

Comparação de Métodos baseados em Redes Convolucionais para Classificação de Fava

Erico Andre da Silva¹

¹Departamento de Estatística e Informática – Universidade Federal Rural de Pernambuco
Rua Dom Manuel de Medeiros, s/n, - CEP: 52171-900 – Recife – PE – Brasil

erico.andre@ufrpe.br

Resumo. *A cultura da fava tem recebido pouca atenção por parte dos órgãos de pesquisa e extensão, o que tem resultado em limitações do conhecimento das características agrônômicas da cultura. Isso tem afetado a precisão em classificá-las. Tal classificação é de grande importância porque a identificação correta de plantas permite boa resposta da cultura em termos de produtividade e comportamento em diferentes condições ambientais. Neste contexto de informações limitadas sobre as características apresentamos uma solução que aplica o poder da visão computacional à agronomia, que visa melhorar a produtividade, reduzir desperdícios e auxiliar na tomada de decisões e na seleção de cultura que mais se adequa a uma região em particular. As técnicas de visão computacional são um conjunto de métodos utilizados para interpretar imagens, extraindo padrões e características. Visando contribuir com esse cenário de desenvolvimento tecnológico do setor do agrônomo, este trabalho compara algumas das abordagens de classificação supervisionada para identificação de forma automática de espécies de favas. O escopo deste trabalho consiste em classificar imagens de mudas geradas por produtores rurais em duas categorias de favas: orelha de vó e cearense. A partir das comparações realizadas entre métodos de classificadores que utilizam redes convolucionais como extratores de características com diferentes classificadores como máquina de vetores de suporte (SVM), árvore de decisão e mlp, para ao final apresentarmos a melhor metodologia para automatizar a classificação das imagens de favas.*

1. Introdução

A fava (*Phaseolus lunatus* L.) também conhecida como feijão-lima ou feijão-fava é uma leguminosa cultivada em quase todas as regiões do mundo, sendo que no Brasil possui ampla distribuição em todo o território, principalmente no Nordeste, é uma das quatro espécies do gênero *Phaseolus* exploradas comercialmente. A espécie foi domesticada na América do Sul ou Central, ou em ambas, e é subtropical [1]. É uma das principais leguminosas cultivadas na região tropical e apresenta potencial para fornecer proteína vegetal à população, diminuindo a dependência quase exclusiva dos feijões do grupo carioca [2].

A humanidade vem busca por meios que maximize o processo agrícola, evitando desperdícios e aumentando a produtividade agrícola, com a classificação das mudas de forma rápida e correta e possível escolher o melhor manuseio e um maior controle da produção.

A importância de se identificar as espécies de fava se dá porque cada variedade tem características que se diferem uma das outras como o tamanho de vagens e quantidade

de grãos que podem variar para mais ou menos por safra, já a variedade orelha de vó tem o maior comprimento de suas vagens e maior peso de sementes que variedade fava-cearense. A escolha dessas duas espécies foi para resolver um problema do curso de agronomia da Universidade Federal Rural de Pernambuco (UFRPE). Que era classificar de forma rápida e precisa as variedades de Fava Orelha de Vó e Fava Cearense.

Espera-se por meio desta pesquisa elaborar um métodos baseados em redes convolucionais com extrator de características na classificação de variedades de fava. Este trabalho foca-se na construção de um agente classificador utilizando redes convolucionais como extrator de características em diferentes abordagem de classificadores para distinguir espécies de feijão-lima ou feijão-fava, dentre elas a fava cearense e a orelha de vó auxiliando o agricultor na escolha do manejo adequado a espécie classificada.

