

XI CONGRESSO DE INICIAÇÃO CIENTÍFICA DO IFSP ITAPETININGA

Itapetininga, 27, 28 e 29 de maio de 2025

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

Campus Itapetininga

DESENVOLVIMENTO DE UMA METODOLOGIA BASEADA EM INTELIGÊNCIA ARTIFICIAL PARA A CLASSIFICAÇÃO AUTOMÁTICA DE GRÃOS DE CAFÉ

Nicholas Mezikas Pego – IFSP¹

Prof. Dr. Arthur Emanuel de Oliveira Carosia - IFSP²

Introdução:

A cultura cafeeira desempenha um papel vital na economia brasileira desde o século XIX, sendo uma das principais atividades agrícolas do país. A produção de café foi responsável por impulsionar o desenvolvimento econômico e social do Brasil, promovendo a construção de infraestrutura como ferrovias e portos (Leme, 2005).

Nesse sentido, uma etapa importante na produção de café é a classificação manual de cafés, feita por agricultores e especialistas. A classificação manual de café é realizada para determinar a qualidade dos grãos, que influencia diretamente o valor de mercado do produto, tornando a classificação uma etapa crucial no processo produtivo. Esse processo envolve a avaliação de diversos fatores, como tamanho, cor, densidade, presença de defeitos e características sensoriais (por exemplo, sabor e aroma). Como a classificação de café é realizada predominantemente de forma manual, essa é uma etapa produtiva que apresenta limitações, como: subjetividade na avaliação, variabilidade entre classificadores e baixa eficiência, o que pode resultar em inconsistências na qualidade e prejuízos econômicos. Além disso, a quantidade de fatores necessários para avaliação dos grãos dificulta ainda mais este processo: análise da qualidade dos grãos, caracteres morfológicos, impurezas, presença de pedras ou partículas de solo, dentre outros (Brasil, 1992).

Ademais, dada a crescente produção do setor cafeeiro e a demanda por produtos de alta qualidade no mercado, a necessidade de análises rápidas e precisas dos atributos físicos e fisiológicos dos grãos está se tornando um ponto crítico no processo produtivo (Bewley et al., 2013).

Tais limitações começaram a ser atenuadas recentemente com o avanço da visão computacional e, posteriormente, inteligência artificial com as técnicas de aprendizado de máquina (Russel, et al., 2010). Nesse sentido, trabalhos começaram a surgir no sentido de classificar grãos de café de acordo com diversos atributos de interesse (Ligar, 2022). Ligar (2022) apresentou uma revisão da literatura de diversos trabalhos considerando os avanços na tecnologia de visão computacional aplicada à identificação e classificação de grãos de

¹Estudante do curso de Bacharelado em Ciência da Computação, Instituto Federal de São Paulo, IFSP – São João da Boa Vista/SP. E-mail: m.nicholas@aluno.ifsp.edu.br. Orcid: <https://orcid.org/0009-0003-6852-5030>.

²Doutor. Instituto Federal de São Paulo, IFSP – São João da Boa Vista/SP. E-mail: arthuremanuel.carosia@ifsp.edu.br. Orcid: <https://orcid.org/0000-0002-6277-0014>.

XI CONGRESSO DE INICIAÇÃO CIENTÍFICA DO IFSP ITAPETININGA

Itapetininga, 27, 28 e 29 de maio de 2025

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

Campus Itapetininga

café verde. O artigo conclui que a visão computacional combinada com técnicas de aprendizado profundo oferece uma solução promissora para a identificação e classificação de grãos de café verde. Essa abordagem também facilita a automação do processo, o que pode ser altamente benéfico para a indústria do café.

Assim, o desenvolvimento de uma metodologia de classificação automatizado de cafés pode auxiliar a produção cafeeira ao fornecer avaliações precisas e rápidas da qualidade dos grãos, que é o objetivo deste trabalho.

