

Classificação De Espécies De Arroz Utilizando Inteligência Artificial

Julio C. Jauer¹, Tiago Lacerda¹

¹Departamento de Computação – Campus Cuiabá – Cel. Octayde Jorge da Silva – Instituto Federal de Mato Grosso (IFMT) - R. Zulmira Canavarros, 95 – Centro, Cuiabá - MT, 78005-390 – Cuiabá – MT – Brasil.

j.jauer@estudante.ifmt.edu.br, tiago.lacerda@ifmt.edu.br

Abstract. *The objective of this article is to classify rice species using Artificial Intelligence and evaluate the method used. The study was based on a dataset containing 75,000 images of rice grains, divided into five varieties. The model used was a pre-trained Convolutional Neural Network (CNN) DenseNet169, to which additional layers were added to adjust the network to the specific problem. The training was conducted using Google Colaboratory, and the results showed an accuracy of 96.77% in training and 94.29% in testing, with some difficulties in generalizing to the Arborio variety. The study highlights the importance of adjusting the model to improve generalization and suggests future works.*

Resumo. *O objetivo deste artigo é classificar espécies de arroz utilizando Inteligência Artificial e avaliar o método utilizado. O estudo foi baseado em um banco de dados contendo 75.000 imagens de grãos de arroz, divididas em cinco variedades. O modelo utilizado foi uma Rede Neural Convolutacional (CNN) pré-treinada DenseNet169, a qual foram adicionadas camadas adicionais para ajustar a rede ao problema específico. O treinamento foi realizado utilizando o Google Colaboratory, e os resultados mostraram uma acurácia de 96,77% no treinamento e 94,29% nos testes, com algumas dificuldades na generalização para a variedade Arborio. O estudo destaca a importância de ajustar o modelo para melhorar a generalização e sugere trabalhos futuros.*

1. Introdução

Observando os números da produção de cereais no mundo, os produtos mais importantes são trigo, milho e arroz (AWIKA,2011). O arroz, por sua vez, é muito rico em carboidrato (amido) e proteína sendo o elemento básico na alimentação de muitas pessoas (NAVES, 2007). No Brasil, sua importância na alimentação fica evidente, pois compõe uma parcela considerável da alimentação no dia a dia do povo (NAVES, 2007). Segundo o IBGE (2004), o arroz compunha 20% do consumo de calorias diário dos brasileiros. Com isso, o arroz move uma grande área da produção e melhorar o seu processo de produção é um fator necessário.

O processo de classificação manual do arroz tem diferentes critérios de avaliação e é um processo custoso, principalmente quando depende da subjetividade e velocidade humana. Diferentes algoritmos do Aprendizado de Máquina podem ser usados para automatizar esse processo e torná-lo mais eficiente, preciso e econômico.

Tratando-se de visão computacional, alguns aspectos físicos como cor, textura, qualidade e tamanho são utilizados para essa classificação.

Nos trabalhos de classificação de arroz, as características morfológicas como área, perímetro, comprimento do eixo maior, comprimento do eixo menor, excentricidade, área convexa e extensão são usadas na classificação (CINAR E KOKLU, 2019). As imagens costumam ser convertidas para tons de cinza e depois para binário (CINAR E KOKLU, 2021). Além disso, é preciso avaliar a classificação do arroz e, para isso, são feitos alguns cálculos estatísticos (CINAR E KOKLU, 2022).

Nesse contexto, vamos testar um algoritmo de aprendizado de máquina com uma base de dados contendo cinco variedades de arroz. Para isso, usaremos Inteligência Artificial (IA) para os dados de forma rápida e precisa.

Assim, o objetivo do artigo é classificar espécies de arroz utilizando Inteligência Artificial e avaliar o método utilizado.

A estrutura do artigo se dá em 5 seções, incluindo esta seção 1: a introdução. A seção 2, que apresenta a fundamentação teórica. A seção 3, onde são expostos os materiais e métodos. A seção 4, que contém os resultados e a discussão. Por fim, a seção 5, contém conclusões finais.

2. Fundamentação Teórica

Nesta seção, são abordados os fundamentos teóricos necessários para a compreensão do artigo.

