

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Thiago Vieira Machado

**Aplicação de Técnicas de Visão Computacional
para Identificação de Ferrugem em Folhas de
Café**

Uberlândia, Brasil

2023

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Thiago Vieira Machado

**Aplicação de Técnicas de Visão Computacional para
Identificação de Ferrugem em Folhas de Café**

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, como requisito exigido parcial à obtenção do grau de Bacharel em Sistemas de Informação.

Orientador: Prof. Dr. Bruno Augusto Nassif Travençolo

Coorientadora: Profa. Me. Larissa Ferreira Rodrigues Moreira

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Sistemas de Informação

Uberlândia, Brasil

2023

Resumo

O café é uma bebida presente na vida de muitas pessoas ao redor do mundo, além de ser de grande importância para a economia de diversos países. A ferrugem dos cafezais é uma doença muito grave que ataca lavouras em todo o mundo e identificá-la de forma rápida e assertiva pode auxiliar no seu controle. O presente trabalho tem como objetivo avaliar o desempenho de seis Redes Neurais Convolucionais (*CNN*) no contexto da detecção de ferrugem em folhas de café, utilizando-se de um banco de imagens disponibilizado publicamente. Todas as *CNNs* foram avaliadas para classificação binária, isto é, para indicar se a folha possui ou não ferrugem. A rede de melhor desempenho na classificação binária foi usada para classificação multiclasse, que inclui quatro diferentes níveis (estágios) de ferrugem. Para o treinamento e validação dos modelos foram aplicadas técnicas de aumento de dados e a validação cruzada *k-fold*. Entre as redes avaliadas, a *ResNet* obteve os melhores resultados, alcançando uma acurácia de 95,19% na classificação binária e 78,03% na classificação multiclasse. O trabalho contribui para a aplicação da Visão Computacional como uma ferramenta auxiliar para os cafeicultores, permitindo a detecção precoce da presença de ferrugem nas folhas de café e auxiliando na tomada de decisões relacionadas ao manejo das lavouras.

Palavras-chave: *Inteligência Artificial, Deep Learning, CNN, folhas de café, classificação de imagens.*

Lista de ilustrações

Figura 1 – Ramo de folhas de café com sintomas de ferrugem.	8
Figura 2 – Plantas de café severamente afetadas pela ferrugem.	9
Figura 3 – Estrutura da <i>CNN AlexNet</i>	11
Figura 4 – Estrutura da <i>CNN DenseNet</i>	12
Figura 5 – Estrutura da <i>CNN ResNet</i>	13
Figura 6 – Estrutura da <i>CNN SqueezeNet</i>	13
Figura 7 – Visão geral da validação cruzada <i>k-fold</i>	14
Figura 8 – Visão geral de uma matriz de confusão.	15
Figura 9 – Exemplo de imagens do conjunto <i>RoCoLe</i>	19
Figura 10 – Método Proposto.	20
Figura 11 – Gráfico de perda (<i>loss</i>) e acurácia durante o treinamento e validação do <i>Fold 0</i> da arquitetura <i>ResNet</i>	22
Figura 12 – Comparação da acurácia das redes testadas.	23
Figura 13 – Variação de desempenho das <i>CNNs</i> ao avaliar folhas saudáveis.	25
Figura 14 – Variação de desempenho das <i>CNNs</i> ao avaliar folhas doentes.	25
Figura 15 – Gráfico de perda (<i>loss</i>) e acurácia durante o treinamento e validação multiclasse do <i>Fold 0</i> da arquitetura <i>ResNet</i>	27

Lista de tabelas

Tabela 1	–	Comparação entre os trabalhos relacionados e o presente trabalho. . . .	18
Tabela 2	–	Avaliação de desempenho da arquitetura <i>Alexnet</i>	23
Tabela 3	–	Avaliação de desempenho da arquitetura <i>Densenet</i>	24
Tabela 4	–	Avaliação de desempenho da arquitetura <i>Inception</i>	24
Tabela 5	–	Avaliação de desempenho da arquitetura <i>ResNet</i>	24
Tabela 6	–	Avaliação de desempenho da arquitetura <i>Squeezenet</i>	24
Tabela 7	–	Avaliação de desempenho da arquitetura <i>VGG</i>	24
Tabela 8	–	Valores das matrizes de confusão de cada modelo <i>CNN</i>	26
Tabela 9	–	Avaliação de desempenho da arquitetura <i>ResNet</i> na classificação multiclasse.	27
Tabela 10	–	Matriz de confusão do <i>Fold 0</i> da classificação multiclasse da arquitetura <i>ResNet</i>	28

Sumário

1	INTRODUÇÃO	6
1.1	Objetivo	7
1.2	Justificativa	7
2	REFERENCIAL TEÓRICO	8
2.1	Conceitos Fundamentais	8
2.1.1	Ferrugem do cafeeiro	8
2.1.2	Visão Computacional	9
2.1.3	<i>Deep Learning</i>	10
2.1.3.1	Redes Neurais Convolucionais	10
2.1.4	Protocolo de Avaliação	14
2.1.5	Validação Cruzada <i>k-fold</i>	14
2.1.6	Matriz de Confusão	14
2.1.7	Métricas de Avaliação	15
2.2	Trabalhos Relacionados	16
3	METODOLOGIA	19
4	RESULTADOS E DISCUSSÕES	21
4.1	Ambiente de Testes	21
4.2	Classificação Binária	21
4.3	Seleção do melhor modelo para Classificação Multiclasse	26
5	CONCLUSÃO	29
	REFERÊNCIAS	30

1 Introdução

O café é uma bebida presente na vida de muitas pessoas ao redor do mundo. Estima-se que sejam consumidas mais de três bilhões de xícaras todos os dias. Mas o café não é importante apenas no cotidiano da população, o cultivo dele também é de grande importância para a economia de mais de 50 países (SILVA et al., 2022).

Os países produtores de café exportam a maior parte do que é produzido, movimentando, somente com as exportações, cerca de 20 bilhões de dólares anualmente, mas se levarmos em consideração todo o setor cafeeiro, as receitas ultrapassam os 220 bilhões de dólares (INTERNATIONAL COFFEE ORGANIZATION, 2019).

Como os valores movimentados com o café são consideráveis, as doenças que atingem a cultura também geram prejuízos consideráveis. Entre as doenças que podem afetar os cafeeiros está a ferrugem da folha, causada pelo fungo *Hemileia Vastatrix*, que provoca a queda precoce das folhas, a seca dos ramos e diminuição considerável de produção da planta (SANTANA et al., 2018).

A ferrugem é uma doença muito severa e que pode afetar a produção de uma lavoura inteira. Seu diagnóstico é feito por meio de um especialista, que faz a análise das folhas para identificar a doença. Porém, em vários lugares do mundo o cultivo do café é feito por pequenos produtores que, muitas vezes, não tem condições de arcar com os custos de um especialista (YEBASSE et al., 2021). Desta forma, a tecnologia pode ser empregada para a identificação das doenças nos cafezais, auxiliando tanto os grande produtores quanto na subsistência dos pequenos.

Entre as tecnologias existentes, o Aprendizado Profundo (do inglês *Deep Learning*), uma subcategoria do aprendizado de máquina, vem ganhando bastante espaço. Ele possibilita que, por meio da utilização de redes neurais, algoritmos sejam treinados para identificação de fala, Visão Computacional e processamento de linguagem natural (GUO et al., 2016).

Dentro do Aprendizado Profundo, as Redes Neurais Convolucionais (do inglês *Convolutional Neural Networks - CNNs*) são, provavelmente, o modelo mais conhecido e usado para resolver tarefas de Visão Computacional e classificação de imagem (PONTI et al., 2017). Sendo que as técnicas de Visão Computacional podem ser uma alternativa de grande valia para auxiliar no processo de identificação da ferrugem do café, visto que ela permite que computadores sejam treinados para compreender e analisar imagens, permitindo um diagnóstico automatizado e preciso (YEBASSE et al., 2021; ONGSULEE, 2017).

Tendo em vista que as *CNNs* têm sido cada vez mais empregadas para reconhecimento visual e que tem atingido níveis altos de eficiência e precisão para detecção de doenças, este trabalho pretende comparar alguns modelos de *CNNs* para ver quais possuem melhor desempenho quanto as métricas de precisão, acurácia, *recall* e *F1-Score* na detecção de ferrugem em folhas de café.

1.1 Objetivo

O presente trabalho tem como objetivo analisar imagens de folhas de café utilizando técnicas de Visão Computacional e aprendizado de máquina para identificar ferrugem nas folhas do cafezal.

Para atingir tal objetivo, o trabalho conterà os seguintes objetivos específicos:

- Comparar diferentes modelos de *CNNs* em termos de acurácia, precisão, *recall* e *F1-Score*.
- Avaliar o desempenho das *CNNs* utilizando validação cruzada *k-fold* para os conjuntos de treinamento e validação.
- Realizar avaliação binária, isto é, para indicar se a folha possui ou não ferrugem.
- Realizar a avaliação multiclasse, que inclui folhas saudáveis e quatro diferentes níveis (estágios) de ferrugem.

1.2 Justificativa

O café é um produto consumido em praticamente todo o mundo e movimentam quantias consideráveis de dinheiro na economia global ([INTERNATIONAL COFFEE ORGANIZATION, 2019](#)). As doenças que atingem esta cultura também possuem potencial de trazerem enormes prejuízos aos produtores, uma vez que as lavouras estão presentes em várias regiões. A ferrugem é uma doença que atinge os cafezais, prejudicando-os severamente e causando muitas perdas aos produtores.

Com a aplicação de técnicas de Visão Computacional, pretende-se auxiliar na identificação da doença presente nas folhas, por meio de imagens captadas nas lavouras. Com isso, o diagnóstico da doença poderá ser feito de forma mais padronizada e ágil, auxiliando tanto o produtor quanto os especialistas em cafezais. Além de auxiliar os produtores de café, o trabalho também pretende auxiliar a área acadêmica, uma vez que a base de dados que foi utilizada, conhecida como *RoCoLe*, foi pouco explorada para tarefas relacionadas a área de classificação.

