

Aplicação de Redes Neurais Convolucionais para Classificação de Imagens de Estágios de Maturação da Banana Prata Catarina

Rafaela S. Francisco¹, Gabriel de S. G. Pedroso¹, Thiago M. Ventura¹

¹Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)
Cuiabá – MT – Brasil

rafaela.souzaf.rsfsf@gmail.com, gabriel.de.s.g.pedroso@gmail.com,
thiago@ic.ufmt.br

Abstract. Bananas play a significant role in fruit farming and, to reduce losses during the post-harvest process and extend their storage life, it is important to identify their ripeness levels, mainly because they have a high rate of deterioration. Artificial intelligence techniques can be applied for this purpose. This work addresses the construction of artificial neural network models for classifying the ripening stages of bananas, using a dataset of images of Bananas Prata Catarina. Data augmentation and convolutional networks were applied to achieve better results.

Resumo. A banana desempenha um papel significativo na fruticultura e, para reduzir as perdas durante o processo pós-colheita e prolongar a vida útil de armazenamento, é importante identificar os seus níveis de maturação, principalmente por apresentarem uma alta taxa de deterioração. Técnicas de inteligência artificial podem ser aplicadas para este fim. Este trabalho aborda a construção de modelos de redes neurais artificiais para a classificação de estágios de maturação das bananas, utilizando-se como conjunto de imagens de Bananas Prata Catarina. Foram aplicadas técnicas de data augmentation e redes convolucionais para alcançar melhores resultados.

1. Introdução

A indústria da banana é uma das maiores do mundo em termos de comércio agrícola, sendo uma das frutas mais cultivadas e consumidas. Portanto, compreender as melhores práticas de cultivo, colheita, transporte e armazenamento é crucial para garantir a segurança alimentar global. Somente em 2023, o Brasil exportou 56,2 mil toneladas de bananas (Duarte e Barbieri, 2024), sendo que a fruta tem grande impacto no setor do agronegócio, setor responsável por 49% das exportações no Brasil (Ministério da Agricultura e Pecuária, 2024). Devido à importância e ao impacto da cadeia da banana, vários trabalhos foram desenvolvidos para resolver os desafios enfrentados pela cadeia de abastecimento da banana, de forma a aumentar a competitividade e a eficiência dos processos de produção agrícola.

Um desses desafios está relacionado a maturidade. A maturação é um tópico importante na colheita da banana devido ao seu impacto na qualidade e no preço da fruta (Mazen e Nashat, 2019). Os agricultores devem identificar os níveis de maturação das bananas para reduzir as perdas durante o processo pós-colheita e prolongar a vida útil de

armazenamento, pois as bananas apresentam uma alta taxa de deterioração (Siddiq et al., 2020).

No quesito maturidade das bananas, utilizar modelos de Inteligência Artificial (IA) no processo de produção para se compreender melhor os diferentes estágios de maturação permitem selecionar o momento ideal para a colheita. Isso assegura que o produto final seja de alta qualidade, satisfazendo as expectativas dos consumidores e contribuindo para a reputação da marca. Sabilla et al. (2019) trabalharam com 3 diferentes níveis de maturação: não madura, para madurar e madura demais. Já Mazen e Nashat (2019) mantiveram a classe madura demais, porém trabalharam com outros 3 estágios anteriores: verde, amarelado e meio maduro. Chen et al. (2018) também definiram quatro estágios de maturação, no entanto para cada estágio definiram critérios, em que não madura era verde; meio madura era verde com amarela; totalmente madura era amarela; e madura demais era amarela com marrom. Zhang et al. (2018) treinaram os modelos utilizando 7 e 14 estágios de maturação, com melhores resultados para 7 classes, medidos através da precisão e recall. Adebayo et al. (2017) também utilizam 7 estágios de maturação, no entanto, o trabalho mediu os estágios de maturação a partir de luzes de laser com cinco diodos de comprimento 532nm, 660nm, 785nm, 830nm e 1060nm.

