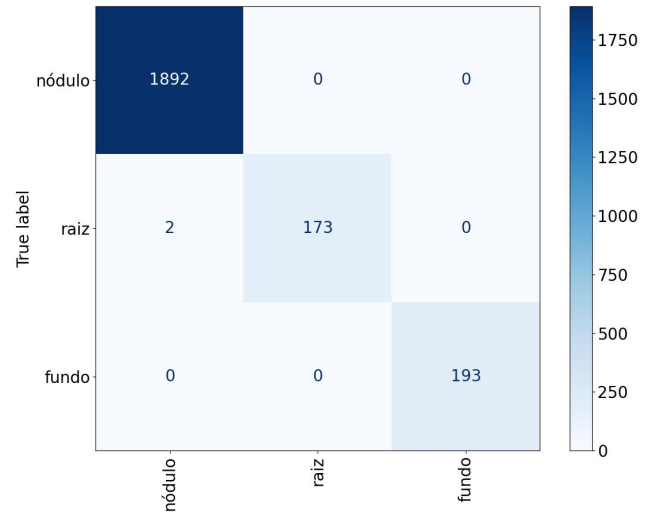
(g)



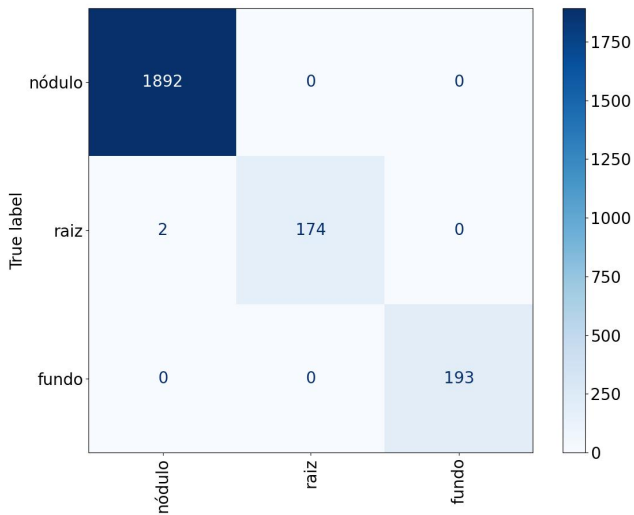
(h)



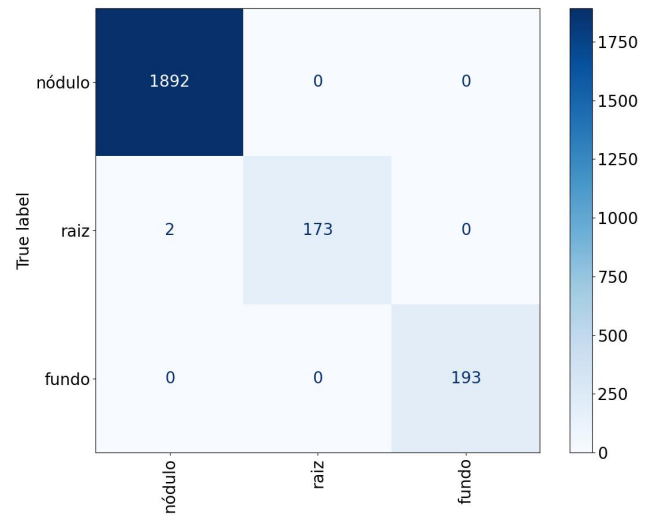
(i)



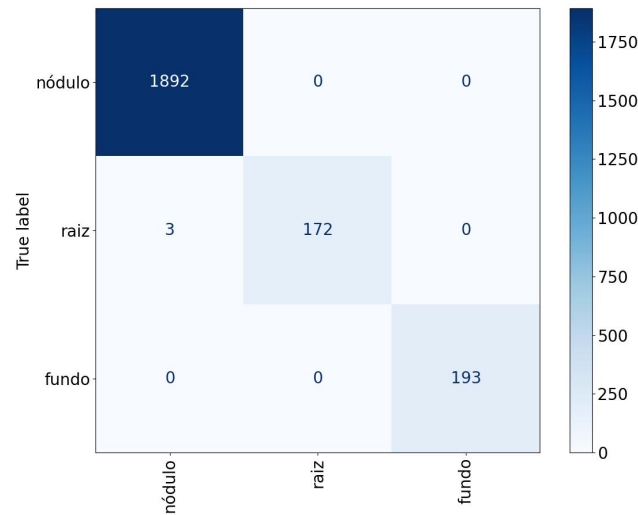
(j)



(k)



(l)

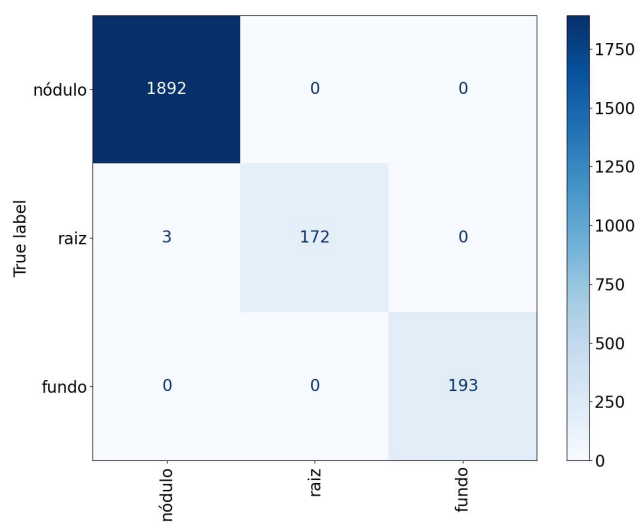


(m)

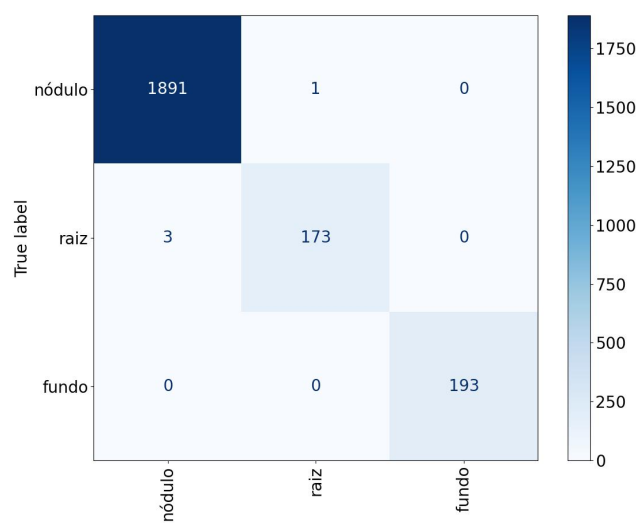
Figura 26: Média das matrizes de confusão obtidas no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas para os melhores resultados (i.e. maiores acurácias médias – Tabela 3). Os modelos foram treinados no conjunto de dados original. a) KNN e ResNet101; b) KNN e ResNet152; c) RF e GCH; d) RF e ResNet101; e) RF e ResNet152; f) RF e ResNet50; g) SVM (KL) e DenseNet121; h) SVM (KL) e ResNet101; i) SVM (KL) e ResNet152; j) SVM (KL) e ResNet50; k) SVM (KPG2) e ResNet101; l) SVM (KPG2) e ResNet152; m) SVM (KPG2) e ResNet50.

Tabela 21: Média das métricas (*loss*, sensibilidade, especificidade, precisão, F1 e *kappa*) \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados no conjunto de dados original.

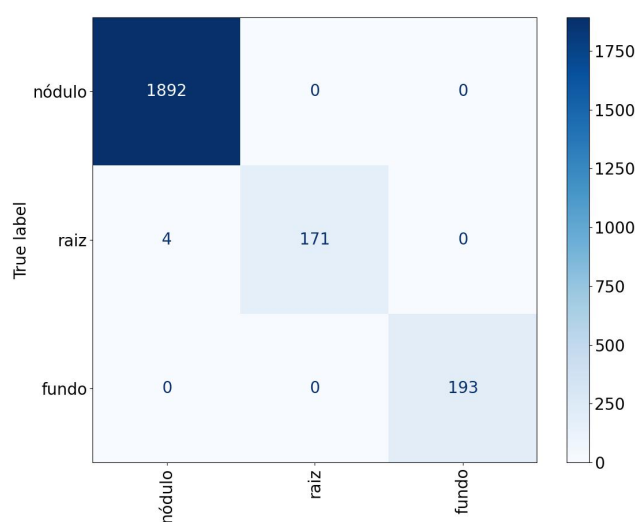
CNN	Conjunto de dados original					
	<i>Loss</i>	Sensibilidade	Especificidade	Precisão	F1	<i>Kappa</i>
VGG16	0,080 \pm 0,057	0,990 \pm 0,018	0,995 \pm 0,009	0,990 \pm 0,018	0,990 \pm 0,018	0,967 \pm 0,054
VGG19	0,171 \pm 0,335	0,976 \pm 0,054	0,988 \pm 0,027	0,976 \pm 0,054	0,976 \pm 0,054	0,937 \pm 0,125
DenseNet121	0,009 \pm 0,001	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,001
MobileNet	0,012 \pm 0,003	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,995 \pm 0,000
NASNetMobile	0,010 \pm 0,002	0,998 \pm 0,000	0,999 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	0,995 \pm 0,001
MobileNetV2	0,024 \pm 0,004	0,996 \pm 0,001	0,998 \pm 0,000	0,996 \pm 0,001	0,996 \pm 0,001	0,988 \pm 0,002
InceptionResNetV2	0,009 \pm 0,002	0,998 \pm 0,000	0,999 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	0,993 \pm 0,001
DenseNet169	0,011 \pm 0,001	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,995 \pm 0,001
DenseNet201	0,008 \pm 0,002	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,002
InceptionV3	0,007 \pm 0,001	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,997 \pm 0,001
ResNet101	0,192 \pm 0,016	0,936 \pm 0,011	0,968 \pm 0,006	0,936 \pm 0,011	0,936 \pm 0,011	0,754 \pm 0,046
ResNet101V2	0,018 \pm 0,003	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,001
ResNet152	0,271 \pm 0,012	0,899 \pm 0,018	0,950 \pm 0,009	0,899 \pm 0,018	0,899 \pm 0,018	0,545 \pm 0,137
ResNet152V2	0,013 \pm 0,003	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,000
ResNet50	0,198 \pm 0,024	0,940 \pm 0,021	0,970 \pm 0,011	0,940 \pm 0,021	0,940 \pm 0,021	0,748 \pm 0,125
ResNet50V2	0,012 \pm 0,002	0,998 \pm 0,000	0,999 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	0,994 \pm 0,001
Xception	0,010 \pm 0,002	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,995 \pm 0,002
NASNetLarge	0,008 \pm 0,002	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,002



