

UNIVERSIDADE ESTADUAL DE PONTA GROSSA
SETOR DE CIÊNCIAS AGRÁRIAS E DE TECNOLOGIA
DEPARTAMENTO DE INFORMÁTICA

GEORGE WILBER DE BESSA OLIVEIRA

REDES NEURAS CONVOLUCIONAIS E AMPLIAÇÃO DE DADOS PARA
DETECÇÃO DE ANTRACNOSE EM FOLHAS DE FEIJOEIRO

PONTA GROSSA
2021

GEORGE WILBER DE BESSA OLIVEIRA

REDES NEURAIS CONVOLUCIONAIS E AMPLIAÇÃO DE DADOS PARA
DETECÇÃO DE ANTRACNOSE EM FOLHAS DE FEIJOEIRO

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Computação Aplicada, curso de Mestrado em Computação Aplicada da Universidade Estadual de Ponta Grossa, como requisito parcial para obtenção do título de Mestre.

Orientador: Prof. Dr. Marcelo Giovanetti Canteri
Coorientador: Prof. Dr. José Carlos Ferreira

PONTA GROSSA
2021

B557 Bessa Oliveira, George Wilber de
Redes neurais convolucionais e ampliação de dados para detecção de antracnose em folhas de feijoeiro / George Wilber de Bessa Oliveira. Ponta Grossa, 2021.
66 f.

Dissertação (Mestrado em Computação Aplicada - Área de Concentração: Computação para Tecnologias em Agricultura), Universidade Estadual de Ponta Grossa.

Orientador: Prof. Dr. Marcelo Giovanetti Canteri.
Coorientador: Prof. Dr. José Carlos Ferreira da Rocha.

1. Redes neurais. 2. Ampliação de dados. 3. Antracnose. 4. Feijão. I. Canteri, Marcelo Giovanetti. II. Rocha, José Carlos Ferreira da. III. Universidade Estadual de Ponta Grossa. Computação para Tecnologias em Agricultura. IV.T.

CDD: 004



UNIVERSIDADE ESTADUAL DE PONTA GROSSA
Av. General Carlos Cavalcanti, 4748 - Bairro Uvaranas - CEP 84030-900 - Ponta Grossa - PR - <https://uepg.br>

TERMO

TERMO DE APROVAÇÃO

George Wilber de Bessa Oliveira

REDES NEURAIAS CONVOLUCIONAIS E AMPLIAÇÃO DE DADOS PARA DETECÇÃO DE ANTRACNOSE EM FOLHAS DE FEIJOEIRO

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Computação Aplicada da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora:

Prof. Dr. Marcelo Giovanetti Canteri - UEPG

Prof. Dr. Hugo Valadares Siqueira - UTFPR

Prof. Dr. Eduardo Fávero Caires - UEPG

Ponta Grossa, 26 de fevereiro de 2021.



Documento assinado eletronicamente por **Hugo Valadares Siqueira, Usuário Externo**, em 26/02/2021, às 17:23, conforme art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por **Eduardo Favero Caires, Professor(a)**, em 26/02/2021, às 17:23, conforme art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por **Marcelo Giovanetti Canteri, Usuário Externo**, em 26/02/2021, às 17:23, conforme art. 1º, III, "b", da Lei 11.419/2006.



A autenticidade do documento pode ser conferida no site <https://sei.uepg.br/autenticidade> informando o código verificador **0414888** e o código CRC **7F3879DD**.

*Aos corações inquietos
Às mentes pensantes
Às mulheres
Operárias
Camponesas
Mães
Aos que acolhem e ensinam
A todo aquele que luta por um mundo melhor
Aos que fortalecem o ensino público, universal,
gratuito e de qualidade*

Dedico.

AGRADECIMENTOS

Aos meus amigos e familiares, pelo incentivo e apoio incondicional.

Aos professores Dr. José Carlos Rocha e Dr. Marcelo Canteri, respectivamente coorientador e orientador deste trabalho, por me guiarem com paciência e receptividade.

Aos professores da banca de avaliação Dr. Eduardo Caires e Dr. Hugo Siqueira pela contribuição com sugestões para aprimoramentos no trabalho final.

Aos professores do Programa de Pós Graduação em Computação Aplicada e aos colaboradores da Universidade Estadual de Ponta Grossa por proporcionarem, através do seu trabalho diário, o desenvolvimento humano e intelectual de milhares de estudantes.

A CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
- pelo apoio financeiro para o desenvolvimento desta pesquisa.

A todos que direta ou indiretamente contribuíram para a conclusão deste trabalho.

RESUMO

Modelos de aprendizagem profunda, em especial as Redes Neurais Convolucionais, são atualmente protagonistas de expressivos avanços nas tarefas de visão computacional. A disponibilidade de computadores robustos, dotados de grande capacidade de processamento, possibilitou a ampliação dos campos de trabalho nessa área e as redes neurais aumentaram ainda mais as possibilidades. Entretanto, os modelos de aprendizagem profunda comumente necessitam de grandes bases de dados para o aprendizado de características relevantes das imagens e a elaboração de bases de dados de doenças de plantas requer o empenho de uma considerável força de trabalho e o acesso a uma vasta coleção de imagens, além de razoáveis condições de captura. Os métodos de ampliação de dados servem ao propósito de aumentar o número de itens de uma base de dados através de modificações nas imagens existentes e, por vezes, a geração de novas imagens com base nas características observadas em parte do conjunto. Este estudo empreendeu a coleta de um conjunto de imagens de folíolos de feijoeiro para a elaboração de bases de dados e a posterior condução de testes com diferentes arquiteturas de redes neurais e configurações de técnicas de ampliação de dados com o propósito de avaliar a influência de tais técnicas na simulação de imagens coletadas em condições distintas, bem como sua influência nas medidas de desempenho de acurácia, sensibilidade, especificidade e sobreajuste no intuito de distinguir imagens de partes de folíolos saudáveis daquelas contendo partes afetadas pela doença fúngica antracnose em diferentes estágios. Os resultados mostram que as técnicas de ampliação de dados possibilitam a simulação de diferentes condições de aquisição das imagens, e os modelos de classificação baseados em redes neurais convolucionais alcançaram resultados variando entre 60 e 90% de acurácia nas diferentes configurações experimentais. A bibliografia consultada, em conjunto com os resultados apresentados, dá margem à sugestão de novos trabalhos na mesma temática, considerando o incremento do número de imagens no conjunto e o uso de outras técnicas de ampliação de dados e arquiteturas de redes neurais.

Palavras-chave: Redes Neurais. Ampliação de dados. Antracnose. Feijão.

ABSTRACT

Deep learning models, especially Convolutional Neural Networks, are currently the protagonists of significant advances in computer vision tasks. The availability of robust computers, with great processing capacity, made it possible to expand the fields of work in this area and neural networks further increased the possibilities. However, deep learning models commonly require large databases to learn relevant characteristics of the images, and building plant disease databases requires the commitment of a considerable workforce and access to a vast collection of information. images, in addition to reasonable capture conditions. Data augmentation methods serve the purpose of increasing the number of items in a database through modifications to existing images and, sometimes, the generation of new images based on the characteristics observed in part of the set. This study undertook the collection of a set of images of bean leaflets for the elaboration of databases and the subsequent conduction of tests with different neural network architectures and configurations of data augmentation techniques for the purpose to evaluate the influence of such techniques in the simulation of images collected under different conditions, as well as their influence on the performance measures of accuracy, sensitivity, specificity and overfitting in order to distinguish images of parts of healthy leaflets from those containing structures affected by anthracnose fungal disease at different stages. The results show that the data expansion techniques allow the simulation of different conditions of image acquisition, and the classification models based on convolutional neural networks achieved results varying between 60 and 90% accuracy in the different experimental configurations. The bibliography consulted in conjunction with the results presented, gives rise to the suggestion of new works on the same theme, considering the expansion of the database and the use of other techniques for data augmentation and neural network architectures.

Keywords: Neural Networks. Data Augmentation. Anthracnose. Common bean.

LISTA DE FIGURAS

Figura 1 – Par de folíolos saudáveis	16
Figura 2 – Par de folíolos com sinais de antracnose.....	16
Figura 3 – Escala diagramática da severidade da antracnose em folíolos de feijão .	17
Figura 4 – Escala diagramática da severidade da antracnose em trifólios de feijão .	18
Figura 5 – Representação esquemática do neurônio biológico.....	22
Figura 6 – Modelo de neurônio base para redes neurais	22
Figura 7 – Exemplo de convolução 2D.....	24
Figura 8 - Filtro 3 x 3 com <i>stride</i> de dois pixels	25
Figura 9 - Exemplo de preenchimento <i>zero-padding</i>	26
Figura 10 – Exemplo de subamostragem (<i>pooling</i>) médio e máximo.....	26
Figura 11 – Camada totalmente conectada	28
Figura 12 – Regularização <i>dropout</i> em rede neural com duas camadas ocultas	29
Figura 13 – Arquitetura da CNN LeNet-5, proposta por LeCun et al. em 1998	30
Figura 14 – Arquitetura da CNN AlexNet, proposta por Krizhevsky et al. em 2012...	31
Figura 15 - Exemplo de um módulo <i>Inception</i>	32
Figura 16 – Comparação entre arquiteturas de redes neurais	33
Figura 17 – Ilustração de variação na iluminação da imagem	35
Figura 18 – Ilustração de variação na rotação da imagem.....	35
Figura 19 – Ilustração de variação na escala da imagem	35
Figura 20 – Imagem de exemplo para a classe positiva da base TESTAntrac	41
Figura 21 – Imagem de exemplo para a classe negativa da base TESTAntrac	41

LISTA DE QUADROS

Quadro 1 – Organização dos elementos na matriz de confusão.....	20
Quadro 2 – Funções de ativação	27
Quadro 3 – Condições experimentais para o Experimento 1	43
Quadro 4 – Procedimento experimental 1	44
Quadro 5 – Procedimento experimental eFF: ampliação de dados.....	45
Quadro 6 – Procedimento experimental eFV: ampliação de dados em Dt.....	45
Quadro 7 – Procedimento experimental eVV: ampliação de dados em DT ,Dv e Dt.....	45
Quadro 8 – Condições experimentais para o Experimento 2	46

LISTA DE TABELAS

Tabela 1 – Média dos índices de comparação entre imagens	48
Tabela 2 – Resultados para eFF no primeiro experimento	49
Tabela 3 – Resultados para eFV no primeiro experimento	50
Tabela 4 – Teste de Friedman para eFV no primeiro experimento	51
Tabela 5 – Resultados eVV no primeiro experimento	52
Tabela 6 – Teste de Friedman para eVV no primeiro experimento	53
Tabela 7 - Resultados para eFF no segundo experimento	54
Tabela 8 - Resultados para eFV no segundo experimento	55
Tabela 9 – Teste de Friedman para eFV no segundo experimento	56
Tabela 10 – Valores de sensibilidade, especificidade e sobreajuste.....	57

LISTA DE SIGLAS

CONAB	Companhia Nacional de Abastecimento
CNN	<i>Convolutional Neural Network</i>
DA	<i>Data Augmentation</i>
GPS	<i>Global Positioning System</i>
MSE	<i>Mean Squared Error</i>
SSIM	<i>Structural Similarity Index</i>
UEPG	Universidade Estadual de Ponta Grossa

SUMÁRIO

1	INTRODUÇÃO	12
2	OBJETIVOS	14
3	FUNDAMENTAÇÃO TEÓRICA	15
3.1	A CULTURA DO FEIJÃO	15
3.1.1	Características da planta.....	15
3.1.2	Doenças no feijoeiro.....	16
3.1.3	Deteção e avaliação da antracnose	17
3.2	CLASSIFICADORES.....	19
3.3	REDES NEURAS ARTIFICIAIS	21
3.4	REDES NEURAS CONVOLUCIONAIS.....	23
3.4.1	Camadas convolucionais.....	24
3.4.2	Camadas de subamostragem	26
3.4.3	Função de ativação	27
3.4.4	Camadas totalmente conectadas	27
3.4.5	Regularização por abandono	28
3.4.6	Treinamento da rede neural convolucional.....	29
3.4.7	Principais arquiteturas de redes neurais convolucionais.....	30
3.5	AMPLIAÇÃO DE DADOS.....	34
3.6	TRABALHOS CORRELATOS	36
3.6.1	Ferramentas computacionais na agricultura.....	36
3.6.2	Visão computacional na deteção de doenças em plantas.....	37
4	METODOLOGIA	39
4.1	BASES DE DADOS UTILIZADAS NOS EXPERIMENTOS	40
4.2	PROCEDIMENTO DE TREINAMENTO DAS REDES NEURAS.....	42
4.3	EXPERIMENTOS.....	42
4.4	ANÁLISE DOS RESULTADOS	46
5	RESULTADOS	48
5.1	AVALIAÇÃO DA SIMULAÇÃO POR AMPLIAÇÃO DE DADOS.....	48
5.2	DESEMPENHO NA AUSÊNCIA DA AMPLIAÇÃO DE DADOS	49
5.3	DESEMPENHO COM AMPLIAÇÃO DE DADOS NA FASE DE TESTE	50
5.4	DESEMPENHO COM AMPLIAÇÃO DE DADOS EM TREINO E TESTE	52
5.5	DESEMPENHO COM PARTIÇÃO DE TESTES DISTINTA	54
6	CONCLUSÃO	58
	REFERÊNCIAS	59

1 INTRODUÇÃO

O feijão comum (*Phaseolus vulgaris L.*) é um importante componente da dieta da maioria da população brasileira. Dados da Companhia Nacional de Abastecimento (CONAB, 2018) relacionam o país entre os maiores produtores e consumidores mundiais do grão, sendo o estado do Paraná o de maior destaque, alcançando quase um quarto da produção nacional. Além das preocupações decorrentes das condições climáticas (a exemplo, a possibilidade de geadas na segunda safra paranaense que podem causar grandes prejuízos agrícolas), e dos tratamentos culturais como adubações e correções de solo, existe ainda a necessidade de se prevenir e remediar as perdas advindas do ataque de pragas e doenças às lavouras. Miranda Filho (2006) relata a possibilidade de perdas decorrentes do aparecimento precoce da doença fúngica antracnose, que podem atingir a totalidade da lavoura, conforme condições ambientais favoráveis ao desenvolvimento desta moléstia.

A detecção e a avaliação visual da severidade da antracnose por técnicos e pesquisadores de doenças em plantas, quando do monitoramento de lavouras ou da coleta de dados em experimentos, são usualmente realizadas com o auxílio de escalas diagramáticas (GODOY et al., 1997). Esta técnica oferece um padrão visual para a identificação da ocorrência e da severidade de doenças em plantas. No entanto, seu emprego é dificultado pela possibilidade de diferentes interpretações quanto à presença e o estágio da moléstia na planta e ainda particularidades nas observações dos avaliadores. Isto pode levar a discrepâncias nos apontamentos de diferentes indivíduos no que tange à presença de lesões e à intensidade destas nas plantas.

Vários autores têm proposto o uso de técnicas de inteligência computacional nas tarefas de detecção e análise da severidade de doenças em plantas como uma forma de reduzir a influência da subjetividade nos apontamentos por parte de operadores e em busca de maiores índices de assertividade. Neste sentido, Amara et al. (2017), Fuentes et al. (2017), Ramcharan et al. (2017) e outros autores propõem o uso das técnicas de classificação com o uso de redes neurais artificiais para tratar o problema. Estes trabalhos exemplificam o interesse para o uso de técnicas de visão computacional nestas tarefas.

O desenvolvimento de tais sistemas classificadores faz uso de bases de dados com séries de imagens rotuladas para o treinamento da rede que será usada para discriminar as possíveis hipóteses de diagnóstico quanto a presença ou não de uma determinada doença. Logo, é necessário considerar a possibilidade de que a base de dados possua um conjunto de imagens com exemplos das lesões que a doença alvo pode causar em diferentes estágios da infecção e diferentes fases de desenvolvimento da planta, bem como em distintas condições de captura fotográfica, em termos de luminosidade, distância focal, ângulo e demais características.

Considerando o exposto, este trabalho avaliou o emprego de técnicas de ampliação de dados (*data augmentation*) no pré-processamento das bases de dados sobre o desempenho de redes neurais convolucionais no treinamento de classificadores para detecção da antracnose em folhas de feijoeiro. Mais detalhadamente, o objetivo foi verificar se procedimentos de ampliação de dados permitem incrementar o desempenho de classificadores em conjuntos de testes cujas condições de iluminação, rotação e escala sejam distintas daquelas registradas na base de treinamento. A consecução deste objetivo está associada à realização de experimentos sobre bases de dados supervisionadas contendo imagens de plantas com e sem antracnose. Duas bases foram usadas na composição dos conjuntos de dados processados nos experimentos: o repositório Digipathos, mantido pela Embrapa (BARBEDO; KOENIGKAN; SANTOS, 2016) e a base de dados UEPGANtrac, estruturada ao longo deste estudo. O uso de bases de dados de imagens e redes neurais convolucionais é demonstrado em trabalhos como Oppenheim e Shani (2017), Picon et al. (2019) e Argüeso et al. (2020).

