

Colher ou não colher? Uma rede neural convolucional para apoiar cafeicultores na decisão da colheita

Anagê C. Mundim Filho¹, Cleyton B. Alvarenga² e Murillo G. Carneiro¹

¹Faculdade de Computação – Universidade Federal de Uberlândia (UFU)

²Instituto de Ciências Agrárias – Universidade Federal de Uberlândia (UFU)

{anage.mundim, cleytonalvarenga, mgcarneiro}@ufu.br

Abstract. *In this article, we investigate deep learning models to assist farmers in the decision-making of the coffee harvest. This problem is barely explored in the literature and is very difficult due to the fruits ripening unevenly in the plant. In this sense, we collected thousands of images of coffee trees which were later classified by specialists in the field. In addition, we increased the database using a data augmentation process. Three architectures of convolutional neural networks were trained and had their predictive performance compared. The results show that our best model achieved an accuracy of 93 % and F1 of 92 % in a separated test set of images.*

Resumo. *Nesse artigo é feito um estudo de modelos de aprendizado profundo para auxiliar os agricultores na melhor tomada de decisão da colheita dos cafeeiros. Este é um problema pouco explorado na literatura e se torna difícil por conta dos frutos amadurecerem de maneira desuniforme na planta. Nesse sentido, coletamos milhares de imagens de cafeeiros as quais foram posteriormente classificadas por especialistas da área, Além disso, triplicamos a base de dados utilizando o processo de aumento de dados. Três arquiteturas de redes neurais convolucionais foram treinadas e tiveram seu desempenho preditivo comparado. Os resultados mostram que nosso melhor modelo obteve uma acurácia de 93% e F1 de 92% em um conjunto separado de imagens de teste.*

1. Introdução

O território brasileiro equivale a 850,3 milhões de hectares. A área explorada com pastagens, cultivo de lavouras e florestas plantadas é de 256,8 milhões de hectares, atividades que ocupam em torno de 30,2% do território nacional. Em relação ao cultivo de lavouras, o café possui grande relevância, sendo o país o maior produtor e exportador do grão, com área plantada de 2,2 milhões de hectares, e o segundo maior consumidor a nível mundial. A produção da safra em 2020 ultrapassou 63 milhões de sacas (60kg), um novo recorde [CONAB 2020].

Apesar do grande progresso em maquinário, irrigação, sistemas de monitoramento dos nutrientes do solo e das plantas, entre outros, muitos aspectos relacionados à tomada de decisão ainda são completamente dependentes da percepção e conhecimento do agricultor [Alves et al. 2015]. O problema abordado nesse artigo contempla o processo de tomada de decisão relacionado à colheita (ou não) dos frutos de um cafeeiro, uma vez que os métodos utilizados atualmente não conseguem fornecer uma resposta precisa sobre o

momento mais adequado para isso, onde a recomendação é realizar a colheita quando o percentual de frutos verdes esteja em proporção menor do que 20% quando considerando todos os frutos da planta [Alves et al. 2015]. Consequentemente, as diversas floradas tornam a determinação do momento ideal para colheita ainda mais complicado em virtude da desuniformidade da maturação dos frutos [Budzinski et al. 2005], a qual é exemplificada pela Figura 1 que evidencia a presença de frutos em diferentes estádios de maturação.



Figura 1. Exemplos da variação de frutos na mesma planta

Diante do exposto, o objetivo desse trabalho é desenvolver um modelo de classificação capaz de caracterizar os estádios de maturação dos frutos de modo a auxiliar os cafeicultores na tomada de decisão relacionada ao momento de colheita do café. Nesse sentido, as principais contribuições do estudo são: 1) a construção de uma base de dados com milhares imagens de plantas de café e seus respectivos estádios de maturação anotados por especialistas; 2) aprendizado de um modelo de RNC baseado na arquitetura MobileNet com alto desempenho preditivo (93% de acurácia e 92% de F1) atestado por um conjunto de dados externo, também anotado por especialistas;

2. Materiais e Métodos

Esta seção descreve as principais etapas conduzidas para a realização desse trabalho, as quais podem ser resumidas no fluxograma apresentado pela Figura 2. O primeiro passo do projeto, denotado por "Entrada" na figura, foi o desenvolvimento de uma metodologia para orientação sobre a captura das imagens bem como para suas anotações. O processo de anotação consistiu na classificação das imagens por cinco especialistas em relação ao estágio de maturação da planta e à colheita. Ao término, obtivemos 1000 imagens anotadas que foram usadas para treinamento, validação e teste das redes neurais convolucionais. Para a etapa seguinte, denotada por "Pré-processamento" na figura, avaliamos a base original das imagens em relação à uma nova base gerada a partir das seguintes estratégias de aumento de dados (*data augmentation*): *flip* horizontal, vertical e a combinação dos dois, as quais são exemplificadas pela Figura 3. Uma vez que o modelo aprendido deverá ser utilizado pelos cafeicultores usando dispositivos móveis (e.g., celular ou tablet), o mesmo não pode demandar muitos recursos computacionais. Dessa forma, na terceira etapa (denotada por "Aprendizado" na figura), foram consideradas arquiteturas de redes neurais convolucionais voltadas para ambientes *mobile*, a saber: MobileNet [Howard et al. 2017], MobileNet2 [Sandler et al. 2018] e NASNet-Mobile [Zoph et al. 2018], as quais são arquiteturas concebidas para otimizar o número de operações. Além do treinamento com-

pleto dos modelos, nessa etapa também foi avaliado o uso de transferência de aprendizado (*transfer learning*), carregando camadas convolucionais pré-treinadas na ImageNET [Deng et al. 2009]. Em seguida, os algoritmos foram treinados sobre as imagens considerando o problema binário de Colher ou Não colher uma planta de café. Para tanto, imagens anotadas nos estádios verde e verde-cana foram atribuídas para a classe Não colher enquanto aquelas anotadas como cereja, passa e seco para a classe Colher.