2. Comparação de Métodos baseados em Redes Convolucionais para Classificação de Fava

2.1. Aquisição das imagens e pré-processamento

As imagens foram disponibilizadas por um aluno do curso de agronomia da Universidade Federal Rural de Pernambuco, que apresentou o seguinte questionamento se era possível classificar as mudas de feijão-fava já que elas eram quase idênticas utilizando métodos de aprendizado computacional. O conjunto possui duas categorias de feijão-fava as Fava Cearense e Fava Orelha de Vó, totalizando 229 imagens, sendo 109 de Fava Cearense figura 1a, e 120 de Fava Orelha de Vó figura 1b.



(a) Cearense

(b) Orelha de Vó

Devido a quantidade limitada de imagens utilizados nesse trabalho, as operações de acréscimo de dados (data augmentation), para aumentar a quantidade e a diversidade dos dados do conjunto de imagens, para evitar o overfitting da rede e garantir a sua generalização. Os métodos descritos foram implementados em Python, utilizando as bibliotecas Opencv (Open Source Computer Vision Library) ¹, uma biblioteca de software de visão computacional e aprendizado de máquina de código aberto disponível em Python. As manipulação nas imagens foram: espelhamento horizontal e vertical, variação do nível

¹<https://opencv.org/about/>

de brilho, foi aplicado cinco transformações mais a copia da imagem original dando um total de seis gerada a partir de uma única imagem. Após esta etapa, o novo conjunto de imagens que foi gerada ficou com 654 imagem de Fava Cearense e 720 imagens de Orelha de Vó, totalizando 1374 imagens ². A tabela 1 mostra total das imagens antes e depois do o acréscimo de dados (data augmentation).



Técnica que ajuda na melhoria do modelo que é o acréscimo de dados (data augmentation) citada anteriormente. Todas as imagens possuem 3456 pixels de largura x 4608 pixels de altura, o que totaliza 47.775.744 atributos, que são provenientes dos três canais. Devido às limitações de hardware disponíveis, para que o treinamento dos algoritmos, foi necessário reduzir o tamanho da imagens em 200 pixels de largura x 267 pixels de altura, este é o menor valor obtido que não afeta a qualidade das imagens, o que resultou em 160.200 atributos para cada imagem, a figura 2c e figura 2d são imagens geradas com o método de acréscimo de dados onde seu tamanho foi reduzidos.

Conjunto	Fava Cearense	Orelha de Vó	Total
Imagens Originais	109	120	229
Imagens Aumentadas	654	720	1374

Tabela 1. Distribuição dos conjuntos de imagens

²https://github.com/ericoandre/reconhecimento_padroes

2.2. Rede Neural Convolucional

Uma Rede Neural Convolucional (CNN do inglês Convolutional Neural network ou ConvNet) é uma variação das rede neural feedforward de Múltiplas Camadas [3], foi inspirada no processo biológico de visão dos seres vivos. Capaz de aplicar filtros aos dados, mantendo a relação de vizinhança entre os pixels da imagem ao longo do processamento da rede (VARGAS; PAES; VASCONCELOS, 2016).

A arquitetura inicia proposta neste estudo consiste em quatro tipos básicos de camadas: 2 camada de convolução (Convolutional Layer) com filtros de tamanho 3×3 e 2x2, 1 camada de subamostragem (Subsampling Layers ou Pooling Layers), 1 camada de Dropout e por fim a totalmente conectadas (Fully Connected Layers), como mostrada na figura 3, a tabela 2 mostra as distribuições dos hiperparâmetros que foram utilizado selecionados pela busca aleatória [4].