Durante os experimentos foram empregadas três arquiteturas de redes neurais convolucionais: VGG16, InceptionV3 e ResNet50. Os procedimentos de ampliação de dados consideraram distintos conjuntos de valores para os parâmetros de iluminação, rotação e escala das imagens. A avaliação foi realizada com a mensuração do desempenho dos classificadores em termos de acurácia, sensibilidade, especificidade e observação do sobreajuste na detecção de lesões da antracnose nas imagens das folhas de feijoeiro. Os resultados obtidos foram analisados estatisticamente a fim de verificar se o emprego das técnicas de ampliação de dados estava relacionado a diferenças significativas no desempenho dos classificadores e a influência dos parâmetros nos resultados obtidos.

2 OBJETIVOS

Este trabalho tem como objetivo geral:

- Avaliar o efeito da ampliação de dados na simulação de alterações nas condições de aquisição de imagens do folíolo do feijoeiro no treinamento de redes neurais convolucionais para detecção de antracnose.

E como objetivos específicos:

- Estimar a eficácia de métodos de ampliação de dados na simulação de diferentes condições de iluminação, rotação e escala das imagens nas bases de dados usadas no treinamento dos classificadores;
- Comparar a acurácia, sensibilidade, especificidade e sobreajuste das arquiteturas de redes neurais convolucionais InceptionV3, ResNet50 e VGG16 em bases de treino, validação e teste obtidas e simuladas em diferentes configurações dos procedimentos de ampliação de dados;
- Comparar a acurácia, sensibilidade, especificidade e sobreajuste das redes neurais convolucionais InceptionV3, ResNet50 e VGG16 em bases de teste cujas imagens foram adquiridas em condições distintas daquelas da base de treinamento.

3 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os principais conceitos associados ao tema apresentado e está organizado de modo a compreender os seguintes tópicos: a cultura do feijão e a moléstia antracnose (1); classificadores em aprendizagem de máquina (2); redes neurais artificiais (3) e redes neurais convolucionais (4); ampliação de dados (4) e, por fim, uma revisão de trabalhos correlatos (5).

3.1 A CULTURA DO FEIJÃO

Um importante produto na segurança alimentar do povo brasileiro, o feijão aparece em mais de 90% dos almoços no país, juntamente com arroz e algum acompanhamento (BARBOSA, 2007). Barbosa e Gonzaga (2012) argumentam que a importância do grão “extrapola o aspecto econômico por sua relevância enquanto fator de segurança alimentar e nutricional”.

A relevância do feijão fica evidente quando se observam dados da Organização das Nações Unidas para Alimentação e Agricultura (FAOSTAT, 2019), que registram o Brasil como o terceiro maior produtor no mundo. A Companhia Nacional de Abastecimento estima a safra brasileira 2018/2019 em 3,104 milhões de toneladas em uma área de 2,977 milhões de hectares (CONAB, 2018). O feijão é produzido em todos os estados do país, com destaque para os estados de Paraná, Minas Gerais, Bahia, São Paulo e Goiás (BARBOSA; GONZAGA, 2012).

3.1.1 Características da planta

Conforme Menezes Júnior (1960) o termo ‘feijão’ abrange as sementes de várias leguminosas como feijão-guandu, feijão-fava, e feijão comum. Silva, Mendes e Kageyama (2009) definem o feijão como uma espécie com “hábito de crescimento determinado ou indeterminado, trepador ou não, formando dossel de 40 a 50 cm de altura”. Este estudo se concentra em observações acerca do feijão-comum (*Phaseolus vulgaris* L).

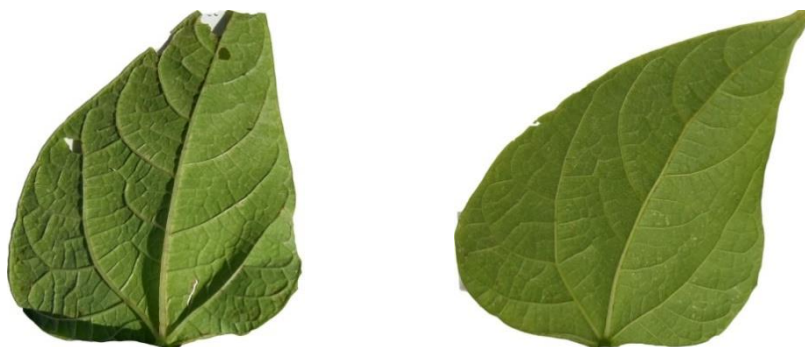
As folhas primárias do feijoeiro são consideradas simples e apenas estas são simples, sendo as demais compostas, formadas por três folíolos triangulares ou ovais ásperos ao tato (MENEZES JÚNIOR, 1960). A planta possui ainda caule áspero e curto, no qual estão inseridas as folhas e de onde saem ramificações.

3.1.2 Doenças no feijoeiro

Várias doenças podem atacar as plantas de feijão durante seu ciclo de vida, causando prejuízos que podem comprometer todo o rendimento econômico da cultura (MIRANDA FILHO, 2006). A antracnose, a ferrugem e a mancha-angular são as doenças da parte aérea (região da planta acima do solo) que causam mais perdas ao feijoeiro (BARBOSA; GONZAGA, 2012).

A antracnose, causada por *Colletotrichum lindemuthianum*, é uma doença fúngica caracterizada por lesões escuras e deprimidas nas nervuras da face abaxial das folhas, podendo ocorrer nos ramos, sementes e vagens, onde apresenta lesões concêntricas e deprimidas, com halo escuro (OLIVEIRA et al., 2018). As Figuras 1 e 2 mostram pares de folíolos de feijoeiro, sendo o primeiro par de folíolos saudáveis e o segundo com sinais de antracnose.

Figura 1 – Par de folíolos saudáveis



Fonte: O Autor

Figura 2 – Par de folíolos com sinais de antracnose

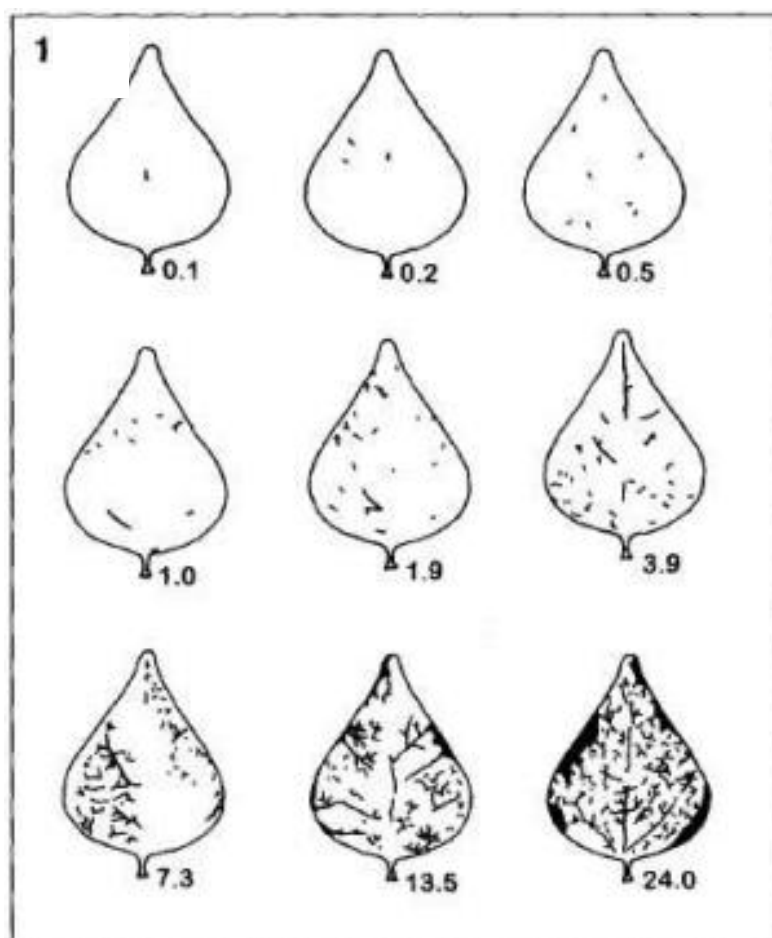


Fonte: O Autor

3.1.3 Detecção e avaliação da antracnose

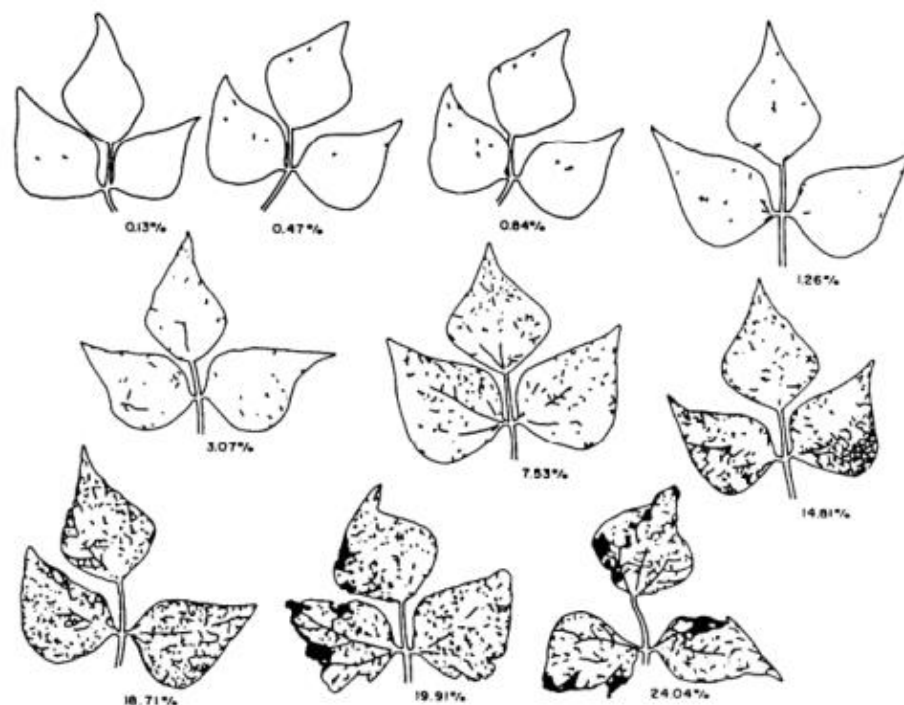
A abordagem usual para detecção da antracnose no feijoeiro é a realização de inspeções visuais rotineiras na parte aérea das plantas, desde a fase de plântula. Durante as inspeções, a detecção da ocorrência e a avaliação do estágio da doença são feitas com auxílio das escalas diagramáticas (CHRISTMANN, 2019). Diversas doenças possuem escalas para quantificação dos níveis de severidade apresentados (AMORIM et al., 1993; BARBOSA; MICHEREFF; MORA-AGUILERA, 2006; LIMA et al., 2013). Godoy et al. (1997) propuseram a escala diagramática para antracnose, alternaria e mancha angular no feijão utilizada no apontamento da severidade de lesões nesta cultura. As Figuras 3 e 4 apresentam imagens de escalas diagramáticas para avaliação da severidade da antracnose no feijão.

Figura 3 – Escala diagramática da severidade da antracnose em folíolos de feijão



Fonte: GODOY, Cláudia V. et al. 1997.

Figura 4 – Escala diagramática da severidade da antracnose em trifólios de feijão



Fonte: GODOY et al., 1997 *apud* HARMS, 2016.

Segundo Bergamin Filho e Amorin (1996), citados por Lazaroto et al. (2012), as escalas diagramáticas são representações ilustradas de plantas inteiras ou em partes com sintomas de doenças em diferentes níveis de severidade. O uso destas escalas é justificado pelo fato delas conferirem um padrão visual para aferição da presença e da severidade da doença (DAMASCENO; FURTADO; FERREIRA FILHO, 2014).

Contudo, conforme observado por Martinelli et al. (2015), apesar do padrão estabelecido, a avaliação de doenças em plantas por métodos visuais está sujeita à interpretações subjetivas dos observadores e às condições ambientais durante a coleta de dados. Tais limitações podem levar a mensurações divergentes de um mesmo quadro da doença.

A fim de contornar tais dificuldades, vários autores têm proposto o emprego de métodos computacionais na mensuração da ocorrência e apontamento da severidade de moléstias em plantas (BARBEDO, 2013). A justificativa subjacente é que o emprego de métodos computacionais e de sensores eletrônicos (câmeras digitais) minimiza o fator subjetivo na realização das leituras. Portanto, pode reduzir a ocorrência de avaliações conflituosas e minorar os erros de situações de apontamentos falso-positivos.

3.2 CLASSIFICADORES

O conceito de classificação em aprendizado de máquina faz parte do problema de reconhecimento de padrões e pode ser lido como a identificação de pertencimento de um objeto x a um conjunto A , a partir da observação de suas características (MORAIS, 2010). Russell e Norvig (2013) definem como sendo o conjunto A uma categoria ou classe de objetos. Segundo Halder; Sarkar e Bahar (2019) o processo de classificação de imagens refere-se à tarefa de extração de classes de informação de uma imagem.

Mais formalmente, Gonçalves (2015) define classificação de dados como: tendo-se um conjunto de entradas $X_i = \{X_1, \dots, X_a\}$ com N instâncias, sendo $i = \{1, \dots, N\}$ e a o número de atributos de X , e um conjunto de rótulos de classificação $R = \{r_1, \dots, r_b\}$, com $b \geq 2$, é montado um conjunto de dados de treinamento $T = \{(x_1, c_1), \dots, (x_i, c_i)\}$ no qual cada entrada x_i é vinculada a um rótulo $c_i \in R$. O autor afirma que o objetivo de um algoritmo de classificação é aprender, a partir de T , uma correlação entre os atributos de entrada tal que dada uma nova entrada não rotulada $X = \{x_1, \dots, x_a\}$, o classificador seja capaz de determinar o rótulo vinculado a ela (GONÇALVES, 2015).

Em complemento, Zimak (2006) define que um classificador binário é uma função f de $X \rightarrow Y$ em que X é um conjunto multivariado em que cada instância representa uma possível configuração dos atributos usados para descrever os objetos de interesse e Y é o conjunto $\{0, 1\}$, no qual os elementos são rótulos que indicam se um objeto pertence ou não ao conjunto A . Assim, dada uma instância x em X , um classificador binário f retorna 0 ou 1 dependendo das características apresentadas, indicando que um objeto pertence ou não a determinada classe.

A construção de um classificador comumente faz uso de métodos de aprendizagem computacional supervisionada. Nesta abordagem considera-se o aprendizado a partir de uma categoria rotulada fornecida em um conjunto de treinamento (DUDA; HART; STORK, 2012). Estes autores comentam ainda que os métodos selecionados para os procedimentos de classificação aprendem como as amostras se distribuem em diferentes classes pelo conjunto e, utilizando um subconjunto para treinamento do algoritmo, encontram regras para a divisão destas amostras.

A avaliação do desempenho dos classificadores, obtida após o processo de aprendizagem, é feita através das medidas de desempenho (RUSSELL; NORVIG, 2013). As seguintes medidas têm sido utilizadas na avaliação de desempenho de classificadores binários (LAVRAČ; FLACH; ZUPAN, 1999):

- a) acurácia: é a razão entre as amostras classificadas corretamente e o total de amostras do conjunto;
- b) especificidade: razão entre amostras corretamente classificadas como negativas sobre o total de amostras negativas;
- c) sensibilidade: total de amostras corretamente classificadas como positivas sobre o total de amostras positivas.

Estas medidas são obtidas a partir da matriz de confusão, um modelo de representação das taxas de desempenho para o problema de classificação em uma base de testes. No caso de classificadores binários, uma matriz de confusão é uma tabela 2×2 (PROVOST; KOHAVI, 1998), na qual os elementos são organizados conforme o Quadro 1:

Quadro 1 – Organização dos elementos na matriz de confusão

Classificação correta	Predição	
	+	-
+	TP	FN
-	FP	TN

Fonte: O Autor

Conforme Tharwat (2018) é possível definir os termos da matriz da confusão como segue: um positivo verdadeiro (TP) como sendo a amostra positiva corretamente classificada como positiva; um negativo falso (FN) como aquele cuja amostra positiva é classificada como negativa; um negativo verdadeiro (TN) sendo a amostra negativa corretamente classificada como negativa e, por fim; um falso positivo (FP), como sendo a amostra negativa que é erroneamente classificada como positiva.

