

Aprimorando a classificação de grãos de café verde por meio do aprendizado profundo

Reinaldo Gonçalves Pereira Neto¹, Pedro Moisés de Sousa¹, Larissa Ferreira Rodrigues Moreira¹,
Pedro Ivo Vieira Good God², João Fernando Mari¹

¹Instituto de Ciências Exatas e Tecnológicas - Universidade Federal de Viçosa (UFV)

²Instituto de Ciências Agrárias - Universidade Federal de Viçosa (UFV)

{reinaldo.neto, pedromois, larissa.f.rodrigues, pedro.god, joaof.mari}@ufv.br

Resumo—O café é a bebida mais consumida no mundo, fundamental para o mercado brasileiro. Avaliar a qualidade dos grãos de café com base em características visuais é essencial para o valor de mercado. No entanto, a análise visual realizada por seres humanos tem limitações. Redes neurais profundas, especialmente redes neurais convolucionais (CNNs – *Convolutional Neural Networks*), oferecem uma solução promissora ao automatizar esse processo. Quatro arquiteturas de CNN foram avaliadas, com otimização de hiperparâmetros e aumento de dados. O EfficientNet B4 se destacou, alcançando 0,8844 de acurácia. Os resultados indicam o potencial do aprendizado profundo na avaliação da qualidade do café, auxiliando na classificação e garantia de qualidade e valor.

Palavras-chave—café verde, grão de café, aprendizado profundo, classificação, aumento de dados, otimização

I. INTRODUÇÃO

O café nos tempos atuais é visto como uma *commodity* global amplamente consumida, o que desperta crescentemente o interesse da população mundial [1]. O Brasil é destaque como o maior produtor e exportador de grãos de café verde, tornando o café uma das principais *commodities* em seu mercado interno [2]. A avaliação da qualidade do café é essencial, com características como cor, morfologia, forma e tamanho influenciando seu valor de mercado [1].

Redes neurais profundas, como CNNs, têm potencial para automatizar a análise de grãos de café, podendo superar especialistas humanos, que estão propensos a erros devido à fadiga e a subjetividade do processo. Quatro arquiteturas de CNN foram avaliadas com otimização de hiperparâmetros e aumento de dados. Os resultados são considerados os melhores até o momento para o conjunto de imagens avaliado, buscando contribuir para pesquisas futuras na avaliação da qualidade do café.

II. MATERIAL E MÉTODOS

A. Conjunto de Dados

Para os experimentos, utilizamos o conjunto de dados USK-Coffee [3]¹. O USK-Coffee é composto por 8.000 imagens de grãos de café verde distribuídas em quatro classes: *peaberry*, *longberry*, *premium* e *defect*. As imagens têm 256 × 256 pixels e estão igualmente distribuídas entre as classes, com 2.000 imagens para cada uma. Amostras de cada classe podem ser visualizadas na Figura 1.

¹<https://comvis.unsyiah.ac.id/usk-coffee/>

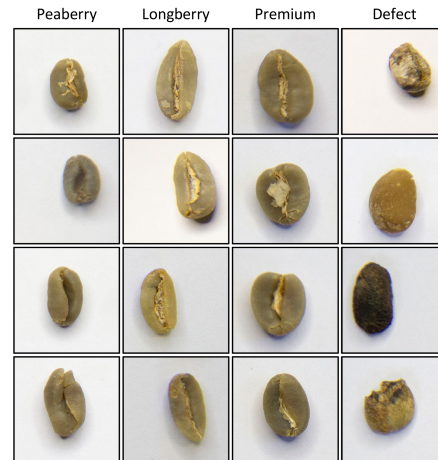


Fig. 1. Amostras de cada classe do conjunto de dados USK-Coffee

B. Arquiteturas

Avaliamos quatro arquiteturas de CNN que foram escolhidas devido ao seu sucesso em tarefas anteriores de classificação de imagens aplicadas na agricultura [3], [4]. A AlexNet foi a primeira CNN a vencer a competição ILSVRC em 2012, composta por camadas de convolução e *pooling*. A ResNet aborda dissipação dos gradientes em redes profundas com conexões residuais, com uma variante chamada ResNet-50 contendo 50 camadas de convolução. A MobileNet é projetada para dispositivos móveis, com diferentes versões introduzindo inovações como convoluções *depth-wise* e estrutura residual invertida. A EfficientNet, uma arquitetura eficiente de CNN, equilibra desempenho com eficiência computacional usando escalonamento, blocos eficientes e *pooling* global. Oferece variantes de B0 a B7 e é popular na classificação de imagens, sendo usada a variante EfficientNet B4 neste trabalho.

C. Projeto do Experimento

O conjunto de dados USK-Coffee é fornecido dividido em conjuntos de treinamento, validação e teste, e a mesma divisão foi mantida para comparação justa com resultados na literatura. Usamos modelos pré-treinados com o conjunto de dados do ImageNet [5] que foram ajustados usando o otimizador Adam, com redução da taxa de aprendizado após

10 épocas sem melhora na perda de validação e interrupção após 21 épocas sem melhora na perda de validação (parada antecipada). O tamanho do lote (*batch size* - BS) e a taxa de aprendizado (*learning rate* - LR) foram otimizados usando *grid-search*, com os intervalos de busca {16, 32, 64, 128} para o BS e {0.01, 0.001, 0.0001, 0.00001} para o LR. Não houve aumento de dados durante a otimização, mas após, os modelos foram treinados com aumento de dados, incluindo espelhamento horizontal, rotação, corte redimensionado, ajuste de cor e remoção aleatória. As transformações de imagens para validação e teste são as mesmas usadas na otimização de hiperparâmetros.

