



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE DESENVOLVIMENTO SUSTENTÁVEL DO SEMIÁRIDO
UNIDADE ACADÊMICA DE TECNOLOGIA DO DESENVOLVIMENTO
CURSO SUPERIOR DE TECNOLOGIA EM AGROECOLOGIA**

GABRIEL DE AZEVEDO BATISTA

**IDENTIFICAÇÃO DE CULTIVARES DE FEIJÃO CAUPI A PARTIR DE
IMAGENS DIGITAIS DE SEMENTES E TÉCNICAS DE
INTELIGÊNCIA ARTIFICIAL**

**SUMÉ - PB
2023**

GABRIEL DE AZEVEDO BATISTA

**IDENTIFICAÇÃO DE CULTIVARES DE FEIJÃO CAUPI A PARTIR DE
IMAGENS DIGITAIS DE SEMENTES E TÉCNICAS DE
INTELIGÊNCIA ARTIFICIAL**

Monografia apresentada ao Curso Superior de Tecnologia em Agroecologia do Centro de Desenvolvimento Sustentável do Semiárido da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Tecnólogo em Agroecologia.

Orientador: Dr. Rener Luciano de Souza Ferraz.

**SUMÉ - PB
2023**



B333i Batista, Gabriel de Azevedo.

Identificação De cultivares de feijão caupi a partir de imagens digitais de sementes e técnicas de inteligência artificial. / Gabriel de Azevedo Batista. - 2023.

32 f.

Orientador: Professor Dr. Renner Luciano de Souza Ferraz.

Monografia - Universidade Federal de Campina Grande; Centro de Desenvolvimento Sustentável do Semiárido; Curso Superior de Tecnologia em Agroecologia.

1. Aprendizagem de máquina. 2. Identificação fenotípica. 3. Feijão caupi. 4. Imagens digitais de sementes. 5. Vigna unguiculata L. 6. Processamento de imagens digitais de sementes. 7. Tecnologia de sementes. I. Ferraz, Renner Luciano de Souza. II. Título.

CDU: 631.53.01(043.1)

Elaboração da Ficha Catalográfica:

Johnny Rodrigues Barbosa
Bibliotecário-Documentalista
CRB-15/626

GABRIEL DE AZEVEDO BATISTA

**IDENTIFICAÇÃO DE CULTIVARES DE FEIJÃO CAUPI A PARTIR DE IMAGENS
DIGITAIS DE SEMENTES E TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL**

**Trabalho de conclusão de curso,
apresentado ao curso de Tecnologia em
Agroecologia, do Centro de
Desenvolvimento Sustentável do Semiárido
da Universidade Federal de Campina
Grande, como requisito parcial para a
obtenção do título de Tecnólogo em
Agroecologia.**

BANCA EXAMINADORA:

**Professor Dr. Rener Luciano de Souza Ferraz
Orientador – UATEC/CDSA/UFPG**

**Professor Me. Miguel Avelino Barbosa Neto
Examinador I – CCA/UFPG**

**Esp. Gabriella Henrique Brandão
Examinadora II – CCHSA/UFPG**

**Me. Igor Eneas Cavalcante
Examinador III – CCA/UFPG**

Trabalho Aprovado em: 16 de Fevereiro de 2023

SUMÉ – PB

Dedico esse trabalho aos meus pais (Risoaldo Batista de Deus e Gilvania de Azevedo Batista), ao meu irmão e minha cunhada (Rafael de Azevedo Batista e Sivoney Cavalcante Batista) e a minha sobrinha (Adriely Cavalcante Batista).

AGRADECIMENTOS

Gratidão é o sentimento que tenho para com Deus, pois ele foi essencial em todas as minhas conquistas e superações.

Ao meu pai, Risoaldo por todo empenho, dedicação e incentivo. Tenho muita admiração pelo senhor e orgulho por ser seu filho!

À minha mãe, Gilvânia, por ser aquela pessoa que me dá forças, me ajuda, estando sempre presente.

Ao meu irmão Rafael e à minha cunhada Sivoney, que me acolheram em sua casa durante esses anos e por serem esses seres maravilhosos que levo comigo no coração para onde for obrigado por todo apoio, carinho e atenção.

Agradeço demais ao meu orientador, Professor Dr. Rener Luciano de Souza Ferraz, um grande amigo que a universidade me deu, pessoa de um coração imenso e que me acolheu nos seus projetos acadêmicos. O seu empenho, dedicação e incentivo foi essencial para a minha motivação à medida que as dificuldades iam surgindo ao longo da realização deste trabalho.

Aos professores, por todos os conselhos, pela ajuda e pela paciência com a qual guiaram meu aprendizado.

Aos amigos que fizeram parte da minha graduação, pessoas que tive o privilégio de conviver durante esse tempo, Joniedson Marcos, José Ilton, Marcelo Eduardo, Lucas Mayan, Rogério Andrade, Matheus Procópio e Lucas Santana. Obrigado a cada um de vocês.

Aos amigos que fiz durante o estágio e em outros projetos acadêmicos, Maria Fernanda, Rosana Lucena, Ana Clara, Davi Siqueira, Deyvid Mendes, além dos funcionários e colaboradores que trabalham no CDSA, em especial aos Srs. Edilson e Paulo, pessoas maravilhosas que tive o prazer em conhecer.

À Pró-Reitoria de Pós-Graduação da UFCG pela concessão de bolsa do Programa Institucional de Voluntário em Iniciação Científica – PIVIC. E a todos que me ajudaram na realização deste trabalho.

RESUMO

O feijão-caupi é uma cultura de grande importância mundial, motivo pelo qual muitas cultivares melhoradas são exploradas. As sementes desses materiais possuem características intrínsecas que os distinguem. Assim, o objetivo –se por meio deste trabalho ajustar os modelos de aprendizagem de máquina para identificação de feijão-caupi a partir do processamento de imagens digitais de sementes utilizando técnicas de inteligência artificial. A pesquisa é do tipo quali-quantitativas, e foi realizada no ambiente de professores, sala 03. Para tanto, imagens digitais de sementes de 10 (dez) cultivares foram obtidas e processadas utilizando-se dos vetorizadores InceptionV3, SqueezeNet, VGG16 e VGG19. Posteriormente, foram testados os algoritmos de aprendizado de Máquina K-Vizinhos Mais Próximos (KNN - number of nearest neighbors), Árvore de Decisão (Tree), Floresta Aleatória (RF - Random Forest), Aumento de Gradiente (GB - Gradient Boosting), Máquina de Suporte de Vetores (SVM - Support Vector Machines) e Rede Neural Artificial (MLP - Multi-Layer Perceptron). Os algoritmos de aprendizado de máquina Rede Neural Artificial e Máquina de Suporte de Vetores (SVM) tiveram melhores indicadores de performance para identificação de feijão-caupi a partir do processamento das imagens digitais de sementes. Esse trabalho vem a contribuir, tanto no âmbito acadêmico, quanto no aspecto prático, tendo em vista que os dados disponíveis podem servir futuramente para criação de aplicativos mobile.

Palavras-chave: *Vigna unguiculata L.*, classificação fenotípica, aprendizagem de máquina.

BATISTA, Gabriel Azevedo de. **Cowpea identification from digital images of seeds and artificial intelligence techniques**. 2023. 32f. (Course Completion Work - Monograph), Curso Superior de Tecnologia em Agroecologia, Centro de Desenvolvimento Sustentável do Semiárido, Universidade Federal de Campina Grande, Sumé – Paraíba – Brazil, 2023.

