

# USO DE MACHINE LEARNING PARA CLASSIFICAÇÃO DE SUCOS DE FRUTAS POR COLORIMETRIA

João Vitor Plakitkem Zarth<sup>1</sup>, Patrick Guilherme Roza<sup>1</sup>, Vanderlei Aparecido de Lima<sup>1</sup>

<sup>1</sup>Universidade Tecnológica Federal do Paraná - Campus Pato Branco

Contato: valima@utfpr.edu.br



*Aprendizado de máquina e processamento de imagens para diferenciar sucos de frutas com precisão e garantir qualidade na indústria de alimentos.*

## INTRODUÇÃO

O uso de métodos computacionais para análise e classificação de alimentos tem ganhado destaque na indústria de alimentos e bebidas, especialmente com o avanço de técnicas de aprendizado de máquina. Essas técnicas possibilitam a distinção entre produtos que apresentam semelhanças visuais, como os sucos de frutas processados, cuja diferenciação pode ser desafiadora devido à proximidade de suas colorações. A classificação correta desses sucos é essencial para garantir a qualidade do produto e evitar possíveis fraudes ou misturas não declaradas. Nesse contexto, o uso de processamento digital de imagem aliado a algoritmos de aprendizado de máquina tem se mostrado uma alternativa promissora, permitindo a identificação e diferenciação de sucos com base em suas características visuais, como a cor, além de outros atributos espectroscópicos e químicos (VAN DER SCHOOT et al., 2020).

A cor é um parâmetro amplamente acessível e indicativo de características intrínsecas dos produtos, sendo particularmente útil para a classificação de sucos. No entanto, diferenças sutis nas tonalidades das dessas matrizes são muitas vezes não perceptíveis a olho nu. Nesse contexto, o uso de técnicas de Machine Learning (ML), em combinação com processamento digital de imagens, surge como uma ferramenta poderosa para identificar padrões e realizar classificações precisas em sucos (HOU et al., 2018).

Este artigo apresenta um modelo de classificação baseado em ML para distinguir amostras de sucos de laranja, abacaxi, manga e maracujá, utilizando atributos extraídos das suas cores.

## DESENVOLVIMENTO

### *Aquisição de imagens e pré-processamento*

Para a geração do modelo de classificação dos quatro sabores de sucos, foram utilizadas imagens das preparações dos sucos de: laranja, abacaxi, manga e maracujá. As imagens adquiridas foram realizadas em condições de iluminação controladas e apenas com smartphone marca Samsung Galaxy S24. Os sucos foram preparados conforme a orientação de seus rótulos de suas respectivas embalagens. As imagens foram processadas em softwares gratuitos. O software GIMP foi utilizado recortar as imagens com dimensões de 200 X 200 pixels e o software ChemoStatV2 foi utilizado para extração dos atributos de cor: tons de cinza dos canais RGB (Red, Green, Blue) e HSV (Hue, Saturation, Value). Esses espaços de cores foram escolhidos pela sua capacidade de capturar variações as quais não são perceptíveis a olho nu e pela capacidade de discriminar tonalidades intrínsecas de cada tipo de suco de fruta (HOU et al., 2018).

### *Modelos de Machine Learning para classificação*

Para a classificação dos sucos foram testados diferentes algoritmos de Machine Learning, incluindo K- nearest Neighbors (KNN), Suport Vector Machine (SVM) e Random Forest. Estes métodos foram escolhidos pela sua eficácia em problemas de classificação e pela capacidade de lidar com conjuntos de dados com múltiplas classes.

**K-Nearest Neighbors (KNN):** Um classificador simples e eficiente para pequenas bases de dados, que atribui a classe de um novo ponto com base nas classes dos pontos mais próximos.

**Support Vector Machine (SVM):** Método popular para classificação que busca encontrar um hiperplano ótimo que maximize a separação entre classes, adequado para diferenciar cores similares entre sucos.

**Random Forest (RF):** Algoritmo de aprendizado conjunto que combina múltiplas árvores de decisão, sendo útil para identificar padrões complexos nas cores dos sucos.

