

Quality Assessment of Coffee Beans Using Convolutional Neural Networks with Wavelet and CLAHE Techniques

Marcus Vinicius Diniz dos Reis¹, Pedro Moises de Sousa²

¹Instituto de Ciências Exatas e Tecnológicas - Universidade Federal de Viçosa - Campus Rio Paranaíba (UFV-CRP)
Rodovia MG230, Km7, Caixa Postal 22 - 38.810-000 - Rio Paranaíba - MG - Brasil
marcus.v.reis@ufv.br, pedromoises@ufv.br

Abstract—This paper presents an analytical study comparing different filtering techniques applied to a Convolutional Neural Network (CNN) for coffee bean classification. The results demonstrated that the CLAHE (Contrast Limited Adaptive Histogram Equalization) filter achieved the highest performance, with an accuracy of 0.8875 on the test set. The findings indicate that applying filtering techniques can enhance the performance of the ResNet18 network. CLAHE's effectiveness is attributed to its ability to improve image details and contrast, leading to superior classification results. This study underscores the potential of advanced filtering methods to boost CNN performance in image classification tasks.

Index Terms—wavelet, clahe, 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 [2].

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 [3] 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 [4], 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 [5]. 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 [6].

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 [7].

O objetivo deste projeto é desenvolver testes para a classificação de grãos de café utilizando redes neurais, como Resnet18 [12], aplicando técnicas de filtragem Wavelet e CLAHE. 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 [8]. Neste artigo, o foco será nos grãos de café Arábica, que podem ser classificados em diferentes categorias, como ilustrado na Figura 1.



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

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 [8].

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 [8]. As imperfeições nos grãos defeituosos não apenas afetam a aparência visual, 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 [5].

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 [5].

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 [6].

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 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

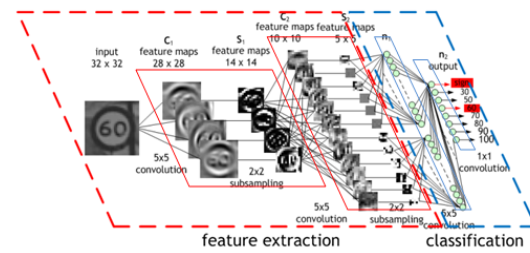


Fig. 2. Exemplo de uma arquitetura de Redes Neurais Convolucionais [16].

padding (preenchimento nas bordas), afetam o tamanho da saída da convolução, controlando a quantidade de informação espacial preservada [5].

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 [5].

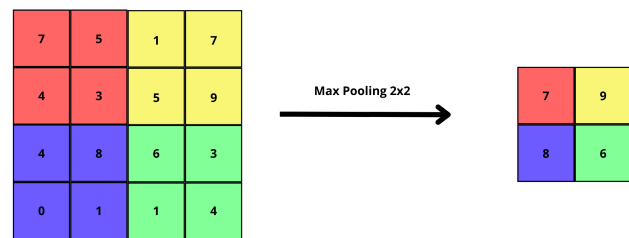


Fig. 3. Exemplo de um *Max Pooling* (Feito pelo autor).

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 [5].

C. Filtragem CLAHE

O CLAHE (*Contrast Limited Adaptive Histogram Equalization*) é uma técnica de aprimoramento de contraste usada em processamento de imagem, que ajusta o contraste da imagem através da equalização adaptativa do histograma [1]. O CLAHE limita a amplificação do contraste para evitar que o ruído da imagem seja excessivamente ampliado, o que é especialmente útil em imagens com áreas de contraste muito baixo. Essa técnica é aplicada em pequenas regiões da imagem, chamadas de *tiles*, e as imagens resultantes são

combinadas utilizando a interpolação bilinear, garantindo uma transição suave entre as regiões [1].

III. TRABALHOS RELACIONADOS

O estudo de [9] 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 [10] 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 [12] demonstrou o melhor desempenho, com uma acurácia de 95,24% ao usar métodos multitarefa.

Em [4], os autores propuseram um aplicativo para dispositivos móveis, utilizando principalmente as arquiteturas UNet [14] e PSPNet [15]. 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 [11], 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 [7] 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 [13] 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.