ABSTRACT

Cowpea is a crop of great importance worldwide, which is why many native varieties and improved cultivars are explored. The seeds of these materials have intrinsic characteristics that distinguish them. Thus, the aim of this work was to adjust the machine learning models for cowpea identification based on the processing of digital images of seeds using artificial intelligence techniques. The research is of the qualitative and quantitative type, and was carried out in the teachers' environment, room 03. Therefore, digital images of seeds of 10 (ten) cultivars were obtained and processed using the vectorizers InceptionV3, SqueezeNet, VGG16 and VGG19. Subsequently, the learning algorithms of K-Nearest Neighbors (KNN - number of nearest neighbors), Decision Tree (Tree), Random Forest (RF), Gradient Boosting (GB), Vector Support Machine (SVM - Support Vector Machines) and Artificial Neural Network (MLP - Multi-Layer Perceptron). The machine learning algorithms Artificial Neural Network and Vector Support Machine (SVM) had better performance indicators for cowpea identification from the processing of digital images of seeds. This work contributes, both in the academic field and in the practical aspect, considering that the available data can serve in the future for the creation of mobile applications.

Keywords: *Vigna unguiculata L.*, phenotypic classification, machine learning.

LISTA DE FIGURAS

Figura 1 - Localização da sala 03 do ambiente de professores do Centro de Desenvolvimento Sustentável do Semiárido da Universidade Federal de Campina Grande.. 16

Figura 2 - Visão parcial da captura das imagens digitais das sementes das cultivares de feijão-caupi.. 17

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 - Estatísticas de desempenho dos Embedder na etapa de modelagem. Sumé, PB, 2022. | 20 |
| Tabela 2 - Matriz de confusão para o desempenho do algoritmo Rede Neural Network na etapa de modelagem para processamento das imagens de sementes de feijão-caupi. Sumé, PB, 2022. | 21 |
| Tabela 3 - Estatísticas de desempenho para o algoritmo Rede Neural Network com suas variações das funções de ativação operando em diferentes Solvers. Sumé, PB, 2022. | 22 |
| Tabela 4 - Matriz de confusão para o desempenho do algoritmo Rede Neural Network, selecionado com a função de ativação Logística e o solver Adam. Sumé, PB, 2022. | 23 |
| Tabela 5 - Estatísticas de desempenho do algoritmo SVM, selecionando o melhor tipo, operando em diferentes Kernel. Sumé, PB, 2022. | 23 |
| Tabela 6 - Matriz de confusão para o desempenho do algoritmo SVM. Sumé, PB, 2022. | 24 |
| Tabela 7 - Dados do algoritmo Rede Neural Artificial no processo de validação. Sumé, PB, 2022. | 25 |
| Tabela 8 - Matriz de confusão para o desempenho do algoritmo Rede Neural Network no processo de validação. Sumé, PB, 2022. | 25 |
| Tabela 9 - Dados do algoritmo SVM no processo de validação. Sumé, PB, 2022. | 26 |
| Tabela 10 - Matriz de confusão para o desempenho do algoritmo SVM, no processo de validação. Sumé, PB, 2022. | 27 |

SUMÁRIO

| | |
|---|-----------|
| 1 INTRODUÇÃO | 12 |
| 2 OBJETIVOS | 13 |
| 2.1 OBJETIVO GERAL | 13 |
| 2.2 OBJETIVOS ESPECÍFICOS | 13 |
| 3 REVISÃO DE LITERATURA..... | 14 |
| 3.1 PREÂMBULO ÀS PESQUISAS COM FEIJÃO-CAUPI..... | 14 |
| 3.2 IDENTIFICAÇÃO TRADICIONAL DE FEIJÃO-CAUPI..... | 14 |
| 3.3 IDENTIFICAÇÃO DE PLANTAS COM INTELIGÊNCIA ARTIFICIAL..... | 15 |
| 4 MATERIAL E MÉTODOS | 16 |
| 5 RESULTADOS E DISCUSSÃO | 20 |
| 5.1 VALIDAÇÃO DOS MODELOS..... | 25 |
| 6 CONCLUSÃO..... | 28 |
| REFERÊNCIAS | 29 |

1 INTRODUÇÃO

A cultura do feijão-caupi (*Vigna unguiculata* (L.) Walp.), também conhecido como feijão-de-corda ou macassar, tem grande importância no mundo, principalmente em regiões áridas e semiáridas onde a espécie constitui uma das principais fontes nutricionais. Por esse motivo, muitos genótipos, variedades, linhagens e cultivares são exploradas nos agroecossistemas (MARTEY *et al.*, 2021; TAVARES *et al.*, 2021; MARINHO *et al.*, 2021; NARAYANA & ANGAMUTHU, 2021).

Tendo em vista a existência de grande diversidade genética da espécie, considerando desde variedades crioulas, até cultivares melhoradas, pode-se inferir a ocorrência de diferenças marcantes nas características das sementes desses materiais. De fato, conforme observado no Catálogo de cultivares de feijão-caupi da Embrapa Meio Norte, as sementes possuem variações expressivas nos aspectos de cor, forma, tamanhos e massa, características que lhes confere identidade (EMBRAPA, 2021).

Pesquisas foram realizadas para comparações entre as características de sementes de cultivares de feijão-caupi, chegando-se à conclusão de que as características avaliadas são bons indicadores para diferenciação dos materiais genéticos (MIQUELONI *et al.*, 2018; SILVA *et al.*, 2020). Com isso, é possível que sejam desenvolvidas ferramentas tecnológicas para identificação rápida e precisa das cultivares dessa espécie a partir de características das sementes.

Com base no exposto, salienta-se que, com o avanço tecnológico, a agricultura tem incorporado novas ferramentas baseadas em computação para diversas finalidades. Com isso, o setor agrícola passou a contar com o uso de inteligência artificial e aprendizagem de máquina para identificação e classificação de plantas a partir de suas características intrínsecas, o que tem ganhado destaque devido ao fato de se basear em análises não-destrutivas das sementes e dos vegetais no campo (HATI & SINGH, 2021).

Técnicas de inteligência artificial e aprendizado de máquina já são conhecidas e aplicadas na região semiárida do Nordeste brasileiro com a finalidade de identificação de plantas (MENEZES NETO *et al.*, 2021). Contudo, ainda não foram relatadas pesquisas para identificação de feijão-caupi utilizando-se de processamento de imagens digitais das sementes e emprego de técnicas de inteligência artificial. Assim, novas pesquisas com essa finalidade podem introduzir desenvolvimento técnico-científico para a agricultura na região.

2 OBJETIVOS

2.1 OBJETIVO GERAL

O presente trabalho teve como objetivo, ajustar modelos de aprendizagem de máquina para identificação de feijão-caupi (*Vigna unguiculata* (L.) Walp) através de imagens digitais de sementes utilizando-se de técnicas de inteligência artificial.

2.2 OBJETIVOS ESPECÍFICOS

Obter sementes de cultivares melhorada de feijão-caupi cultivadas no semiárido brasileiro.

Obter imagens digitais de sementes de cultivares melhoradas de feijão-caupi em dois ângulos e quatro posições diferentes.

Processar imagens digitais de sementes e ajustar modelos de inteligência artificial para identificação de cultivares melhoradas de feijão-caupi.

Avaliar o desempenho dos modelos ajustados para identificação de cultivares melhoradas de feijão-caupi.

3 REVISÃO DE LITERATURA

3.1 PREÂMBULO ÀS PESQUISAS COM FEIJÃO-CAUPI

Atualmente, existem muitas cultivares melhoradas de feijão-caupi identificadas pelos agricultores familiares e lançadas pelos programas de melhoramento genético (GOMES *et al.*, 2021; MARINHO *et al.*, 2021; NARAYANA & ANGAMUTHU, 2021). Esses materiais são cultivados de forma experimental em diversos agroecossistemas e submetidos a variados sistemas de manejo, sobretudo para aumento de rendimento em condições de estresses bióticos e abióticos (SÁ *et al.*, 2021; TOYINBO *et al.*, 2021).