2 Referencial Teórico

Neste capítulo são abordados os principais conceitos técnicos e teóricos que serviram como base para o desenvolvimento do trabalho. São apresentados os aspectos da ferrugem da folha do café, bem como as técnicas computacionais baseadas em Visão Computacional para detecção da doença.

2.1 Conceitos Fundamentais

Nesta seção são descritos os principais conceitos que formam a base para a melhor compreensão do tema e do trabalho.

2.1.1 Ferrugem do cafeeiro

A ferrugem do cafeeiro é a doença mais importante para a cultura do café e atinge lavouras do mundo inteiro. Ela é causada pelo fungo *Hemileia Vastatrix* e provoca sérios danos às plantas, especialmente ao cafeeiro da espécie arábica (SILVA et al., 2022).

Os primeiros sintomas da doença se manifestam por pequenas lesões amareladas na face inferior da folha. Conforme elas crescem passam a adquirir sua distinta cor laranja e é possível ver manchas marrons na parte superior da folha. Durante o último estágio da doença as lesões se tornam necróticas (AVELINO et al., 2015). Na Figura 1 é possível ver um ramo de folhas de café infectado pela ferrugem do cafeeiro e os sintomas da doença em vários dos níveis descritos.



Figura 1 – Ramo de folhas de café com sintomas de ferrugem.
Fonte: Avelino et al. (2018).

Seus principais efeitos são a desfolha severa, como pode ser visto na Figura 2, que reduz a capacidade fotossintética da planta e, conseqüentemente, a capacidade de produção de frutos, e a morte dos ramos (AVELINO et al., 2018), resultando em perdas significativas para a lavoura no curto e médio prazo.



Figura 2 – Plantas de café severamente afetadas pela ferrugem.
Fonte: Avelino et al. (2018).

O diagnóstico da doença é feito por um especialista por meio de análise visual das folhas do cafeeiro e, se necessário, também são realizados testes laboratoriais.

2.1.2 Visão Computacional

A Visão Computacional pode ser definida como a área de estudo que possibilita as máquinas a capacidade de enxergarem o mundo, ou seja, dar a elas o “sentido da visão”. Ela combina técnicas de Processamento Digital de Imagens e de Inteligência Artificial para possibilitar que as máquinas possam discernir o conteúdo de uma imagem, não apenas captá-la. Isso quer dizer que, além de captar a imagem, os computadores realizam o tratamento dela, eliminando ruídos, separando os objetos de interesse e extraíndo as informações dependendo de cada imagem e, ao final, relacionando com imagens vistas previamente (BALLARD; BROWN, 1982; BACKES; JUNIOR, 2016).

Como dito anteriormente, a Visão Computacional utiliza do processamento de imagem, que são um conjunto de técnicas aplicadas para “melhorarem” a imagem final, retirando ruídos, salientando bordas e suavizando a imagem, por exemplo (BACKES; JUNIOR, 2016). Vale ressaltar que uma imagem digital monocromática pode ser descrita matematicamente por uma função $f(x,y)$, gerando uma matriz bidimensional, e uma matriz com mais dimensões para o caso de imagens coloridas. Para as imagens coloridas, a

intensidade das cores varia de acordo com a sensibilidade ao vermelho (R - *red*), ao verde (G - *green*) e ao azul (B - *blue*) (MARQUES-FILHO; VIEIRA-NETO, 1999).

2.1.3 *Deep Learning*

O *Machine Learning*, aprendizado de máquina em português, pode ser entendido como a capacidade de um computador aprender sem ser, necessariamente, programado. A partir da observação de um conjunto de dados monta-se um modelo que será utilizado como hipótese sobre o mundo e, também, um software com a capacidade de resolver problemas complexos (LECUN; BENGIO; HINTON, 2015; RUSSELL; NORVIG, 2022).

O *Deep Learning* é um subcampo do *Machine Learning*, havendo bastante semelhança entre eles, sendo que se distinguem pela forma como cada algoritmo aprende. O *Deep Learning* necessita de pouca intervenção humana, pois ele consegue automatizar grande parte do processo de extração de características, aprendendo a partir dos dados, e necessita de grandes bases de dados (PONTI et al., 2017; KAUFMAN, 2022).

O termo *deep* (“profundo”) faz referência ao grande número de camadas que podem haver, indicando que os caminhos de computação das entradas e das saídas possuem diversas etapas. A entrada que cada uma dessas etapas receberá será a saída das camadas nos níveis superiores (PONTI et al., 2017; RUSSELL; NORVIG, 2022).

Atualmente, o aprendizado profundo é amplamente utilizado para aplicações de reconhecimento de voz, tradução automática, reconhecimento de objetos e reconhecimento de imagens, sendo que, para este último, um dos modelos de *deep learning* que tem se destacado são as Redes Neurais Convolucionais (PONTI et al., 2017; ESGARIO; KROHLING; VENTURA, 2020; RUSSELL; NORVIG, 2022)

2.1.3.1 Redes Neurais Convolucionais

Entre os modelos de *Deep Learning*, as Redes Neurais Convolucionais (*CNNs*) são, provavelmente, os mais conhecidos para resolverem tarefas de Visão Computacional. A principal característica da *CNN* é a existência de camadas convolucionais que agem como campos receptivos de neurônios e tem como principal aplicação o processamento de informações visuais (PONTI et al., 2017; ESGARIO; KROHLING; VENTURA, 2020).

Além das camadas convolucionais, a estrutura principal da uma *CNN* também é composta por mais duas camadas neurais, sendo as camadas de *pooling* e as camadas totalmente conectadas. Geralmente as camadas são organizadas de forma hierárquica, cujas camadas convolucionais se alternam com as camadas de *pooling* e são seguidas por camadas totalmente conectadas (GUO et al., 2016).

A camada de convolução é a responsável por aplicar uma operação de convolução a imagem de entrada. O objetivo principal de uma convolução é extrair características da

entrada por meio da aplicação de filtros (ou *kernel*) aos *pixels* da imagem. Os filtros são representados por uma matriz, normalmente de 3×3 ou 5×5 , com pesos (ou valores) que serão aplicados ao vetor de entrada (GUO et al., 2016; PONTI et al., 2017; RUSSELL; NORVIG, 2022).

A camada de *pooling* é comumente aplicada após uma camada convolucional, com o objetivo de reduzir a dimensão da imagem produzida nas camadas anteriores. Com a redução das dimensões da imagem há um grande ganho de velocidade, pois ela diminui o poder computacional necessário para processar os dados (PONTI et al., 2017; RUSSELL; NORVIG, 2022)

A última camada da *CNN* é a camada totalmente conectada, que é responsável por converter os mapas de características 2D em um único vetor de características 1D. Elas contém cerca de 90% do total de parâmetros de uma *CNN* e são responsáveis pela maior parte de todo o custo computacional para o treinamento (GUO et al., 2016; PONTI et al., 2017).

Neste trabalho serão avaliadas seis arquiteturas *CNNs*: *AlexNet* (KRIZHEVSKY; SUTSKEVER; HINTON, 2017), *DenseNet* (HUANG et al., 2016), *Inception* (SZEGEDY et al., 2015), *ResNet* (HE et al., 2016), *SqueezeNet* (IANDOLA et al., 2016) e *VGG* (SIMONYAN; ZISSERMAN, 2014). Cada uma delas foi escolhida devido ao seu sucesso em classificação de imagens.

- *AlexNet*: é composta por cinco camadas convolucionais, três camadas totalmente conectadas e *max-pooling* após cada convolução. Utiliza-se de conexões *dropout* para diminuir o *overfitting* e a função de ativação *ReLU* para acelerar o treinamento. Importante mencionar que ela foi a campeã do *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* de 2012. A Figura 3 demonstra a estrutura do modelo *AlexNet*.

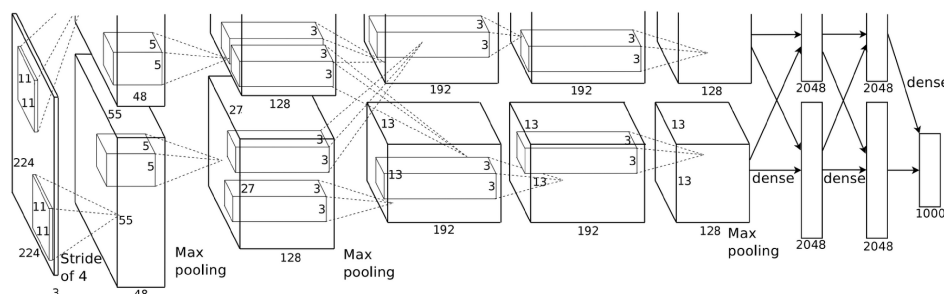


Figura 3 – Estrutura da *CNN AlexNet*.
Fonte: Krizhevsky, Sutskever e Hinton (2017).

- *DenseNet*: utiliza-se de estruturas de blocos densos para concatenar camadas convolucionais. O fluxo de informações entre as camadas é feita utilizando-se de conexões

diretas de qualquer camada para todas as subsequentes. Desta forma a *CNN* reduz o número de parâmetros necessários, incentiva a reutilização de recursos e fortalece a propagação de recursos (HUANG et al., 2016; MOREIRA et al., 2023). No trabalho foi utilizada a *DenseNet-121*. A Figura 4 demonstra a estrutura do modelo *DenseNet* de cinco camadas.

