FACULDADE DE TECNOLOGIA, CIÊNCIAS E EDUCAÇÃO

GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Sistema de classificação de imagens de flores através do uso de Rede Neural Convolucional

Rodrigo Coelho da Silva Dias
Adinovam Henriques de Macedo Pimenta (Orientador)

RESUMO

A classificação botânica, embora seja uma atividade antiga, não dispõe hoje de ferramentas avançadas para tal tarefa, onerando o processo de criação de taxonomias e comprometendo a acurácia da classificação realizada. Para suprir esta lacuna, foi desenvolvido neste trabalho um protótipo chamado RE25 que tem por objetivo fazer a classificação de flores, podendo alcançar não apenas o meio científico, mas também, ser utilizado em ambiente escolar e acadêmico, bem como para os entusiastas da flora. O protótipo desenvolvido utiliza técnicas de processamento de imagens e Redes Neurais Convolucionais para fazer a classificação das imagens. O RE25 alcançou uma acurácia de 88,04% nos testes, superando um dos aplicativos mais populares para a mesma finalidade, fortalecendo a hipótese de que futuramente o banco de imagens poderá ser expandido por meio de redes cooperativas de usuários.

Palavras-chave: *deep learning,* classificação, rede neural, taxonomia

ABSTRACT

Although botanical classification is an old activity, it does not have advanced tools for such task, burdening the taxonomy creation process and compromising the accuracy of the classification. To fill this gap, a prototype called RE25 was developed in this work that aims to classify flowers and can reach not only the scientific environment, but also be used in school and academic environment, as well as for flora enthusiasts. The developed prototype uses image processing techniques and Convolutional Neural Networks to classify the images. The RE25 achieved 88.04% accuracy in testing, surpassing one of the most popular applications for the same purpose, strengthening the hypothesis that the image bank could eventually be expanded through cooperative user networks.

Keywords: *deep learning, classification, neural network, taxonomy*

Introdução

Dada a importância de objetos encontrados na fauna e na flora e em sua biodiversidade, pesquisas foram realizadas para a identificação e reconhecimento destes objetos como forma de classificá-los, segundo Nicolau (2017) torna-se clara a razão e importância do processo classificativo das plantas e seres vivos:

> A classificação biológica, ou o modo como organizamos os organismos vivos, facilita a nossa compreensão da enorme diversidade biológica e das relações evolutivas entre espécies. Esta compreensão, para além do seu interesse biológico intrínseco, pode ser também visto através de um prisma utilitário, de interligação com outras ciências, como a medicina ou outras.

Além da classificação, estas pesquisas proporcionam a divulgação dos novos objetos que veem sendo classificados e inseridos em algum grupo botânico já conhecido ou em um novo grupo botânico recém-criado, proporcionando a disseminação e o acesso a estes conhecimentos para a sociedade de forma a acrescentar e facilitar o uso na rotina ou estudos mais aprofundados na área, de forma que o cunho social e acadêmico sejam alcançados.

Em meio aos avanços da tecnologia na sociedade, já é quase que inato o uso de alguma ferramenta digital para solucionar dúvidas e buscar alguma

resposta imediata, como por exemplo, uma situação hipotética em que um indivíduo, ao deparar-se com uma flor, sente interesse em saber a identificação da mesma. A partir deste ponto, fomentou-se neste trabalho a ideia de incorporar os recursos tecnológicos já existentes nas áreas da Ciência da Computação, aos materiais de pesquisa e estudos da botânica, também já existentes.

Considerando as contribuições para este processo de classificação, destaca-se o trabalho do sueco Carl von Linné (1707-1778) que aprimorou o processo de classificação botânica aplicando um novo conceito às nomenclaturas, melhoria essa que segue até os dias atuais sem grandes alternâncias (PRESTES et al., 2009).

O objetivo de Lineu consistia em facilitar a identificação das plantas sem que ocorresse confusão e esquecimento, dessa maneira instaurou a nomenclatura binomial, que mantinha um padrão para dar nome às novas plantas que fossem categorizadas. O êxito dessa sua ideia foi alcançado e formalizado no século XVIII.

> A taxonomia é uma das áreas de pesquisa da biologia que tem como objetivo identificar, descrever e classificar a diversidade de seres vivos. A taxonomia vegetal é a ciência responsável pela correta organização, classificação e nomenclatura das espécies de plantas [...] (MORA et al., 2011).

Partindo do ponto que classificação botânica ainda é feita de maneira rudimentar, vê-se como uma oportunidade promissora a automatização deste processo por meio do uso de técnicas de processamento de imagens e Redes Neurais Artificiais (RNA).

Objetivo Geral:

Ponderando sobre estes aspectos vinculados entre a taxonomia botânica e as RNA, este trabalho tem como objetivo geral construir um protótipo para a identificação automática de flores, aperfeiçoando o processamento de classificação das imagens.

Objetivos Específicos:

- Identificar as técnicas mais promissoras para a extração de dados a partir das imagens de entrada;
- Identificar as técnicas mais promissoras para a classificação das flores contidas nas imagens;

- Definir métricas de comparação de desempenho para o escopo escolhido;
- Comparar o desempenho entre o sistema proposto e outras abordagens.

1. Classificação botânica

A sistemática botânica dedica-se a estimar e representar a história evolutiva, promovendo agrupamentos das espécies de acordo com o grau de parentesco. No entanto, nem sempre foi assim. Os sistemas de classificação de seres vivos são muito anteriores às teorias evolutivas do século XIX. Ao invés de serem fundamentados na dimensão temporal, eles foram construídos a partir das semelhanças e diferenças aparentes entre os organismos (NICOLAU, 2017).

Durante o século XVIII, começam a ser reconhecidos alguns estudiosos e suas contribuições, com destaque para o sueco Carl von Linné (1707 - 1778), visto que seus estudos voltavam-se para uma nova forma de categorização, que aspirava por organizar normativamente as plantas atribuindo os nomes científicos das espécies. Esta forma descritiva de nomear era composta pelo nome da espécie e por sua característica específica, os quais deveriam ser de origem latina, estabelecendo uma composição que é aceita até hoje.

Esta mudança na nomeação das plantas tinha o intuito de facilitar a identificação e o conhecimento global das espécies entre especialistas e impossibilitar a duplicação de registros de mesmas espécies devido ao nome diferente que era atribuído, como é possível confirmar em Costa et al. (2011, p. 13):

> Uma dessas soluções foi estabelecer como padrão a nomenclatura binomial – até então as espécies eram descritas por expressões polinomiais. Outra foi adotar o latim como língua padrão universal para os textos onde as espécies seriam descritas – os autores costumavam usar seu próprio idioma, o que só dificultava a comunicação e o entendimento entre autores de línguas diferentes.