De fato, muitas pesquisas são desenvolvidas com feijão-caupi, por exemplo, utilizando doses de fósforo para homeostase iônica sob estresse salino (SÁ *et al.*, 2021), avaliando respostas da cultura à fertilização orgânica e mineral (SÁNCHEZ-NAVARRO *et al.*, 2021), respostas ecofisiológicas à deficiência hídrica (FERREIRA *et al.*, 2021) e biofortificação agrônômica de genótipos (BARBOSA *et al.*, 2021; SILVA *et al.*, 2021). Ainda são escassos os trabalhos na literatura que contempla um banco de dados de imagens digitais ou modelos de inteligência artificial para identificação e classificação de feijão-caupi a partir das características das sementes.

3.2 IDENTIFICAÇÃO TRADICIONAL DE FEIJÃO-CAUPI

Em relação às sementes, para identificação das cultivares, devem ser levantadas características como a cor da semente - determinada em sementes recém-colhidas e secas, de acordo com as seguintes cores: branca com hilo marrom, branca com hilo preto, branco vermelha, marrom-clara, marrom-escura, vermelha; cor do halo determinada em sementes recém-colhidas e secas, de acordo com as seguintes cores: amarelo, marrom, vermelho; brilho da semente - determinado em sementes secas: opaco, intermediário, brilhante; peso de 100 sementes (g) - peso médio em gramas, obtido em uma amostragem de quatro repetições de 100 sementes a 13,5% de umidade; produtividade - medida em kg ha⁻¹ (MARINHO *et al.*, 2001).

Com base nas características das plantas, tradicionalmente, a identificação de genótipos de feijão-caupi é realizada a partir do levantamento de dados relativos aos caracteres fenológicos, morfológicos e agrônômicos, nos estádios de germinação, plântulas, floração, maturação, colheita e pós-colheita (MARINHO *et al.*, 2001). Nesta perspectiva, os

estudos realizados por Freire Filho *et al.* (1981) e Fonseca *et al.* (1986), reportam os principais caracteres a serem levantados para fins de fenotipagem dos genótipos, a saber: Emergência de plântulas, cor do hipocótilo, floração média, cor da flor, porte da planta, pigmentação da haste principal, número de nós da haste principal, comprimento do folíolo central, largura do folíolo central, corda vagem durante a maturação, número de vagens por planta e número de sementes por vagem.

De acordo com Oliveira *et al.* (2015), a identificação clássica dos genótipos abrange, primeiramente, classificação segundo o Regulamento Técnico do Feijão, Instrução Normativa Número 12 de 28 de fevereiro de 2008, instituído pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA) (BRASIL, 2008), adaptado por Freire Filho *et al.* (2005). Em complemento, também são utilizados os descritores recomendados pelo Biodiversity International (2007), anteriormente conhecido como International Board for Plant Genetic Resources (IBPGR), considerando a abrangência mundial do documento em relação aos descritores estabelecidos pelo Ato Número 4 de 19 de agosto de 2010, instituído pelo MAPA (BRASIL, 2010).

3.3 IDENTIFICAÇÃO DE PLANTAS COM INTELIGÊNCIA ARTIFICIAL

A identificação de espécies por inteligência artificial é um processo de fundamental importância para suporte à agricultura de precisão (SUN *et al.*, 2021). Nos últimos anos, o aprendizado de máquina alcançou avanços significativos no reconhecimento de imagens digitais. Consequentemente, essa tecnologia se torna cada vez mais popular e constitui um ramo importante da agricultura inteligente com aplicação na identificação de espécies de plantas de interesse econômico, social e ambiental (XIONG *et al.*, 2021).

A tecnologia de identificação utilizando-se de aprendizagem de máquina tem sido utilizada pela comunidade científica para relacionar marcadores morfológicos, fisiológicos e moleculares às características de imagens digitais de plantas submetidas a estresses abióticos (SOLTABAYEVA *et al.*, 2021); para detecção de doenças em folhas de plantas (SUJATHA *et al.*, 2021); e reconhecimento e identificação de espécies (HATI & SINGH, 2021). Apesar da importância da cultura do feijão-caupi e da inteligência artificial para identificação de plantas no Nordeste do Brasil (MENEZES NETO *et al.*, 2021), ainda não foram realizadas pesquisas para identificação de cultivares melhoradas dessa espécie a partir de processamento de imagens digitais de sementes.

4 MATERIAL E MÉTODOS

A presente pesquisa é de caráter exploratório, em que por meio dela pode preencher lacunas em pesquisas futuras com esse tema.

As fotos das sementes foi realizada na sala 03 (três) do ambiente de professores e o processamento das imagens no programa foi realizada no laboratório de informática do CDSA. As sementes das cultivares melhoradas de feijão-caupi foram obtidas do banco de germoplasma do Instituto Agrônomo de Pernambuco (IPA). Foram utilizados os genótipos da Subclasse Comercial Branco rugoso (BRS Novaera), Subclasse Comercial Branco liso (BRS Guariba, BRS Potengi, BRS Cauamé), Subclasse Comercial Cores – Mulato (BRS Pajeú, BRS Miranda, BRS Pingo-de-Ouro e IPA 206), Subclasse comercial Preto-Brilhoso (BRS Tapahium).

As sementes das cultivares melhoradas de feijão-caupi foram transportadas para o Laboratório de Ecologia e Botânica do Centro de Desenvolvimento Sustentável do Semiárido (CDSA) da Universidade Federal de Campina Grande (UFCG), localizado no município de Sumé – PB, nas coordenadas de Latitude 7° 40' 18" S; Longitude 36° 52' 54" W e altitude de 518 m acima do nível do mar. O clima local, segundo a classificação de Köppen-Geiger, é do tipo Bsh (semiárido quente), com temperatura média anual de 26 °C e precipitação pluviométrica média anual de 600 mm (VITAL et al., 2020) (Figura 1).

Figura 1 - Localização da sala 03 do ambiente de professores do Centro de Desenvolvimento Sustentável do Semiárido da Universidade Federal de Campina Grande.

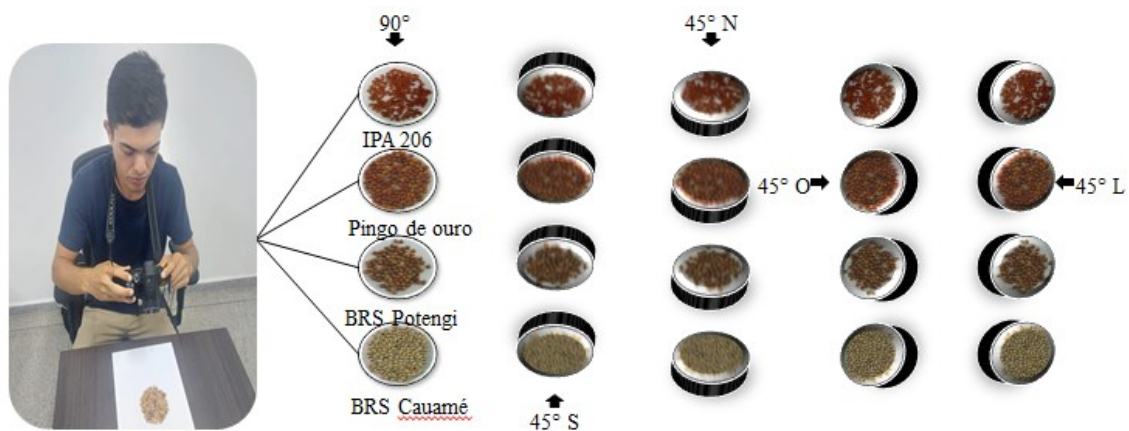


Fonte: Wikipedia e Google Maps. Sumé, PB, 2022.