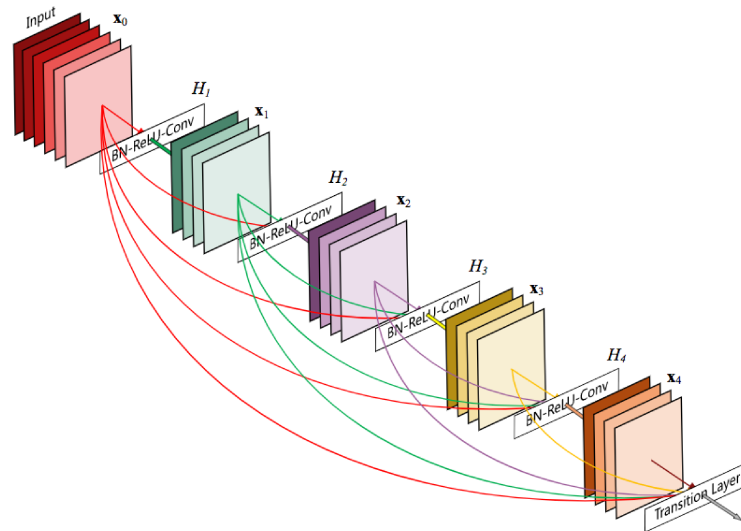


Figura 4 – Estrutura da *CNN DenseNet*.

Fonte: Huang et al. (2016).

- *Inception-V3*: a *Inception-V3* foi a arquitetura vencedora do *ILSVRC* 2014. Houve a introdução dos módulos *Inception*, que consistem em realizar múltiplas operações de convolução em paralelo, utilizando filtros de tamanhos diferentes, e concatenando as saídas. Desta forma, é possível que a rede capture características em diferentes escalas espaciais, melhorando a eficiência computacional e a capacidade de representação da rede (SZEGEDY et al., 2015).
- *ResNet*: foi a vencedora do *ILSVRC* 2015. A principal inovação apresentada por ela foi a forma como as camadas de uma rede neural profunda são formuladas e treinadas. A principal característica dessa abordagem é a introdução de conexões residuais, permitindo que as informações transitem diretamente entre camadas não adjacentes, ajudando a resolver o problema de degradação do desempenho que ocorre em redes neurais mais profundas. Neste trabalho foi utilizada a *ResNet* com 18 camadas e a sua arquitetura pode ser vista na Figura 5.

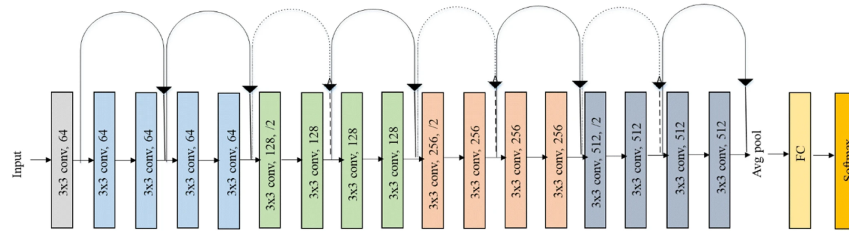


Figura 5 – Estrutura da *CNN ResNet*.

Fonte: Ramzan et al. (2020).

- *SqueezeNet*: é uma *CNN* leve, com 50 vezes menos parâmetros que a *AlexNet*, porém com precisão semelhante. Sua arquitetura é composta por uma camada de convolução independente, seguida por oito módulos “*fire*” e uma camada final de convolução. O módulo “*fire*” é uma camada de compressão e alimenta uma camada de expansão que possui uma mistura de filtros de convolução 1×1 e 3×3 (IANDOLA et al., 2016). A Figura 6 ilustra a estrutura da *CNN SqueezeNet*.

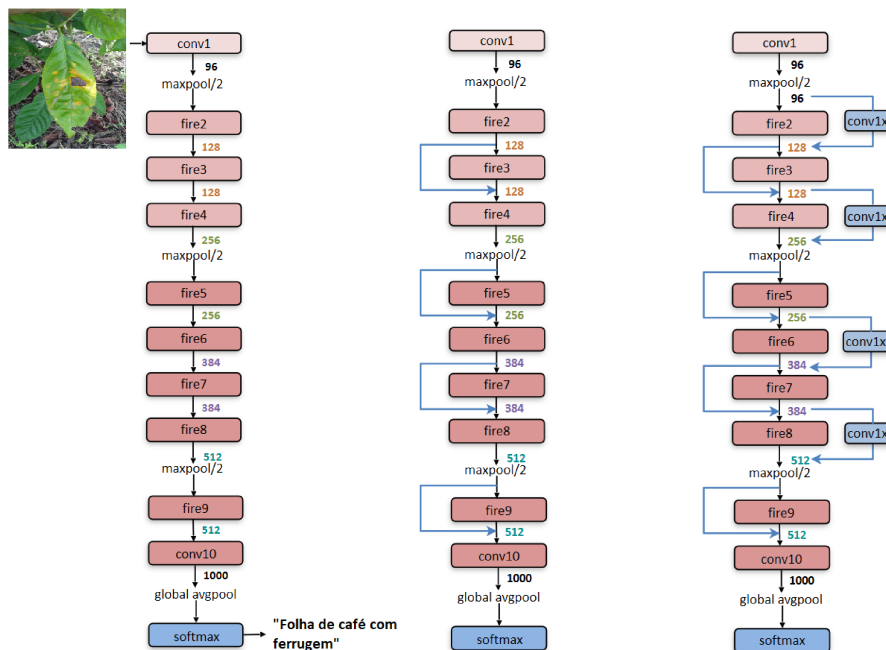


Figura 6 – Estrutura da *CNN SqueezeNet*.

Fonte: Adaptado de Iandola et al. (2016).

- *VGG*: foi a arquitetura vencedora nas tarefas de identificação e classificação do (*ILSVRC*) 2014. O modelo possui algumas variações, como a *VGG-11*, *VGG-16* e *VGG-19*, sendo que para esse trabalho foi considerada apenas a arquitetura *VGG-11*. Esse modelo possui oito camadas convolucionais e três camadas totalmente conectadas e foi o primeiro a substituir os filtros maiores, que exigem maior custo computacional, por uma sequência de filtros de tamanho 3×3 .

2.1.4 Protocolo de Avaliação

Os protocolos de avaliação são de grande importância para medir o desempenho das *CNNs* que trabalham com tarefas de reconhecimento de imagem. Neste trabalho foram utilizados os métodos de validação cruzada *k-fold* e as métricas de avaliação de acurácia, precisão, *recall* e *F1-Score*.

2.1.5 Validação Cruzada *k-fold*

Na validação cruzada *k-fold*, o conjunto de imagens é dividido aleatoriamente em subconjuntos de tamanhos iguais. Em cada um dos subconjuntos k , uma amostra é utilizada para validação dos testes e as demais amostras ($k - 1$) utilizadas para o treinamento (KOHAVI, 1995). Com o valor de k igual a cinco, por exemplo, ocorrerão cinco treinamentos e em cada um deles um subconjunto será utilizado para teste e os demais para treinamento, garantindo que todos os cinco subconjuntos serão utilizados nas duas fases. Utilizando-se a validação cruzada *k-fold* pode-se minimizar as incertezas dos testes de dados e o *overfitting* (VU et al., 2022)

A Figura 7 ilustra a validação cruzada *k-fold* em que o k tem tamanho cinco e especifica qual subamostra é utilizada para teste (marcado de verde) e quais serão utilizadas para treinamento (em branco) em cada iteração.

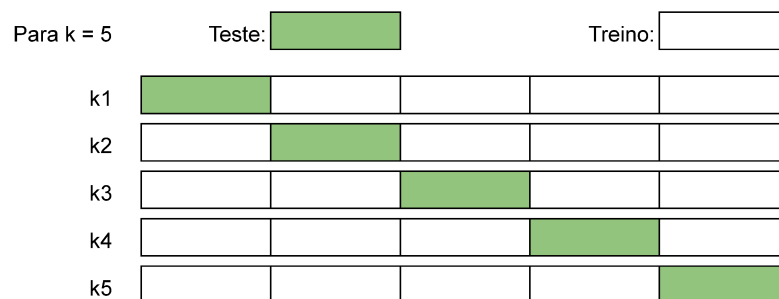


Figura 7 – Visão geral da validação cruzada *k-fold*.

2.1.6 Matriz de Confusão

A matriz de confusão é uma ferramenta útil para visualização de problemas de classificação. Ela é composta por uma tabela bidimensional que ilustra o número de predições corretas e incorretas em cada classe, fornecendo informações detalhadas sobre o desempenho de um classificador em um conjunto de dados. Com base na matriz de confusão, é possível examinar, de forma quantitativa, quais as classes o algoritmo de aprendizado tem maior dificuldade em classificar (FACELI et al., 2021; RUSSELL; NORVIG, 2022).

A matriz baseia-se no número de verdadeiros positivos, verdadeiros negativos, falso positivos e falsos negativos. Será considerado como verdadeiro positivo (*VP*) o elemento

positivo que é classificado como positivo; como verdadeiro negativo (*VN*) o elemento que é negativo e é classificado corretamente como negativo; como falso positivo (*FP*) o elemento de entrada falso que é definido pelo classificador como positivo; e, por fim, no falso negativo (*FN*) o elemento de entrada é positivo mas é classificado como negativo (NETTO; MACIEL, 2021).

A Figura 8 demonstra uma visão geral de uma matriz de confusão, a qual é composta por linhas e colunas que representam as instâncias da classe real e as instâncias da classe predita.

		Classe Predita	
		Positivo	Negativo
Classe real	Positivo	VP Verdadeiro Positivo	FN Falso Negativo
	Negativo	FP Falso Positivo	VN Verdadeiro Negativo

Figura 8 – Visão geral de uma matriz de confusão.