Atualmente os processos de classificação seguem em desenvolvimento buscando alcançar melhorias nos métodos usados para tal tarefa. Porém, um obstáculo no aprimoramento desses processos de classificação é o baixo número de profissionais na área da taxonomia, fazendo com que ainda hoje encontremos processos manuais, conforme estudos recentes feitos por Massucatto (2018, p. 4):

> "[...] o número de profissionais taxonomistas é escasso e a imensa variedade de espécies existentes no mundo acaba tornando difícil o trabalho a ser realizado. Um taxonomista pode chegar a conhecer algumas centenas de espécimes, um número relativamente baixo se comparado ao número de espécies existentes."

Atualmente, um recurso indispensável para as pesquisas na classificação das folhas e flores é o herbário, uma coleção de plantas dessecadas que servem como orientação e base para estes estudos, onde é possível alterar, retirar e acrescentar informações à respeito da planta que está sendo analisada. Entretanto, como mencionado, tudo é feito manualmente, o que acarreta em uma demanda demorada, tardando o desenvolvimento conforme Massucatto (2018, p. 9) afirma ao citar Vicente et al (1999):

> Apesar da existência dos herbários, todo o trabalho de identificação e caracterização continua sendo feito manualmente (VICENTE et al., 1999), sendo um trabalho difícil, o qual demanda muito tempo para classificação da folha em vista ao imenso número de espécimes.

O processo de identificação em si faz uso das características morfológicas e anatômicas dos aparelhos reprodutivos e vegetativos a partir de observação e comparação entre as espécies, sendo o resultado diretamente dependente do conhecimento acadêmico e profissional do especialista.

Outro recurso usado neste processo de classificação botânica recebe o nome de Cladística, que é uma sistemática filogenética (taxonômica), que categoriza as plantas através das semelhanças dentro da cadeia evolutiva, ou seja, as espécies, através de seu antepassado mais próximo. O resultado obtido através dessa análise evolutiva é chamado de *cladograma*, onde é criado um parâmetro gráfico para hipotetizar as possíveis origens e ramificações das espécimes que posteriormente serão submetidas a testes para confirmação. Assim, conforme Santos (2008, p. 2) "A biogeografia cladística enfatiza a procura por padrões de distribuição congruentes na história filogenética dos táxons estudados" (Rosen, 1978; Nelson & Platnick, 1981).

Tendo em vista esses dois métodos de classificação mais recorrentes na atualidade, é perceptível a necessidade de aplicar recursos tecnológicos que automatize estes processos de classificação das plantas, como forma de

acelerar e proporcionar um cenário favorável para que haja um salto no desenvolvimento nesta área.

Por conta destes métodos serem realizados de forma manual, onde a análise feita depende apenas do conhecimento humano do especialista envolvido, acaba por limitar o alcance em relação ao tempo de pesquisa e muitas vezes interferir no resultado. Desta maneira, baseado em todos estes dados históricos da evolução da taxonomia, o protótipo deste trabalho apresentará uma alternativa para a da inclusão da automatização através do processamento e classificação de imagens fazendo o uso de Redes Neurais Convolucionais que identificará uma espécie de flor a partir de uma amostra de imagem, poupando tempo de busca na análise manual e permitindo direcionar os esforços para o crescimento e avanço na classificação das espécies que ainda não foram reconhecidas.

2. Classificação de imagens

O processo de classificação de imagens no escopo da computação acontece por meio de técnicas variadas que quando combinadas trazem resultados específicos conforme o objetivo que o cientista deseja alcançar.

Dentre as diversas técnicas existentes com este propósito podemos destacar: *k-Nearest Neighbours* (MOREIRA, 2016), *Suport Vector Machine* (GAROFALO et al., 2015), algoritmo de *Otsu* artigo desenvolvido por Torok (2016) e Redes Neurais Artificiais fundamentado em Fonseca (2017).

2.1 K-*Nearest Neighbours* (KNN):

O KNN é um algoritmo de classificação baseado no distanciamento entre o objeto a ser analisado e aqueles que estão no grupo de treinamento, gerando uma busca ordenada entre os elementos mais próximos àquele que foi inserido para análise.

Este algoritmo usa objetos armazenados por um conjunto de características (contínua, binária ou categórica), em que cada objeto $\boldsymbol{x}$ de um conjunto de dados possui d características tal que $\boldsymbol{x}=\{x_1, x_2, x_3, ..., x_d\}$. O reconhecimento do objeto é feito a partir de características descritivas comparando o vetor (objeto) $\boldsymbol{x}$ com as características de uma classe $\boldsymbol{c}=\{c_1, c_2, c_3, ..., c_d\}$.

O processo de classificação com uso da KNN se dá a partir de três etapas:

1. É feita a comparação entre o objeto de teste $x_t = \{x_{t1}, x_{t2}, \ldots, x_{td}\}$ com todos os *N* objetos ***x*** pertencentes a um conjunto de treinamento por meio da distância Euclidiana;
2. A organização das distâncias é feita em ordem crescente de valores: distancia(x_t, x_i) < distancia(x_t, x_j), sendo *i* diferente de *j*;
3. Quando *k=1* a classe ***c*** do objeto contido no conjunto de dados de menor distância é concedido ao objeto de teste; mas se *k>1* a classe maior dos vizinhos mais próximos será atribuída ao x_t.

Embora seja feita uma análise entre o objeto de teste e o conjunto de treinamento, ela não legitima um aprendizado, sendo preciso realizar uma comparação entre o conjunto de dados sempre que um novo objeto é inserido para classificação sem existir um modelo e, por isso, recebe o nome de *Lazzy Learning* (Aprendizado Preguiçoso) (WU; KUMAR, 2019; MAIMON; ROKACH, 2010).

Desta maneira o uso da KNN acarreta a desvantagem de necessitar de um grande período de tempo para o processo classificatório. Para protótipo proposto neste trabalho, usar o KNN não satisfará o objetivo de agilizar o método atual de classificar flores, onde também o objetivo deste protótipo está direcionado ao aprendizado e não à métodos de comparação.

2.2 Suport Vector Machine (SVM)

Esta técnica tem suas bases enraizadas na estatística com foco na generalização do aprendizado. Esta técnica pertence a uma categoria de aprendizado de máquina usada para classificar informações que podem ser lineares ou não-lineares, sendo um dos métodos de aprendizado usados em problemas de reconhecimento de padrões.

Para este feito, é necessária a elaboração de um hiperplano que será o responsável pela divisão dos dados obtidos em classes diferentes, esta ideia baseia-se na busca de um hiperplano ótimo, através de conceitos matemáticos.