Foram obtidas imagens de 10 genótipos de um total de 10 genótipos usados nesta pesquisa ou 125 imagens de cada cultivar, o que representa 100% de imagens obtidas. Para tanto, foi utilizada uma câmera digital da marca Nikon, modelo COOLPIX P530 V1.0,

configurada para ISSO de 400 e resolução de 16 MB em dimensões da imagem de 4608 x 3456 de comprimento e largura, respectivamente e 300 DPI (*Dots Per Inch* ou pontos por polegadas). As imagens foram capturadas no modo RGB de 8 bits e salvas no formato .jpeg. Foram obtidas 125 imagens de cada genótipo (amostras de 30 g de sementes), sendo 25 fotos nas seguintes posições: 25 fotos no ângulo de 90 graus, 25 fotos no ângulo de 45 graus na direção norte, 25 fotos no ângulo de 45 graus na direção sul, 25 fotos no ângulo de 45 graus na direção leste, e 25 fotos no ângulo de 45 graus na direção oeste, todas a uma distância média de 30 cm (Figura 2).

Figura 2 - Visão parcial da captura das imagens digitais das sementes das cultivares de feijão-caupi



. Fonte: Acervo da equipe de pesquisa. Sumé, PB, 2022.

Para processamento das imagens foi utilizado o software *Orange Data Mining* v. 3.29.3. Inicialmente, das 125 imagens obtidas, 100 foram importadas utilizando-se do *widget* “Import Images” contido no *add-on* “Image Analytics”. Posteriormente, as imagens foram processadas e vetorizadas utilizando-se do *widget* “Image Embedding”. A vetorização das imagens foi realizada utilizando-se dos “Embedders” InceptionV3, SqueezeNet, VGG16 e VGG19 (GODEC *et al.*, 2019).

Para ajuste dos modelos de identificação das cultivares de feijão caupi, serão testados algoritmos de aprendizagem de máquina e suas respectivas variações até que melhores valores de performance fossem obtidas para cada modelo, a saber:

O algoritmo “K-vizinhos mais próximos (KNN - number of nearest neighbors)” procura *K* exemplos de treinamento mais próximos no espaço de recursos e usa sua média como previsão. Serão testadas como métricas dos modelos as distâncias Euclidiana, Manhattan, Máxima e Mahalanobis, além de variações entre pesos uniformes e distância como critérios dos modelos. O pré-processamento padrão do algoritmo será mantido.

O algoritmo “Árvore de decisão (Tree)” divide os nós por pureza de classes a partir de dados contínuos e discretos e é utilizado para classificação e regressão. Serão testadas variações dos parâmetros de indução binária da árvore, número mínimo de instâncias nas folhas, limite para não divisão de subconjuntos, limite da profundidade máxima da árvore e limite de parada para divisão da maioria dos nós.

O algoritmo “Floresta aleatória (RF - Random Forest)” é um método de aprendizagem por conjunto usado para classificação, regressão e outras tarefas. Serão testadas variações das propriedades, como número de árvores, número de árvores considerado em cada divisão, treinamento replicável e distribuição das classes de equilíbrio. Para controle de crescimento da floresta, serão testadas variações de limite de profundidade de árvores individuais e limite para não divisão de subconjuntos.

O algoritmo de “aumento de gradiente (GB - Gradient Boosting)” é uma técnica de aprendizado de máquina para problemas de regressão e classificação, a partir do aumento do gradiente de árvores de decisão. Serão testadas variações do método de aumento de gradiente (scikit-learn, xgboost e catboost); das propriedades básicas, como número de árvores, taxa de aprendizagem, treinamento replicável e regularização. Para controle de crescimento da floresta, serão testadas variações de limite de profundidade de árvores individuais e limite para não divisão de subconjuntos. Também serão testadas variações das subamostragens, como fração de instâncias de treinamento, fração de recursos para cada árvore, fração de recursos para cada nível e fração de recursos para cada divisão.

O algoritmo “Máquina de vetores de suporte (SVM - Support Vector Machines)” é uma técnica de aprendizado de máquina que separa o espaço de atributos com um hiperplano, maximizando a margem entre as instâncias de diferentes classes. Serão testadas variações de tipo (SVM e v-SVM); do Kernel (linear, polinomial, RBF - função de base radial e sigmoide); e número de limite de interações.

O algoritmo “Rede neural (MLP - *multi-layer perceptron*)” pode aprender a partir de modelos lineares e não lineares. Serão testadas variações dos parâmetros dos modelos, como neurônios por camada oculta, função de ativação para a camada oculta (identidade, logística, tan hiperbólica e ReLu), solver para otimização de peso (L-BFGS-B, SGD e Adam), alfa e máximo de interações. Outros parâmetros foram definidos como os padrões do sklearn.

A performance dos modelos foi avaliada utilizando-se de diferentes métodos de amostragem do *Orange Data Mining*, a saber: validação cruzada (Cross-validation). Para validação dos modelos, foi utilizado o método teste em dados de teste (Test on test data)

utilizando-se de vetores de 25 imagens das 125 obtidas inicialmente e que não foram utilizadas na etapa de ajuste dos modelos.

Tanto para a avaliação do desempenho quanto para a validação, foram utilizadas estatísticas de desempenho. Para modelos de classificação foram utilizadas: a área sob a curva ROC de operação do receptor (Area under ROC), precisão da classificação (Classification accuracy), média harmônica ponderada de precisão e recall (F-1), precisão (Precision), recordação (Recall), especificidade (Specificity), perda de entropia cruzada (LogLoss), tempo de treinamento (Train time) e tempo de teste (Test time).

5 RESULTADOS E DISCUSSÃO

No processamento das imagens de sementes de feijão-caupi, verificou-se que, dentre os *Embedders* testados, ou seja, InceptionV3, SqueezeNet, VGG16 e VGG19, os que obtiveram melhor performance na etapa de modelagem foi o InceptionV3 e VGG16 com o algoritmo Rede Neural Artificial (Neural Network), com acurácia de 1,00, área abaixo da curva ROC de 1, tempo de teste de 6,091 s, tempo de treino de 18,110 s, F1 de 1, precisão de 0,990, recordação de 0,99 e especificidade de 0,999. Com o empate em vários parâmetros e diferenças mínimas em outros, optou-se por escolher o embedder InceptionV3, isso porque os pesquisadores que utilizam essa ferramenta de processamento de imagens optam por trabalhar com esse embedder, também por ser o embedder padrão desse algoritmo dentro do programa (Tabela 1).

Tabela 1 - Estatísticas de desempenho dos Embedder na etapa de modelagem. Sumé, PB, 2022.

