

**INSTITUTO FEDERAL GOIANO - CAMPUS CERES**  
**BACHARELADO EM BACHARELADO EM SISTEMAS DE INFORMAÇÃO**  
**ANDREY SILVA RIBEIRO**

**RECONHECIMENTO DE NERVURAS DE FOLHAS EM PLANTAS**  
**UTILIZANDO REDES NEURASIS CONVOLUCIONAIS**

**CERES - GO**  
**2019**

**ANDREY SILVA RIBEIRO**

**RECONHECIMENTO DE NERVURAS DE FOLHAS EM PLANTAS  
UTILIZANDO REDES NEURAS CONVOLUCIONAIS**

Trabalho de curso apresentado ao curso de Bacharelado em Sistemas de Informação do Instituto Federal Goiano – Campus Ceres, como requisito parcial para a obtenção do título de bacharel em Bacharelado em Sistemas de Informação, sob orientação do Prof. Me. André Luiz Nascente Ferreira e co-orientação do Prof. Me. Ronneesley Moura Teles.

**CERES - GO**

**2019**

Sistema desenvolvido pelo ICMC/USP  
Dados Internacionais de Catalogação na Publicação (CIP)  
**Sistema Integrado de Bibliotecas - Instituto Federal Goiano**

R484r           Ribeiro, Andrey Silva  
                  Reconhecimento de Nervuras de Folhas em Plantas  
                  utilizando Redes Neurais Convolucionais / Andrey  
                  Silva Ribeiro; orientador André Luiz Nascente  
                  Ferreira; co-orientador Ronneesley Moura Teles. --  
                  Ceres, 2019.  
                  29 p.

                  Monografia ( em Bacharelado em Sistemas de  
                  Informação) -- Instituto Federal Goiano, Campus  
                  Ceres, 2019.

                  1. Redes Neurais Artificiais. 2. Reconhecimento  
                  de Nervuras. 3. Visão Computacional. 4. Identificação  
                  de Plantas. I. Ferreira, André Luiz Nascente,  
                  orient. II. Teles, Ronneesley Moura, co-orient. III.  
                  Título.



**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR PRODUÇÕES TÉCNICO-CIENTÍFICAS NO REPOSITÓRIO INSTITUCIONAL DO IF GOIANO**

Com base no disposto na Lei Federal nº 9.610/98, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia Goiano, a disponibilizar gratuitamente o documento no Repositório Institucional do IF Goiano (RIIF Goiano), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, em formato digital para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IF Goiano.

**Identificação da Produção Técnico-Científica**

- Tese
- Dissertação
- Monografia - Especialização
- TCC - Graduação
- Produto Técnico e Educacional - Tipo: \_\_\_\_\_
- Artigo Científico
- Capítulo de Livro
- Livro
- Trabalho Apresentado em Evento

Nome Completo do Autor:

Matrícula:

Título do Trabalho:

**Restrições de Acesso ao Documento**

Documento confidencial:  Não  Sim, justifique: \_\_\_\_\_

Informe a data que poderá ser disponibilizado no RIIF Goiano: \_\_\_\_/\_\_\_\_/\_\_\_\_

- O documento está sujeito a registro de patente?  Sim  Não
- O documento pode vir a ser publicado como livro?  Sim  Não

**DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA**

O/A referido/a autor/a declara que:

1. o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
2. obteve autorização de quaisquer materiais incluídos no documento do qual não detém os direitos de autor/a, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia Goiano os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
3. cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia Goiano.

Goias Local, 04/12/2019 Data

*Carla Regina de Souza*  
Assinatura do Autor e/ou Detentor dos Direitos Autorais

Ciente e de acordo:

*André Luiz M. Ferraz*  
Assinatura do(a) orientador(a)

#### ANEXO IV - ATA DE DEFESA DE TRABALHO DE CURSO

Ao(s) 21 dia(s) do mês de NOVEMBRO do ano de dois mil e 2019, realizou-se a defesa de Trabalho de Curso do(a) acadêmico(a) ANDREY SILVA RIBEIRO, do Curso de BACHARELADO EM SISTEMAS DE INFORMAÇÃO matrícula \_\_\_\_\_, cujo título é "RECONHECIMENTO DE NERVURAS DE FOLHAS EM PLANTAS UTILIZANDO REDES NEURAIS ARTIFICIAIS". A defesa iniciou-se às 20 horas e 00 minutos, finalizando-se às 20 horas e 50 minutos. A banca examinadora considerou o trabalho APROVADO com média 8,6 no trabalho escrito, média 8,9 no trabalho oral, apresentando assim média aritmética final 8,7 de **pontos**, estando o(a) estudante APROVADO para fins de conclusão do Trabalho de Curso.