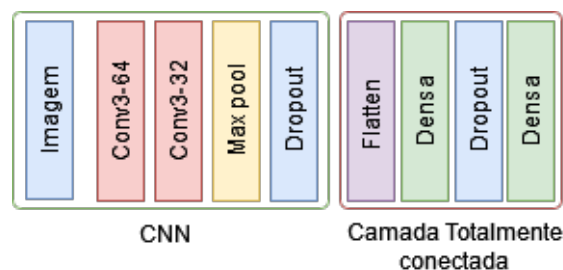


Figura 3. Arquitetura da CNN

Fonte: Elaborada pelo autor (2022)

Tabela 2. Hiperparâmetros do Modelo Inicial

Algoritimos	Hiperparâmetros	Distribuição de Probabilidade
CNN MLP	dropout	uniforme em {0.5, 0.6, 0.7, 0.8, 0.9, 0.10, 0.11, 0.12}
CNN MLP	neurônios	uniforme em {32, 64, 128, 256}
CNN MLP	ativação	uniforme em {relu, sigmoid}
CNN MLP	épocas	uniforme em [50; 500]
CNN MLP	tamanho do batch	uniforme em {16, 32, 64, 128}
CNN MLP	otimizador	uniforme em {adam, sgd, rmsprop, nadam}

Foi separada uma sub amostra de 30% do conjunto de 1374 imagens totalizando 413 imagens, que foram utilizadas na validação dos modelos. Após isso, avaliação da rede convolucional, através da matriz de confusão notamos que das 413 acertou 182 da classe fava Cearense e 212 de fava Orelha de vó e teve instâncias teve 19 erros obtendo uma precisão de 0.94, com uma cobertura de 0.96 e a Área sob a curva AUC de 0.966.

Na tabela 3, mostra os hiperparâmetros relacionado a melhor CNN, obtidos através da seleção de modelos.

Tabela 3. Hiperparâmetros do Melhor Modelo de CNN

Algoritmos	Hiperparâmetros	Variação
CNN MLP	épocas	364
CNN MLP	batch size	16
CNN MLP	otimizador	adam
CNN MLP	ativação	relu
CNN MLP	neurônios	128
CNN MLP	dropout	0.1

2.3. CNN como Extrator de Características

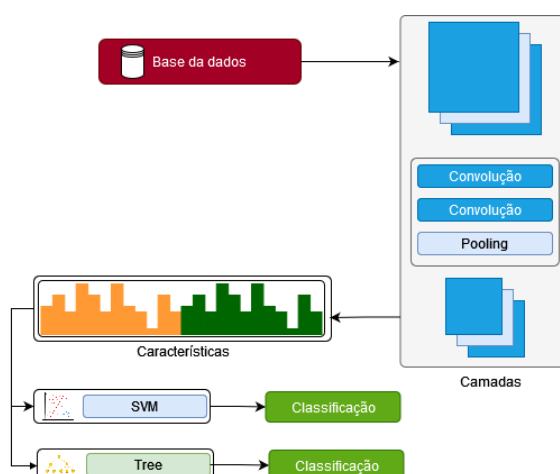


Figura 4. Representação do método proposto nos experimentos

Fonte: Elaborada pelo autor (2022)

1. Nesta abordagem construímos um modelo de rede neural convolucional tradicional CNN sua arquitetura é mostrada na figura 3, foi treinada para classificar as instancias de feijão-fava, encontrando ótimos resultados com mínimo de erro, obtendo valores ótimos de classificação de feijão-fava.
2. Na segunda abordagem o modelo de CNN gerado inicialmente foi utilizado como extrator de características, onde foi removida a camada de classificação softmax. Assim, os valores de ativação dos neurônios da rede que estariam ligados nesta camada são utilizados como entrada à Máquina de Vetores de Suporte (SVM) e também para a árvore de decisão como mostra na figura 4.

2.4. SVM

O SVM tem o objetivo encontrar o hiperplano de separação ideal o qual maximiza a margem da base de treinamento buscando o equilíbrio entre ambos os erros, minimizando o excesso de ajustes (overfitting) e melhorando a capacidade de generalização [5].

Utilizando o modelo de CNN descrito a cima como extrator de característica, onde foi removendo sua última camada a de classificação softmax. Assim, os valores

de ativação dos neurônios da rede que estariam ligados nesta camada são utilizados como entrada em uma Máquina de Vetores de Suporte (SVM) para a classificação. A tabela 4 mostra as distribuições dos hiperparâmetros que foram utilizado selecionados pela busca aleatória [4].