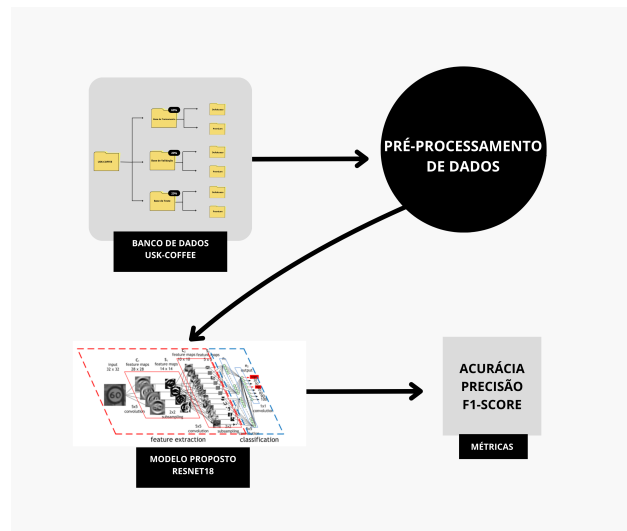


Fig. 4. Diagrama geral da Metodologia (Feito pelo autor).

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. Pré-processamento de imagens

A etapa de pré-processamento das imagens envolveu a aplicação de um recorte de imagem, retirando o fundo da imagem, como ilustrado na Figura 5, e de diferentes filtros à base de dados, com o objetivo de melhorar o desempenho do modelo. Para entender o impacto de cada tipo de filtragem no comportamento do modelo, foram testadas diversas abordagens, incluindo a filtragem wavelet, a filtragem de clahe, e o uso das imagens originais, sem nenhum tipo de filtragem.

Essa análise comparativa permitiu avaliar como cada técnica de pré-processamento influencia os resultados, contribuindo para a otimização do desempenho do modelo.

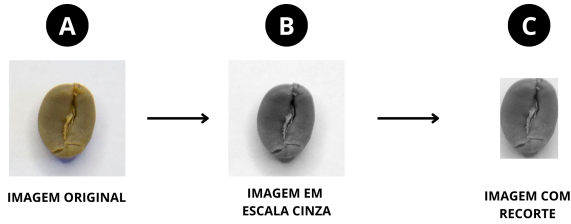


Fig. 5. Recorte de imagens em um grão de café (Feito pelo autor).

1) *Filtragem Wavelet*: A transformada Wavelet é uma técnica matemática usada para decompor sinais, como imagens, em componentes de diferentes frequências. Em processamento de imagens, ela permite analisar e processar detalhes em várias escalas, capturando tanto características globais quanto locais da imagem [17]. Essa abordagem é especialmente útil para capturar variações sutis na superfície dos grãos que podem ser indicativas de sua qualidade, como podemos notar na Figura 6.

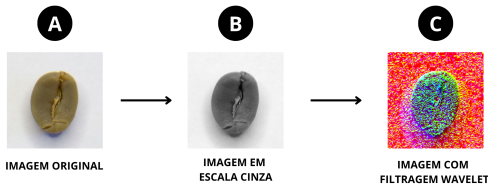


Fig. 6. Uso da Filtragem Wavelet em um grão de café (Feito pelo autor).

2) *Filtragem CLAHE*: Para este projeto, o uso da filtragem CLAHE visa melhorar a qualidade visual das imagens de entrada, destacando características texturais importantes que podem ser essenciais para a diferenciação entre grãos de qualidade variada, como podemos observar na Figura 7. Ao melhorar a visibilidade dos detalhes na imagem, o CLAHE pode potencialmente facilitar o trabalho da rede neural em aprender padrões relevantes.

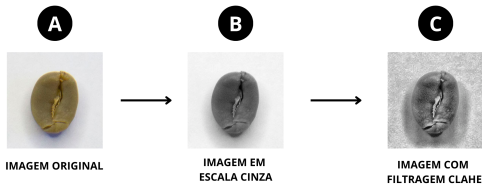


Fig. 7. Uso da Filtragem Clahe em um grão de café (Feito pelo autor).

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 8, com intuito de validar a generalização do modelo da melhor maneira.