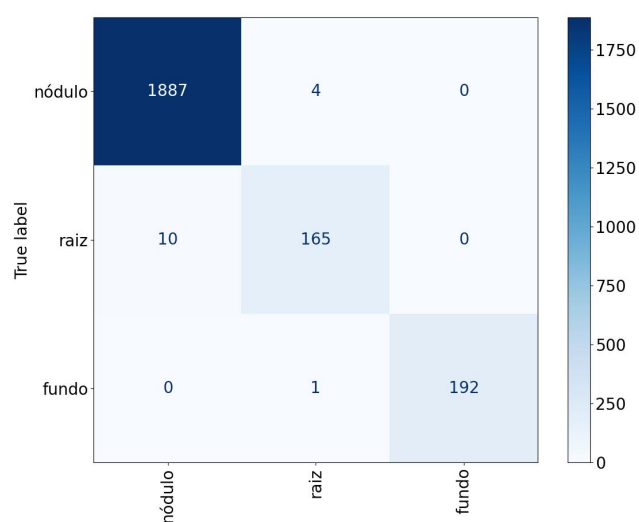
(a)



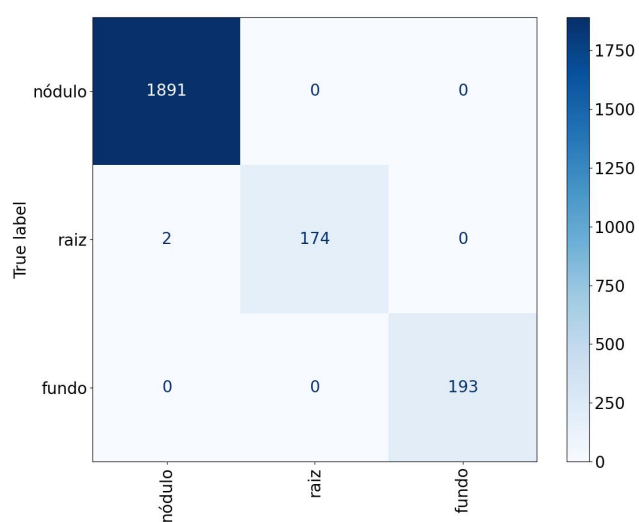
(b)



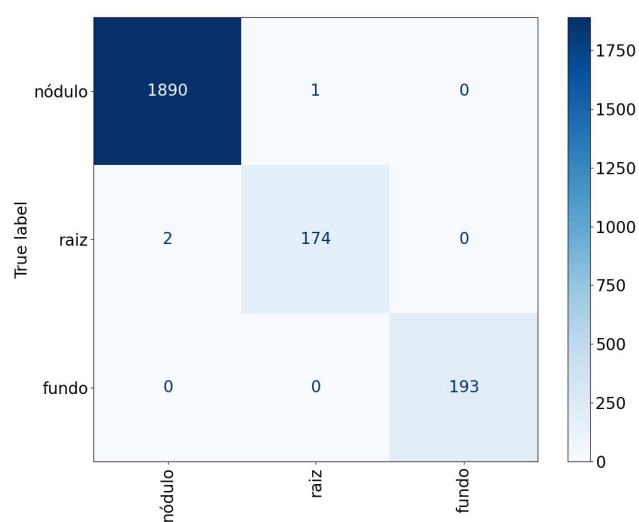
(c)



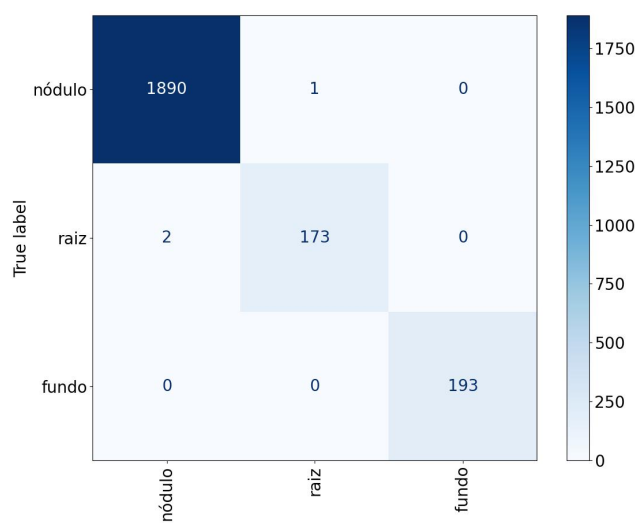
(d)



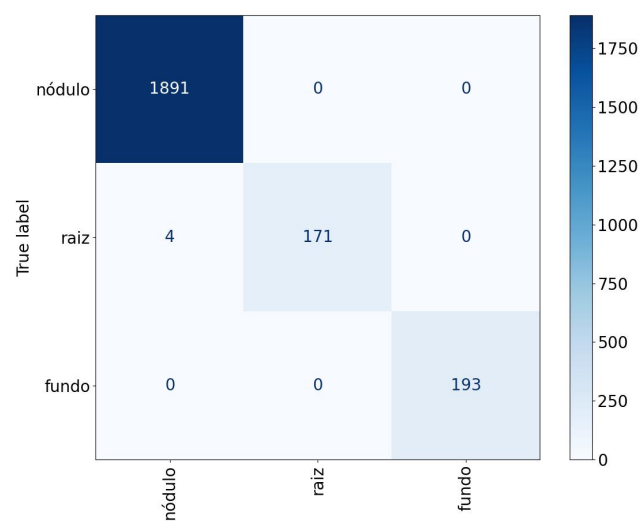
(e)



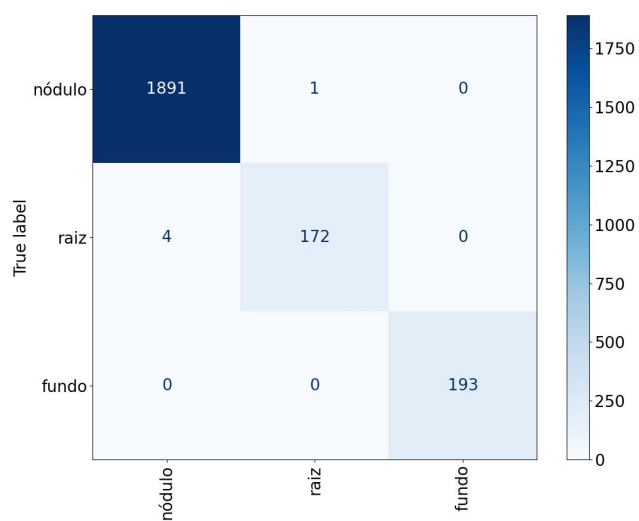
(f)



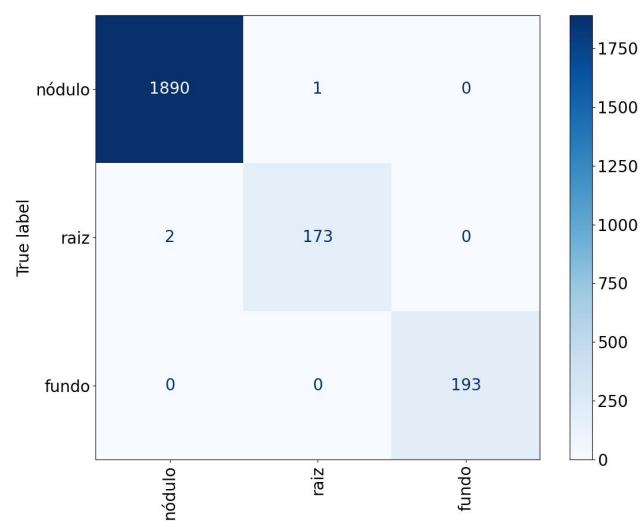
(g)