A equação que rege o hiperplano definida por ▫(▫) = ▫ ▪ ▫ + b.

Dado que *w* é um vetor de peso e *b* o bias que se adapta conforme o fornecimento de dados durante o treinamento, chamado aprendizado

supervisionado, que é representado por *x*, esse processo se repete até que a ação de separação dos dados em classes distintas seja possível de ser feita pelo hiperplano.

A resolução se restringe a uma classificação linear binária, ou seja, apenas duas possibilidades. Desta maneira com esta técnica o conjunto de dados usado será classificado para, então, quando o usuário utilizar, poder reconhecer se aquela informação é ou não algum dos elementos pré categorizados.

2.3 Algoritmo de Otsu

Proposto por Nobuyuki Otsu (Otsu, 1975), esta técnica consiste na limiarização de imagens, ou seja, classifica as imagens transformando-as em escalas de cinza, através de um valor limite (*threshold*), separando os elementos da imagem entre fundo e frente em dois grupos (*clusters),* a aplicação desta técnica apresenta melhor funcionamento com imagens que usam histogramas.

A ideia proposta consiste em encontrar todos os valores possíveis para o *threshold* na imagem utilizada, até encontrar o menor índice de variação da imagem garantindo o melhor limite, dissociando a frente e o fundo atribuindo uma cor para cada classe (preto e branco).

Para encontrar o melhor posicionamento de um *threshold* de **t** pode ser calculado com a seguinte fórmula:

$$\sigma_w^2 = q_b \sigma_b^2 + q_f \sigma_f^2 \quad (1)$$

onde *w* é o peso de cada classe, que condiz com a probabilidade que um pixel tem de pertencer à classe *b* (*fundo)* ou *f* (*frente)*.

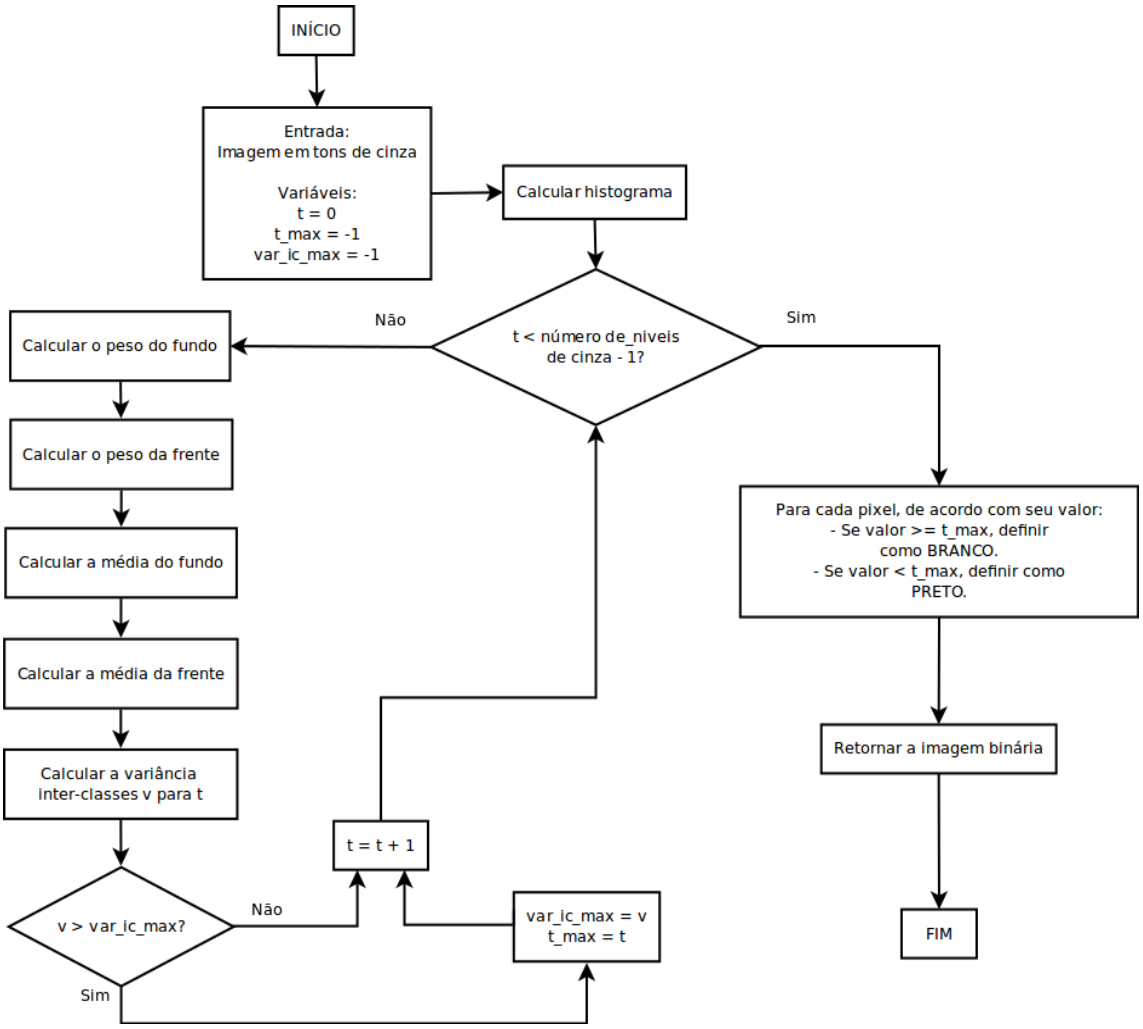
Sendo este o cálculo que será sempre realizado para encontrar todos os *thresholds* possíveis, sendo que aquele que apresentar menor variação é o que será selecionado para a binarização da imagem.

A Figura 1 exibe o fluxograma do funcionamento do algoritmo de Otsu.

- O primeiro passo calcula o histograma da imagem em tons de cinza de entrada;
- Para cada *threshold* serão calculados pesos e medidas para as classes de frente e fundo da imagem, onde cada valor é usado no cálculo de variância;

- Após todo o cálculo, o *threshold* escolhido será aquele que atingiu o maior valor para a variância;
- Se o valor for igual ou superior ao *threshold*, a cor do pixel torna-se branca, mas se o tom for inferior à cor será preta conforme Ilustra o exemplo da Figura 2.

Figura 1- Fluxograma do Algoritmo de Otsu.



Fonte: TOROK (2016)

Com base no exposto sobre o método de Otsu e refletindo na proposta deste trabalho, sua aplicação limitaria a funcionalidade do nosso protótipo devido à monocromatização das imagens, afetando o reconhecimento das imagens de flores, tendo em vista que para o reconhecimento idealizado no protótipo o padrão de cores é um elemento importante para aprendizagem do sistema.