Após atender às considerações da banca e respeitando o prazo disposto em calendário acadêmico, o(a) estudante deverá fazer a submissão da versão corrigida em formato digital (.pdf) no Repositório Institucional do IF Goiano – RIIF, acompanhado do Termo Ciência e Autorização Eletrônico (TCAE), devidamente assinado pelo autor e orientador.

Os integrantes da banca examinadora assinam a presente.

André Luiz M. Arruda

Assinatura Presidente da Banca

Lucas José de Souza

Assinatura Membro 1 Banca Examinadora

Luiz Fernando Calça Silva

Assinatura Membro 2 Banca Examinadora

## **AGRADECIMENTOS**

Agradeço ao Instituto Federal Goiano - Campus Ceres, sua administração, direção e corpo docente que sempre buscaram oferecer oportunidades únicas e um ensino de qualidade.

Ao meu orientador, Prof. Me. André Luiz Nascente Ferreira pelo constante apoio, disposição e vontade de ajudar, assim como suas correções e incentivos que foram essenciais para este trabalho.

Ao meu co-orientador, Prof. Me. Ronneesley Moura Teles pelo apoio, confiança e sua enorme vontade de ensinar e aprender, sempre será uma grande inspiração.

A Prof. Dra. Jaqueline Alves Ribeiro por toda a sua ajuda e confiança incondicional todos esses anos.

A todos os professores do curso de Bacharelado em Sistemas de Informação pela disposição em ajudar, ensinar e colaborar em todos os momentos.

A todos meus amigos que sempre estiveram ao meu lado, mesmo em momentos difíceis.

A minha família por sempre acreditar em mim.

'As escolhas mais difíceis requerem as vontades  
mais fortes.'

Thanos - Vingadores: Guerra Infinita

## RESUMO

O presente trabalho busca a segmentação de imagens para identificação de nervuras de folhas de plantas disponíveis no *dataset* Flavia por meio de Redes Neurais Artificiais. As imagens passaram por um pré-processamento para criação de máscaras de segmentação, conversão em escala de cinza, redimensionamento e normalização.

A criação de arquiteturas para a segmentação utilizando Redes Neurais Convolucionais tem como base hiperparâmetros como a quantidade de épocas de treinamento, tamanho do lote, função de ativação, algoritmo de aprendizagem, dentre outros. A quantidade de parâmetros da rede neural e o valor da função de perda foram utilizados para escolha do melhor modelo. Os resultados mostram que o melhor modelo foi utilizando a arquitetura U-Net com função de perda binary crossentropy no valor de perda de 0.0209.

**Palavras-chave:** Redes Neurais Artificiais. Reconhecimento de Nervuras. Visão Computacional. Identificação de Plantas



## **ABSTRACT**

The present work seeks the segmentation of images to identify leaf veins of plants available on the Flavia dataset through Convolutional Neural Networks. Images have been preprocessed for mask segmentation, grayscale conversion, rescaling and normalization.

The creation of segmentation architectures using Convolutional Neural Networks is based on hyperparameters such as the number of training times, batch size, activation function, learning algorithm, among others. The number of neural network parameters and the loss function value were used to choose the best model. The results show that the best model is using the U-Net architecture with binary crossentropy loss function in the value of 0.0209.

**Keywords:** Convolutional Neural Networks. Vein Recognition. Computer Vision. Plant identification.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Máscara da nervura da folha de planta . . . . .	16
Figura 2 – Sumário do modelo sequencial . . . . .	20
Figura 3 – Predição do modelo U-Net . . . . .	21
Figura 4 – Função de perda no conjunto de teste e Função de perda no conjunto de validação. . . . .	22
Figura 5 – Função de perda nos conjuntos de teste e validação com 1000 épocas. . . .	22

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
1.1	Objetivo Geral	11
1.2	Objetivos específicos	11
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>12</b>
2.1	Redes Neurais Artificiais	12
2.2	Classificação de folhas do ponto de vista botânico	14
2.3	Utilização de RNAs	14
<b>3</b>	<b>MATERIAIS E MÉTODO</b>	<b>16</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>19</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>23</b>
	REFERÊNCIAS	24

# 1 INTRODUÇÃO

A identificação da nervura da folha de uma planta é um dos critérios para a determinação de sua espécie. A classificação de folhas permite aos pesquisadores analisarem diversas estruturas, dimensões, disposição e conectividade das mesmas, se tornando um campo de estudo amplo e diversificado que busca compreender a história evolutiva do nosso planeta em períodos geológicos distintos por meio dos diferentes tipos de plantas, sempre considerando os vários passos e critérios para essa classificação (VIDAL; VIDAL, 2006).

Os autores Junior e Covolan (2018) dizem que o domínio de técnicas para a classificação de folhas de plantas na área agrícola é importante para que seja possível obter dados relevantes do ponto de vista científico com a obtenção de um elevado nível de conhecimento sobre a área de interesse e do ponto de vista econômico, gerando empregos com a mão de obra.