As expressões mostram como as medidas de desempenho são calculadas a partir da matriz de confusão (THARWAT, 2018):

$$\text{a) acurácia} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\text{b) especificidade} = \frac{TN}{FP+TN}$$

$$\text{c) sensibilidade} = \frac{TP}{TP+FN}$$

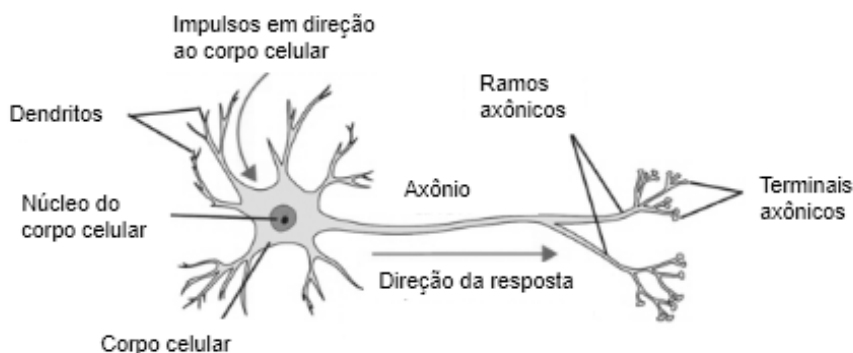
Dentre as técnicas computacionais utilizadas para as tarefas de classificação citam-se, conforme Aggarwal (2015): as árvores de decisão, métodos probabilísticos, máquinas de vetores de suporte (SVM), e as redes neurais artificiais.

3.3 REDES NEURAIS ARTIFICIAIS

Redes neurais artificiais são ferramentas computacionais que permitem a resolução de problemas em áreas diversas como a matemática, a estatística e a computação. Haykin (2001) define rede neural como “um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torna-lo disponível para uso”, inspirado no sistema nervoso dos organismos superiores pelo aspecto de adquirir conhecimento por um processo de aprendizagem e pela força de conexão entre os neurônios, conhecida como peso sináptico, que é utilizada para armazenar o conhecimento adquirido.

Conforme Haykin (2001), a unidade básica de processamento da rede neural é o neurônio artificial e esse é uma expressão matemática simplificada da maneira com que os neurônios biológicos manuseiam informação (SIQUEIRA, 2009). Neurônios biológicos são unidades de processamento do sistema nervoso. Segundo Tetteroo (2018), o neurônio biológico consiste de dendritos, que coletam sinais de outros neurônios, o corpo celular, que processa o sinal recebido e gera a resposta apropriada, e o axônio, uma extensão responsável pela propagação do sinal-resposta para outros neurônios e demais estruturas celulares, sendo esse o processo denominado sinapse. A Figura 5 traz uma representação esquemática do neurônio biológico.

Figura 5 – Representação esquemática do neurônio biológico

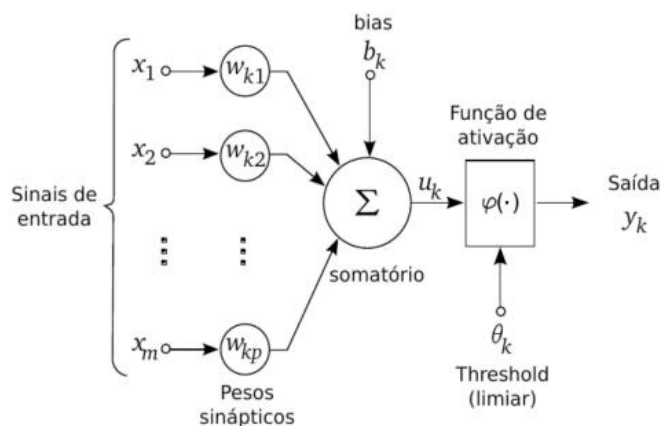


Fonte: Adaptado de Tetteroo (2018)

Na Figura 5 pode-se observar uma representação do neurônio biológico mostrando os dendritos, responsáveis por receber sinais; o sentido dos impulsos em direção ao corpo celular; o corpo celular propriamente dito e o núcleo deste; o axônio, e o direcionamento da resposta processada pelo corpo celular; os ramos axônicos e seus terminais, de onde são enviadas as respostas processadas.

As sinapses são estruturas funcionais que intermediam as interações entre os neurônios, que por sua vez são unidades de processamento de informação fundamentais para uma rede neural (HAYKIN, 2001). Em cada neurônio artificial, ilustrado na Figura 6, existem quatro elementos a serem considerados: um conjunto de sinapses ou conexões, com pesos próprios, em que um sinal x_j é multiplicado pelo peso w_{kj} para a sinapse j e o neurônio k ; uma unidade somadora; um valor de *bias* externo b_k que altera o valor de entrada da função de ativação; e a função de ativação, que restringe a amplitude de saída e introduz não linearidade à resposta.

Figura 6 – Modelo de neurônio base para redes neurais



Fonte: Adaptado de Haykin (2001)

Haykin (2001) define o neurônio em termos matemáticos como a equação (1):

$$y_k = f(u_k) = \varphi \left(\sum_{j=1}^m w_{kj} x_j + b_k \right) \quad (1)$$

Zurada (1992) relata que os neurônios artificiais descritos operam normalmente em camadas e existem diferentes modelos dessas estruturas.

3.4 REDES NEURAIAS CONVOLUCIONAIS

As Redes Neurais Convolucionais (*Convolutional Neural Network* - CNN) são modelos de redes neurais especializadas no processamento de dados que possuem uma topologia de grade conhecida (GOODFELLOW; BENGIO; COURVILLE, 2016), perfazendo desde a extração de características do objeto alvo até sua classificação. Uma CNN consiste em camadas de entrada e saída, além de múltiplas camadas ocultas entre estas, incluindo convolução, *pooling* e normalização (ARIVAZHAGAN; LIGI, 2018). Vargas; Paes e Vasconcelos (2016) relatam que uma CNN é capaz de aplicar filtros em dados visuais mantendo a relação de vizinhança entre os *pixels* da imagem ao longo das etapas de processamento.

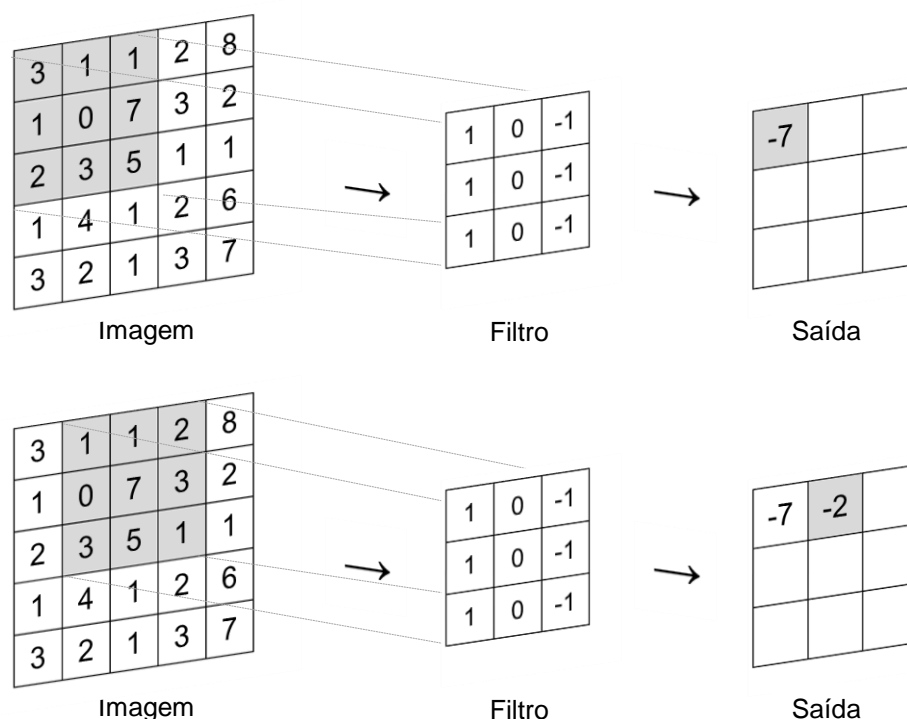
Na proposta da CNN, LeCun et al. (1989), basearam-se no fato de que “extratores de características ideais para resolver problemas que lidam com a variabilidade e a riqueza de dados naturais podem ser incorporados na rede neural automaticamente durante o processo de aprendizagem” (FERNANDES, 2013). Russell e Norvig (2013) definem a tarefa de aprendizagem como um conjunto de treinamento que compreende pares de exemplos de entrada e saída, em que cada saída é gerada por uma função desconhecida. Dessa maneira, tende-se a descobrir uma função que se aproxime da função exata que produziu aquela saída.

Vargas; Paes e Vasconcelos (2016) observam ainda que uma CNN é inspirada nos processos biológicos de processamento de dados visuais e que esse modelo de rede neural vem sendo amplamente utilizado nas tarefas de classificação, detecção e reconhecimento em imagens e vídeos. A seguir são descritas as principais camadas que compõe uma rede neural convolucional.

3.4.1 Camadas convolucionais

A camada convolucional usa filtros de aprendizado que consistem em estruturas de tamanho determinado que ‘deslizam’ (ou convolucionam) por sobre a entrada – no caso, a imagem. Ao deslizar por sobre a imagem de entrada, os filtros geram mapas de características, conforme suas especificidades. LeCun et al. (1995) citam que uma camada de convolução é normalmente composta por diversos filtros que extraem mapas de características e que múltiplas características podem ser extraídas de cada localidade da imagem. Os filtros consistem em matrizes de dimensão conhecida, como ‘5 x 5 x 3’, sendo o último número referente aos canais de cores (*Red*, *Green* e *Blue* – RGB). Na Figura 7 pode-se observar um exemplo do processo de convolução de uma imagem de dimensão 5 x 5 por um filtro 3 x 3.

Figura 7 – Exemplo de convolução 2D

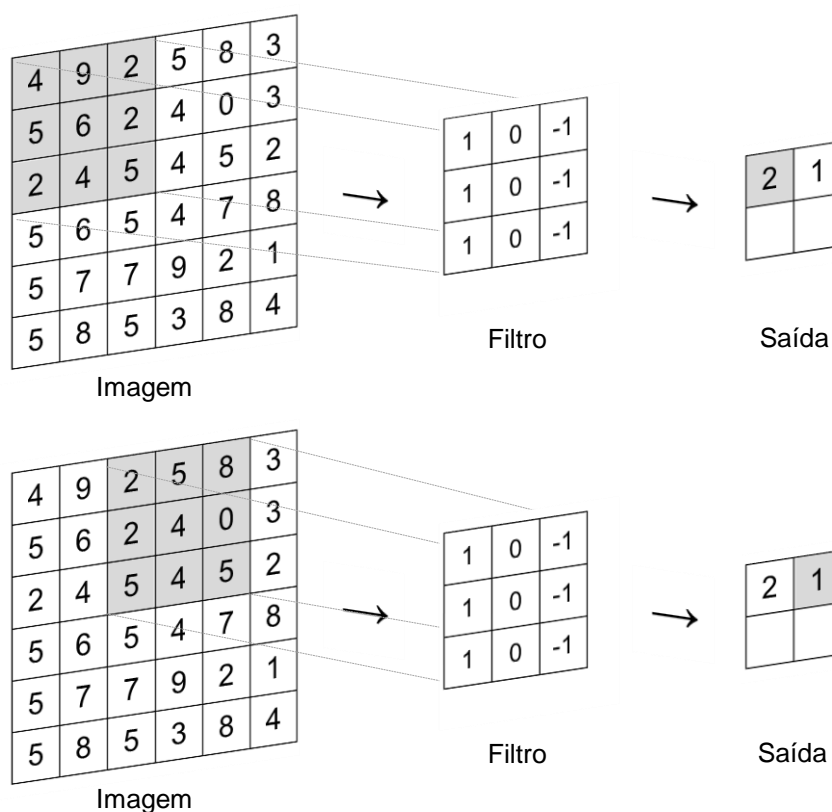


Fonte: O Autor

No exemplo da Figura 7, a convolução é demonstrada quando da passagem do filtro por uma área de pixels da imagem denominada campo receptivo local, desde a parte superior esquerda até a parte inferior direita. Os valores dos pixels são multiplicados pelos valores do filtro e então somados, obtendo-se um resultado de saída correspondente ao mapa de características da imagem de entrada.

Além do tamanho do filtro, outros parâmetros são considerados no processo de convolução. O tamanho do passo do deslocamento ao se deslizar por sobre a imagem é denominado de *stride* e, na Figura 7, esse parâmetro tem o valor de um pixel. A Figura 8 apresenta um exemplo de filtro 3 x 3 com *stride* de dois pixels sobre uma matriz de dimensão 6 x 6 como imagem de entrada.

Figura 8 - Filtro 3 x 3 com *stride* de dois pixels



Fonte: O Autor

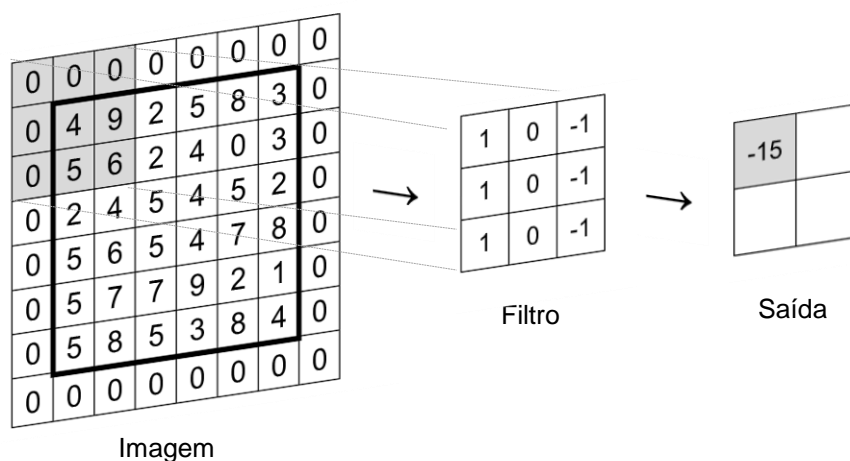
Existem ainda os parâmetros relacionados ao preenchimento (*padding*) da borda da matriz que representa a imagem de entrada. O preenchimento é feito quando se deseja obter um mapa de características com um tamanho específico, uma vez que a dimensão deste pode ser calculada conforme:

$$Tm = \frac{M + 2p - k}{s} + 1$$

sendo Tm o tamanho do mapa de características, M o tamanho da imagem de entrada, p o tamanho do preenchimento da borda (*padding*), k o tamanho do filtro, e s o deslocamento (*stride*) (EBERMAN; KROHLING, 2018).

A Figura 9 mostra um exemplo de preenchimento denominado *zero-padding* em que a borda da imagem é preenchida com zeros.

Figura 9 - Exemplo de preenchimento *zero-padding*

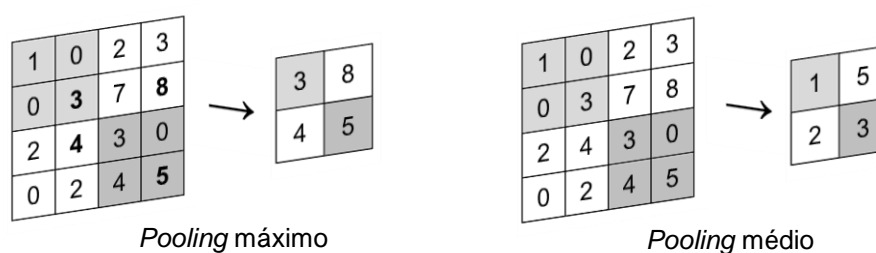


3.4.2 Camadas de subamostragem

As camadas de subamostragem, conhecidas como camadas de *pooling*, são usualmente imediatas às camadas de convolução e servem ao objetivo de reduzir a dimensionalidade dos resultados da matriz obtida na convolução, permitindo a diminuição do tempo de processamento das entradas e criando invariância espacial. Esta camada substitui a saída da rede com um resumo estatístico das saídas das camadas anteriores (GOODFELLOW; BENGIO; COURVILLE, 2016).

A camada de *pooling* agrupa um conjunto de valores proveniente da camada de convolução e os representa com um valor específico. Dentre os valores de representação tem-se, entre outros, o valor máximo (*max pooling*) e médio (*average pooling*) da região analisada. Na Figura 10 observa-se uma matriz 4 x 4 como saída de uma convolução e os resultados de *pooling* médio e máximo em dimensão 2 x 2.

Figura 10 – Exemplo de subamostragem (*pooling*) médio e máximo



3.4.3 Função de ativação

As funções de ativação em redes neurais determinam a ativação de um respectivo neurônio da rede. A ativação do neurônio acontece de acordo com o limiar (*threshold*) proposto na função de ativação. Esse processo ocorre em sequência às camadas de convolução e subamostragem.

