



XII Congresso Brasileiro de Agroinformática

11 a 14 de Novembro de 2019

ANAIS

INDAIATUBA
Novembro de 2019

ANAIS DO XII CONGRESSO BRASILEIRO DE AGROINFORMATICA

ORGANIZADORES

MARIA FERNANDA MOURA

JAYME GARCIA ARNAL BARBEDO

ALAINÉ MARGARETE GUIMARÃES

VALTER CASTELHANO DE OLIVEIRA

FATEC/EMBRAPA/UEPG-PPGCA

11 A 14 DE NOVEMBRO DE 2019

Ficha Catalográfica

Congresso Brasileiro de Agroinformática (11.: 2019: Indaiatuba, SP)

C749 Anais eletrônicos do XII Congresso Brasileiro de Agroinformática. 11 a 14 de novembro de 2019. Indaiatuba, SP; Universidade Estadual de Ponta Grossa/ Maria Fernanda Moura et al. (Org.). Ponta Grossa: SBIAGRO, 2019.

ISBN: 978-65-00-10242-0

1. Agroinformática. 2. Agricultura digital. 3. Internet das coisas. I. Moura, Maria Fernanda (Org.). II. Barbedo, Jayme Garcia Arnal (Org.). III. Guimarães, Alaine Margarete (Org.). IV. Oliveira, Valter Castelhana de (Org.). V. SBIAgro. VI. T.

CDD: 004.22





SVM com DenseNet para classificação de mudas do guaraná a partir da nervura central do folíolo

Brenda Talyne Costa Martins¹, Marcos Filipe Alves Salame²

¹Bolsista de Iniciação Científica FAPEAM, Embrapa Amazônia Ocidental
Manaus, Amazonas, Brasil, brendatalyne@gmail.com

²Analista de Tecnologia da Informação, Embrapa Amazônia Ocidental
Manaus, Amazonas, Brasil, marcos.salame@embrapa.br

RESUMO

O guaranazeiro (*Paullinia cupana var. sorbilis*) é um importante e tradicional cultivo do Estado do Amazonas. Com o aumento da demanda por guaraná, a Embrapa Amazônia Ocidental desenvolveu e registrou 20 cultivares e lançou 18 até o momento, com o objetivo de aumentar a produtividade, reduzir o tempo de formação, fornecer alto teor de cafeína e mais resistência às principais pragas. Entretanto, o procedimento de identificação desses cultivares é difícil e feito somente por especialistas. Diante deste cenário, foi criado um *dataset* de mudas de guaraná, composto por 10 cultivares, através de um protótipo desenvolvido para dispositivos móveis, de forma a evitar a necessidade de qualquer pré-processamento manual e visando a proximidade com o ambiente real do produtor. Foram realizados experimentos com 5 dos cultivares coletados, utilizando aprendizado de máquina com SVM e DenseNet para classificação a partir de imagens da nervura central dos folíolos e obteve-se como resultado uma acurácia de 88%.

PALAVRAS-CHAVE: Aprendizado de Máquina, guaranazeiro, cultivares.

ABSTRACT

The guarana plant (*Paullinia cupana var. Sorbilis*) is an important and traditional crop of state of Amazonas. Due to increase of demand for guarana, Embrapa Amazônia Ocidental developed and registered 20 cultivars and launched 18 so far, aiming to increase productivity, reduce formation time, provide high caffeine content and more resistance to major pests. However, the identification procedure of these cultivars is difficult and it is made only by specialists. Based on this scenario, we created a dataset of guarana seedlings, composed by 10 cultivars, through a prototype developed for mobile devices, in order to avoid the need for any manual preprocessing and aiming at proximity to the environment of the rural producer. Experiments

were carried out on 5 of the cultivars collected, using machine learning with SVM and Dense-Net for classification from images of the central vein of the leaflets and we obtained, as a result, an accuracy of 88%.

KEYWORDS: Machine Learning, guarana, cultivars.

INTRODUÇÃO

O guaranazeiro (*Paullinia cupana var. sorbilis*) dentre muitas espécies amazônicas, é reconhecido pelo seu potencial econômico (GONÇALVES et al., 2006). Com o aumento da demanda e importância, a Embrapa Amazônia Ocidental desenvolveu e registrou 20 cultivares ao longo dos anos através do programa de melhoramento genético, que visa selecionar e recomendar cultivares com grande potencial de produtividade, alto teor de cafeína e com mais resistência às principais pragas. As informações dos cultivares podem ser encontradas em (MAPA, 2019).

