

Classificação de estresses bióticos em folhas de café utilizando visão computacional e aprendizagem profunda

Luís Henrique Vieira*

Prof. Dr. Gabriel Marcelino Alves†

6 de Dezembro de 2023

RESUMO

Os estresse bióticos podem ser definidos como dano sofrido pelas plantas, causado por outros organismos vivos. Tendo isso em vista o Brasil como o maior produtor cafeeiro do mundo, lidar com esses estresses na lavoura é de suma importância. Portanto, este trabalho procura identificar os estresses bióticos e o estágio de injúria na folha, para essa tarefa foi proposto a utilização de redes neurais convolucionais (CNN), sendo considerada um modelo de *deep learning* e trabalha com o processamento de imagens. Além disso, foi utilizada uma técnica de aumento de dados para o sistema ser mais robusto e conseguisse lidar com os problemas de generalização. Ademais, os experimentos conduzidos que fizeram a utilização de arquiteturas pré-treinadas apresentaram resultados satisfatórios.

Palavras-chave: Estresses bióticos. Café. Aprendizado de máquina. Redes neurais. Aprendizado profundo.

ABSTRACT

Biotic stress can be defined as damage suffered by plants caused by other living organisms. Bearing in mind that Brazil is the largest coffee producer in the world, dealing with these stresses in the field is of paramount importance. Therefore, this work aims to identify the biotic stresses and the stage of injury in the leaf, for this task it was proposed to use convolutional neural network (CNN), which is considered a *deep learning* model and works with the processing of images. In addition, a data augmentation technique was used to make the system more robust and able to deal with generalization problems. Furthermore, the experiments conducted using pre-trained architectures showed satisfactory results.

Keywords: Biotic stresses. Coffee. Machine learning. Neural network. Deep learning.

*Graduando do Curso Superior em Bacharelado em Ciência da Computação.

† Docente do Curso Superior em Ciência da Computação.

1 INTRODUÇÃO

A agricultura brasileira tem demonstrado um crescimento ininterrupto no quesito produtividade e nos últimos anos os agricultores rurais estão adicionando cada vez mais tecnologias em seus meios de produção, o que contribui para o sucesso da agricultura no Brasil, bem como possibilita oportunidades de crescimento (Contini; Gasques; ALVES; Bastos, 2010). Conforme o Ministério da Agricultura, Pecuária e Abastecimento, o Brasil é o maior produtor e exportador de café no mundo e segundo maior consumidor de sua bebida, logo é inegável que a cultura cafeeira é um grande motor de renda e produto relevante para economia nacional, que gera mais de 8 milhões de empregos no país, além de que em algumas das regiões cafeeiras são oferecidas projetos de inclusão digital. Ademais, ano após ano cada vez mais investimento em certificações, que auxiliam na produção e contato com trabalhadores têm sido colocadas em prática, além de, novas tecnologias que tem se mostrado também capazes de auxiliar os produtores rurais (BLISKA et al., 2009; Agricultura, 2017).

Tendo em vista este panorama descrito, a produção da lavoura de café é acometida por estresses bióticos, tais como: bicho mineiro, ferrugem, mancha de *phoma* e cercosporiose. Importante observar que para evitar doenças uma estratégia é aplicar defensivos agrícolas mesmo a planta não contendo doenças. No entanto, esta é uma prática que tem sido desestimulada a fim de reduzir e otimizar o uso de produtos químicos na lavoura e, portanto, o monitoramento de doenças tem sido priorizado. Porém, o monitoramento é geralmente realizado de maneira manual e depende da experiência e acuidade visual do especialista responsável. Um exemplo desse monitoramento manual consiste em dividir a lavoura em talhões e colher uma quantidade de folhas, por planta, para depois se realizar a inspeção e avaliação do plantio de café (Fernandes; Junior, 2015).

O uso da tecnologia atual, principalmente com auxílio do ramo da inteligência artificial e visão computacional, com os métodos de *deep learning* podem facilitar a detecção desses estresses bióticos. Desta forma, deixando de ser uma tarefa manual, dentre os métodos do *deep learning* temos as Redes Neurais Convolucionais (CNN), que se destacam no reconhecimento de imagens, pois as CNN são capazes de aprender automaticamente as características apropriadas da base dados que lhe foi entregue, compostas por camadas de convolução, agrupamento e totalmente conectada (Voulodimos; Doulamis; Doulamis; Protopapadakis, 2018).

Em seu trabalho, utilizando-se das CNN Esgario, de Castro, Tassis e Krohling (2022) desenvolveram um aplicativo capaz de classificar para seu usuário os estresses presentes e sua severidade na folha do café, os estresses que sua rede CNN foi capaz de detectar são os mesmos quatro citados anteriormente, porém isso só foi possível porque foi criado uma base de dados com várias folhas saudáveis e acometidas por estresses bióticos. Tal base de dados desenvolvida no Espírito Santo por Esgario, Krohling e Ventura (2020) nomeada BRACOL, contém fotos tiradas em ambiente controlado a fim de facilitar o processamento do modelo. Em seu trabalho foi utilizada essa base de dados com adição de algumas imagens disponíveis por outro conjunto, que possui imagens de folhas de diferentes plantas, esse conjunto é chamado PlantVillage, porém existem outras bases de dados como a RoCole do Equador e LilCole de Filipinas ambas possuem imagens com ambientes não controlados e foram desenvolvidas respectivamente por, Parraga-Alava, Cusme, Loor e Santander (2019) e Montalbo e Hernandez (2020).

Neste sentido, o objetivo deste trabalho é identificar e classificar os estresses bióticos

em folhas de café utilizando técnicas de visão computacional e aprendizagem profunda. Como objetivos específicos tem-se: 1) implementar a rede neural CNN; 2) aplicar e avaliar a rede implementada em uma base de imagens de folhas de café existente e, 3) construir uma base adicional de imagens de folhas de café.

Por fim, este trabalho está organizado como segue: após esta introdução é apresentada uma revisão da literatura, seguindo para a metodologia adotada neste trabalho. Posteriormente os resultados bem como as conclusões são expostas.

2 REVISÃO DA LITERATURA

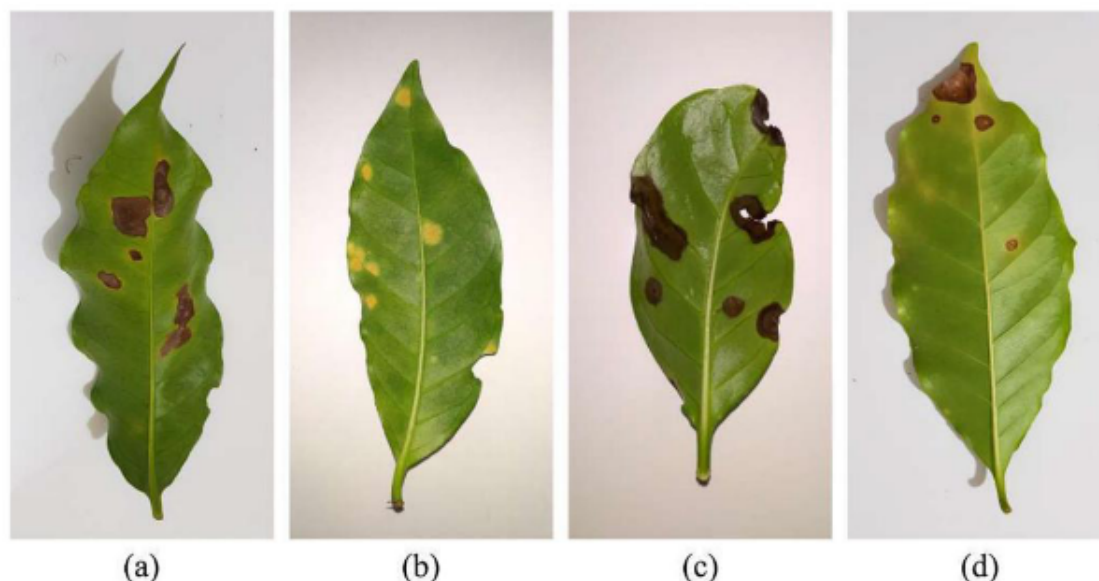
A classificação dos estresses bióticos em folhas de café implica em conhecimento sobre a cultura de café, principalmente da região e país deste trabalho e um campo de estudo da computação, a inteligência artificial (IA), com subcampos de suma importância para o entendimento da pesquisa, sendo eles a *machine learning* e *deep learning*.