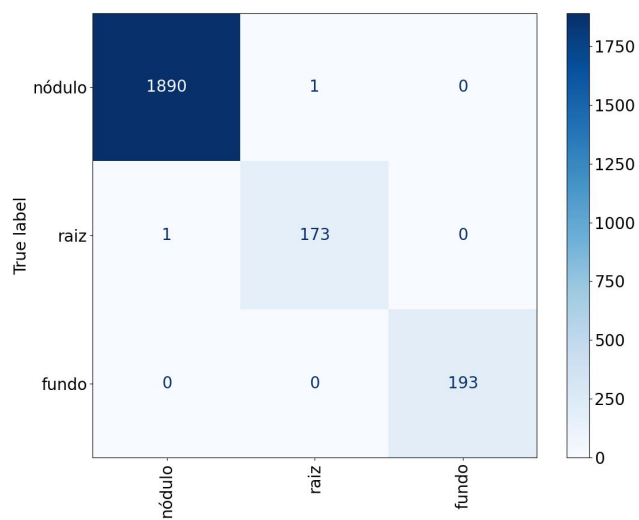
(h)



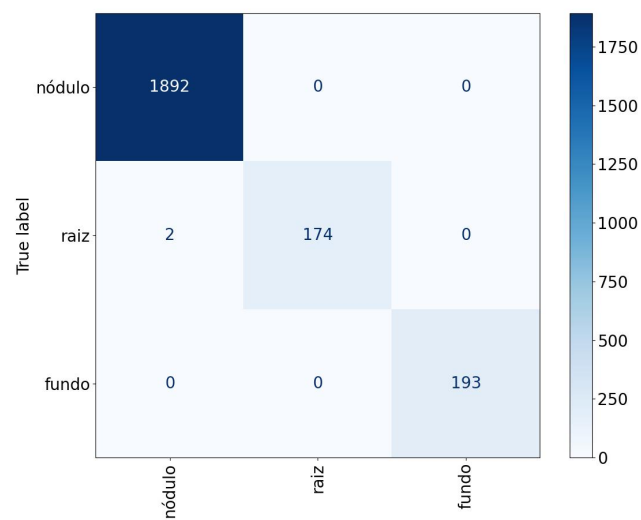
(i)



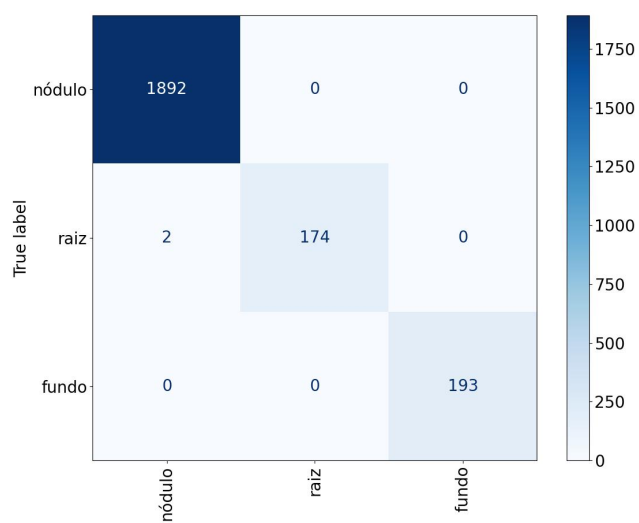
(j)



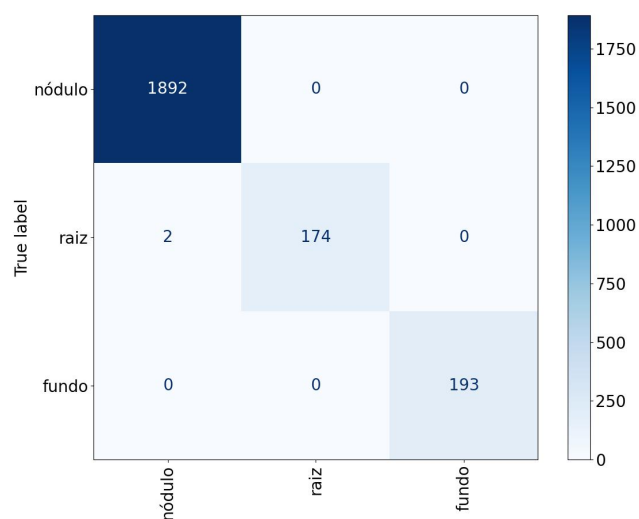
(k)



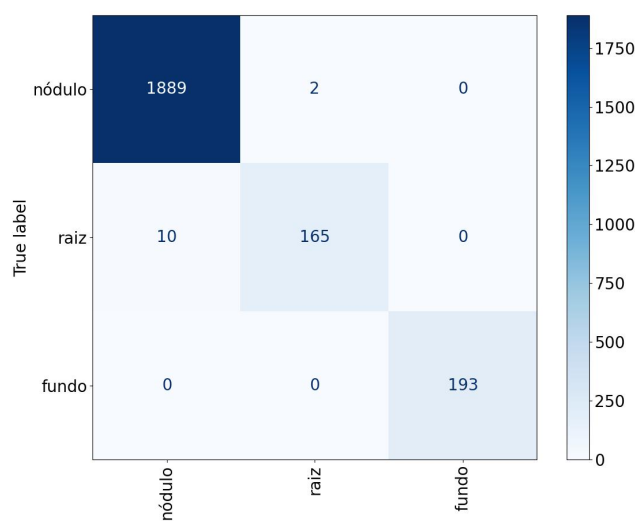
(l)



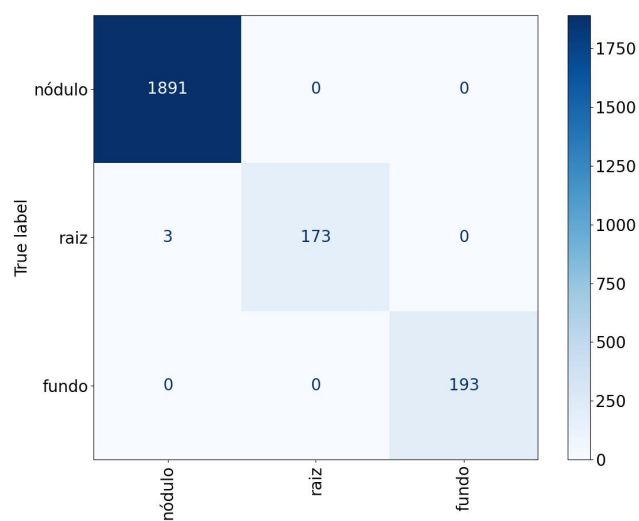
(m)



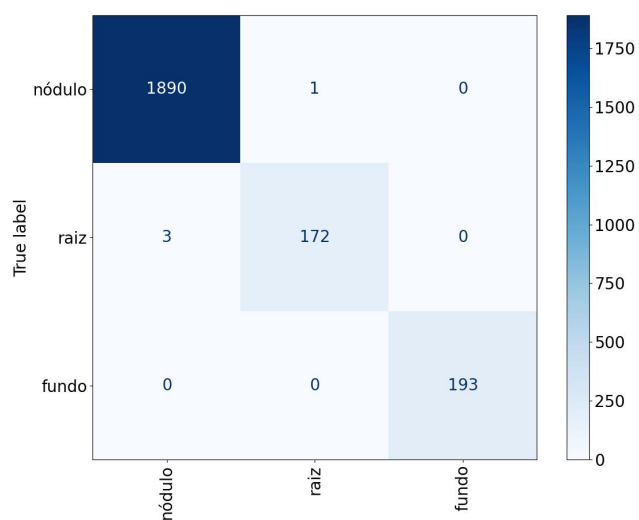
(n)



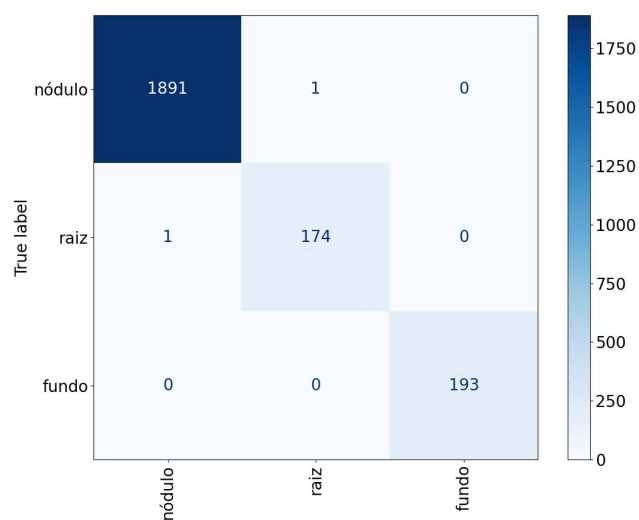
(o)



(p)



(q)



(r)