2.1.7 Métricas de Avaliação

O desempenho das *CNNs* testadas serão medidos com base nas seguintes métricas de avaliação:

- **Acurácia:** o proposito é estimar a quantidade de classes classificadas corretamente, para isso é feita a divisão de todos os acertos pelo total de classes (COSTA; CESAR-JR., 2009). A fórmula matemática está representada na Equação 2.1, onde *VP* representa os verdadeiros positivos, o *VN* são os verdadeiros negativos, o *FP* se refere ao falso positivo e, por fim, o *FN* é o falso negativo.

$$Acurácia = \frac{VP + VN}{VP + FP + FN + VN} \quad (2.1)$$

- **Precisão:** é indicador que relaciona as previsões positivas realizadas corretamente com o somatório de todas as previsões positivas, incluindo as falsas (COSTA; CESAR-JR., 2009). A Equação 2.2 corresponde à fórmula matemática da precisão.

$$Precisão = \frac{VP}{VP + FP} \quad (2.2)$$

- **Recall:** faz uma relação entre a quantidade de amostras que foram classificadas corretamente como positivas e aquelas que são de fato positivas (COSTA; CESAR-JR., 2009). O cálculo é feito pela razão dos valores de verdadeiro positivo com o somatório da quantidade de verdadeiros positivos e falso negativos, como pode ser visto na Equação 2.3.

$$Recall = \frac{VP}{VP + FN} \quad (2.3)$$

- *F1-Score*: é uma relação entre precisão e *recall*. O cálculo é feito por meio da média harmônica entre essas duas métricas (COSTA; CESAR-JR., 2009), como pode ser visto na Equação 2.4.

$$F1-Score = \frac{Precisão \times Recall}{Precisão + Recall} \quad (2.4)$$

2.2 Trabalhos Relacionados

O uso do processamento de imagem para detecção de doenças do café e outras plantas tem sido estudado por diversos pesquisadores. Neste sentido, Marcos, Rodovalho e Backes (2019) propuseram uma *CNN* para segmentação de imagens, que contém duas camadas convolucionais, seguidas pelo filtro *ReLU* e uma camada *max-pooling*. Também foi aplicada uma normalização a cada saída das camadas de *max-pooling*. Para avaliar a *CNN* foi utilizado um banco de imagens próprio, composto por 159 imagens de folhas de café, as quais foram classificadas manualmente por um especialista, que delimitou todos os pontos contendo ferrugem. Para mensurar o desempenho da *CNN* foi utilizado o coeficiente de Dice, sendo que os resultados obtidos foram de 0,79 como média e 0,82 como mediana.

Yebasse et al. (2021) usaram como banco de imagens o *RoCoLe* (2019), composto por 1560 imagens de folhas de café, sendo 791 saudáveis e 769 doentes. As folhas doentes foram divididas e classificadas da seguinte forma: 167 com ácaro-rajado-vermelho, 344 com ferrugem nível 1, 166 com ferrugem nível 2, 62 com ferrugem nível 3 e 30 com ferrugem nível 4. O estudo avaliou três técnicas de visualização de imagem e duas abordagens diferentes para detecção da ferrugem na folhas de café. A primeira foi a abordagem ingênua, que utilizou a *ResNet*, com algumas modificações, como *backbone* e o *Grad-CAM* como técnica de visualização. Nesta abordagem, o modelo atingiu 99% de acurácia no treinamento mas somente 77% nos testes. Já na abordagem guiada foi utilizada técnica de segmentação de imagem antes de realizar o treinamento, neste caso, obtiveram 98% de acurácia nos testes.

Dutta e Rana (2021) usaram o conjunto de dados *Bracol* (2019), composto por 1747 imagens, contendo folhas saudáveis e folhas afetadas pelas principais doenças do café. O conjunto é dividido em duas partes, uma contendo imagens das folhas completas e outro com imagens cortadas exibindo apenas uma das doenças. A *CNN MobileNet V2* foi utilizada para treinamento e transferência de aprendizado. Com o modelo proposto obtiveram uma acurácia de 98,51%.

Suparyanto et al. (2022) avaliaram a *CNN ResNet-18* para a classificação da ferrugem nas folhas de café, que foi treinada utilizando o *Stochastic Gradient Descent (SGD)*. No estudo foram avaliadas 100 imagens de folhas de café, obtidas exclusivamente para o projeto, as quais foram avaliadas por especialistas em doenças e pestes, que dividiram as folhas entre doentes e saudáveis. Por conta da base utilizada ser limitada, a *CNN* teve um sobreajuste (*overfitting*) muito grande e a acurácia atingida, após o treinamento, foi de apenas 59%.

Pandian et al. (2022) propuseram a criação de uma *CNN* de 14 camadas para detecção de doenças de plantas. Para treinamento, validação e testes foram utilizadas imagens de cinco diferentes repositórios públicos, dentre eles o *Bracol* (2019) e o *RoCoLe* (2019), resultando em 61.459 imagens de folhas de plantas. Após o treinamento, foram submetidas imagens reais de folhas de diversas plantas e a *CNN* foi capaz de identificar a qual espécie de planta pertence a folha e qual a doença presente nela. O modelo obteve uma acurácia de 99.96%, precisão de 99.79%, *recall* de 99.79% e *F1-score* de 99.79%.

Montalbo (2022) apresentaram um modelo composto por três *CNNs* agregadas como uma única unidade, sendo elas a *DenseNet-121*, *VGG-16* e *EfficientNetB0*. A base de dados utilizada contém 4.675 imagens de folhas de café, divididas em sete categorias diferentes, a depender do nível e do tipo de doença. Essa base foi construída a partir da união de três outras bases de imagens públicas, sendo elas a *Bracol* (2019), a *RoCoLe* (2019) e a *LiCoLe* (2020). O modelo proposto atingiu uma precisão geral de 95,98%.

O presente trabalho se diferencia dos trabalhos citados anteriormente ao comparar diferentes arquiteturas *CNNs*, explorando técnicas de aumento de dados e validação cruzada *k-fold*. Além disso, a base de dados *RoCoLe* foi pouco explorada para tarefas de classificação de imagens e o presente trabalho pode trazer oportunidades de pesquisa neste campo, o que pode auxiliar, diretamente, os cafeicultores que sofrem com a ferrugem dos cafezais.

A Tabela 1 sumariza os trabalhos relacionados e descritos neste capítulo, permitindo encontrar, de forma rápida e resumida, as informações mais importantes sobre eles.

Tabela 1 – Comparação entre os trabalhos relacionados e o presente trabalho.

Trabalho	Conjunto de Imagens	Método	Métrica de Avaliação
Marcos, Rodovalho e Backes (2019)	Conjunto de imagens próprio (159 imagens de folhas de café)	Segmentação com <i>CNN</i>	Coefficiente de Dice
Yebasse et al. (2021)	<i>RoCoLe</i>	<i>CNN: ResNet</i> Abordagem ingênua e treinamento com imagens segmentadas	Acurácia
Dutta e Rana (2021)	<i>Bracol</i>	<i>CNN MobileNet V2</i> otimizada com <i>AdaMax</i>	Acurácia
Suparyanto et al. (2022)	Conjunto de imagens próprio (100 imagens de folhas de café)	<i>CNN ResNet-18</i> com <i>SGD</i>	Acurácia
Pandian et al. (2022)	<i>Bracol</i> e <i>RoCoLe</i>	Propôs uma <i>CNN</i> com 14 camadas	Acurácia Precisão <i>Recall</i> <i>F1-Score</i>
Montalbo (2022)	<i>Bracol</i> , <i>RoCoLe</i> e <i>LiCoLe</i>	Agregação de três <i>CNNs</i> : <i>DenseNet-121</i> , <i>VGG-16</i> , <i>EfficientNetB0</i>	Acurácia Precisão <i>Recall</i> <i>F1-Score</i>
Presente Trabalho	<i>RoCoLe</i>	<i>CNNs: AlexNet, DenseNet, Inception, ResNet, SqueezeNet</i> e <i>VGG</i> validação cruzada k-fold e aumento de dados	Acurácia Precisão <i>Recall</i> <i>F1-Score</i>

3 Metodologia

Para o desenvolvimento do trabalho foi utilizada a linguagem *Python* e o *Framework Pytorch* (PASZKE et al., 2019). Este é um *framework* de código aberto, voltado para aplicativos de Visão Computacional e possui diversas arquiteturas *CNNs* para classificação de imagens.

Periodicamente é realizada a competição *ImageNet*¹ para avaliar algoritmos de detecção de objetos e classificação de imagens. Nesses desafios surgiram várias arquiteturas *CNNs* que alcançaram robustos resultados e estão presentes no *framework Pytorch*.

Para a realização do trabalho foram comparadas seis arquiteturas *CNNs*: *AlexNet* (KRIZHEVSKY; SUTSKEVER; HINTON, 2017), *DenseNet* (HUANG et al., 2016), *Inception* (SZEGEDY et al., 2015), *ResNet* (HE et al., 2016), *SqueezeNet* (IANDOLA et al., 2016) e *VGG* (SIMONYAN; ZISSERMAN, 2014).

Foi utilizado o conjunto de dados (*dataset*) *RoCoLe*², disponibilizado publicamente, contendo 1.560 imagens de folhas de café, sendo 791 de folhas saudáveis e 769 folhas doentes. Na Figura 9 é possível ver um exemplo de folha saudável e outra com ferrugem que estão disponíveis no conjunto de imagens *RoCoLe*.



(a) Folha saudável.

(b) Folha doente.

Figura 9 – Exemplo de imagens do conjunto *RoCoLe*.

Foram utilizadas técnicas de aumento de dados sobre as imagens do *dataset* a fim de aumentar artificialmente o conjunto de dados e evitar que ocorra o *overfitting*. As imagens também foram redimensionadas para se adequarem ao tamanho de entrada requerido pelas *CNNs* utilizadas neste trabalho. Além disso, o conjunto de dados foi dividido utilizando-se a técnica de validação cruzada *k-fold*, com valor de *k* igual a cinco.