Figura 2 - Imagem em tons de cinza a esquerda e seu resultado após limiarização.

Fonte: TOROK (2016)

2.4 Redes Neurais Convolucionais

Mais recentemente, as Redes Neurais Convolucionais (CNN, do inglês *Convolutional Neural Network*) veem apresentando resultados promissores em diversas áreas.

As CNN são um tipo de Rede Neural Artificial (RNA) específica para o processo de classificação de imagens, sendo uma especificidade das redes neurais que através dos dados de entrada extrai de forma arquitetada características de uma imagem de forma que estas informações retiradas sejam menos numerosas, porém precisas para o resultado e aplicações computacionais, como é possível constatar através de Fonseca (2017):

> Cada neurônio realiza um produto escalar em suas entradas e opcionalmente também realiza uma função não linear antes de gerar uma saída. Mais precisamente, as CNNs são modelos de Deep Neural Networks, ou DNNs, que funcionam melhor para o processamento de imagens pois as convoluções realizadas por elas representam as informações que um pixel contém relacionado com seus vizinhos.

Nesta técnica é feita a classificação de imagens a partir de camadas que têm por objetivo retirar as características relevantes e, assim, seguir afunilando até atingir os objetivos desejados.

Esse processo ocorre em algumas etapas. A primeira etapa é através da *segmentação*, que consiste na divisão da imagem em pequenas partes chamadas de regiões (conjunto de pixels) ou objetos, que podem ser feitas a partir da descontinuidade (separação da imagem através do contorno). Na segunda etapa será obtido um conjunto de informações referentes à essa imagem (pixels) baseando-se na similaridade dos pixels seguindo também algum critério ou padrão. A forma exata na qual a segmentação acontece é determinante para o êxito ou a falha da análise que está sendo feita, a imagem passa por filtros existentes na entrada das camadas convolucionais. Esse filtro é responsável pela separação e mapeamento de recursos da imagem como, por exemplo, bordas ou nitidez.

Este processo acontece na CNN resultando na classificação da imagem através da extração de dados relevantes, conforme Fonseca (2017) explica que foram usados para os testes de comparação e identificação da classe na qual pertence o objeto.

> "Isso significa que as CNNs conseguem extrair de forma inteligente atributos de uma determinada imagem, diferentemente das RNA convencionais. Portanto, as CNNs extraem menos atributos, e ao mesmo tempo, atributos mais importantes para as tarefas de visão computacional."

Diante das técnicas citadas e levando em consideração o objetivo do trabalho, que consiste na identificação dos padrões de flores para classificá-las, a técnica que melhor se aplica e garante resultados satisfatórios é a CNN, pois está diretamente relacionada cm classificação de imagens e também ao número de categorias utilizadas neste protótipo. Esta técnica apresenta menor perda de informações, garantindo que tudo seja usado em prol de melhores resultados.

3. Aprendizado profundo

Em meio a tantos avanços já alcançados através de inúmeros estudos, em dado momento pesquisadores interessados pela área da Inteligência Artificial (IA) passaram a direcionar seus estudos em busca de respostas para uma pergunta: *seriam as máquinas capazes de aprender como o homem aprende?*

As funções computacionais, tais quais como conhecemos, são ações que acontecem devido a construção programada de maneira estática. Por exemplo,

se um programa foi desenvolvido para reconhecer apenas flores do tipo *margaridas* e o usuário inserir qualquer outra espécie de flor, o algoritmo retornará que o dado fornecido não existe na sua base de conhecimento. Por sua vez, através de técnicas de IA, o aprendizado do sistema continua sendo feita. No entanto, o grande diferencial se dá pelo fato de que o programa será desenvolvido para adaptar-se e aprender as informações que ainda não possuem na base de conhecimento sem que precise acontecer alguma intervenção ou alteração humana no código.

Esta forma de execução em que a máquina aprende de maneira automática leva a definição de *machine learning*, ou seja, a máquina passa de fato a aprender. As ideias que levaram os estudiosos a desenvolver este tipo de ação teve como motivação a intenção de simular o cérebro humano, usando para isso a ideia de Rede Neural Artificial.

Partindo deste ponto, estudos trazem o mais recente avanço neste contexto de Rede Neural Artificial, o *Deep Learning*, representando uma forma de aprendizagem mais profunda que a *machine learning*, em que a compreensão de dados acontece de forma mais simplificada e natural. De acordo com Deep Learning Book (2019), seu funcionamento baseia-se no treinamento de dados para apurar as informações necessárias para o armazenamento do aprender, baseando-se nas taxas de erros e acertos, aprimorando a capacidade da máquina em classificar, reconhecer, detectar e descrever conforme o contexto no qual seja aplicado.

Com o uso do *Deep Learning*, o reconhecimento de imagens acontece através de um mapeamento autônomo de características semelhantes. Por exemplo, se o conjunto de dados usado contém imagens de flores, o algoritmo aprenderá o padrão dos objetos contidos nestas imagens, ou seja, perceberá que as imagens apresentam características comuns como as pétalas e, em seguida, começará a extrair os padrões de cada uma das espécies que estão alimentando o conjunto de dados.

Assim, a aplicação do *Deep Learning* facilita e torna mais ágil o processamento e a classificação de imagens, aumentando o retorno positivo em relação ao desempenho, garantindo maior eficiência nos resultados em que quanto maior a precisão da validação maior a taxa de acertos e,

consequentemente, maior credibilidade e confiança serão acrescidos ao programa utilizado.

4. Método proposto

O modelo implementado neste trabalho envolve as seguintes etapas: Aquisição da imagem; Pré-processamento; Segmentação; Extração de informação, Classificação e Informação para o usuário. A Figura 3 ilustra estas etapas.

Figura 3 - Fluxograma das etapas

Fonte: Elaborado pelo próprio autor.