A análise automática de folhas de plantas é um dos grandes desafios na determinação das espécies. Estudos recentes afirmam que as propriedades das nervura de uma folha pode ser de alta importância ao realizar o reconhecimento da planta (LARESE et al., 2014). Segundo Garcia-Garcia et al (2018), as Redes Neurais Convolucionais (CNN) vem se destacando entre essas técnicas, por sua abordagem com uma boa margem de precisão e, na maioria das vezes, de eficiência.

## 1.1 Objetivo Geral

O presente trabalho busca o reconhecimento, identificação e segmentação de nervuras de folhas.

## 1.2 Objetivos específicos

- Pré-processamento das imagens com máscaras. Normalização e leitura;
- Avaliar arquiteturas de redes neurais;
- Ajustar os hiperparâmetros da rede;
- Escolha do melhor modelo.

## 2 REFERENCIAL TEÓRICO

### 2.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são sistemas computacionais inspirados no sistema neural biológico humano, utilizam cálculos de funções lineares e não lineares através de unidades de processamento simples que são distribuídas em camadas e estas camadas são interligadas por conexões, onde cada conexão possui um peso que, por sua vez, é utilizado para ponderar as entradas recebidas (VIEIRA; PARREIRA; SILVEIRA, 2016).

Um dos objetivos da Visão Computacional é reproduzir a visão humana em máquinas, se tornando uma tecnologia de sensoriamento extremamente importante. Sua vasta aplicabilidade pode ser implementada por diversas áreas que utilizam digitalização, processamento e análise da imagem e interação com a aplicação (RIOS; NETTO, 2016). Com esses métodos a Visão Computacional vem ganhando destaque nas últimas décadas, pois possibilita a extração de informações úteis das imagens analisadas (FERREIRA, 2017).

As RNAs vem sendo amplamente estudadas e utilizadas em diversas áreas do conhecimento devido a sua capacidade de extrair informações e padrões através de exemplos que não foram apresentados de forma explícita (BISI; NETO; BONINI, 2015). Sua aplicabilidade em áreas agrícolas colabora em alguns desafios como a detecção de sementes defeituosas, plantas daninhas que prejudicam o desenvolvimento de vegetações, dentre outros (JUNIOR; COVOLAN, 2018).

Durante o processo de aprendizado, a rede extrai do conjunto de treinamento um número suficiente de padrões para generalizar o problema, para ao final, ser capaz de produzir respostas para dados desconhecidos que façam parte do mesmo problema para o qual foi treinada. Este treinamento pode ser dividido em treinamentos supervisionados e não-supervisionados. No primeiro caso, cada exemplo trás a classe a que pertence, já no treinamento não-supervisionado, essa informação de classe não é considerada ou conhecida (VIEIRA; PARREIRA; SILVEIRA, 2016).

Além do treinamento, outra característica importante para a implementação de uma RNA é a sua arquitetura, que é definida pelo número de camadas e o tipo de conexão entre os neurônios. As RNAs com camadas múltiplas são utilizadas para problemas complexos com análise de diversas combinações de entradas, que diferente das redes de camada única, possuem uma ou mais camadas entre as de entrada e saída, conhecidas como camadas ocultas. Quanto

as conexões entre os neurônios, são definidas como *feedforward* e *feedback*.

Quando a rede é estruturada em camadas e suas entradas não são realimentadas pelas suas saídas, ela é uma rede *feedforward* ou não recorrente. Dentre os algoritmos de treinamentos com mais eficiência nesse tipo de rede está o *backpropagation*. As RNAs *feedback* se diferenciam ao possuir uma estrutura de realimentação da saída para a entrada (VIEIRA; PARREIRA; SILVEIRA, 2016).

As Redes Neurais Convolucionais (CNN) são um tipo particular de rede neural profunda que se mostrou eficaz em seus treinamentos e generalização de imagens, sendo assim, amplamente adotadas em estudos de Visão Computacional (FERREIRA, 2017).

Para melhor compreender uma CNN é necessário entender como funciona uma imagem. Imagens são computacionalmente representada por matrizes onde cada posição da matriz representa um pixel da imagem que geralmente são em escala de cinza (*grayscale*) ou coloridas (CLAPPIS, 2012).

Uma imagem em *grayscale* é representada em uma matriz 2D com cada valor variando de 0 (preto) até 255 (branco). Na imagem colorida, a representação é feita por uma matriz 3D para que seja possível armazenar valores da escala de vermelho, verde e azul, conhecida como *RGB* (CLAPPIS, 2012).

Em alguns casos, quando o interesse da análise está apenas nos padrões associados a morfologia da imagem, a mesma é convertida em *grayscale* para que as informações de cores sejam desconsideradas (LARESE et al., 2014).

Uma das tarefas da Visão Computacional é a segmentação de imagens que ocorre quando regiões específicas da imagem são rotuladas de acordo com o que esta sendo mostrado. Mais especificamente, a segmentação semântica mostra o que existe na imagem e onde está localizado, criando classes e rotulando cada *pixel* (JORDAN, 2018).

