

# Quality Assessment of Coffee Beans Using Convolutional Neural Networks

1<sup>st</sup> Marcus Vinicius Diniz dos Reis  
Universidade Federal de Viçosa - UFV  
Rio Paranaíba, Brasil  
marcus.v.reis@ufv.br

2<sup>nd</sup> Brenno Alves Silva  
Universidade Federal de Viçosa - UFV  
Rio Paranaíba, Brasil  
brenno.alves@ufv.br

**Abstract**—The classification of coffee beans into defect and premium categories is a critical task in ensuring quality control and standardization in the coffee industry. This work evaluates the performance of three deep learning architectures—AlexNet, ResNet-18, and ResNet-50—in addressing this challenge. The models were trained and tested on a dataset of coffee beans, called USK-Coffee, achieving promising results. AlexNet demonstrated the best balance between validation and test accuracy, with an F1-Score of 0.9232 and a recall of 0.9175. ResNet-50 achieved the highest validation accuracy (0.9750), but with a noticeable performance drop in the test set, suggesting room for optimization. ResNet-18, while simpler, achieved a recall of 0.925, excelling in identifying high-quality beans. These results highlight the potential of convolutional neural networks in automating coffee bean classification and emphasize the importance of fine-tuning and dataset diversity to enhance model generalization. This study contributes to the advancement of artificial intelligence in quality control for the coffee industry.

**Index Terms**—alexnet, resnet50, resnet18, coffee bean, CNN.

## I. INTRODUÇÃO

O mercado de café no Brasil é um dos mais importantes e influentes no mundo, sendo o maior exportador de café globalmente. Graças ao clima e solo favorável, o Brasil, mesmo com oscilações nas condições climáticas, sempre representa uma parte significativa da oferta de mercado, com o ano de 2022 gerando uma receita de exportação de mais de 50 bilhões de reais [1].

Apesar do clima e solos favoráveis, o café está suscetível a várias doenças causadas por vermes, fungos e bactérias, além de estresses abióticos, como geada, seca, e estresse por nutrientes, que alteram a qualidade do produto final [2] e que até hoje, em muitos casos, a qualidade do grão de café é identificada de forma manual, assim, com uma baixa precisão. Além disso, há o limite humano, um fator limitante para o número de grãos a serem analisados.

Com o desenvolvimento das tecnologias de inteligência artificial, é possível criar novas formas para a análise dos grãos de café, que podem ser uma grande ferramenta, principalmente para os produtores, como foi o trabalho de [3], usando *deep learning*, foi criado um aplicativo de celular para essa identificação, contribuindo para muitos agricultores. Assim, as fazendas terão um método com menos custo e mais eficiente.

No campo da inteligência artificial, as redes neurais são modelos computacionais inspirados no córtex visual de seres vivos e tem tido grande avanço no uso de redes neurais

artificiais [4]. Para o projeto de análise de qualidade de grãos de café, usamos as redes neurais convolucionais, um tipo de arquitetura de rede neural projetada para processar dados como imagens. Elas são amplamente utilizadas em tarefas de visão computacional, como reconhecimento de objetos, classificação, detecção de padrões e segmentação de imagens [5].

O café é uma das maiores *commodities* mais consumidas no mundo, com cada vez mais demanda por alta qualidade. O cheiro, forma e tamanho dos grãos são de importância para a qualidade e valor no mercado. Geralmente o processo de análise dos grãos é feito por seres humanos, que além de um bom tempo de estudos e experiência, ainda dá abertura para qualquer tipo de erro, além da velocidade da análise, onde um sistema seria mais rápido, sem pausas e com mais precisão [6].

O objetivo deste projeto é desenvolver testes para a classificação de grãos de café utilizando redes neurais, como Resnet18 [11], Resnet50 [11] e AlexNet [17]. Com a implementação dessas metodologias, buscamos aprimorar a acurácia e a precisão do processo de classificação dos grãos. Esperamos que os resultados obtenham benefícios para os agricultores e inspirem novas pesquisas e avanços na área de análise da qualidade do café.

## II. REFERENCIAL TEÓRICO

### A. Grãos de café

Cada tipo de grão de café possui características únicas de forma, textura e cor, que variam de acordo com o método de torra utilizado, além de apresentar um sabor característico [7]. Neste artigo, o foco será nos grãos de café Arábica, que podem ser classificados em diferentes categorias, como ilustrado na Figura 1.