Objetivo:

Como destacado, a classificação manual do café enfrenta limitações significativas, como a subjetividade na avaliação, a variabilidade entre classificadores e a baixa eficiência, levando a inconsistências na qualidade e possíveis prejuízos econômicos. Nesse sentido, este trabalho aborda a criação de uma metodologia automatizada para a classificação de grãos de café utilizando técnicas de inteligência artificial, em especial algoritmos de Aprendizado de Máquina, como Support Vector Machines, Logistic Regression e a rede neural artificial do tipo Multilayer Perceptron.

Metodologia:

No contexto deste trabalho, o uso de técnicas de IA aliadas com visão computacional possui um potencial significativo para transformar a classificação de grãos de café, oferecendo um método eficiente e preciso para avaliar a qualidade dos grãos (Oyama et al., 2012). Os sistemas de visão computacional podem analisar origem, tipos de café, características visuais dos grãos, como cor, tamanho, forma e presença de defeitos, dentre outros. Isso permite uma classificação rápida e consistente, minimizando a subjetividade e a necessidade de intervenção humana.

Para o desenvolvimento do classificador proposto neste trabalho, exploramos abordagens tradicionais de Aprendizado de Máquina, como: Support Vector Machines, Logistic Regression e redes neurais do tipo Multilayer Perceptron.

Para que o objetivo deste trabalho seja atingido, os seguintes objetivos específicos devem ser alcançados:

- **Busca de base de dados de imagens de café:** Nessa etapa, o aluno procurou pela internet por bases de dados que pudessem se encaixar nas necessidades deste trabalho. A base de dados utilizada consiste em imagens de grãos de café previamente classificadas por especialistas da área e disponíveis abertamente na internet (Disponível em: <https://www.kaggle.com/datasets/gpiosenka/coffee-bean-dataset-resized-224-x-224>). As imagens representam grãos de café torrados: café **Laos Typica Bolaven**, levemente torrado, café **Doi Chaang**, torra média, e **Brazil Cerrado**, que recebe uma torra escura. O banco de imagens apresenta mil e seiscentas imagens de grãos de café classificadas em quatro grupos: (1) preto; (2) branco; (3) verde; e (4) médio. Um exemplo de um dos grãos presentes na base de dados pode ser visualizado na Figura 1.

XI CONGRESSO DE INICIAÇÃO CIENTÍFICA DO IFSP ITAPETININGA

Itapetininga, 27, 28 e 29 de maio de 2025

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

Câmpus Itapetininga



Figura 1. Imagem de um grão de café torrado (verde) presente na base de dados

- **Pré-Processamento dos dados:** Nesta etapa, os dados são carregados e ocorre a extração de características das imagens de café e preparação dos dados antes do treinamento dos modelos. Além disso, as imagens foram distribuídas aleatoriamente em dois grupos, para treinamento e teste do modelo, respeitando a seguinte proporção: 80% das imagens para treinamento e 20% das imagens para teste.
- **Desenvolvimento de modelos baseados em técnicas de Aprendizado de Máquina tradicionais:** neste trabalho, selecionamos os seguintes modelos para comparação: Support Vector Machines, Regressão Logística e Multilayer Perceptron. Tais modelos foram selecionados por serem frequentemente utilizados na literatura (García et al., 2019). A técnica Support Vector Machines pertence à categoria de aprendizado supervisionado e busca encontrar um hiperplano ótimo para separar classes maximizando a margem entre os dados; Logistic Regression é um modelo estatístico para classificação que usa a função sigmoide para estimar a probabilidade de uma classe; e Multilayer Perceptron é uma rede neural artificial composta por múltiplas camadas de neurônios interconectados que aprendem padrões complexos por meio de pesos ajustáveis (Simon, 2009).
- **Análise dos resultados:** Para avaliar a eficiência dos classificadores, atualmente utilizamos a métrica de Acurácia, que mede a porcentagem de acertos obtida pelos modelos de Aprendizado de Máquina considerados.