Uma Rede Neural Convolucional utiliza operação de convolução em suas camadas. No conceito da matemática, a convolução é o processo de utilizar uma função linear que se baseia na multiplicação de matrizes. Uma CNN possui uma 'janela' de multiplicação que percorre a imagem a um passo definido para encontrar a característica na imagem de entrada (SPOLTI, 2018).

A complexidade de uma CNN não depende do número de amostras, porém é necessário um conjunto de dados relativamente grande para seu treinamento. Uma rede neural convolucional é mais robusta, com aprendizado orientado através dos pesos, que a torna capaz de distinguir

o grau de relevância das características (JUNIOR; COVOLAN, 2018).

Os autores Ronneberger et al (2015) realizaram um estudo onde apresentam uma nova arquitetura chamada de U-Net para treinamento de Redes Neurais Convolucionais que se baseia em uma contração e expansão da imagem durante o treinamento, permitindo uma segmentação precisa das características que supera outros modelos.

A arquitetura U-Net foi desenvolvida inicialmente para segmentação de imagens biomédicas em casos onde o número de amostras é reduzido devido a complexidade e dificuldade em obter um banco de dados grande o suficiente para o treinamento, sendo útil em diversos casos (SPOLTI, 2018).

## **2.2 Classificação de folhas do ponto de vista botânico**

Existe uma grande variedade de plantas e cada uma delas tem uma função importante na ecologia do planeta, a classificação de uma planta é importante para a criação de grupos para fácil reconhecimento. Essa classificação é essencial para a botânica, medicina e diversas outras indústrias, e tem como base diferentes características como casca do caule, flores, tamanho, textura e folhas (YADAV et al., 2017).

A folha é um dos principais elementos para essa classificação por possuir informações e características taxonômicas da planta. A folha possui a vantagem de estar disponível por vários meses em um ano, as flores e frutas por exemplo, ficam disponíveis apenas por algumas semanas na planta. Dentre as características para classificação da folha está a cor, forma, textura e nervuras (VIJAYALAKSHMI; MOHAN, 2016).

## **2.3 Utilização de RNAs**

No Brasil, existe um aumento gradativo de estudos com a utilização de RNAs, Sena Júnior et al (2001) desenvolveram um método de processamento de imagens para identificar sintomas em plantas de milho atacadas pela lagarta do cartucho, as Redes Neurais Artificiais foram utilizadas para processamento rápido e eficiente das imagens digitais. Os autores obtiveram resultados excelentes em folhas saudáveis, ocorrendo poucos erros nas bordas das folhas, já em folhas atacadas ou raspadas, o resultado não foi o ideal pois o programa classificou áreas com tonalidades mais claras como folhas atacadas.

O estudo dos autores Vieira, Parreira, Silveira (2016) aborda um sistema que utiliza RNAs treinadas para buscar casos de pessoas com diabetes em uma determinada base de dados

e auxiliar profissionais da área da saúde a realizar o diagnóstico da doença. O trabalho foi desenvolvido com RNAs de múltiplas camadas através da linguagem de programação Java, tecnologias Web e *open source*, dentre eles a linguagem de marcação de texto Hypertext Markup Language (HTML5), uma vez que se trata de um sistema Web, além do Sistema de Gerenciamento de Banco de Dados MySQL. O índice de acertos da RNA foi de 85,7%.

Já no exterior, os autores Garcia-Garcia et al (2018) mostram um estudo sobre a segmentação semântica aplicada a imagens e vídeos, uma vez que a compreensão desse tema é importante para o quadro geral da visão computacional em um mundo onde aplicações se alimentam de imagens com por exemplo, direção automática de carros e motores de busca de imagens. O trabalho abrange 28 conjuntos de dados e 29 métodos, dentre eles AlexNet, VGG, ResNet e traz um conhecimento básico e necessário ao tema.

Em uma pesquisa realizada por Yadav et al (2017) discute o método de um sistema automatizado baseado em visão computacional para processamento e classificação de folhas de plantas. Tendo em vista que uma folha pode ter diferentes parâmetros e características que a identificam como única, um algoritmo automático para identificar e processar parâmetros de textura, geometria, dentre outras, se faz necessário.

O autor Ferreira (2017) utiliza RNA para detecção e classificação de ervas daninhas em imagens de lavouras de soja, com o resultado obtido pela rede é possível direcionar o herbicida específico ao tipo de erva daninha encontrado pelo processamento. As imagens foram capturadas de uma plantação de soja e tiveram as características extraídas manualmente para posteriormente, ser criado um banco de imagens para o treinamento dos classificadores. Os resultados obtidos pelo autor foram de 99.5% de precisão média nas imagens analisadas com um conjunto de 15 mil imagens e uma precisão superior a 98% na classificação de todas as classes.

