

Classificação de Sementes de Cacau no Teste de Corte por Meio de CNNs como Extratoras de Características e SVMs como Classificadoras

Antonio Ramos, Joelma Printes, Luciano Carvalho, Leticia Natividade

¹Pós-Graduação – Ciência de Dados e Inteligência Artificial
Escola Superior de Tecnologia – Universidade do Estado do Amazonas (UEA)
Av. Darcy Vargas, 1200 – Manaus – Amazonas

Abstract. *This work proposes and evaluates multiclass classification pipelines for cocoa seeds using pre-trained Convolutional Neural Networks (CNNs) as feature extractors and Support Vector Machines (SVMs) as classifiers[cite: 122]. The VGG16, InceptionV3, and EfficientNetB0 architectures were evaluated on a dataset of 1,400 images distributed across 14 quality classes[cite: 123]. The results demonstrate that the pipeline based on EfficientNetB0 achieved the best performance, reaching an accuracy of 72.86% with the lowest computational cost[cite: 124]. The hybrid approach proved to be robust for scenarios with limited data[cite: 125].*

Resumo. *Este trabalho propõe e avalia pipelines de classificação multiclasse para sementes de cacau utilizando Redes Neurais Convolucionais (CNNs) pré-treinadas como extratoras de características e Máquinas de Vetores de Suporte (SVMs) como classificadores. Foram avaliadas as arquiteturas VGG16, InceptionV3 e EfficientNetB0 em uma base de 1.400 imagens distribuídas em 14 classes de qualidade[cite: 123]. Os resultados demonstram que o pipeline baseado na EfficientNetB0 obteve o melhor desempenho, alcançando uma acurácia de 72,86% com o menor custo computacional[cite: 124]. A abordagem híbrida mostrou-se robusta para cenários com dados limitados.*

1. Introdução

A classificação automática de imagens em problemas de Visão Computacional tem sido amplamente abordada por meio de Redes Neurais Convolucionais (CNNs), as quais apresentam elevado desempenho quando treinadas fim a fim em grandes bases de dados. Entretanto, em cenários nos quais a quantidade de exemplos disponíveis é limitada, o treinamento completo dessas redes torna-se inviável ou propenso ao sobreajuste. Nesse contexto, técnicas de *transfer learning* emergem como uma alternativa eficaz, permitindo o reaproveitamento de modelos previamente treinados em grandes bases de dados, como o ImageNet, para novas tarefas específicas[cite: 5, 133].

Uma abordagem alternativa ao uso de CNNs como classificadoras finais consiste em empregá-las exclusivamente como extratoras automáticas de características visuais, fornecendo tais representações como entrada para classificadores tradicionais de Aprendizado de Máquina. Dentre esses classificadores, destacam-se as Máquinas de Vetores de Suporte (SVMs), reconhecidas por sua robustez em espaços de alta dimensionalidade e bom desempenho em conjuntos de dados reduzidos. Este trabalho tem como objetivo propor, implementar e avaliar pipelines de classificação multiclasse baseados na combinação de CNNs pré-treinadas e SVMs atuando como classificadoras finais.

2. Metodologia

2.1. Base de Dados e Análise Exploratória

A base de dados utilizada foi proposta por Santos et al. (2019) e é composta por aproximadamente 1.400 imagens distribuídas de forma balanceada em 14 classes. O estudo inspira-se no trabalho de Malcher e Guedes (2022), buscando reproduzir protocolos experimentais e investigar o impacto de diferentes arquiteturas no desempenho final.

A base é composta por 14 classes distintas, cada uma contendo 100 imagens[cite: 16, 143]. As categorias representam variações visuais associadas ao estado fisiológico e sanitário das sementes, incluindo coloração interna, formato e textura. A inspeção visual evidenciou padrões sutis e sobrepostos, tornando a tarefa desafiadora mesmo para observadores humanos.