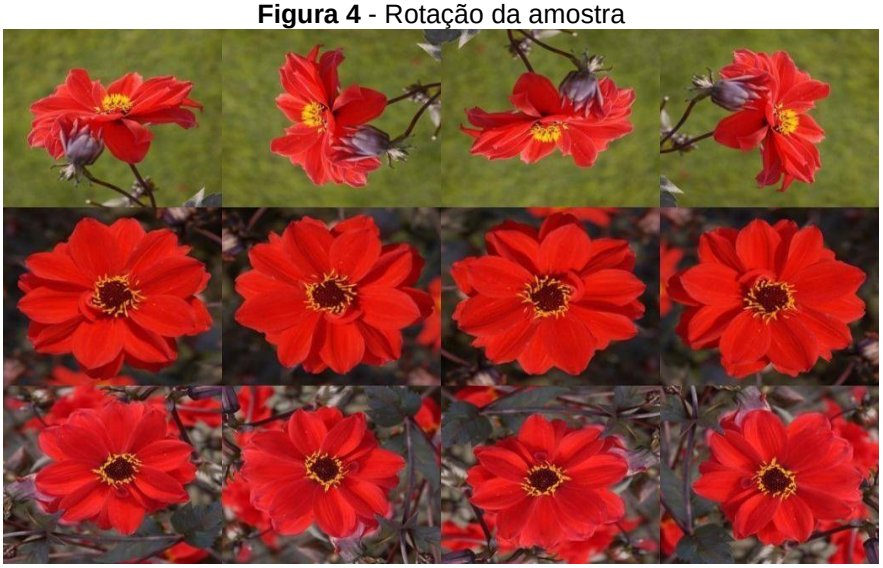
Inicialmente, várias imagens de flores são inseridas no conjunto de dados (aquisição de imagem). A partir destas imagens, durante o pré-processamento, novos lotes de imagens são gerados a partir das imagens que já existem, fazendo transformações como: distorções, zoom, rotação. Estas transformações são feitas utilizando o framework Keras[1] e a função *ImageDataGenerator*.

Optou-se por utilizar esta técnica porque o conjunto de dados utilizado neste projeto apresenta uma quantidade relativamente limitada de amostras, acarretando em um aprendizado menos amplo dificultando um melhor

[1] Disponível em keras.io

desempenho. A Figura 4 ilustra rotação feita pelo *ImageDataGenerator* gerando, assim, novas amostras.

Figura 4 - Rotação da amostra

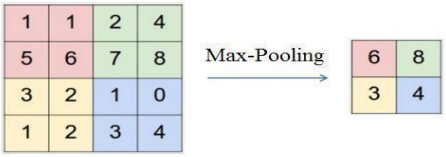


Fonte: Elaborado pelo próprio autor.

Em seguida é criada a CNN. Neste trabalho a CNN criada possui quatro camadas convolucionais, sendo aplicada nas duas primeiras 32 filtros e nas duas últimas utilizam-se 64 filtros. Esses filtros são aplicados por toda a imagem e são usados para gerar mapa de características e também é responsável pela nitidez e reconhecimento de borda, fazendo apenas pequenas alterações na matriz da imagem. A quantidade de filtros foi escolhida empiricamente.

Após a leitura dos filtros é aplicada a função de ativação, que é uma função que decide se ativa ou desativa os neurônios conforme a relevância da informação para o aprendizado da rede. A função escolhida foi a *ReLu,* que é uma função não linear responsável pela ativação de várias camadas de neurônios simultaneamente. Em seguida, aplicou-se a função *max pooling,* que é encarregada por reduzir cada *feature map* em segmentos de dimensão 2x2, ou seja, essa função é responsável por gerar uma amostra em baixa resolução para a camada convolucional fazendo uso da ativação máxima dos filtros. Na Figura 5 temos um exemplo da função *max pooling*.

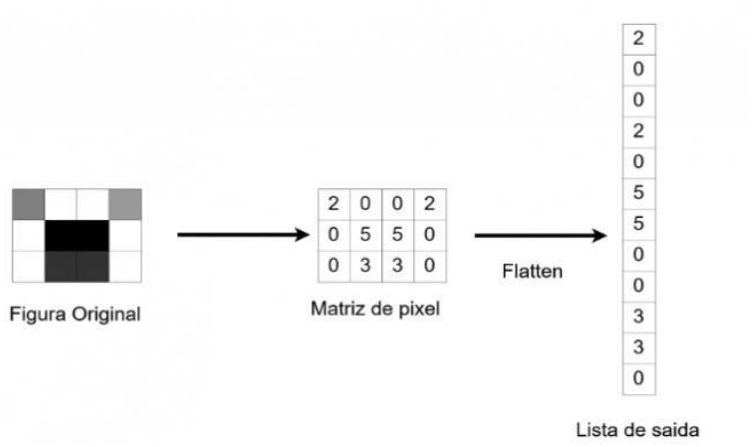
Figura 5 - Max Pooling 2x2



Fonte: PENHA (2018)

Em seguida é aplicado o *Dropout*, que é uma função usada nos neurônios ocultos para evitar *overfitting*, ou seja, evitar que a rede fique enviesada com os dados de treinamento. Assim, quando finalizada às 4 convoluções, concretiza-se *o flatten,* função incumbida de transformar os vetores 2D em vetores 1D, que nada mais é do que a transformação em representação linear como foi ilustrada na Figura 6:

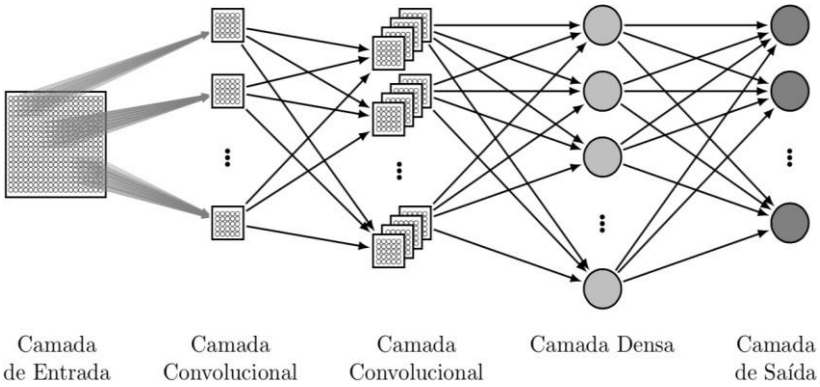
Figura 6 - Convertendo vetor 2D em 1D



Fonte: https://letscode-academy.com/blog/redes-convolucionais-reconhecer-objetos-em-imagens/

Em seguida esses vetores são inseridos na camada *dense* de entrada, contendo 512 neurônios e, então, aplica-se novamente a função de ativação *ReLu*. Em um próximo passo, novamente usa-se o *Dropout* e uma camada de saída com 102 neurônios, que representam cada uma das 102 classes de flores, e assim é aplicado uma função softmax, que transforma as saídas em 0 e 1 para auxiliar na definição das classes. A Figura 7 ilustra as camadas da CNN.