2.1 Cultura do café no Brasil e os estresses bióticos

O café surgiu no Brasil no ano de 1727, no estado do Pará, posteriormente a cultura do café foi responsável pela construção de novos caminhos e áreas Brasil afora. O café brasileiro atravessou e continua atravessando dois ciclos econômicos durante todos esses anos, o ciclo de plantio e produção, e o de abandono das lavouras e redução de safras. O parque cafeeiro é baseado na população, pelo número de cafeeiros e pela área dos cafezais, atualmente em um âmbito regional o estado de São Paulo tem 9,5% de área plantada com café, e o líder dessa estatística é o estado de Minas Gerais, que tem como as duas principais espécies plantadas o conilon, ou robusta e a arábica. Ademais, a cultura do café é responsável por gerar renda a diversos pequenos produtores (Matiello; Santinato; Almeida; Garcia, 2016; Agricultura, 2017).

O cafeeiro pode apresentar diversos distúrbios na planta, que por sua vez afetam economicamente a produção, denominados estresses bióticos (Mesquita et al., 2016). Exemplos de estresses bióticos que acometem as folhas do cafeeiro: bicho mineiro, ferrugem, mancha de phoma e cercosporiose. A Figura 1 apresenta imagens de folhas acometidas por esses estresses bióticos.

Figura 1 – Estresses bióticos nas folhas (a) bicho mineiro, (b) ferrugem, (c) mancha de phoma, (d) cercosporiose



Fonte: Elaborado por Esgario, Krohling e Ventura (2020)

Segundo Mesquita et al. (2016), os estresses bióticos mencionados são descritos de maneira breve das seguintes formas.

O bicho mineiro é uma praga que ataca somente o cafeeiro, é percebido nas folhas mediante lesões escurecidas com contorno irregular e tamanho incerto, as lesões começam na face dorsal (superior) da folha e atinge a face ventral (inferior) da folha.

A ferrugem é causada por um fungo, nomeado *Hemileia vastatrix* Berk. et Br., que se alimenta de células vivas da planta, percebido nas folhas através da presença de esporos de cor laranja ou amarela, na face inferior.

A mancha de phoma é causada por um fungo, nomeado *Phoma costaricensis*, que se desenvolve em tecidos vivos, matando as células no processo de dominação é percebido nas folhas por manchas negras ou marrons que provocam deformações na parte principal da folha.

A cercosporiose, também conhecida como mancha de cercóspora causado por um fungo, nomeado *coffeicola* Berk & Cook é percebido nas folhas como lesões pequenas, próximas de serem circulares, com a cor marrom a pardo claro, com a parte central cinza, tendo em sua volta um halo amarelado parecendo um olho.

O trabalho de identificação desses estresses bióticos é geralmente realizado de maneira visual por um profissional, que pode não ser certa (Esgario; Krohling; Ventura, 2020). Desta maneira, o aprendizado de máquina se apresenta junto a visão computacional como uma via para a modernização deste processo.

2.2 Machine Learning

O *machine learning* (ML), ou aprendizagem de máquina, adentrou tanto a vida pessoal, quanto o mundo empresarial de maneira impactante, remodelando as interações na rede de várias formas diferentes (Janiesch; Zschech; Heinrich, 2021a). Tendo em vista

este impacto, o que descreve a causa ML? O ML permite resolver tarefas complexas com programas determinísticos escritos por seres humanos, por exemplo, com um programa determinístico pode-se fazer um robô se movimentar manualmente, ou programamos para o robô aprender a se movimentar com ML, para isso serão necessários exemplos de entradas para que modelo de aprendizado processe e alcance a capacidade de movimentação. Os exemplos representam uma coleção de características, no caso de imagens as características são seus píxeis (Goodfellow; Bengio; Courville, 2016).

Os algoritmos de aprendizado de máquina podem ser definidos de maneira geral como: não supervisionados e supervisionados, devido à experiência durante o processo de aprendizagem. Ambos os algoritmos necessitam de um conjunto de dados, que se refere a uma coleção de exemplos (Janiesch; Zschech; Heinrich, 2021b).

O algoritmo de aprendizagem não supervisionada experimenta um conjunto de dados contendo muitas características e, por conseguinte, aprende atributos relevantes do conteúdo do conjunto fornecido. O algoritmo de aprendizagem supervisionada experimenta um conjunto, que contém características, porém, estas estão associadas a um rótulo, por exemplo, em conjunto referente a estresses bióticos contido em folhas de café, as folhas devem ser anotados com o estresse presente. Logo, um algoritmo de aprendizagem supervisionada pode estudar e classificar qual estresse apresenta a folha. Pode-se entender que a aprendizagem supervisionada possui como um professor que contém a resposta, já no aprendizado não supervisionado envolve uma observação mais severa dos exemplos para distribuição das probabilidades (Goodfellow; Bengio; Courville, 2016).

Existem outros algoritmos de aprendizagem como o semi-supervisionado, o qual, contém muitos exemplos rotulados, mas, outros não. O algoritmo de aprendizagem por reforço, que interagem com o ambiente, havendo um ciclo de *feedback* entre o sistema e suas experiências (Janiesch; Zschech; Heinrich, 2021b).

O ML como dito está muito presente no cotidiano dos seres humanos, para isso ter recorrido cada vez mais as técnicas da classe de aprendizagem profunda (LeCun; Bengio; Hinton, 2015).

2.2.1 Deep Learning

O ML por décadas viu-se limitado no processamento de dados naturais em sua forma bruta, pois conceber um bom extrator de característica é um trabalho difícil, a fim de transformar, como, por exemplo, os valores dos píxeis de imagem em uma representação interna que pudesse ser processada por um modelo de ML.

Os métodos pertencentes da aprendizagem de representações, são métodos de aprendizagem profunda, que nos permite a extração de característica, a partir, de dados brutos e por meio das transformações consiga aprender funções complexas. Em tarefas de classificação, dada uma imagem representada por uma matriz de píxeis, a camada superior de representação amplifica os aspectos importantes e suprime variações, ao fim deve retornar a presença ou ausência de arestas em determinadas localidades, a segunda camada seguinte detecta os padrões de presença ou ausência de arestas, a terceira camada pode reunir esses padrões em combinações maiores pertencentes a um objeto comum, já as próximas camadas detectam objetos como sendo combinações dessas partes. Ademais, o fundamental em relação à aprendizagem profunda, é que o processo de extração de características não é concebido por mãos humanas.