2.1. Rede Neural Convolutacional

Redes Neurais Convolucionais, do inglês *Convolutional Neural Networks* (CNNs), são Redes Neurais Artificiais (RNAs). Quando chamamos uma camada de uma RNA de “camada completamente conectada” (“fully connected layer”), significa que cada neurônio da camada anterior é conectado a todos os neurônios da próxima camada. Em uma CNN, por outro lado, na camada seguinte, os neurônios recebem informações apenas de uma subárea da camada anterior. Além disso, podemos aplicar o conceito de *pooling* que consiste em reduzirmos a dimensão da rede de uma camada para a outra conservando as informações necessárias para o treinamento (DELL’AVERSANA, 2019).

2.2. Camadas de Rede Neural

Algumas das camadas utilizadas em RNAs são: *Flatten*, *Dense* e *Dropout*. A camada *Flatten* converte uma matriz em uma vetor, geralmente sendo a camada intermediária entre as camadas CNN e as camadas mais tradicionais. Nas camadas *Dense* todos os neurônios da camada anterior estão ligados a todos os neurônios da camada seguinte e elas podem ter algumas funções específicas como a *rectified linear unit* (ReLU), que converte os números negativos em zeros, e a *softmax*, que geralmente é utilizada como a camada de saída por converter os valores de todos os neurônios da camada em probabilidades que tem a somatória igual a 1. E, as camadas de *Dropout*, utilizadas como camadas intermediárias, que desativam alguns neurônios aleatoriamente com o objetivo de evitar o *overfitting* do modelo nos dados de treinamento.

2.3. Estatísticas

Entre as principais estatísticas para avaliar as redes neurais estão: acurácia, perda e matriz de confusão. Em um cenário binário, há dois resultados: verdadeiro ou falso. Se o resultado verdadeiro estiver certo ele é um verdadeiro positivo (Vp), se estiver errado ele é um falso positivo (Fp). Se o resultado for falso e estiver correto ele é um verdadeiro negativo (Vn), se estiver errado ele é um falso negativo (Fn). Assim, a acurácia mede a proporção de resultados verdadeiros em relação à totalidade de resultados. A matriz de confusão é uma matriz com as variáveis Vp, Fp, Fn e Vn. E a perda é uma quantização da margem de erro; nesse caso, fala-se da perda de entropia cruzada categórica que é o valor negativo do somatório dos logaritmos naturais das probabilidades previstas para as classes verdadeiras.

3. Metodologia

Aqui são explicadas as etapas da pesquisa.

3.1. Definição da Base de Dados e ambiente de treinamento

Iniciamos encontrando uma base de dados adequada ao projeto: um grande conjunto de imagens de grãos de arroz. A base de dados utilizada foi extraída do *dataset Rice Image* (KÖKLÜ; ÇINAR; TAŞPINAR, 2021). Ele é um conjunto total de 75.000 imagens de grãos de arroz, distribuídas em cinco variedades diferentes, sendo 15.000 imagens de cada uma das variedades: Arborio, Basmati, Ipsala, Jasmine e Karacadag. Essas são imagens de variedades de arroz muito utilizadas na Turquia que possuem diferentes características físicas e foram selecionadas pela facilidade de acesso ao conjunto de dados. Os dados adquiridos foram armazenados no Google Drive em arquivo zip.

3.2. Processamento das Imagens

As imagens de grãos de arroz foram pré-processadas antes do treinamento porque, segundo Çinar, Köklü e Taşpinar (2019), isso afeta diretamente nos resultados. Assim, as imagens foram, em uma primeira etapa, convertida para tons de cinzas e, em seguida, convertidas para valores binários.

3.3. Separação das Imagens

Após o processamento, as imagens foram separadas em três conjuntos: treino, validação e teste. O conjunto de treino é utilizado para o treinamento em si, o conjunto de validação é utilizado em paralelo com o conjunto de treinamento para avaliar o desempenho do aprendizado durante a fase de treinamento, e o conjunto de teste é utilizado após o treino para avaliar a capacidade de generalização do modelo, simulando o comportamento do modelo no “mundo real”. Além disso, foi importante garantir que os conjuntos de dados tivessem quantidades equilibradas de imagens de cada uma das espécies para evitar um treinamento tendencioso.

Com as imagens separadas, foram criados *data sources* para os grupos. Nenhum dado novo foi gerado nessa etapa, porém ela é importante pois são os *data sources* que alimentam o treinamento, a validação e os testes enviando um pequeno pacote de dados de cada vez para otimizar o uso da memória. Além disso, os *data sources* garantem que os dados passados adiante estarão normalizados no padrão definido. Nesse caso, foi feito o redimensionamento das imagens para 128x128 pixels com 3 canais de cores.