Figura 7 - Camadas da CNN



Fonte: http://rafaelsakurai.github.io/cnn-mapreduce/

Na próxima etapa, temos a compilação utilizando o otimizador *Adadelta*, que tem por objetivo manter a rede aprendendo constantemente, incluindo as atualizações geradas pela mesma. Com o uso desta técnica não é necessário definir uma taxa de aprendizagem inicial. Nessa função não são armazenadas todas as partes do gradiente, mas sim uma média exponencial conforme mostra a Equação 2, em que *Y* é um parâmetro de ajuste.

$$\square_{\square}(\square^{2}_{\square}, \square) = \square_{\square}(\square^{2}_{\square-1}, \square) + (1 - \square)\square^{2}_{\square} \quad (2)$$

Também utilizou-se a função *loss* com a função *categorical* que é encarregada pela perda e categorização, onde apenas uma categoria é aplicada para cada classe. Outra função aplicada foi a *metrics accuracy*, que diz respeito às taxas de acerto, ou seja, ela é responsável por medir a quantidade de dados classificadas corretamente.

Em seguida, foi feita a execução do treinamento da rede com 6.550 passos em 8 *épocas*, sendo que cada *época* representa cada passagem completa do algoritmo por todo conjunto de dados.

Na última etapa, finalizando o processo de aprendizagem, é gerado um gráfico contendo o resultado do treinamento e teste realizados, que em seguida salva toda a rede desenvolvida.

O método apresentado passa por um treinamento, sendo que o mesmo lê constantemente as imagens para aprender a identificar os padrões de cada objeto, que em nosso caso são flores, e com isso se tornar capaz de classificá-las corretamente.

Durante esse processo a rede é otimizada arrumando os pesos até que se consiga atingir uma taxa onde o erro seja mínimo, e é neste momento que o método caminha para aprendizagem.

5. Experimentos

Em conformidade com a proposta deste trabalho, para o desenvolvimento do projeto foram utilizadas alguns instrumentos essenciais para a boa execução do mesmo, sendo eles: um notebook Asus Intel Core i7 com suporte para a elaboração das funcionalidades que foram desenvolvidas; e o framework Keras, responsável pelo fornecimento das funções que foram utilizadas para a construção do programa, como gráficos e a criação de novas imagens para aumentar a capacidade do conjunto de dados aplicado.

O Keras foi escolhido devido a agilidade que o mesmo apresenta na hora do desenvolvimento e por sua forma de dispor as funções que facilitam a implementação das CNN.

Utilizou-se também o IDE Spyder 3.5, programa próprio para trabalhar com Python, além de trazer a possibilidade de lidar em um único ambiente sem precisar migrar para outras ferramentas. A linguagem de programação Python 3.7 foi escolhida devido ao framework usado para a implementação ser escrito nesta mesma linguagem. Além disso, como o projeto se trata de uma *deep learning*, onde todas as bibliotecas necessárias para o desenvolvimento também foram feitas em Python, tornando esta uma referência entre as linguagens de programação Inteligência Artificial (IA).

Foi usado para o treinamento da CNN um conjunto de dados gratuito disponível através da plataforma Robots[2] disponibilizado pelos pesquisadores de Oxford, Maria Helena Nilsback e Andrew Zisserman, contendo 8.196 imagens de flores, sendo separadas por 102 categorias de espécies.

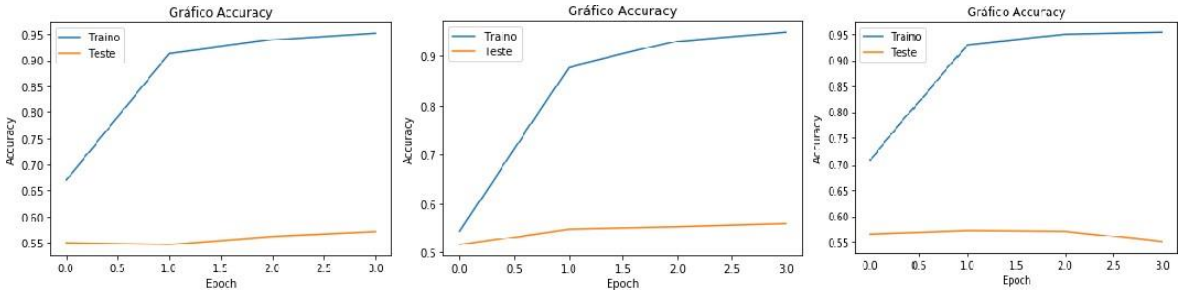
Durante o processo de aprendizado da CNN foi utilizado o método de validação cruzada (ARLOT, 2010) reservando 80% das imagens de cada espécie de flor para treino e 20% das imagens de cada espécie de flor para teste.

Na fase de testes foram utilizadas duas camadas convolucionais. Entretanto, a taxa de acerto se manteve muito baixa, sendo necessário aumentar o número destas para observar se obteria um aumento na acurácia. O aumento da acurácia foi obtido quando foram aplicadas quatro camadas.

Visando ampliar os índices alcançados pelo acurácia, uma nova tentativa foi feita, mas desta vez utilizando seis camadas. Porém, o resultado contrariou as expectativas, pois apresentou-se inferior ao obtido com o uso de 4 camadas. Dessa forma, foi mantido o uso de 4 camadas para o melhor aprendizado, conforme podemos ver nos gráficos da Figura 8. Nesta figura são mostradas as taxas de acurácia com 2, 4 e 6 camadas, respectivamente.

[2] https://www.robots.ox.ac.uk/~vgg/data/flowers/

Figura 8 - Gráfico comparativo entre camadas



Fonte: Elaborado pelo próprio autor.

Na etapa seguinte realizaram-se testes usando funções de ativação a fim de encontrar qual função se adequava melhor para o domínio deste problema. Foram usadas as seguintes funções nos experimentos:

- **Sigmóide**: forma um gráfico em formato de “S” e trabalha de forma binária entre 0 (não ativado) e 1 (ativado);
- **TanH (Tangente Hiperbólica)**: também apresenta um gráfico em formato de “S”. Porém, a diferença se dá pelo fato de variar no intervalo [-1; 1] para decidir a ativação ou não dos neurônios;
- **ReLu**: que é a ativação linear retificada, que não ativa todos os neurônios simultaneamente, aumentando desta forma a eficiência da rede usada. Por este motivo esta última função foi a escolhida para aplicação no protótipo desenvolvido neste trabalho, pois obteve melhor resultado dentre as demais funções.

  A Tabela 1 mostra o desempenho das funções de ativação.

Tabela 1 - Tabela de desempenho das funções de ativação.

Ativadores	Percentual de acerto (%)
Sigmóde	59 %
TanH	62%
ReLu	64%

Fonte: Elaborado pelo próprio autor.

Foram também testadas duas funções otimizadoras, que são responsáveis pela atualização dos pesos e diminuição do erro. A primeira função testada foi a função *Adam*, que trabalha com cálculo de gradiente e atualização

das variáveis. No entanto, usá-la neste protótipo manteve acurácia em baixo nível, o que afastava do objetivo esperado. Em seguida testou-se a função *Adadelta* que, como já foi citado, conseguiu proporcionar um aumento na taxa de acerto. A Tabela 2 resume o desempenho destas funções nos testes.