Dito isso, o *deep learning* está realizando avanços em soluções que a inteligência

artificial não encontrava (LeCun; Bengio; Hinton, 2015). Por exemplo, alavancou grande ajuda na resolução de problemas de visão computacional, desde o reconhecimento de objetos ao monitoramento de movimento (Voulodimos; Doulamis; Doulamis; Protopapadakis, 2018).

Todavia, mesmo com todos os avanços realizados, ainda se encontram limitações, o melhor cenário para o aprendizado de máquina é aquele com abundância de dados rotulados para o conjunto, porém, esta é uma tarefa muito difícil e custosa. Logo, uma maneira de contornar parcialmente esta limitação é o uso do aprendizado semi-supervisionado, que relaxa a necessidade de termos dados somente rotulados, mas mesmo a obtenção de dados não rotulados pode ser complexa. A partir disso, a transferência de aprendizagem se concentra transferência de conhecimento entre domínios (Zhuang et al., 2021).

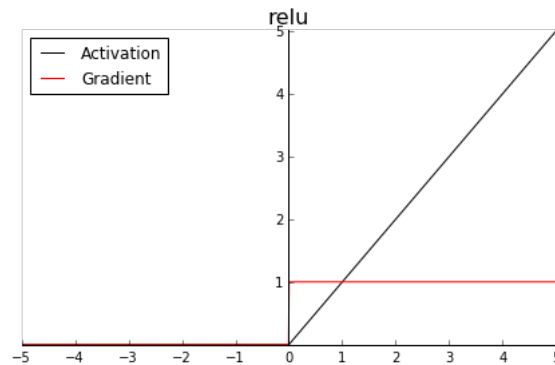
2.2.2 Transfer Learning

O *transfer learning*, ou transferência de aprendizagem, trata-se de uma possível solução que consiste na transferência de conhecimento entre um domínio de origem para o de destino. Fazendo um paralelo ao aprendizado das pessoas, por exemplo, uma pessoa que aprendeu a tocar violino encontra mais facilidade para aprender piano, desde, que generalize sua experiência e devido os dois serem instrumento musicais, ou seja, pertencentes a um mesmo domínio. A transferência de aprendizagem busca impulsionar conhecimento de um domínio relacionado, conhecido como origem, para com que tal domínio tenha seu conhecimento impulsionado para o domínio de destino, visando a diminuição da necessidade de dados rotulados. Contudo, nem sempre o impacto gerado é positivo, isso pode acontecer em casos de domínios não relacionados. Por fim, pode ser dividida em duas categorias, a homogênea e heterogênea, os quais são propostas para lidar com domínios do mesmo espaço de características e domínios de espaços diferentes (Zhuang et al., 2021).

2.2.3 Redes Neurais Convolucionais (CNN)

As redes neurais convolucionais(CNN) têm tido grande sucesso na última década, principalmente nos ramos de reconhecimento de imagem e de voz. Sua arquitetura é composta por camadas de convolução, não linearidade, *pooling* e Camada totalmente conectada. A camada de convolução é a responsável por extrair as características dos dados brutos que lhe foi dado, nessa mesma camada ainda são aplicados filtros que possuem tanta quantidade quanto tamanho, na camada não linearidade é utilizada para ajustar a saída gerada na camada anterior, ou seja, na camada de convolução, atualmente a função utilizada tem sido a ReLU, pois possui uma representação mais difundida, onde o zero do gradiente resulta no ganho de um zero completo (Albawi; Mohammed; Al-Zawi, 2017). A Figura 2 apresenta o gráfico da função ReLU:

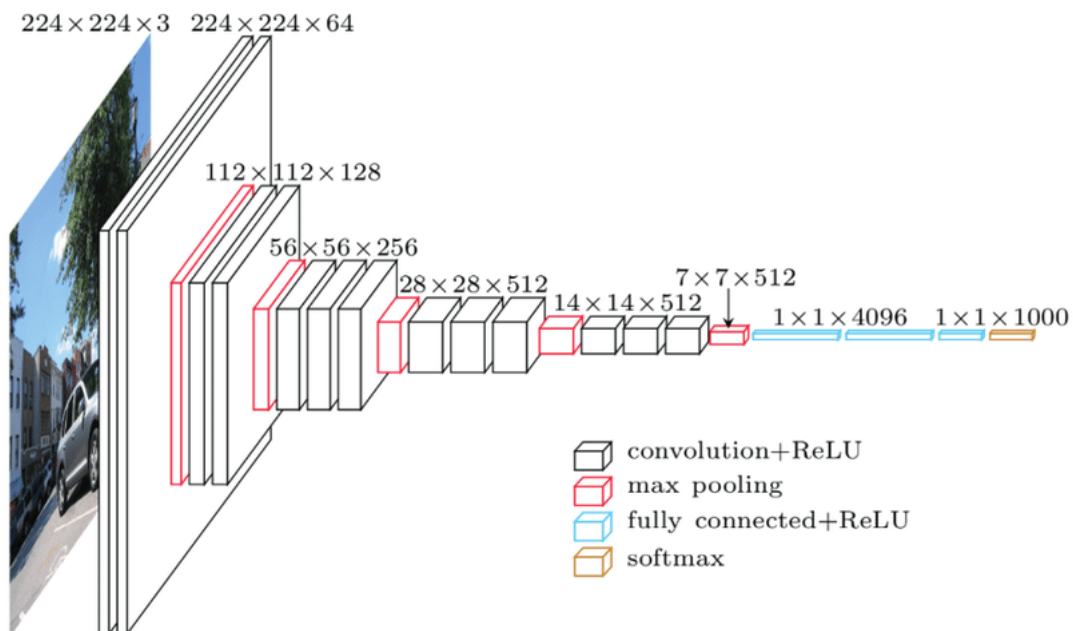
Figura 2 – Gráfico da função ReLU e seu gradiente



Fonte: Elaborado por (Godoy, 2018)

A próxima camada de *pooling*, tem como principal método o *Max-pooling*, esse recorta a imagem em retângulos e retorna apenas o valor máximo dentro dele, o tamanho comumente utilizado para esse método é 2×2 . Ademais, a última camada conhecida como camada totalmente conectada funciona de maneira semelhante a como os neurônios são organizados em um modelo de rede neural tradicional. Desta forma, os nós desta camada estão conectados diretamente a cada nó da camada anterior e da próxima. Por fim, a técnica de *dropout* satisfaz a desvantagem dessa camada que é possuir muitos parâmetros que necessitam de computação complexa em exemplos de treino (Albawi; Mohammed; Al-Zawi, 2017). A Figura 3 apresenta o modelo conceitual da rede VGG16.

Figura 3 – Arquitetura VGG16



Fonte: Elaborado por Bezdán e Džakula (2019).

2.3 Trabalhos Correlatos

As redes CNN já foram utilizadas para identificação de doenças em diferentes plantas, no caso do trabalho de Liu, Zhang, He e Li (2018) foi desenvolvido uma rede neural CNN com base no modelo de rede pré-treinada já existente conhecido como AlexNet, para a identificação de doenças nas folhas de macieiras fazendo a utilização de uma base dados com 13.689 imagens de folhas doentes. O resultado obtido por esta rede se mostrou melhor do que a seu modelo de criação, a rede AlexNet.