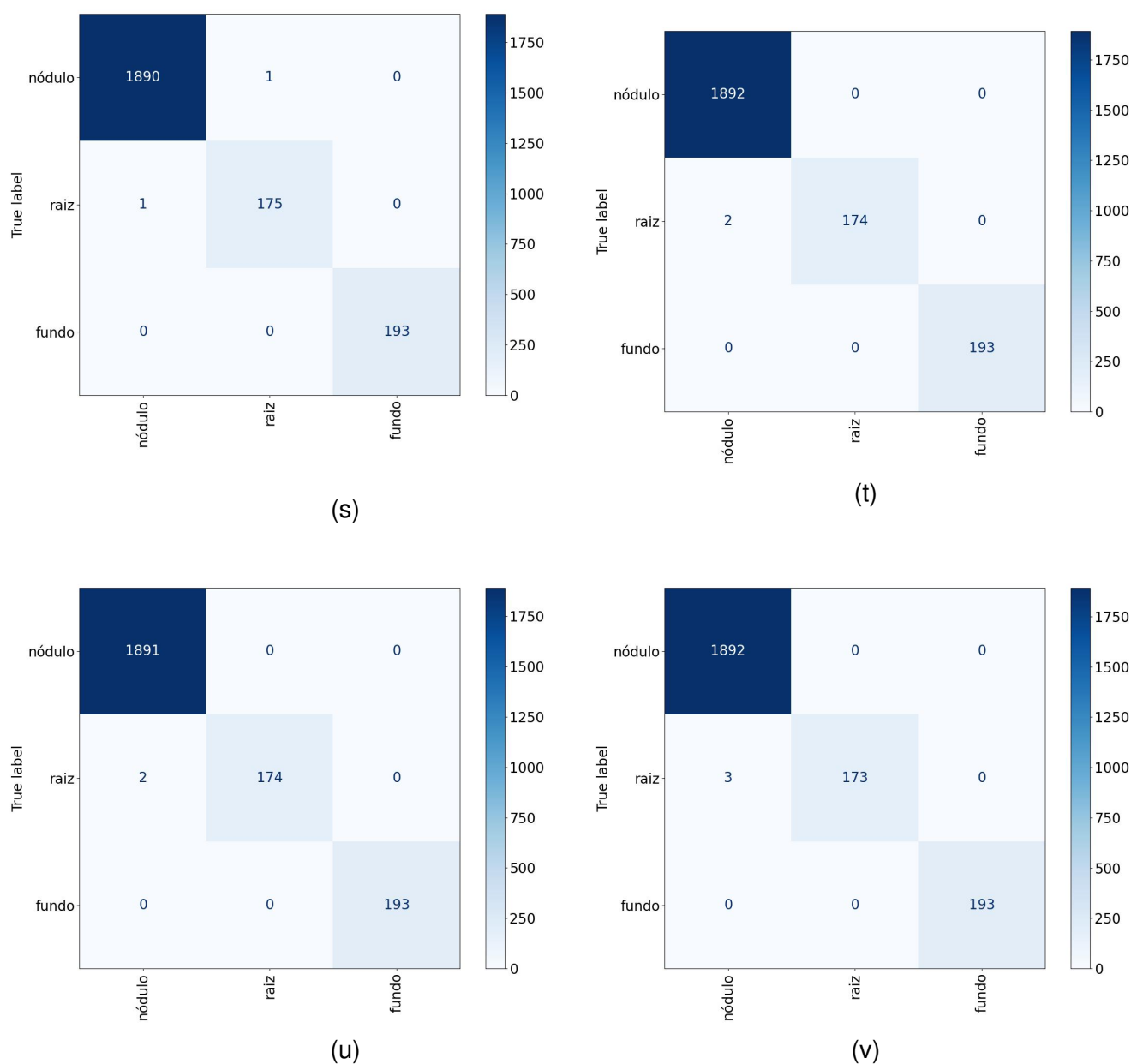
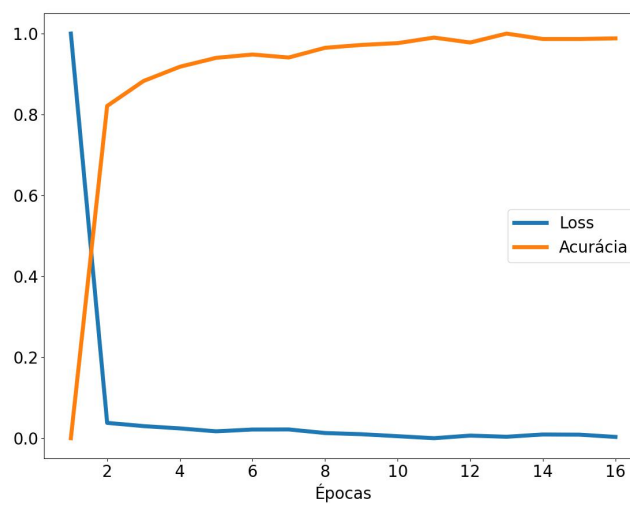


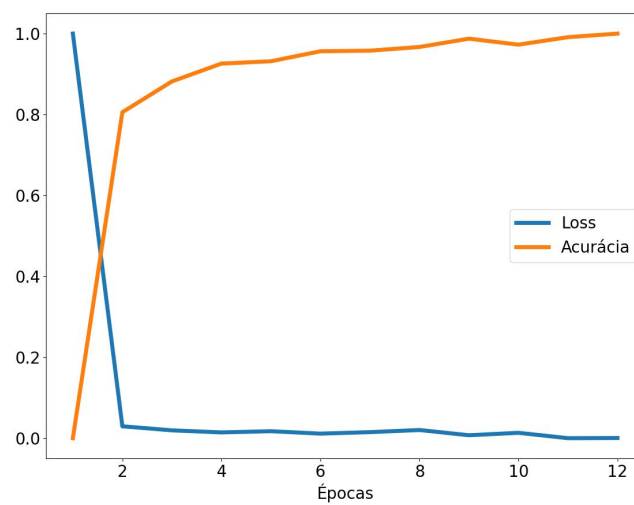
Figura 27: Média das matrizes de confusão obtidas no conjunto de dados de teste pelos classificadores KNN, RF, SVM (KL) e SVM (KPG2), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas para os melhores resultados (i.e. maiores acurácias médias – Tabela 4). Os modelos foram treinados no conjunto de dados aumentado. a) KNN e ResNet101; b) KNN e ResNet152; c) KNN e ResNet50; d) RF e FOM (Color); e) RF e ResNet101; f) RF e ResNet152; g) RF e ResNet50; h) SVM (KL) e VGG16; i) SVM (KL) e VGG19; j) SVM (KL) e MobileNet; k) SVM (KL) e MobileNetV2; l) SVM (KL) e ResNet101; m) SVM (KL) e ResNet152; n) SVM (KL) e ResNet50; o) SVM (KPG2) e FOM (Color); p) SVM (KPG2) e VGG16; q) SVM (KPG2) e VGG19; r) SVM (KPG2) e MobileNet; s) SVM (KPG2) e MobileNetV2; t) SVM (KPG2) e ResNet101; u) SVM (KPG2) e ResNet152 e v) SVM (KPG2) e ResNet50.

Tabela 22: Média das métricas (*loss*, sensibilidade, especificidade, precisão, F1 e *kappa*) \pm desvio padrão obtidos no conjunto de dados de teste usando classificação por redes neurais convolucionais. Os modelos foram treinados no conjunto de dados aumentado.

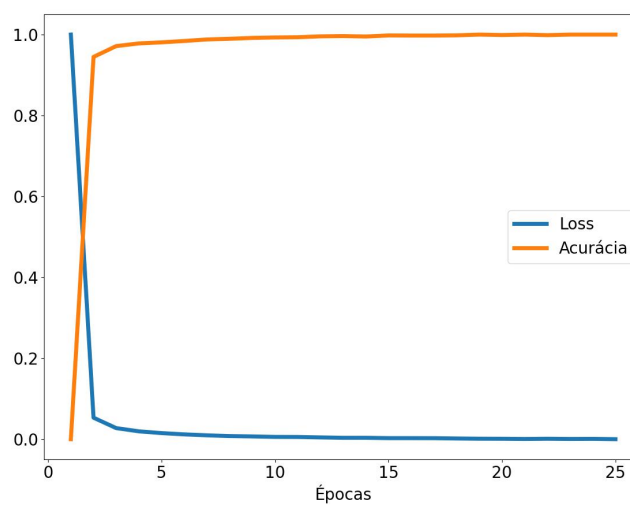
CNN	Conjunto de dados aumentado					
	<i>Loss</i>	Sensibilidade	Especificidade	Precisão	F1	<i>Kappa</i>
VGG16	0,159 \pm 0,139	0,995 \pm 0,004	0,998 \pm 0,002	0,995 \pm 0,004	0,995 \pm 0,004	0,983 \pm 0,013
VGG19	0,303 \pm 0,324	0,974 \pm 0,043	0,987 \pm 0,021	0,974 \pm 0,043	0,974 \pm 0,043	0,925 \pm 0,110
DenseNet121	0,010 \pm 0,001	0,998 \pm 0,000	0,999 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	0,995 \pm 0,001
MobileNet	0,016 \pm 0,003	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,995 \pm 0,001
NASNetMobile	0,009 \pm 0,001	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,002
MobileNetV2	0,034 \pm 0,004	0,996 \pm 0,000	0,998 \pm 0,000	0,996 \pm 0,000	0,996 \pm 0,000	0,985 \pm 0,001
InceptionResNetV2	0,011 \pm 0,002	0,998 \pm 0,000	0,999 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	0,994 \pm 0,001
DenseNet169	0,014 \pm 0,002	0,998 \pm 0,000	0,999 \pm 0,000	0,998 \pm 0,000	0,998 \pm 0,000	0,995 \pm 0,001
DenseNet201	0,008 \pm 0,001	1,000 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000
InceptionV3	0,010 \pm 0,001	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,997 \pm 0,000
ResNet101	0,374 \pm 0,126	0,851 \pm 0,069	0,926 \pm 0,034	0,851 \pm 0,069	0,851 \pm 0,069	0,627 \pm 0,103
ResNet101V2	0,018 \pm 0,003	0,999 \pm 0,000	1,000 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,997 \pm 0,000
ResNet152	0,592 \pm 0,152	0,744 \pm 0,080	0,872 \pm 0,040	0,744 \pm 0,080	0,744 \pm 0,080	0,451 \pm 0,083
ResNet152V2	0,018 \pm 0,002	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,001
ResNet50	0,307 \pm 0,071	0,898 \pm 0,032	0,949 \pm 0,016	0,898 \pm 0,032	0,898 \pm 0,032	0,700 \pm 0,065
ResNet50V2	0,011 \pm 0,002	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,995 \pm 0,001
Xception	0,011 \pm 0,002	1,000 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000	1,000 \pm 0,000	0,998 \pm 0,000
NASNetLarge	0,014 \pm 0,002	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,999 \pm 0,000	0,996 \pm 0,001