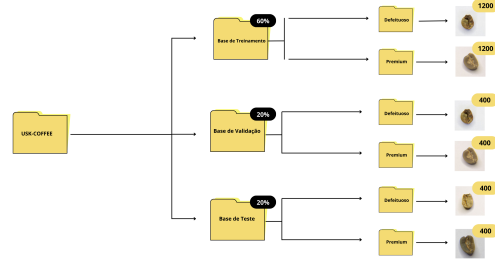


Fig. 8. Distribuição do banco de dados (Feito pelo autor).

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) [18].

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

Precisão: Mede a proporção de previsões positivas corretas em relação ao total de previsões positivas feitas. É útil para entender a eficácia do modelo ao classificar exemplos como positivos [18].

$$\text{Precisão} = \frac{TP}{TP + FP}$$

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) [18].

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

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 abordagens para pré-processamento de imagens e aplicação de técnicas de *deep learning*. Testamos três métodos principais: conversão para escala de cinza, aplicação do filtro *CLAHE* e utilização de *Wavelets*. A seguir, apresentamos e comparamos os resultados.

A. Resultados com CLAHE

Na primeira abordagem, aplicamos o filtro *CLAHE* para melhorar o contraste das imagens, especialmente em regiões de baixa iluminação. O modelo treinado com estas imagens obteve uma acurácia de 0.9038 na validação e 0.8875 no teste. A *Loss* no conjunto de validação foi de 0.3411, indicando uma boa adaptação do modelo aos dados. O melhor F1-Score obtido foi 0.8130. Na Figura 9, a matriz de confusão mostra que a filtragem *CLAHE* permitiu uma boa generalização para a classe *defect*, aumentando os verdadeiros negativos (TN).

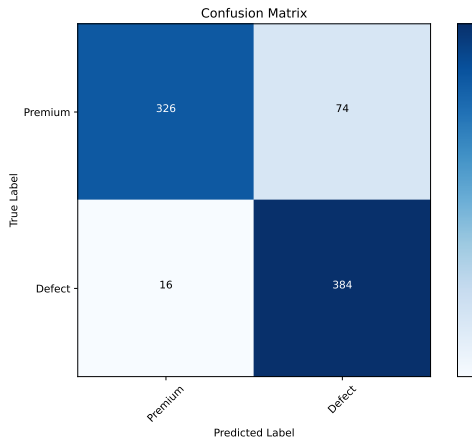


Fig. 9. Matriz de confusão CLAHE (Feito pelo autor).

B. Resultados Escala de Cinza

A segunda abordagem converteu as imagens para escala de cinza e aplicou recorte antes do treinamento. Essa técnica reduziu a complexidade das imagens, mantendo apenas informações de intensidade. O modelo alcançou uma acurácia de 0.8863 na validação e 0.8525 no teste, mostrando uma boa generalização para os dados de teste (Figura 10). O melhor F1-Score foi de 0.8517, com bom equilíbrio entre Precisão e *Recall*. A *Loss* foi de 0.5037, sugerindo que a performance poderia ser melhorada.

C. Resultados com Wavelet

A terceira abordagem utilizou o filtro *Wavelet* com a família db1, oferecendo uma análise mais detalhada dos padrões das imagens. O modelo alcançou uma acurácia de 0.8765 na validação e 0.7688 no teste, menor que as outras abordagens. A *Loss* foi a mais alta, 0.6284, indicando maior dificuldade de adaptação. O F1-Score também foi inferior, 0.7931, mostrando um desequilíbrio entre precisão e *recall*. A matriz de confusão na Figura 11 ilustra este desempenho.

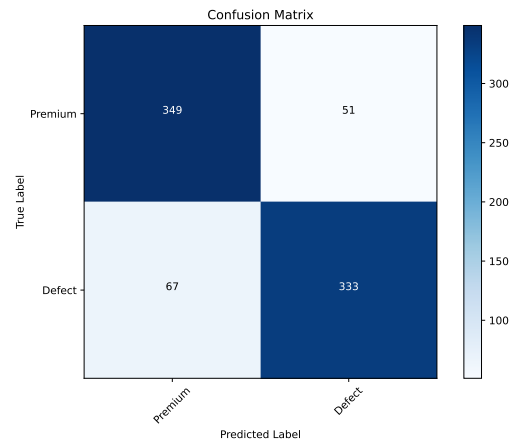


Fig. 10. Matriz de confusão Escala de Cinza (Feito pelo autor).