O uso de técnicas e modelos de IA para compreender os estados de maturação da banana é fundamental para a promoção de processos mais eficientes, redução de desperdícios, melhores práticas agrícolas e uma indústria de alimentos mais sustentável. Desta forma, este trabalho tem como objetivo a implementação de técnicas de Redes Neurais Artificiais para o desenvolvimento de um modelo que possa determinar qual o estado de maturação em que uma dada banana está a partir da sua imagem, considerando situações reais em que o modelo poderá ser utilizado e sem a necessidade de equipamentos específicos que não seja a câmera de um dispositivo móvel.

2. Metodologia

Este trabalho utilizou como fonte de dados o conjunto de imagens de Banana Prata Catarina disponibilizado em Martins Neto et al. (2023). O repositório é o primeiro conjunto de dados brasileiro de acesso público com imagens de bananas rotuladas em 8 estágios de maturação. O conjunto de dados conta com 1.000 imagens com resolução que variavam de 2248×4000 a 3120×4160, e foram classificadas em 8 estágios de maturação.

A determinação dos estágios seguiu as normas de classificação de banana do Programa Brasileiro para a Modernização da Horticultura & Produção Integrada de Frutas (CEAGESP, 2006). Os 8 estágios foram separados de acordo com a coloração da casca da banana: completamente verde, verde com traços amarelos, mais verde do que amarelo, mais amarelo do que verde, amarelo com as pontas verdes, amarelo, amarelo salpicado com pontos marrons e amarela com casa muito marrom. Na Figura 1 é possível ver exemplos dos oito graus de maturação da banana.

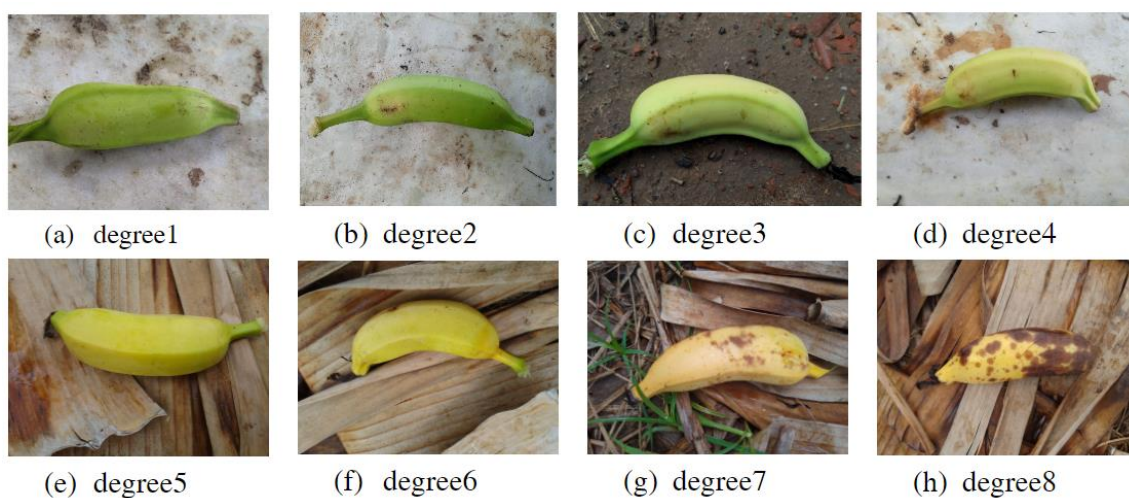


Figura 1. Exemplos dos 8 graus de maturação da banana, desde a coloração (a) “completamente verde” até (h) “amarela com casca muito marrom (Martins Neto et al., 2023).

Trabalhos correlatos citados na seção anterior mostravam diferentes categorizações de maturação. Assim, para uma melhor comparação dos resultados, além da separação dos estágios em 8 classes, foram testados a separação em 5 classes e em apenas 2 classes. Para o teste com 5 classes, houve a separação em bananas verdes (grau 1 e 2), verdes próximas a amarelo (grau 3 e 4), amarelas (grau 5 e 6) e amarelas pouco (grau 7) e muito salpicadas (grau 8). E para o teste de apenas 2 classes, houve a separação em bananas verdes (grau 1 ao 4) e bananas amarelas (grau 5 ao 8).