Em geral a função de ativação Unidade Retificadora Linear (*Rectified Linear Unit* - ReLU) é a mais utilizada nos modelos de redes neurais convolucionais, devido ao fato de gerar menor erro de treinamento e depender menor custo computacional para sua execução (JARRETT et al., 2009; GLOTZ et al., 2010; BORDES; BENGIO, 2011; NAIR; HINTON, 2010; KRIZHEVSKY et al., 2012).

Além da ReLU, outras funções comumente utilizadas para os propósitos de ativação em modelos de redes neurais são a Sigmoidal ou Logística, Tangente Hiperbólica e Softmax, sendo essa última comumente utilizada após a camada totalmente conectada para fins de normalização em redes neurais convolucionais. As funções de ativação citadas são apresentadas no Quadro 2.

Quadro 2 – Funções de ativação

Nome da função	Equação
Unidade Retificadora Linear (ReLU)	$f(x) = \max(0, x)$
Sigmoidal ou Logística	$f(x) = \frac{1}{1 + e^{-x}}$
Tangente Hiperbólica	$f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$
Softmax	$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$

Fonte: O Autor

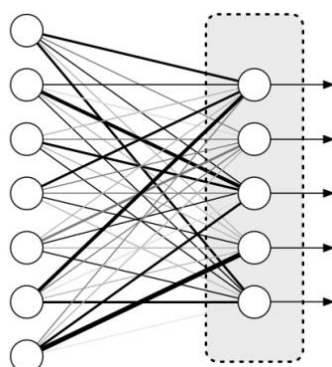
3.4.4 Camadas totalmente conectadas

As camadas totalmente conectadas são usualmente instaladas como camadas finais das redes neurais convolucionais e servem ao propósito de classificação. Tais redes apresentam uma ou mais camadas deste tipo em suas arquiteturas.

Ramsundar e Zadeh (2018) afirmam que nas camadas totalmente conectadas qualquer valor de entrada pode afetar qualquer valor de saída. Essa camada tem por objetivo usar as características obtidas nas camadas de convolução e *pooling* anteriores para a classificação dos dados com base no conjunto de treinamento.

A saída da última camada de subamostragem (*pooling*) é a entrada da camada totalmente conectada. Na saída da camada totalmente conectada é utilizada a função de ativação Softmax, conforme apresentado no capítulo 3.4.3 e a equação no Quadro 2. A saída gerada é um fator de probabilidade para cada classe de predição. A Figura 11 apresenta um modelo da camada totalmente conectada.

Figura 11 – Camada totalmente conectada



Fonte: Ramsundar e Zadeh (2018)

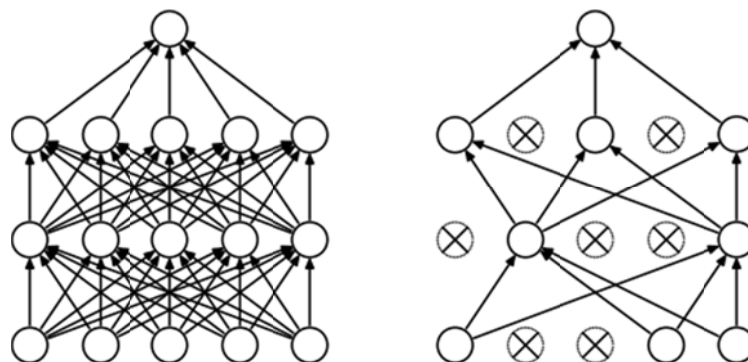
3.4.5 Regularização por abandono

A regularização por abandono, conhecida pelo termo em inglês “*dropout*”, foi proposta por Hinton et al. (2012) e consiste em descartar parte dos detectores de atributos em cada caso de treinamento. Dessa maneira, evita-se o chamado *overfitting*, quando há um sobreajuste do modelo aos dados fornecidos no processo de treinamento, diminuindo a possibilidade de generalização. Hawkins (2004) define o *overfitting* como o uso de modelos ou procedimentos que incluem mais termos ou abordagens que o necessário para a resolução do problema.

O *dropout* aleatoriamente abandona - ou descarta - unidades (neurônios e suas conexões) durante o processo de treinamento (SRIVASTAVA et al., 2014) reduzindo a complexidade da rede e impedindo a influência e contribuição dos neurônios então desativados sobre outros neurônios.

Na Figura 12 é possível visualizar um modelo de *dropout* de uma rede com duas camadas ocultas na esquerda e após o processo de regularização à direita.

Figura 12 – Regularização *dropout* em rede neural com duas camadas ocultas



Fonte: Srivastava et al. (2014)

3.4.6 Treinamento da rede neural convolucional

O treinamento de redes neurais convolucionais, dada a característica de complexidade desse modelo, requer um grande número de exemplos de entrada (SRIVASTAVA et al., 2014; TAJBAKSH, et al., 2016; KRIZHEVSKY et al., 2012). Alguns trabalhos propõem técnicas para lidar com a necessidade de grandes bases de dados para treinamento, como o aprendizado das estruturas e da força dos filtros (KESHARI et al., 2018), ajustes finos em redes pré-treinadas (TAJBAKSH, et al., 2016) bem como o uso de técnicas de regularização (COURBARIAUX; BENGIO; DAVID, 2015; HINTON, et al., 2012).

O algoritmo *backpropagation* (RUMELHART, 1988) é há muito o mais utilizado no treinamento de redes neurais (LECUN, et al., 1989; WERBOS, et al., 1990). Esse algoritmo consiste em duas fases, a propagação (*forward*) e propagação reversa (*backward*), ou retropropagação. Na primeira fase a entrada passa por todas as camadas da rede até a produção de sua respectiva saída, que é comparada ao resultado desejado e a diferença apresentada é o valor de erro da rede. Na fase *backward* os pesos são ajustados conforme uma regra de correção, levando em consideração a taxa de erro obtida na fase anterior de modo a permitir que a rede neural aprenda a mapear corretamente uma determinada entrada para a respectiva saída (HAYKIN, 2001; RUSSELL; NORVIG, 2013).

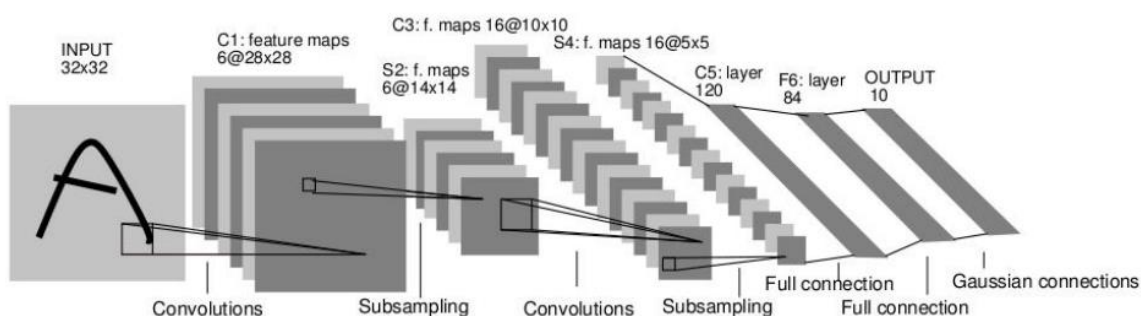
3.4.7 Principais arquiteturas de redes neurais convolucionais

As arquiteturas das redes neurais convolucionais variam conforme as necessidades apresentadas em suas utilizações e atualmente os modelos tornaram-se consideravelmente profundos – isto é, com várias camadas. As diferentes arquiteturas dos dias atuais existem e estão acessíveis devido a fatores como a maior disponibilidade de estrutura computacional de alto desempenho, as competições de reconhecimento e detecção de imagens, e o desenvolvimento de pesquisas acadêmicas (SZEGEDY et al., 2015).

LeCun et al. (1998) propuseram a arquitetura de rede neural convolucional LeNet-5, composta por duas camadas convolucionais e de subamostragem e três camadas inteiramente conectadas. Conforme os autores, a utilização inicial dessa arquitetura de rede era o reconhecimento de caracteres escritos à mão. Szegedy et al. (2015) comentam que a partir desse modelo formou-se o padrão estrutural para redes neurais convolucionais, contando com camadas convolucionais empilhadas, opcionalmente seguidas por camadas de subamostragem e então camadas totalmente conectadas.

A Figura 13 apresenta um modelo da LeNet-5, em que é possível observar a entrada (*input*) da rede, seguida pelas camadas convolucionais (*convolutions*) e de subamostragem (*subsampling*), de onde se extraem os mapas de características e, ao fim, as camadas inteiramente conectadas (*full connection*) onde acontece a classificação, e por fim a saída da rede (*output*).

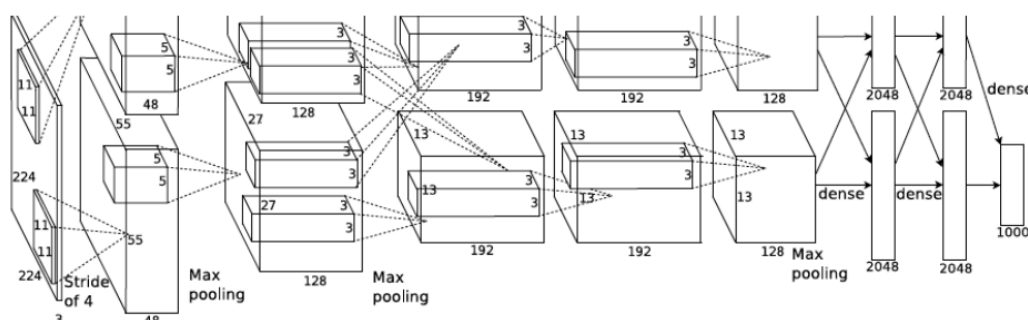
Figura 13 – Arquitetura da CNN LeNet-5, proposta por LeCun et al. em 1998



Fonte: LeCun et al. (1998)

Krizhevsky et al. (2012) propuseram uma rede chamada AlexNet para classificar 1,2 milhão de imagens em alta resolução da base ImageNet, que consiste em 15 milhões de imagens categorizadas pertencentes a 22 mil categorias. A AlexNet é organizada em cinco camadas convolucionais, algumas seguidas por camadas de subamostragem, e três camadas completamente conectadas, além de usar ReLU como função de ativação. A descrição da rede AlexNet é ilustrada na Figura 12. O modelo alcançou resultados superiores aos do estado da arte para a base citada até então e propôs a divisão das camadas para se aperfeiçoar o modo de trabalhar em unidades de processamento gráfico (KRIZHEVSKY et al., 2012).

Figura 14 – Arquitetura da CNN AlexNet, proposta por Krizhevsky et al. em 2012

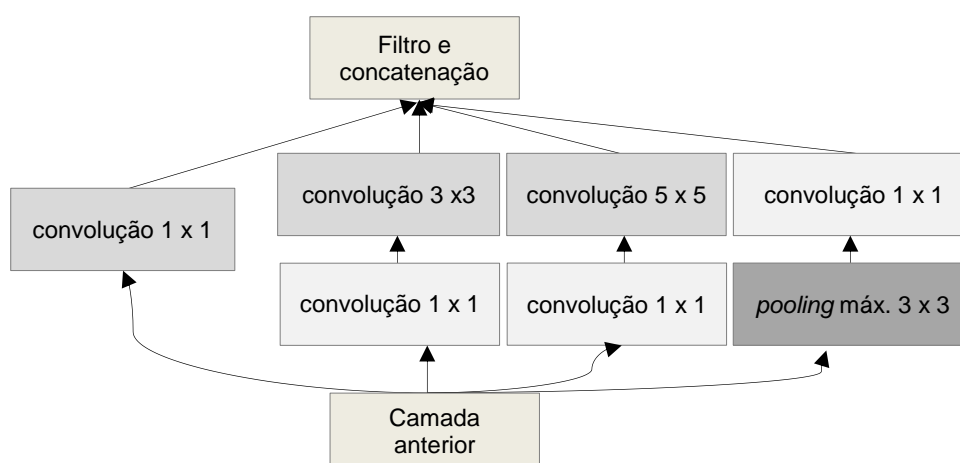


Fonte: Krizhevsky et al. (2012)

Conforme Szegedy et al. (2015), a maneira mais direta de aperfeiçoar o desempenho de redes neurais é aumentar o seu tamanho, em se tratando de número de camadas e de neurônios nas camadas. Neste sentido, Simonyan e Zisserman (2014) investigaram o efeito da profundidade de redes neurais convolucionais e propuseram o modelo VGG16, com treze camadas convolucionais e três camadas totalmente conectadas, utilizando como função de ativação a função ReLU. Eles também propuseram o modelo VGG19, uma variação do anterior, com dezesseis camadas convolucionais. A sigla que dá nome a essa arquitetura vem de 'Visual Geometry Group' (Grupo de Geometria Visual - em tradução livre), ligado ao Departamento de Ciência da Engenharia, da Universidade de Oxford. Com essas arquiteturas os autores foram vencedores da competição ImageNet de 2014, obtendo resultados superiores aos do estado da arte, demonstrando que a profundidade das redes é benéfica aos propósitos de classificação de imagens.

A partir dos trabalhos de Lin; Chen e Yan (2013) que previam a construção de 'redes de redes', em que eram empilhadas pequenas redes neurais com complexas estruturas para abstrair os dados em uma rede maior, Szegedy et al. (2015) desenvolveram uma rede com um novo modelo organizacional na forma do que chamaram de 'módulo *inception*', em que eram utilizados blocos (ou módulos) contendo camadas convolucionais e esses blocos individuais eram utilizados na construção de uma rede completa. Dessa maneira, enxerga-se esta estrutura como redes dentro de uma rede. A Figura 15 exemplifica o módulo *inception*.

Figura 15 - Exemplo de um módulo *Inception*



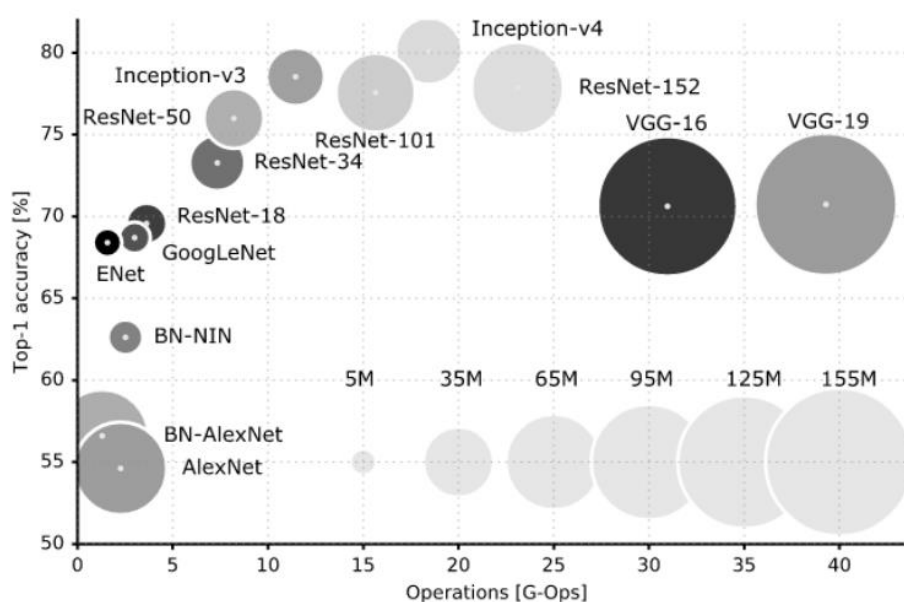
Fonte: Adaptado de Szegedy et al. (2015)

Mais detalhadamente, a rede Inception, desenvolvida por Szegedy et al. (2015), consiste de uma estrutura de módulos empilhados totalizando quarenta e duas camadas e, apesar dessa dimensão, argumenta-se que esta rede é mais eficiente, em se tratando de custo computacional, que o modelo VGG. Os autores salientam que a principal conquista desta arquitetura é o aperfeiçoamento da utilização dos recursos computacionais dentro da rede (SZEGEDY et al., 2016).

Posteriormente, pesquisadores da Microsoft (HE et al., 2016) e do Facebook (XIE et al., 2017) introduziram o conceito de redes residuais, que consistia em uma estrutura para facilitar o treinamento de redes de maior profundidade. Nesses modelos foi aplicada a técnica de '*skip connections*', de modo que a entrada de uma camada seja a saída de uma camada prévia não imediata, permitindo o aprendizado residual e diminuindo a degradação da acurácia alcançada por modelos de redes neurais convolucionais com camadas profundas.