1) *Grãos de Café Premium*: Os grãos *premium* são geralmente de cor verde-azulada. Além disso, os grãos *premium* tendem a ser mais arredondados e de maior tamanho, o que é um indicativo de um processamento adequado e criterioso [7].

2) *Grãos de Café Defeituosos*: Em contraste, os grãos defeituosos são aqueles que apresentam falhas no processamento ou que foram danificados por animais ou insetos. Esses grãos podem ser facilmente identificados pela sua forma irregular, quebradiça, oca e de tamanho desigual [7]. As imperfeições nos grãos defeituosos não apenas afetam a aparência visual,



Fig. 1. Exemplo das classes de grãos de café (Feito pelo autor).

mas também comprometem a qualidade do café produzido, impactando negativamente o sabor e aroma do produto final.

## B. Deep Learning

O *deep learning* (aprendizado profundo) é uma subcategoria do aprendizado de máquina, caracterizada pelo uso de redes neurais artificiais com múltiplas camadas de unidades de processamento interconectadas, chamadas de neurônios artificiais. Diferentemente dos métodos tradicionais, o *deep learning* aprende representações complexas dos dados em vários níveis de abstração sem a necessidade de engenharia manual de características. O treinamento de um modelo envolve a otimização dos pesos das conexões entre neurônios para que ele aprenda a tomar decisões com base nos dados de entrada [4].

1) *Redes Neurais Convolucionais (CNN)*: As redes neurais convolucionais (CNNs) são uma arquitetura de *deep learning* especialmente eficaz para dados estruturados em forma de grade, como imagens. Inspiradas no córtex visual, onde diferentes neurônios reagem a estímulos visuais específicos, as CNNs processam padrões visuais em vários níveis, como bordas e texturas, sendo amplamente utilizadas em reconhecimento de objetos e classificação de imagens [4].

As CNNs têm demonstrado resultados notáveis em visão computacional, com um processo de extração que começa identificando características simples em camadas iniciais e combinando-as em padrões mais complexos nas camadas intermediárias e finais, que então são processadas para permitir a classificação final dos objetos [5].

2) *Arquitetura Geral*: A estrutura de uma CNN é composta por várias camadas que desempenham funções específicas: entrada, extração de características e classificação. Inicialmente, os dados de entrada, como uma imagem, passam por camadas convolucionais que aplicam filtros (ou *kernels*) para detectar padrões visuais, seguidos de camadas de *pooling* que reduzem a dimensionalidade, mantendo as informações mais relevantes, como mostrado na Figura 2. Por fim, as características extraídas são achatadas em um vetor 1D pela camada *flatten* e passadas a uma rede totalmente conectada para a classificação final dos dados processados.

3) *Camadas Convolucionais*: As camadas convolucionais, essenciais na extração de características, utilizam filtros para detectar padrões específicos em regiões da imagem de entrada. Cada filtro possui um tamanho específico (ex.: 3x3, 5x5) e

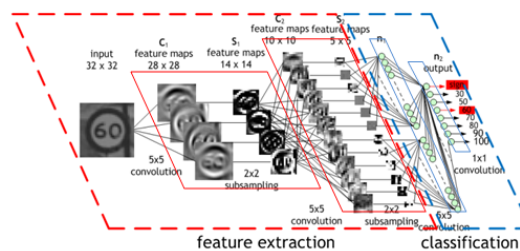


Fig. 2. Exemplo de uma arquitetura de Redes Neurais Convolucionais [15].

gera um *mapa de características* que representa um aspecto distinto da imagem, como bordas ou texturas. Os parâmetros de configuração, como o *stride* (passo de deslocamento) e o *padding* (preenchimento nas bordas), afetam o tamanho da saída da convolução, controlando a quantidade de informação espacial preservada [4].

4) *Camadas de Pooling*: As camadas de *pooling* têm como objetivo reduzir a quantidade de parâmetros e a complexidade computacional da rede. Um exemplo comum é o *max pooling*, que considera o valor máximo em uma região específica (ex.: 2x2) do mapa de características, como mostrado na Figura 3. Outro tipo é o *average pooling*, que calcula a média dos valores na mesma região, útil em situações que priorizam uma análise mais abrangente dos padrões da imagem [4].