Desses cultivares, 19 foram obtidos pelo método de estaquia (reprodução assexuada). Uma das grandes diferenças desses cultivares para os que são produzidos por sementes (reprodução sexuada) é que segundo (TRICAUD; PINTON; PEREIRA, 2016), o tempo de formação das mudas é de 7 meses, enquanto que a da muda tradicional por semente, demora 12 meses até ficar pronta para ir ao campo.

De acordo com (CORRÊA, 1982), é recomendável que as mudas, antes do plantio, possuam no mínimo de quatro a seis folhas completas, com folíolos bem desenvolvidos, maduros e de coloração normal. Até a formação das primeiras folhas, as mudas no viveiro recebem tratamento através da irrigação intermitente, ou seja, o fornecimento controlado de água para as plantas em quantidade suficiente e no momento que precisam de irrigação, assegurando a produtividade e a sobrevivência da plantação. Segundo (PEREIRA, 2005), após esse período, a planta é levada para um segundo viveiro que lhe proporcionará melhores resultados na propagação vegetativa com a nebulização, que lhe proporciona umidade necessária para o bom desempenho do enraizamento das estacas, promovendo uma proteção eficiente da superfície até a formação das mudas para que possam ser levadas para o seu local definitivo de plantio.

No entanto, a distinção desses cultivares, principalmente em mudas, que é quando ocorre o período de transporte dessas plantas para o local definitivo, ainda é feito por mecanismos manuais (PEREIRA, 2005). Isso resulta em um processo árduo e muitas vezes impreciso, podendo resultar numa aquisição incorreta por parte do agricultor. Esse cenário estimula a pesquisa em novas tecnologias para identificação dessas variedades genéticas, uma vez que as folhas apresentam muita semelhança entre si e geram resultados bem distintos.

Nos experimentos realizados por (SOUSA et al., 2017), foram utilizadas imagens de 2 cultivares que foram obtidas de 2 ângulos diferentes, sendo 20 imagens por cultivar de folhas inteiras, uma de cada amostra. O *dataset* utilizado pode ser considerado pequeno quando comparado a outros trabalhos relacionados de classificação por imagens, portanto, eles expandiram através de *data augmentation*, gerando também mais variações. Com a utilização de Redes

Neurais Convolucionais foi possível obter o resultado de 97% de acurácia neste experimento.

Entretanto, houve um pré-processamento realizado manualmente com redimensionamento, retirada de fundo, e centralização da folha na imagem, inviabilizando de poder ser usado de forma mais prática.

Em (SOUSA; SALAME, 2019) é possível identificar resultados para análises com 3 cultivares considerados de alta produtividade. Os autores utilizaram 434 imagens de plantas adultas e aplicaram técnicas de aprendizado de máquina supervisionado para a classificação realizada pelo formato dos folíolos e pelos padrões de venação, que são mais visíveis na superfície abaxial das folhas. A precisão de reconhecimento para este experimento foi de 89%. Ou seja, reduziu a acurácia, mas o processo foi facilitado, devido a ser apenas o folíolo, reduzindo os trabalhos de pré-processamento manual.

Dessa forma, este trabalho propõe uma abordagem computacional para identificação dos cultivares de guaraná a partir de imagens dos folíolos das mudas de guaranazeiros utilizando aprendizado de máquina. As imagens foram coletadas de forma prática, o mais próximo do ambiente real de produção, sem necessidade de qualquer tipo de pré-processamento manual e utilizando apenas a nervura central dos folíolos, com o intuito de que a solução, futuramente, possa ser usada pelos produtores rurais através do celular.