Tabela 4. Hiperparâmetros

Algoritmos	Hiperparâmetros	Distribuição de Probabilidade
CNN SVM	C	uniforme em [0.0001; 100.00]
CNN SVM	kernel	uniforme em {poly, rbf, linear}
CNN SVM	Grau do polinômio	uniforme em [2; 5]
CNN SVM	γ	uniforme em [0.0001; 100.00]

Os Resultados obtidos através da matriz de confusão, que das 413 instâncias acertou 197 da classe fava cearense e 201 de fava orelha de vó e teve 6 erros obtendo uma precisão de 0.98, com uma cobertura de 0.99 e a Área sob a curva AUC de 0.988.

Na tabela 5, mostra os hiperparâmetros relacionado ao modelo que obteve os melhores resultados, obtidos através da seleção de modelos.

Tabela 5. Hiperparâmetros do Melhor Modelo SVM

Algoritmos	Hiperparâmetros	Variação
CNN SVM	C	2.91×10^1
CNN SVM	kernel	poly
CNN SVM	Grau do polinômio	2
CNN SVM	γ	4.78×10^{-4}

2.5. Árvore de decisão

As árvores são formadas por nós, que armazenam informação (perguntas). O nó raiz é o nó que possui maior nível hierárquico e, a partir dele, ramificam-se os nós filhos. O nó que não possui filhos é conhecido como nó folha ou terminal. Para se gerar um novo galho é necessário que ocorra um separação dos dados no plano. No entanto, essa separação precisa melhorar a separação anterior, caso contrário não faz sentido separar mais os dados.

Da mesma forma, Utilizando CNN com extrator de característica, removendo sua última camada a de classificação softmax. Assim, os valores de ativação dos neurônios da rede que estariam ligados nesta camada são utilizados na entrada de uma Árvore de Decisão para a classificação como mostra na figura 4. A tabela 6 mostra as distribuições dos hiperparâmetros que foram utilizado selecionados pela busca aleatória [4].

Tabela 6. Hiperparâmetros

Algoritmos	Hiperparâmetros	Distribuição de Probabilidade
CNN TREE	critério	uniforme em {gini, entropy}
CNN TREE	divisor	uniforme em {best, random}
CNN TREE	amos mín dividida	uniforme em [0; 1]
CNN TREE	amos mín por folha	uniforme em [0.00000; 0.5]
CNN TREE	máx de nós por folha	uniforme em [1; 1000]
CNN TREE	profundidade máx	uniforme em [1; 1000]

Os Resultados obtidos através da matriz de confusão, que das 413 instâncias acertou 186 da classe fava Cearense e 216 de fava Orelha de Vó e teve 11 erros obtendo uma precisão de 0.96, com uma cobertura de 0.98 e a Área sob a curva AUC de 0.979.

Na tabela 7, mostra os hiperparâmetros relacionado ao modelo que obteve os melhores resultados, obtidos através da seleção de modelos.

Tabela 7. Hiperparâmetros do Melhor Modelo Arvore

Algoritmos	Hiperparâmetros	Variação
CNN TREE	criterion	gini
CNN TREE	max depth	436
CNN TREE	max leaf nodes	861
CNN TREE	min samples leaf	9.17×10^{-2}
CNN TREE	min samples split	7.79×10^{-1}
CNN TREE	splitter	best

2.6. VGGNet

Este experimento foi realizado de forma adicional, utilizando Transferência de Aprendizado (transfer learning) como arquitetura vgg16 treinados com os pesos da imagenet [6], com extrator de característica, da mesma forma como foi feito com o modelo de CNN descrito a cima, removendo a camada totalmente conectada e substituindo por uma arquitetura de MLP, com a mesma configuração utilizada nas camadas totalmente conectadas da CNN. A tabela 8 mostra as distribuições dos hiperparâmetros que foram utilizado selecionados pela busca aleatória [4].