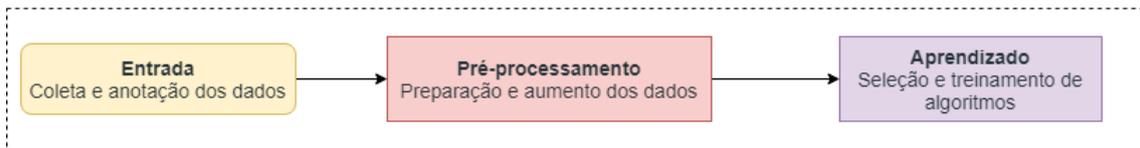


Figura 2. Fluxograma do projeto

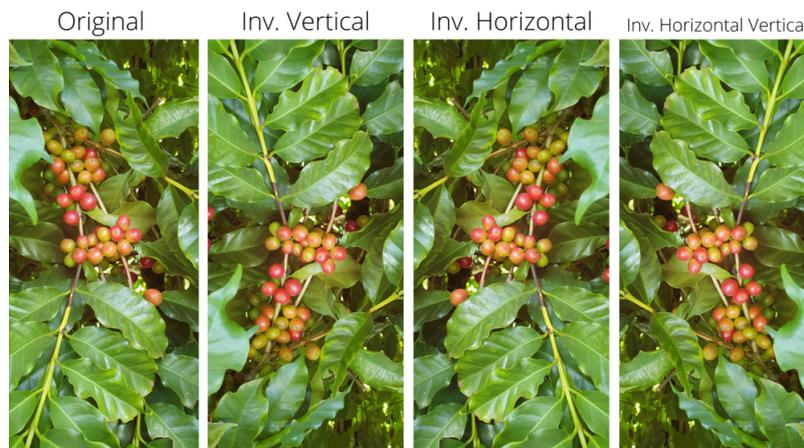


Figura 3. Imagens geradas como o aumento do banco de dados

3. Avaliação dos modelos aprendidos e Resultados

As três arquiteturas de redes neurais convolucionais possuem vários parâmetros. Nesse sentido, foram realizados testes empíricos a partir dos quais foram definidos os seguintes parâmetros:

- Entrada de arquivos em 224x224.
- 15 Épocas de treino, 55 Passos por época e com tamanho do batch de 64.
- Otimizador: Adam learning rate = 0.0001
- 3040 (Aumento de dados) imagens para treino.
- 760 (Original) imagens para validação.
- Foram usadas 150 imagens para o teste (75 de cada classe) .

A Tabela 1 apresenta o resultado das arquiteturas do estado-da-arte após o seu treinamento sobre a base original e sobre a base de dados resultante das estratégias de aumento de dados. Analisando os resultados da tabela, observa-se que o melhor desempenho preditivo foi da MobileNet, seguido pela MobileNetV2. Por outro lado, a NASNetMobile se mostrou inferior às outras duas arquiteturas em ambos os testes. Importante observar ainda que as estratégias de aumento de dados melhoraram o desempenho de todas as arquiteturas investigadas.

Tabela 1. Comparação dos modelos e desempenhos em duas bases de dados.

Modelos	Original			Aumento de dados		
	Ac.(%)	F1(%)	Tam.(KB)	Ac.(%)	F1(%)	Tam.(KB)
MobileNet	0.90	0.89	23.079	0.93	0.92	20.489
MobileNetV2	0.86	0.86	20.483	0.92	0.91	20.280
NASNetMobile	0.74	0.78	28.709	0.87	0.88	28.709

4. Conclusão

Este artigo retrata os métodos e resultados obtidos em nossa investigação para desenvolver um sistema capaz de auxiliar cafeicultores na escolha do momento adequado para a colheita dos frutos. Para tanto, foi construída uma base de dados com milhares de imagens de cafeeiros, as quais tiveram seus estádios de maturação anotados por especialistas. Sobre esses dados foram treinadas e avaliadas três importantes arquiteturas de redes neurais convolucionais, com a MobileNet retornando resultados muito promissores (93% de acurácia e 92% de F1). Os próximos passos da pesquisa incluem a migração desse modelo aprendido para um aplicativo móvel, a fim de que possa ser avaliado e testado por especialistas e cafeicultores no campo. Também pretendemos melhorar nosso desempenho considerando um conjunto maior de parâmetros de treinamento para as arquiteturas. Além disso, ainda existem outras arquiteturas a serem exploradas como redes neurais convolucionais por região (RCNN), a qual acreditamos que poderá ajudar na classificação direta dos frutos da planta.

Referências

- Alves, E. A., Costa, J. N., and Santos, J. C. F. (2015). Procedimentos de colheita do café. In Marcolan, AL; Espindula, M., editor, *Café na Amazônia*, pages 345–358.
- Budzinski, I. G., Cação, S., Carneiro, C. E., Pereira, L. F. P., and Vieira, L. G. E. (2005). Análise de genes expressos durante estádios finais da maturação de frutos de café. In *Simpósio de Pesquisa dos Cafés do Brasil*, pages 1–5.
- CONAB (2020). Acompanhamento da safra brasileira de café, v. 6 – Safra 2020, n.4 - Quarto levantamento.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255. Ieee.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, pages 1–7.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 4510–4520.
- Zoph, B., Vasudevan, V., Shlens, J., and Le, Q. V. (2018). Learning transferable architectures for scalable image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 8697–8710.