MATERIAL E MÉTODOS

Pelo fato de não haver um *dataset* público dos cultivares de guaranazeiro criados pela Embrapa e também pelo intuito de transformar a solução em algo que possa ser usado por produtores rurais no dia-a-dia, foi desenvolvido um protótipo em Ionic para dispositivos móveis com sistema operacional Android, com o propósito de coletar as imagens dos folíolos das mudas para compôr um *dataset*, de forma a evitar qualquer tipo de pré-processamento manual. Para a criação do mesmo, foram utilizados os *frameworks* Ionic e nodeJs com as linguagens *javaScript* e *typescript*

O protótipo também pode ser usado para validar a classificação em ambiente de produção, podendo vir, futuramente, a se tornar um produto. Ele contém a funcionalidade de câmera com filtro e corte e também possui informações referentes aos cultivares, obtidas a partir de publicações científicas e entrevistas com especialistas da área.

Criação do dataset

A coleta teve enfoque na nervura central dos folíolos e foi utilizado um smartphone com uma câmera de 16 *megapixels* e Android 7.0. Ao todo foram coletados 10 cultivares, onde em cada cultivar foram capturados 2 folíolos de cada folha, sendo 10 fotos por folíolo e 10 folhas, totalizando 200 imagens por cultivar. A captura das imagens foi feita dos folíolos 2 e 3 de cada folha da parte abaxial (parte inferior), conforme Figura 1. Foram fotografadas várias áreas, com o cuidado para que a captura não excedesse o perímetro dos folíolos. Na coleta foi observada a

distância entre o *smartphone* e o folíolo, além de variações de tamanhos dos folíolos conforme Figura 2.

Os cultivares coletados foram desenvolvidos pelo método de estaquia (reprodução assexuada), sendo eles: BRS-Luzéia, BRS-Maués, BRS CG612, BRS Andirá, BRS Cereçaporanga, BRS-Amazonas, BRS CG 372, BRS CG608, BRS CG610 e BRS CG850.

Figura 1: Nervura central dos folíolos capturadas com uso do protótipo para criação do dataset



Figura 2: Tamanhos dos folíolos e aproximação do *smartphone* na captura das imagens.

Tam. (Altura)	Aprox. da Câmera	Período
10 à 12 cm	06 à 08 cm	Jan à Fev
13 à 16 cm	08 à 09 cm	Jan à Fev
17 à 20 cm	08 à 10 cm	Maio
21 à 22 cm	08 à 11 cm	Maio

Os cultivares BRS Mundurucânia, BRS CG611, BRS Saterê e BRS Onhiamuaçabê, não foram coletados pois não haviam mudas no viveiro. Já os cultivares BRS Marabitaná, BRS CG505, BRS CG648, BRS CG189, BRS CG882, ainda serão coletados para serem acrescentados ao *dataset*. Segundo (KARASAWA, 2009) há diferenças entre os sistemas de reprodução de plantas. O BRS Noçoquém (ATROCH; FILHO; PEREIRA, 2015) foi o primeiro cultivar lançado de reprodução via semente. Este cultivar não foi capturado por não fazer parte dos experimentos de reprodução pelo método de estaquia, como é o caso dos outros 19 cultivares registrados.

Cultivares selecionados

Para os experimentos deste trabalho, devido ao tempo de processamento e à importância, foram escolhidos 5 cultivares que apresentam alta produtividade: BRS-Luzéia, BRS-Maués, BRS-

Andirá, BRS-Cereçaporanga e BRS-CG612, com idades variando entre 8 meses e 2 anos, mas ainda consideradas mudas, plantadas dentro de sacos.

De acordo com (JAIN; DUIN; MAO, 2000), o reconhecimento de padrões é o estudo de como máquinas podem observar o ambiente, aprender a distinguir padrões e tomar decisões sobre as categorias dos padrões. Sendo assim, podemos concluir que a importância do reconhecimento desses padrões está ligada a encontrar métodos que automatizem atividades, que possam ser trabalhosas e também sujeitas a falhas.

Técnicas de classificação

Foi utilizada uma CNN (*Convolutional neural network*) como extrator de características. Conforme (ABDEL-HAMID et al., 2012), a CNN consiste em um ou mais pares de camadas de convolução. As camadas convolucionais aplicam filtros que processam pequenos locais de uma imagem, e são replicados por toda a imagem.

Para a implementação dos algoritmos, foi utilizada a linguagem de programação Python e bibliotecas como NumPy, *OpenCv*, *Scikit-learn*. Para o aprendizado de máquina, foram utilizadas diversas técnicas, desde técnicas clássicas até *deep learning*. O melhor resultado obtido foi com SVM (*Support Vector Machine*) e DenseNet (Dense Convolutional Network). O melhor conjunto de parâmetros encontrado no conjunto de desenvolvimento nesta classificação foi: C:1, gamma: 0.1 e Kernel: linear.