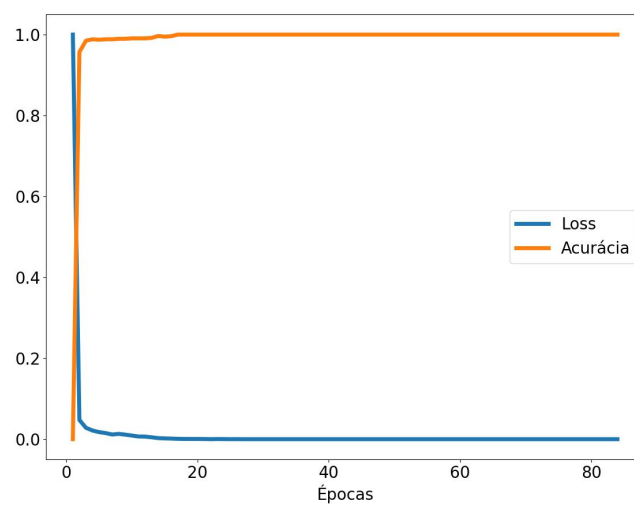
(a)



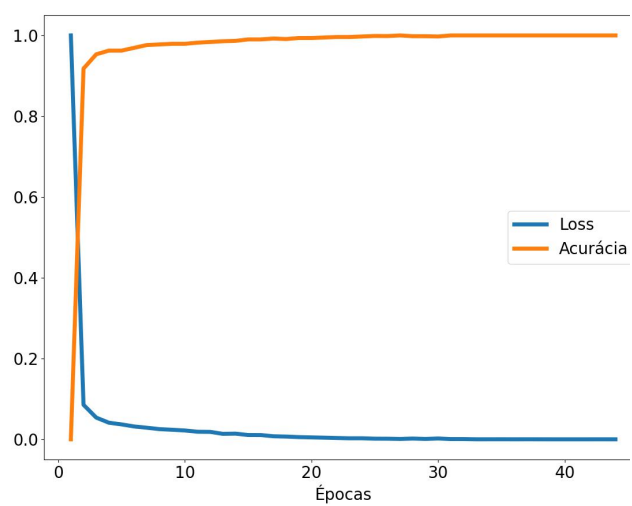
(b)



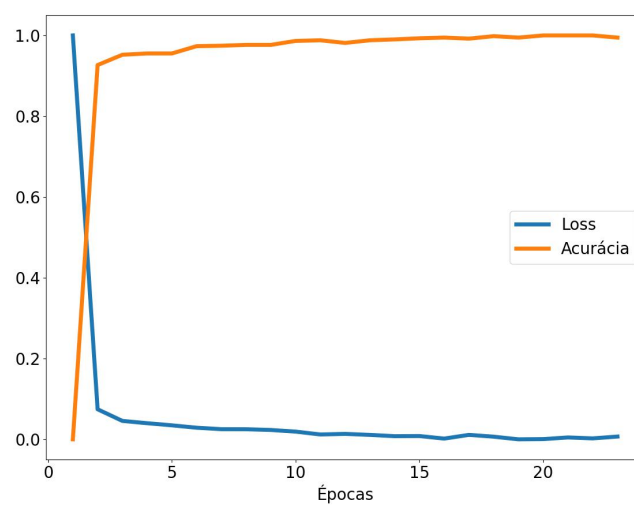
(c)



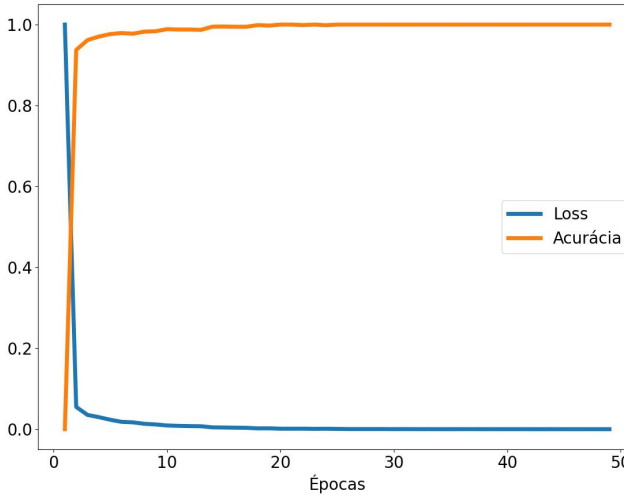
(d)



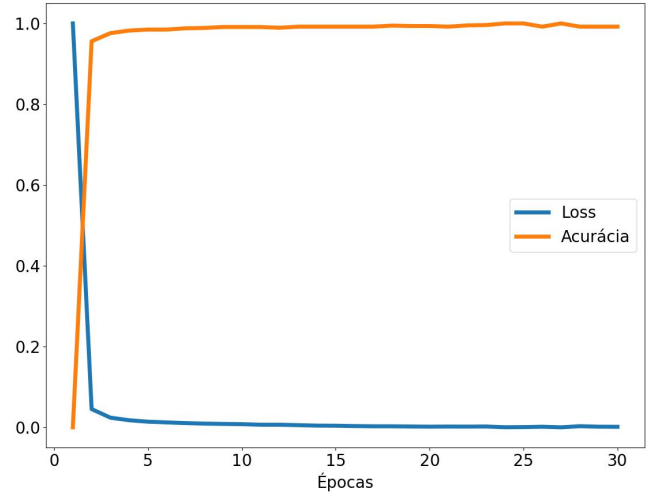
(e)



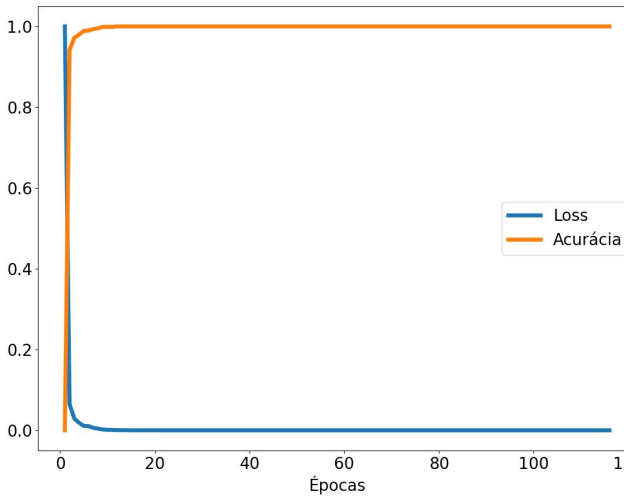
(f)



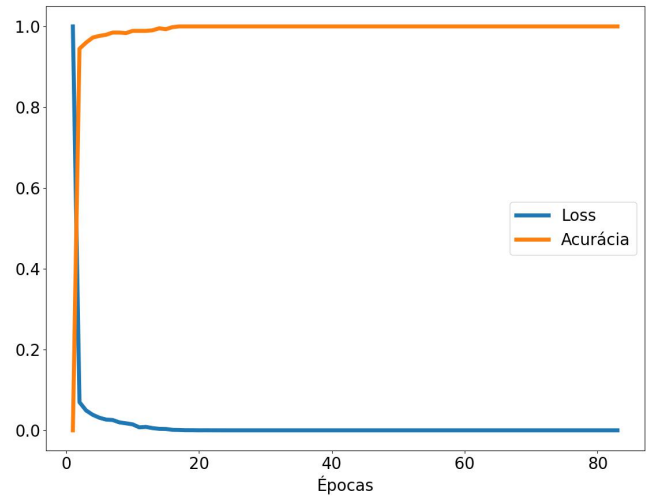
(g)



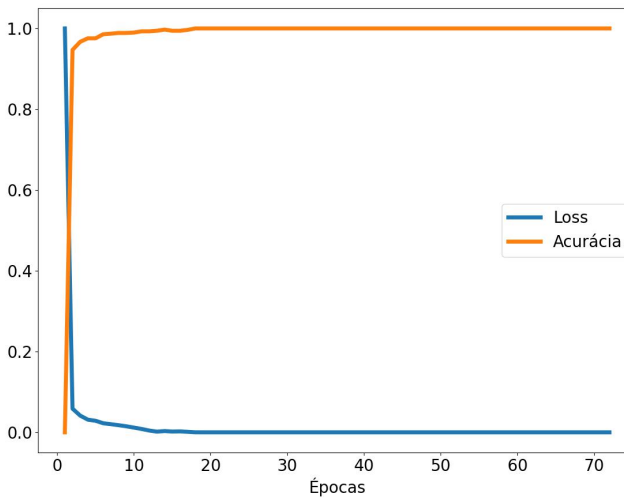
(h)