Seguindo a mesma linha de trabalho, os autores Vijayalakshmi (2016) utilizam uma técnica diferente das RNAs. Conhecida como *Fuzzy Relevance Vector Machine* (FRVM), essa abordagem utiliza regras baseadas na lógica Fuzzy e analisa propriedades como textura, forma e cor para realizar a classificação. Ao final do pré-processamento o método FRVM é empregado para caracterizar o tipo de folha.



### 3 MATERIAIS E MÉTODO

A abordagem proposta neste trabalho para a segmentação das nervuras das folhas em plantas é composto por etapas. A primeira etapa é manual e consiste na especificação da característica a ser analisada diretamente na imagem da folha, redimensionamento, conversão das imagens originais em escala de cinza, conversão das máscaras em binário, alocação de espaço e carregamento das imagens na memória RAM. As imagens digitalizadas tem como fonte o *dataset* Flavia.

A especificação da característica, consiste na criação de máscaras das nervuras. Por meio do software Photoshop CS 6, uma camada transparente é criada por cima da imagem original, sempre preservando suas dimensões. Com a utilização da ferramenta Lápis ou equivalente do software, o traçado da nervura de cada folha é realizado na camada transparente, tendo em vista que a imagem original não pode ser alterada.

Após a finalização do traçado, a imagem da folha é retirada e em seu lugar é adicionado o fundo em branco, o resultado final é uma máscara do traçado da nervura em preto, com o fundo em branco do mesmo tamanho da imagem da folha. Durante esse processo foi considerado apenas a nervura principal e as secundárias, quem tem sua origem partindo da nervura principal, como mostra a Figura 1:

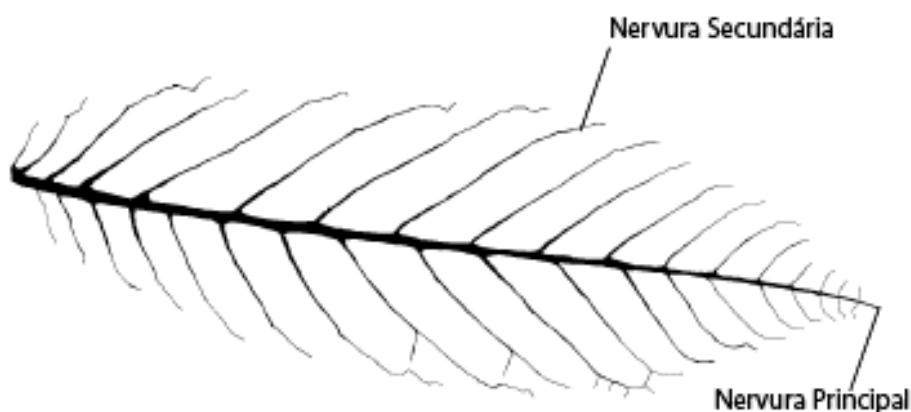


Figura 1 – Máscara da nervura da folha de planta

O redimensionamento da imagem é necessário para reduzir o poder computacional necessário para análise e treinamento das imagens, a conversão em escala de cinza é utilizada para que as informações de cores não sejam consideradas. Utilizando a linguagem de programação Python na versão 3.6, uma função foi codificada para redimensionar as imagens para um tamanho de 224 pixels de altura por 224 pixels de largura e ao mesmo tempo converter as imagens para escala de cinza. Todos os treinamentos foram realizados com as imagens das folhas de plantas em escala RGB e em escala de cinza. Para conversão, redimensionamento e manipulação geral de matrizes foi utilizado a biblioteca Numpy em sua versão 1.16 juntamente com a biblioteca OpenCV 2.

Com o redimensionamento e conversão feitos, é realizado a alocação do espaço na memória RAM, leitura e carregamento das imagens no espaço reservado. O mesmo processo de redimensionamento, conversão, leitura e carregamento é realizado com as máscaras criadas, porém são convertidas em binário e não em escala de cinza. A função para conversão em binário garante que nas máscaras existam apenas pixels no valor de 0 (preto) e 255 (branco) sem nenhum outro valor intermediário. Durante a leitura, as máscaras são transformadas em um vetor para possibilitar a utilização da função de ajuste dos pesos das classes. Também é realizado a normalização da imagem, com uma divisão por 255, para as funções de ativação não operarem nos intervalos de saturação.

A próxima etapa é a separação dos conjuntos de treinamento, teste e validação. Do conjunto total de 214 imagens, 70% das imagens foram para o treinamento, 15% para o teste e 15% para a validação. Nessa etapa é importante a separação correta dos conjuntos e que as imagens não se repitam entre eles, isso garante uma melhor eficácia da aprendizagem da rede neural. Na arquitetura U-Net a divisão dos conjuntos é 50% para treino e 50% para validação, seguindo o padrão do modelo.