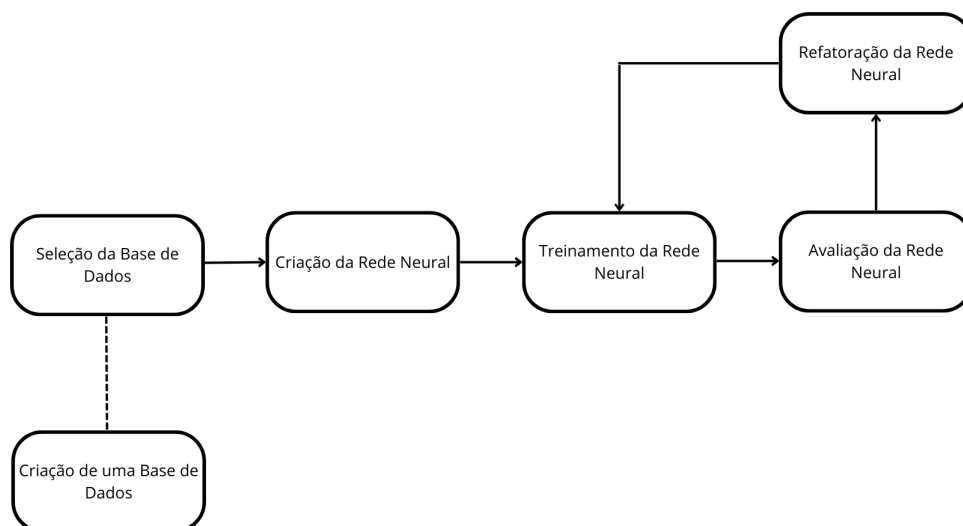
Em outro trabalho, agora tendo como finalidade a detecção de estresses bióticos nas folhas de café, Esgario, Krohling e Ventura (2020) desenvolveram algumas redes neurais CNN seguindo um processo semelhante ao que pode ser visto na próxima seção de metodologia. Para isso foi feita a criação de uma base de dados rotulada para alimentar os modelos, submetida a um pré-processamento e aumento de dados, que posteriormente realiza o uso da técnica de aprendizagem de transferência, junto a outra técnica conhecida como *Multi-Task*, para classificação dos estresses bióticos presentes na folha de café e a sua severidade. Essas são algumas das arquiteturas utilizadas: VGG16, AlexNet e ResNet50.

Em 2022, Esgario, de Castro, Tassis e Krohling (2022) fizeram a utilização dos modelos desenvolvidos no trabalho citado acima, para o desenvolvimento de aplicativo capaz com uma foto da folha do café retornasse aos seus usuários os estresses bióticos contidos e sua severidade.

3 METODOLOGIA

Esta seção demonstra os passos seguidos para a construção das redes neurais citadas neste trabalho, de início temos seleção e criação de uma base dados, em seguida o que foi utilizado para desenvolvimento e execução, estas são as duas etapas iniciais para o desenvolvimento. A Figura 4 apresenta as etapas definidas neste trabalho.

Figura 4 – Metodologia adotada no trabalho.



Fonte: Elaborado pelo autor.

3.1 Criação da base de dados

A base de dados da região do leste paulista, deve ser construída seguindo passos de coleta e rotulagem dos dados, tendo como exemplo a base BRACOL desenvolvida no estado Espírito Santo, Brasil.

Os passos de coleta dos dados: 1) Selecionar os municípios da região do leste paulista, 2) Coletar folhas nas fazendas da região e, 3) Separar as folhas de acordo com estresse biótico contido ou não. Após a coleta, deve executar os passos de rotulagem.

Os passos de rotulagem de dados: 1) Com as folhas separadas, escolher o local adequado para fotografar das imagens, 2) Fotografar as folhas, 3) Rotular as imagens produzidas e, 4) Criar um repositório com as imagens devidamente rotuladas.

Esse é o processo que deve ser seguido na construção da base, esta se propõe fornecer o máximo de imagens dos estresses bióticos encontrados e se possível uma classificação em estágios, além das folhas saudáveis.

3.2 Seleção da base de dados

Para a seleção temos alguns conjuntos de dados já disponíveis, os conjuntos conhecidos são: BRACOL, LilCole, RoCole, para o desenvolvimento deste trabalho a escolha foi pelo conjunto BRACOL construído por Esgario, Krohling e Ventura (2020), que contém folhas de café arábica afetadas pelos seguintes estresses bióticos: bicho mineiro, ferrugem, mancha de phoma e cercospora, além das que possuem estresses contém folhas saudáveis, com um número total de 1747, com folhas coletadas de forma aleatória em Santa Maria no Espírito Santo, Brasil. As fotos foram tiradas da parte abaxial (inferior) das folhas por smartphones de diferentes marcas, posteriormente o conjunto foi dividido em dois diferentes conjuntos, o *leaf* e o *symptom*.

O conjunto *leaf* conta com as imagens coletadas originalmente rotuladas a partir do estresse biótico predominante, já o conjunto *symptom* contém imagens recortadas dos estresses bióticos presentes nas folhas, ou seja, nesse ponto uma folha podia apresentar mais do que um estresse, dessa forma, o conjunto teve seu número de dados aumentado para 2147. Ademais, o conjunto está particionado em treinamento, validação e teste.

3.3 Tecnologias e ambiente

Para a estrutura da rede neural foi definido a linguagem de programação, as bibliotecas desta linguagem e o ambiente do desenvolvimento. Como linguagem de programação foi escolhida Python, por ser uma linguagem de simples aprendizado e possuir bibliotecas muito poderosas em *deep learning*, dentre as principais bibliotecas utilizadas temos tensorflow, que foi desenvolvida para produção de modelos de aprendizado de máquina, keras, que foi projetada para facilitar os processos de construção de um modelo de aprendizado de máquina e scikit-learn, que nos fornece métodos simples para avaliação do modelo (TensorFlow, 2022; Keras..., 2022; scikit-learn..., 2022). Por fim, temos a biblioteca openCV, que foi desenvolvida para o ramo de visão computacional, possuindo ferramentas para o processamento de imagens (Barelli, 2019).

O ambiente escolhido consoante o que estava disponível, logo para o desenvolvimento e execução das redes neurais utilizamos o Google Colab (plano gratuito) e um notebook equipado com uma GPU 940m, ambos contam com uma configuração limitada

principalmente para a etapa de execução da rede, desta forma, fazendo com que a execução seja demorada, por conseguinte, dificultando o desenvolvimento do projeto.

3.4 Pré-processamento e aumento de dados

O conjunto leaf, possui as folhas recortadas com destaques para os estresses bióticos e está com as medidas de (224 x 224 x 3), logo não houve a necessidade de um processo anterior para modificar o tamanho das imagens.

Sendo assim, parte-se para uma técnica de aumento de dados, devido à quantidade de dados limitada, que se possui, mas o importante desse aumento de dados não se trata apenas de suprir uma falta de dados e sim evitar o *overfitting*, que é o sobre-ajuste dos dados, ou seja, o modelo se adequa de mais a um problema, como exemplo se entregarmos a este modelo uma imagem da folha de café sempre da horizontal ou de um mesmo tamanho, pode-se que quando quisermos que ele classifique uma folha na vertical e de outro tamanho será feito um julgamento errado, assim caímos no problema de generalização. Desta forma, para este não ocorrer este problema usa-se seguimos a técnica de *data augmentation*.