Canziani; Paske e Culurciello (2016) analisaram diferentes arquiteturas de redes neurais e consideraram que, apesar do constante incremento nos valores de acurácia obtidos por arquiteturas de redes participantes do desafio computacional ImageNet, a utilização de recursos não era apropriadamente levada em consideração. Os autores então avaliaram métricas como: acurácia, utilização de memória, número de parâmetros, contagem de operações e consumo de energia. A Figura 16 traz uma comparação entre as taxas de acurácia (*accuracy*), número de operações (*operations*) e dimensão da rede. O tamanho da bolha referente a cada arquitetura é relativo ao número de parâmetros da rede.

Figura 16 – Comparação entre arquiteturas de redes neurais



Fonte: Canziani; Paske e Culurciello (2016)

Na Figura 16 se observa a comparação entre operações (eixo x) e acurácia (eixo y) e as diferentes arquiteturas de redes neurais convolucionais. É possível notar valores de acurácia superiores a 70% para as arquiteturas de redes Inception, ResNet e VGG. Nota-se ainda um maior número de operações nas arquiteturas VGG, bem como maior número de parâmetros. A maior acurácia apontada na figura é a da arquitetura Inception, no modelo InceptionV4.

As redes neurais convolucionais devem ter a capacidade de generalizar a partir dos exemplos de treino a fim de obter maiores taxas de acerto quando da execução de testes. Objetivando incrementar a quantidade e a variedade dos itens da partição de treino, considera-se o uso de técnicas como a ampliação de dados.

3.5 AMPLIAÇÃO DE DADOS

A generalização de um modelo de classificação refere-se ao desempenho deste quando avaliado em dados conhecidos (base de treinamento) em comparação a um conjunto de dados não anteriormente visto (base de testes). Modelos com pouca capacidade de generalização tendem a produzir sobreajuste: quando o desempenho do classificador resulta em maiores níveis de acurácia na base de treinamento em comparação com a amostra em geral (ELKAN, 2012). A ampliação de dados é uma técnica que previne a ocorrência do sobreajuste (conhecido pelo termo *overfitting*) ao aumentar a base de dados e permitir maior generalização dos modelos (TONG et al., 2018; SIMARD et al., 2003; LAKHANI, 2017).

Conforme Van Dyk e Meng (2001), o termo 'ampliação de dados' refere-se a métodos para otimização iterativa através de dados não observados. Perez e Wang (2017) comentam que uma abordagem genérica e efetiva para a condução deste método é o uso de transformações geométricas e de colorização nas imagens como rotações, espelhamentos, cortes e ampliações e ainda mudanças nas cores. Os dados ampliados representarão um conjunto mais abrangente de possibilidades, minimizando a distância entre os conjuntos de treino e de testes (SHORTEN; KHOSHGOFTAAR, 2019).

McLaughlin; Del Rincon e Miller (2015) utilizaram métodos de ampliação de dados para simular a variabilidade presente em dados reais, mas pouco representados na partição de treinamento disponível. Os autores aplicaram variações de brilho, de pose, no ângulo de rotação e outras mais no intuito de ampliar a generalização de redes neurais convolucionais, concluindo que transformações geométricas, como as citadas, possibilitam incremento no desempenho dos modelos. Perez et al. (2018) investigaram o impacto de métodos de ampliação de dados para a classificação de melanoma com o uso de redes neurais convolucionais, modificando valores de saturação, contraste e brilho das imagens em busca de simular mudanças nas cores devido a condições de ajuste da câmera e de características das lesões em estudo. Outros parâmetros também foram modificados no experimento, como rotações, escala e cortes. Estes autores apresentaram resultados confirmando a importância da ampliação de dados tanto na partição de treino quanto na de testes.

Dessa forma, considerando os objetivos deste estudo, descrevem-se os procedimentos de ampliação de dados por transformações geométricas que incluem mudanças nos parâmetros de iluminação, rotação e escala das imagens: mudanças na iluminação constituem variações no mapa de cor da imagem, aumentando ou diminuindo valores de brilho; em termos de rotação, trata-se de mudanças no ângulo de giro sob o eixo da imagem, em variações entre 1° e 359° (SHORTEN; KHOSHGOFTAAR, 2019). O parâmetro escala refere-se a alterações nos níveis de aproximação (*zoom in*) e distanciamento (*zoom out*) do objeto em foco. As Figuras 17, 18 e 19 ilustram as transformações comparadas com a imagem original.

Figura 17 – Ilustração de variação na iluminação da imagem



Fonte: O Autor

Figura 18 – Ilustração de variação na rotação da imagem



Fonte: O Autor

Figura 19 – Ilustração de variação na escala da imagem



Fonte: O Autor

3.6 TRABALHOS CORRELATOS

3.6.1 Ferramentas computacionais na agricultura

Em artigo publicado na revista *Science* em 1985, Holt (1985) descreve como os computadores começaram a ser importantes personagens na produção agrícola, de uma maneira geral coletando dados e os processando, e agindo como ferramentas de apoio à decisão. O autor narra um cenário de possíveis benefícios da combinação de monitoramento, comunicação, análise, modelos de simulação e controles automáticos que permitiriam a agricultura ser ainda mais eficiente e produtiva que era à época, indicando a necessidade da popularização das ferramentas computacionais para o uso pessoal.

Putler e Zilberman (1988) destacam sobre os microcomputadores de baixo custo, que permitiriam o uso de tecnologias computacionais para o desenvolvimento de tarefas rotineiras da agricultura como manutenção de registros, planejamento financeiro, controle de pragas e outras. Nesse trabalho se observa uma maior tendência de uso de computadores pessoais por grandes propriedades rurais e profissionais agrícolas de maior nível técnico e educacional.

O aprimoramento do uso de ferramentas computacionais no meio agrícola se deu em meados dos anos 2000 com a efetivação da agricultura de precisão, que está diretamente ligada à coleta de informações e manipulação de dados. Stafford (2000) relata sobre a implementação da agricultura de precisão, o conjunto de informações adquiridas e os dispositivos que coletam dados e o sistema de posicionamento global (GPS), que permite atrelar a informação capturada a um ponto geográfico específico.

Arroqui et al. (2012) abordam o rápido desenvolvimento das tecnologias de informação e comunicação aplicadas à agricultura, citando o sensoriamento remoto georreferenciado, sensores *online* e bases de dados públicas, como de clima e de mercados. Os autores comparam tecnologias de transmissão de dados em um simulador agrícola baseado em serviços *web*, com foco na plataforma Android devido a sua crescente popularidade entre os usuários (BUTLER, 2011, p. 4-7 apud ARROQUI et al., 2012, p. 15). Para Cranston (2010) citado por Arroqui et al. (2012, p. 15), as aplicações móveis têm sido apontadas como a nova alternativa para melhorias econômicas e sociais nas comunidades rurais.

Delgado; Kowalski e Tebbe (2013) relatam o aprimoramento de ferramenta computacional, antes exclusivamente destinada a computadores domésticos, para uso em dispositivos móveis, mencionando que *smartphones* e *tablets* permitem o uso de *softwares* portáteis, que facilitam a comunicação no campo entre agricultores, técnicos de nutrição e outros; já Montoya et al. (2013) lembram que *smartphones* são uma comum e importante parte da vida cotidiana, sendo portáteis, onipresentes, pequenos e leves, sendo ainda extremamente relevantes pela perspectiva social.

3.6.2 Visão computacional na detecção de doenças em plantas

Barbedo (2016) apresenta uma revisão dos principais desafios na identificação de doenças em plantas em imagens no espectro visível. O autor relata que, apesar da atenção destinada a esses trabalhos nas últimas décadas, as técnicas ainda são limitadas e dependem de condições ideais de captura em termos de ângulo da fotografia, iluminação e sombreamento, e equipamento utilizado. O autor também destaca que os seguintes fatores afetam negativamente o desempenho de sistemas de detecção de doenças em imagens de plantas:

- A existência de objetos de fundo nas imagens registradas: alguns objetos podem ter características visuais que se confundem com a classe alvo;
- Variações na forma de apresentação dos sintomas;
- A definição dos limites destes do sintoma na imagem;
- A composição da base de dados supervisionada a fim de se obter conjuntos de imagens, destinadas a aprendizado e teste dos classificadores, que sejam representativas para a tarefa alvo.

Melo (2015) relata o desenvolvimento de aplicativo para dispositivos móveis para o reconhecimento automático de índices de severidade da ferrugem asiática da soja utilizando processamento digital de imagens e redes neurais artificiais, no intuito de auxiliar no diagnóstico precoce e na tomada de decisão quanto ao manejo da doença. O trabalho alcançou taxas de acerto superiores a 70% para um dos cenários experimentais, sendo considerado, conforme a autora, “viável e eficaz”.

Similarmente, Wang et al. (2012) utilizaram modelos de redes neurais e técnicas de processamento de imagens para a detecção de doenças em trigo (ferrugem estriada e ferrugem da folha) e uva (míldio e oídio), alcançando acurácias superiores a 94%. Estes autores efetuaram procedimentos de extração de características, como cores e formas, e posteriormente utilizaram o método de análise de componentes principais (PCA) para a seleção dos atributos a serem considerados na classificação.

Em se tratando do uso de modelos de redes neurais convolucionais, Fuentes et al. (2017) demonstraram sua aplicação na detecção de pragas e doenças em uma base de dados de tomate, alcançando resultados efetivos para a detecção de nove tipos de pragas e doenças com as arquiteturas de rede VGG e ResNet. As imagens neste trabalho foram coletadas em distintas condições de iluminação, temperatura, umidade e localização formando uma base heterogênea, considerando ainda diferentes estádios de desenvolvimento das plantas alvo, partes infectadas e nível de doença.

Arivazhagan e Ligi (2018) apresentaram um modelo para detecção de doenças nas folhas de manga, incluindo antracnose, utilizando arquitetura de três camadas ocultas inspirada nos modelos VGG e Alexnet. Em uma base de cinco mil imagens, entre plantas saudáveis e doentes, e efetuando procedimentos de ampliação de dados como rotações e recortes, foi possível alcançar resultados de acurácia superiores a 95% na etapa de testes.

Ainda sobre redes neurais convolucionais, Ramcharan et al. (2017) treinaram uma CNN para a identificação de doenças e ataque de pragas na mandioca, alcançando uma acurácia geral de 93% na execução dos testes. Ferentinos (2018) apresentou a detecção de doenças e diagnóstico usando imagens de folhas de plantas saudáveis e doentes com uso de CNN, treinando os modelos em uma base de dados com 25 plantas e mais de 87 mil registros. Finalmente, nesse sentido, considera-se o trabalho de Sladojevic et al. (2016) que também usaram CNN para detecção de doenças em treze diferentes tipos de plantas e Aji et al. (2013) que implementaram uma rede neural para a detecção de doenças em plantas de óleo de palma para a detecção de antracnose, entre outras doenças.

4 METODOLOGIA

Como observado por Tong et al. (2018), o desenvolvimento de classificadores para operar em ambientes em que a aquisição das imagens não é controlada em termos de iluminação, distância do alvo e pose é uma tarefa complexa. Isto porque, nesta situação, o conjunto de dados usados no treinamento do classificador pode não permitir a captura do padrão de comportamento observado nas imagens a serem processadas no futuro. Neste contexto, Krizhevsky et al. (2012) observam que procedimentos aplicados nas técnicas de ampliação de dados em imagens proveem uma estratégia para simular o efeito de diferentes condições de aquisição destas e, assim, operam para reduzir a possibilidade de sobreajuste dos modelos de classificação, um dos problemas que podem ser derivados do emprego de bases de dados pouco representativas no treinamento de classificadores.

Este capítulo descreve dois conjuntos de experimentos que mensuram o desempenho de diferentes arquiteturas de redes neurais convolucionais na detecção da doença antracnose em estruturas foliares de feijoeiros. Durante estes experimentos, procedimentos de ampliação de dados foram utilizados para simular diferenças entre os conjuntos de treinamento e os conjuntos de testes. O primeiro experimento é composto por um grupo de testes tal que, cada teste está associado a uma configuração relativa à aplicação da ampliação de dados nas bases de treinamento e/ou teste durante o processo de aprendizagem das redes neurais convolucionais. O segundo experimento estende o primeiro, com o teste de redes neurais convolucionais treinadas nas bases de dados do primeiro experimento em uma base de dados adquirida em condições ambientais e temporais distintas.

O desempenho dos processos de classificação obtido a partir das redes neurais convolucionais durante os procedimentos de treinamento e teste dos modelos são comparados e mensurados em termos de acurácia, sensibilidade e especificidade, assim como através da observação da ocorrência de sobreajuste. Para este último são consideradas as diferenças observadas entre os valores de acurácia obtidos na partição de treinamento em comparação com o da partição de testes, conforme Elkan (2012).

4.1 BASES DE DADOS UTILIZADAS NOS EXPERIMENTOS

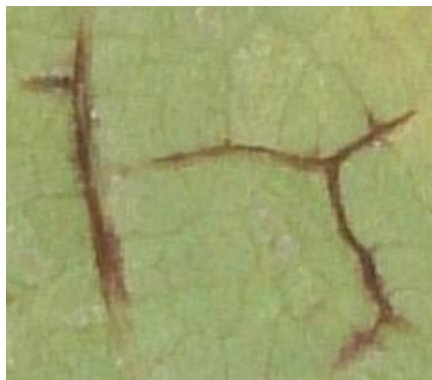
Os experimentos descritos neste capítulo assumem o uso de dois conjuntos de dados principais. A base de dados TESTAntrac e a base de dados UEPGANtrac. A primeira base é formada por duas mil imagens, sendo mil delas associadas a um diagnóstico positivo para a antracnose no feijoeiro e outras mil relacionadas a um diagnóstico negativo. Nesta base, parte das imagens de estruturas das folhas de feijoeiro com sinais de antracnose são provenientes do repositório Digipathos¹ (2019) da Embrapa (BARBEDO; KOENIGKAN; SANTOS, 2016), mais especificamente do conjunto de itens "Feijao (*Dry bean*) - Antracnose (*Anthracnose*) - *Cropped*", que compreende 561 imagens de folíolos lesionados do feijoeiro, seccionadas digitalmente em áreas com ocorrência da doença, destacando especificamente as nervuras da face inferior (abaxial) dos folíolos das plantas.

Outras 439 imagens foram extraídas da base UEPGANtrac1, adquiridas durante o ano de 2018, no contexto dos trabalhos desenvolvidos junto ao Departamento de Fitotecnia e Fitossanidade da Universidade Estadual de Ponta Grossa, em experimentos desenvolvidos na cultura do feijão com a moléstia antracnose, em áreas de cultivo desta instituição. Este subconjunto de imagens da base TESTAntrac está relacionado ao rótulo da classe POSITIVA. A classe NEGATIVA, ou seja, que não apresenta a desordem antracnose, foi constituída de imagens de plantas saudáveis em registros da base UEPGANtrac1 e imagens do repositório Digipathos, também no estilo *cropped*, com outras desordens distintas que não a antracnose.

O processo de aquisição do conjunto de dados TESTAntrac, bem como da base UEPGANtrac1 citada anteriormente, foi conduzido de forma que as imagens capturadas tivessem como alvo a face abaxial dos folíolos da planta, onde se observa a incidência da doença com mais facilidade (CHRISTMANN, 2019). As imagens foram coletadas em diferentes estágios de evolução das plantas e da doença e foram digitalmente cortadas - *cropped* - nas áreas lesionadas, conforme o padrão apresentado na base Digipathos. As Figuras 17 e 18 mostram duas imagens da base TESTAntrac, respectivamente das classes positiva e negativa.

¹ <https://www.digipathos-rep.cnpia.embrapa.br/jspui/handle/123456789/1097>

Figura 20 – Imagem de exemplo para a classe positiva da base TESTAntrac



Fonte: Digipathos (2019).

Figura 21 – Imagem de exemplo para a classe negativa da base TESTAntrac



Fonte: O Autor

A segunda base de dados, UEPGANtrac, também foi estruturada a partir dos trabalhos desenvolvidos junto ao Departamento de Fitotecnia e Fitossanidade da Universidade Estadual de Ponta Grossa, com antracnose na cultura do feijão, tanto em casa de vegetação quanto em campo, na Fazenda Experimental Capão da Onça. O período de aquisição das imagens deu-se junto ao cultivo da safra, que aconteceu em meados do mês de novembro de 2019, até a colheita dos grãos, por volta do final do mês de janeiro de 2020. A aquisição de imagens nesse período pretendeu compreender distintas condições em termos de ambiente (dias nublados, ensolarados, em diferentes horários, etc) e equipamento, ao utilizar diferentes dispositivos para captura das fotos.

