UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Botucatu

FCA
FACULDADE DE CIÊNCIAS AGRONÔMICAS
UNESP | BOTUCATU

MATEUS DE CAMPOS LEME

INTELIGÊNCIA ARTIFICIAL UTILIZADA NA IDENTIFICAÇÃO DE ESPÉCIES E PROGNÓSTICO VISUAL DE MUDAS FLORESTAIS

Botucatu

2021

MATEUS DE CAMPOS LEME

INTELIGÊNCIA ARTIFICIAL UTILIZADA NA IDENTIFICAÇÃO DE ESPÉCIES E PROGNÓSTICO VISUAL DE MUDAS FLORESTAIS

Dissertação apresentada à Faculdade de Ciências Agronômicas da Unesp Câmpus de Botucatu, para a obtenção do título de Mestre no Programa de Pós-Graduação em Agronomia – Energia na Agricultura.

Orientador: Prof. Dr. Sérgio Campos
Coorientador: Prof. Dr. Osvaldo César Pinheiro

Botucatu

2021

L551i

Leme, Mateus de Campos

Inteligência artificial utilizada na identificação de espécies e prognóstico visual de mudas florestais / Mateus de Campos Leme. -- Botucatu, 2021

48 p. : tabs., fotos

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agronômicas, Botucatu

Orientador: Sérgio Campos

Coorientador: Osvaldo César Pinheiro de Almeida

1. inteligência artificial. 2. redes neurais convolucionais. 3. viveiros. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agronômicas, Botucatu. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Botucatu

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: INTELIGÊNCIA ARTIFICIAL UTILIZADA NA IDENTIFICAÇÃO DE ESPÉCIES E PROGNÓSTICO DE MUDAS FLORESTAIS

AUTOR: MATEUS DE CAMPOS LEME
ORIENTADOR: SÉRGIO CAMPOS
COORIENTADOR: OSVALDO CESAR PINHEIRO DE ALMEIDA

Aprovado como parte das exigências para obtenção do Título de Mestre em AGRONOMIA (ENERGIA NA AGRICULTURA), pela Comissão Examinadora:

Prof. Dr. SÉRGIO CAMPOS (Participaçao Virtual)
Engenharia Rural e Socioeconomia / Faculdade de Ciências Agronômicas de Botucatu - UNESP

Prof. Dr. LUCIANO NARDINI GOMES (Participaçao Virtual)
Geociências / Universidade Estadual de Londrina

Prof. Dr. JOÃO RICARDO FAVAN (Participaçao Virtual)
Big Data no Agronegócio / Faculdade de Tecnologia de Pompéia

Botucatu, 03 de março de 2021

Aos meus amados pais, irmãos e avós, Ademar e Silvia, Tiago e Carol e Izelda, Wilson, Leontina e Deoclécio dedico.

AGRADECIMENTOS

A palavra agradecimento é sinônima de gratidão, possuindo como semântica: reconhecimento por um benefício recebido ou agradecimento; dar provas de gratidão. Gostaria de agradecer a todas as pessoas que se envolveram de maneira direta e indireta no decorrer do desenvolvimento deste trabalho, realizado, pensando nos possíveis benefícios que a sociedade brasileira poderá obter com esses resultados. À minha família, pais Aldemar e Silvia, irmãos Carol e Tiago, por serem a base de tudo. À minha amiga Flávia, companheira, amiga de grupo LINEAR, *startup* e “pau para toda obra”. Ao Ivan, parceiro de grupo, pedal e cerveja, também a Katiane, Aline, Professor César, Professor Rafael, Professor Sérgio que, senão dessem o amparo científico e pessoal na hora certa, o trabalho não teria fluído, meus agradecimentos. À minha república de graduação, a saudosa e eterna Monte Olimpo. Aos amigos de república Power Guido de pós-graduação.

O agradecimento especial, e que mudou minha vida nesses últimos tempos, é direcionado a um grupo de pessoas que jamais esquecerei, o Grupo de Pesquisas Avançadas em Inteligência artificial no Setor AgroFlorestal – LINEAR. Por meio desse trabalho, pude experimentar um universo de informações que transformaram a minha vida tanto pessoal quanto profissional de maneira geral.

Agradecimentos à Universidade Estadual Paulista “Júlio de Mesquita Filho”, por ser a base de toda a formação profissional, e uma dedicatória a todos os professores que trabalham na instituição, meus sinceros agradecimentos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001

“E um velho preceito meu que é de capital importância para um jovem autor publicar (se for com nome dele) apenas o que for muito bom e novo; de forma que o público creia nele e leia o que ele escreve.”

MOORE, J; DESMOND, A. **Darwin**: a vida de um evolucionista atormentado. Tradução de Cynthia Azevedo. 4. ed. rev. e ampl. São Paulo: Geração Editorial, 2001. p. 616.

RESUMO

A aplicação de técnicas de inteligência artificial é ampla e envolve ciências políticas, econômicas, geopolíticas, sociais, agrárias, ambientais, legais e outras. Nas agrárias, elas são utilizadas e experimentadas no reconhecimento de pragas, classificação de culturas, contagem de indivíduos arbóreos visando a otimização e redução de custos. Dentre estas técnicas, o *deep learning,* ou aprendizado profundo de máquinas, é uma opção interessante. Também conhecida como redes neurais profundas convolucionais, tem por base modelos matemáticos capazes de realizar o processo de classificação de imagens. No setor florestal, essa metodologia pode ser aplicada visando a melhoria de produtividade e auxiliando na conservação. O presente estudo teve como objetivo aplicar *deep learning* para classificar imagens de mudas florestais no processo de identificação de mudas e prognóstico visual. O estudo envolveu a criação de dois *datasets*: um com 610 imagens, sendo 400 para treino, 200 para teste e 10 para validação, balanceado entre as classes: pau d'alho e ipê-verde; e o outro com imagens de cabreúva, pau d'alho, brinco-de-índio, jatobá e ipê-verde, composto por 1298 imagens sendo 1022 para treino, 256 para teste e 20 para validação. As bibliotecas importadas para realização do processamento e classificação foram *Keras,* com o *framework* do *Tensorflow,* e *Matplotlib,* todas usando a linguagem de programação *Python*. A avaliação da performance preditiva do modelo utilizou o *10-fold cross validation*. Na classificação entre as mudas de pau d'alho e ipê-verde, a acurácia média foi de 96,84%. Para o teste de prognóstico visual, as classes foram divididas em mudas com anomalia e mudas padrão, e a acurácia média foi de 49,18%. Assim, para diferenciar as duas espécies, o modelo pode ser incorporado em uma *API* ou *Software* para aplicação em viveiros florestais. Para utilização com fins de realização do prognóstico visual, serão necessárias adaptações.

Palavras-chave: inteligência artificial; redes neurais convolucionais; viveiros.

ABSTRACT

The application of artificial intelligence techniques is broad and involves political, economic, geopolitical, social, agrarian, environmental, legal, and other sciences. In the agrarian sciences, they are used and experimented with within the recognition of pests, classification of crops, counting of tree individuals aiming at optimization, and cost reduction. Among these techniques, deep learning or deep machine learning is an interesting option. Also known as convolutional deep neural networks, it is based on mathematical models capable of performing the image classification process. In the forestry sector, this methodology can be applied to improve productivity and aid in conservation. The present study aimed to apply deep learning to classify images of forest seedlings in the process of seedling identification and visual prognosis. The study involved the creation of two datasets: one with 610 images, 400 for training, 200 for testing and 10 for validation, balanced between the classes: pau d'alho and ipê-verde and the other with images of cabreúva, pau d'alho, brinco- de-índio, jatobá, and ipê-verde, composed of 1298 images, 1022 for training, 256 for testing and 20 for validation. The libraries imported for processing and classification were Keras, with the Tensorflow framework, and matplotlib, all using the python programming language. The evaluation of the predictive performance of the model used 10-fold Cross-Validation. In the classification between pau d'alho and ipê verde seedlings, the average accuracy was 96.84%. For the visual prediction test, the classes were divided into seedlings with anomaly and standard seedlings, and the average accuracy was 49.18%. Thus, to differentiate the two species, the model can be incorporated into an API or Software for application in forest nurseries. For use for visual prediction purposes, adaptations will be required.

Keywords: artificial intelligence; convolutional neural networks; nurseries.

SUMÁRIO

1	**INTRODUÇÃO.....**	**17**
2	**REVISÃO DE LITERATURA.....**	**18**
2.1	**Identificação Botânica.....**	**18**
2.2	**Visão Computacional.....**	**20**
2.3	**Aprendizado de Máquinas ou *Machine learning*.....**	**21**
2.4	**Aprendizado Profundo de Máquinas ou *Deep Learning*....**	**22**
2.4.1	**Redes Neurais Profundas Convolucionais**.....	**23**
2.5	**Viveiros Florestais.....**	**24**
2.6	**Prognóstico Visual.....**	**26**
3	**MATERIAL E MÉTODOS.....**	**28**
3.1	**Estudos Preliminares.....**	**28**
3.2	**Open Source.....**	**33**
3.3	**Experimento Final.....**	**33**
3.4	**Avaliação da performance preditiva do modelo.....**	**36**
4	**RESULTADOS E DISCUSSÃO.....**	**37**
4.1	**Aplicação da metodologia na identificação botânica e seus possíveis impactos.....**	**39**
4.2	**Aplicação de inteligência artificial e seus possíveis impactos.....**	**40**
5	**CONCLUSÕES.....**	**41**
	REFERÊNCIAS	**43**

1 INTRODUÇÃO

As recentes aplicações em aprendizagem profunda de máquinas, conhecida como *deep learning,* demonstram a relevância dessa área de estudo. Alguns campos de sua aplicação são: estratégias de negócio; experimentos nas áreas médicas; política; ciências sociais e ambientais; economia, além da sua potencialidade em demais ciências (WANG, 2019).