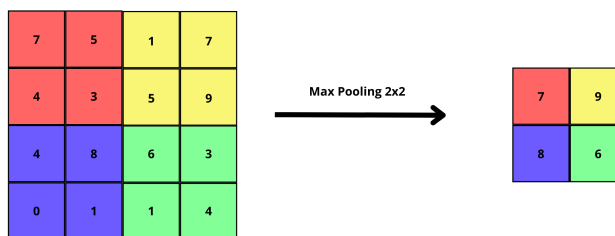


Fig. 3. Exemplo de um *Max Pooling* (Feito pelo autor).

5) *Camadas Totalmente Conectadas e Classificação*: Ao final da rede convolucional, os mapas de ativação (ou características) são achatados em um vetor 1D e passados a uma rede totalmente conectada para a classificação final. Diferente das camadas convolucionais e de *pooling*, a MLP conecta cada neurônio da camada anterior a cada neurônio da próxima camada, permitindo aprender representações mais complexas dos dados. Essas camadas geralmente são usadas para realizar a decisão final sobre a classe do objeto detectado [4].

## C. AlexNet

A AlexNet, é uma arquitetura pioneira de redes neurais convolucionais (CNNs) que revolucionou a área de visão computacional. Essa rede introduziu várias inovações, como o uso de ReLU (Rectified Linear Unit) como função de ativação, dropout para evitar overfitting, e treinamento em GPUs para acelerar o processamento. A arquitetura consiste em cinco

camadas convolucionais seguidas por três camadas totalmente conectadas, e é otimizada com o algoritmo de descida do gradiente estocástico com momentum [17].

#### D. ResNet

As redes ResNet, introduzidas por He et al. em 2015, revolucionaram a área de aprendizado profundo ao propor conexões residuais para superar o problema de degradação de desempenho em redes muito profundas. O conceito principal das conexões residuais é permitir que os gradientes fluam mais facilmente ao longo da rede, introduzindo conexões de atalho que ignoram uma ou mais camadas. Dentre as variantes, a ResNet-50 é uma arquitetura com 50 camadas, que utiliza blocos residuais profundos compostos por convoluções 1x1, 3x3 e 1x1, otimizados para reduzir a dimensionalidade e melhorar a eficiência computacional. Por outro lado, a ResNet-18 é uma variante mais leve, composta por 18 camadas e projetada para cenários onde há menor capacidade computacional ou menor quantidade de dados disponíveis. Ambas as arquiteturas alcançaram avanços significativos em tarefas de classificação, como no ImageNet, estabelecendo novos padrões de desempenho em benchmarks [11].

### III. TRABALHOS RELACIONADOS

O estudo de [8] avaliou a qualidade de café arábica utilizando métodos baseados em CNNs. Um dos métodos focou na análise da coloração dos grãos, demonstrando que essa é uma característica importante para a avaliação do café, identificando padrões para diferentes classes. A base de dados, criada pelos autores, contém 480 imagens de alta resolução (4928x3264), que passaram por um processo de corte, reduzindo a resolução para 678x678 para diminuir o processamento necessário. A CNN utilizada obteve uma acurácia de 81%. Os autores indicaram um plano de implementar inteligência artificial para auxiliar no reconhecimento da qualidade dos grãos de café em produção.

O trabalho de [9] centrou-se na identificação e classificação de doenças em plantas de café, também utilizando CNNs. Como diferencial, os autores utilizaram técnicas de aumento da base de dados, sistema multitarefa para estimar a gravidade das doenças, e compararam vários modelos de *deep learning*. O modelo ResNet [11] demonstrou o melhor desempenho, com uma acurácia de 95,24% ao usar métodos multitarefa.

Em [3], os autores propuseram um aplicativo para dispositivos móveis, utilizando principalmente as arquiteturas UNet [13] e PSPNet [14]. O objetivo foi desenvolver um sistema de segmentação e classificação de lesões em folhas de café para estimar a gravidade do estresse causado por agentes bióticos. O método de segmentação semântica permitiu identificar e delimitar as lesões nas folhas, e a classificação dos sintomas possibilitou a identificação do tipo de estresse biótico. O aplicativo, desenvolvido para *Android*, fornece uma ferramenta prática para especialistas e agricultores monitorarem o estresse biótico nas plantas de café.