4.2 PROCEDIMENTO DE TREINAMENTO DAS REDES NEURAIAS

Muitos estudos têm sido propostos com a utilização de redes neurais convolucionais para diferentes objetivos, demonstrando sucesso na utilização desta tecnologia. As arquiteturas de redes utilizadas nos experimentos descritos neste trabalho possuem características distintas, conforme apresentado na seção 3.4.7. Desse modo, segue a configuração para o desenvolvimento da pesquisa.

Todo o procedimento experimental foi desenvolvido em linguagem de programação Python utilizando o ambiente de desenvolvimento Google Colab. Foi utilizado ainda o Google Drive para armazenamento das imagens para treinamento, validação e teste. Como biblioteca de redes neurais utilizou-se o Tensorflow, e, por fim, a biblioteca de código aberto para redes neurais Keras como API de alto nível.

Em cada teste, os conjuntos de imagens foram particionados na proporção 7:2:1 (70% dos dados do conjunto para treinamento, 20% para validação e 10% para testes). O treinamento da rede foi feito de acordo com estudos preliminares e amparado pela literatura especializada, utilizando o algoritmo *backpropagation*, com trinta épocas e taxa de aprendizado de 0,0001 (MOHANTY; HUGHES; SALATHÉ, 2016; SUN et al., 2017; TOO et al., 2019). As redes InceptionV3, ResNet50 e VGG16 foram carregadas sem pesos pré-definidos e com uma camada de classificação padronizada adicionada posteriormente. Dessa maneira, a estrutura das redes compreende, em suma, as camadas de extração de características, sendo a classificação efetuada por camadas de redimensionamento e abandono, de acordo com as duas classes propostas.

4.3 EXPERIMENTOS

Inicialmente, o desempenho das redes neurais convolucionais avaliadas neste trabalho foi mensurado em três condições experimentais diferentes. A definição de cada condição foi determinada com a especificação de dois parâmetros, $p1$ e $p2$. O parâmetro $p1$ é uma variável proposicional cujo valor verdadeiro (V) indica que as bases de dados usadas no treinamento e validação de classificadores são submetidas à ampliação de dados. O parâmetro $p2$ se refere ao emprego da ampliação de dados sobre o conjunto de teste; quando a variável $p2$ tem valor verdadeiro (V) denota a execução do procedimento de ampliação de dados sobre a partição de testes da base.

Dado o exposto, o primeiro experimento, identificado por eFF, se refere a um conjunto de testes que avaliou o desempenho das redes InceptionV3, VGG16 e ResNet50 assumindo que ambos os parâmetros $p1$ e $p2$ eram falsos. O segundo, cujo rótulo é eFV, diz respeito a uma série de testes em que $p1 = F$ e $p2 = V$. O terceiro experimento, cuja notação é eVV, trata de testes em que tanto $p1$ quanto $p2$ eram verdadeiros. O Quadro 3 resume cada uma das condições experimentais descritas.

Quadro 3 – Condições experimentais para o Experimento 1

Experimento	Os conjuntos de treinamento e validação foram submetidos à ampliação de dados?	O conjunto de teste foi submetidos à ampliação de dados?
eFF	Não	Não
eFV	Não	Sim
eVV	Sim	Sim

Fonte: O Autor

Durante o procedimento experimental, e de acordo com os valores dos parâmetros $p1$ e $p2$, fez-se o uso combinado de métodos de ampliação de dados baseados nas operações de escala (ampliação/redução) da imagem (MIKOŁAJCZYK; GROCHOWSKI, 2018), rotação (LAKHANI, 2017) e variação de brilho (BHATT et al., 2019). As seguintes combinações foram consideradas na realização dos testes:

- Ampliação de dados 1: aplicação de escala de até 20%, rotação de até 90° e variação de brilho de até 10%;
- Ampliação de dados 2: aplicação de escala de até 20%, rotação de até 90° e variação de brilho de até 20%;
- Ampliação de dados 3: aplicação de escala de até 20%, rotação de até 90° e variação de brilho de até 30%.

O Quadro 4 apresenta uma descrição do primeiro procedimento experimental empregado neste estudo na forma de um pseudocódigo.

Quadro 4 – Procedimento experimental 1

Procedimento Experimental 1

- Entrada:
 - um modelo de Rede Neural Convolutacional: R ;
 - o método de ampliação de dados: A ,
 - booleano: p_1, p_2 ;
- Definições:
 - D denota a base de dados utilizada

início

1. Para $i = 1$ até 10:

- a. gere três partições do conjunto D, D_T, D_V, D_t para treinamento, validação e teste de R ;
- b. se $p_1 = V$, então execute a rotina de ampliação de dados A sobre D_T e D_V ;
- c. se $p_2 = V$, então execute a rotina de ampliação de dados A sobre D_t ;
- d. treine R sobre os dados de D_T ;
- e. Medir o desempenho da rede na base de validação D_V ;
- f. Medir o desempenho da rede na base de teste D_t ;
- g. registrar os resultados mensurados;

2. fim-para

fim

Fonte: O Autor

Como pode ser observada, a rotina descrita no Quadro 4 tem quatro argumentos: a rede neural convolutacional a ser treinada e avaliada (denotada por m); a especificação da ampliação de dados a ser considerada no teste (indicado por A) e os parâmetros p_1 e p_2 . O parâmetro R tem três valores, I, R e V , que indicam o uso das redes InceptionV3, ResNet50 ou VGG16, respectivamente. O parâmetro A pode assumir os valores 0, 1, 2, 3. A especificação $A = 0$ informa que nenhum procedimento de ampliação de dados foi utilizado. Já A igual 1, 2 ou 3 determina o emprego de uma das três ampliações de dados combinada listadas anteriormente.

Após a inicialização dos parâmetros a base de dados foi particionada em conjuntos de treino, validação e teste. Em seguida a operação de ampliação de dados ocorreu conforme: se p_1 é verdadeiro, a ampliação de dados era aplicada sobre as partições de treino e validação; se p_2 é verdadeiro a ampliação de dados era aplicada sobre a partição de testes. A rede foi treinada sobre a respectiva base e os desempenhos aferidos. Por fim os resultados foram registrados em arquivo.

As variações dos testes realizados no primeiro experimento, em relação às arquiteturas de redes, procedimentos de ampliação de dados e a aplicação destes, estão listadas nos Quadros 5, 6 e 7. As siglas que nomeiam os experimentos foram compostas pela inicial da arquitetura de rede utilizada, seguida pelo valor numérico que identifica o procedimento de ampliação de dados aplicado e uma dupla de letras V e F indicando o particionamento em que foi aplicada a ampliação de dados.

Quadro 5 – Procedimento experimental eFF: ampliação de dados

Experimento	Arquitetura de rede	Ampliação de dados	Ampliação de dados em D_T, D_V	Ampliação de dados em D_t
I0FF	InceptionV3	Não	Não	Não
R0FF	ResNet50	Não	Não	Não
V0FF	VGG16	Não	Não	Não

Fonte: O Autor

Quadro 6 – Procedimento experimental eFV: ampliação de dados em D_t

Experimento	Arquitetura de rede	Ampliação de dados	Ampliação de dados em D_T, D_V	Ampliação de dados em D_t
I1FV	InceptionV3	A1	Não	Sim
I2FV	InceptionV3	A2	Não	Sim
I3FV	InceptionV3	A3	Não	Sim
R1FV	ResNet50	A1	Não	Sim
R2FV	ResNet50	A2	Não	Sim
R3FV	ResNet50	A3	Não	Sim
V1FV	VGG16	A1	Não	Sim
V2FV	VGG16	A2	Não	Sim
V3FV	VGG16	A3	Não	Sim

Fonte: O Autor

Quadro 7 – Procedimento experimental eVV: ampliação de dados em D_T, D_V e D_t

Experimento	Arquitetura de rede	Ampliação de dados	Ampliação de dados em D_T, D_V	Ampliação de dados em D_t
I1VV	InceptionV3	A1	Sim	Sim
I2VV	InceptionV3	A2	Sim	Sim
I3VV	InceptionV3	A3	Sim	Sim
R1VV	ResNet50	A1	Sim	Sim
R2VV	ResNet50	A2	Sim	Sim
R3VV	ResNet50	A3	Sim	Sim
V1VV	VGG16	A1	Sim	Sim
V2VV	VGG16	A2	Sim	Sim
V3VV	VGG16	A3	Sim	Sim

Fonte: O Autor

O segundo experimento seguiu as características apresentadas no primeiro, sendo a base de testes fixa UEPGANtrac e as condições FF e FV, conforme o Quadro 8. Cada configuração experimental foi repetida dez vezes

Quadro 8 – Condições experimentais para o Experimento 2

Experimento	Os conjuntos de treinamento e validação foram submetidos à ampliação de dados?	O conjunto de teste foi submetidos à ampliação de dados?
eFF	Não	Não
eFV	Não	Sim

Fonte: O Autor

4.4 ANÁLISE DOS RESULTADOS

Inicialmente, as imagens do conjunto TESTAntrac, pré e pós-ampliação de dados, foram comparadas através dos índices de erro quadrático médio (*mean squared error* - MSE) e de similaridade estrutural (*structural similarity* - SSIM). Conforme Winkler (2005), o erro quadrático médio é a média das diferenças quadráticas entre os valores no nível de cinza dos pixels em duas figuras ou sequências. O índice de similaridade estrutural, proposto por Wang et al. (2004), leva em consideração características estruturais da imagem. Testes estatísticos foram empregados para verificar se os diferentes procedimentos de ampliação de dados produziram conjuntos de imagens que diferiam significativamente entre si.

Após o cálculo da acurácia, da sensibilidade, da especificidade e da observação acerca do sobreajuste dos classificadores nos experimentos, foi realizado o cômputo da média e do desvio padrão destes resultados. Os dados foram agrupados considerando as diferentes condições experimentais simuladas, arquiteturas de redes e configuração dos procedimentos de ampliação de dados. Na sequência, as informações obtidas foram usadas para realizar análises estatísticas para avaliar o desempenho dos classificadores e a eficácia da ampliação de dados. Realizou-se uma análise de variância (ANOVA) para verificar a existência de diferenças nas médias dos escores de desempenho e sobreajuste das redes InceptionV3, ResNet50 ou VGG16, fixadas as condições de teste e o procedimento de ampliação de dados. O objetivo desta análise foi determinar em que condições os classificadores testados tinham um desempenho distinto aos demais no que se refere aos índices apurados.

Em seguida, cada indicador de desempenho foi submetido a um teste de Friedman com *post hoc* dentro de cada condição experimental. Em cada uma das três análises as combinações usadas na ampliação de dados definiram os blocos do experimento e as arquiteturas de rede dos diferentes tratamentos. A motivação para execução deste teste estatístico foi detectar o efeito da ampliação de dados sobre o desempenho dos classificadores.

Finalmente, realizou-se a comparação do desempenho de cada classificador nas diferentes condições experimentais (eFF, eFV,eVV). Para tanto, tomou-se os escores de desempenho de cada classificador, considerando cada um dos esquemas de ampliação de dados, e realizou-se um teste de teste de Friedman com *post hoc* em que as configurações de ampliação de dados definiram os grupos.

5 RESULTADOS

Neste capítulo são apresentados os resultados obtidos no decorrer desta pesquisa, desde a avaliação da simulação por ampliação de dados (5.1) que apresenta uma comparação entre as imagens pré e pós ampliação; os desempenhos do primeiro experimento, que levou em consideração ausência de ampliação de dados (5.2), ampliação de dados exclusiva na fase de testes (5.3), e ampliação de dados em ambas as fases – treino e testes – (5.4); e, por fim, os resultados do segundo experimento, em que a partição de testes do primeiro experimento foi substituída por uma outra coletada em situação distinta (5.5).

5.1 AVALIAÇÃO DA SIMULAÇÃO POR AMPLIAÇÃO DE DADOS

Foi verificado se a aplicação dos procedimentos de ampliação de dados produziu conjuntos de imagens com conteúdo significativamente diferente daquele observado nas imagens da base original. Para tanto, computou-se os índices de erro quadrático médio (MSE) e de similaridade estrutural (SSIM) entre cada imagem gerada e sua correspondente após os distintos processos de ampliação de dados.

Considerando que a imagem original comparada com ela própria resulta em valores de 0,000 e 1,000 para os índices MSE e SSIM respectivamente, tem-se que os valores de MSE aumentam conforme é maior a diferença entre as imagens; enquanto os valores do SSIM diminuem na mesma situação (WANG et al., 2004). Os valores computados estão resumidos na Tabela 1, que mostra a média do MSE e do SSIM, bem como o desvio padrão, para cada parametrização do procedimento de ampliação de dados (DA).

Tabela 1 – Média dos índices de comparação entre imagens

DA	MSE		SSIM	
	Média	Desvio Padrão	Média	Desvio Padrão
1	2455,862	1915,652	0,328	0,176
2	1670,476	1320,517	0,406	0,185
3	2268,269	1532,052	0,305	0,129

Os valores da tabela mostram que as diferenças entre os conjuntos pré e pós ampliação de dados foram maiores para a primeira e terceira parametrização e menores na segunda. O teste de Friedman com *post hoc* mostra que não há diferença estatística entre os valores de MSE nas distintas ampliações de dados. Já para os valores de SSIM existe diferença estatística na comparação entre os procedimentos de ampliação de dados 1 e 2.

5.2 DESEMPENHO NA AUSÊNCIA DA AMPLIAÇÃO DE DADOS

A Tabela 2 apresenta os valores de acurácia registrados pelos classificadores baseados nas arquiteturas de redes neurais convolucionais InceptionV3, ResNet50 e VGG16 na categorização de instâncias nos conjuntos de treino e de teste. Também estão listados os valores de desvio padrão para essa medida de desempenho. Estes resultados foram obtidos quando da condução experimental sem aplicação de métodos de ampliação de dados nas bases.

Tabela 2 – Resultados para eFF no primeiro experimento

Experimento	Acurácia no conjunto de treino		Acurácia no conjunto de testes	
	Média	Desvio padrão	Média	Desvio padrão
I0FF	0,966	0,009	0,919	0,036
R0FF	0,947	0,040	0,889	0,032
V0FF	0,945	0,009	0,922	0,018

Fonte: O Autor

Não houve diferença estatística na acurácia da fase de teste entre as redes neurais VGG16 e InceptionV3, baseando-se no teste de Friedman com *post hoc*. Em termos relativos, a diferença da acurácia média da arquitetura VGG16 em relação a ResNet50 foi de 3,7%. Em relação à comparação de acurácia entre as fases de treino e de teste, houve diferença estatística nos três experimentos, conforme o Teste T. A queda no desempenho entre a fase de treinamento e a fase de testes foi de 5,2%, 6,6% e 2,5% para InceptionV3, ResNet50 e VGG16.

Os resultados obtidos são consistentes com os observados na literatura, alcançando valores de acurácia superiores a 90% para o propósito especificado, ainda que sem o uso de procedimentos de ampliação de dados. Abas et al. (2018) utilizaram a arquitetura de rede VGG16 para classificação de plantas, baseando-se em imagens de flores, e alcançaram resultados de acurácia de 96,25% e 89,96% para treino e teste respectivamente, considerando o treinamento do modelo em uma base com 2800 imagens. Houve uma diminuição no desempenho da acurácia de 6,9% entre as fases, similar aos valores observados neste estudo.

5.3 DESEMPENHO COM AMPLIAÇÃO DE DADOS NA FASE DE TESTE

A Tabela 3 lista os resultados dos experimentos que avaliaram o desempenho dos classificadores nos ensaios em que apenas o conjunto de testes foi submetido aos procedimentos de ampliação de dados, conforme descrito na seção 4.3. Deste modo, como observado na seção 5.1, as imagens de teste usadas neste experimento são distintas das utilizadas no treinamento devido a diferenças na iluminação, rotação e escala oriundas dos procedimentos de ampliação de dados.

A primeira coluna da Tabela 3 informa a configuração experimental aplicada. A segunda e a terceira colunas enumeram os valores de média e desvio padrão da acurácia obtidos na fase de treinamento. Por fim, nas duas últimas colunas tem-se os valores de média e desvio padrão da acurácia durante a fase de testes.

Tabela 3 – Resultados para eFV no primeiro experimento

Experimento	Acurácia no conjunto de treino		Acurácia no conjunto de testes	
	Média	Desvio padrão	Média	Desvio padrão
I1FV	0,969	0,006	0,777	0,045
I2FV	0,961	0,012	0,782	0,044
I3FV	0,967	0,008	0,796	0,039
R1FV	0,926	0,098	0,673	0,070
R2FV	0,958	0,015	0,691	0,067
R3FV	0,962	0,007	0,690	0,062
V1FV	0,944	0,007	0,843	0,051
V2FV	0,939	0,007	0,845	0,043
V3FV	0,941	0,008	0,843	0,023

Observa-se que, de uma maneira geral, a arquitetura de rede neural convolucional VGG16, com os experimentos V1FV, V2FV e V3FV, apresentou maiores valores de acurácia para a fase de testes que as demais redes na mesma fase. A arquitetura ResNet50 obteve os menores valores na mesma comparação (R1FV, R2FV e R3FV). Na média, a arquitetura de rede VGG16 teve uma acurácia na partição de testes 23% superior a ResNet50. A análise de variância (ANOVA) mostrou que as diferentes parametrizações de ampliação de dados não levaram a diferenças estatísticas significativas na acurácia das redes durante a fase de testes.