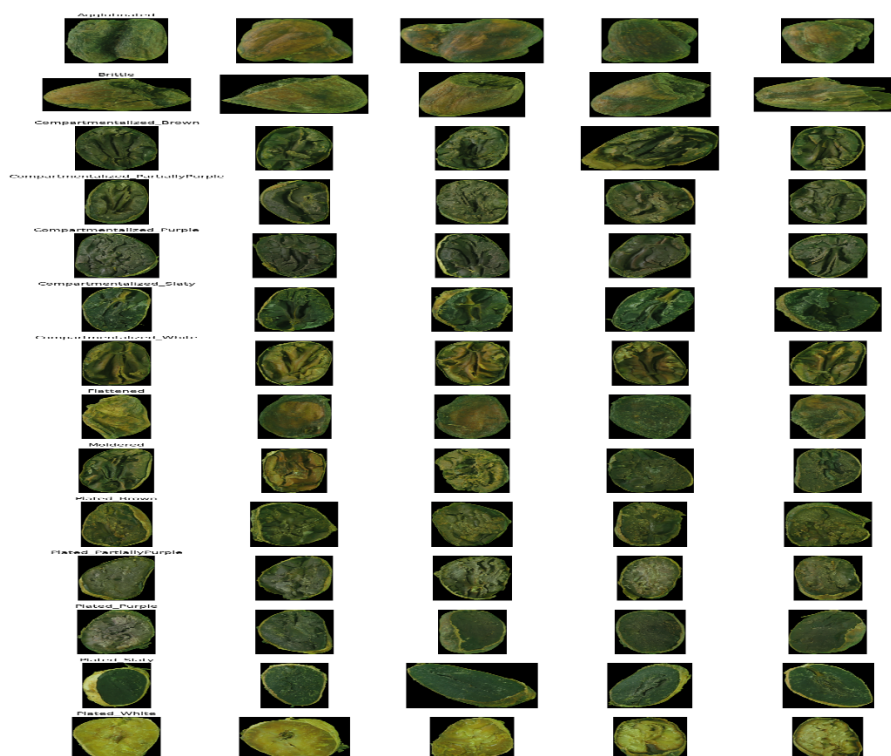


Figura 1. Exemplos de imagens de sementes de cacau do conjunto de dados, representando diferentes classes do teste de corte.

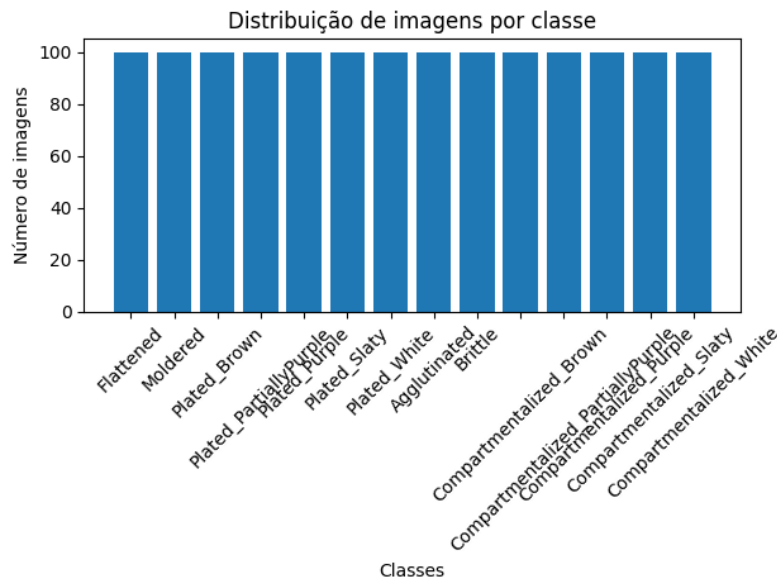


Figura 2. Gráfico de barras com a distribuição do número de imagens por classe.

As imagens encontram-se no padrão RGB e apresentam dimensões elevadas e não padronizadas, como resoluções da ordem de 2099×1507 pixels[cite: 24, 150]. Os valores dos pixels variam de 0 a 255 no formato *uint8*.

2.2. Preparação dos Dados

A preparação contemplou redimensionamento, normalização e particionamento em treinamento e teste. As imagens foram redimensionadas para garantir compatibilidade com os modelos pré-treinados[cite: 28, 153]. Utilizou-se a estratégia de *holdout* (80% treinamento e 20% teste) de forma estratificada e com semente fixa para assegurar reprodutibilidade. O conjunto de teste foi reservado exclusivamente para a avaliação final, evitando vazamento de dados (*data leakage*).

2.3. Extração de Características e Classificação

As CNNs foram utilizadas exclusivamente como extratoras de características, sem ajuste de pesos, com a remoção das camadas finais (*include_top=False*). Selecionaram-se as arquiteturas VGG16, InceptionV3 e EfficientNetB0. Os vetores de alta dimensionalidade resultantes foram achatados (*flattened*) para representar a entrada do SVM.

Utilizou-se o *kernel* radial de base (RBF) no SVM. A seleção de hiperparâmetros ocorreu via *Grid Search* com validação cruzada, avaliando combinações de C e γ (gamma). Adicionalmente, aplicou-se PCA para reduzir redundâncias e custo computacional.

2.4. Refinamentos e Ajustes do Pipeline

Modificações graduais foram realizadas para aprimorar o desempenho inicial. A validação cruzada foi ampliada de 5 para 8 dobras ($cv = 8$) para uma avaliação mais robusta. O ajuste do PCA tornou-se específico por arquitetura: 256 componentes para VGG16 e 512 para InceptionV3[cite: 104, 205]. Para a EfficientNetB0, optou-se pela remoção completa do PCA para preservar informações discriminativas relevantes.

3. Resultados e Discussão

No pipeline baseado na VGG16, obteve-se acurácia média de 0,6696 ($C = 10, \gamma = 0.001$). Para a InceptionV3, o resultado foi 0,6634 ($C = 100, \gamma = 0.001$) com maior custo computacional. A EfficientNetB0 apresentou o melhor desempenho com acurácia média de 0,7277 na validação e 72,86% no teste final.