Em se tratando de pré-processamento, foi gerado uma variação dos dados com a aplicação de data augmentation (DA), técnica que consiste em gerar dados sintéticos a partir de dados existentes, aumentando a quantidade final de dados (Maharana et al., 2022). Foram utilizadas 4 operações diferentes: flip horizontal, flip vertical, rotação e zoom, com probabilidade de 50%, 50%, 70% e 50% respectivamente da operação ocorrer para cada imagem. Após a técnica, foram mantidas as mesmas imagens para cada uma das outras etapas com 8, 5 ou 2 classes, o que gerou 150 imagens a mais para o treinamento. As três primeiras operações foram escolhidas por não alterarem a estrutura da banana, já que elas podem ser encontradas na natureza. A ação de alteração de zoom foi feita para simular as diferentes formas de realizar a captura da imagem.

Por fim, foram criados modelos do tipo Redes Neurais Convolucionais (Convolutional Neural Network - CNN). Para tanto, foi utilizada a linguagem de programação Python e, principalmente, as bibliotecas Keras para construção dos modelos de deep learning (Chollet et al., 2015) e Augmentor para data augmentation (Bloice et al., 2017). A comparação do resultado dos modelos criados foi realizada com o uso da métricas de acurácia, precisão e recall, demonstradas nas Equações 1, 2 e 3, respectivamente, onde TA é o número de acertos do modelo, N é o total de amostras previstas pelo modelo, TP é o número de acertos para a classe, FP é o número de falsos positivos para a classe e FN é o número de falsos negativos. Vale salientar que a precisão e o recall foram calculados a partir da média não ponderada da precisão e recall para cada classe.

$$\text{Acurácia} = \frac{TA}{N} \quad (1)$$

$$\text{Precisão} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

Em resumo, seis experimentos foram realizados para avaliar o uso de CNN na classificação de maturidade da banana prata Catarina, considerando três separações de níveis de maturidade e utilizando ou não a técnica de data augmentation.

3. Resultados e Discussões

3.1. Preparação dos dados

Como mencionado, foram utilizados todos os dados disponibilizados pelo conjunto de Martins Neto et al. (2023). As imagens originais possuem duas resoluções diferentes, ambas em alta resolução. Foi necessária a padronização da resolução, para que os modelos tenham uma entrada única, além da redução da resolução, para agilidade computacional e remoção de detalhes desnecessários para classificação das classes. Depois de testes variando as resoluções, chegou-se a conclusão de utilizar imagens de 200x200 pixels.

Além da mudança de resolução, dependendo do experimento, foi aplicado data augmentation. A Figura 2 mostra exemplos da aplicação das quatro técnicas citadas: flip horizontal, flip vertical, rotação e zoom.



Figura 2. Exemplos de data augmentation para as ações de flip horizontal, flip vertical, rotação e zoom.

Para construção e avaliação dos modelos, os dados foram separados em 70% para treino, 20% para validação e 10% para teste. A técnica de data augmentation foi utilizada apenas no conjunto de treinamento.

3.2. Construção dos modelos

Diversos testes foram realizados para a definição das camadas do modelo CNN para este trabalho. Os testes passaram por modelos com 3 a 5 blocos convolucionais (Convolutional+ReLU e Max Pooling), mantendo ou não a dimensão da saída igual à entrada para comparar captura de regiões mais largas ou atribuição a pontos específicos a partir da movimentação da janela de convolução nas imagens. Depois de variações para alcançar um resultado satisfatório, fixou-se o mesmo modelo para todos os experimentos,

variando apenas a camada final de classificação devido ao número diferente de classes entre os experimentos (conforme descrito na seção de Metodologia).