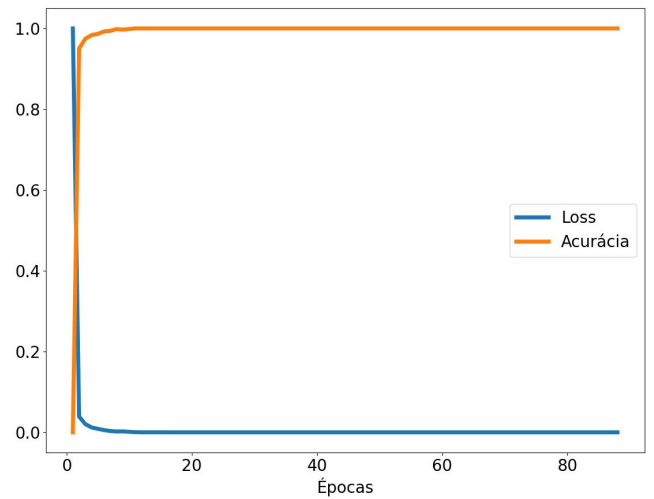
(i)



(j)



(k)



(l)

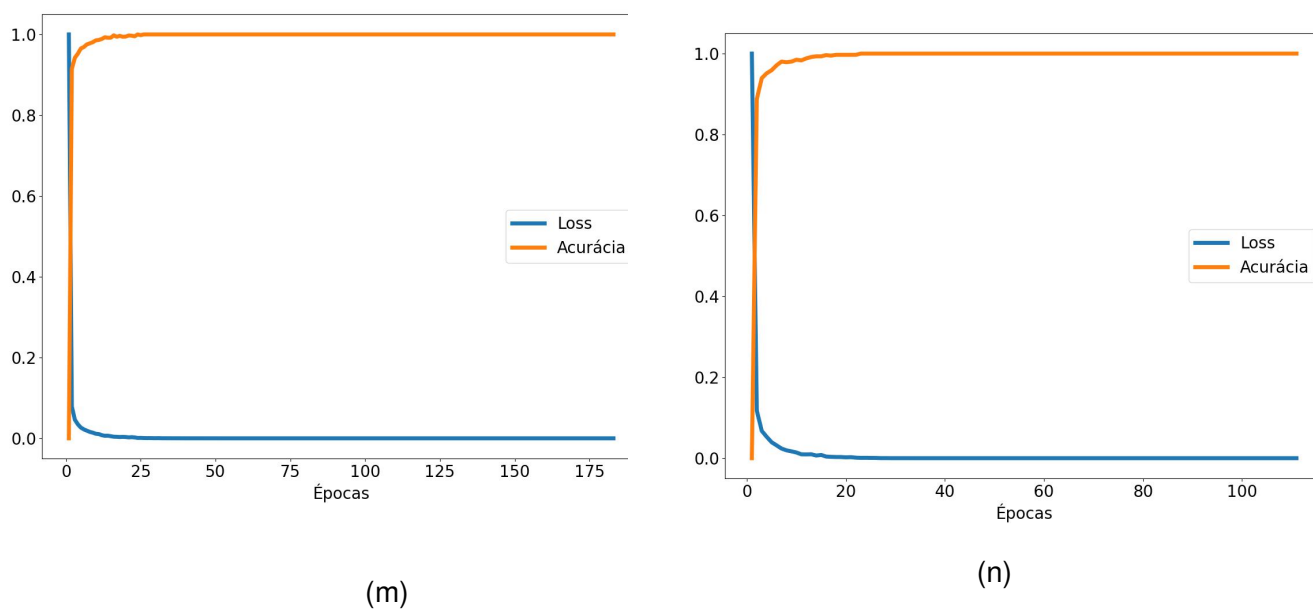
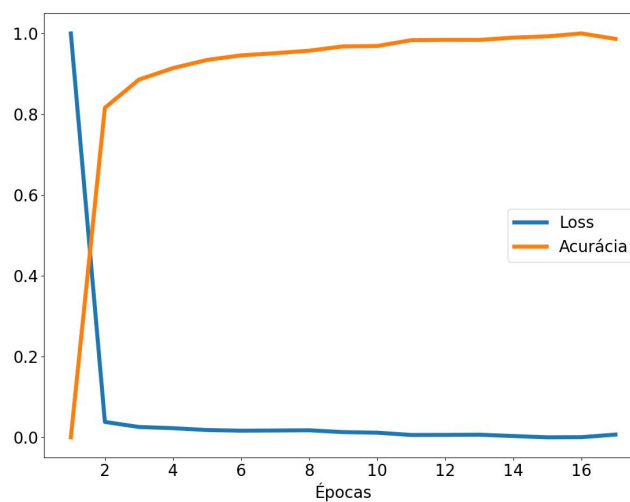
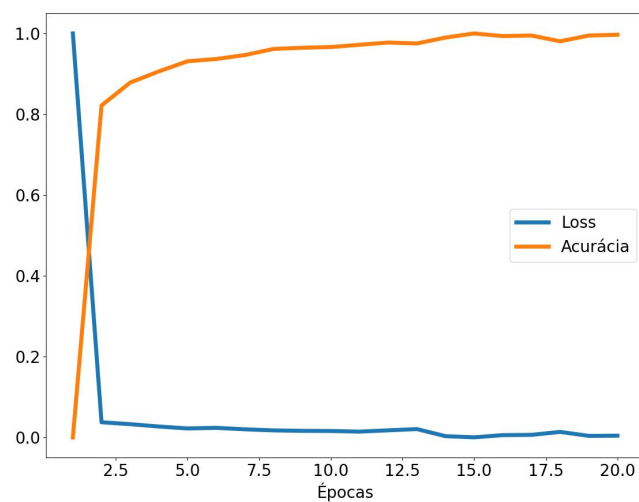


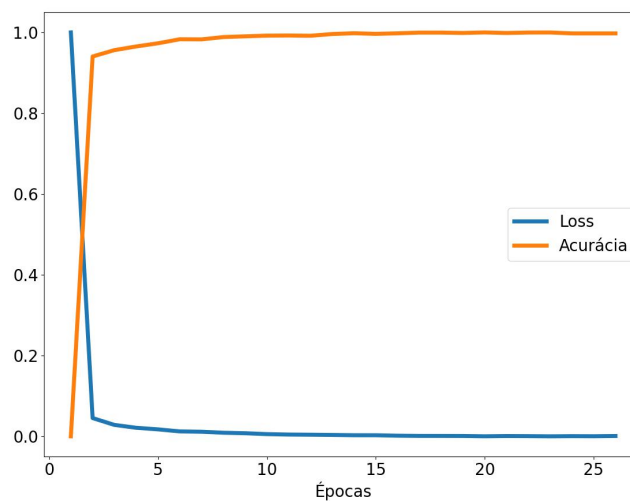
Figura 28: Progressão média das métricas acurácia x *loss* obtidas durante o treinamento das redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados no conjunto de dados original. a) VGG16; b) VGG19; c) DenseNet121; d) MobileNet; e) NASNetMobile; f) InceptionResNetV2; g) DenseNet169; h) DenseNet201; i) InceptionV3; j) ResNet101V2; k) ResNet152V2; l) ResNet50V2; m) Xception e n) NASNetLarge.



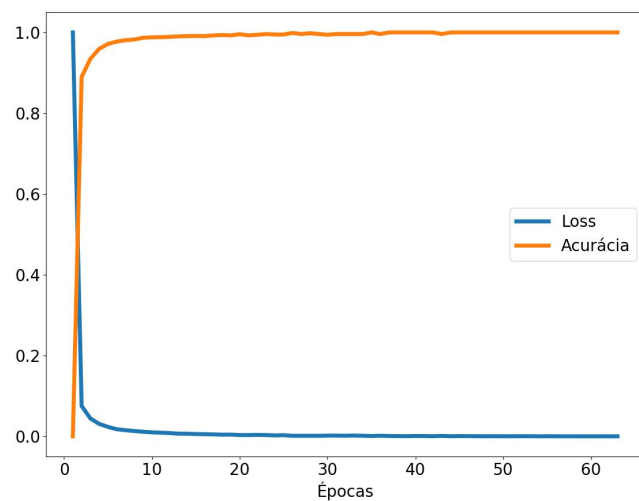
(a)



(b)

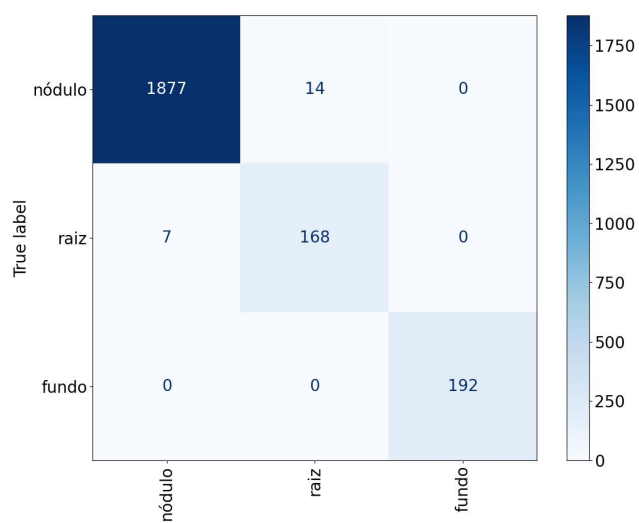


(c)

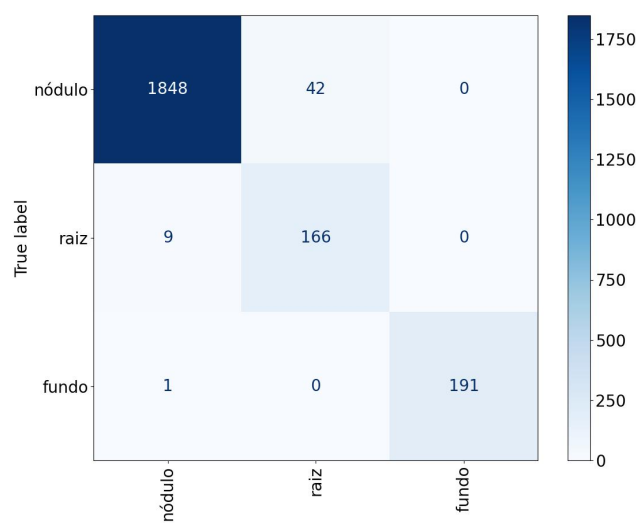


(d)