No estudo [10], foi proposta uma abordagem de *deep learning* com etapas de pré-processamento de imagem, extração de

características e classificação. Foram utilizadas as arquiteturas GoogLeNet e ResNet, com uma técnica de filtragem gaussiana para aumentar o banco de dados. As características extraídas foram então classificadas usando Perceptrons de Múltiplas Camadas (MLPs) e classificadores de conjunto. O modelo alcançou uma precisão de teste de 99,08%, superando outros classificadores.

O trabalho de [6] propôs um modelo de *machine learning* para detecção de grãos defeituosos, visando melhorar a qualidade do café. Os autores utilizaram a técnica de *knowledge distillation* (KD), que transfere o aprendizado de um modelo mais complexo para um modelo mais simples, tornando-o mais leve. Embora o progresso do *deep learning* tenha sido significativo nos últimos anos, o estudo não apresentou confiabilidade suficiente. No entanto, o modelo foi capaz de classificar com precisão a qualidade dos grãos, proporcionando uma maneira de acesso facilitado a esse tipo de avaliação.

O artigo [12] desenvolveu um algoritmo de *deep learning* para detectar grãos defeituosos antes da torrefação, visando uma seleção de grãos de alta qualidade. Atualmente, a maioria das análises é realizada de forma manual, o que pode gerar resultados inconsistentes. A base de dados com 3621 amostras foi criada pelos próprios autores. O modelo, baseado na arquitetura AlexNet, foi desenvolvido em quatro fases: processamento de imagem, construção do modelo, avaliação e modificação do modelo. O estudo concluiu que o modelo apresentou uma precisão de 95,1% para detecção e classificação de grãos defeituosos, com acurácia de 100% ao focar apenas na detecção. Esse trabalho mostra grande potencial para automação na indústria de café.

### IV. METODOLOGIA

A metodologia adotada envolve a preparação e tratamento do banco de dados, com foco no aprimoramento das imagens e na correta distribuição dos dados para garantir a eficácia do treinamento do modelo e técnicas de pré-processamento, como ilustrado na Figura 4 enquanto o desempenho do modelo foi avaliado através de métricas adequadas.

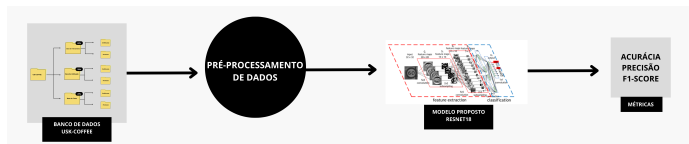


Fig. 4. Diagrama geral da Metodologia (Feito pelo autor).

#### A. Banco de dados

Neste estudo, utilizamos o banco de dados público USK-COFFEE, que disponibiliza imagens com resolução de 256x256 pixels. As classes trabalhadas são *defect* e *premium*, com cada classe contendo 2.000 imagens de grãos de café em formato PNG e codificação RGB.

### B. Arquiteturas de redes utilizadas

Para este trabalho, foram utilizadas as arquiteturas AlexNet, ResNet18 e ResNet50 devido à sua eficácia comprovada em tarefas de classificação de imagens. Essas arquiteturas foram adaptadas para atender ao escopo do problema, que envolve a classificação com base em entradas binárias.

As modificações realizadas concentraram-se nas últimas camadas das redes, responsáveis pela saída e pela interpretação das características extraídas. As camadas finais foram ajustadas para trabalhar com o formato binário.

Os modelos foram treinados e avaliados no conjunto de dados utilizando técnicas de ajuste fino, aproveitando a transferência de aprendizado de pesos pré-treinados. O desempenho de cada arquitetura foi comparado, analisando-se métricas como acurácia, eficiência no treinamento e capacidade de generalização para o problema em questão.

### C. Distribuição do banco de dados

Para a realização dos experimentos, o banco de dados foi dividido da seguinte forma: 60% foram utilizadas para treinamento, 20% para teste e 20% para validação, como ilustrado na Figura 5, com intuito de validar a generalização do modelo da melhor maneira.

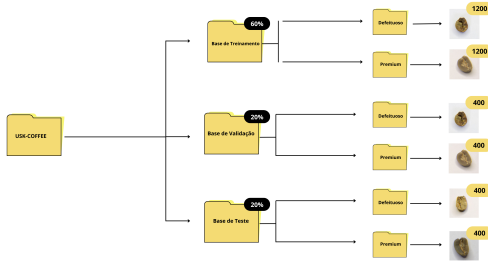


Fig. 5. Distribuição do banco de dados (Feito pelo autor).