A economia mundial, geopolíticas e relações internacionais serão cada vez mais afetadas com a utilização em larga escala dos recursos de inteligência artificial, como *deep learning* (OSEI, 2020). No setor florestal não é diferente, o uso desses recursos tem sido importante nas áreas produtivas e em ações de conservação. No combate ao desmatamento ilegal, a ONG *Rainforest Connection* realiza o monitoramento de aproximadamente 300ha de floresta amazônica, captando os sons da floresta com celulares que seriam descartados, servindo, essa sonoridade, de dados de entrada para que o sistema reconheça padrões e prognostique possíveis ações ilegais dentro da área (DAUVERGNE, 2020). Na produtividade, o caso da Fibria da unidade de Três lagoas (MS) confirma essa questão. A corporação inaugurou, em 2017, o viveiro de mudas de eucalipto com processos autônomos baseados em inteligência artificial, havendo, com isso, um diferencial na produção final de celulose (FIBRIA, 2017).

Os viveiros florestais são a base do setor, as mudas produzidas neles podem ser utilizadas para plantios comerciais, recuperação de áreas degradas, arborização urbana, entre outras. A otimização de alguns ou todos processos produtivos de mudas florestais pode gerar uma redução de custos e melhorias no manejo. Dentre esses processos, citamos a identificação de mudas e prognóstico visual, visando as condições de qualidade da planta. A identificação correta resulta em assertividade no manejo final e o prognóstico visual reduz perdas e, se realizado no momento adequado, pode garantir a qualidade das mudas e a sua sobrevivência em campo.

Assim, o presente trabalho teve como objetivo aplicar *deep learning* na classificação de imagens de mudas florestais, visando a diferenciação de mudas de pau d’alho e ipê-verde e, realização de prognóstico visual de mudas de jatobá, pau d’alho, ipê-verde, brinco-de-índio e cabreúva, classificando as espécies em duas classes: anomalia ou padrão.

2 REVISÃO DE LITERATURA

2.1 Identificação Botânica

O estudo da botânica se originou a partir de registros de plantas medicinais nos livros dos templos egípcios, como o *Livro dos Mortos* e o *Livro dos Vivos*. No *Livro dos Mortos*, os registros são sobre as aplicações de plantas no embalsamento de cadáveres, enquanto no *Livro dos Vivos*, há descrições e usos de plantas no combate a diversas doenças. Também existem relatos sobre gregos que registraram observações primitivas sobre plantas. Assim, à medida que o conhecimento aumentou, exigiu-se uma organização maior dessas informações (GEMTCHÚJNICOV, 1976).

A Identificação Botânica é um ramo da Sistemática Vegetal que tem por objetivo determinar se uma planta pertence a um grupo já conhecido de plantas. O conceito de identificação é muitas vezes confundido com o da classificação, que consiste em localizar os organismos em um sistema de táxons e de categorias hierárquicas. Essas categorias são compostas de grupos grandes e inclusivos de organismos, como Reino, Filo, Classe, Ordem, Família, Gênero e Espécies (JUDD et al., 2008).

O homem sempre teve a necessidade de classificar, por meio de caracteres similares, seres e objetos para facilitar seu manuseio ou reconhecimento (QUINE, 1987). Esse ato de classificar é uma capacidade inata que a mente humana realiza involuntariamente e que tem registro desde o homem primitivo, que classificava as plantas em comestíveis e não comestíveis.

A necessidade de agrupar seres com características comuns, e nomeá-los, deu-se a partir de observações e estudo dos seres vivos, criando sistemas hierárquicos de categorias denominados sistemas de classificação biológica, que podem ser Artificiais, Naturais e Filogenéticos, cada qual construído por diferentes objetivos, princípio filosóficos, normativas e conjuntos de caracteres distintos (BARROSO et al., 2002; BRITO; AGUIAR; LIMA, 2013).

Na Botânica, os sistemas de classificação se desenvolveram em quatro períodos: o primeiro agrupava plantas com hábitos semelhantes, como árvore, arbustos, ervas e trepadeiras. Esse sistema era pouco elaborado e atendia as necessidades humanas como alimentação, medicina, construção, entre outros.

Dentre o principal nome do período tem-se Theophratus (370 a.C.), considerado o pai da Botânica (BARROSO et al., 2002; JUDD et al, 2008).

O segundo período é caracterizado pelos sistemas artificiais, baseados em caracteres numéricos, sem a preocupação de mostrar relações de afinidade. Nesse sistema é utilizado poucos caracteres para construir grupos, empregando, por exemplo, número de estames e sua posição na flor e número de estilete para agrupar indivíduos em diferentes categorias taxonômica. Lineu (1707-1778) foi o principal nome desse período. Ele é considerado o pai da taxonomia moderna e do sistema de nomenclatura que utilizamos na atualidade (BARROSO et al., 2002; BRITO; AGUIAR; LIMA, 2013).

O terceiro período reflete os sistemas naturais, que utilizam muitos caracteres para construir grupos. Ele surge com o crescimento da morfologia vegetal, de novas coleções originadas dos trópicos e da melhoria dos instrumentos ópticos. Dentre os principais nomes do período têm-se a Família Jussieu, que reconheceu 100 ordens de plantas, hoje consideradas famílias, e a Família De Candolle, que tentou descrever todas as espécies vegetais conhecidas (BARROSO et al., 2002).

Com a teoria Evolucionista proposta por Darwin e Wallace na segunda metade do século XVII, os sistemas naturais de classificação são finalizados, pois, a partir desse ponto, os agrupamentos dos organismos dão-se pela afinidade, que está baseada em ancestralidade e descendência. Assim, o quarto período é caracterizado pelo sistema filogenético, que procura usar todas as informações disponíveis no momento a respeito dos táxons envolvidos. Engler, Bessey e Cronquist são os principais nomes desse sistema de classificação (BARROSO et al., 2002).

Além de agrupar os sistemas de classificação, o homem sentiu a necessidade de nomear esses grupos, facilitando o reconhecimento e estudos dos seres. Assim, foram criados o nome científico e as categorias taxonômicas, tornando, para comunidade científica, uma linguagem universal. Lineu criou o sistema binomial de nomenclatura que é utilizado na atualidade e que, para escrevê-lo corretamente, faz se o uso do Código Internacional de Nomenclatura para algas, fungos e plantas (ICN), que abrange um conjunto de normas e recomendações e assegura que não haja duplicidade nominal, tornando o táxon reconhecível e aceito em todo mundo (VIDE et al., 2018).

Assim, para identificar um grupo taxonômico vegetal é necessário utilizar um sistema de classificação. A identificação pode ser realizada por meio de chaves de identificação, comparação com táxons nos herbários e por especialista da área (SOUZA; LORENZI, 2014). No entanto, novas tecnologias estão sendo elaboradas no sentido de classificar e/ou identificar os táxons a partir de imagens aéreas, visão computacional e inteligência artificial (KLOTII, 2015; KOGAN, 2013). A interação entre essas ferramentas pode transformar de forma positiva a diferenciação de táxons, melhorando o processo de levantamento, acelerando e tornando-o menos oneroso, quando comparado com os métodos atuais.

Os métodos atuais são manuais e se referem à confecção de chaves de identificação elaboradas pelo especialista da área. Esse processo pode ter uma baixa eficácia, pois os caracteres apresentados na chave são determinados pela experiência do pesquisador e, quando utilizada por outras pessoas, pode gerar dúvidas com relação aos caracteres aplicados para separar os táxons em questão e, assim, não poder chegar corretamente à categoria taxonômica desejada. Essas dificuldades, presentes na taxonomia tradicional, refletem-se na criação de métodos automatizados existentes, embora existam poucos trabalhos em visão computacional para classificação vegetal que utilizem imagens. Logo, esse método pode vir a contribuir no processo de identificação vegetal em diferentes áreas (CASANOVA, 2008).

Dessa forma, esse novo modelo de identificação botânica permite, a curto prazo e com baixo custo, o reconhecimento de espécimes botânicos de um determinado bioma e, consequentemente, trará medidas imediatas visando a mitigação com relação à perda de nossa biodiversidade vegetal.

2.2 Visão Computacional

A visão computacional é uma área de estudos que busca, por meio de programas, identificar objetos representados em imagens digitalizadas (BRITANNICA, 2020). Também pode ser definida como uma área especializada no processamento de imagens, onde o dado de entrada (*input*) é uma imagem e a saída (*output*) é um conjunto de valores numéricos que são utilizados na interpretação dessa imagem (GONZALEZ, 2010). Szeliski (2010) definiu que, o objetivo da visão computacional é obter imagens por meio de um dispositivo e, a

partir delas, realizar a extração de características necessárias para efetuar-se o processamento, possibilitando, assim, uma decisão assertiva. Como, por exemplo, um braço robótico seletor de resíduos em uma linha industrial.

A detecção, classificação e identificação de objetos em uma imagem por meio de algoritmos é a área conhecida na visão computacional como reconhecimento de padrões (*pattern recognition*) (BALLARD; BROWN, 1982). De acordo com Prince (2012), há a possibilidade de se utilizar pré-processamentos para que se atinja o objetivo da visão computacional, devido ao grande número de elementos que uma imagem possui. Cita-se, como exemplos, o recorte e adequação de dimensões da imagem de entrada, com o intuito de extrair e/ou identificar os elementos desejados, por meio de algoritmos teorizados e testados.