¹ Disponível em: <https://www.image-net.org>

² Disponível em: <https://doi.org/10.17632/c5yv32dzg.2>

As CNNs foram treinadas e avaliadas de acordo com critérios de acurácia, precisão, *recall* e *F1-score*. Duas análises foram consideradas: (i) binária, em que as imagens foram classificadas em duas classes, com e sem ferrugem; e (ii) multiclasse, em que as imagens foram classificadas em cinco classes, uma para imagens sem ferrugem e as demais para diferentes graduações de ferrugens. O método proposto pelo trabalho pode ser visto na Figura 10.

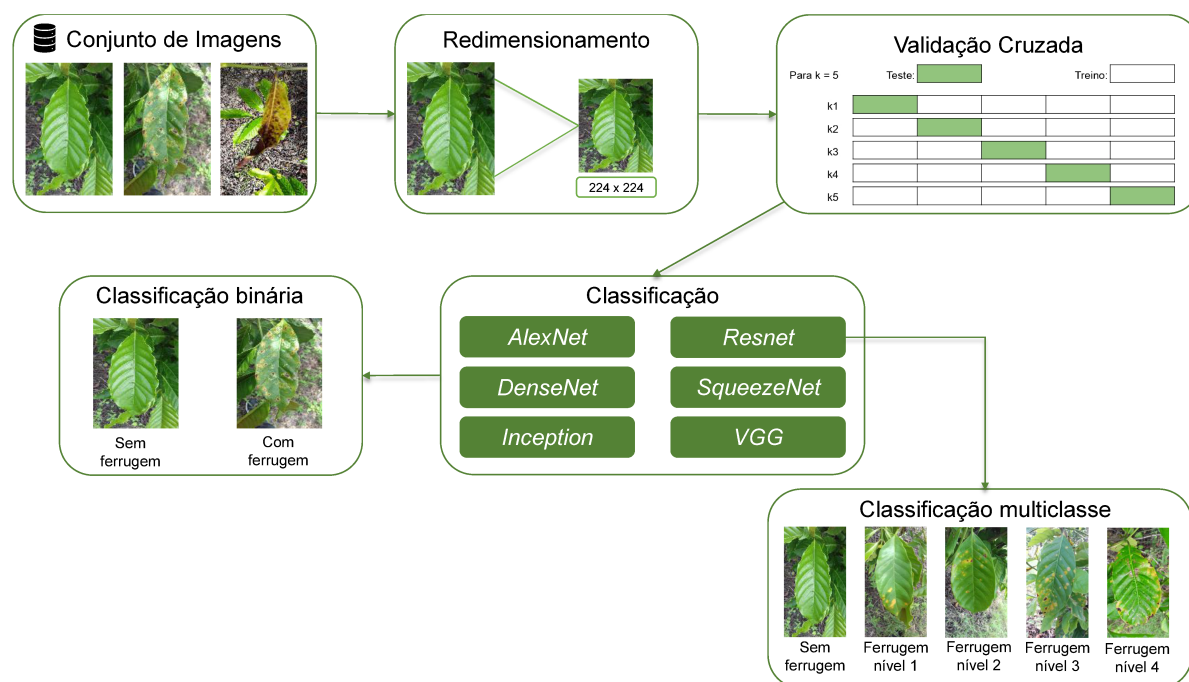


Figura 10 – Método Proposto.

4 Resultados e discussões

Neste capítulo são apresentados os resultados e discussões deste trabalho, avaliando o desempenho das seis Redes Neurais Convolucionais (*CNNs*) selecionadas para a tarefa de classificação de imagens com o objetivo de detectar a presença de ferrugem em folhas de café. Inicialmente, as *CNNs* foram testadas em uma classificação binária, para distinguir as folhas saudáveis das folhas doentes. Posteriormente, a *CNN* com a melhor acurácia na classificação binária foi selecionada para um teste de classificação multiclasse no qual as folhas foram divididas nas seguintes classes: saudáveis e níveis de ferrugem que variam entre um (sintomas leves) e quatro (sintomas severos). Os resultados obtidos e as análises realizadas fornecerão informações valiosas sobre a capacidade das *CNNs* de detectar e diferenciar ferrugem nas folhas do cafeeiro, contribuindo para o desenvolvimento de técnicas eficientes de detecção de doenças na cultura do café.

4.1 Ambiente de Testes

Os testes foram executados em uma máquina com processador *AMD Ryzen 5 5600X*, operando a uma frequência máxima de 4,5 GHz, 32 GB de *RAM*, *GPU NVIDIA RTX 3060 Ti* de 8 GB de memória e sistema operacional *Windows 10*. Todos os modelos foram programados utilizando a linguagem *Python* na versão 3.8, juntamente com o *framework PyTorch* (versão 2.0) com *CUDA* versão 11.7 e *cuDNN* 8.0.

4.2 Classificação Binária

Para a classificação binária, foram classificadas imagens de folhas de café como saudáveis ou doentes. Foram utilizadas as 1.560 imagens presentes no *dataset RoCole* (PARRAGA-ALAVA et al., 2019), que contém 791 imagens de folhas saudáveis e 769 doentes. Cabe salientar que entre as folhas doentes, há a divisão por nível de ferrugem, que varia do nível um ao nível quatro e também uma pequena quantidade de imagens (apenas 167) de folhas atacadas pelo ácaro-rajado-vermelho. Houve aplicação de aumento de dado nas imagens de treinamento por rotação aleatória (variando entre -30° e 30°), além das inversões verticais e horizontais.

A Figura 11 apresenta, como exemplo, os gráficos de perda e acurácia da *ResNet* gerados durante o treinamento e a validação do *fold* 0. Como pode ser visto, ao decorrer das épocas a perda vai diminuindo, o que indica que o modelo está melhorando. Há uma diminuição rápida no início e uma estabilização no final da curva, em um nível baixo, sugerindo que o modelo está convergindo e alcançando um bom desempenho.

Já a acurácia vai aumentando ao longo das épocas, indicando que a taxa de acertos do modelo está crescendo e melhorando. Ao final da curva da acurácia, podemos ver que ela se estabiliza em um nível alto, indicando que o modelo está aprendendo efetivamente os padrões nos dados. Também é importante mencionar que, tanto para a perda quanto para a acurácia, as curvas de treinamento e validação estão próximas e seguem um padrão semelhante, indicando que o modelo está generalizando bem e o desempenho no conjunto de validação é semelhante ao de treinamento, demonstrando que não está ocorrendo o sobreajuste.

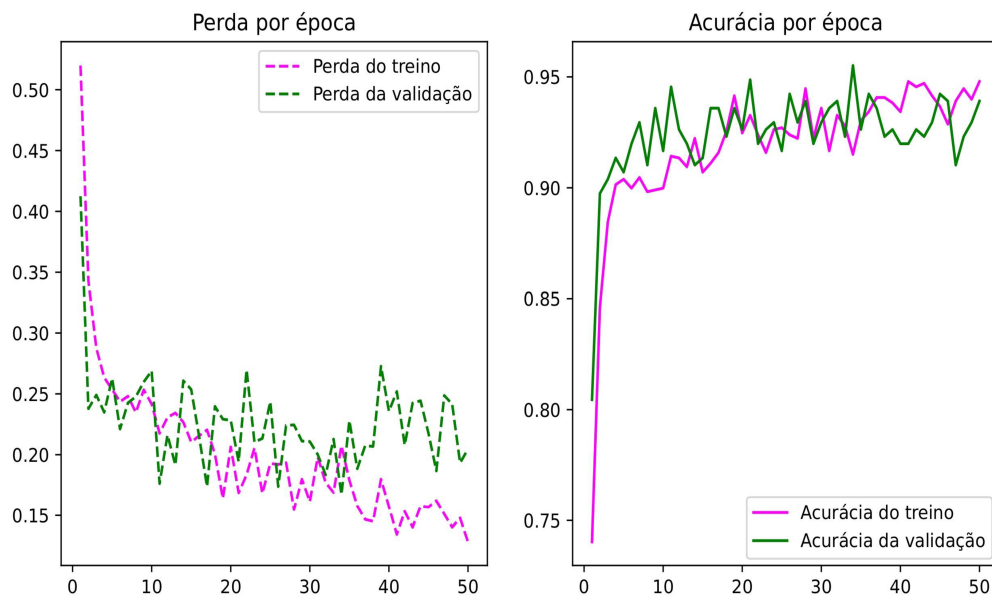


Figura 11 – Gráfico de perda (*loss*) e acurácia durante o treinamento e validação do *Fold 0* da arquitetura *ResNet*.

As médias de desempenho das *CNNs* quanto a métrica de acurácia, que é calculada dividindo o número de previsões corretas pelo número total de imagens de folhas de café no conjunto de teste, são apresentadas na Figura 12. É interessante observar que todas as *CNNs* alcançaram uma média de acurácia acima de 90%. Destaca-se a *ResNet*, que obteve o melhor desempenho, com uma acurácia de 95,19%. Em seguida, temos a *VGG* com 94,87% e a *Inception* com 94,55%.

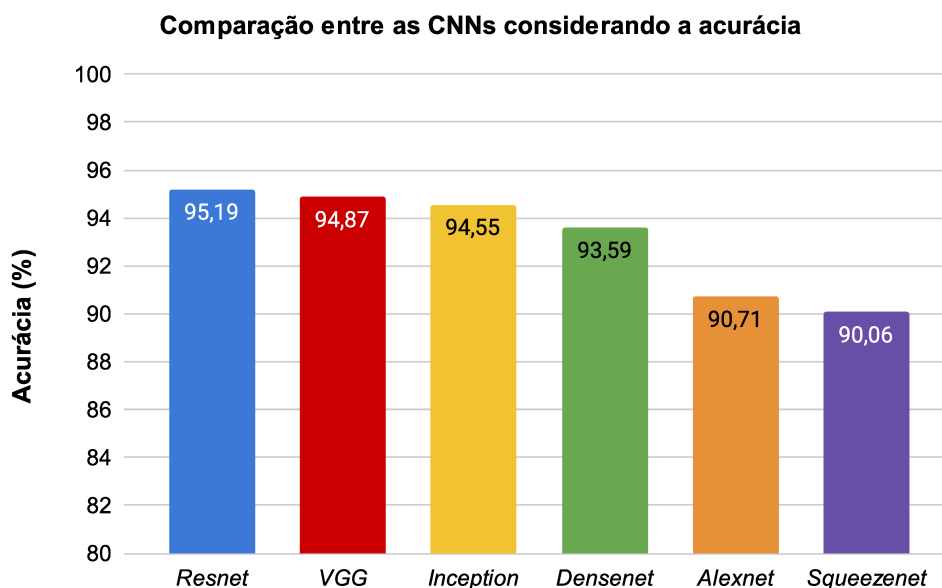


Figura 12 – Comparação da acurácia das redes testadas.