### D. Métricas de Avaliação

A acurácia na validação foi usada para escolher o melhor conjunto de hiperparâmetros durante a otimização. Para avaliar o modelo final, consideramos o F1-score nos conjuntos de validação e teste. Comparar esses índices entre a validação e o teste nos ajuda a entender a capacidade do modelo de aplicar o que aprendeu durante o treinamento em novos dados.

**Acurácia:** Medida que reflete a proporção de previsões corretas em relação ao total de previsões feitas. É calculada considerando tanto verdadeiros positivos (TP) quanto verdadeiros negativos (TN) [16].

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN}$$

**Recall:** É a capacidade do modelo de identificar corretamente os casos positivos. Essa métrica é especialmente útil em cenários onde é mais importante evitar falsos negativos (FN) do que falsos positivos (FP), como no diagnóstico de doenças ou detecção de fraudes. [16].

$$\text{Recall} = \frac{TP}{TP + FN}$$

**F1-Score:** É a média da precisão e da sensibilidade. Essa métrica é valiosa em casos de classes desbalanceadas, pois oferece um único valor que considera tanto falsos positivos (FP) quanto falsos negativos (FN) [16].

$$\text{F1-Score} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

**Loss:** É uma função que mede o quão distante as previsões do modelo estão dos valores reais. Essa métrica é essencial para guiar o treinamento, ajustando os pesos do modelo para minimizar os erros e melhorar sua performance [16].

$$\text{Loss}_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log(\sigma(\hat{y}_i)) + (1 - y_i) \cdot \log(1 - \sigma(\hat{y}_i))]$$

## V. AMBIENTE DE HARDWARE

Os experimentos foram realizados em um PC equipado com um processador Ryzen 5 5500 3.6 GHz e 32 GB de RAM, rodando Windows 10 PRO, e equipado com uma GPU NVIDIA RTX 2060. Os experimentos foram desenvolvidos utilizando Python 3.8, PyTorch 2.3, torchvision 0.18 com CUDA Toolkit 11.8.

## VI. RESULTADOS

Nesta seção, discutimos os resultados obtidos a partir das diferentes arquiteturas de *deep learning*. Testamos três arquiteturas principais: Alexnet, Resnet18 e Resnet50:

### A. Resultados com Alexnet

Na primeira abordagem, utilizamos a arquitetura AlexNet, que alcançou uma acurácia de 0.9550 na validação e 0.9150 no teste. A *Loss* no conjunto de validação foi de 0.1824, evidenciando uma boa adaptação aos dados. O modelo obteve um F1-Score de 0.9232 e um recall de 0.9175. A matriz de confusão do conjunto de teste (Figura 6) confirma que a AlexNet generalizou bem para as classes.

### B. Resultados Resnet18

Com a arquitetura ResNet18, o modelo alcançou uma acurácia de 0.9400 na validação e 0.8762 no teste, demonstrando boa generalização para os dados de teste (Figura 7). O F1-Score foi de 0.8820, indicando um bom equilíbrio entre precisão e *recall*, com um recall de 0.925. A *Loss* de 0.3356 sugere que ainda há margem para melhorias na performance.

### C. Resultados com Resnet50

Na terceira abordagem, foi empregada a arquitetura ResNet-50, que alcançou uma acurácia de 0.9750 no conjunto de validação e 0.8812 no teste, indicando bom desempenho na generalização para dados desconhecidos. A *Loss* foi de 0.3723, sugerindo que o modelo conseguiu ajustar-se bem aos dados, mas ainda com espaço para refinamento. O F1-Score de 0.8948 reflete um equilíbrio sólido entre precisão

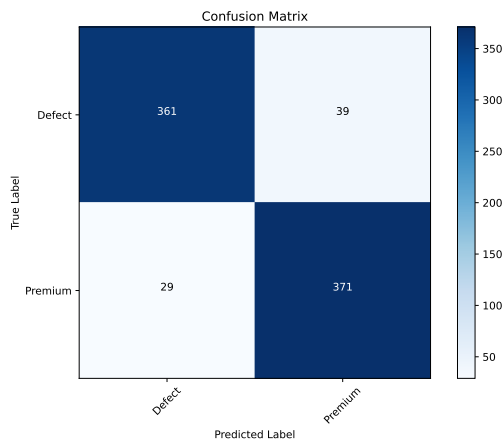


Fig. 6. Matriz de confusão Alexnet (Feito pelo autor).