3.4. Definição do Modelo

O modelo tem como base a CNN pré treinada *DenseNet169*. Removendo a última camada, podemos adicionar nossas próprias camadas na sequência, adequando o modelo pré treinado ao nosso problemas. As camadas adicionadas foram: uma camada *Flatten*; uma camada *Dense* com ReLU; uma camada de *Dropout*; uma camada *Dense* com ReLU; e uma camada *Dense* com *Softmax*. O treinamento foi realizado utilizando o ambiente do Google Colab com 12.7 GB de RAM.

4. Resultados e Análises

Após vários testes, as imagens de entrada foram definidas na forma 128x128 pixels com 3 canais de cores (RGB), pois com 256x256 pixels o tempo de treinamento estava três vezes maior, o que dificultava a conclusão das etapas de treinamento sem erros na conexão. A última camada foi removida e em seguida adicionamos a camada *Flatten* para vetorizar a matriz das imagens. Foram adicionadas uma camada *Dense* com 256 neurônios e ReLU, uma camada com *Dropout* de 25%, uma camada *Dense* com 64 neurônios e ReLU, e uma camada de saída *Dense* com 5 neurônios e softmax. Utilizando uma taxa de aprendizado de 0.001, a função de perda de entropia cruzada categórica e a métrica acurácia para avaliar o desempenho.

Em um primeiro momento o objetivo era conseguir uma acurácia de pelo menos 80%, mas os resultados foram muito melhores do esperado. A acurácia do primeiro treinamento foi de 96,77%, com uma perda de 0.1099, enquanto a acurácia da primeira validação foi de 94,48% e a perda foi de 0.4079. A diferença entre a perda do treinamento e da validação indicam que o modelo está sofrendo *overfitting*, perdendo sua capacidade de generalizar os dados. Para melhorar a generalização, pode ser útil aumentar a porcentagem de *Dropout* ou adicionar outra camada de *Dropout*.

Os resultados foram semelhantes aos da validação e tivemos uma acurácia de 94,29%, com uma perda de 0.4141. Além disso, a análise da matriz de confusão do teste, a Figura 1, revelou que os grãos de arroz Arborio tiveram muitos falsos positivos. Isso sugere que, com o conjunto de treinamento deste teste, a rede neural teve dificuldades em generalizar essa espécie.

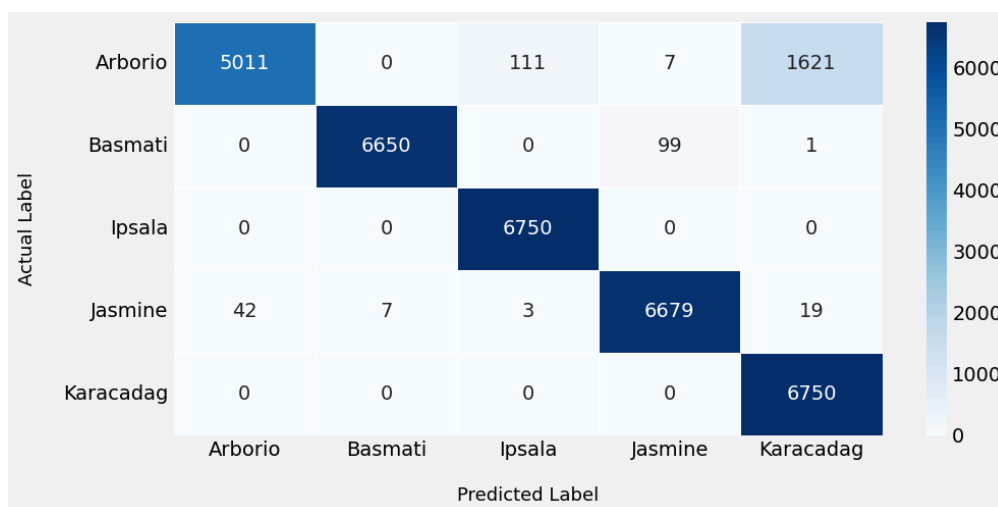


Figura 1. Matriz de confusão dos testes do modelo.