Um sistema de visão computacional envolve algumas etapas, como: a aquisição da imagem; extração de características; detecção de segmentação; pré-processamento e processamento de alto nível (FACON, 2006). Neste trabalho, foram aplicadas a aquisição e extração de características.

2.3 Aprendizado de Máquinas ou *Machine Learning*

O aprendizado de máquinas é uma subárea da inteligência artificial com o objetivo de estudar algoritmos e técnicas que permitem que o computador aprenda (MONARD; BARANAUSKAS, 2003). Esse método é baseado no princípio da inferência por indução. Assim sendo, com um determinado conjunto de dados é possível obter conclusões (LORENA, 2006). Para obter esses resultados, existem três tipos de aprendizado indutivo:

- Supervisionado: a máquina realiza o aprendizado baseando-se em informações já conhecidas, ou seja, por meio de saídas predeterminadas que viabilizam a indução, durante o aprendizado, e a obtenção do resultado esperado (HAYKIN, 1999). Dessa forma, o processo consiste na seleção de alguns exemplos e, se necessárias, informações adicionais para o indutor. Após a indução, o resultado deve ser avaliado, para identificar se a saída gerada corresponde à saída esperada, ocorrendo, se preciso, a repetição do processo.
- Semi-supervisionado: rotular um grande número de dados pode ser uma tarefa onerosa e delongada. Portanto, é possível solicitar ao especialista no

domínio, que classifique apenas alguns exemplos do conjunto de dados, os quais ele tenha certeza sobre os rótulos (*labels*) associados, utilizando, desse modo, algoritmos semi- supervisionados.

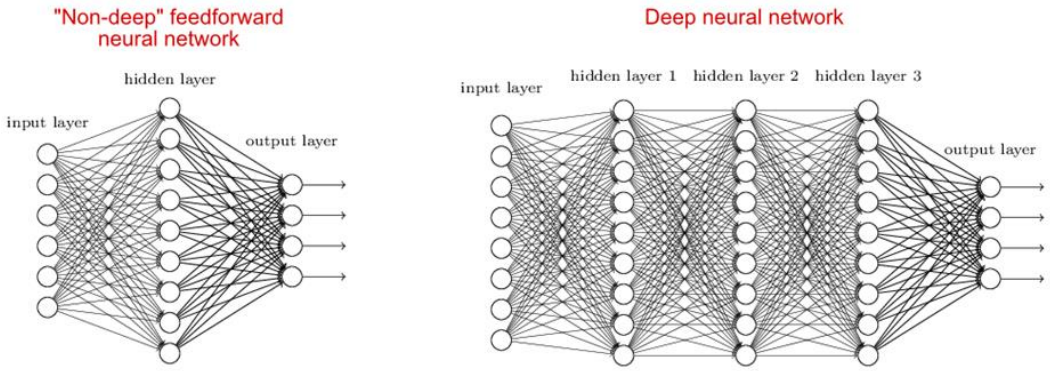
- Não-supervisionado: para esse aprendizado, a máquina é treinada com uma base de informações sem saídas predeterminadas. Por isso, o sistema agrupa, por similaridade, as amostras, formando os *clusters*. Essa técnica é utilizada principalmente para reconhecer padrões ou tendências que auxiliem no entendimento dos dados (LORENA, 2006). Para tratar do aprendizado com uma grande quantidade de informações de entrada, nasceu o conceito de aprendizado profundo de máquinas ou *deep learning*.

2.4 Aprendizado Profundo ou *Deep Learning*

O aprendizado profundo de máquinas é um ramo do aprendizado de máquinas que vem se desenvolvendo nos últimos anos (HINTON, 2007), tendo seus estudos intensificados após desenvolvimento de pesquisas da Google em reconhecimento de imagens (LE et al., 2012).

O principal objetivo dessa técnica é utilizar classificadores, como as redes neurais profundas (*Deep Neural Networks*), para possibilitar a detecção de padrões mais abstratos e, portanto, complexos. As redes neurais artificiais surgem como um esforço de imitar o funcionamento do cérebro humano e adquirir a capacidade de executar tarefas. A *deep learning* baseia-se em uma grande quantidade de camadas de redes neurais artificias (Figura 1) em sequência que permite ao computador detectar padrões e realizar tarefas de classificação ou regressão após treinamento com um conjunto de dados adequado (GOODFELLOW, BENGIO; COURVILLE, 2016).

Figura 1 - Diferença entre redes profundas e não profundas



Fonte: Stackexchange (2015).

A técnica pode ser aplicada na classificação de imagens, onde a rede neural recebe como entrada imagens, como as RGB, e produz como saída a probabilidade desse exemplo pertencer a um conjunto possível de classes, sendo possível diferenciar cachorros, gatos, tartarugas e corujas, por exemplo. Esse resultado na classificação gera uma função, temos f(x) = y, onde x representa a imagem e y a classe mais provável para x (MENDOZA; VELASTÍN, 2017). Assim, aliada ao sensoriamento remoto, pode ser uma técnica importante na identificação de espécies por meio de imagens digitais.

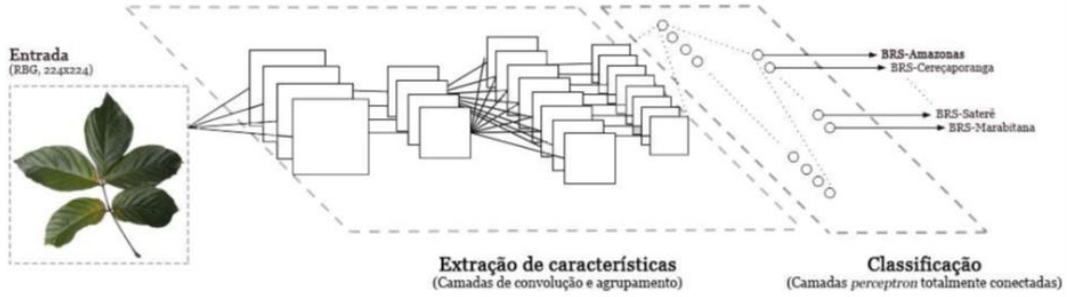
2.4.1 Redes Neurais Convolucionais (CNN's)

O conceito de aprendizado profundo usando redes neurais se depara com limitações computacionais ligadas ao processo de treinamento dessas redes. Isso porque as topologias das redes neurais no aprendizado profundo necessitam de um grande número de ligações entre cada uma de suas camadas. Essas limitações são tratadas por meio de diversas soluções, como o uso de redes neurais convolucionais.

As Redes Neurais Convolucionais (*Convolutional Neural Networks* – CNNs), são um modelo matemático de classificação inspirado no sistema biológico de processamento de dados visuais. Foram obtidos resultados satisfatórios com esses algoritmos para classificação vegetal entre outras problemáticas (GU et al., 2015; JASSMAN et al., 2015). As CNNs aplicam diversos filtros nos dados de entrada, construindo uma espécie de hierarquia de características (Figura 2) que resultam em descritores precisos da imagem

(KRIZHEVSKY et al., 2012; CIRESAN et al., 2011). As informações da imagem são inseridas diretamente no modelo (Figura 2), que utiliza modelos matemáticos para a extração das características (valores) dessa imagem. Dessa maneira, não é necessário aplicar o processo prévio de extração de características, como quando aplicado à uma rede neural comum.

Figura 2 – Processamento de imagem em rede neural convolucional



Fonte: Vargas et al. (2016).

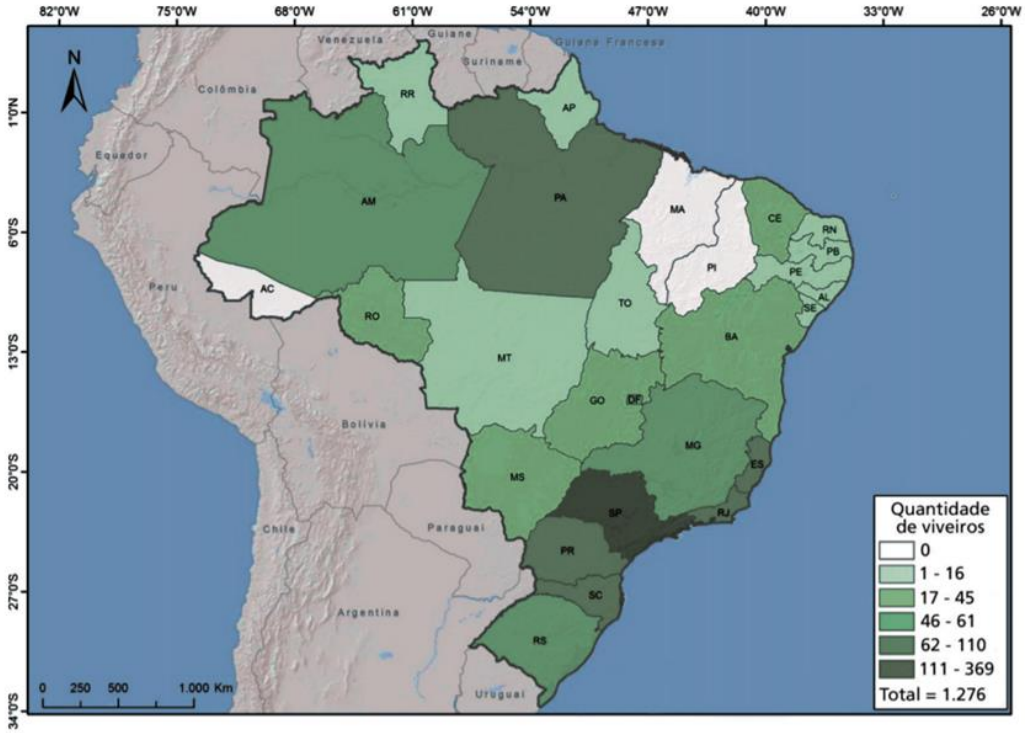
2.5 Viveiros Florestais

O viveiro de mudas florestais é a base da cadeia produtiva, sendo responsável pela produção de mudas que abastecem projetos de recuperação de áreas degradadas, arborização urbana, plantio de povoamentos florestais, restauração florestal entre outras finalidades (WALKER et al., 2011). O Decreto nº 5.153, de 23 de julho de 2004, da Lei nº 10.711/2003, define que o viveiro é uma área demarcada e tecnicamente adequada para a manutenção e produção de mudas (BRASIL, 2004). Segundo Wendling (2001), além da produção e manutenção, o viveiro de produção também realiza o acompanhamento do desenvolvimento da muda até o tamanho ideal e qualidade adequada, para que tenha a maior taxa de sobrevivência em campo.