Tabela 8. Hiperparâmetros

Algoritmos	Hiperparâmetros	Distribuição de Probabilidade
CNN VGG16	dropout	uniforme em {0.5, 0.6, 0.7, 0.8, 0.9, 0.10, 0.11, 0.12}
CNN VGG16	neurônios	uniforme em {32, 64, 128, 256}
CNN VGG16	ativação	uniforme em {relu, sigmoid}
CNN VGG16	épocas	uniforme em [50; 500]
CNN VGG16	tamanho do batch	uniforme em {16, 32, 64, 128}
CNN VGG16	otimizador	uniforme em {adam, sgd, rmsprop, nadam}

Os Resultados obtido com está abordagem, através da matriz de confusão, que das 413 instâncias acertou 186 da classe fava cearense e 212 de fava Orelha de Vó e teve 15 erros obtendo uma precisão de 0.95, com uma cobertura de 0.96 e a Área sob a curva AUC de 0.974.

Na tabela 9, mostra os hiperparâmetros relacionado ao modelo que obteve os melhores resultados, obtidos através da seleção de modelos.

Tabela 9. Hiperparâmetros do Melhor Modelo VGG16

Algoritmos	Hiperparâmetros	Variação
VGG16 MLP	épocas	166
VGG16 MLP	tamanho do batch	32
VGG16 MLP	otimizador	nadam
VGG16 MLP	ativação	relu
VGG16 MLP	neurônios	64
VGG16 MLP	dropout	0.1

3. Conclusão

Foi possível perceber que todos os métodos testados erraram poucas instância, indicando uma boa precisão nas classificações no experimento como mostra na tabela 10. De modo geral, os resultados com SVM foram superiores aos com o classificador baseado em Softmax e Árvore de decisão. Onde a máquina de vetores de suporte (SVM) obteve uma precisão de 0.98 com 6 erros apenas 6 instâncias foi onde se obteve os melhores resultados seguido pela Árvore de decisão que apresentou uma precisão de 0.96.

Tabela 10. Contem as métricas precisão e Sensibilidade

Algoritmos	Precisão	Cobertura	AUC-ROC
CNN	0.946429	0.968037	0.966
CNN SVM	0.981308	0.990566	0.988
CNN TREE	0.964286	0.986301	0.979
VGG16 MLP	0.958762	0.963730	0.974

Os sistemas de visão computacional têm sido cada vez mais utilizados na agricultura facilitando a tomada de decisão dos agricultores. O presente trabalho apresentou um método para identificação de espécies e classificação de mudas de feijão-fava. Neste trabalho, quatro modelos foram comparados com a finalidade de determinar qual seria o melhor classificador. Como visto no capítulo anterior, a junção de SVM com as camadas de convolução na etapa de extração de características se mostrou eficaz, com apenas 6 erros obtendo uma precisão de 0.98 e cobertura de 0.99.

Como trabalhos futuros, pretendemos explorar mais variedades como boca-de-moça, branquinha, mororó, olho-de-ovelha, olho-de-peixe, raio-de-sol, rajada vermelha, rajada preta. Obter recursos de hardware mais robustos, para estender a seleção de modelos.

Referências

- [1] MJ de O Zimmermann and MG Teixeira. Origem e evolução. *ARAÚJO, R*, 1996.
- [2] RF Vieira. A cultura do feijão-fava. *Informe Agropecuário*, 16(174):30–37, 1992.
- [3] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [4] Bergstra, James and Bengio, Y. Random Search for Hyper-Parameter Optimization. *The Journal of Machine Learning Research*, 13:281–305, 03 2012.
- [5] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer, New York, 1995.
- [6] ILSVRC-201. ImageNet. <https://www.image-net.org/challenges/LSVRC/2014/>, 2022. Accessed: 2022-04-01.