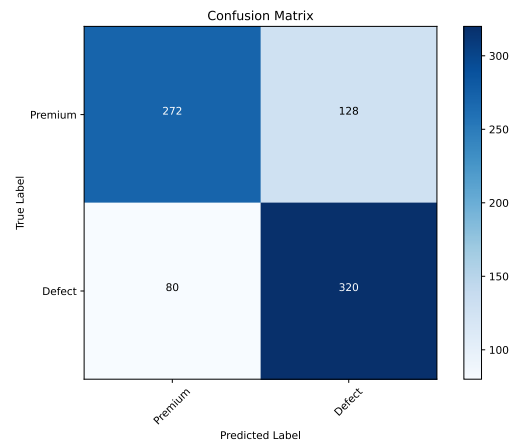


Fig. 11. Matriz de confusão Wavelet (Feito pelo autor).

D. Comparação Geral dos Resultados

A Tabela I compara as métricas das três abordagens com os resultados do estudo de [1]. Notamos que a técnica *CLAHE* obteve a melhor acurácia no teste e na validação, destacando-se como a abordagem mais eficaz para esta base de dados. Entretanto, a escala de cinza apresentou o maior F1-Score, sugerindo melhor equilíbrio entre as classes. A abordagem com *Wavelet* teve o desempenho inferior em todas as métricas, indicando necessidade de ajustes adicionais para esse tipo de aplicação.

TABLE I
COMPARAÇÃO DE RESULTADOS ENTRE DIFERENTES TÉCNICAS

Métrica	Escala Cinza	CLAHE	Wavelet	[8], 2022
Loss	0.5037	0.3411	0.6284	-
Validação Acurácia	0.8863	0.9038	0.8765	0.9506
F1-Score	0.8517	0.8130	0.7931	-
Teste Acurácia	0.8525	0.8875	0.7688	0.8113
Precisão	0.89	0.81	0.68	0.8114

VII. CONCLUSÃO

Neste estudo, avaliamos o desempenho da rede ResNet18 aplicada à classificação da qualidade de grãos de café, utilizando diferentes técnicas de filtragem de dados. Os experimentos foram conduzidos em um conjunto de dados abrangente, onde comparamos o impacto de filtros como o *Wavelet* e o *CLAHE*, em relação aos dados sem filtragem.

Os resultados obtidos mostraram que a aplicação de técnicas de filtragem podem melhorar o desempenho da rede ResNet18, especialmente em termos de acurácia e F1-score. Especificamente, o filtro *CLAHE* demonstrou ser mais eficaz em realçar características relevantes dos grãos de café, resultando em uma melhor taxa de acertos em comparação com os dados em escala cinza e com o filtro *Wavelet*.

Esses estudos reforçam a importância de pré-processamento de imagens em tarefas de classificação, sugerindo que a escolha adequada de técnicas de filtragem pode ser decisiva para melhorar a performance de redes neurais profundas em aplicações específicas.

Para trabalhos futuros pretende-se explorar o uso combinado de técnicas de filtragem e pré-processamento de maneira mais sofisticada, com o objetivo de potencialmente melhorar a acurácia e o desempenho geral do modelo.

REFERENCES

- [1] MUSA, Purnawarman; AL RAFI, Farid; LAMSANI, Missa. A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition. In: 2018 third international conference on informatics and computing (ICIC). IEEE, 2018. p. 1-6.
- [2] 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
- [3] AgroPós (agosto de 2017). Principais Doenças do Café: Diagnose e Controle! Acesso em: <https://agropos.com.br/doencas-do-cafe/>
- [4] 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>
- [5] 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>
- [6] Awari (5 de março de 2023) Visão computacional: o que é e como funciona esta inteligência artificial Acesso em: www.awari.com.br
- [7] 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>
- [8] 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
- [9] 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>
- [10] 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>
- [11] 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>
- [12] 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.
- [13] 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>
- [14] 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
- [15] 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>
- [16] 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>
- [17] JANSSEN, M. *Noise reduction by wavelet thresholding*. [S.l.]: Springer Science & Business Media, v. 161, 2012
- [18] BELCIUG, S. *Artificial intelligence in cancer: diagnostic to tailored treatment*. Cambridge: Academic, 2020.