O *data augmentation*, que trata as invariâncias a partir de transformações de amostras, podendo ser transformação simples como inversão horizontal, aumento de cor no espaço, cortes aleatórios e até mesmo algo mais complexo como mistura de imagens.(Shorten; Khoshgoftaar, 2019)

3.5 Configuração e treinamento do modelo

Com o ambiente a base dados BRACOL e as arquiteturas pré-treinadas, sendo: VGG16, RESNET50V2 e MOBILENETV2, o uso do método de aprendizagem de transferência, com pesos carregados a partir da base de dados imageNet, que possui milhares de imagens disponibilizadas gratuitamente para pesquisa (ImageNet, 2021). O total de seis redes neurais, duas para cada arquitetura, sendo uma com todas as camadas e outra com a tentativa de ajuste fino, ou seja, congelando algumas camadas.

A estrutura com a leitura da base de dados, instâncias do modelo inicializadas, juntamente as duas camadas totalmente conectadas, seguidas camada de saída que utiliza o mesmo parâmetro para todas as diferentes arquiteturas, o softmax, utilizada em problemas com múltiplas saídas (probabilidades para cada estresse biótico).

Todos os ensaios utilizam os mesmos hiper parâmetros dispostos na Tabela 1 e com as proporções de 70-15-15, respectivamente para treinamento, validação e teste.

Tabela 1 – Hiper parâmetros de treinamento das redes neurais

Hipe parâmetro	Valor
Otimizador	Adam
Função de perda	Entropia Cruzada
Tamanho de batch	32
Épocas	100
Taxa de aprendizagem	0,0001

Fonte: Elaborada pelo autor

3.6 Avaliação do modelo

A avaliação das redes neurais foi feita com a utilização da biblioteca scikit-learn, na etapa de treinamento foi utilizado o método ModelCheckpoint permitindo que salvemos os pesos de treinamento quando encontramos a menor perda no conjunto de validação. A partir dos modelos salvos foi realizada a avaliação, desta forma medindo Acurácia, Precisão e *Recall*.

- **Acurácia:** Utilizado para calcular o número total de previsões corretas.

$$Acurácia = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

- **Precisão:** Utilizado para calcular a fração de padrões positivos, corretamente previstos entre todos os padrões previstos na classe positiva.

$$Precisão = \frac{TP}{TP + FP} \quad (2)$$

- **Recall:** Utilizado para calcular a fração de positivos reais (TP), que o foi identificado corretamente.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

4 RESULTADOS E DISCUSSÃO

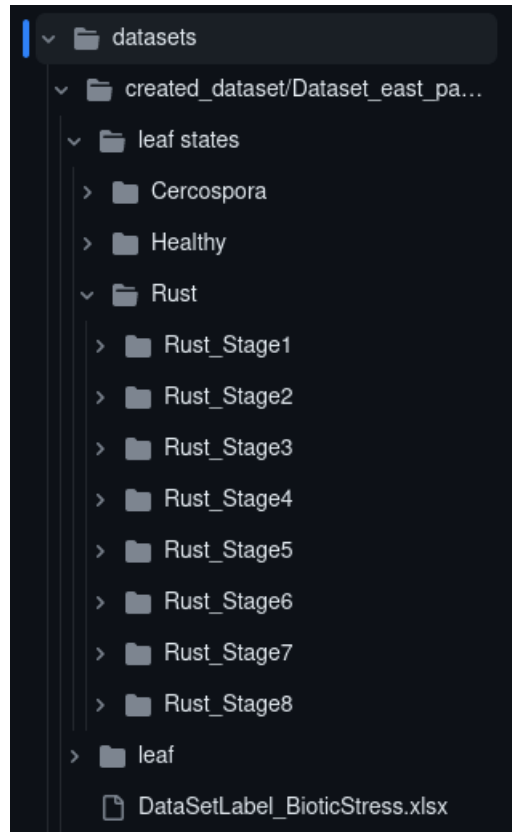
4.1 Base de Dado Leste Paulista

A base de foi construída seguindo processo citado na metodologia, sendo assim, satisfazendo os passo de coleta de dados: 1) Divinolândia, São Sebastião da Grama, Caconde e Santo Antônio do Jardim, somaram quatro municípios da região leste paulista, 2) As folhas foram coletadas de forma aleatória para aumentar diversidade da base e, 3) Dividas em sacolas com sua devido estresse para facilitar para o processo de rotulagem, tudo apoiado de um servidor do Instituto Federal de Ciências e Tecnologia de São Paulo, campus São João da Boa Vista. Desta forma, finalizou-se a etapa de coleta e seguiu-se para rotulagem.

Assim, satisfazendo o processo de rotulagem de dados: 1) Ambiente escolhido era parcialmente controlado, dispostos com um fundo branco junto a uma iluminação mínima, 2) A captura das folhas realizada com um smartphone, Motorola (Moto G9 Play) seguindo o lado abaxial das folhas, 3) para a rotulagem as imagens foram dividas em pastas referentes ao estresse contido na folha, a nomenclatura esta em inglês e segue nome do estresse, número iniciado em 0), caso contenha estágio segue a nomenclatura anterior, porém o número passa para depois da palavra *stage* junto ao nível do estágio do estresse (iniciado em 1). Ademais, uma planilha representa os estresses bióticos na folha, com 0 e 1, que respectivamente se traduz em não contém e contém o estresse, sendo as folhas dispostas pelas linhas e os estresses pelas colunas, para a associação com a planilha foi usado outra pasta nomeada *leaf*, onde a imagens levam o nome da pasta junto ao número (iniciado em 0), 4) A rotulagem trouxe os números finais da base, com um total de 264 imagens, sendo as folhas, 51 saudáveis, 20 com cercosporiose, 193 com ferrugem, destas dispostas em 8 estágios, contendo respectivamente: 20, 45, 14, 9, 13, 27, 24 e 41, posteriormente

disponibilizada em um projeto do GitHub. Veja na figura 5 como ficou a estruturas das pastas da base:

Figura 5 – Estrutura da base de dados



Fonte: Elaborado pelo autor

A base de imagens com 264 folhas coletadas não foi utilizada neste trabalho, pois o número inicialmente previsto de 1000 folhas, não foi alcançado, devido impossibilidade de mais visitas aos municípios, logo impossibilitando a coleta de folhas para inserção na base. Veja na Figura 6 apresenta algumas das imagens presentes na base de dados.