O desempenho das redes quanto as métricas de precisão, *recall* e *F1-Score* estão sumarizados nas Tabelas 2, 3, 4, 5, 6 e 7 e as Figuras 13 e 14 ilustram as variações de cada uma das *CNNs*. Dessas métricas, podemos destacar a *recall*, que, quando alta, indica que o modelo tem uma alta capacidade de identificar corretamente as folhas doentes, ou seja, um *recall* alto indica que a *CNN* tem uma menor taxa de falsos negativos, com uma menor probabilidade de não identificar corretamente uma folha doente como tal. Para o cafeicultor, identificar precocemente as folhas afetadas pela doença é crucial para controle e prevenção, permitindo que ele tome medidas adequadas de forma mais assertiva.

Nesse sentido, todas as *CNNs* utilizadas no trabalho alcançaram um *recall* acima de 90%. Podemos destacar os três melhores desempenhos, sendo que a *VGG* obteve o melhor, identificando corretamente 95,74% das amostras com doença. Em seguida a *Inception* com 95,04% e a *ResNet* com 94,33%.

Tabela 2 – Avaliação de desempenho da arquitetura *Alexnet*.

<i>Classe</i>	<i>Precisão (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
Saudável	92,77	90,06	91,39
Doente	88,36	91,49	89,90
Média	90,56	90,77	90,65

Tabela 3 – Avaliação de desempenho da arquitetura *Densenet*.

<i>Classe</i>	<i>Precisão (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
Saudável	94,67	93,57	94,12
Doente	92,31	93,62	92,96
Média	93,49	93,59	93,54

Tabela 4 – Avaliação de desempenho da arquitetura *Inception*.

<i>Classe</i>	<i>Precisão (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
Saudável	95,83	94,15	94,99
Doente	93,06	95,04	94,04
Média	94,44	94,59	94,51

Tabela 5 – Avaliação de desempenho da arquitetura *ResNet*.

<i>Classe</i>	<i>Precisão (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
Saudável	95,35	95,91	95,63
Doente	95,00	94,33	94,66
Média	95,17	95,12	95,14

Tabela 6 – Avaliação de desempenho da arquitetura *Squeezenet*.

<i>Classe</i>	<i>Precisão (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
Saudável	91,18	90,64	90,91
Doente	88,73	89,36	89,05
Média	89,95	90,00	89,98

Tabela 7 – Avaliação de desempenho da arquitetura *VGG*.

<i>Classe</i>	<i>Precisão (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
Saudável	96,41	94,15	95,27
Doente	93,10	95,74	94,41
Média	94,76	94,95	94,84

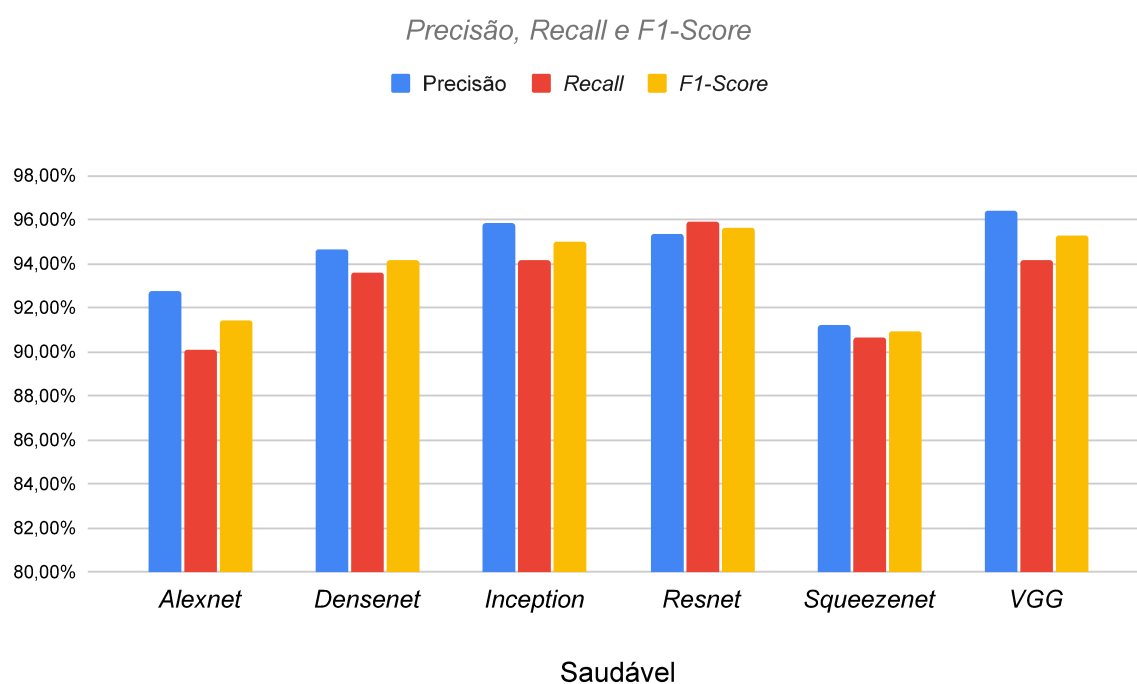


Figura 13 – Variação de desempenho das *CNNs* ao avaliar folhas saudáveis.

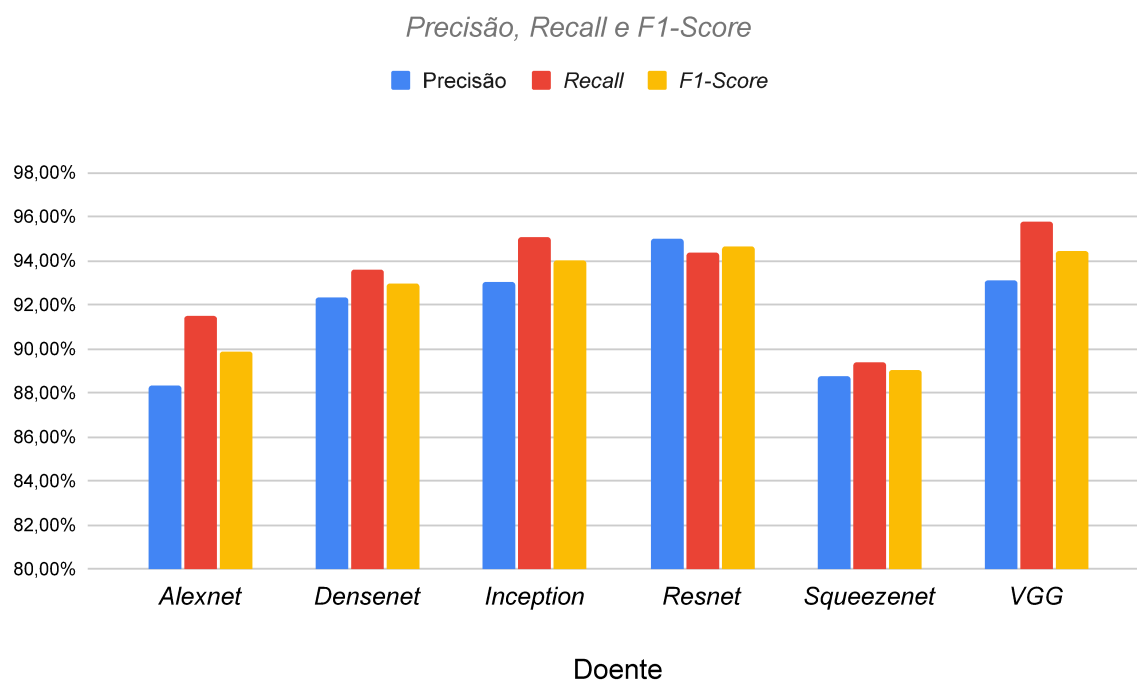


Figura 14 – Variação de desempenho das *CNNs* ao avaliar folhas doentes.

Analisando a Tabela 8 é possível notar, a partir das matrizes de confusão do *fold* 0, que as *CNNs* analisadas no trabalho conseguiram diferenciar, com bastante assertividade, as imagens que contém as folhas saudáveis das folhas doentes. Em especial, as arquiteturas

DenseNet e *VGG* foram capazes de classificar uma quantidade maior de folhas saudáveis corretamente enquanto para as folhas doentes, a *DenseNet* e a *ResNet* se destacaram.

Tabela 8 – Valores das matrizes de confusão de cada modelo *CNN*.

		<i>Alexnet</i>		<i>DenseNet</i>	
		Saudável	Doente	Saudável	Doente
Saudável		151	20	165	6
Doente		11	130	4	137

		<i>Inception</i>		<i>ResNet</i>	
		Saudável	Doente	Saudável	Doente
Saudável		161	10	159	12
Doente		9	132	7	134

		<i>Squeezenet</i>		<i>VGG</i>	
		Saudável	Doente	Saudável	Doente
Saudável		160	11	162	9
Doente		11	130	11	130

4.3 Seleção do melhor modelo para Classificação Multiclasse

Para a classificação multiclasse, utilizou-se a arquitetura *ResNet*, por ter obtido a melhor acurácia na classificação binária. Foram classificadas as imagens do *dataset RoCoLe* (PARRAGA-ALAVA et al., 2019) separadas em cinco classes: saudável, ferrugem nível 1, ferrugem nível 2, ferrugem nível 3 e ferrugem nível 4. O número de imagens por classe não está dividida de forma homogênea, sendo separadas da seguinte forma: 791 saudáveis, 344 com ferrugem nível 1, 166 com ferrugem nível 2, 62 com ferrugem nível 3 e 30 com ferrugem nível 4. O conjunto de dados também possui 167 imagens de folhas atacadas pelo ácaro-rajado-vermelho, mas não foram consideradas para a classificação multiclasse. Houve aplicação de aumento de dados nas imagens de treinando por rotação aleatória (variando entre -30° e 30°), além das inversões verticais e horizontais.