A etapa seguinte trata da definição do modelo, onde é declarado a quantidade de camadas, função de ativação, *padding*, dentre outros parâmetros que variam de acordo com o modelo. Em seguida, os pesos das classes são especificados por meio de uma função que realiza a contagem dos pixels na imagens do conjunto de treinamento, o total de pixel pode ser utilizado para que a soma dos pesos seja igual a 1, sempre dando maior importância para a classe menos representada. Os pesos das classes bem definidos tem grande impacto no ajuste e penalização do algoritmo de treinamento em caso de classes desbalanceadas durante o treinamento.

Todos os modelos foram treinados em 300 épocas, com um *dropout* de 20%, *padding* que preserva as bordas da imagem, função de ativação sigmoide, função de perda *binary crossentropy*, algoritmo de aprendizagem Adam, e um tamanho do lote de 5 imagens. O tamanho do lote precisou ser reduzido para o total de 5 devido a problemas com limite de memória da GPU.

Os *framework* utilizados foram o Keras versão 2.3.1 e Tensorflow em sua versão 1.14.0. O Tensorflow disponibiliza opção de instalação para CPU e GPU, e devido a eficiência do processamento, foi escolhido a utilização da GPU em uma placa de vídeo *Nvidia GeForce GTX 750 Ti* com 640 núcleos CUDA, 2048 *megabytes* de memória total e 128-bit de memória de interface. O sistema operacional escolhido foi o Ubuntu 18.04.3 LTS de 64 bits.

Um total de três diferentes modelos foram definidos, sendo eles um modelo sequencial com quatro camadas com o *kernel* no tamanho de 3x3 em cada camada. Uma arquitetura U-Net por ser indicada em casos onde existe uma dificuldade em obter um base de dados grande o suficiente para o treinamento da rede e um modelo com a utilização da técnica de *Transfer Learning*, onde a rede usa um modelo pré-treinado para o problema definido. Nesse contexto, o modelo DenseNet foi escolhido.

Após a definição dos modelos e do dicionário com pesos das classes que é utilizado somente durante o treinamento que por sua vez é iniciado com a utilização do algoritmo de aprendizagem Adam, função de ativação sigmoide, quantidade de 300 épocas de treinamento e *droupout* de 20%.

## 4 RESULTADOS E DISCUSSÃO

O pré-processamento das imagens foi a base para os conjuntos utilizados por todos os modelos. As imagens das folhas de plantas passaram pelo redimensionamento para um tamanho de  $224 \times 224$ , conversão em escala de cinza para as imagens das folhas e conversão em binário para as máscaras criadas, todo o resultado desse processo é utilizado como parâmetro de entrada dos modelos.

Durante a fase do treinamento, percebeu-se a necessidade de utilizar a função de balanceamento de pesos das classes do *framework* Keras, uma vez que a quantidade de *pixels* brancos na imagem era superior a quantidade de *pixels* pretos. A contagem dos *pixels* foi realizada por meio de uma função que contabiliza a quantidade de *pixels* diferentes de 0 (pretos) nas máscaras criadas. Como no pré-processamento as máscaras foram convertidas em binário, o número total de 1 (brancos) foi obtida através da subtração da quantidade total de *pixels* pela quantidade de 0 (pretos) obtidas pela função.

A função de pesos das classes gerou uma dificuldade em sua utilização devido a documentação pequena e sem exemplos no *framework*. O resultado da contagem de *pixels* foi transformado em porcentagem e utilizado inversamente, de forma que a soma dos pesos fosse igual a 1 onde a classe 0 recebeu 0.97 e a classe 1 recebeu 0.03.

O modelo sequencial foi estruturado conforme o sumário da Figura 2, sendo o primeiro a ser treinado e não apresentou resultados satisfatórios, a função de perda que apesar de obter bons valores não passava confiabilidade ao modelo. A predição final do modelo sequencial resultou em uma imagem escura da folha sem características da nervura. O mesmo resultado foi obtido no treinamento com as imagens em RGB.

Model: "sequential\_1"

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 224, 224, 64)	640
conv2d_2 (Conv2D)	(None, 224, 224, 32)	18464
conv2d_3 (Conv2D)	(None, 224, 224, 64)	18496
conv2d_4 (Conv2D)	(None, 224, 224, 1)	577
dropout_1 (Dropout)	(None, 224, 224, 1)	0
flatten_1 (Flatten)	(None, 50176)	0
activation_1 (Activation)	(None, 50176)	0
Total params: 38,177		
Trainable params: 38,177		
Non-trainable params: 0		

Figura 2 – Sumário do modelo sequencial

O segundo treinamento foi o modelo por meio de *Transfer Learning*. O DenseNet121 foi escolhido e seguiu os mesmos parâmetro já pré-definidos. A acurácia e a função de perda obtiveram bons valores porém não foram o suficiente para uma boa predição, sendo o resultado uma imagem sem nenhum aspecto visual desejável. O treinamento com as imagens em RGB obtiveram a mesma saída.