RESULTADOS E DISCUSSÃO

As técnicas abordadas no experimento apresentaram resultados satisfatórios na classificação dos cultivares de guaranazeiros. O resultado de algumas métricas utilizadas pode ser visto na Figura 3. Através de SVM com DenseNet para classificação a partir de imagens da nervura central dos folíolos do guaraná, obteve-se como resultado uma acurácia de 88%.

Figura 3: Resultado da classificação

Cultivares	Precision	Recall	F1-score
BRS Andirá	0.85	1.00	0.92
BRS Luzéia	0.83	0.96	0.89
BRS Cereçaporanga	0.84	0.76	0.79
BRS Maués	0.95	0.78	0.86
BRS CG612	0.96	0.92	0.94

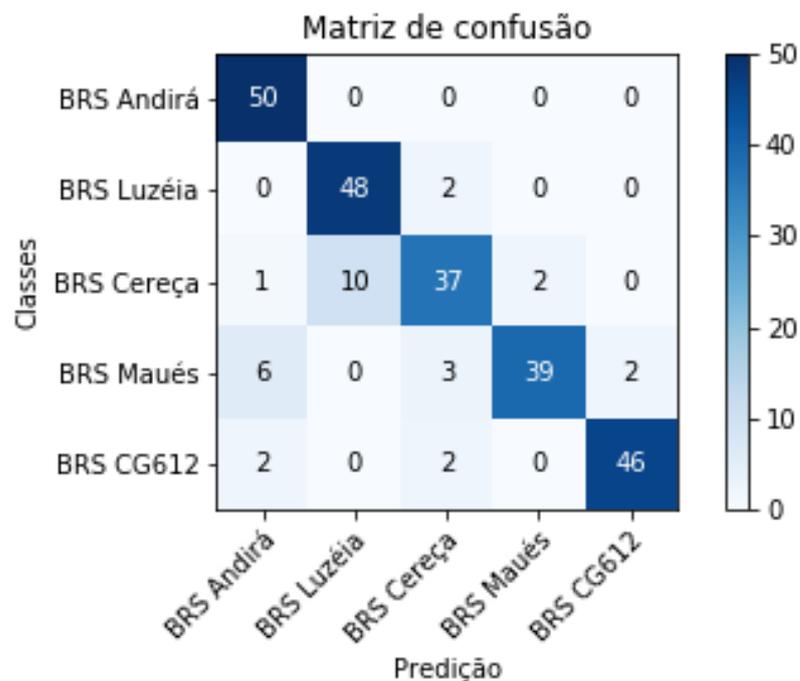
Segundo (FACELI et al., 2011), com princípios apoiados na teoria de aprendizado estatístico, o SVM destaca-se por apresentar uma boa capacidade de generalização. Uma outra ca-

racterística atrativa é a convexidade do problema da otimização formulado por seu treinamento, que implica a existência de um único mínimo global.

Conforme (JÉGOU et al., 2017), a DenseNet tem mostrado excelentes resultados em tarefas de classificação de imagens. A ideia da DenseNet baseia-se na observação de que, se cada camada estiver diretamente conectada para todas as outras camadas em um feed-forward, a rede será mais precisa e mais fácil de treinar. A DenseNet requer menos parâmetros do que uma CNN tradicional equivalente, já que não há necessidade de aprender mapas de recursos.

Na Figura 4, é apresentada a matriz de confusão do experimento. Nela, pode ser observado que o BRS Andirá apresentou 100% de verdadeiro positivo, sendo o único com tal feito. No entanto, falsos positivos ocorreram. Dentre os outros cultivares, podemos observar também que o BRS Luzéia obteve um bom resultado e o BRS Cereçaporanga foi o que obteve menos acertos.

Figura 4: Matriz de confusão da arquitetura de classificação



CONCLUSÕES

Com os resultados do experimento, é possível verificar que há viabilidade na implementação de uma solução para reconhecer cultivares de guaraná a partir na nervura central dos folíolos, que pode ser considerado o dna da planta. As técnicas utilizadas, SVM com DenseNet, apresentaram um resultado promissor na classificação.