A EMBRAPA (1998) classifica os viveiros florestais em permanentes ou temporários, sendo os permanentes aqueles onde a produção de mudas ocorre por vários anos e as instalações são maiores e melhor planejadas. Nos viveiros temporários, as mudas são produzidas para suprir uma determinada demanda, sendo instalados próximos ao local de plantio, que geralmente são de difícil acesso. No diagnóstico da produção de mudas florestais nativas no Brasil, foi

gerada uma lista com 1.276 viveiros produtores de espécies nativas, desses, 369 se encontram no estado de São Paulo (IPEA, 2015).

Figura 3 - Quantidade de Viveiros produtores de espécies nativas levantados por unidade federativa (UF)



Fonte: Ipea (2015).

Atualmente o estado de São Paulo é composto por 296 viveiros florestais, concentrados nas regiões do Sudeste e Centro do estado, com 72,6% (VIDAL; RODRIGUES, 2019), demonstrando sua produtividade e potencial de mercado. Existem algumas dificuldades na produção de mudas nativas florestais, como a falta de referências bibliográficas para produção (MARTINS, 2011), falta de intercâmbio de experiência entre os produtores, organização e gestão empresarial, concorrência com viveiros informais e cumprimento da legislação (VIDAL; RODRIGUES, 2019).

Segundo o IPEA (2015), os principais problemas relatados pelos produtores entrevistados foram: a dificuldade técnica para cultivo (adubação, pragas e doenças); dificuldade em armazenar e germinar sementes; falta de organização do setor; suprimentos de sementes; comercialização das mudas e falta de mão de obra capacitada.

Com o intuito de melhorar a produtividade e administração de viveiros florestais, há o surgimento de alguns *softwares*. Um exemplo que pode ser citado é o do viveiro *web*, que tem como objetivo: viabilizar a gestão de mudas, por meio de um planejamento à distância da produção; gerenciar fornecedores; rastrear lotes de produção; controlar vendas; auditar atividades e usuários; entre outras funções (CAMPOS, 2020). Nesse contexto, um sistema inteligente, que seja capaz de diferenciar mudas nativas e realizar prognósticos sobre o estado fitossanitário, estado nutricional, estresse hídrico e outras variáveis de qualidade da muda, baseado em modelos matemáticos, pode trazer benefícios ao setor.

2.6 Prognóstico visual

Monitorar o desenvolvimento das mudas nos viveiros florestais pode reduzir os riscos com perdas ocasionadas por: ataque de pragas e doenças; estresse hídrico; deficiência nutricional; entre outros. Fatores que, quando diagnosticados e controlados, podem favorecer na obtenção da “qualidade de muda”. Fonseca (2002) relata que, para que a muda tenha qualidade, são considerados alguns atributos ideais para o sucesso do plantio no campo.

A qualidade de mudas é um fator intrínseco à espécie e a escolha adequada, somada aos tratos ideias, favorecerão a adaptação e sobrevivência das mudas no campo (CARNEIRO, 1995 *apud* SOUZA, 2017). Para determinar esse fator, é necessário utilizar conceitos relacionados à morfologia e fisiologia da planta (GOMES, 2002). Os parâmetros morfológicos são os mais utilizados nessa determinação (MULA, 2011), visto os atributos morfológicos da planta serem passíveis de análise visual, o que facilita a observação. Há estudos que objetivam mostrar a importância dos critérios visuais, morfológicos, no desenvolvimento e sobrevivência das mudas após plantio (FONSECA, 2002 *apud* SOUZA, 2017).

Para o bom desenvolvimento da floresta, é importante garantir que as mudas estejam adequadas para o campo, assim o controle da qualidade da mudas deve ser feito antes da expedição (GOMES et al., 2002).

De tal modo, o desenvolvimento de um sistema capaz de prognosticar qualquer indicador de perda na qualidade da planta, tais como: pragas visíveis a olho nu; deficiência nutricional; estresse hídrico e outros, pode auxiliar no controle da qualidade das mudas. Sladojevic et al. (2016) propuseram a

detecção de doenças utilizando *deep learning* e redes neurais convolucionais. Foram classificadas 13 doenças, tendo o modelo atingido 96,3% de acurácia. Para classificação de cultivares de guaranazeiro, foi desenvolvido o sistema WAVA, que utiliza redes neurais convolucionais para classificar imagens, obteve-se, com esse sistema, 97% de acurácia (SOUSA; SALAME, 2017) .

Uma aplicação móvel que utiliza rede neural convolucional foi desenvolvida para classificar doenças foliares (Figura 4) da videira e auxiliar agricultores. O *dataset* foi composto por 4000 imagens distribuído nas classes: Saudável, *Black rot*, *Esca*, *Leaf Blight* ou *Isariopsis* entre outros. O melhor resultado em acurácia, dentre os modelos testados, foi de 97% (CARVALHO; ZOBY, 2018).

Figura 4 - Interface do aplicativo para identificação de doença foliar

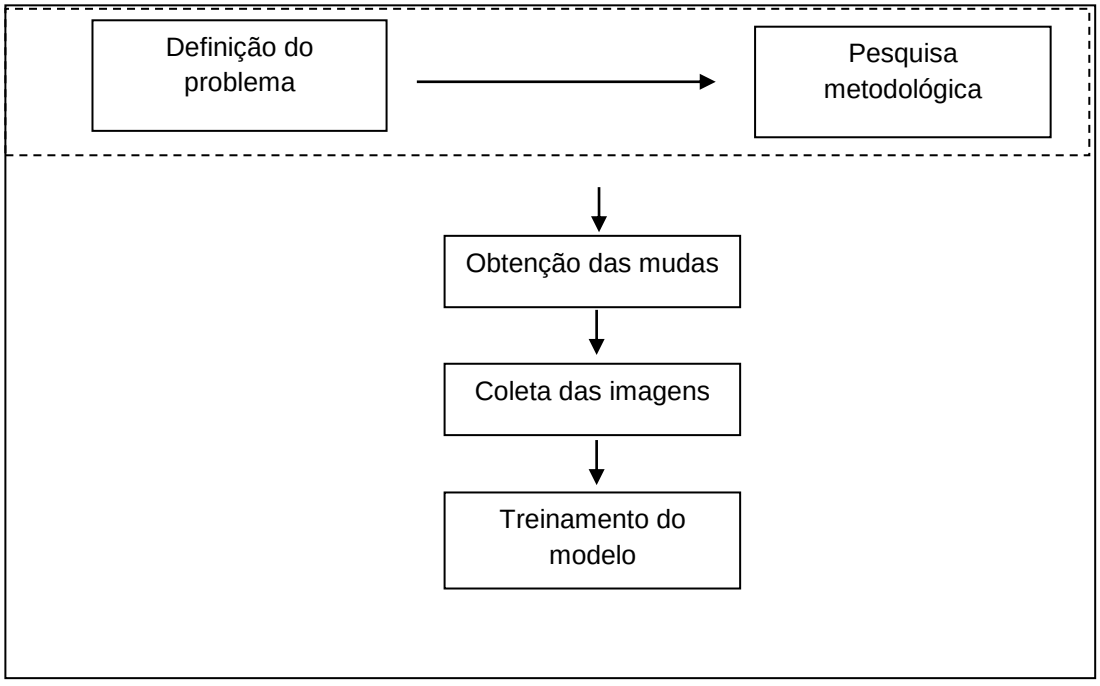
Fonte: Carvalho e Zoby (2018).

3 MATERIAL E MÉTODOS

3.1 Estudos Preliminares

Foram feitas imagens das mudas florestais de canafístula (*Peltophorum dubium* (Spreng.) Taub.) e pau d'alho (*Gallesia integrifolia* (Spreng.) Harms), cedidas pelo Viveiro de Pesquisa em Produção de Mudas Florestais, pertencente ao Departamento de Ciência Florestal da Faculdade de Ciências Agronômicas (FCA) - UNESP, no município de Botucatu – SP. As principais etapas do trabalho (Figura 5) contaram com a definição do problema, pesquisa metodológica, obtenção das mudas, coleta das imagens e treinamento do modelo.

Figura 5 - Fluxo de trabalho



As imagens foram padronizadas na resolução de 1028 x 768 pixels e feitas com um celular *Smartphone*. Para aumentar a quantidade de exemplos, foram utilizados 3 indivíduos de cada espécie, totalizando 6 mudas, que foram fotografadas individualmente (Figura 6) e agrupadas (Figura 7) entre os indivíduos de mesma espécie. Além disso, as imagens foram coletadas em ângulos e horários diferentes, visando o aumento de características para conjunto de dados (*Dataset*). Para Deng et al. (2009) e Sun et al. (2017), a

utilização de um grande volume de dados para aplicações em *deep learning* é necessária para se obter bons resultados.