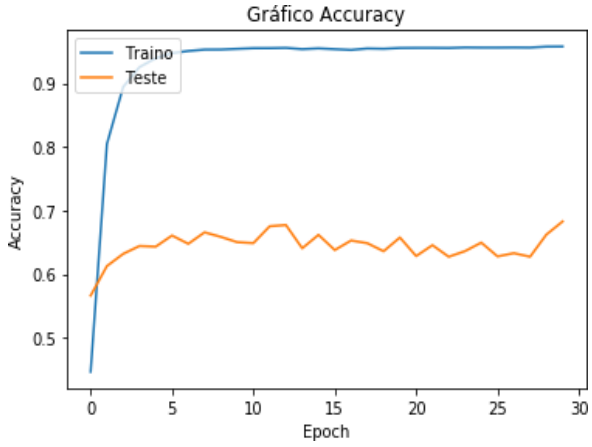
Tabela 2 - Tabela de desempenho das funções de otimização

Otimizador	Percentual de acerto (%)
Adam	63%
Adadelta	68%

Fonte : Elaborado pelo próprio autor.

Com base nas tabelas e gráficos apresentados a respeito dos resultados adquiridos entre treinamentos e testes considerando o número de camadas, tipos de funções de ativação e otimizadores, chegou-se as seguinte escolhas: utilizar 4 camadas de rede convolutiva aplicando o ativador *ReLu* e o otimizador *Adadelta*, podendo ser visto o resultado final desta fase de testes na Figura 9 abaixo.

Figura 9 - Gráfico desempenho após a configuração final da CNN.



Fonte: Elaborado pelo próprio autor.

6. Resultados

Finalizada a etapa de aprendizagem e testes, inicia-se agora o processo de comparação e validação do protótipo em execução com dois aplicativos presentes no mercado, que também trabalham com a identificação e classificação de flores.

Inicialmente, primeiro aplicativo usado foi o PlantNet (Figura 10), disponível tanto para o sistema Android como para o iOS de forma gratuita. Este aplicativo identifica flores através do uso de imagens que são inseridas pelo usuário. Este aplicativo é uma iniciativa educacional apoiada pela Fundação Agropolis[3] desde 2009, contendo catalogadas, até a data dos testes, mais de 1 milhão de amostras de flores.

Figura 10 - Tela Inicial do PlantNet

Fonte: Aplicativo PlantNet

O outro aplicativo usado na validação foi o *Seek* (Figura 11), também com a funcionalidade de identificação de plantas através de imagens fornecidas pelo usuário, e está disponível apenas para o sistema operacional *Android*, também de forma não paga. Este aplicativo foi desenvolvido pela *iNaturalist*[4], uma iniciativa conjunta entre a *California Academy of Sciences* e a *National Geographic Society*. Atualmente conta com mais de 150.000 espécies catalogadas em sistema. Porém, observou-se nos testes que mesmo contendo um banco de imagem tão vasto, o aplicativo apresenta falhas na hora de classificar, classificando de forma errada objetos fora do contexto como se fosse flor.

[3] https://www.capes.gov.br/bolsas-e-auxilios-internacionais/pais/205-franca/9608-programa- capesfundacao-agropolis
[4] https://www.inaturalist.org/

Figura 11 - Inicialização Aplicativo Seek

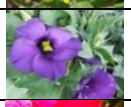
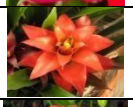
Fonte: Aplicativo Seek

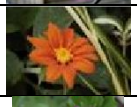
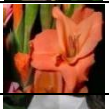
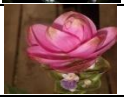
Ambos os aplicativos citados foram utilizados como comparativo de desempenho com o protótipo desenvolvido neste trabalho (que a partir deste ponto vamos identifica-lo como RE25).

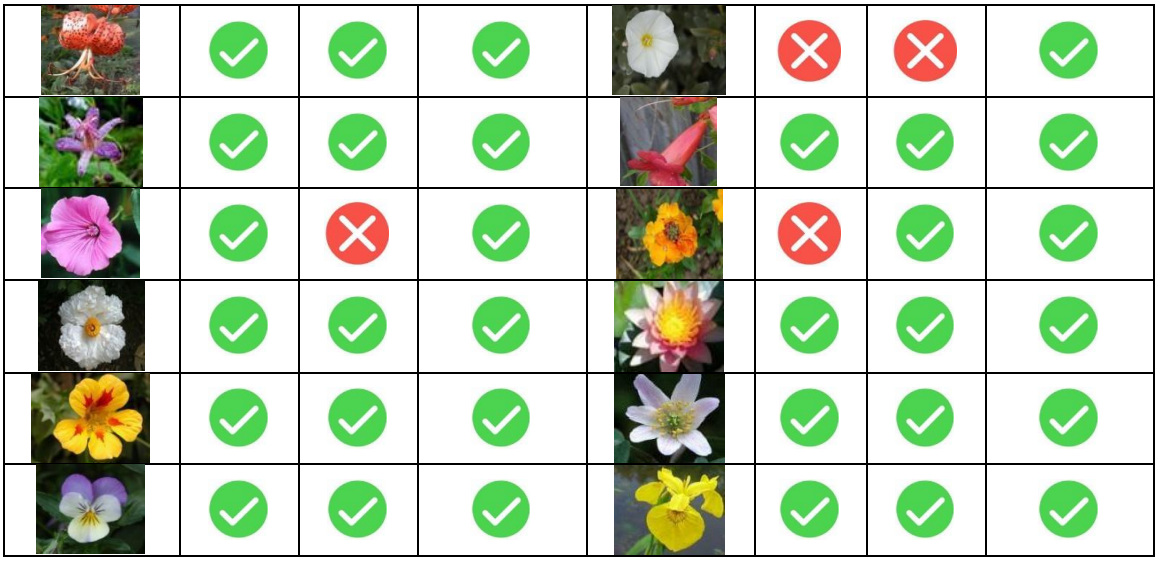
A comparação foi feita colocando os dois aplicativos e o protótipo RE25 em contato com as 102 imagens de espécies distintas de flores, selecionadas de forma aleatória do banco de imagens para não haver comprometimento do resultado. A partir do desempenho de cada um dos três classificadores foi construído um quadro (Quadro 1) de comparação para visualizar o desempenho específico de cada classificador.