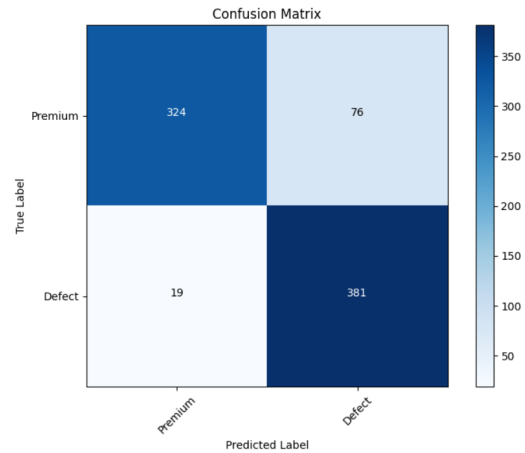


Fig. 8. Matriz de confusão Resnet50 (Feito pelo autor).

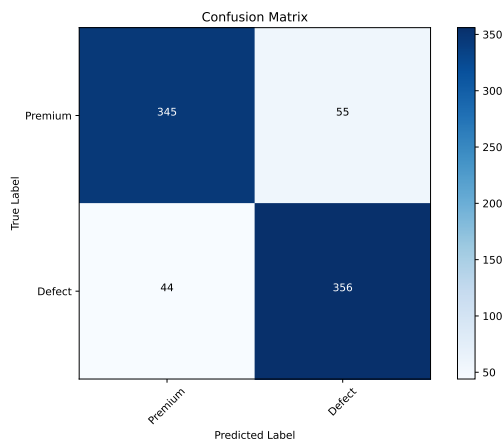


Fig. 7. Matriz de confusão Resnet18 (Feito pelo autor).

e *recall*, enquanto o recall de 0.915 destaca sua capacidade de identificar corretamente as classes. A matriz de confusão apresentada na Figura 8 detalha a distribuição dos acertos e erros do modelo no conjunto de testes.

#### D. Comparação Geral dos Resultados

Comparando as três abordagens, a AlexNet apresentou um equilíbrio interessante entre acurácia e *Loss*, com um F1-Score mais elevado e boa generalização, destacando-se como uma escolha sólida para o problema abordado. A ResNet-18, apesar de sua simplicidade e menor profundidade em relação à ResNet-50, apresentou resultados consistentes, mas com uma *Loss* um pouco mais elevada, sugerindo maior potencial para otimização. Já a ResNet-50 alcançou a maior acurácia na validação (0.9750), porém apresentou uma diferença maior de desempenho entre validação e teste, apontando para possíveis ajustes necessários no treinamento. Esses resultados indicam que arquiteturas mais profundas, como a ResNet-50, podem obter desempenhos superiores, mas exigem maior atenção em relação à regularização e ajuste fino dos hiperparâmetros.

TABLE I  
COMPARAÇÃO DOS RESULTADOS ENTRE AS ARQUITETURAS TESTADAS.

Métrica	AlexNet	ResNet-18	ResNet-50
Acurácia Val.	0.9550	0.9400	<b>0.9750</b>
Acurácia Teste	<b>0.9150</b>	0.8762	0.8812
Loss Val.	<b>0.1824</b>	0.3356	0.3723
F1-Score	<b>0.9232</b>	0.8820	0.8948
Recall	0.9175	<b>0.9250</b>	0.9150
Precisão	<b>0.9275</b>	0.8625	0.8100

## VII. CONCLUSÃO

Os resultados obtidos indicam que todas as arquiteturas testadas (AlexNet, ResNet-18 e ResNet-50) apresentam potencial para a tarefa de classificação de grãos de café entre as categorias defect e premium. A AlexNet destacou-se por apresentar um equilíbrio consistente entre acurácia nos conjuntos de validação e teste, obtendo um F1-Score de 0.9232, o mais alto entre as arquiteturas. Por outro lado, a ResNet-50 alcançou a maior acurácia na validação (0.9750), mas apresentou uma diferença notável de desempenho entre validação e teste, indicando a necessidade de ajustes finos para melhorar sua generalização.