Figura 6 - Imagens de canafístula (A) e pau d'alho (B) individuais na bandeja de tubetes

Figura 7 - Imagens de canafístula (A) e pau d'alho (B) agrupadas na bandeja de tubetes

O experimento foi realizado em uma máquina com sistema *Linux mint* baseado em *Ubuntu* versão 14.1 com algumas ferramentas instaladas, que otimizam a execução do treinamento. Foi utilizada uma arquitetura de rede neural profunda convolucional (CNN) da competição *Kaggle dogs and cats*

(Figura 8). Na sequência foi adaptada e treinada com as imagens das mudas florestais para a classificação dos indivíduos.

Figura 8 - Importação do modelo *dogs and cats* no *Jupyter Notebook*

A arquitetura do algoritmo foi adquirida no maior repositório de código aberto, para Akula, Yousefi e Garibay (2019), o *GitHub* usa a API escrita em *Python* para redes neurais, o *Keras.* Essa API utiliza um *framework* de *deep learning* como base, o *Tensorflow* da *Google Brain*.

O ambiente de programação utilizado para o treinamento e visualização dos resultados foi o *Jupyter Notebook,* presente nas bibliotecas do *Anaconda. Jupyter* é uma ferramenta interativa da *web* de código aberto e gratuito, conhecida como notebook computacional. A ambiente web é preparado para combinar códigos em *Python* e obter de forma visual e lógica a execução dos scripts de programação linear, sendo o processo de solução iterativa impresso, após a execução de um código (SUÁREZ, 2018).

A distribuição Anaconda é um ambiente de desenvolvimento de código aberto para ciência de dados em *Python* e *R*, baseada no conceito de notebook interativo (ANACONDA, 2020). Inclusas nas ferramentas utilizadas estão o *Jupyter Notebook*, *TensorFlow*, *Numpy* e *matplotlib* (Figura 12).

A importação do *Matplotlib* foi realizada pela necessidade de plotar em 2D do *Python,* números de qualidade de publicação no ambiente interativo *Jupyter* (MATPLOTLIB, 2020), permitindo a visualização dos dados por meio da sua interface gráfica. Outra característica importante deste recurso é a sua capacidade de funcionamento para muitos sistemas operacionais e *beck-end* gráficos (VANDERPLAS,2017).

Alguns parâmetros são importantes para a realização do treinamento, sendo eles o conjunto de treino, mini *batch size* e número de épocas (*epochs*) (Figura 9). O conjunto de treino, ou *dataset,* descrito acima foi dividido em 3, sendo uma para treino, outra para teste e outra para validação. O conjunto de teste contém o maior número de amostras, com 154 imagens, sendo 98 imagens de Canafístula (*Peltophorum dubium*) e 56 imagens de pau d'alho (*Gallesia integrifólia*). O conjunto do experimento continha 6 indivíduos de cada espécie. As imagens utilizadas para treino, teste e validação não foram repetidas,

portanto, os dados utilizados no conjunto de treino são diferentes do conjunto de teste e de validação.

Figura 9 - Tamanho do *Batch size*, número de classes e número de *epochs* utilizadas no teste

```
In [2]: batch_sz = 8 # Batch size
        nb_class = 2  # Número de classes
        nb_epochs = 40 # Número de epochs de treinamento
        img_h, img_w = 64,64  # Altura e largura das imagens
        dir_treino = './novoteste/train/'
        dir_teste = './novoteste/test/'
```

O *mini batch size* ou tamanho do lote é o número de exemplos de treino executados para que os pesos sejam utilizados, devendo ser maior que 1, mas menor que o tamanho total do conjunto de dados. No presente estudo foi atribuído um valor de 8.

Uma época é uma iteração passando por todos os dados do conjunto de treino e foi atribuído o número de 40 para esse experimento.

As funções das bibliotecas de código aberto da *Open Source Computer Vision* (*Open cv*) são responsáveis por realizar a filtragem de ruídos, reconhecimento de padrões e isolar objetos de interesse na imagem (MARENGONI; STRINGHININI, 2009).

As camadas convolucionais extraem características locais para formar o mapa de características (*feature maps*), por meio de um filtro bidimensional (*kernel)* que é deslizado por toda a entrada (imagem) (Figura 10). Neste experimente, a convolução foi realizada usando um filtro de 3 x 3. As entradas e as camadas convolucionais têm ligações com parâmetros (pesos), que são ajustados de forma autônoma durante o aprendizado (SOTHE et al., 2020). Esses pesos, assim como o *bias*, são compartilhados por cada uma das entradas que faz parte desse filtro. As convoluções são acompanhadas de uma função de ativação, como a *Rectified Linear Units* (ReLU), sendo um tipo de função de não-linearidade, utilizadas em *deep learning* por acelerar a aprendizagem (NAIR; HINTON, 2010).

Figura 10 - Descrição das camadas e execução no *Jupyter Notebook*

In [6]:
```
def conv3x3(input_x,nb_filters):
    # Prepara a camada convolucional
    return Conv2D(nb_filters, kernel_size=(3,3), use_bias=False,
                  activation='relu', padding="same")(input_x)
```

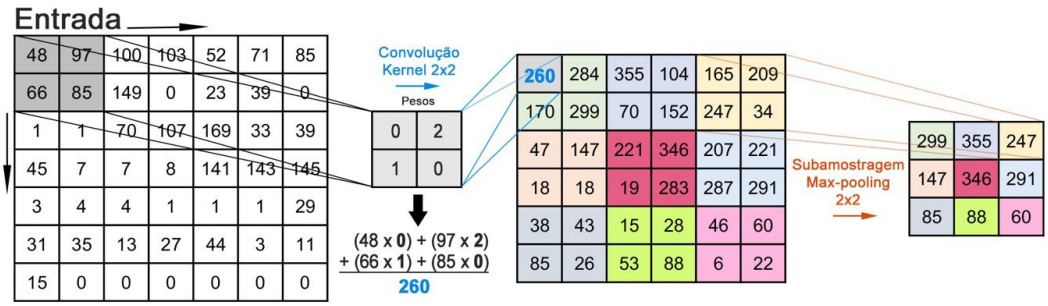
In [7]:
```
# Normaliza os valores dos pixels
x_treino = x_treino.astype('float32')
x_teste = x_teste.astype('float32')
x_treino /= 255.0
x_teste /= 255.0

# Converte os rótulos para "One-hot encoding":
y_treino = keras.utils.to_categorical(y_treino, nb_class)
y_teste = keras.utils.to_categorical(y_teste, nb_class)

# Cria o modelo executando um treino e avaliação:
inputs = Input(shape=(img_h, img_w, 1))
x = conv3x3(inputs, 32)
x = conv3x3(x, 32)
x = MaxPooling2D(pool_size=(2,2))(x)
x = conv3x3(x, 64)
x = conv3x3(x, 64)
x = MaxPooling2D(pool_size=(2,2))(x)
x = conv3x3(x, 128)
x = MaxPooling2D(pool_size=(2,2))(x)
x = Flatten()(x)
x = Dense(128, activation="relu")(x)
preds = Dense(nb_class, activation='softmax')(x)
model = Model(inputs=inputs, outputs=preds)
```

Uma subamostragem foi feita por uma camada típica, a *maxpooling*, que transmite o valor mais significativo (máximo de ativação) de determinada região da entrada. A aplicação desse recurso faz com que a dimensão dos dados de entrada seja reduzida, por exemplo, em uma entrada de 24 x 24 após a aplicação de um Maxpooling de 2 x 2, a saída será de 12 x 12.

Após as convoluções, são acrescentadas duas ou mais camadas totalmente conectadas (RNAs com ReLU**),** além da função *Softmax* do Tensorflow utilizada para normalizar os valores reais dos pontos na distribuição de probabilidades e, por fim, uma camada de saída, onde é realizada a classificação final ativada por um sigmoide, no caso de classificação binária (0 ou 1).

A topologia da Rede Neural Convolucional (Figura 11) foi determinada com uma entrada de 64 x 64 (dimensão da imagem), sendo aplicados na sequência: duas convoluções de 3 x 3, com 32 mapas de características, um Maxpooling de 2 x 2, mais duas convoluções de 3 x 3, com 64 mapas de características, seguidas de um Maxpooling de 2 x 2, uma nova convolução de 3 x 3, com 128 mapas de características, um novo Maxpooling de 2 x 2 e uma camada oculta tradicional com 128 neurônios. A camada de saída foi projetada com 2 neurônios para determinar as duas classes de resultado: canafístula ou pau d’alho.

Figura 11 - Exemplo de Convolução e *Maxpooling* como processos de amostragem e subamostragem para demonstrar a ativação das regiões de entrada



Fonte: Schmidt (2018).

3.2 Open Source

O código foi adquirido no maior repositório de códigos abertos, o *GitHub*, que conta com mais de 3 milhões de desenvolvedores registrados (Thung et al., 2013). Junto ao *StackOverflow*, uma plataforma de perguntas e respostas direcionada para desenvolvedores de *machine learning*, foi possível realizar as adaptações necessárias para a classificação, por meio da importação de bibliotecas preparadas para realização do teste (Figura 12).