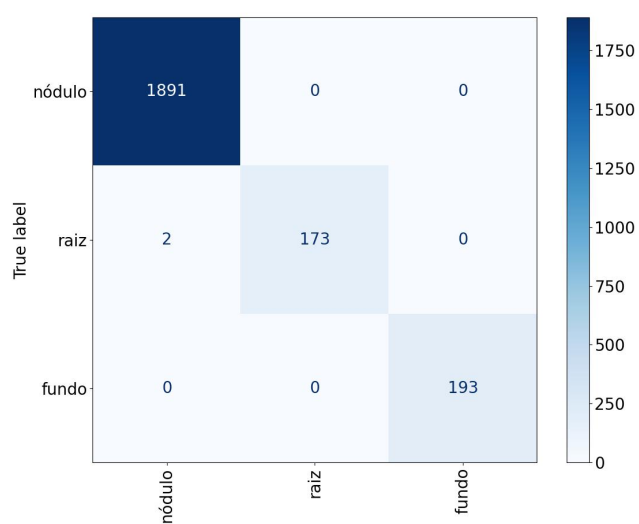
Figura 29: Progressão média das métricas acurácia x *loss* obtidas durante o treinamento das redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados no conjunto de dados aumentado. a) VGG16; b) VGG19; c) DenseNet201 e d) Xception.



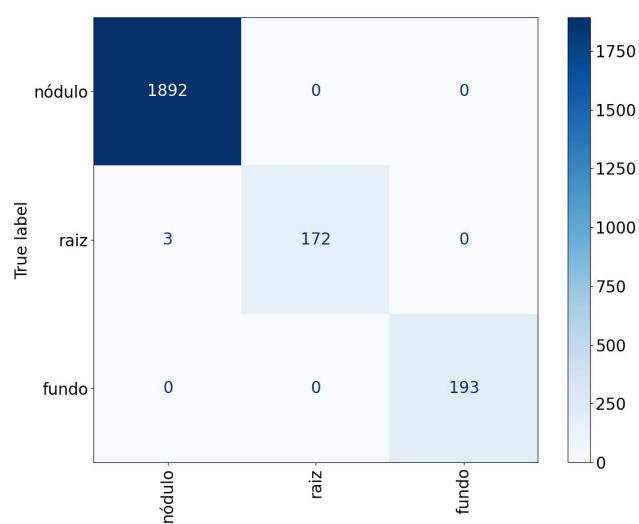
(a)



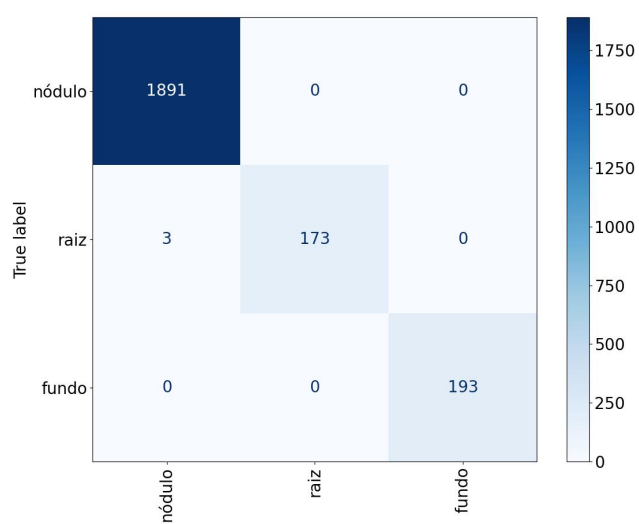
(b)



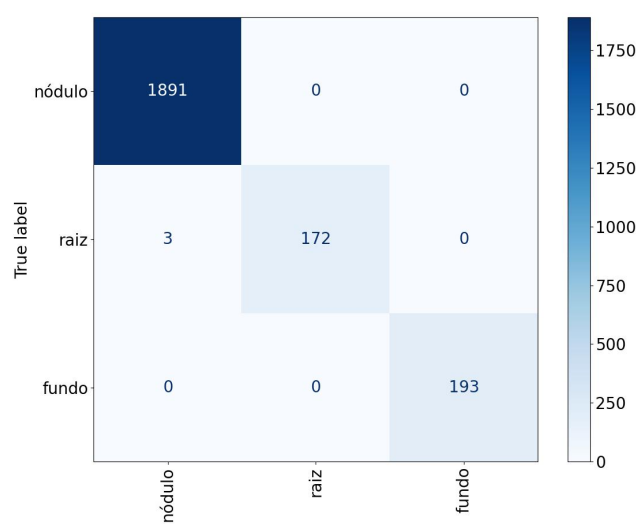
(c)



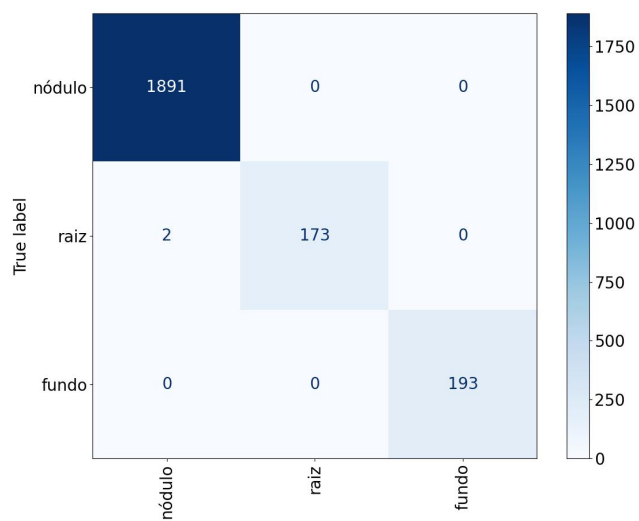
(d)



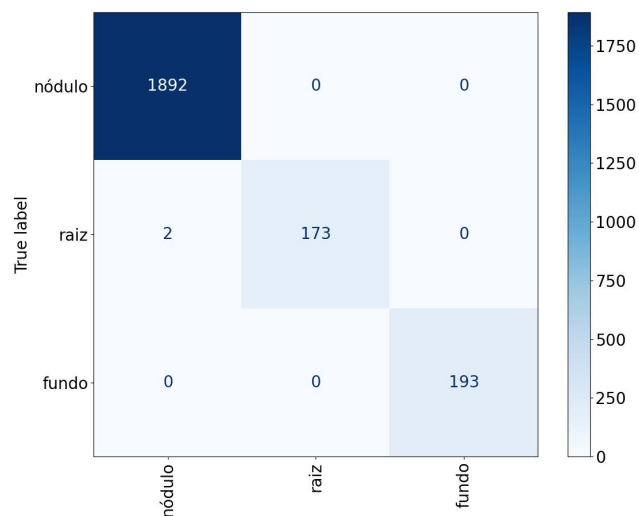
(e)



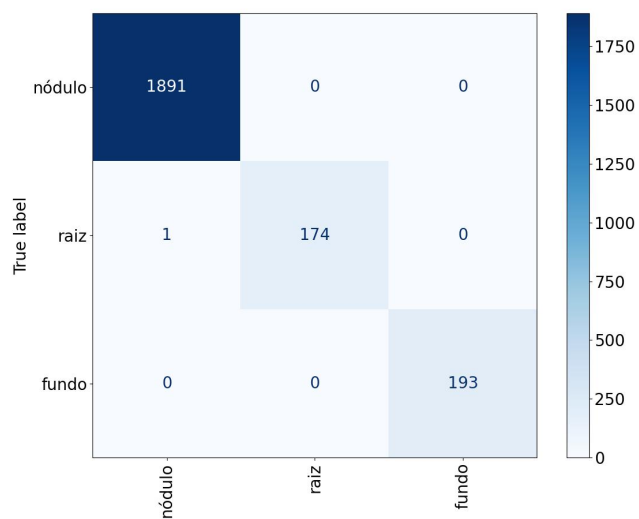
(f)



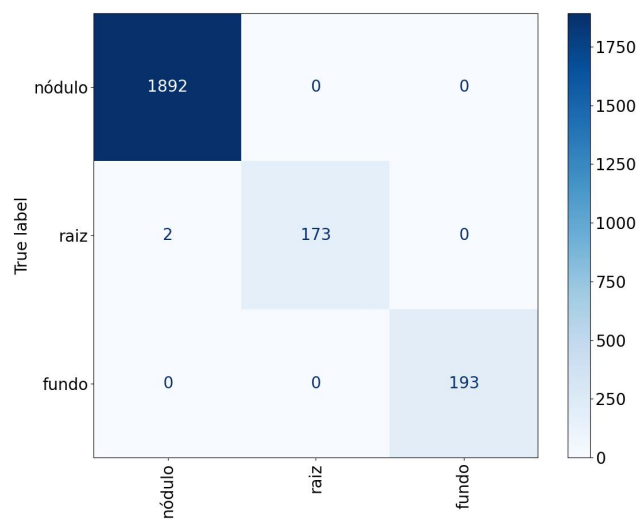
(g)