O processamento das três etapas (treinamento, validação e testes) durou aproximadamente 3 horas, com o treinamento sendo feito em apenas uma etapa.

5. Considerações Finais

O objetivo do artigo era classificar espécies de arroz utilizando Inteligência Artificial e avaliar o método utilizado. Foi encontrada uma base de dados adequada e feita a busca de informações a respeito da classificação de arroz. Houve problemas durante o treinamento com a falta de hardware que gerou longos tempos de treinamento que ocasionalmente eram perdidos devido a oscilações na conexão da rede de internet. Isso trouxe a busca por simplificar o modelo buscando diminuir o tempo de processamento que, com sucesso, reduziu o tempo para um terço dos testes iniciais, além de aumentar a qualidade do treinamento, devido ao modelo ter ficado mais específico para o objetivo.

Entre as propostas de trabalhos futuros está o treinamento de mais modelos, com outros modelos pré-treinados como base, para que seja possível comparar os resultados entre cada um deles. Além de buscar um ambiente de desenvolvimento mais estável onde seja possível executar treinamentos mais robustos; fazer as alterações observadas nesse artigo e avaliar os novos resultados; testar outros métodos de treinamento; buscar um banco de dados das espécies de arroz mais utilizadas no Brasil ou produzir um banco de dados delas; e, colocar a utilização do modelo treinado em prática. Para, com isso, contribuir para o desenvolvimento dos estudos sobre IAs e com o estudo sobre arroz no Brasil.

Referências

- AWIKA, J. M. Major Cereal Grains Production and Use around the World. In: Awika et al. *Advances in Cereal Science: Implications to Food Processing and Health Promotion*. ACS Symposium Series, Vol. 1089, American Chemical Society: Washington, DC, 2011. P. 1-13 Disponível em: <https://pubs.acs.org/doi/full/10.1021/bk-2011-1089.ch001>. Acesso em: 02 set. 2024.
- CINAR, I.; KOKLU, M. Determination of Effective and Specific Physical Features of Rice Varieties by Computer Vision In Exterior Quality Inspection. *Selcuk Journal of Agriculture and Food Sciences*, 35(3), 229-243, 25 Dec. 2021. Disponível em: <https://doi.org/10.15316/SJAFS.2021.252>. Acesso em: 02 set. 2024.
- CINAR, I.; KOKLU, M. Identification of Rice Varieties Using Machine Learning Algorithms. *Journal of Agricultural Sciences*. 28(2), 307-325, 25 abr 2022. Disponível em: <https://doi.org/10.15832/ankutbd.862482>. Acesso em: 02 set. 2024.
- ÇINAR, I.; KÖKLÜ, M.; TAŞPINAR, Y. S. Rice Image Dataset. 2021. Disponível em: <https://www.muratkoklu.com/datasets/>. Acesso em: 09 set. 2024.
- CINAR, I.; KOKLU, M. Classification of Rice Varieties Using Artificial Intelligence Methods. *International Journal of Intelligent Systems and Applications in Engineering*, 7(3), 188-194, 30 nov. 2019 . Disponível em: <https://doi.org/10.18201/ijisae.2019355381>. Acesso em: 02 set. 2024.
- Dell'aversana, P. Artificial Neural Networks And Deep Learning. A Simple Overview. in DELL'AVERSANA, P. "A Global Approach to Data Value Maximization": Integration Machine Learning and Multimodal Analysis. 1. ed. Lady Stephenson

- Library, Newcastle upon Tyne, NE6 2PA, UK: Cambridge Scholars Publishing, 2019. cap. 5. 184-196. Disponível em: https://www.researchgate.net/profile/Paolo-Dellaversana/publication/337678793_ARTIFICIAL_NEURAL_NETWORKS_AND_DEEP_LEARNING_A_SIMPLE_OVERVIEW/links/5ebed4ed299bf1c09ac09260/ARTIFICIAL-NEURAL-NETWORKS-AND-DEEP-LEARNING-A-SIMPLE-OVERVIEW.pdf . Acesso em: 15 set. 2024.
- NAVES, M. M. V. (2007) Características Químicas E Nutricionais Do Arroz. B.CEPPA, Curitiba, v. 25, n. 1. P. 51-60, jan./jun. 2007. Disponível em: <https://core.ac.uk/download/pdf/328054375.pdf>. Acesso em: 02 set. 2024.