| Model | Estatísticas de desempenho | | | | | | | | |
|---------------------|----------------------------|-----------|-------|-------|-------|-----------|--------|-----------|-------------|
| | Embedder InceptionV3 | | | | | | | | |
| | Train time | Test time | AUC | CA | F1 | Precision | Recall | LogLoss | Specificity |
| Neural Network | 18,110 | 6,091 | 1,000 | 1,000 | 0,990 | 0,990 | 0,990 | 0,097 | 0,999 |
| SVM | 19,198 | 3,167 | 1,000 | 0,980 | 1,000 | 0,996 | 0,980 | 0,021 | 1,000 |
| kNN | 2,420 | 2,068 | 1,000 | 0,980 | 0,980 | 0,980 | 0,980 | 0,105 | 0,998 |
| Gradient Boosting | 1777,54 | 2,122 | 1,000 | 0,960 | 0,960 | 0,958 | 0,960 | 0,135 | 0,995 |
| Random Forest | 3,272 | 1,732 | 0,990 | 0,910 | 0,910 | 0,912 | 0,910 | 0,547 | 0,990 |
| Tree | 31,950 | 0,004 | 0,890 | 0,780 | 0,780 | 0,777 | 0,780 | 6253,196 | 0,975 |
| Embedder SqueezeNet | | | | | | | | | |
| SVM | 10,625 | 1,640 | 1,000 | 0,990 | 0,990 | 0,994 | 0,990 | 0,024 | 0,999 |
| Neural Network | 6,768 | 2,201 | 1,000 | 0,990 | 0,990 | 0,988 | 0,990 | 0,106 | 0,999 |
| kNN | 1,379 | 0,986 | 1,000 | 0,990 | 0,990 | 0,987 | 0,990 | 0,080 | 0,999 |
| Gradient Boosting | 849,495 | 0,911 | 1,000 | 0,960 | 0,960 | 0,962 | 0,960 | 0,149 | 0,996 |
| Random Forest | 1,694 | 0,974 | 1,000 | 0,960 | 0,960 | 0,958 | 0,960 | 0,404 | 0,995 |
| Tree | 11,377 | 0,001 | 0,920 | 0,840 | 0,840 | 0,842 | 0,840 | 443548,39 | 0,982 |
| Embedder VGG16 | | | | | | | | | |
| Neural Network | 36,607 | 7,152 | 1,000 | 1,000 | 1,000 | 0,999 | 1,000 | 0,007 | 1,000 |
| SVM | 6,153 | 3,586 | 1,000 | 0,990 | 0,990 | 0,987 | 0,990 | 0,134 | 0,999 |
| Gradient Boosting | 2307,719 | 3,739 | 1,000 | 0,980 | 0,980 | 0,982 | 0,980 | 0,090 | 0,998 |
| Random Forest | 6,824 | 3,212 | 1,000 | 0,980 | 0,980 | 0,978 | 0,980 | 0,217 | 0,998 |
| kNN | 31,898 | 11,988 | 1,000 | 0,970 | 0,970 | 0,974 | 0,970 | 0,133 | 0,997 |

| Tree | 34,001 | 0,003 | 0,950 | 0,910 | 0,910 | 0,908 | 0,910 | 2841,923 | 0,990 |
|-------------------|----------|--------|-------|-------|-------|-------|-------|-----------|-------|
| Embedder VGG19 | | | | | | | | | |
| Neural Network | 39,081 | 6,959 | 1,000 | 0,990 | 0,990 | 0,990 | 0,990 | 0,040 | 0,999 |
| SVM | 2033,605 | 3,729 | 1,000 | 0,980 | 0,980 | 0,983 | 0,980 | 0,078 | 0,998 |
| Gradient Boosting | 6,458 | 3,199 | 1,000 | 0,980 | 0,980 | 0,976 | 0,980 | 0,229 | 0,997 |
| kNN | 5,902 | 3,198 | 1,000 | 0,970 | 0,970 | 0,975 | 0,970 | 0,145 | 0,997 |
| Random Forest | 28,998 | 11,150 | 1,000 | 0,970 | 0,970 | 0,971 | 0,970 | 0,137 | 0,996 |
| Tree | 32,953 | 0,002 | 0,960 | 0,920 | 0,920 | 0,921 | 0,920 | 22430,927 | 0,991 |

IA: inteligência artificial; Train time: tempo de treinamento; Test time: tempo de teste; AUC: área sob a curva ROC de operação do receptor; CA: precisão da classificação; F1: média harmônica ponderada de precisão e recall; Precision: precisão; Recall: recordação; LogLoss: perda de entropia cruzada; Specificity: especificidade. Fonte: Elaboração da equipe de pesquisa.

Verificou-se que a Rede Neural Artificial classificou corretamente em 100% das amostras de sementes das cultivares BRS Guariba, BRS Miranda, BRS Novaera, BRS Pajeú, BRS Pujante, BRS Tapaihum, IPA 206, Pingo de ouro, enquanto que para as cultivares BRS Cauamé e BRS Potengi a taxa de classificação correta foi de 98%, isso pode acontecer por conta da coloração muito semelhante entre as cultivares (Tabela 2).

Tabela 2 - Matriz de confusão para o desempenho do algoritmo Rede Neural Network na etapa de modelagem para processamento das imagens de sementes de feijão-caupi. Sumé, PB, 2022.

| Cultivares | | Valores preditos pelo modelo | | | | | | | | | | Σ |
|------------|---------------|------------------------------|-------------|-------------|-------------|-----------|-------------|-------------|--------------|---------|---------------|----------|
| | | BRS Cauamé | BRS Guariba | BRS Miranda | BRS Novaera | BRS Pajeú | BRS Pontegi | BRS Pujante | BRS Tapaihum | IPA 206 | Pingo de ouro | |
| Valor real | BRS Cauamé | 98 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 100 |
| | BRS Guariba | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Miranda | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Novaera | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pajeú | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pontegi | 0 | 2 | 0 | 0 | 0 | 98 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pujante | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100 |
| | BRS Tapaihum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 100 |
| | IPA 206 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 100 |
| | Pingo de ouro | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| | Σ | 98 | 102 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 1000 |

Fonte: Elaboração da equipe de pesquisa.

Dentre as funções de ativação testadas, dentro do algoritmo Rede Neural Artificial verificou-se que a Logística (Logistic) operando com solver Adam, obteve maior acurácia (Tabela 3).

Tabela 3 - Estatísticas de desempenho para o algoritmo Rede Neural Network com suas variações das funções de ativação operando em diferentes Solvers. Sumé, PB, 2022.

| Função de ativação | Estatísticas de desempenho | | | | | | | | |
|--------------------|----------------------------|-----------|-------|-------|-------|-----------|--------|---------|-------------|
| | Solver L-BFGS-B | | | | | | | | |
| | Train time | Test time | AUC | CA | F1 | Precision | Recall | LogLoss | Specificity |
| Identidade | 29,137 | 6,993 | 1,000 | 0,997 | 0,997 | 0,997 | 0,997 | 0,010 | 1,000 |
| Logística | 27,055 | 6,966 | 1,000 | 0,990 | 0,990 | 0,990 | 0,990 | 0,025 | 0,999 |
| Tan hiperbólica | 25,852 | 6,371 | 1,000 | 0,999 | 0,999 | 0,999 | 0,999 | 0,010 | 1,000 |
| ReLu | 14,930 | 4,310 | 1,000 | 0,996 | 0,996 | 0,996 | 0,996 | 0,008 | 1,000 |
| Solver SGD | | | | | | | | | |
| Identidade | 57,733 | 3,732 | 1,000 | 0,993 | 0,993 | 0,993 | 0,993 | 0,039 | 0,999 |
| Logística | 78,314 | 3,447 | 0,998 | 0,947 | 0,947 | 0,948 | 0,947 | 0,330 | 0,994 |
| Tan hiperbólica | 64,768 | 2,946 | 1,000 | 0,992 | 0,992 | 0,992 | 0,992 | 0,064 | 0,999 |
| ReLu | 59,203 | 2,939 | 1,000 | 0,994 | 0,994 | 0,994 | 0,994 | 0,049 | 0,999 |
| Solver Adam | | | | | | | | | |
| Identidade | 21,687 | 3,365 | 1,000 | 0,994 | 0,994 | 0,994 | 0,994 | 0,019 | 0,999 |
| Logística | 45,371 | 2,877 | 1,000 | 0,999 | 0,999 | 0,999 | 0,999 | 0,015 | 1,000 |
| Tan hiperbólica | 23,044 | 2,843 | 1,000 | 0,996 | 0,996 | 0,996 | 0,996 | 0,016 | 1,000 |
| ReLu | 15,116 | 2,787 | 1,000 | 0,997 | 0,997 | 0,997 | 0,997 | 0,018 | 1,000 |

IA: inteligência artificial; Train time: tempo de treinamento; Test time: tempo de teste; AUC: área sob a curva ROC de operação do receptor; CA: precisão da classificação; F1: média harmônica ponderada de precisão e recall; Precision: precisão; Recall: recordação; LogLoss: perda de entropia cruzada; Specificity: especificidade. Fonte: Elaboração da equipe de pesquisa.