Ao todo, foram coletadas 2 mil imagens. Todavia, no experimento, foram utilizadas mil imagens de 5 cultivares e obteve-se como resultado uma acurácia de 88%. Nenhum pré-

processamento manual foi necessário para o *dataset*, tornando real a possibilidade da solução vir a se tornar um produto.

Como trabalhos futuros, pretende-se coletar imagens dos cultivares restantes e realizar vários experimentos com todos os 19, desenvolvidos por método de estaquia, até conseguir um resultado satisfatório. Para dessa forma, posteriormente, ajustar o algoritmo no servidor, realizar validação com o protótipo em ambiente de produção e, se tudo der certo, disponibilizar para o público, podendo oferecer relevante complemento aos métodos atuais utilizados na identificação de cultivares.

AGRADECIMENTOS

Aos pesquisadores em guaranaicultura Firmino José do Nascimento Filho e André Luiz Atroch, cuja disponibilidade e conhecimentos sobre o guaranzeiro transmitidos foram fundamentais; aos especialistas Orbélio Mota Campos Silva e Luciano Simões Malcher pela ajuda no viveiro; aos colegas Alex de Lima Sousa e Richardson Allan Ferreira de Souza pelo auxílio nos estudos e à Fundação de Amparo à Pesquisa do Estado do Amazonas pela disponibilização da bolsa de iniciação científica.

REFERÊNCIAS

- ABDEL-HAMID, O. et al. Applying convolutional neural networks concepts to hybrid nn-hmm model for speech recognition. In: IEEE. *2012 IEEE international conference on Acoustics, speech and signal processing (ICASSP)*. [S.l.], 2012. p. 4277–4280.
- ATROCH, A.; FILHO, F. do N.; PEREIRA, J. Brs noçoquém: primeira cultivar de guaranzeiro de reprodução via sementes para cultivo no estado do amazonas. *Embrapa Amazônia Ocidental-Comunicado Técnico*, Manaus: Embrapa Amazônia Ocidental, 2015., 2015.
- CORRÊA, M. *Formação de mudas de guaraná*. [S.l.]: EMBRAPA-UEPAE de Manaus, 1982.
- FACELI, K. et al. Inteligência artificial: Uma abordagem de aprendizado de máquina. 2011.
- GONÇALVES, J. F. de C. et al. Características fisiológicas e anatômicas de folhas de dois clones de guaraná. *Pesquisa Agropecuária Brasileira*, v. 41, n. 3, p. 393–398, 2006.
- JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, Ieee, v. 22, n. 1, p. 4–37, 2000.
- JÉGOU, S. et al. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. [S.l.: s.n.], 2017. p. 11–19.
- KARASAWA, M. Diversidade reprodutiva de plantas, uma perspectiva evolutiva e bases genéticas. *Editora SBG, Sociedade Brasileira de Genética, Ribeirão Preto, SP: p. 113*, 2009.

MAPA. *Registro Nacional de Cultivares - RNC*. 2019. Disponível em:
<<http://www.agricultura.gov.br/guia-de-servicos/registro-nacional-de-cultivares-rnc>>.
Acesso em: 1 jul. 2019.

PEREIRA, J. C. Cultura do guaranazeiro no amazonas. *Embrapa Amazônia Ocidental-Sistema de Produção*, Manaus: Embrapa Amazônia Ocidental, 2005., 2005.

SOUSA, A. d. L. et al. Redes neurais convolucionais aplicadas ao processo de classificação de cultivares de guaranazeiros. In: IN: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL E COMPUTACIONAL, 14. [S.l.], 2017.

SOUSA, A. de L.; SALAME, M. F. A. Classificação de guaranazeiros através do padrão de venação e formato de folíolos utilizando aprendizagem de máquina. *Anais do Computer on the Beach*, p. 676–685, 2019.

TRICAUD, S.; PINTON, F.; PEREIRA, H. dos S. Saberes e práticas locais dos produtores de guaraná (*paullinia cupana kunth var. sorbilis*) do médio amazonas: duas organizações locais frente à inovação. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, Museu Paraense Emílio Goeldi, v. 11, n. 1, p. 33–53, 2016.