Quadro 1 - Tabela de resultado de comparação de métodos

IMAGEM	RE25	SEEK	PLANTNET	IMAGEM	RE25	SEEK	PLANTNET
	✓	✗	✓		✓	✓	✓
	✓	✓	✓		✓	✗	✓
	✓	✓	✓		✓	✗	✗
	✓	✓	✓		✓	✓	✓
	✓	✗	✓		✓	✓	✓

	✗	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✗	✓
	✓	✗	✓		✓	✓	✓
	✗	✓	✓		✓	✗	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✗	✗	✓
	✓	✗	✓		✓	✓	✓
	✓	✗	✓		✓	✓	✓
	✓	✓	✓		✓	✗	✓
	✓	✓	✓		✓	✗	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✗	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✗	✗	✓		✓	✓	✓

	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✗	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✗	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✗	✓
	✓	✓	✓		✓	✓	✓
	✓	✗	✓		✗	✓	✓
	✓	✓	✓		✗	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✗	✓	✓		✓	✓	✓
	✓	✗	✓		✓	✓	✓
	✓	✓	✓		✓	✓	✓
	✓	✓	✓		✓	✗	✓



Fonte: Elaborado pelo próprio autor.

A Tabela 3 sumariza o Quadro 1. Por meio desta podemos observar a acurácia de cada um dos classificadores. Destacamos a acurácia superior do RE25 em relação ao SEEK, uma vez que este último possui aproximadamente 150.000 espécies catalogadas em sua base de conhecimento, ao passo que o RE25 foi treinado com apenas 102 espécies. Quando comparado com o PLANTNET, a acurácia do RE25 obteve aproximadamente 11% a menos. Porém, é importante resgatar a informação de que o PLANTNET possui uma base de conhecimento com mais de 1 milhão de imagens, ao passo que neste trabalho contamos com 8.196 imagens, usando apenas 6.557 (80%) para o treino do RE25.

Desta forma, os resultados mostram que o uso de CNN para tal fim é uma proposta promissora. Considerando o aumento na quantidade de imagens para treinamento, o RE25 é capaz de aumentar a acurácia com uma base de conhecimento menor que a dos classificadores utilizados.

Tabela 4 - Tabela de Percentual de Acerto

Método	**Acurácia (%)**
RE25	88.04%
SEEK	79.41%
PLANTNET	99.01%

Fonte: Elaborado pelo próprio autor.

Considerações Finais

Neste trabalho foi desenvolvido o protótipo de um sistema de classificação, chamado RE25, que propõe fazer a classificação de imagens de flores. A tarefa de classificação botânica é uma tarefa que, apesar de poder contar com alguns aplicativos que auxiliam este processo, nem todos estes aplicativos realizam o processo de classificação de forma totalmente automática ou com uma taxa aceitável e acerto.

O protótipo desenvolvido foi comparado com dois dos aplicativos mais precisos encontrados no mercado, obtendo superioridade sobre um deles. Apesar de não ter obtido acurácia superior quando comparado com o PLANTNET, o RE25 obteve um resultado competitivo necessitando muito menos imagens para treinamento.

Com trabalho futuro considera-se a criação de um aplicativo mobile e um sistema de cooperação entre usuários para aumentar a quantidade de imagens que alimenta a base de conhecimento do RE25, permitindo que profissionais da área de botânica e entusiastas possam ter acesso à uma ferramenta colaborativa e precisa.

Referências

ARLOT, S. et al. A survey of cross-validation procedures for model selection. Statistics surveys, v. 4, p. 40-79, 2010.

CARVALHO, F. O."aplicação de support vector machine na detecção de falhas em sensores de uma coluna debutanizadora.", Congresso Brasileiro de Engenharia Química em Iniciação Científica. 21-24 jul. 2019, Uberlândia-MG.

COSTA, FELIPE APL, Marinês Eiterer, and Lucia Maria Paleari. "Capítulo 4 Classificação biológica: desafios na história da biologia."

Data Science Academy. Deep Learning Book, 2019. Disponível em: <http://www.deeplearningbook.com.br/>. Acesso em: 08 out. 2019

JÚNIOR, Sena, et al. Algoritmo para classificação de plantas de milho atacadas pela lagarta do cartucho (Spodoptera frugiperda, Smith) em imagens digitais. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 2001, 5.3: 502-509.

GAROFALO, D. F. T.; MESSIAS, C. G.; LIESENBERG, V.; BOLFE, E. L.; FERREIRA, M. C. Análise comparativa de classificadores digitais em imagens do Landsat-8 aplicados ao mapeamento temático. Pesquisa Agropecuária Brasileira, v. 50, n. 7, p. 593-604, 2015.

KUMAR, Neeraj, et al. Leafsnap: A computer vision system for automatic plant species identification. In: *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2012. p. 502-516.

MASSUCATTO, Jean Daniel Prestes. *Aplicação de conceitos de redes neurais convolucionais na classificação de imagens de folhas*. 2018. Bachelor's Thesis. Universidade Tecnológica Federal do Paraná.

MOREIRA, Lenadro Juvêncio, et al. Classificação de dados combinando mapas auto-organizáveis com vizinho informativo mais próximo. 2016.

NICOLAU, Paula Bacelar. "História da classificação biológica." (2017): 1-28.

PENHA, Deyvison de Paiva, et al. Rede neural convolucional aplicada à identificação de equipamentos residenciais para sistemas de monitoramento não-intrusivo de carga. 2018.

PRESTES, M. E. B.; OLIVEIRA, P.; JENSEN, G. M. As origens da classificação de plantas de Carl von Linné no ensino de Biologia. Filosofia e História da Biologia, Ribeirão Preto, v.4, p. 101-137, 2009.

PROCÓPIO, Lílian Costa; SECCO, Ricardo de Souza. A importância da identificação botânica nos inventários florestais: o exemplo do “tauari”(Couratari spp. e Cariniana spp.-Lecythidaceae) em duas áreas manejadas no estado do Pará. 2008.

PONTE M. J. M, “Referencial Semântico no suporte da identificação botânica de Espécies Amazônicas”, Tese de doutorado – UFOPA/UNL -Santarém – Pará, Abril 2017.

RIGHETTO, Guilherme. *O uso da rede neural convolucional como extrator de recursos aplicados ao problema de identificação de escritores* . 2016. Tese de Bacharel. Universidade Tecnológica Federal do Paraná.

Robots. Conjunto de dados da flor da categoria 102. Disponível em: <http://www.robots.ox.ac.uk/~vgg/data/flowers/102/index.html>. Acessado em 19 jun. 2019.

TOROK, L. Método de Otsu. Instituto de Computação (UFF). 2016. Disponível em:
< http://www2.ic.uff.br/~aconci/OtsuTexto.pdf>. Acesso em: 13 nov. 2019.