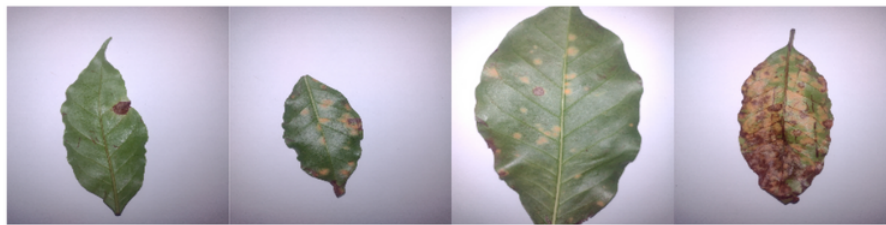
Figura 6 – Imagens contidas na base de dados criada



(a) Saudáveis



(b) Cercosporiose

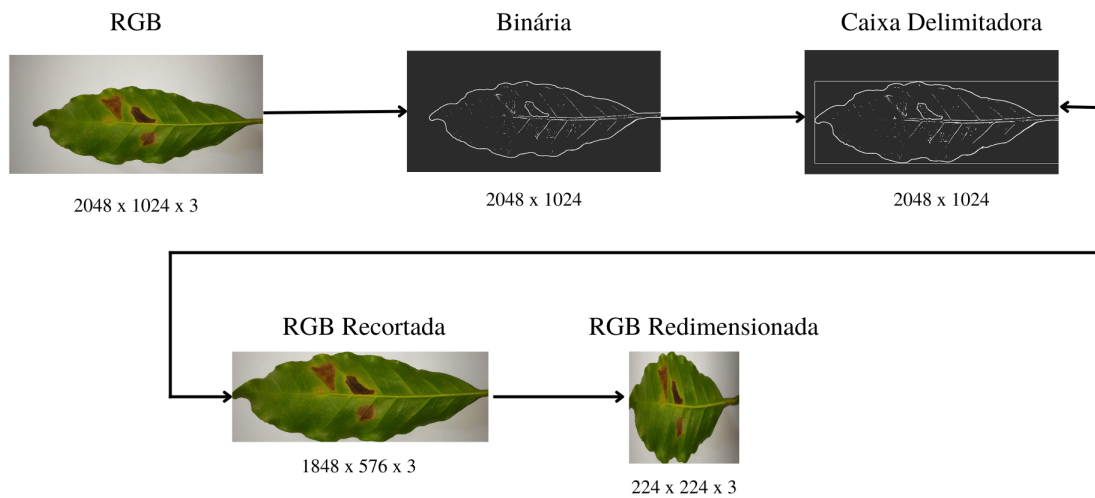


(c) Ferrugem

Fonte: Elaborado pelo autor

Por fim, a maneira como estão as imagens da base há necessidade de um pré-processamento, para as arquiteturas CNN que estão sendo utilizadas nesse trabalho possuem requisitos de entrada, sendo assim com a utilização dos métodos disponíveis com openCV pode-se redimensionar a imagem a $224 \times 224 \times 3$ (largura, altura e número de canais) ou um valor aceito na arquitetura, um algoritmo de o pré-processamento se faz necessário, este faz o recorte da folha conforme a caixa delimitadora criada ao redor da folha e por fim seu tamanho é modificado, o processo pode ser vista na Figura 7.

Figura 7 – Pré-processamento para redimensionamento da imagem



Fonte: Elaborado pelo autor

4.2 Resolução das Redes Neurais

A partir dos experimentos realizados com o conjunto de dados *Symptom* da base BRACOL, a fim de classificar os estresses bióticos presentes nas folhas, com uso da rede CNN e de redes neurais pré-treinadas sendo: VGG16, RESNET50V2 e MOBILENETV2, obteve-se um total de três tabelas dispondo os resultados, de acordo, com a rede construída, traduzidos nas colunas de Acurácia, Precisão e Recall. A Tabela 2 apresenta do modelo CNN.

Tabela 2 – Resultados do modelo CNN

Acurácia	Precisão	Recall
84,8%	85,3%	84,8%

Fonte: Elaborada pelo autor

Os resultados conquistados pela rede CNN são muito satisfatório, pois, se tratando de uma rede construída e testada do zero, sem uso de uma arquitetura pré-treinada, ainda foi possível chegar a ótimos valores, todo via, na Tabela 3 que apresenta os resultados das redes pré-treinadas, pose-se observar um ganho nos percentuais das métricas de avaliação, veja:

Tabela 3 – Resultados das arquiteturas

Arquitetura	Acurácia	Precisão	Recall
RESNET50V2	93,4%	93,4%	93,4%
VGG16	91,9%	92,1%	91,9%
MOBILENETV2	89,9%	90,8%	89,9%

Fonte: Elaborada pelo autor

Analisando os resultados das arquiteturas todos são satisfatórios dada a construção e execução, a RESNET50V2 se destaca com as melhores porcentagens, a arquitetura possui número intermediário e parâmetros em relação as outras e terminou o treino no menor tempo, MOBILENETV2 se encaixa em sua posição, pois é justamente a rede com menor número de parâmetros devido a sua natureza ser voltada para dispositivos móveis onde o poder de processamento é menor, por fim a VGG16 que possui o maior número de parâmetros, porém espera-se que sua porcentagem de acerto seja menor que a da RESNET50V2, tendo em vistas os valores das documentações do Keras.

Os resultados acima vem da utilização completa das arquiteturas, ou seja, todas as camadas das arquiteturas pré-treinadas estão sendo utilizadas, porém, no processo de construção alguns testes congelando camadas, foram feitos na busca de um ajuste fino. Diante disso, a tabela 4 os dispõe os melhores resultados.

Tabela 4 – Resultados das arquiteturas com ajuste fino

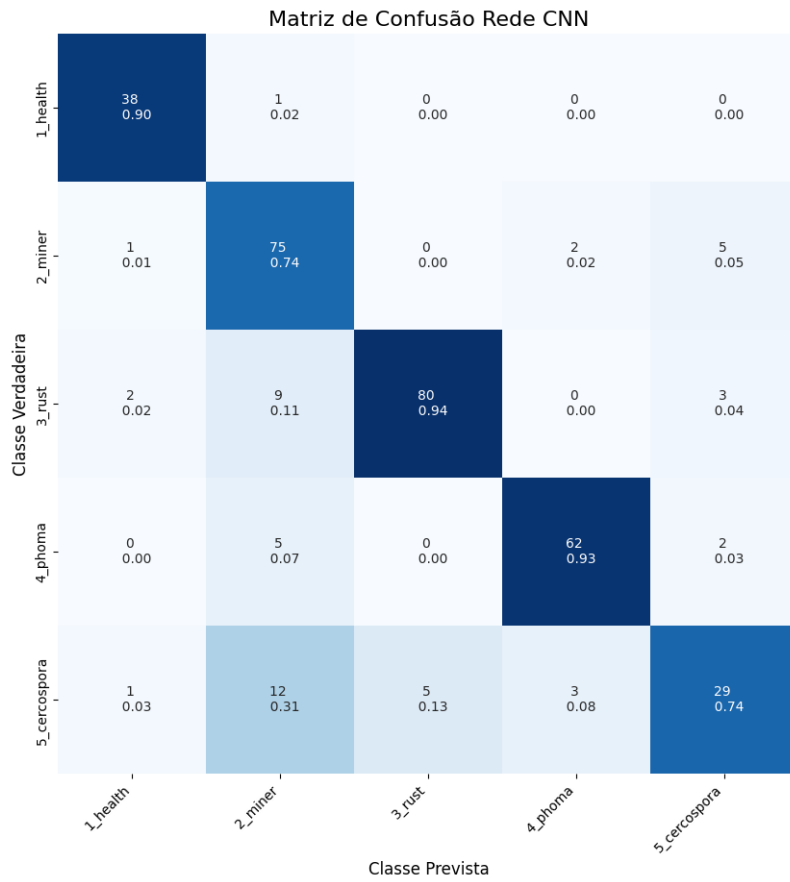
Arquitetura	Acurácia	Precisão	Recall	Número de camadas
RESNET50V2	72,8%	74,7%	72,8%	53/105
VGG16	90,7%	90,8%	90,7%	2/16
MOBILENETV2	77,6%	78,6%	77,6%	3/103

Fonte: Elaborada pelo autor

A interpretação dos resultados demonstra que na tentativa do ajuste fino todas as redes perderam desempenho. Entretanto, a VGG16 nos trouxe valores próximos aos que foram vistos na utilização de todas as camadas, fazendo jus a sua arquitetura robusta, já ambas as restantes se esperava uma queda maior devido à queda no número de camadas em relação ao total.