Figura 12 - Importação das bibliotecas computacionais

```
import keras
import keras.backend as K
import os, random
from keras.layers import Dense, Conv2D, Input, MaxPooling2D, Flatten, Dropout
from keras.models import Model
from keras.datasets import fashion_mnist
from keras.callbacks import ModelCheckpoint
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import cm
import cv2
from keras.preprocessing import image
%matplotlib inline
```

3.3 Experimento Final

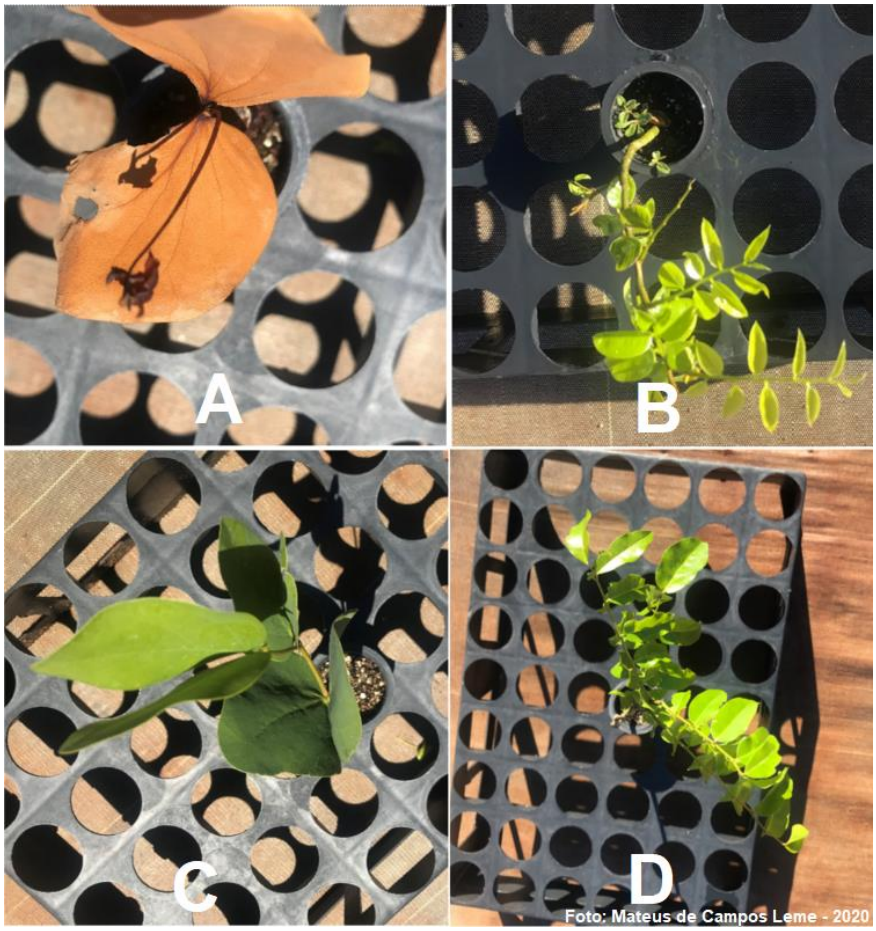
Os estudos preliminares evidenciam toda a metodologia utilizada para realização do experimento. Para o experimento final, foram montados dois *datasets* diferentes para os objetivos descritos. Para diferenciação de espécies, foram fotografadas mudas de ipê-verde (*Cybistax antisyphilitica* (Mart.) Mart.) e

pau d’alho (*Gallesia integrifolia* (Spreng.) Harms). Para o segundo conjunto, com objetivo de predição do prognóstico visual, foram selecionadas 5 espécies sendo pau d’alho, brinco-de-índio (*Cojoba arborea* (L.) Britton & *Rose*, cabreúva (*Myroxylon peruiferum* L. f.), ipê-verde e jatobá (*Hymenaea courbaril* L.).

Para ambos conjuntos de dados, todas as imagens foram padronizadas na resolução 1028 x 768 pixels. Na classificação de imagens para diferenciação das mudas, as imagens receberam os *labels* (Figura 13) de ipê e pau d’alho e o *dataset* foi composto por 610 imagens, sendo 400 para treino, 200 para teste e 10 para validação, sendo balanceado entre as classes.

Figura 13 - Imagens utilizadas no *dataset* de classificação binária entre mudas de pau d’alho e ipê-verde

Figura 14 - Mudas florestais com anomalia (A) jatobá e (B) cabreúva e mudas florestais padrão (C) jatobá e (D) Cabreúva



Para classificação de imagens com o objetivo de prognóstico, os *labels* foram padrão e anomalia (Figura 14) e o conjunto de dados foi formado por 1298 imagens sendo 1022 para treino, 256 para teste e 20 para validação.

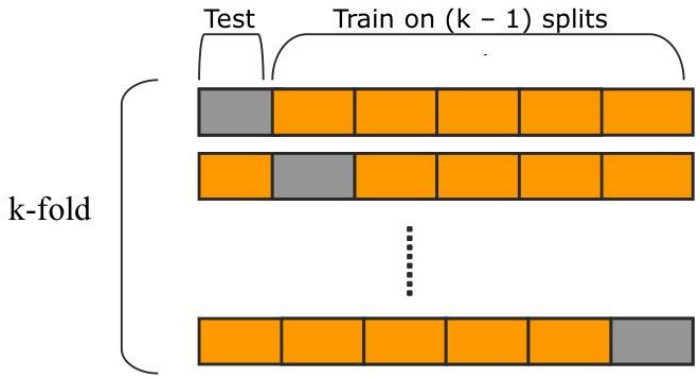
O tamanho do *batch* e número de *epochs*, que obtiveram os melhores resultados quanto ao poder preditivo do modelo utilizados nos testes, foram de 6 e 150 para a classificação das espécies e de 6 e 1000 para o prognóstico visual. Para obtenção desses valores, foram experimentados valores entre 2 e 8 para o batch e valores de 50, 150, 200, 500, 1000, 1500 e 2000 para as *epochs*.

Em *machine learning* é importante testarmos a capacidade de generalização ou poder preditivo do modelo experimentado e, para isso, foi feito o teste de *Cross Validation k- fold.*

3.4 Avaliação da performance preditiva do modelo

A avaliação da acurácia do algoritmo foi realizada por meio da validação cruzada *k-fold*, utilizada em experimentos de *machine learning*. Para Nurhayati et al. (2014), a validação cruzada é adequada para avaliar a performance do modelo e por meio dela é possível simular o desempenho preditivo. Com k=10 ou 10-fold, a base de dados é dividida em dez partes, onde 9 partes são utilizadas para treino ou aprendizagem e a outra para teste (Figura 15). Isso é repetido 10 vezes, de modo que toda a base de dados tenha servido como treino e teste (HELMA et al., 2004). Após a validação cruzada, foi feita a média aritmética dos valores obtidos para a obtenção da *performance* do modelo.

Figura 15 - Segmentação do conjunto de dados para avaliação do poder preditivo do modelo, *Cross Validation* (*k-fold*)



Fonte: Coelho (2020).

4 RESULTADOS E DISCUSSÃO

Os resultados iniciais mostraram a necessidade de adaptações no trabalho, como o balanceamento entre as classes amostradas, tamanho e tipo de *dataset*. Assim, é significativo ressaltar a realização dos testes iniciais para o desenvolvimento do projeto.

A utilização do modelo matemático em *software* ou *API* pode auxiliar no manejo do viveiro quanto ao tipo de muda produzida. O sistema binário, capaz de diferenciar as duas classes de dados (imagens de ipê-verde e pau d'alho), pode reduzir erros de identificação de espécies produzidas em viveiros. Visando estudos futuros e sua aplicação nas rotinas produtivas, há a necessidade do aumento de classes, nova configuração do conjunto de dados, testagem de outros modelos matemáticos e avaliação da aplicação.

Além da identificação, o prognóstico é parte das práticas de processo de produção de mudas. O diagnóstico precoce de pragas e doenças em mudas florestais evita perdas produtivas para o viveiro.

Tabela 1 - Resultados do *cross validation* da classificação das imagens de ipê-verde e pau d'alho

Teste	Acurácia (%)	Perda
k-1	95,0	0,7405
k-2	96,0	0,6250
k-3	95,0	0,7405
k-4	98,0	0,3223
k-5	99,4	0,0895
k-6	95,5	0,6326
k-7	95,0	0,7405
k-8	99,5	0,8050
k-9	97,5	0,4030
k-10	97,5	0,4030
Média aritmética	96,84	0,4777

A acurácia é utilizada em estudos de *machine learning* para medição do poder preditivo do modelo (KUMAR, 2020). Para diferenciação das mudas de ipê – verde e pau d'alho, o modelo apresentou valor médio de 96,84%. O valor máximo de 99,5% no k-8 e mínimo de 95% no k-1 e k-3. O valor é significativo, considerando o tamanho do conjunto de dados e o tempo médio de treinamento que foi de 28 minutos por k. Com esse resultado, pode-se aplicar o modelo para a finalidade de diferenciação dessas duas mudas, seguindo a mesma metodologia utilizada.

Para os valores de perda, quanto mais próximos de 1, maior a capacidade preditiva do modelo para o determinado conjunto de dados. Significa que maior será a probabilidade dos *labels* pertencerem a uma determinada classe ou não. A perda também é um indicador do potencial preditivo do modelo. No algoritmo foi utilizado o *Categorical Cross Entropy Loss Function* do *Keras* (KERAS, 2020). O valor médio de perda foi de 0,4777, com valores de máxima 0,7405 no k1, k3 e k7 e valor de mínimo de 0,0895 no k-5.