A diferença entre os conjuntos de imagens, derivada dos procedimentos de ampliação de dados, gerou maiores discrepâncias entre os valores de acurácia das fases de treino e de testes, quando da ampliação exclusiva na partição de testes (eFV). Basicamente, o desempenho de uma rede na etapa de teste ser inferior àquele registrado durante o treinamento é um indicativo de sobreajuste (ELKAN, 2012). A acurácia da arquitetura ResNet50 apresentou queda média de 38% no desempenho entre treino e teste nas três configurações de ampliação de dados, enquanto que a rede InceptionV3 apresentou diminuição média de 23%. A arquitetura de rede VGG16 apresentou os menores valores de diminuição do desempenho entre as fases de treino e teste, com média de 11% queda para as mesmas condições experimentais.

A Tabela 4 lista o valor de *p-value* para as comparações que apresentaram diferenças estatísticas para os valores de acurácia do teste neste grupo experimental (eFV), segundo o teste de Friedman.

Tabela 4 – Teste de Friedman para eFV no primeiro experimento

H_0	<i>p-value</i>
R1FV = V1FV	0,00043
R1FV = V3FV	0,00043
R1FV = V2FV	0,00061
V1FV = R3FV	0,00073
R3FV = V3FV	0,00073
V2FV = R3FV	0,00104
V1FV = R2FV	0,00206
R2FV = V3FV	0,00206
R2FV = V2FV	0,00287

O desempenho das arquiteturas de rede ResNet50 e VGG16 diferem entre si em todas as configurações dos procedimentos de ampliação de dados nesta configuração experimental. Não foram detectadas diferenças estatísticas de desempenho entre as redes InceptionV3 e VGG16.

5.4 DESEMPENHO COM AMPLIAÇÃO DE DADOS EM TREINO E TESTE

A Tabela 5 apresenta os resultados dos experimentos que avaliaram o desempenho dos classificadores em ensaios onde as partições de treino e de testes foram simultaneamente submetidas aos procedimentos de ampliação de dados.

Tabela 5 – Resultados eVV no primeiro experimento

Experimento	Acurácia no conjunto de treino		Acurácia no conjunto de testes	
	Média	Desvio padrão	Média	Desvio padrão
I1VV	0,906	0,015	0,915	0,028
I2VV	0,901	0,009	0,907	0,020
I3VV	0,897	0,007	0,910	0,026
R1VV	0,547	0,120	0,546	0,104
R2VV	0,631	0,165	0,618	0,161
R3VV	0,536	0,106	0,548	0,125
V1VV	0,888	0,011	0,875	0,062
V2VV	0,886	0,013	0,882	0,030
V3VV	0,879	0,011	0,891	0,027

Fonte: O Autor

Nesse conjunto de resultados, referente à condição experimental de aplicação da ampliação de dados sobre as partições de treino e teste simultaneamente (eVV), observa-se maiores níveis de acurácia para a arquitetura de rede neural convolucional InceptionV3. A arquitetura ResNet50 apresentou resultados mínimos para acurácia nas diferentes condições de ampliação de dados. Observou-se ainda maior consistência entre os valores de acurácia para treino e teste nos diferentes casos experimentais, com variações inferiores a 2%.

Em média, os valores de acurácia obtidos na fase de testes da configuração experimental eVV foram superiores àqueles observados na configuração eFV. Conforme Shorten e Khoshgoftaar (2019) pode-se alcançar um nível de predição mais robusto ao ampliar as imagens de teste do mesmo modo que as de treino, apesar do custo computacional envolvido nessa tarefa. Esse argumento é corroborado por Valle et al. (2019) quando reportam que métodos de ampliação de dados permitem alcançar melhor desempenho quando aplicados simultaneamente às partições de treino e teste de um conjunto.

A análise de variância (ANOVA) mostrou que as diferentes parametrizações de ampliação de dados deste conjunto experimental (eVV) não levaram a diferenças estatísticas significativas na acurácia das redes durante a fase de testes.

A Tabela 6 lista o *p-value* para as comparações que apresentaram diferenças estatísticas para os valores de acurácia do teste neste grupo experimental (eVV), segundo o teste de Friedman.

Tabela 6 – Teste de Friedman para eVV no primeiro experimento

H_0	<i>p-value</i>
R1VV = I3VV	0,00024
I1VV = R1VV	0,00029
I3VV = R3VV	0,00043
I1VV = R3VV	0,00051
R1VV = I2VV	0,00147
R2VV = I3VV	0,00147
I1VV = R2VV	0,00174
I2VV = R3VV	0,00243
I2VV = R2VV	0,00736

Fonte: O Autor

Observa-se que o desempenho das arquiteturas de rede InceptionV3 e ResNet50 diferiram entre si em todas as parametrizações dos procedimentos de ampliação de dados. Não foram detectadas diferenças estatísticas de desempenho entre as redes InceptionV3 e VGG16. No que se refere às redes VGG16 e ResNet50, as únicas diferenças foram observadas nos testes R1VV e V3VV.

5.5 DESEMPENHO COM PARTIÇÃO DE TESTES DISTINTA

O segundo conjunto de experimentos desenvolvidos é relativo ao uso da base de dados UEPGANtrac substituindo a partição de testes dos experimentos anteriores. Dessa forma, a etapa de testes é conduzida utilizando uma base de dados com características distintas daquela utilizada para treinamento. Os resultados, conforme o Quadro 7, referem-se a condução experimental com ampliação de dados na partição de teste (eFV) e sem essa técnica em ambas partições (eFF).

Em se tratando da ausência de ampliação, a Tabela 7 apresenta os valores de acurácia obtidos pelos classificadores nas partições de treino e de teste nesta configuração, bem como os respectivos valores de desvio padrão.

Tabela 7 - Resultados para eFF no segundo experimento

Experimento	Acurácia no conjunto de treino		Acurácia no conjunto de testes	
	Média	Desvio padrão	Média	Desvio padrão
I0FF	0,968	0,005	0,803	0,115
R0FF	0,977	0,008	0,682	0,093
V0FF	0,947	0,004	0,867	0,042

Fonte: O Autor

Nestas condições é possível observar que os valores de acurácia apresentados no conjunto de treino são similares aos observados no experimento conduzido no capítulo 5.2 (desempenho com uma única partição subdividida para treino e testes, sem uso de técnicas de ampliação de dados). Entretanto, os valores de acurácia obtidos neste conjunto de testes são, em geral, inferiores aos observados naquele experimento. Em relação à comparação dos valores de acurácia entre as fases de treino e de teste, houve diferença estatística nos três experimentos, conforme o Teste t. A queda de desempenho entre as fases de treinamento e testes foi de 20,5%, 43,2% e 9,2% para InceptionV3, ResNet50 e VGG16, respectivamente. Os resultados apresentados são consistentes com os anteriormente observados, bem como com os da literatura especializada.

A Tabela 8 apresenta os valores de acurácia e desvio padrão relativos à avaliação de desempenho dos classificadores nas partições de treino e de teste em ensaios com o uso da base de dados UEPGANtrac substituindo a partição de testes. Nesta configuração experimental houve também a aplicação das diferentes técnicas de ampliação de dados propostas na partição de treino.

Tabela 8 - Resultados para eFV no segundo experimento

Experimento	Acurácia no conjunto de treino		Acurácia no conjunto de testes	
	Média	Desvio padrão	Média	Desvio padrão
I1FV	0,969	0,005	0,712	0,109
I2FV	0,972	0,005	0,670	0,107
I3FV	0,972	0,010	0,739	0,071
R1FV	0,972	0,021	0,612	0,107
R2FV	0,977	0,005	0,661	0,073
R3FV	0,973	0,011	0,680	0,063
V1FV	0,946	0,005	0,747	0,047
V2FV	0,946	0,005	0,695	0,091
V3FV	0,952	0,007	0,675	0,070

Fonte: O Autor

Os resultados de acurácia no conjunto de testes para a arquitetura de rede neural convolucional ResNet50, com os experimentos R1FV, R2FV e R3FV, apresentaram, de uma maneira geral, os valores mais baixos nessa parametrização experimental. As arquiteturas InceptionV3 e VGG16 apresentaram resultados similares de acurácia para a fase de testes. Quanto à da acurácia no conjunto de treino, as três arquiteturas de rede apresentaram resultados médios aproximados.

É possível também observar nesta configuração experimental uma diminuição nos valores de acurácia do conjunto de testes para o conjunto de treinamento. Em se tratando das técnicas de ampliação de dados, a análise de variância (ANOVA) mostrou que as diferentes parametrizações não levaram a diferenças estatísticas significativas na acurácia das redes durante a fase de testes.

A Tabela 9 lista o valor de *p-value* para as comparações que apresentaram diferenças estatísticas para os valores de acurácia do teste neste grupo experimental, segundo o teste de Friedman.

Tabela 9 – Teste de Friedman para eFV no segundo experimento

H_0	<i>p-value</i>
R1FV = V1FV	0,00043
R1FV = V3FV	0,00043
R1FV = V2FV	0,00061
V1FV = R3FV	0,00073
R3FV = V3FV	0,00073
V2FV = R3FV	0,00104
V1FV = R2FV	0,00206
R2FV = V3FV	0,00206
R2FV = V2FV	0,00287

Fonte: O Autor

O desempenho das arquiteturas de rede ResNet50 e VGG16 diferem entre si em todas as configurações experimentais deste ensaio. Não foram detectadas diferenças estatísticas no desempenho das redes InceptionV3 e VGG16.

A utilização de conjuntos distintos de imagens nas etapas de treinamento e de testes foi abordada por Goncharov et al. (2018), observando que mesmo utilizando uma base de grandes dimensões (base de dados *PlantVillage*, com 86147 imagens de plantas saudáveis e doentes) para treinamento, o modelo classificador não demonstrava maior capacidade de generalização resultando, quando da condução de testes em imagens coletadas em condições distintas, em valores de acurácia de 48%. A mesma base de dados *PlantVillage* particionada e utilizada simultaneamente para as etapas de treinamento e de testes resulta em um modelo capaz de alcançar 99% de acurácia.

Os valores médios de sensibilidade, especificidade e sobreajuste para este ensaio foram próximos entre as diferentes configurações experimentais, com algum destaque para a arquitetura de rede ResNet50. A Tabela 10 apresenta os valores para sensibilidade, especificidade e sobreajuste.

Tabela 10 – Valores de sensibilidade, especificidade e sobreajuste

Experimento	Sensibilidade	Especificidade	Sobreajuste
I1FV	78,10%	68,30%	25,78%
I2FV	69,60%	67,00%	30,21%
I3FV	76,80%	67,50%	23,33%
R1FV	61,70%	60,20%	36,06%
R2FV	76,40%	56,30%	31,60%
R3FV	75,40%	58,10%	29,28%
V1FV	78,10%	72,00%	19,90%
V2FV	70,50%	69,00%	25,11%
V3FV	65,50%	69,50%	27,79%

Fonte: O Autor

A arquitetura de rede neural ResNet50 apresentou, em média, valores mais expressivos para sobreajuste e valores ligeiramente mais baixos para a medida de especificidade.

6 CONCLUSÃO

Este estudo analisou o uso de ampliação de dados para simular o efeito de alterações nas condições de aquisição de imagens do folíolo do feijoeiro no treinamento de redes neurais convolucionais para a detecção da moléstia antracnose. Os resultados indicam que a ampliação de dados permite a condução de testes em condições diferentes daquelas do treinamento das redes neurais convolucionais. Isto, como mostra a seção 5.3, pode levar ao chamado sobreajuste. Neste contexto, observou-se queda média de 38% no desempenho entre treino e teste nas três configurações de ampliação de dados para a arquitetura de rede ResNet50, 23% para InceptionV3 e de 11% em média para VGG16.

Os resultados apresentados na seção 5.4 também mostram que a aplicação da ampliação de dados sobre o conjunto de treinamento pode reduzir ou eliminar o sobreajuste. Nesse caso, notou-se diminuição inferior a 2% no desempenho entre treino e teste em todos os casos experimentais. A seção 5.5 apresentou os resultados quando da utilização de uma base de dados distinta na partição de testes, situação em que o classificador foi submetido a imagens fora do escopo daquelas do treinamento. Nessa situação os resultados foram inferiores àqueles da seção 5.4.

A condução deste trabalho propiciou a coleta de imagens e consequente elaboração de uma base de dados de sintomas de antracnose em folíolos de feijoeiro. A utilização de arquiteturas de redes neurais convolucionais, concomitante ao artifício da ampliação de dados, mostrou-se conveniente aos fins de permitir maior generalização dos conjuntos de dados.

Em consonância com a bibliografia consultada e de acordo com os resultados aqui apresentados, sugere-se a condução de demais experimentos para ampliar as propostas trazidas neste estudo, como a ampliação da base de dados utilizada através de esforços colaborativos, como nos trabalhos de Bhat; Sarangi e Pappula (2019) e Chen et al. (2020); a utilização de outras arquiteturas de redes neurais como Arivazhagan; Ligi (2018) e Zhang et al. (2019); e ainda , a utilização de refinadas técnicas de ampliação de dados, como nos trabalhos de Cap et al. (2020) e Nazki et al. (2020).

REFERÊNCIAS

- ABAS, Mohamad Aqib Haqmi et al. VGG16 for plant image classification with transfer learning and data augmentation. **International Journal of Engineering and Technology (UAE)**, v. 7, p. 90-94, 2018.
- AGGARWAL, Charu C. **Data Classification: Algorithms and Applications**. Estados Unidos: CRC Press, 2015.
- AJI, Alham F. et al. Detection of palm oil leaf disease with image processing and neural network classification on mobile device. **International Journal of Computer Theory and Engineering**, v. 5, n. 3, p. 528, 2013.
- AMARA, Jihen et al. A Deep Learning-based Approach for Banana Leaf Diseases Classification. *In: DATENBANKSYSTEME FÜR BUSINESS, TECHNOLOGIE UND WEB (BTW 2017) – WORKSHOPBAND. Proceedings [...]*, 2017, p. 79-88.
- AMORIM, L. et al. Clorose variegada dos citros: uma escala diagramática para avaliação da severidade da doença. **Fitopatologia Brasileira**, v. 18, n. 2, p. 174-180, 1993.
- ARGÜESO, David et al. Few-Shot Learning approach for plant disease classification using images taken in the field. **Computers and Electronics in Agriculture**, v. 175, p. 105542, 2020.
- ARIVAZHAGAN, S.; LIGI, S. Vineth. Mango Leaf Diseases Identification Using Convolutional Neural Network. **International Journal of Pure and Applied Mathematics**, v. 120, n. 6, p. 11067-11079, 2018.
- ARROQUI, Mauricio et al. RESTful Web Services improve the efficiency of data transfer of a whole-farm simulator accessed by Android smartphones. **Computers and electronics in agriculture**, v. 87, p. 14-18, 2012.
- BARBEDO, Jayme Garcia Arnal. Digital image processing techniques for detecting, quantifying and classifying plant diseases. **SpringerPlus**, v. 2, n. 1, p. 660, 2013.
- BARBEDO, Jayme Garcia Arnal. A review on the main challenges in automatic plant disease identification based on visible range images. **Biosystems Engineering**, v. 144, p. 52-60, 2016.
- BARBEDO, J. G. A.; KOENIGKAN, L. V.; SANTOS, T. T. Identifying multiple plant diseases using digital image processing. **Biosystems Engineering**, v. 147, p. 104-116, 2016.
- BARBOSA, Flávia Rabelo; GONZAGA, AC de O. **Informações técnicas para o cultivo do feijoeiro-comum na Região Central-Brasileira: 2012-2014**. Goiânia: Embrapa Arroz e Feijão, 2012.
- BARBOSA, Livia. Feijão com arroz e arroz com feijão: o Brasil no prato dos brasileiros. **Horizontes Antropológicos**, v. 13, n. 28, p. 87-116, 2007.

BARBOSA, Maria Angélica G.; MICHEREFF, Sami J.; MORA-AGUILERA, Gustavo. Elaboração e validação de escala diagramática para avaliação da severidade da ferrugem branca do crisântemo. **Summa Phytopathologica**, v. 32, n. 1, p. 57-62, 2006.