D. Avaliação do Modelo

A acurácia de validação foi considerada para selecionar o melhor conjunto de hiperparâmetros durante a otimização de hiperparâmetros. Para avaliar o modelo final, treinado com e sem aumento de dados, utilizamos, além da acurácia, a precisão, o *recall* e o *F1-Score* para os conjuntos de validação e teste. Ao comparar os índices entre os conjuntos de validação e teste, é possível acessar a capacidade do modelo de extrapolar o conhecimento adquirido durante o treinamento para dados desconhecidos.

III. RESULTADOS E DISCUSSÃO

Considerando os hiperparâmetros obtidos por meio da estratégia de otimização de busca em grade descrita na Seção II-C, treinamos novamente os modelos sem e com operações de aumento de dados.

TABELA I
RESULTADOS DO EXPERIMENTO SEM E COM AUMENTO DE DADOS

Estratégia	Arquitetura	Acc. Val.	Acc. Teste	Épocas
Sem aumento de dados	AlexNet	0.9169	0.8663	13
	ResNet-50	0.9475	0.8663	21
	MobileNet V3	0.9606	0.8506	50
	EfficientNet B4	0.9469	0.8588	24
Com aumento de dados	AlexNet	0.8788	0.8581	13
	ResNet-50	0.9644	0.8694	41
	MobileNet V3	0.9394	0.8650	16
	EfficientNet B4	0.9563	0.8844	30

Visualizamos os valores de acurácia na Tabela I para melhor compreensão dos resultados apresentados, onde é possível comparar o desempenho entre os modelos treinados nos conjuntos de validação e teste sem e com aumento de dados. Considerando o conjunto de teste, nosso melhor resultado ao treinar os modelos sem aumento de dados foi alcançado por AlexNet e ResNet-50, com acurácia de 0,8663. Ao treinar os modelos com aumento de dados, EfficientNet B4 obteve a melhor acurácia no teste, com 0,8844.

É interessante observar que AlexNet e ResNet-50 alcançaram o mesmo valor de acurácia no teste sem usar o aumento de dados, mas valores diferentes usando o aumento de dados (0,8581 e 0,8694, respectivamente) dando uma ligeira vantagem para ResNet-50 (cerca de 1,13%). Uma possível

conclusão é que a ResNet-50 tem uma arquitetura mais profunda e complexa do que a AlexNet [6]–[8], o que permite que ela aprenda melhor os padrões das imagens e tenha uma maior capacidade de generalização de dados. A rede AlexNet, sendo mais simples, provavelmente se beneficia menos do aumento de dados, pois já atinge um limite de desempenho com os dados originais.

Por fim, nossas descobertas indicam que o melhor resultado obtido para o conjunto de dados USK-Coffee foi uma acurácia de 88,44% alcançada com o EfficientNet B4 com aumento de dados, superando o desempenho da técnica publicada anteriormente (88,31% em Febriana et al. [3]), que havia sido o melhor até agora.

IV. CONCLUSÕES

Neste trabalho, comparamos arquiteturas de aprendizado profundo para classificar grãos de café verde com base em imagens, usando modelos pré-treinados do ImageNet e otimização de hiperparâmetros. Apenas o AlexNet não se beneficiou do aumento de dados, enquanto os outros modelos tiveram desempenho melhor com ele, com destaque para o EfficientNet B4, que teve a melhor acurácia no teste com aumento de dados (0,8844), mas sem aumento de dados, foi somente o terceiro melhor, superando apenas o MobileNet V3. Este estudo mostra o potencial do aprendizado profundo na avaliação da qualidade do café e sugere aplicações futuras na cadeia de produção de café, com possíveis melhorias, como testar mais arquiteturas e estratégias de treinamento, bem como explorar conjuntos de dados adicionais.

AGRADECIMENTOS

Este estudo foi financiado em parte pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código Financeiro 001. R. G. Pereira Neto foi bolsista do PIBIC/FAPEMIG.

REFERÊNCIAS

- [1] M. García, J. E. Candelo-Becerra, and F. E. Hoyos, “Quality and defect inspection of green coffee beans using a computer vision system,” *applied sciences*, vol. 9, no. 19, p. 4195, 2019.
- [2] Y. M. Guimarães, J. H. P. P. Eustachio, W. Leal Filho, L. F. Martinez, M. R. do Valle, and A. C. F. Caldana, “Drivers and barriers in sustainable supply chains: The case of the brazilian coffee industry,” *Sustainable Production and Consumption*, vol. 34, pp. 42–54, 2022.
- [3] A. Febriana, K. Muchtar, R. Dawood, and C.-Y. Lin, “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)*, pp. 469–473, IEEE, 2022.
- [4] E. L. da Rocha, L. Rodrigues, and J. F. Mari, “Maize leaf disease classification using convolutional neural networks and hyperparameter optimization,” in *Anais do XVI Workshop de Visão Computacional*, pp. 104–110, SBC, 2020.
- [5] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.
- [6] R. Wightman, H. Touvron, and H. Jégou, “Resnet strikes back: An improved training procedure in timm,” 10 2021.
- [7] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, vol. 25, 2012.