Esses resultados reforçam a viabilidade do uso de redes neurais convolucionais para a classificação de grãos de café, com destaque para o desempenho consistente da AlexNet e o potencial da ResNet-50 em cenários mais otimizados. No entanto, os valores de *Loss* e as diferenças entre validação e teste sugerem que melhorias podem ser alcançadas com ajustes nos hiperparâmetros, técnicas de regularização e aumento da diversidade no conjunto de dados. A implementação de modelos de inteligência artificial como os analisados neste trabalho contribui significativamente para a automatização e padronização na classificação de grãos de café, promovendo maior eficiência e precisão no processo.

## REFERENCES

- [1] Thiago Farah Cavaton. (23 de janeiro de 2023). Faturamento das lavouras dos Cafés do Brasil alcança R\$ 56 bilhões em 2022. Acesso em: [www.embrapa.br](http://www.embrapa.br)

- [2] AgroPós (agosto de 2017). Principais Doenças do Café: Diagnose e Controle! Acesso em: <https://agropos.com.br/doencas-do-cafe/>
- [3] ESGARIO, José GM et al. 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. Disponível em: <https://doi.org/10.1016/j.inpa.2021.01.004>
- [4] WANI, M. A. et al. *Advances in deep learning*. [S.l.]: Springer, v. 57, 2020. Disponível em: <https://doi.org/10.1007/978-981-13-6794-6>
- [5] Awari (5 de março de 2023) Visão computacional: o que é e como funciona esta inteligência artificial Acesso em: [www.awari.com.br](http://www.awari.com.br)
- [6] WANG, Ping et al. Deep convolutional neural network for coffee bean inspection. *Sensors and Materials*, v. 33, n. 7, p. 2299-2310, 2021. Disponível em: <https://doi.org/10.18494/SAM.2021.3277>
- [7] FEBRIANA, Alifya et al. USK-COFFEE dataset: a multi-class green arabica coffee bean dataset for deep learning. In: 2022 IEEE international conference on cybernetics and computational intelligence (CyberneticsCom). IEEE, 2022. p. 469-473. Disponível em: [10.1109/CyberneticsCom55287.2022.9865489](https://doi.org/10.1109/CyberneticsCom55287.2022.9865489)
- [8] PRZYBYŁ, Krzysztof et al. Application of machine learning to assess the quality of food products—case study: Coffee bean. *Applied Sciences*, v. 13, n. 19, p. 10786, 2023. Disponível em: <https://doi.org/10.3390/app131910786>
- [9] ESGARIO, José GM; KROHLING, Renato A.; VENTURA, José A. Deep learning for classification and severity estimation of coffee leaf biotic stress. *Computers and Electronics in Agriculture*, v. 169, p. 105162, 2020. Disponível em: <https://doi.org/10.1016/j.compag.2019.105162>
- [10] ABUHAYI, Biniyam Mulugeta; MOSSA, Abdela Ahmed. Coffee disease classification using Convolutional Neural Network based on feature concatenation. *Informatics in Medicine Unlocked*, v. 39, p. 101245, 2023. Disponível em: <https://doi.org/10.1016/j.imu.2023.101245>
- [11] HE, Kaiming et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. p. 770-778.
- [12] HANG, Shyang-Jye; HUANG, Chien-Yu. Deep learning model for the inspection of coffee bean defects. *Applied Sciences*, v. 11, n. 17, p. 8226, 2021. Disponível em: <https://doi.org/10.3390/app11178226>
- [13] RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18. Springer International Publishing, 2015. p. 234-241. Disponível em: [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [14] ZHAO, Hengshuang et al. Pyramid scene parsing network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. p. 2881-2890. Disponível em: <https://doi.org/10.48550/arXiv.1612.01105>
- [15] Petsi (19 de abril de 2017) Redes neurais convolucionais aplicadas à análise de gestos considerando os problemas estudados na área de Estudos dos Gestos. Acesso em: <http://www.each.usp.br/petsi/?p=7325>
- [16] BELCIUG, S. *Artificial intelligence in cancer: diagnostic to tailored treatment*. Cambridge: Academic, 2020.
- [17] KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. ImageNet classification with deep convolutional neural networks. *Advances in neural information processing systems*, v. 25, 2012.