BHATT, Prakruti V.; SARANGI, Sanat; PAPPULA, Srinivasu. Detection of diseases and pests on images captured in uncontrolled conditions from tea plantations. *In: AUTONOMOUS AIR AND GROUND SENSING SYSTEMS FOR AGRICULTURAL OPTIMIZATION AND PHENOTYPING IV. Proceedings [...]* 2019. p. 1100808.

CANZIANI, Alfredo; PASZKE, Adam; CULURCIELLO, Eugenio. An analysis of deep neural network models for practical applications. **arXiv preprint arXiv:1605.07678**, 2016.

CAP, Quan Huu et al. LeafGAN: An Effective Data Augmentation Method for Practical Plant Disease Diagnosis. **arXiv preprint arXiv:2002.10100**, 2020.

CHEN, Junde et al. Using deep transfer learning for image-based plant disease identification. **Computers and Electronics in Agriculture**, v. 173, p. 105393, 2020.

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira: grãos**, v. 8 Safra 2017/18 - Oitavo levantamento, Brasília, p. 1-145 maio 2018.

COURBARIAUX, Matthieu; BENGIO, Yoshua; DAVID, Jean-Pierre. BinaryConnect: Training deep neural networks with binary weights during propagations. **Advances in neural information processing systems**. p. 3123-3131, 2015.

CHRISTMANN, Polyana Elvira Tobias Pinto. **Produtos alternativos aplicados na cultura do feijão para controle da antracnose**. 2019. Dissertação (Mestrado em Agronomia) - Universidade Estadual de Ponta Grossa, Ponta Grossa, 2019.

DAMASCENO, Vítor Flávio Furtado; FURTADO, Edson Luiz; FERREIRA FILHO, Pedro José. Comparação de dois métodos de elaboração e validação de escala diagramática para a quantificação da severidade da mancha de *Cylindrocladium* em eucalipto. **Summa Phytopathologica**, v. 40, p. 248-255, 2014.

DELGADO, Jorge A.; KOWALSKI, Kevin; TEBBE, Caleb. The first Nitrogen Index app for mobile devices: Using portable technology for smart agricultural management. **Computers and electronics in agriculture**, v. 91, p. 121-123, 2013.

DIGIPATHOS **Base de Imagens de Sintomas de Doenças de Plantas PDDB: Feijao (Dry bean) - Antracnose (Anthracnose) – Cropped** Disponível em: <https://www.digipathos-rep.cnptia.embrapa.br/jspui/handle/123456789/1097>. Acesso em: 23 nov. 2019.

DUDA, Richard O.; HART, Peter E.; STORK, David G. **Pattern classification**. Alemanha: John Wiley & Sons, 2012.

ELKAN, Charles. **Evaluating classifiers**. San Diego: University of California, 2012.

FAOSTAT. **Crops**. Disponível em: <http://www.fao.org/faostat/en/#data/QC>. Acesso em: 23 nov. 2019.

FERENTINOS, Konstantinos P. Deep learning models for plant disease detection and diagnosis. **Computers and Electronics in Agriculture**, v. 145, p. 311-318, 2018.

FERNANDES, Bruno José Torres. **Redes neurais com extração implícita de características para reconhecimento de padrões visuais**. 2013. Tese (Doutorado em Ciência da Computação). Universidade Federal de Pernambuco, Pernambuco, 2013.

FERREIRA, Janderson et al. Extreme learning machine autoencoder for data augmentation. *In*: 2017 IEEE LATIN AMERICAN CONFERENCE ON COMPUTATIONAL INTELLIGENCE (LA-CCI). **Proceedings** [...] 2017. p. 1-6.

FUENTES, Alvaro et al. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. **Sensors**, v. 17, n. 9, p. 2022, 2017.

GODOY, Cláudia V. et al. Diagrammatic scales for bean diseases: development and validation. **Journal of Plant Diseases and Protection**, p. 336-345, 1997.

GONÇALVES, Eduardo C. **Novel classifier chain methods for multi-label classification based on genetic algorithms**. 2015. Tese (Doutorado em Computação) – Universidade Federal Fluminense, Niterói, 2015.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning**. Massachusetts: MIT press, 2016. Disponível em: <http://www.deeplearningbook.org>. Acesso em: 03 out. 2019.

GLOROT, Xavier; BORDES, Antoine; BENGIO, Yoshua. Deep sparse rectifier neural networks. *In*: FOURTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND STATISTICS. **Proceedings** [...] 2011. p. 315-323.

HALDER, Monishanker; SARKAR, Ananya; BAHAR, Habibullah. Plant disease detection by image processing: a literature review. **SDRP Journal of Food Science & Technology**. v. 3, ed. 6, 2019.

HARMS, Mônica Gabrielle. **Controle da antracnose em feijão com produtos alternativos**. 2016. Tese (Doutorado em Agricultura) - Universidade Estadual de Ponta Grossa, Ponta Grossa, 2016.

HAWKINS, Douglas M. The problem of overfitting. **Journal of chemical information and computer sciences**, v. 44, n. 1, p. 1-12, 2004.

HAYKIN, Simon. **Redes neurais: princípios e prática**. Porto Alegre: Bookman, 2001.

HE, Kaiming et al. Deep residual learning for image recognition. *In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. Proceedings [...]* 2016. p. 770-778.

HINTON, Geoffrey E. et al. Improving neural networks by preventing co-adaptation of feature detectors. **arXiv preprint arXiv:1207.0580**, 2012.

HOLT, Donald A. Computers in production agriculture. **Science**, v. 228, n. 4698, p. 422-427, 1985.

JARRETT, Kevin et al. What is the best multi-stage architecture for object recognition?. *In: 12TH INTERNATIONAL CONFERENCE ON COMPUTER VISION. Proceedings [...]* 2009. p. 2146-2153.

KESHARI, Rohit et al. Learning structure and strength of CNN filters for small sample size training. *In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. Proceedings [...]* 2018. p. 9349-9358.

KRIZHEVSKY, Alex et al. Imagenet classification with deep convolutional neural networks **Advances in neural information processing systems**, v. 25, p. 1097-1105, 2012.

LAKHANI, Paras. Deep convolutional neural networks for endotracheal tube position and X-ray image classification: challenges and opportunities. **Journal of digital imaging**, v. 30, n. 4, p. 460-468, 2017.

LAVRAČ, Nada; FLACH, Peter; ZUPAN, Blaz. Rule evaluation measures: A unifying view. *In: INTERNATIONAL CONFERENCE ON INDUCTIVE LOGIC PROGRAMMING. Proceedings [...]* Berlin: Heidelberg, 1999. p. 174-185.

LAZAROTO, Adriana et al.; Escala diagramática para avaliação de severidade da helmintosporiose comum em milho. **Ciência Rural**, v. 42, n. 12, p. 2131-2137, 2012.

LECUN, Yann et al. Backpropagation applied to handwritten zip code recognition. **Neural computation**, v. 1, n. 4, p. 541-551, 1989.

LECUN, Yann et al. Convolutional networks for images, speech, and time series. **The Handbook of Brain Theory and Neural Networks**, v. 3361, n. 10, p. 1995, 1995.

LECUN, Yann et al. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278-2324, 1998.

LIMA, Hyanameyka E. et al. Elaboração e validação de escalas diagramáticas para avaliação da severidade da mancha-bacteriana do feijão-caupi em cultivares com trifólios morfológicamente distintos. **Ciência Rural**, v. 43, n. 10, 2013.

LIN, Min; CHEN, Qiang; YAN, Shuicheng. Network in network. **arXiv preprint arXiv:1312.4400**, 2013.

MARTINELLI, Federico et al. Advanced methods of plant disease detection: A review. **Agronomy for Sustainable Development**, v. 35, n. 1, p. 1-25, 2015.

MCLAUGHLIN, Niall; DEL RINCON, Jesus Martinez; MILLER, Paul. Data-augmentation for reducing dataset bias in person re-identification. *In: 12TH IEEE INTERNATIONAL CONFERENCE ON ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE (AVSS)*. **Proceedings** [...] 2015. p. 1-6.

MELO, Geisla de Albuquerque. **Utilização de técnicas de processamento digital de imagens e redes neurais artificiais para o reconhecimento automático da severidade da ferrugem asiática da soja**. 2015. Dissertação (Mestrado em Computação Aplicada) – Universidade Estadual de Ponta Grossa, Ponta Grossa, 2015.

MENEZES JÚNIOR, JB Ferraz. O feijão comum: Taxinomia, morfologia, histologia, parasitologia, microbiologia, composição química e usos. **Revista do Instituto Adolfo Lutz**, v. 20, n. 1, p. 83-104, 1960.

MIKOŁAJCZYK, Agnieszka; GROCHOWSKI, Michał. Data augmentation for improving deep learning in image classification problem. *In: 2018 INTERNATIONAL INTERDISCIPLINARY PHD WORKSHOP*. **Proceedings** [...] 2018. p. 117-122.

MIRANDA FILHO, Reinaldo José de. **Perda de produtividade em feijoeiro comum cultivar Pérola causada por *Curtobacterium flaccumfaciens* pv. *flaccumfaciens***. 2006. Dissertação (Mestrado em Fitopatologia) – Universidade de Brasília, Brasília, 2006.

MOHANTY, Sharada P.; HUGHES, David P.; SALATHÉ, Marcel. Using deep learning for image-based plant disease detection. **Frontiers in plant science**, v. 7, p. 1419, 2016.

MONTOYA, Francisco G. et al. A monitoring system for intensive agriculture based on mesh networks and the android system. **Computers and Electronics in Agriculture**, v. 99, p. 14-20, 2013.

MORAIS, Emerson Cordeiro. **Reconhecimento de padrões e redes neurais artificiais em predição de estruturas secundárias de proteínas**. 2010. Tese (Doutorado em Engenharia de Sistemas e Computação) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

NAIR, Vinod; HINTON, Geoffrey E. Rectified linear units improve restricted boltzmann machines. *In: 27TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML-10)*. **Proceedings** [...] 2010. p. 807-814.

NAZKI, Haseeb et al. Unsupervised image translation using adversarial networks for improved plant disease recognition. **Computers and Electronics in Agriculture**, v. 168, p. 105117, 2020.

OLIVEIRA, Márcia Gonzaga de Castro. et al. **Conhecendo a fenologia do feijoeiro e seus aspectos fitotécnicos**. 2. ed. Brasília: Embrapa Arroz e Feijão, 2018.

OPPENHEIM, Dor; SHANI, Guy. Potato disease classification using convolutional neural networks. **Advances in Animal Biosciences**, v. 8, n. 2, p. 244, 2017.

PEREZ, Fábio et al. Data augmentation for skin lesion analysis. *In: FIRST INTERNATIONAL WORKSHOP ON OR 2.0 CONTEXT-AWARE OPERATING THEATERS. Proceedings [...]* 2018. p. 303-311.

PEREZ, Luis; WANG, Jason. The effectiveness of data augmentation in image classification using deep learning. **arXiv preprint arXiv:1712.04621**, 2017.

PICON, Artzai et al. Crop conditional Convolutional Neural Networks for massive multi-crop plant disease classification over cell phone acquired images taken on real field conditions. **Computers and Electronics in Agriculture**, v. 167, p. 105093, 2019.

PUTLER, Daniel S.; ZILBERMAN, David. Computer use in agriculture: evidence from Tulare County, California. **American Journal of Agricultural Economics**, v. 70, n. 4, p. 790-802, 1988.

PROVOST, Foster; KOHAVI, R. Glossary of terms. **Journal of Machine Learning**, v. 30, n. 2-3, p. 271-274, 1998.

RAMCHARAN, Amanda et al. Deep learning for image-based cassava disease detection. **Frontiers in plant science**, v. 8, p. 1852, 2017.

RAMSUNDAR, Bharath; ZADEH, Reza Bosagh. **TensorFlow for deep learning: from linear regression to reinforcement learning**. 1. ed. USA: O'Reilly, 2018.

RUMELHART, David E. et al. Learning representations by back-propagating errors. **Cognitive modeling**, v. 5, n. 3, p. 1, 1988.

RUSSEL, Stuart; NORVIG, Peter. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Campus, 2013.

SHORTEN, Connor; KHOSHGOFTAAR, Taghi M. A survey on image data augmentation for deep learning. **Journal of Big Data**, v. 6, n. 1, p. 60, 2019.

SILVA, B. B.; MENDES, FBG; KAGEYAMA, P. Y. **Desenvolvimento econômico, social e ambiental da agricultura familiar pelo conhecimento agroecológico**. São Paulo: Universidade de São Paulo, 2009.

SIMARD, Patrice Y. et al. Best practices for convolutional neural networks applied to visual document analysis. *In: INTERNATIONAL CONFERENCE ON DOCUMENT ANALYSIS AND RECOGNITION (ICDAR). Proceedings [...]* 2003, p. 958–962.

SIMONYAN, Karen; ZISSERMAN, Andrew. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.

SIQUEIRA, Hugo Valadares. **Previsão de series de vazões com redes neurais artificiais e modelos lineares ajustados por algoritmos bio-inspirados**. 2009. Tese (Mestrado em Engenharia Elétrica) – Universidade Estadual de Campinas, Campinas, 2009.

SLADOJEVIC, Srdjan et al. Deep neural networks based recognition of plant diseases by leaf image classification. **Computational Intelligence and Neuroscience**, v. 2016, p. 1-11, 2016.

SRIVASTAVA, Nitish et al. Dropout: a simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, n. 1, p. 1929-1958, 2014.

STAFFORD, John V. Implementing precision agriculture in the 21st century. **Journal of Agricultural Engineering Research**, v. 76, n. 3, p. 267-275, 2000.

SUN, Yu et al. Deep learning for plant identification in natural environment. **Computational Intelligence and Neuroscience**, v. 2017, 2017.

SZEGEDY, Christian et al. Going deeper with convolutions. *In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. Proceedings [...]* 2015. p. 1-9.

SZEGEDY, Christian et al. Rethinking the inception architecture for computer vision. *In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. Proceedings [...]* 2016. p. 2818-2826.

TAJBAKSHI, Nima et al. Convolutional neural networks for medical image analysis: Full training or fine tuning?. **IEEE Transactions on Medical Imaging**, v. 35, n. 5, p. 1299-1312, 2016.

TETTEROO, Jonathan. **Exploring Convolutional Neural Networks on the p-VEX architecture**. 2018. Tese (Master of Science in Computer Engineering), Delft University of Technology, The Netherlands, 2018.

THARWAT, Alaa. Classification assessment methods. **Applied Computing and Informatics**, v. 17, n. 1, p. 168-192, 2018.

TONG, Xin-Yi et al. Learning Transferable Deep Models for Land-Use Classification with High-Resolution Remote Sensing Images. **arXiv preprint arXiv:1807.05713**, 2018.

TOO, Edna Chebet et al. A comparative study of fine-tuning deep learning models for plant disease identification. **Computers and Electronics in Agriculture**, v. 161, p. 272-279, 2019.

VALLE, Eduardo et al. Data, depth, and design: learning reliable models for melanoma screening. **arXiv preprint arXiv:1711.00441**, v. 1, 2019.

VAN DYK, David A.; MENG, Xiao-Li. The art of data augmentation. **Journal of Computational and Graphical Statistics**, v. 10, n. 1, p. 1-50, 2001.

VARGAS, Ana Caroline Gomes; PAES, A. V. C. N.; VASCONCELOS, Cristina Nader. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. *In: XXIX CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES. Proceedings [...]* 2016. p. 1-4.

XIE, Saining et al. Aggregated residual transformations for deep neural networks. *In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. Proceedings [...]* 2017. p. 1492-1500.

WANG, Haiguang et al. Image recognition of plant diseases based on principal component analysis and neural networks. *In: 8TH INTERNATIONAL CONFERENCE ON NATURAL COMPUTATION. Proceedings [...]* 2012. p. 246-251.

WANG, Zhou et al. Image quality assessment: from error visibility to structural similarity. **IEEE transactions on image processing**, v. 13, n. 4, p. 600-612, 2004.

WERBOS, Paul J. et al. Backpropagation through time: what it does and how to do it. **Proceedings of the IEEE**, v. 78, n. 10, p. 1550-1560, 1990.

WINKLER, Stefan. **Digital video quality: vision models and metrics**. England: John Wiley & Sons, 2005.

ZHANG, Shanwen et al. Cucumber leaf disease identification with global pooling dilated convolutional neural network. **Computers and Electronics in Agriculture**, v. 162, p. 422-430, 2019.

ZIMAK, Dav A. **Algorithms and Analysis for Multi-Category Classification**. 2006. Tese (Doctor of Philosophy in Computer Science) – University of Illinois, Urbana-Champaign, 2006.

ZURADA, Jacek M. **Introduction to Artificial Neural Systems**. St. Paul: West, 1992.