Foram definidos cinco blocos convolucionais, com uma camada de Dropout de 20% dos pesos antes de *flattening*, finalizando com uma camada totalmente conectada com função de ativação ReLU, com as 8 classes de saída variáveis para cada experimento. Um exemplo da arquitetura do modelo utilizado encontra-se na Figura 3.

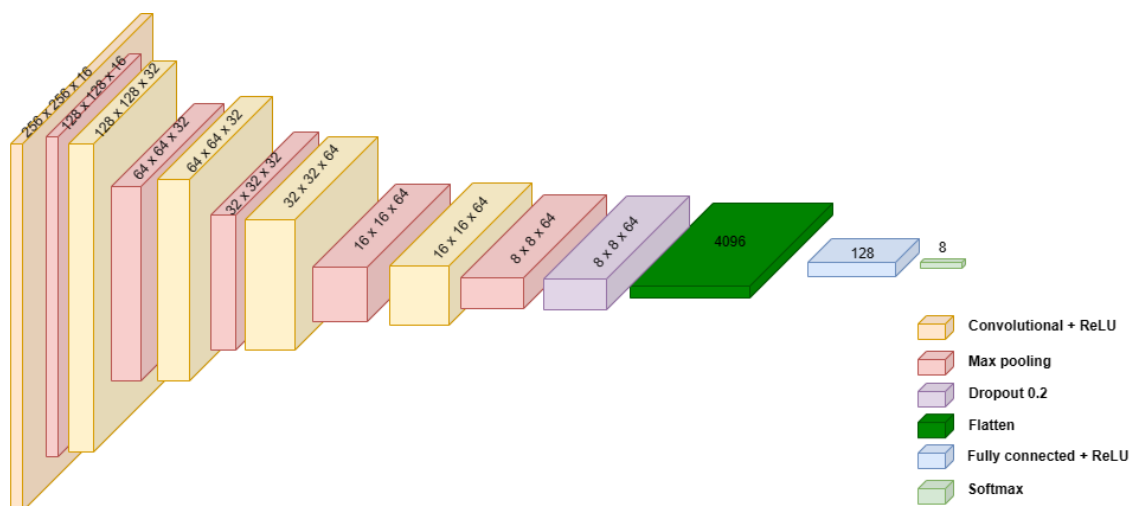


Figura 3. Arquitetura do modelo CNN criado, tendo como exemplo o experimento com 8 classes.

Todos os modelos foram treinados com 10 épocas e com batch size de 32 amostras, com otimização dos parâmetros através do otimizador Adam (Kingma e Ba, 2014) e função de perda Entropia categórica cruzada esparsa.

3.3. Desempenho dos modelos

Conforme mencionado, houve a construções de diferentes modelos e avaliados com as métricas de acurácia, precisão e recall. Os resultados obtidos tanto no treinamento quanto com os dados de teste estão apresentados na Tabela 1 (treinamento) quanto na Tabela 2 (teste).

Tabela 1. Resultado dos modelos criados para classificar os experimentos utilizando 8, 5 e 2 classes, com os dados de treinamento, utilizando ou não data augmentation.

Resultados apresentados em percentual (%).

Número de classes	Acurácia		Precisão		Recall	
	sem DA	com DA	sem DA	com DA	sem DA	com DA
8 classes	52,3	46,0	51,0	49,0	40,0	43,4
5 classes	65,3	69,2	64,3	72,4	61,0	63,0
2 classes	88,6	78,6	87,9	79,4	87,7	81,7

Tabela 2. Resultado dos modelos criados para classificar os experimentos utilizando 8, 5 e 2 classes, com os dados de teste, utilizando ou não data augmentation. Resultados apresentados em percentual (%).