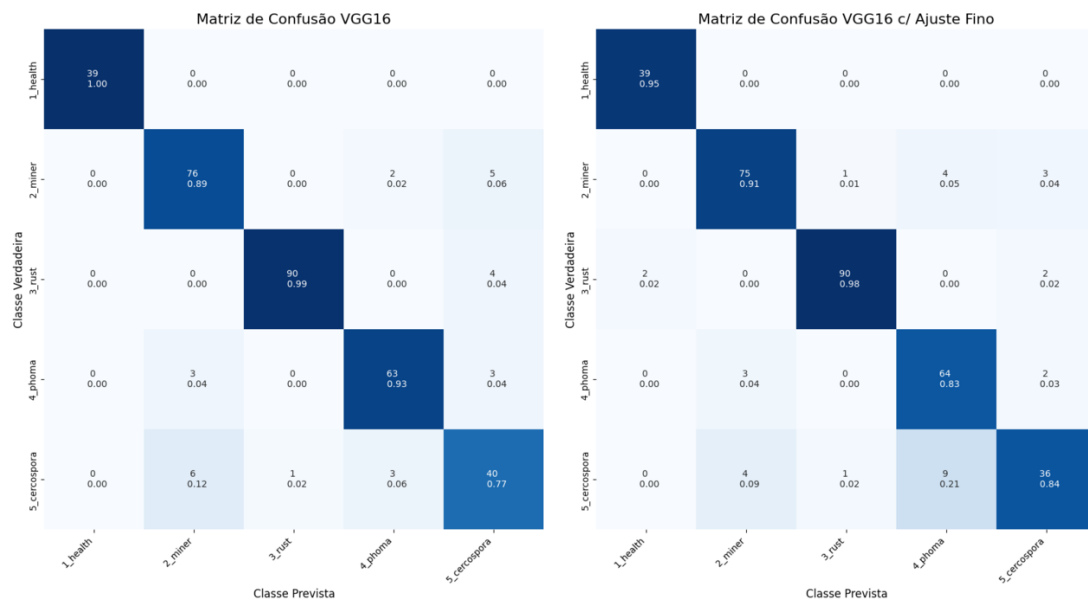
Ademais, as matrizes de confusão dispostas da Figura 8 e Figura 11 apresentam o resultado dos modelos para cada classe, o número inteiro visto nos quadrados da matriz representa o número de imagens predita como pertencente àquela classe e o número flutuante logo abaixo representa sua precisão.

Figura 8 – Matrizes de confusão rede CNN



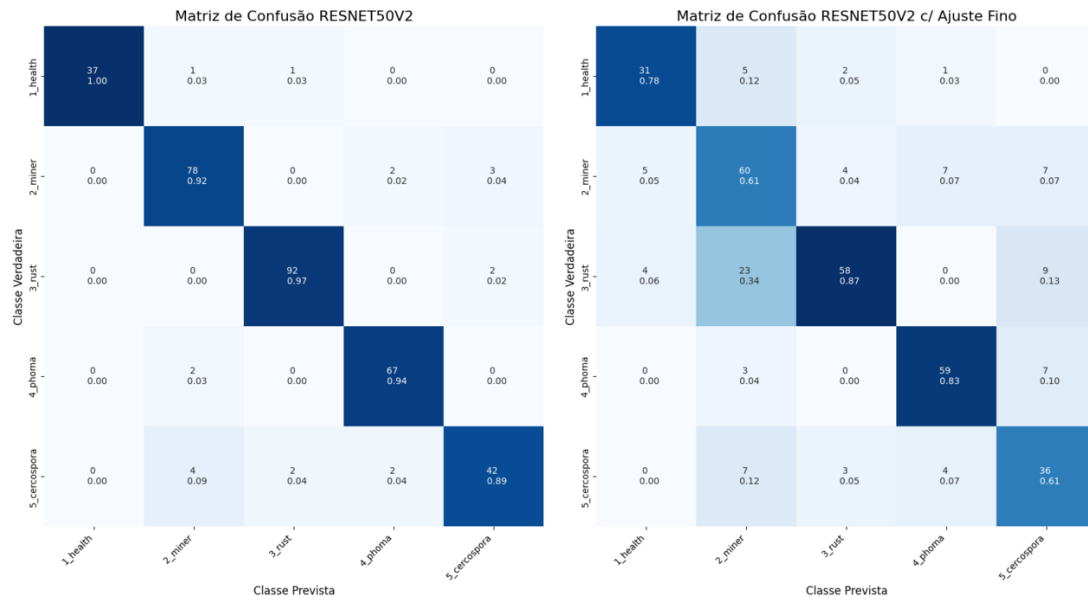
Fonte: Elaborado pelo autor

Figura 9 – Matrizes de confusão VGG16



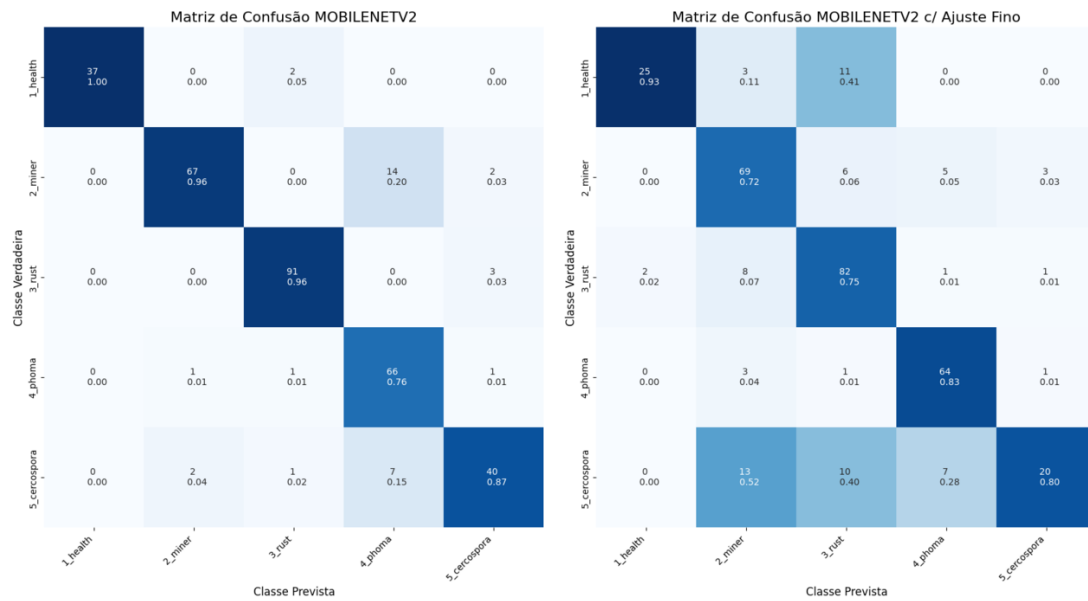
Fonte: Elaborado pelo autor

Figura 10 – Matrizes de confusão RESNET50V2



Fonte: Elaborado pelo autor

Figura 11 – Matrizes de confusão MOBILENETV2



Fonte: Elaborado pelo autor

Ao visualizar as matrizes podemos concluir, que as redes neurais pré-treinadas, que fizeram uso de todas as camadas das arquiteturas, obtiveram resultados bem definidos. Todavia, com a tentativa de um ajuste fino, a VGG16 foi a que trouxe o melhor resultado, até mesmo próxima às redes construídas sem ajuste fino.