A Figura 15 apresenta o gráfico de perda e acurácia da *ResNet* durante o treinamento e validação para o *fold* 0. É possível notar que durante a fase de treinamento, a perda cai rapidamente e estabiliza no final da curva em um nível baixo, sugerindo que, ao decorrer das épocas, o modelo está alcançando um bom desempenho. Também é possível notar essa evolução no gráfico da acurácia, que vai subindo a medida que as épocas vão aumentando e estabiliza em um nível alto no final da curva. Porém, diferentemente do gráfico demonstrado na classificação binária da Seção 4.2, podemos notar um afastamento da curva de validação, tanto para o gráfico da perda quanto para a acurácia, o que sugere que, apesar do modelo estar aprendendo, está ocorrendo um sobreajuste e ele não está conseguindo generalizar de forma adequada.

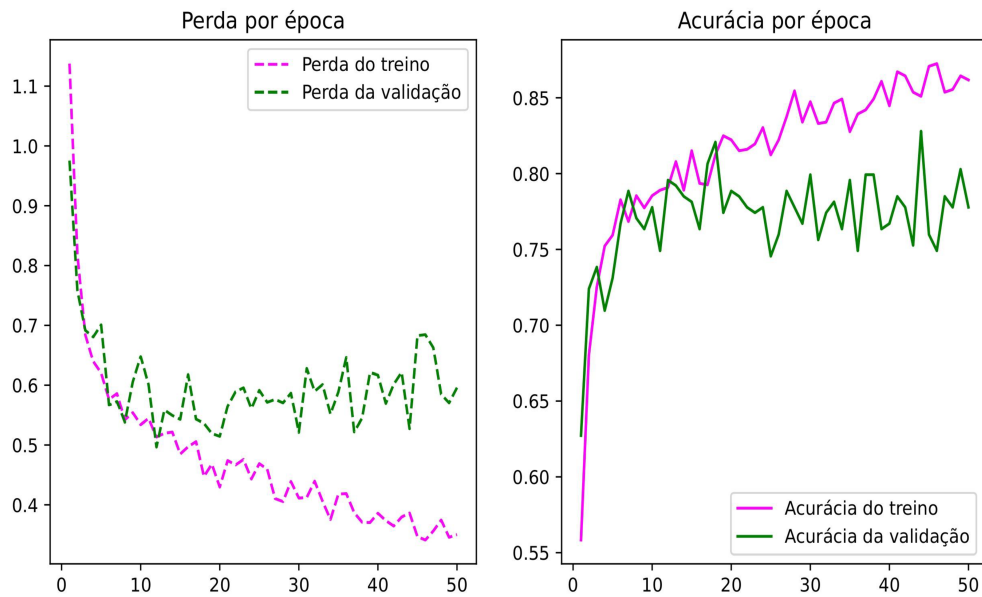


Figura 15 – Gráfico de perda (*loss*) e acurácia durante o treinamento e validação multi-classe do *Fold 0* da arquitetura *ResNet*.

Para a classificação multiclasse a *ResNet* atingiu 78,03% de acurácia, mostrando que a abordagem é promissora para a classificação da ferrugem, mas é evidente que utilizando-se as cinco classes é bem mais desafiador para a *CNN* e precisa ser melhor explorada em pesquisas futuras.

O desempenho das demais métricas podem ser visualizados na Tabela 9, na qual é possível perceber que para as folhas saudáveis o modelo obteve bons resultados, mas a partir do nível 2 de ferrugem o desempenho cai drasticamente. Isso pode ter ocorrido por conta do desbalanceamento da base de dados. Enquanto as folhas saudáveis representam 56,78% das imagens, as folhas com ferrugem nível três e quatro representam apenas 4,45% e 2,15% respectivamente. Mesmo com o aumento de dados aplicados, ainda foi insuficiente para que o modelo conseguisse aprender de forma adequada.

Tabela 9 – Avaliação de desempenho da arquitetura *ResNet* na classificação multiclasse.

<i>Classe</i>	<i>Precisão</i>	<i>Recall</i>	<i>F1-Score</i>
Saudável	91,93%	94,74%	93,31%
Ferrugem nível 1	67,99%	69,81%	68,77%
Ferrugem nível 2	53,95%	45,62%	48,28%
Ferrugem nível 3	23,37%	20,69%	21,30%
Ferrugem nível 4	52,21%	35,52%	34,42%
Média	57,89%	53,28%	53,21%

É possível notar, por meio da matriz de confusão apresentada na Tabela 10, que a *ResNet* conseguiu classificar corretamente quase todas as folhas saudáveis e a maioria

das folhas com ferrugem nível 1, mas a medida que a quantidade de imagens por classe vai diminuindo, a capacidade da rede classificar de forma correta também fica reduzida, fazendo com que ela confunda a classe na qual cada imagem pertence.

Tabela 10 – Matriz de confusão do *Fold 0* da classificação multiclasse da arquitetura *ResNet*

	Saudável	Ferrugem nível 1	Ferrugem nível 2	Ferrugem nível 3	Ferrugem nível 4
Saudável	150	8	0	0	0
Ferrugem nível 1	19	44	12	1	0
Ferrugem nível 2	0	8	11	2	1
Ferrugem nível 3	0	0	11	2	0
Ferrugem nível 4	0	1	2	1	6

5 Conclusão

O café está presente na vida de inúmeras famílias ao redor do mundo. É inegável a importância que ele tem na vida e no dia a dia das pessoas. Também é inegável que a ferrugem traz grandes desafios e prejuízos à cafeicultores ao redor do mundo, sejam eles produtores em grande escala ou em cultivo familiar. Nesse sentido, buscar métodos que possam auxiliar na identificação da doença, de forma rápida, assertiva e de baixo custo é muito importante. Para isso, o uso de técnicas de Visão Computacional podem ser uma alternativa promissora.

No trabalho foram avaliadas seis Redes Neurais Convolucionais para classificação de imagens de folha de café, com o objetivo de detectar a presença de folhas doentes em um conjunto de dados. Foram exploradas duas estratégias diferentes de classificação: binária (folha saudável ou doente) e multiclasse (folha saudável ou com algum dos quatro níveis de ferrugem). Também foram aplicadas técnicas de aumento de dados para tentar lidar com o desequilíbrio da base.

As arquiteturas utilizadas no trabalho foram: *AlexNet*, *DenseNet*, *Inception*, *ResNet*, *SqueezeNet* e *VGG*. Todas tiveram seus desempenhos avaliados de acordo com as métricas de acurácia, precisão, *recall* e *F1-Score*. Para a estratégia de classificação binária, o melhor resultado foi de 95,19% de acurácia, obtido com a *CNN ResNet*. Por ter alcançado o melhor desempenho na classificação binária, a *ResNet* foi selecionada para a classificação multiclasse, porém a tarefa se mostrou bem desafiadora. Por se tratar de uma base desbalanceada e haver poucas imagens para os níveis mais severos da ferrugem, a *CNN* teve um pouco mais de dificuldade e obteve a acurácia de 78,03%. Os resultados obtidos sugerem que a utilização das Redes Neurais Convolucionais testadas podem auxiliar o agricultor na identificação da doença, permitindo que ele possa efetuar o controle dela e prevenir que se espalhe para o restante do cafezal.

Em trabalhos futuros, pretende-se aumentar a base de dados, investigar novas técnicas de aumento de dados ([SANTOS et al., 2023](#)), testar diferentes arquiteturas *CNNs*, realizar a otimização dos hiper-parâmetros ([RODRIGUES et al., 2022](#)) e desenvolver uma aplicação móvel.

Referências

AVELINO, J.; ALLINNE, C.; CERDA, R.; WILLOCQUET, L.; SAVARY, S. Multiple-disease system in coffee: From crop loss assessment to sustainable management. **Annual review of phytopathology**, Annual Reviews, v. 56, p. 611–635, 2018. Citado 2 vezes nas páginas 8 e 9.

AVELINO, J.; CRISTANCHO, M.; GEORGIU, S.; IMBACH, P.; AGUILAR, L.; BORNEMANN, G.; LÄDERACH, P.; ANZUETO, F.; HRUSKA, A. J.; MORALES, C. The coffee rust crises in colombia and central america (2008–2013): impacts, plausible causes and proposed solutions. **Food security**, Springer Netherlands, Dordrecht, v. 7, n. 2, p. 303–321, 2015. ISSN 1876-4517. Citado na página 8.

BACKES, A.; JUNIOR, J. J. d. M. S. **Introdução à Visão Computacional Usando MATLAB**. Rio de Janeiro: Alta Books Editora, 2016. ISBN 978-85-508-0023-3. Citado na página 9.

BALLARD, D. H.; BROWN, C. M. **Computer Vision**. 1st. ed. New Jersey: Prentice Hall Professional Technical Reference, 1982. ISBN 0131653164. Citado na página 9.

COSTA, L. d. F.; CESAR-JR., R. M. **Shape Classification and Analysis: Theory and Practice**. Boca Raton, FL, USA: CRC Press, Inc., 2009. Citado 2 vezes nas páginas 15 e 16.

DUTTA, L.; RANA, A. K. Disease Detection Using Transfer Learning In Coffee Plants. In: IEEE. **2021 2nd Global Conference for Advancement in Technology (GCAT)**. Bangalore, India, 2021. p. 1–4. Citado 2 vezes nas páginas 16 e 18.