Número de classes	Acurácia		Precisão		Recall	
	sem DA	com DA	sem DA	com DA	sem DA	com DA
8 classes	45,3	34,7	30,3	23,3	36,7	31,1
5 classes	57,9	60,0	59,3	48,6	56,5	52,3
2 classes	89,5	76,8	87,2	77,0	89,6	81,3

Os resultados com as 8 classes para o treinamento sem e com data augmentation obtiveram, respectivamente, 52,3% e 46% de acurácia durante o treinamento, enquanto a precisão foi de 51% e 49%, e o recall foi maior com data augmentation, 43,4% em comparação a sem data augmentation, com 40%. No entanto, para os dados de teste a acurácia foi para 45,3% e 34,7% para sem e com data augmentation respectivamente, enquanto a precisão foi de 30,3% e 23,3%, com recall de 36,7% e 31,1%. Para esse experimento, o melhor desempenho foi de 45,3% considerando os 8 diferentes estágios de maturação e apenas treinando com os dados originais.

Os resultados para 5 classes tiveram comportamento diferente do modelo treinado com 8 classes. O desempenho geral foi superior, conforme era esperado por se tratar de menos classes, em que a acurácia foi de 65,3% e 69,2% para o treinamento sem e com data augmentation, enquanto no teste a acurácia foi 57,9% e 60%, respectivamente. A precisão foi de 64,3% e 72,4% respectivamente para sem data augmentation e com data augmentation no treinamento, com recall de 61% e 63%. Já no teste, a precisão foi de 59,3% e 48,6%, enquanto o recall foi 56,5% e 52,3% respectivamente sem data augmentation e com data augmentation.

Por fim, ao utilizar somente 2 classes, o modelo demonstrou maior capacidade de generalização, pois a acurácia chegou a 88,6% e 78,6%, sem e com data augmentation, respectivamente, durante o treinamento. Com os dados de teste, a acurácia foi de 89,5% sem data augmentation, enquanto com data augmentation a acurácia chegou a 76,8%. A precisão foi 87,9% e 79,4% no treinamento enquanto o recall foi de 87,7% e 81,7% respectivamente, sem e com data augmentation, respectivamente. No teste, a precisão foi de 87,2 e 77%, enquanto o recall foi 89,6% e 81,3% para sem e com data augmentation.

3.4. Comparações com trabalhos correlatos

Outros trabalhos atuaram na tentativa de criar modelos para classificar a maturidade de bananas. Adebayo et al. (2017) e Chen et al. (2018) conseguiram boas estimativas. Entretanto, as características utilizadas nos métodos de aprendizado de máquina não vieram de fotos comuns, mas sim de equipamentos específicos no qual exige que os mesmos aparatos sejam utilizados em produção.

Há também trabalhos que atuam com imagens de fotos comuns, como Mazen e Nashat (2019) e Sabilla et al. (2019). Mas, nesses casos, as imagens tinham um fundo controlado. Desta forma, há uma extração de características da fruta mais estável. Por outro lado, não possibilita que o mesmo modelo seja utilizado fora de um mesmo ambiente controlado.

Por fim, há o trabalho de Zhang et al. (2018), no qual foi proposto um novo classificador de CNN para bananas em sete estágios de maturação. O algoritmo aprendeu um conjunto de recursos de imagem refinados e obteve desempenho de classificação superior a 90%. Um diferencial neste trabalho é que o conjunto de dados continha 17.312, o que provavelmente afetou diretamente e positivamente o desempenho do modelo.

4. Considerações Finais

Dada a importância do cultivo da banana no cenário nacional e global, este estudo teve como objetivo a implementação de modelo CNN para a classificação de estados de maturação da banana a partir de um conjunto de imagens de bananas Prata Catarina disponibilizado por meio do trabalho de Martins Neto et al. (2023).

Foram elaborados vários modelos, de forma que a acurácia com dados de teste variou de 45% a 89,5%, com a melhor precisão e recall de respectivamente 87,2% e 89,6%, no qual os melhores resultados variam de acordo com o número de classes a serem detectadas. Quanto maior o número de classes, mais semelhante são os exemplos entre si, o que dificulta a tarefa. A base de dados utilizada representa com fidelidade a classificação do fruto em situações reais, o que torna a tarefa ainda mais desafiadora.