Tabela 2 - Resultados do *cross validation* no teste classificação de imagens do prognóstico visual

Teste	Acurácia (%)	Perda
k-1	50,00	0,0805
k-2	50,00	0,0805
k-3	50,00	0,0805
k-4	50,00	0,0805
k-5	50,00	0,0805
k-6	50,00	0,0805
k-7	49,80	0,7405
k-8	42,01	0,0934
k-9	50,00	0,0805
k-10	50,00	0,0805
Média aritmética	49,18	0,1478

Na avaliação do poder preditivo do modelo para prognóstico visual de mudas, o k-médio foi de 49,18% com valor de máximo de 50% e mínimo de

42,01%. Esses valores não significam que o modelo não pode ser utilizado para tal finalidade, mas que serão necessários novos testes para obtenção da melhora do poder preditivo, como indicado.

Os valores de perda não contrapõem o valor de acurácia, sendo a média 0,1478, distante de 1. A máxima foi de 0,7405 no k-7 e mínima de 0,0805.

Os testes de classificação de imagens de mudas florestais para predição do tipo de muda (pau d'alho x ipê verde) e prognóstico visual, utilizando o modelo de *machine learning,* evidenciaram o potencial de aplicação da técnica na rotina de viveiros florestais.

Com os resultados apresentados, a viabilidade prática do modelo deve sofrer alterações no prognóstico visual. Para novos testes, modificações no conjunto de dados, testagem de novos modelos e a identificação da anomalia podem ser considerados.

4.1 Aplicação da metodologia na identificação botânica e seus possíveis impactos.

Os resultados obtidos nessa amostragem mostram potencial aplicação na classificação botânica. Solucionando e otimizando os processos já citados no trabalho. Assim, a continuação dos estudos na área se faz importante, uma vez existente a possibilidade de fomento às ferramentas nessa linha de pesquisa.

Gonzáles (2009) realizou um levantamento sobre a quantificação de custo e tempo no processo de informatização das coleções biológicas brasileiras, estudo realizado no maior jardim botânico do Rio de Janeiro, e que está diretamente ligada a este trabalho. Segundo o autor, o valor da informatização fica em R$ 1,56 por espécie, valor estimado de forma adequada também por Peixoto (2006), com um valor de R$ 1,51. O mesmo autor, em 2006, levou em consideração a informatização de 332 mil espécimes de cerca de 17 instituições.

Os Impactos na classificação e identificação botânica serão significativos por meio da classificação de imagens utilizando técnicas de inteligência artificial, como realizado no trabalho. A otimização de alguns processos como ida a campo, armazenamento e ordenamento de informações, precisão e custo baixo pode ocorrer com a aplicação da inteligência artificial no processo.

4.2 Aplicação da inteligência artificial em viveiros e seus possíveis impactos.

Em 2017 a Fibria, empresa do setor florestal, divulgou parte do seu relatório anual. A companhia investiu no viveiro de mudas florestais, para produção de eucalipto, com o maior número de processos autônomos quando comparado aos sistemas produtivos de mudas florestais pelo mundo. A unidade fica na cidade de Três Lagoas (MS) e como resultado foi obtido um acréscimo de quase 2 milhões de toneladas/ano de celulose à capacidade produtiva da empresa. A capacidade produtiva do viveiro é de 43 milhões de mudas por ano (FIBRIA, 2017).

Dentro dos processos de automação da companhia, são utilizados alguns sistemas inteligentes no controle de qualidade das mudas. Esses sistemas são capazes de otimizar processos, devido a capacidade de processamento de um grande número de dados em pouco tempo. Assim, serão necessários estudos para validação dos reais impactos da utilização de técnicas de inteligência artificial em viveiros florestais.

5 CONCLUSÕES

A aplicação de técnicas computacionais é abrangente e no presente trabalho se mostrou satisfatória em ambos testes. O teste preliminar demonstrou a necessidade de alteração no balanceamento dos dados e da avaliação do poder preditivo do modelo pelo método do *cross validation*.

Nos experimentos finais, com novos conjuntos de dados, sendo um para identificação de duas espécies e outro para prognóstico visual de mudas florestais, o poder preditivo médio do modelo foi de 96,84% para classificação e 49,18% para prognóstico.

Assim, para utilização do modelo, visando uma *API* ou *Software* na identificação, o valor de acurácia indicou resultados satisfatórios, sem a necessidade de alterações quanto a quantidade e tipo de dados. Para a aplicação, com o objetivo de prognóstico visual, serão necessários novos testes.

REFERÊNCIAS

AKULA, R.; YOUSEFI, N.; GARIBAY, I. DeepFork: Supervised Prediction of Info Diffusion**. ArXiv.org. Cornell University.** 2019.

ANACONDA. **Data science technology for human sense making**. Disponível em < https://www.anaconda.com/>. Acesso em 12 de ago. 2018.

ARCO-VERDE, M. F.; MOREIRA, M. A. B. Viveiros Florestais: construção, custos, cuidados e atividades desenvolvidas para a produção de mudas. Boa Vista: **Embrapa CPAF-Roraima**, p. 18, 1998.

BARROSO, G. M.; PEIXOTO, A. L.; ICHASO, C. L. F.; GUIMARÃES, E. F. **Sistemática de Angiospermas do Brasil**. Viçosa, v.1, 2 ed. Viçosa, p. 309, 2002.

BRASIL. **Ministério da Agricultura, Pecuária e Abastecimento**. Lei n. 10.711, de 5 de agosto de 2003. Decreto nº 5.153, de 23 de julho de 2004. Brasília, DF, p. 122, 2004.

BRITO, R. A.; AGUIAR, W. J.; LIMA, I. M. **Materiais didáticos para o ensino de botânica, confeccionados a partir de materiais alternativos**, p. 3, 2013.

BROWN, C. M.; BALLARD, D. H.; KIMBALL, O. A. Constraint Interaction in Shape-from-Shading Algorithms. **Defense Technical Information Center**. p.11. 1982.

CAMPOS, F. **O desafio é viabilizar a gestão de mudas a quaisquer viveiros florestais**. Viveiro web, 2020. Disponível em < http://viveiroweb.com.br/#produto> Acesso em 15 jan. 2021.

CARNEIRO, J. G. A. **Produção e controle de qualidade de mudas florestais.** Curitiba: UFPR/FUPEP, p. 451, 1995.

CARVALHO, R. G.; ZOBY, L. T. M. Convolutional Neural Networks for Leaf Disease Classification. *In*: **ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL E COMPUTACIONAL (ENIAC),** 15 ed. São Paulo**.** Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, p. 332-342, 2018. DOI: https://doi.org/10.5753/eniac.2018.4428.

CASANOVA, D.; **Identificação de espécies vegetais por meio da análise de textura foliar**. 2008. 112 p. Dissertação (Mestrado em Ciência de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2008.

CIREŞAN, D. C.; MEIER, U.; MASCI, J.; GAMBARDELLA, L. M.; SCHMIDHUER, J. High-Performance Neural Networks for Visual Object Classification. **arXiv:1102.0183 [cs]**, 1 fev. 2011.

COELHO, G. **Como fazer cross validation**. Disponível em< https://gusrabbit.com/code/cross_validate/>. Acesso em 20 jan. 2020.

DAUVERGNE, P. AI in the Wild: Sustainability in the age of artificial intelligence. **The MIT Press**, Cambridge, Massachusetts. p. 1-8. 2020.

DENG, J.; DONG, W.; SOCHER, R.; LI, L-J; LI, K.; FEI-FEI, L. ImageNet: A Large-Scale Hierarchical Image Database. **IEEE Conference on Computer Vision and Pattern Recognition**. p. 248-255, 2009.

FACON, J. **Técnicas de processamento digital de imagens aplicadas à área da saúde**. xiii escola regional de informática da sbc-paraná. 2006.

FIBRIA. Gestão Florestal. **Relatório Fibria 2017**. Disponível em <http://r2017.fibria.com.br/nossos-processos-produtivos/gestao-florestal/>. Acesso em 20 dez. 2020.

FONSECA, E. P.; VALÉRI, S. V.; MIGLIORANZA, E.; FONSECA, N. A. N.; COUTO, L. Padrão de qualidade de mudas de *Trema micrantha* (L.) Blume, produzidas sob diferentes períodos de sombreamento. **Revista Árvore,** Viçosa-MG, v. 26, n. 4, p. 515-523, 2002.

GAUR, A. Computer vision. In: **Encyclopedia Britannica**. Disponível em <https://academicebbritannica.ez87.periodicos.capes.gov.br/levels/collegiate/article/computer-vision/472098> 2020.

GEMTCHÚJNICOV, I. D. **Manual de taxonomia vegetal**. Agronômica Ceres, São Paulo, p. 368. 1976.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Introduction: Historical Trends in Deep Learning. In: GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge: Mit Press, p. 12-18, 2017. Disponível em: <https://www.deeplearningbook.org>. Acesso em: 11 abr. 2019.

GOMES, J. M. et al. **Parâmetros Morfológicos na Avaliação da Qualidade de Mudas Florestais**. p. 10, 2002.

GONZALEZ, I.; BENGIO, Y.; COURVILLE, A. Introduction: Historical Trends in Deep Learning. In: GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge: Mit Press, 2017. p. 12-18. Disponível em: <https://www.deeplearningbook.org>. Acesso em: 11 abr. 2019.

GU, Y.; LIU, J.; CHEN, Y.; JIANG, X.; YU, H. TOSELM: Timeliness online sequential extreme learning machine. **Neurocomputing,** v. 128, p. 119–127, 2014.

HAYKIN, S. Neural Networks - A Compreensive Foundation.In: **Prentice-Hall**, p. 8. 1999.