5 CONCLUSÕES

Este trabalho alcançou seu objetivo ao identificar e classificar os estresses bióticos (bicho mineiro, ferrugem, mancha de *phoma*, cercosporiose) por meio de aprendizagem profunda.

O estudo abordou diferentes técnicas de *deep learning*, a fim de construir redes neurais capazes de classificar diferentes estresses bióticos presentes nas folhas de café e, para isso, foi utilizada uma base de dados de imagens disponíveis na Internet. Entre as arquiteturas avaliadas, a ResNet50V2 apresentou os melhores resultados para o conjunto de dados considerado. Observou-se, ainda, que as redes neurais que fizeram uso de todas as camadas de suas respectivas arquiteturas obtiveram resultados bem definidos, pois a variação foi mínima. Com as arquiteturas estudadas e implementadas pode-se concluir que a utilização das técnicas de transferência de aprendizagem melhorou os resultados. Além disso, durante o desenvolvimento deste trabalho foi iniciada a construção de uma base de imagens de folhas de café de plantações da região leste paulista.

Como trabalhos futuros, pretende-se finalizar a construção da base de imagens de folhas de café, bem como aplicar as redes neurais para avaliação dos estresses bióticos na região leste paulista.

REFERÊNCIAS

- Agricultura, P. e. A. Ministério da. *Café no Brasil*. 2017. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/politica-agricola/cafe/cafeicultura-brasileira>>. 2, 3
- Albawi, S.; Mohammed, T. A.; Al-Zawi, S. Understanding of a convolutional neural network. In: *2017 International Conference on Engineering and Technology (ICET)*. [S.l.: s.n.], 2017. p. 1–6. 6, 7
- Barelli, F. *Introdução À Visão Computacional - Uma abordagem prática com Python e OpenCV*. 1. ed. [S.l.]: Casa do código, 2019. 9
- Bezdan, T.; Džakula, N. B. Convolutional Neural Network Layers and Architectures. In: *Proceedings of the International Scientific Conference - Sinteza 2019*. Novi Sad, Serbia: Singidunum University, 2019. p. 445–451. ISBN 9788679127037. Disponível em: <<http://portal.sinteza.singidunum.ac.rs/paper/700>>. 7
- BLISKA, F. d. M.; Mourão, E. A. B.; Júnior, P. C. A.; Vegro, C. L. R.; Pereira, S. P.; Giomo, G. S. Dinâmica fitotécnica e socioeconômica da cafeicultura brasileira. *Embrapa Café-Artigo em periódico indexado (ALICE)*, In.: Informações Econômicas, São Paulo, v. 39, n. 1, jan. 2009., 2009. 2
- Contini, E.; Gasques, J. G.; ALVES, E. d. A.; Bastos, E. T. Dinamismo da agricultura brasileira. *Área de Informação da Sede-Artigo em periódico indexado (ALICE)*, Revista de Política Agrícola, Brasília, DF v. 19, Edição Especial, p. 42-64 ... , 2010. 2
- Esgario, J. G.; de Castro, P. B.; Tassis, L. M.; Krohling, R. A. An app to assist farmers in the identification of diseases and pests of coffee leaves using deep learning. *Information*

- Processing in Agriculture*, v. 9, n. 1, p. 38–47, 2022. ISSN 2214-3173. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2214317321000044>>. 2, 8
- Esgario, J. G.; Krohling, R. A.; Ventura, J. A. Deep learning for classification and severity estimation of coffee leaf biotic stress. *Computers and Electronics in Agriculture*, v. 169, p. 105162, 2020. ISSN 0168-1699. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0168169919313225>>. 2, 4, 8, 9
- Fernandes, C. d. F.; Junior, J. R. V. Doenças do cafeeiro. 2015. Disponível em: <<http://www.infoteca.cnptia.embrapa.br/handle/doc/1040720>>. 2
- Godoy, D. *Hiper-parâmetros em Ação! Parte I - Funções de Ativação*. 2018. Disponível em: <<https://medium.com/ensina-ai/hiper-par%C3%A2metros-em-a%C3%A7%C3%A3o-parte-i-fun%C3%A7%C3%B5es-de-ativa%C3%A7%C3%A3o-cd37600e46c6>>. 7
- Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <<http://www.deeplearningbook.org>>. 5
- ImageNet. 2021. Disponível em: <<https://www.image-net.org/>>. 10
- Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electronic Markets*, v. 31, n. 3, p. 685–695, Sep 2021. ISSN 1422-8890. Disponível em: <<https://doi.org/10.1007/s12525-021-00475-2>>. 4
- Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electronic Markets*, v. 31, n. 3, p. 685–695, set. 2021. ISSN 1019-6781, 1422-8890. Disponível em: <<https://link.springer.com/10.1007/s12525-021-00475-2>>. 5
- Keras: the Python deep learning API. 2022. Disponível em: <<https://keras.io/>>. 9
- LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *nature*, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015. 5, 6
- Liu, B.; Zhang, Y.; He, D.; Li, Y. Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry*, v. 10, n. 1, 2018. ISSN 2073-8994. Disponível em: <<https://www.mdpi.com/2073-8994/10/1/11>>. 8
- Matiello; Santinato; Almeida; Garcia. *Cultura de Café no Brasil: Manual de recomendações*. 2015. ed. São Paulo: Futurama editora, 2016. 3
- Mesquita, C.; Rezende, J.; Carvalho, J.; Júnior, M.; Moraes, N.; Dias, P.; Carvalho, R.; Araújo, W. *MANUAL DO CAFÉ Distúrbios fisiológicos, pragas e doenças do cafeeiro*. [s.n.], 2016. (Ciências Agrárias). Disponível em: <http://www.sapc.embrapa.br/arquivos/consorcio/publicacoes_tecnicas/livro_disturbios_fisiologicos_pragas_doen%C3%A7as.pdf>. 3, 4
- Montalbo, F. J.; Hernandez, A. Classifying barako coffee leaf diseases using deep convolutional models. *International Journal of Advances in Intelligent Informatics*, v. 6, n. 2, p. 197–209, 2020. ISSN 2548-3161. Disponível em: <<http://ijain.org/index.php/IJAIN/article/view/495>>. 2
- Parraga-Alava, J.; Cusme, K.; Loor, A.; Santander, E. Rocol: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition. *Data in Brief*, v. 25, p. 104414, 2019. ISSN 2352-3409. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352340919307693>>. 2

scikit-learn: machine learning in Python — scikit-learn 1.1.1 documentation. 2022. Disponível em: <<https://scikit-learn.org/stable/>>. 9

Shorten, C.; Khoshgoftaar, T. M. A survey on image data augmentation for deep learning. *Journal of big data*, SpringerOpen, v. 6, n. 1, p. 1–48, 2019. 10

TensorFlow. 2022. Disponível em: <<https://www.tensorflow.org/?hl=pt-br>>. 9

Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, v. 2018, p. 1–13, 2018. ISSN 1687-5265, 1687-5273. Disponível em: <<https://www.hindawi.com/journals/cin/2018/7068349/>>. 2, 6

Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, v. 109, n. 1, p. 43–76, 2021. 6