Novas adequações ao modelo e ao conjunto de dados podem ser realizadas em trabalhos futuros para melhorar esses resultados, como diferentes funções de ativação ou números de camadas, e validação cruzada em *k-folds* para entender o enviesamento do modelo para diferentes subconjuntos dos dados em diferentes quantidades de classes. Este trabalho, além de mostrar a possibilidade e dificuldades de classificação deste fruto, serve também como um baseline para outros trabalhos da área e, principalmente, ao utilizar a base de dados disponibilizada publicamente.

Referências Bibliográficas

- Adebayo S., Hashim N., Abdan, K., Hanafi, M., Zude-Sasse, M. (2017). Prediction of banana quality attributes and ripeness classification using artificial neural network. *Acta Horticulturae*, 1152(1152):335–344.
- Bloice, M. D., Stocker, C., & Holzinger, A. (2017). Augmentor: an image augmentation library for machine learning. Disponível em arXiv preprint arXiv:1708.04680. Acessado em fevereiro de 2024.
- CEAGESP (2006) “PBMH & PIF — Programa Brasileiro para a Modernização da Horticultura & Produção Integrada de Frutas”. Normas de Classificação de Banana. Disponível em <https://ceagesp.gov.br/wp-content/uploads/2015/07/banana.pdf>. Acessado em fevereiro de 2024.
- Chen, LY., Wu, CC., Chou, TI., Chiu, SW., Tang, KT. (2018). Development of a dual MOS electronic nose/camera system for improving fruit ripeness classification. *Sensors*, 18(10):3256.

- Chollet, F et al. (2015). Keras. Disponível em <https://keras.io>. Acessado em fevereiro de 2024.
- Duarte, A. G., Barbieri, M. (2024). BANANA/CEPEA: Exportações fecham 2023 em queda. Disponível em <https://www.hfbrasil.org.br/br/banana-cepea-exportacoes-fecham-2023-em-queda.aspx>. Acessado em fevereiro de 2024.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. Disponível em arXiv preprint arXiv:1412.6980. Acessado em fevereiro de 2024.
- Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1), 91-99.
- Ministério da Agricultura e Pecuária. (2024). Exportações do agronegócio fecham 2023 com US\$ 166,55 bilhões em vendas. Disponível em <https://www.gov.br/agricultura/pt-br/assuntos/noticias/exportacoes-do-agronegocio-fecham-2023-com-us-166-55-bilhoes-em-vendas>. Acesso em fevereiro de 2024.
- Martins Neto, J. L., Neves, J. P. H., Xavier, M. V. T. N., Freitas, E. D. G. de, Gomes, D. G. Dataset de Imagens de Banana Prata Catarina Rotuladas em Oito Estados de Maturação. In: *WORKSHOP DE COMPUTAÇÃO APLICADA À GESTÃO DO MEIO AMBIENTE E RECURSOS NATURAIS (WCAMA)*, 14. , 2023, João Pessoa/PB. Porto Alegre: Sociedade Brasileira de Computação, 2023 . p. 155-158. ISSN 2595-6124.
- Mazen, F. M. A. and Nashat A. A. (2019). Ripeness classification of bananas using an artificial neural network. *Arabian Journal for Science and Engineering* 44(8):6901–6910.
- Sabilla, I. A., Wahyuni, C. S., Fatichah, C., Herumurti, D. (2019). Determining banana types and ripeness from image using machine learning methods. In: *IEEE International Conference of Artificial Intelligence and Information Technology (ICAIIIT)*. Proceedings.
- Siddiq, M., Ahmed, J., Lobo, M. G. (2020). *Handbook of banana production, postharvest science, processing technology, and nutrition*. Hoboken, Wiley, 284 p.
- Zhang, Y., Lian, J., Fan, M., Zheng, Y. (2018). Deep indicator for fine-grained classification of banana's ripening stages. *EURASIP Journal on Image and Video Processing* 2018(1):1–10.