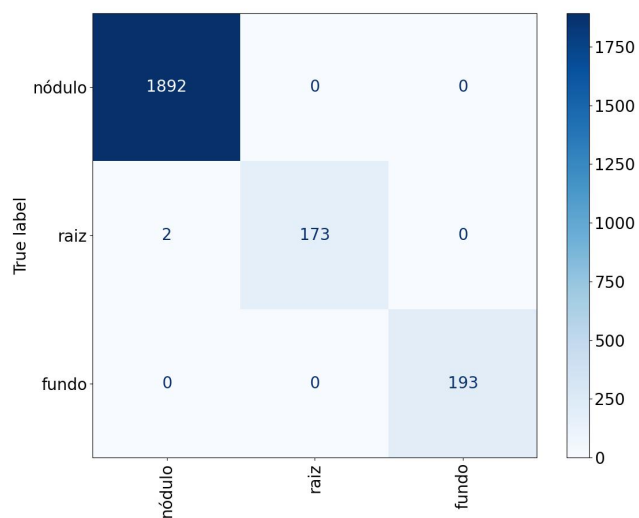
(h)



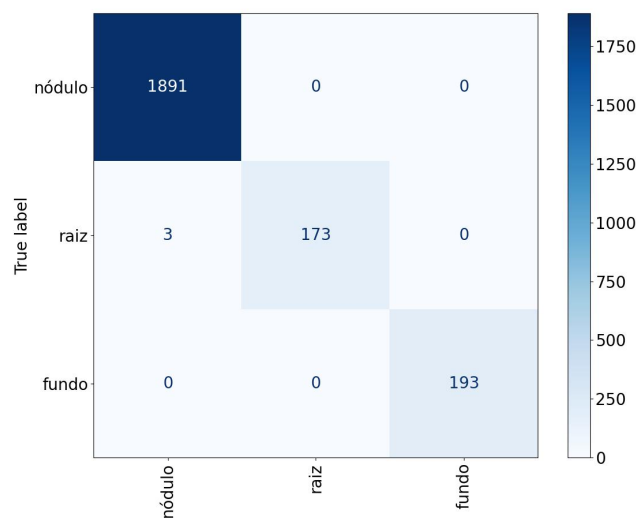
(i)



(j)



(k)



(l)

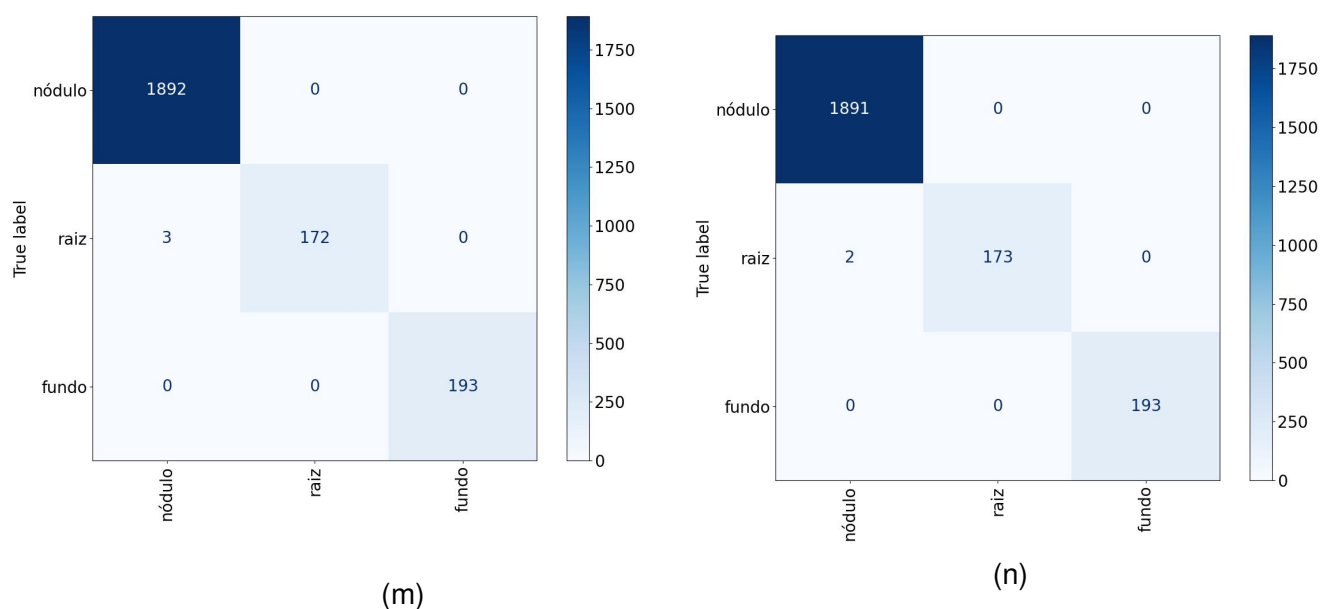


Figura 30: Média das matrizes de confusão obtidas no conjunto de dados de teste usando classificação por redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados no conjunto de dados original. a) VGG16; b) VGG19; c) DenseNet121; d) MobileNet; e) NASNetMobile; f) InceptionResNetV2; g) DenseNet169; h) DenseNet201; i) InceptionV3; j) ResNet101V2; k) ResNet152V2; l) ResNet50V2; m) Xception e n) NASNetLarge.

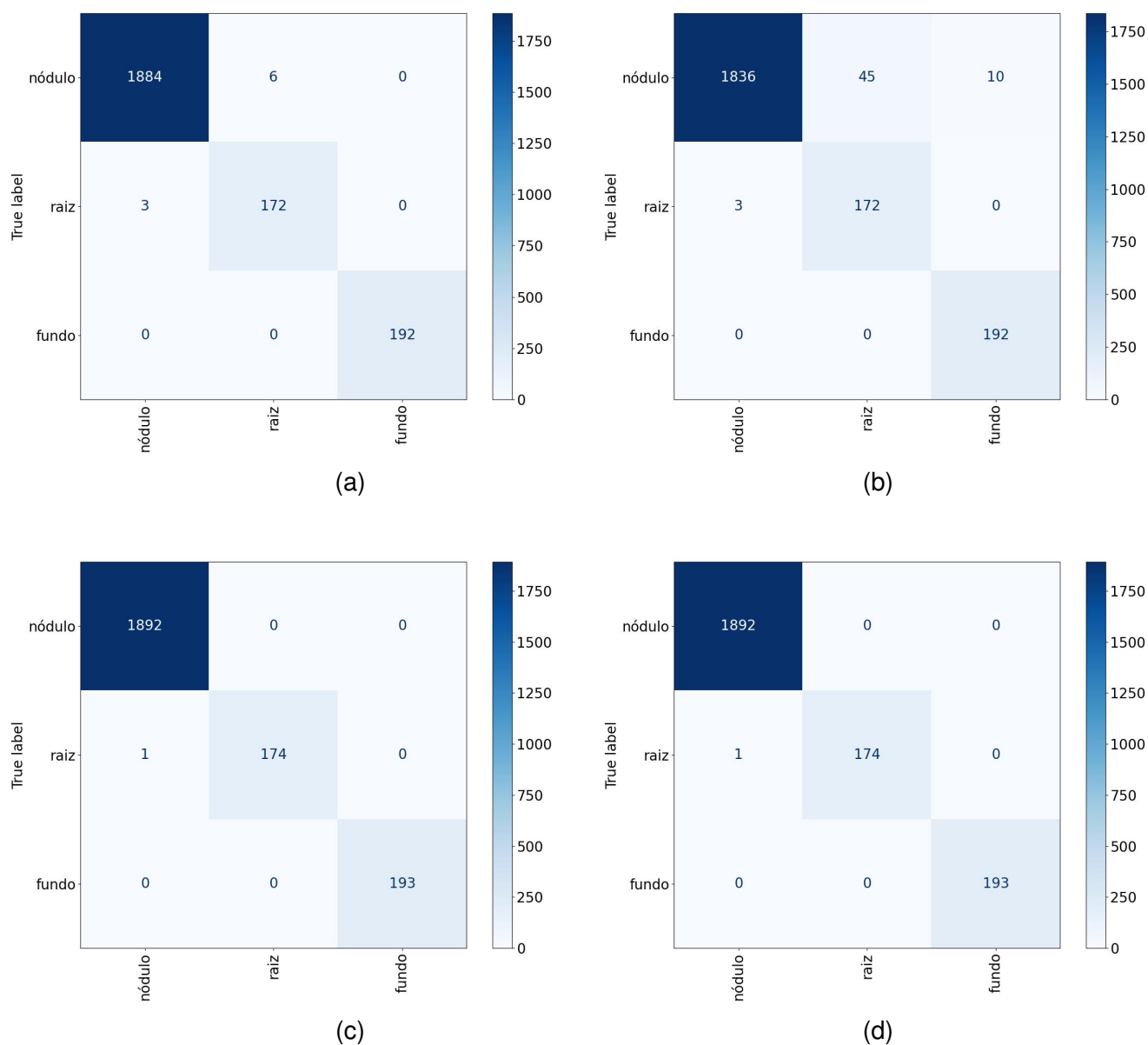


Figura 31: Média das matrizes de confusão obtidas no conjunto de dados de teste usando classificação por redes neurais convolucionais para os melhores resultados (i.e. maiores acurácias médias – Tabela 6). Os modelos foram treinados no conjunto de dados aumentado. a) VGG16; b) VGG19; c) DenseNet201 e d) Xception.