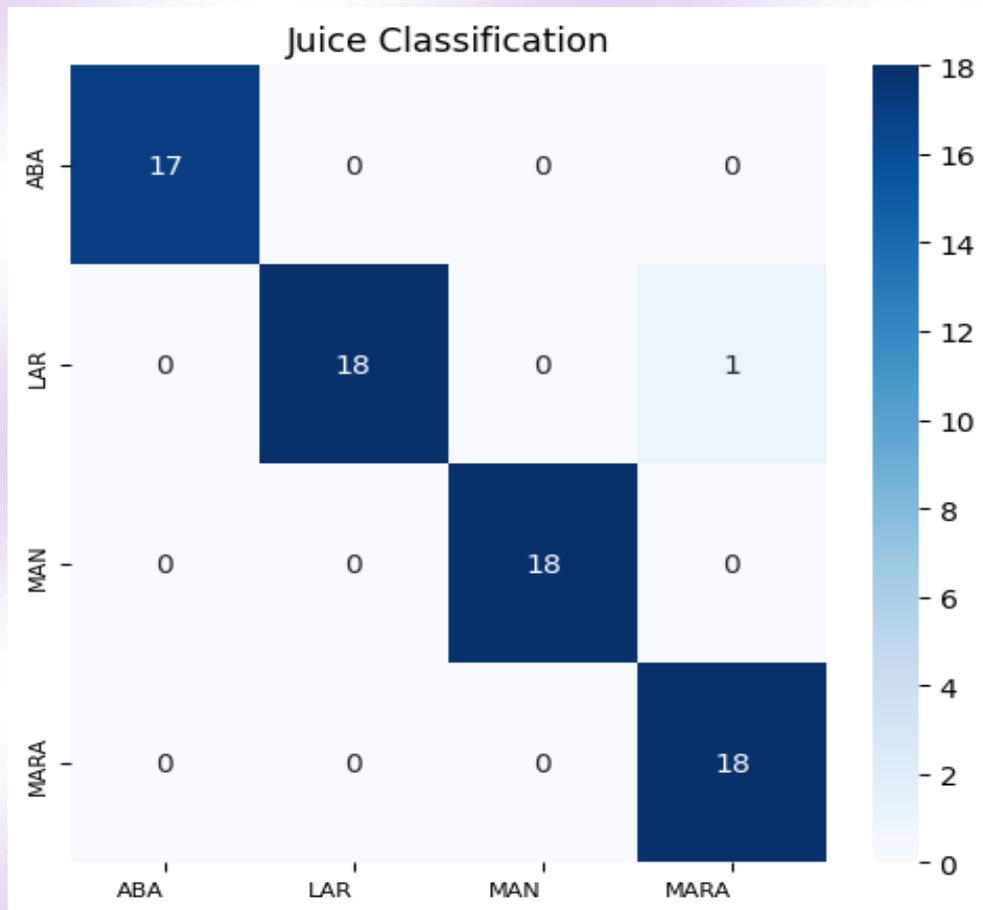
Cada modelo de classificação foi gerado utilizando- se conjuntos de treino (70% dos dados) e teste (30% dos dados), comparando as precisões de cada modelo na classificação de cada tipo de suco. As modelagens foram realizadas por meio da linguagem de programação Python em ambiente do Google Colab. Apenas os resultados do modelo de classificação de sucos utilizando-se o algoritmo Random Forest foram apresentados nesse artigo.

### *Avaliação dos resultados*

Os resultados gerados por cada modelo foram avaliados utilizando métricas estatísticas as quais avaliam a precisão do modelo e sua acurácia. Dos modelos testados, o algoritmo Random Forest se mostrou o algoritmo mais eficaz na classificação dos tipos de sucos, com uma precisão global de 99%. O modelo foi capaz de classificar corretamente todas as amostras do suco de abacaxi e classificando erroneamente apenas uma amostra de suco de laranja, classificando-o como o suco de maracujá. Esses resultados obtidos pelo modelo de classificação são apresentados em forma de uma Matriz de Confusão



(Figura 1). Na matriz de confusão, um eixo representa as classes esperadas e o outro eixo apresenta as classes previstas pelo modelo. A partir do número de acertos e erros, são calculadas as métricas que avaliam o desempenho do modelo (Tabela 1).



**Figura 1.** Matriz de confusão para com o algoritmo Random Forest.

**Tabela 1.** Resultado das métricas da performance do modelo Random Forest.

Tipos de sucos de frutas	Precisão	Recall	F1-score	Suporte
<b>0-ABA</b>	100%	100%	100%	17
<b>1-LAR</b>	100%	95%	97%	19
<b>2-MAN</b>	100%	100%	100%	18
<b>3-MARA</b>	95%	88%	97%	18
<b>Acurácia</b>			99%	72

## CONSIDERAÇÕES FINAIS

A classificação de sucos de frutas por meio de machine learning oferece uma alternativa prática e eficiente para a classificação de sucos e se mostra uma técnica com grande potencial para ser utilizada como ferramenta de controle de qualidade e análise de adulteração em indústria de sucos e bebidas. O modelo de classificação, especificamente quando utilizado o algoritmo Random Forest se mostra eficiente e apresenta grande potencial para detectar possíveis fraudes no setor alimentício.



Ainda, futuras pesquisas poderão explorar a combinação de análise cromática com dados espectroscópicos, visando aumentar ainda mais a robustez e sensibilidade dos modelos na classificação de alimentos e bebidas. Além disso, a implementação de sistemas automatizados para captura e análise de imagens pode facilitar a aplicação comercial dessas tecnologias.

## REFERÊNCIAS

---

HOU, Z. et al. Machine learning approaches for food quality and safety analysis based on near-infrared and visual imaging techniques. *Food Chemistry*, 270, 2018.

VAN DER SCHOOT, L. et al. Authentication of fruit juice using infrared spectroscopy and chemometrics: A review. *Food Control*, 120, 2020.