O terceiro modelo foi de arquitetura U-Net seguindo os mesmo parâmetro dos outros dois treinados anteriormente, sendo este o que apresentou os melhores resultados. O modelo foi capaz de generalizar o problema e fazer uma predição qualitativa das nervuras das folhas de plantas, com os melhores resultados nas imagens das folhas em RGB. A predição após o treinamento foi satisfatória e adequada para a saída esperada, como mostra a Figura 3:

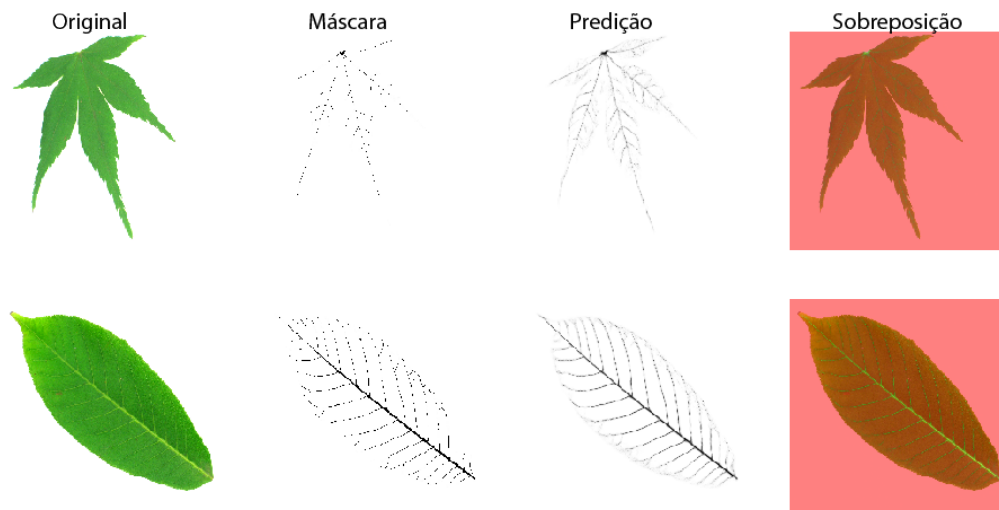


Figura 3 – Predição do modelo U-Net

O resultado da predição na Figura 3 mostra a imagem original, a máscara criada, a predição feita pelo modelo e a imagem sobreposta com a predição, respectivamente.

Para a escolha do melhor modelo a função de perda (*loss function*) foi considerada na análise. A função de perda é uma fonte básica de informações sobre o processo de aprendizado da rede. Ela mostra o valor de cada interação ao longo das épocas e o quanto a rede errou na detecção de padrões. Quanto menor o erro, maior é o acerto da predição da rede em relação ao dados de entrada.

Na arquitetura apresentada pelo modelo U-Net, a ponderação de perda é feita para cada pixel, de modo que exista um peso maior na borda dos objetos segmentados, isso permite uma ponderação de forma descontínua, facilitando a identificação da característica no mapa de segmentação binária, sendo um diferencial aos outros modelos.

O valor da função de perda no modelo U-Net treinado foi de 0.0209 no conjunto de treino ao final das 300 épocas de treinamento. Este resultado mostra que a rede conseguiu minimizar o seu erro e consequentemente realizar predições qualitativas para cada amostra do conjunto.

O gráfico da relação entre o valor da função de perda no conjunto de treinamento e o conjunto de validação pode ser visualizado na Figura 4:

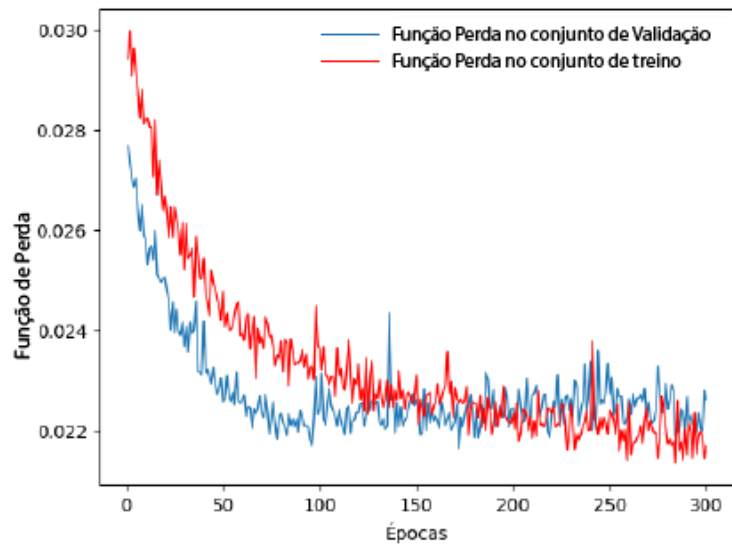


Figura 4 – Função de perda no conjunto de teste e Função de perda no conjunto de validação.