A Rede Neural Network selecionada com a função de ativação Logística e o solver Adam teve taxa de acerto de 100% para todos os genótipos, exceto, BRS Pingo de ouro que a taxa foi de 99% (Tabela 4).

Tabela 4 - Matriz de confusão para o desempenho do algoritmo Rede Neural Network, selecionado com a função de ativação Logística e o solver Adam. Sumé, PB, 2022.

| Cultivares | | Valores preditos pelo modelo | | | | | | | | | Σ | |
|------------|-------------------|------------------------------|-------------|-------------|-------------|-----------|-------------|-------------|--------------|---------|----------|-------------------|
| | | BRS Cauamé | BRS Guariba | BRS Miranda | BRS Novaera | BRS Pajeú | BRS Pontegi | BRS Pujante | BRS Tapaihum | IPA 206 | | BRS Pingo de ouro |
| Valor real | BRS Cauamé | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Guariba | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Miranda | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Novaera | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pajeú | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pontegi | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pujante | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100 |
| | BRS Tapaihum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 100 |
| | IPA 206 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 100 |
| | BRS Pingo de ouro | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 99 | 100 |
| | Σ | 100 | 100 | 101 | 100 | 100 | 100 | 100 | 100 | 100 | 99 | 1000 |

Fonte: Elaboração da equipe de pesquisa.

O segundo melhor algoritmo para processar as imagens das sementes dos genótipos de feijão caupi foi o SVM (Support Vector Machines). E fazendo a comparação após processar as imagens nos tipos de SVM: SVM e v-SVM, o que obteve melhores resultados foi o SVM operando no Kernel Linear (Tabela 5).

Tabela 5 - Estatísticas de desempenho do algoritmo SVM, selecionando o melhor tipo, operando em diferentes Kernel. Sumé, PB, 2022.

| Kernel | Estatísticas de desempenho | | | | | | | | |
|------------|----------------------------|-----------|-------|-------|-------|-----------|--------|---------|-------------|
| | Tipos de Máquina | | | | | | | | |
| | SVM | | | | | | | | |
| | Train time | Test time | AUC | CA | F1 | Precision | Recall | LogLoss | Specificity |
| Linear | 11,286 | 3,914 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 0,068 | 1,000 |
| Polinomial | 13,405 | 4,107 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 0,067 | 1,000 |
| RBF | 19,887 | 6,595 | 1,000 | 0,990 | 0,990 | 0,990 | 0,990 | 0,097 | 0,999 |
| Sigmoide | 13,334 | 4,587 | 0,994 | 0,906 | 0,905 | 0,908 | 0,906 | 0,288 | 0,990 |
| | v-SVM | | | | | | | | |
| Linear | 18,509 | 5,386 | 0,999 | 0,953 | 0,953 | 0,954 | 0,953 | 0,174 | 0,995 |
| Polinomial | 19,071 | 4,895 | 0,999 | 0,963 | 0,962 | 0,965 | 0,963 | 0,141 | 0,996 |
| RBF | 24,186 | 6,572 | 1,000 | 0,971 | 0,971 | 0,971 | 0,971 | 0,144 | 0,997 |
| Sigmoide | 19,205 | 5,061 | 0,994 | 0,914 | 0,914 | 0,914 | 0,914 | 0,294 | 0,990 |

Fonte: Elaboração da equipe de pesquisa.

A Máquina de Suporte de Vetores (SVM) operando com o Kernel Linear e o tipo de máquina SVM teve taxa de acerto de 100% para todos os genótipos (Tabela 6).

Tabela 6 - Matriz de confusão para o desempenho do algoritmo SVM. Sumé, PB, 2022.

| Cultivares | | Valores preditos pelo modelo | | | | | | | | | | |
|-------------------|-------------------|------------------------------|-------------|-------------|-------------|-----------|-------------|-------------|--------------|---------|-------------------|----------|
| | | BRS Cauamé | BRS Guariba | BRS Miranda | BRS Novaera | BRS Pajeú | BRS Pontegi | BRS Pujante | BRS Tapaihum | IPA 206 | BRS Pingo de ouro | Σ |
| Valor real | BRS Cauamé | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Guariba | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Miranda | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Novaera | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pajeú | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pontegi | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 100 |
| | BRS Pujante | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100 |
| | BRS Tapaihum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 100 |
| | IPA 206 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 100 |
| | BRS Pingo de ouro | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| | Σ | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 1000 |

Fonte: Elaboração da equipe de pesquisa.

Os resultados supracitados, evidenciando o bom desempenho dos modelos de inteligência artificial, se devem ao fato de que as sementes contêm características físicas que as distinguem uma das outras, sendo esses aspectos que os algoritmos Rede Neural e SVM conseguem fazer a distinção de cada genótipo. Isso pode ser ratificado pelas informações disponibilizadas pela Embrapa (2021) em seu Catálogo de cultivares de feijão-caupi. No referido catálogo, características como formato, cor, tamanho e peso, por exemplo, são apontadas como descritores importantes para garantia da acurácia dos modelos.

A inteligência artificial para identificação de imagens de sementes tem boa eficiência, Segundo Dönmez (2022), relatou em seu artigo sobre Identificação de Sementes de Milho Haploides, o algoritmo SVM obteve uma eficiência 96,74%%, isso mostra que os sistemas baseados em inteligência artificial na análise de dados em sementes apresentam resultados notáveis, consolidando cada vez mais a inteligência artificial na seleção de sementes.

5.1 VALIDAÇÃO DOS MODELOS

Utilizando-se de vetores de 25 imagens das 125 obtidas inicialmente e que não foram utilizadas na etapa de ajuste dos modelos, verificou-se que houve boa performance para o algoritmo Rede Neural Network, haja vista os elevados valores de acurácia e área abaixo da curva $> 0,7$ (Tabela 7).

Tabela 7 - Dados do algoritmo Rede Neural Artificial no processo de validação. Sumé, PB, 2022.

| Melhor Modelo | Estatísticas de desempenho | | | | | | | | |
|---------------|----------------------------|-----------|------|-------|------|-----------|--------|---------|-------------|
| | Train time | Test time | AUC | CA | F1 | Precision | Recall | LogLoss | Specificity |
| Rede Neural | N/A | N/A | 0,94 | 0,768 | 0,72 | 0,715 | 0,768 | 3,805 | 0,974 |

Fonte: Elaboração da equipe de pesquisa.

A Rede Neural Artificial classificou corretamente com taxa de acerto de 100% as cultivares BRS Miranda, BRS Novaera, BRS Pujante, BRS Tapaihum e BRS Pingo de ouro (Tabela 8).

Tabela 8 - Matriz de confusão para o desempenho do algoritmo Rede Neural Network no processo de validação. Sumé, PB, 2022.

| Cultivares | Valores preditos pelo modelo | | | | | | | | | | Σ |
|-------------------|------------------------------|-------------|-------------|-------------|-----------|-------------|-------------|--------------|---------|-------------------|----------|
| | BRS Cauamé | BRS Guariba | BRS Miranda | BRS Novaera | BRS Pajeú | BRS Pontegi | BRS Pujante | BRS Tapaihum | IPA 206 | BRS Pingo de ouro | |
| BRS Cauamé | 14 | 0 | 0 | 5 | 0 | 4 | 0 | 0 | 2 | 0 | 25 |
| BRS Guariba | 10 | 0 | 0 | 8 | 0 | 6 | 0 | 0 | 1 | 0 | 25 |
| BRS Miranda | 0 | 0 | 23 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 25 |
| BRS Novaera | 0 | 0 | 0 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 25 |
| BRS Pajeú | 0 | 0 | 9 | 0 | 9 | 0 | 1 | 0 | 6 | 0 | 25 |
| BRS Pontegi | 0 | 0 | 0 | 3 | 0 | 22 | 0 | 0 | 0 | 0 | 25 |
| BRS Pujante | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 0 | 0 | 0 | 25 |
| BRS Tapaihum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 0 | 0 | 25 |
| IPA 206 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 24 | 0 | 25 |
| BRS Pingo de ouro | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 25 |
| Σ | 24 | | 32 | 41 | 10 | 32 | 26 | 25 | 35 | 25 | 250 |

Fonte: Elaboração da equipe de pesquisa.