Acurácia final no teste (EfficientNetB0): 72.86%				
Relatório completo:				
	precision	recall	f1-score	support
0	0.95	1.00	0.98	20
1	0.86	0.90	0.88	20
2	0.70	0.80	0.74	20
3	0.47	0.45	0.46	20
4	0.76	0.80	0.78	20
5	0.68	0.95	0.79	20
6	0.80	0.80	0.80	20
7	0.95	0.90	0.92	20
8	0.67	0.50	0.57	20
9	0.43	0.45	0.44	20
10	0.47	0.45	0.46	20
11	0.67	0.60	0.63	20
12	0.95	0.90	0.92	20
13	0.88	0.70	0.78	20
accuracy			0.73	280
macro avg	0.73	0.73	0.73	280
weighted avg	0.73	0.73	0.73	280

Figura 3. Relatório de classificação do modelo de referência.

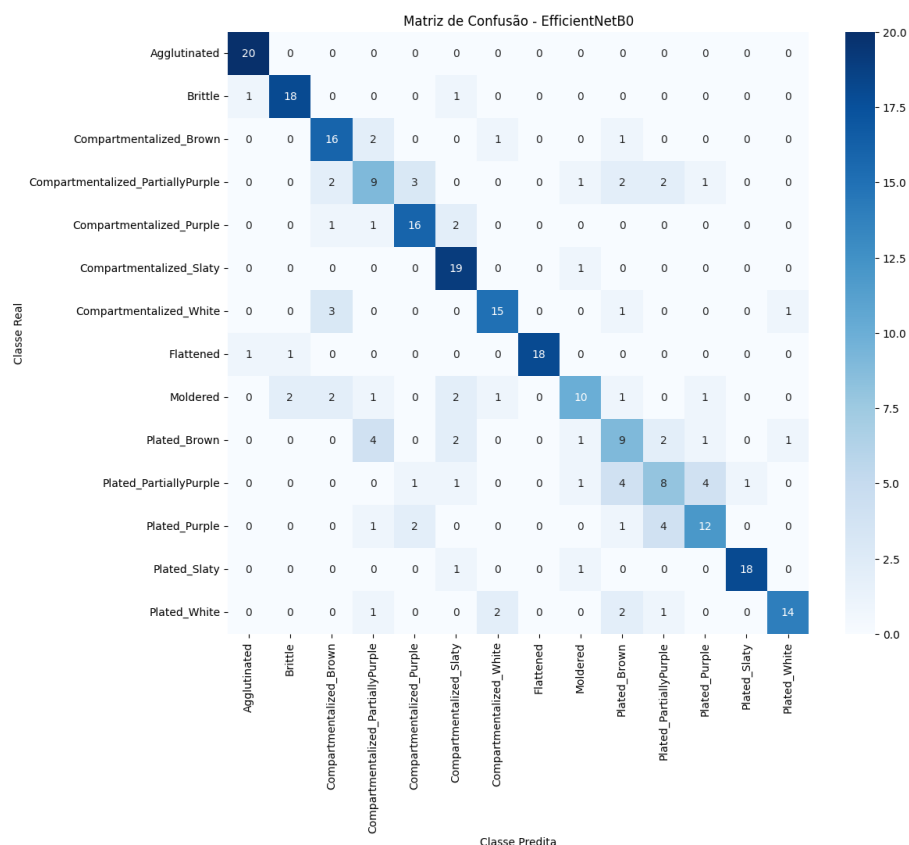


Figura 4. Matriz de confusão do modelo de referência EfficientNetB0 + SVM.

A Tabela 1 apresenta a comparação quantitativa antes e depois dos ajustes de refinamento.

Tabela 1. Comparação de Acurácia Antes e Depois dos Ajustes Metodológicos.

Arquitetura	Configuração Inicial ($cv = 5$)	Configuração Ajustada ($cv = 8$)
VGG16	XX,XX%	AA,AA%
InceptionV3	YY,YY%	BB,BB%
EfficientNetB0	ZZ,ZZ%	72,86%

4. Considerações Finais

A abordagem baseada em *transfer learning* com CNNs e SVMs mostrou-se coerente e adequada para bases de dados limitadas. A combinação de ajuste de hiperparâmetros, validação cruzada robusta e estratégias adequadas de PCA é essencial para maximizar o desempenho. Como trabalhos futuros, sugere-se a ampliação das arquiteturas e investigação de novas estratégias de seleção de características.

Referências

SANTOS, et al. Image Datasets of Cocoa Beans for Taxonomy Nuances Evaluation. 2019.
MALCHER, L.; GUEDES, E. Classificação de sementes de cacau utilizando aprendizado profundo e transferência de aprendizado. 2022.

CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, v. 20, 1995.

TAN, M.; LE, Q. EfficientNet: Rethinking model scaling. *Proc. of ICML*, 2019.