HELMA, C.; CRAMER, T.; KRAMER, S.; RAEDT, L. Data Mining and Machine Learning Techniques for the Identification of Mutagenicity Inducing Substructures and Structure Activity Relationships of Noncongeneric Compounds. **Journal of**

Chemical Information and Computer Sciences, v. 44, n. 4, p. 1402–1411, jul. 2004.

HINTON, G. E. Learning multiple layers of representation. **Trends in Cognitive Sciences**, p. 428-434, 2007.

JASSMANN, T. J.; TASHAKKORI, R.; PARRY, R. M. Leaf classification utilizing a convolutional neural network. In: **SoutheastCon**, p. 1-3, 2015.
JUDD, W. S.; CAMPBELL, C. S.; KELLOGG, E. A.; STEVENS, P. F.; DONOGHUE, M. J. Plant Systematics. A Phylogenetic Approach. **Sinauer Association,** Sunderland, 2008.

KERAS. **Accuracy metrics**. 2020. Disponível em <https://keras.io/api/metrics/accuracy_metrics/> Acesso dia 29 jan. 2021.

KOGAN, F.; KUSSUL, N.; ADAMENKO, T.; SKAKUN, S.; KRAVCHENKO, O.; KRYVOBOK, O.; SHELESTOV, A.; KOLOTII, A.; KUSSUL, O.; LAVRENYUK, A. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models. **Int.J. Appl. Earth Observat. Geoinf**. vol. 23, p. 192–203, 2013.

KOLOTII, A.; KUSSUL, N.; SHELESTOV, A.; SKAKUN, S.; YAILYMOV, B.; BASARAB, R.; LAVRENIUK, M.; OLIINYK, T.; OSTAPENKO, V. Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine. In: **The 36th Symposium on Remote Sensing of Environment**, p. 35, 2015.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. **In NIPS**, p. 1106-1114, 2012.

KUMAR, A. **Keras – Categorical Cross Entropy Loss Function. Data Analytics.** 2020. Disponível em <https://vitalflux.com/keras-categorical-cross-entropy-loss-function/> Acesso dia 28 jan de 2021.

LE, Q.; RANZATO, M.; MONGA, R.; DEVIN, M.; CHEN, K.; CORRADO, G.; DEAN, J.; NG, A. Building high-level features using large scale unsupervised learning. In: **International Conference in Machine Learning**, 2012.

LORENA, A. C. **Investigação de estratégias para a geração de máquinas de vetores suporte multiclasses**. 2006. 203 f. Tese (Doutorado em Ciência da Computação e Matemática Computacional) — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2006.

MARENGONI, M.; STRINGHINI, S. Tutorial: Introdução à Visão Computacional usando OpenCV. **Revista de Informática Teórica e Aplicada**, v. 16, n. 1, p. 125–160, 2010.

MARTINS, A. P. M.; DEBASTIANI, A. B.; PELISSARI, A. L.; MACHADO, S. A.; SANQUETTA, C. R. Estimativa do Afilamento do Fuste de Araucária Utilizando Técnicas de Inteligência Artificial. **Floresta e Ambiente**, v. 24, p.363-370, 2017.

MATPLOTLIB. **Matplotlib: Visualization with Python**. Documentation. Disponível em < https://matplotlib.org/stable/index.html>. Acesso em 20 set. 2018.

MENDOZA, M.; VELASTÍN, S. (EDS.). Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. Cham: **Springer International Publishing**, v. 10657, 2018

MULA, H. C. A. **AVALIAÇÃO DE DIFERENTES SUBSTRATOS NA PRODUÇÃO DE MUDAS DE *Sebastiania commersoniana* (Baillon) L. B. Smith & R. J. Downs.** UNIVERSIDADE FEDERAL DO PARANÁ**,** p. 117, 2011.

NAIR, V.; HINTON, G. E. **Rectified Linear Units Improve Restricted Boltzmann Machines**, p. 8, 2019.

NURHAYATI et al. A study of hold-out and k-fold cross validation for accuracy of groundwater modeling in tidal lowland reclamation using extreme learning machine. In. **2014 Ieee 2nd International Conference On Technology, Informatics, Management, Engineering & Environment (Time-E).** Bandung, Indonesia: IEEE, ago. 2014. Disponível em: <http://ieeexplore.ieee.org/document/7011623/>. Acesso em 11 jan. 2021.

OSEI, K. B. TURBULENCE ON THE GLOBAL ECONOMY INFLUENCED BY ARTIFICIAL INTELLIGENCE AND FOREIGN POLICY INEFFICEINCIES. **Journal of Liberty and International Affairs, Institute for Research and European Studies - Bitola**, v. 2, p. 113–122, 2020.

PRINCE, S. J. D. **Computer vision: models, learning and inference**. Cambridge University Press, p. 13, 2012.

QUINE, W. Quiddities: An Intermittently Philosophical Dictionary. **Cambridge: Harvard University Press**, p. 4. 1987.

SCHMIDT, D.; SILVA, D. B.; COSTA, C. A.; RIGHI, R. R. Um Modelo de Predição de Mortalidade em Unidades de Terapia Intensiva Baseado em Deep Learning. In: **Anais do XVIII simpósio Brasileiro de computação aplicada à saúde (SBCAS)**, 2020.

SILVA A. P. M.; MARQUES, H. R.; SANTOS, T. V. M. N.; TEIXEIRA, A. M. C.; LUCIANO, M. S. F.; SAMBUICHI, R. H. R. Diagnóstico da Produção de Mudas Florestais Nativas no Brasil. Brasília: **IPEA Relatório**. p. 58. 2015.

SLADOJEVIC, S.; ARSENOVIC, M.; ANDERLA, A.; CULIBRK, D.; STEFANOVIC, D. Deep neural networks based recognition of plant diseases by leaf image classification. **Computational Intelligence and Neuroscience**, p. 12, 2016.

STACKEXCHANGE. **Cross Validated.** What is the difference between a neural network and a deep neural network, and why do the deep ones work better?**.** Disponível em <https://stats.stackexchange.com/questions/182734/what-is-the-difference-between-a-neural-network-and-a-deep-neural-network-and-w>. Acesso em 22 dez. 2019.

SOTHE, C.; ALMEIDA, C. M.; SCHIMALSKI, M. B.; LA ROSA, L. E. C.; CASTRO, J. D. B.; FEITOSA, R. Q.; DALPONTE, M.; LIMA, C. L.; LIESENBERG, V.; MIYOSHI, G. T.; TOMMASELLI, A. M. G. Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data. **GIScience & Remote Sensing**, v. 57, n. 3, p. 1–26, 2020.

SOUSA, A. L.; SALAME, M. F. A.; NASCIMENTO FILHO, F. J.; ATROCH, A.L. Redes neurais convolucionais aplicadas ao processo de classificação de cultivares de guaranazeiros. **Embrapa Amazônia Ocidental**, p. 12, 2017.

SOUSA, A. L.; SALAME, M. F. A. Indução de hipóteses a partir de experiência passada com imagens para identificação de espécimes vegetais. **Anais do XLVI Seminário Integrado de Software e Hardware**, v. 46, p. 125–136, 2019.

SOUZA, R. E. **Aplicação de ferramentas da qualidade na expedição de mudas de espécies nativas da Mata Atlântica**. 2017. 30 f. Trabalho de conclusão de curso-Universidade Federal Rural do Rio de Janeiro.

SOUZA, V. C.; LORENZI, H. **Chave de Identificação: para as principais famílias de Angiospermas e Gimnospermas nativas e cultivadas do Brasil.** 3 ed. São Paulo: Instituto Plantarum de Estudos da Flora, 2014.

SUÁREZ, A.; ALVAREZ-FEIJOO, M. A.; GONZÁLEZ, R. F.; ARCE, E. Teaching optimization of manufacturing problems via code components of a Jupyter Notebook. **Computer Applications in Engineering Education**, v. 26, n. 5, p. 1102–1110, 2018.

SUN, J. et al. ExCAPE-DB: an integrated large scale dataset facilitating Big Data analysis in chemogenomics. **Journal of Cheminformatics**, v. 9, n. 17, 2017.

SZELISKI, R. Computer vision: algorithms and applications. **Springer Science & Business Media**. p. 18 2010.

THUNG, F.; BISSYANDÉ, T. F.; LO, D.; JIANG, L. Network structure of social coding in GitHub. **Proceedings of the European Conference on Software Maintenance and Reengineering, CSMR**, p. 323–326, 2013.

TURLAND, N. J., et al. International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code). In: **Nineteenth International Botanical Congress Shenzhen**. 2018.

VANDERPLAS, J. **Python data science handbook.** Essential Tools for working with data. *Handbook.* 1 ed. Editora: O'Reilly Media. p. 548, 2017.

VARGAS, A. C. G.; CARVALHO, A. M. P.; VASCONCELOS, C. N. **Um Estudo sobre Redes Neurais Convolucionais e sua Aplicação em Deteccão de Pedestres**, p. 4, 2016.

VIDAL, C. Y.; RODRIGUES, R. R. **Restauração da diversidade: os viveiros do estado de São Paulo**. Universidade de São Paulo. Escola Superior de Agronomia Luiz de Queiroz, 2019.

WALKER, C.; ARAÚJO, M. M.; MACIEL, C. G.; MARCUZZO, S. B. Viveiro florestal: Evolução tecnológica e legalização. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, p. 7, 2011.

WANG, P. On Defining Artificial Intelligence. **Journal of Artificial General Intelligence**, v. 10, n. 2, p. 1–37. 2019.

WENDLING, I.; GATTO, A.; PAIVA, H. N. D.; GONÇALVES, W. Planejamento e instalação de viveiros. Viçosa: **Aprenda fácil**. Coleção Jardinagem e Paisagismo, v. 1, p. 106, 2001.