Verificou-se que o algoritmo SVM teve boa performance, haja vista os elevados valores de acurácia e área abaixo da curva $> 0,9$ (Tabela 9).

Tabela 9 - Dados do algoritmo SVM no processo de validação. Sumé, PB, 2022.

| Modelo | Estatísticas de desempenho | | | | | | | | |
|--------|----------------------------|-----------|-------|-------|-------|-----------|--------|---------|-------------|
| | SVM | | | | | | | | |
| | Train time | Test time | AUC | CA | F1 | Precision | Recall | LogLoss | Specificity |
| SVM | N/A | N/A | 0,941 | 0,776 | 0,729 | 0,739 | 0,776 | 3,842 | 0,975 |

Fonte: Elaboração da equipe de pesquisa.

A Máquina de Suporte de Vetores (SVM) operou corretamente com taxa de acerto de 100% nas cultivares BRS Pujante, BRS Tapaihum, IPA 206, BRS Pingo de ouro e BRS Novaera (Tabela 10). Face ao exposto, levando em consideração a pesquisa de Wang e Song (2023), em que eles trabalharam a identificação de milho doce com imagens hiperespectrais combinadas com aprendizado profundo, deve-se destacar que os resultados da pesquisa mostraram que a identificação artificial de sementes de milho doce obteve precisão de 95% nos conjuntos de treinamento e teste. Portanto, com os dados da validação, em que são utilizadas novas fotos dos genótipos no processamento, os algoritmos conseguiram distinguir boa parte das cultivares, logo pode-se afirmar que a inteligência artificial conseguiu diferenciar de acordo com traços marcantes das sementes, foi como também afirmaram Miqueloni *et al.* (2018) e Silva *et al.* (2020).

Tabela 10 - Matriz de confusão para o desempenho do algoritmo SVM, no processo de validação. Sumé, PB, 2022.

| Cultivares | Valores preditos pelo modelo | | | | | | | | | | Σ |
|-------------------|------------------------------|-------------|-------------|-------------|-----------|-------------|-------------|--------------|---------|-------------------|----------|
| | BRS Cauamé | BRS Guariba | BRS Miranda | BRS Novaera | BRS Pajeú | BRS Pontegi | BRS Pujante | BRS Tapaihum | IPA 206 | BRS Pingo de ouro | |
| BRS Cauamé | 17 | 0 | 0 | 2 | 0 | 4 | 0 | 0 | 2 | 0 | 25 |
| BRS Guariba | 5 | 0 | 0 | 6 | 0 | 14 | 0 | 0 | 0 | 0 | 25 |
| BRS Miranda | 0 | 0 | 22 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 25 |
| BRS Novaera | 0 | 0 | 0 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 25 |
| BRS Pajeú | 0 | 0 | 8 | 0 | 8 | 0 | 0 | 0 | 9 | 0 | 25 |
| BRS Pontegi | 0 | 0 | 0 | 3 | 0 | 22 | 0 | 0 | 0 | 0 | 25 |
| BRS Pujante | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 0 | 0 | 0 | 25 |
| BRS Tapaihum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 0 | 0 | 25 |
| IPA 206 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 0 | 25 |
| BRS Pingo de ouro | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 25 |
| Σ | 22 | | 30 | 36 | 8 | 40 | 25 | 25 | 39 | 25 | 250 |

Fonte: Elaboração da equipe de pesquisa.

De modo geral a inteligência artificial tem grande potencial na identificação de plantas, pois, segundo estudos realizados por Menezes Neto *et al.* (2021), a inteligência artificial conseguiu diferir as diferentes plantas alimentícias não convencionais em hortas do Laboratório Horta Comunitárias Nutrir (LabNutri), vinculado ao Departamento de Nutrição do Centro de Ciências da Saúde (CCS) da UFRN. O que leva a afirmar que a identificação das sementes de feijão-caupi através da inteligência artificial, tem bases relacionadas que atestam a veracidade dos resultados. Deve-se ressaltar que os achados desta pesquisa terão importante contribuição para o avanço técnico-científico da cadeia produtivas da cultura de feijão-caupi, sobretudo na microrregião do Cariri Ocidental paraibano.

6 CONCLUSÃO

Os algoritmos de aprendizado de máquina Rede Neural Network e Máquina de Suporte de Vetores (SVM) tiveram melhores indicadores de performance para identificação de feijão-caupi respectivamente, a partir do processamento das imagens digitais de sementes.

REFERÊNCIA

BARBOSA, E. P. A.; SODRÉ, D. N.; BRAUN, H.; VIEIRA, R. F. Seeds enriched with molybdenum improve cowpea yield in sub-humid tropical regions of Brazil. **Agronomy Journal**, v. 113, p. 2044-2052, 2021.

BIOVERSITY INTERNATIONAL. Descritores para Feijão frade ou caupi (*Vigna unguiculata* (L.) Walp.). IN: BETTENCOURT, E. (Ed.). **Bioversity International**. Roma-Itália, 2007.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Instrução Normativa nº 12 de 28 mar. 2008**. Diário Oficial da República Federativa do Brasil, Brasília, DF, 31 mar. 2008. p. 11-14.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Secretaria de Desenvolvimento Agropecuário e Cooperativismo. Serviço Nacional de Proteção de Cultivares**. Ato nº 4, de 19 de agosto de 2010. Diário Oficial da República Federativa do Brasil, Brasília, 2010. Seção 1, p. 6-7.

DEMŠAR, J.; CURK, T.; ERJAVEC, A.; GORUP, C.; HOCEVAR, T.; MILUTINOVIC, M.; MOZINA, M.; POLAJNAR, M.; TOPLAK, M.; STARIC, A.; STAJDOHAR, M.; UMEK, L.; ZAGAR, L.; ZBONTAR, J.; ZITNIK, M.; ZUPAN, B. Laranja: Caixa de ferramentas de mineração de dados em Python, **Journal of Machine Learning Research**, v. 14, p. 2349-2353, 2013.

DÖNMEZ, E. Enhancing classification capacity of CNN models with deep feature selection and fusion: A case study on maize seed classification. **Data & Knowledge Engineering**, v. 141, p. e102075, 2022.

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. **Catálogo de cultivares de feijão-caupi**. Embrapa Meio-Norte, 2021. 23p.

FERREIRA, D. P.; SOUSA, D. P.; NUNES, H. G. G. C.; PINTO, J. V. N.; FARIAS, V. D. S.; COSTA, D. L. P.; MOURA, V. B.; TEIXEIRA, E.; SOUSA, A. M. L.; PINHEIRO, H. A.; SOUZA, P. J. O. P. Cowpea Ecophysiological Responses to Accumulated Water Deficiency

during the Reproductive Phase in Northeastern Pará, Brazil. **Horticulturae**, v. 7, p. 116, 2021.

FONSECA, J. R.; SARTORATO, A.; RAVA, C. A.; COSTA, J. E. C.; FREIRE, M. S.; ANTUNES, I. F.; TEIXEIRA, M. G.; SILVA, J. G. **Características botânicas, agronômicas e fenológicas de cultivares regionais de feijão coletadas na região do Recôncavo Baiano**. (Boletim de Pesquisa, 4). Goiânia: EMBRAPA-CNPAP, 1986. p. 27.

FREIRE FILHO, F. R.; CARDOSO, M. J.; ARAÚJO, A. G.; SANTOS, A. A.; SILVA, P. H. S. **Características botânicas e agronômicas de cultivares de feijão macassar (*Vigna unguiculata* (L.) Walp)**. (Boletim de Pesquisa, 4). Teresina: EMBRAPA-UEPAE, 1981. p. 45.