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**, Elsevier, v. 169, p. 105162, 2020. Citado na página 10.

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. d. L. F. d. **Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina**. Grupo GEN, 2021. ISBN 9788521637509. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9788521637509/>>. Citado na página 14.

GUO, Y.; LIU, Y.; OERLEMANS, A.; LAO, S.; WU, S.; LEW, M. S. Deep learning for visual understanding: A review. **Neurocomputing**, v. 187, p. 27–48, 2016. ISSN 0925-2312. Recent Developments on Deep Big Vision. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231215017634>>. Citado 3 vezes nas páginas 6, 10 e 11.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep Residual Learning for Image Recognition. In: **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Las Vegas, NV, USA: IEEE, 2016. p. 770–778. Citado 2 vezes nas páginas 11 e 19.

- HUANG, G.; LIU, Z.; MAATEN, L. van der; WEINBERGER, K. Q. **Densely Connected Convolutional Networks**. arXiv, 2016. Disponível em: <<https://arxiv.org/abs/1608.06993>>. Citado 3 vezes nas páginas 11, 12 e 19.
- LANDOLA, F. N.; HAN, S.; MOSKEWICZ, M. W.; ASHRAF, K.; DALLY, W. J.; KEUTZER, K. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and < 0.5 mb model size. **arXiv preprint arXiv:1602.07360**, 2016. Citado 3 vezes nas páginas 11, 13 e 19.
- INTERNATIONAL COFFEE ORGANIZATION. **Coffee Development Report 2019**. London, 2019. Disponível em: <<https://www.ico.org/documents/cy2018-19/ed-2318e-overview-flagship-report.pdf>>. Citado 2 vezes nas páginas 6 e 7.
- KAUFMAN, D. **Desmistificando a Inteligência Artificial**. Grupo Autêntica, 2022. ISBN 9786559281596. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9786559281596/>>. Citado na página 10.
- KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: **Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995. (IJCAI'95), p. 1137–1143. ISBN 1558603638. Citado na página 14.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Commun. ACM**, Association for Computing Machinery, New York, NY, USA, v. 60, n. 6, p. 84–90, may 2017. ISSN 0001-0782. Disponível em: <<https://doi-org.ez34.periodicos.capes.gov.br/10.1145/3065386>>. Citado 2 vezes nas páginas 11 e 19.
- KROHLING, R. A. **BRACOL - A Brazilian Arabica Coffee Leaf images dataset to identification and quantification of coffee diseases and pests**. Mendeley, 2019. Type: dataset. Disponível em: <<https://data.mendeley.com/datasets/yy2k5y8mxg/1>>. Citado 2 vezes nas páginas 16 e 17.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015. Citado na página 10.
- MARCOS, A. P.; RODOVALHO, N. L. S.; BACKES, A. R. Coffee Leaf Rust Detection Using Convolutional Neural Network. In: **2019 XV Workshop de Visão Computacional (WVC)**. São Bernardo do Campo, Brazil: IEEE, 2019. p. 38–42. Disponível em: <<http://doi.org/10.1109/WVC.2019.8876931>>. Citado 2 vezes nas páginas 16 e 18.
- MARQUES-FILHO, O.; VIEIRA-NETO, H. **Processamento digital de imagens**. Rio de Janeiro: Brasport, 1999. Citado na página 10.
- 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>>. Citado na página 17.

- MONTALBO, F. J. P. Automated diagnosis of diverse coffee leaf images through a stage-wise aggregated triple deep convolutional neural network. **Machine Vision and Applications**, v. 33, n. 1, p. 19, Jan 2022. ISSN 1432-1769. Disponível em: <<https://doi.org/10.1007/s00138-022-01277-y>>. Citado 2 vezes nas páginas 17 e 18.
- MOREIRA, L. F. R.; MOREIRA, R.; TRAVENÇOLO, B. A. N.; BACKES, A. R. An Artificial Intelligence-as-a-Service Architecture for deep learning model embodiment on low-cost devices: A case study of COVID-19 diagnosis. **Applied Soft Computing**, v. 134, p. 110014, 2023. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494623000327>>. Citado na página 12.
- NETTO, A.; MACIEL, F. **Python para Data Science e Machine Learning Descomplicado**. Alta Books, 2021. ISBN 9786555203172. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9786555203172>>. Citado na página 15.
- ONGSULEE, P. Artificial intelligence, machine learning and deep learning. In: **IEEE. 2017 15th international conference on ICT and knowledge engineering (ICT&KE)**. Bangkok, Thailand, 2017. p. 1–6. Citado na página 6.
- PANDIAN, J. A.; KUMAR, V. D.; GEMAN, O.; HNATIUC, M.; ARIF, M.; KANCHANADEVI, K. Plant disease detection using deep convolutional neural network. **Applied Sciences**, v. 12, n. 14, 2022. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/12/14/6982>>. Citado 2 vezes nas páginas 17 e 18.
- PARRAGA-ALAVA, J.; CUSME, K.; LOOR, A.; SANTANDER, E. Rocolo: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition. **Data in brief**, Elsevier, v. 25, p. 104414, 2019. Citado 4 vezes nas páginas 16, 17, 21 e 26.
- PASZKE, A.; GROSS, S.; MASSA, F.; LERER, A.; BRADBURY, J.; CHANAN, G.; KILLEEN, T.; LIN, Z.; GIMELSHEIN, N.; ANTIGA, L.; DESMAISON, A.; KOPF, A.; YANG, E.; DEVITO, Z.; RAISON, M.; TEJANI, A.; CHILAMKURTHY, S.; STEINER, B.; FANG, L.; BAI, J.; CHINTALA, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: **Advances in Neural Information Processing Systems 32**. Curran Associates, Inc., 2019. p. 8026–8037. Disponível em: <<http://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>>. Citado na página 19.
- PONTI, M. A.; RIBEIRO, L. S. F.; NAZARE, T. S.; BUI, T.; COLLOMOSSE, J. Everything You Wanted to Know about Deep Learning for Computer Vision but Were Afraid to Ask. In: **2017 30th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T)**. Niterói, Brazil: IEEE, 2017. p. 17–41. Citado 3 vezes nas páginas 6, 10 e 11.
- RAMZAN, F.; KHAN, M. U. G.; REHMAT, A.; IQBAL, S.; SABA, T.; REHMAN, A.; MEHMOOD, Z. A deep learning approach for automated diagnosis and multi-class classification of Alzheimer’s disease stages using resting-state fMRI and residual neural networks. **Journal of medical systems**, Springer, v. 44, p. 1–16, 2020. Citado na página 13.

- RODRIGUES, L. F.; BACKES, A. R.; TRAVENÇOLO, B. A. N.; OLIVEIRA, G. M. B. de. Optimizing a deep residual neural network with genetic algorithm for acute lymphoblastic leukemia classification. **Journal of Digital Imaging**, Springer Science and Business Media LLC, v. 35, n. 3, p. 623–637, feb 2022. Citado na página 29.
- RUSSELL, S. J.; NORVIG, P. **Inteligência Artificial: Uma Abordagem Moderna**. Grupo GEN, 2022. ISBN 9788595159495. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9788595159495/>>. Citado 3 vezes nas páginas 10, 11 e 14.
- SANTANA, M. F.; ZAMBOLIM, E. M.; CAIXETA, E. T.; ZAMBOLIM, L. Population genetic structure of the coffee pathogen hemileia vastatrix in minas gerais, brazil. **Tropical Plant Pathology**, Springer, v. 43, n. 5, p. 473–476, 2018. Citado na página 6.
- SANTOS, D. F. D. dos; FARIA, P. R. de; TRAVENÇOLO, B. A. N.; NASCIMENTO, M. Z. do. Influence of data augmentation strategies on the segmentation of oral histological images using fully convolutional neural networks. **Journal of Digital Imaging**, Springer Science and Business Media LLC, apr 2023. Citado na página 29.
- SILVA, M. d. C.; GUERRA-GUIMARÃES, L.; DINIZ, I.; LOUREIRO, A.; AZINHEIRA, H.; PEREIRA, A. P.; TAVARES, S.; BATISTA, D.; VÁRZEA, V. An overview of the mechanisms involved in coffee-hemileia vastatrix interactions: Plant and pathogen perspectives. **Agronomy**, v. 12, n. 2, 2022. ISSN 2073-4395. Disponível em: <<https://www.mdpi.com/2073-4395/12/2/326>>. Citado 2 vezes nas páginas 6 e 8.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014. Citado 2 vezes nas páginas 11 e 19.
- SUPARYANTO, T.; FIRMANSYAH, E.; CENGGORO, T. W.; SUDIGYO, D.; PARDAMEAN, B. Detecting Hemileia vastatrix using Vision AI as Supporting to Food Security for Smallholder Coffee Commodities. In: . IOP Publishing, 2022. v. 998, n. 1, p. 012044. Disponível em: <<https://dx.doi.org/10.1088/1755-1315/998/1/012044>>. Citado 2 vezes nas páginas 17 e 18.
- SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCHE, V.; RABINOVICH, A. Going deeper with convolutions. In: **2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Boston, MA, USA: IEEE, 2015. p. 1–9. Citado 3 vezes nas páginas 11, 12 e 19.
- VU, H. L.; NG, K. T. W.; RICHTER, A.; AN, C. Analysis of input set characteristics and variances on k-fold cross validation for a recurrent neural network model on waste disposal rate estimation. **Journal of Environmental Management**, v. 311, p. 114869, 2022. ISSN 0301-4797. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S030147972200442X>>. Citado na página 14.
- YEBASSE, M.; SHIMELIS, B.; WARKU, H.; KO, J.; CHEOI, K. J. Coffee disease visualization and classification. **Plants**, MDPI, v. 10, n. 6, p. 1257, 2021. ISSN 2223-7747. Citado 3 vezes nas páginas 6, 16 e 18.