Para o desenvolvimento deste trabalho, as seguintes ferramentas foram utilizadas. Linguagem de Programação Python, que possui diversas bibliotecas que facilitam a implementação dos algoritmos aqui citados; interface de desenvolvimento PyCharm (<https://www.jetbrains.com/pycharm/>); e a biblioteca para aprendizado de máquina Scikit-Learn (<https://scikit-learn.org/stable/>).

Resultados:

XI CONGRESSO DE INICIAÇÃO CIENTÍFICA DO IFSP ITAPETININGA

Itapetininga, 27, 28 e 29 de maio de 2025

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

Campus Itapetininga

A Tabela 1 apresenta os parâmetros utilizados em cada algoritmo de Aprendizado de Máquina para a condução dos experimentos..

Tabela 1: Parâmetros dos algoritmos de Aprendizado de Máquina

Algoritmo de Aprendizado de Máquina	Parâmetros
Support Vector Machines	kernel: rbf; C:1; degree = 3
Logistic Regression	solver=lbgfs
Multilayer Perceptron	hidden_layer_sizes=(100,); activation=relu; solver=adam; learning_rate_init=0.001

Os resultados detalhados dos experimentos são apresentados na Tabela 2. É possível observar que a acurácia obtida variou entre 93% e 95% no conjunto de testes. Ou seja, para o conjunto de dados de teste, os modelos foram capazes de acertar até 95% das classes de café. O melhor modelo dentre os avaliados foi o modelo *Logistic Regression*.

Tabela 2: Resultados obtidos nos experimentos

Algoritmo de Aprendizado de Máquina	Acurácia
Support Vector Machines	93,63%
Logistic Regression	95,45%
Multilayer Perceptron	93,18%

Conclusão:

O desenvolvimento de um classificador automático baseado em Inteligência Artificial pode representar um avanço significativo na classificação de grãos de café, proporcionando maior padronização e confiabilidade ao processo, além de impulsionar a competitividade da produção cafeeira nacional.

Espera-se que a proposta desenvolvida apresente acurácia comparável ou superior à classificação manual, garantindo maior consistência na avaliação dos grãos. Atividades futuras de automação nesse processo possuem potencial de aumentar a produtividade neste ramo da economia nacional.

Como trabalhos futuros, iremos experimentar técnicas de Aprendizado Profundo, como redes neurais convolucionais, e outras métricas de avaliação, como F1-Score. Pretendemos também explicar outras configurações para os algoritmos de Aprendizado de Máquina tratados neste artigo. Além disso, pretendemos também avaliar o desempenho deste trabalho na prática, ou seja, em um contexto real de classificação de cafés.

XI CONGRESSO DE INICIAÇÃO CIENTÍFICA DO IFSP ITAPETININGA

Itapetininga, 27, 28 e 29 de maio de 2025

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

Campus Itapetininga

Referências:

Bewley, J. D., & Black, M. (2013). *Seeds: physiology of development and germination*. Springer Science & Business Media.

Brasil, Regras para análise de sementes. Brasília, DF: Ministério da Agricultura e Reforma Agrária, SNDA/DNDV/CLAV, 1992.

García, M., Candelo-Becerra, J. E., & Hoyos, F. E. (2019). Quality and defect inspection of green coffee beans using a computer vision system. *applied sciences*, 9(19), 4195.

Leme, J. A. (2005). *A Expansão do Café no Brasil*. São Paulo: Editora Unesp.

Ligar, B. W. (2022). Computer Vision for Identifying and Classifying Green Coffee Beans: A Review. *Jurnal CoreIT*, 8(1), 11-18.

Oyama, P. D. C., JORGE, L. D. C., Rodrigues, E. L. L., & Gomes, C. C. (2012). Sistema para classificação automática de café em grãos por cor e forma através de imagens digitais.

Russel, S., Norvig, P. *Artificial intelligence—a modern approach*. Prentice Hall. 2010.

Simon, H. (2009). *Neural networks and learning machines*. McMaster University.