FREIRE FILHO, F. R.; ROCHA, M. M.; RIBEIRO, V. Q.; LOPES, A. C. A. Adaptabilidade e estabilidade produtiva de feijão-caupi. **Ciência Rural**, v. 35, p. 24-30, 2005.

GODEC, P.; PANČUR, M.; ILENIČ, N.; ČOPAR, A.; STRAŽAR, M.; ERJAVEC, A.; PRETNAR, A.; DEMŠAR, J.; STARIČ, A.; TOPLAK, M.; ŽAGAR, L.; HARTMAN, J.; WANG, H.; BELLAZZI, R.; PETROVIČ, U.; GARAGNA, S.; ZUCCOTTI, M.; PARK, D.; SHAULSKY, G.; ZUPAN, B. Democratized image analytics by visual programming through integration of deep models and small-scale machine learning. **Nat Commun**, v. 10, p. 4551, 2019.

GOMES, A. M. F.; DRAPER, D.; NHANTUMBO, N.; MASSINGA, R.; RAMALHO, J. C.; MARQUES, I.; RIBEIRO-BARROS, A. I. Diversity of Cowpea [*Vigna unguiculata* (L.) Walp] Landraces in Mozambique: New Opportunities for Crop Improvement and Future Breeding Programs. **Agronomy**, v. 11, p. 991, 2021.

HATI, A. J.; SINGH, R. R. Artificial Intelligence in Smart Farms: Plant Phenotyping for Species Recognition and Health Condition Identification Using Deep Learning. **AI**, v. 2, p. 274-289, 2021.

KLUNNIKOVA, Y. V.; ANIKEEV, M. V.; FILIMONOV, A. V.; KUMAR, R. Machine learning application for prediction of sapphire crystals defects. **Journal of Electronic Science and Technology**, v. 18, p. e100029, 2020.

MARINHO, J. T. S.; LESSA, L. S.; COSTA, C. R. Agronomic performance of cowpea genotypes in southwestern Brazilian Amazon. **Pesquisa Agropecuária Brasileira**, v. 56, p. e02046, 2021.

MARTEY, E.; ETWIRE, P. M.; ADOGOBA, D. S.; TENGEY, T. K. Farmers' preferences for climatesmart cowpea varieties: implications for crop breeding programmes. **Climate and Development**, 2021.

MENEZES NETO, E. J.; LIMA, D. G.; FEITOSA, I. S.; GOMES, S. M.; JACOB, M. C. M. Plant Identification Using Artificial Intelligence: Innovative Strategies for Teaching Food Biodiversity. IN: JACOB, M. C. M.; ALBUQUERQUE, U. P. (Eds.). **Local Food Plants of Brazil**. Springer, 2021.

MIQUELONI, D. P.; SANTOS, V. B.; LIMA, S. R.; MESQUITA, D. N.; FURTADO, S. S. F. Descrição e discriminação de variedades crioulas de feijão-caupi na Amazônia Ocidental brasileira. **Acta Iguazu**, v. 7, p. 49-61, 2018.

NARAYANA, M.; ANGAMUTHU, M. Cowpea. IN: PRATAP, A.; GUPTA, S. (Eds.). **The Beans and the Peas: From Orphan to Mainstream Crops**. Woodhead Publishing, 2021.

OLIVEIRA, E.; MATTAR, E. P. L.; ARAÚJO, M. L.; JESUS, J. C. S.; NAGY, A. C. G.; SANTOS, V. B. Descrição de cultivares locais de feijão-caupi coletados na microrregião Cruzeiro do Sul, Acre, Brasil. **Acta Amazonica**, v. 45, p. 243-254, 2015.

SÁ, F. V. S.; SILVA, I. E.; FERREIRA NETO, M.; LIMA, Y. B.; PAIVA, E. P.; GHEYI, H. R. Phosphorus doses alter the ionic homeostasis of cowpea irrigated with saline water. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 25, p. 372-379, 2021.

SÁNCHEZ-NAVARRO, V.; ZORNOZA, R.; FAZ, A.; FERNÁNDEZ, J. A. Cowpea Crop Response to Mineral and Organic Fertilization in SE Spain. **Processes**, v. 9, p. 822, 2021.

SANTOS, A. S.; SILVA, E. D.; TAVARES, E. D.; CURADO, F. F.; SALES, J. F. S.; PEREIRA, L. **Desempenho de Variedades Crioulas e Comerciais de Feijão-Macassar ou Feijão Caupi no Agreste Paraibano**. (Comunicado Técnico 186), 2016.

SILVA, N. V.; LINHARES, P. C. F.; SOUSA, R. P.; ASSIS, J. P.; CARDOSO, E. A.; PEREIRA, J. O.; MARTINS, W. R.; ALVES, L. S.; SANTOS, M. F. A.; NOGUEIRA, G. S.

L. R.; FARIAS, J. R. S. Biometry of Seeds of Caupi Beans Cream Varieties Cultivated in Northeast Brazil. **Journal of Agricultural Science**, v. 12, p. 239-244, 2020.

SILVA, V. M.; NARDELI, A. J.; MENDES, N. A. C.; ROCHA, M. M.; WILSON, L.; YOUNG, S. D.; BROADLEY, M. R.; WHITE, P. J.; REIS, A. R. Agronomic biofortification of cowpea with zinc: Variation in primary metabolism responses and grain nutritional quality among 29 diverse genotypes. **Plant Physiology and Biochemistry**, v. 162, p. 378-387, 2021.

SOLTABAYEVA, A.; ONGALTAY, A.; OMONDI, J. O.; SRIVASTAVA, S. Morphological, Physiological and Molecular Markers for Salt-Stressed Plants. **Plants**, v. 10, p. 243, 2021.

SUJATHA, R.; CHATTERJEE, J. M.; JHANJHI, N.; BROHI, S. N. Performance of deep learning vs machine learning in plant leaf disease detection. **Microprocessors and Microsystems**, v. 80, p. e103615, 2021.

SUN, Q.; XIA, X.; CHAI, X. Machine Vision Based Phenotype Recognition of Plant and Animal. IN: Chinese Academy of Sciences (Eds). **China's e-Science: Blue Book 2020**. Springer, Singapore. 2021.

TAVARES, D. S.; FERNANDES, T. E. K.; RITA, Y. L.; ROCHA, D. C.; SANT'ANNA-SANTOS, B. F.; GOMES, M. P. Germinative metabolism and seedling growth of cowpea (*Vigna unguiculata* (L.) Walp) under salt and osmotic stress. **South African Journal of Botany**, v. 139, p. 399-408, 2021.

TOYINBO, J. O.; FATOKUN, C.; BOUKAR, O.; FAKOREDE, M. A. B. Genetic variability and trait association under thrips (*Megalurothrips sjostedti* Trybom) infestation in cowpea (*Vigna unguiculata* [L.] Walp.). **Euphytica**, v. 217, 2021.

VITAL, A. F. M.; SOUZA, T.; SILVA, L. J. R.; SANTOS, R. V.; SILVA, S. I. A.; NASCIMENTO, G. S.; SANTOS, D. Biomass production and macronutrient content in *Pennisetum glaucum* (L.) R. Brown as affected by organic fertilization and irrigation. **Revista Brasileira de Ciências Agrárias**, v. 15, p. e8576, 2020.

XIONG, J.; YU, D.; LIU, S.; SHU, L.; WANG, X.; LIU, Z. A Review of Plant Phenotypic Image Recognition Technology Based on Deep Learning. **Electronics**, v. 10, p. 81, 2021.

WANG, Y.; SONG, S. Variety identification of sweet maize seeds based on hyperspectral imaging combined with deep learning. **Infrared Physics & Technology**, v. 130, p. 104611, maio 2023.