O modelo U-Net obteve resultados melhores em comparação aos outros modelos devido a sua conexão de camadas semelhantes durante todo o processo de contração da imagem, não necessitando de utilizar a classe de ajuste de pesos, sua ponderação de perda para cada pixel, além de sua grande capacidade de segmentação em banco de dados onde a quantidade de amostras não é suficientemente grande para o treinamento da rede por meio de outros modelos.

Após a definição do modelo U-Net como melhor modelo, outro treinamento com 1000 épocas foi realizado. O gráfico resultante na Figura 5 mostra a função de perda no conjunto de treino e validação indicando uma alta taxa de acerto da rede com o valor da função de perda no conjunto de treino em 0.0299.

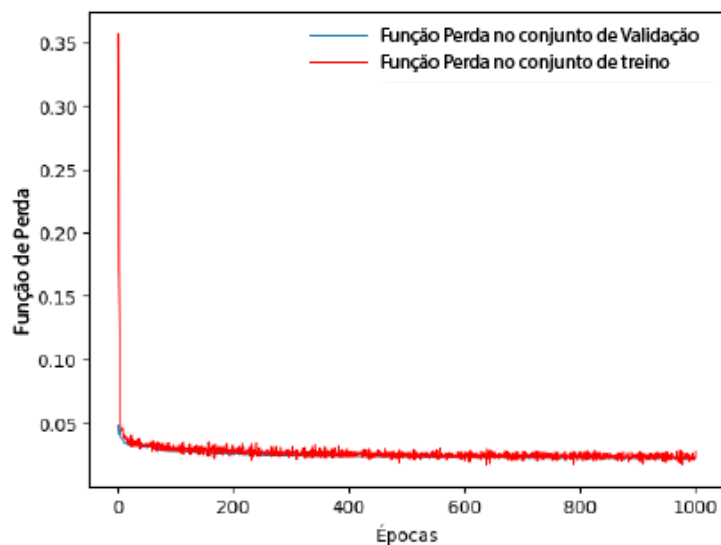


Figura 5 – Função de perda nos conjuntos de teste e validação com 1000 épocas.

## 5 CONCLUSÃO

Neste trabalho foram implementados e testados três diferentes tipos de modelos de Redes Neurais Convolucionais para a segmentação de nervuras em folhas de plantas através de imagens digitalizadas do *dataset* Flavia. As imagens passaram por todo o pré-processamento necessário para se realizar o treinamento da rede.

Os resultados obtidos mostram que o modelo sequencial e o modelo DenseNet por *transfer learning* não apresentaram resultados satisfatórios e não generalizaram o problema ao ponto de extrair a segmentação da nervura da folha de planta. A rede por U-Net foi capaz de obter o valor de 0.0209 em sua função de perda durante as 300 épocas de treino em uma base de dados considerada pequena para problemas de visão computacional. As predições resultantes do modelo mostram a capacidade da arquitetura em solucionar o problema proposto neste trabalho.

Uma das maiores dificuldades dos modelos foi a função de ajuste dos pesos de classes, uma vez que sua documentação falha ao não exemplificar o uso da mesma. Porém, no modelo U-Net não foi necessária sua utilização, isso mostra sua maior simplicidade e eficiência em relação aos outros modelos.

O modelo U-Net foi considerado o melhor modelo por conseguir realizar uma boa predição qualitativa das nervuras de folhas de plantas com o valor da função de perda sendo minimizado ao longo das épocas de treinamento.

A qualidade e padronização das imagens e dos seus rótulos são indispensáveis para bons resultados em diferentes modelos e arquiteturas. Com uma base de dados suficientemente grande e diferentes técnicas que podem ser desenvolvidas para análise de segmentação de imagens, novos modelos podem apresentar uma solução rápida e eficiente para diferentes problemas.



## REFERÊNCIAS

BISI, B. S.; NETO, A. B.; BONINI, C. d. S. B. Redes Neurais Artificiais : Utilização Do Algoritmo Retropropagação Para Classificação De Grupos Em Biosistemas , Parte 1: Introdução Teórica. **Fórum Ambiental da Alta Paulista**, v. 11, n. 1980-0827, p. 345–354, 2015.

CLAPPIS, A. M. **Uma introdução as redes neurais convolucionais utilizando o Keras**. 2012. Disponível em: <https://medium.com/data-hackers/uma-introducao-as-redes-neurais-convolucionais-utilizando-o-keras-41ee8dcc033e>.

FERREIRA, A. d. S. Redes Neurais Convolucionais Profundas na Detecção de Plantas Daninhas em Lavoura de Soja. 2017.

GARCIA-GARCIA, A. et al. **A survey on deep learning techniques for image and video semantic segmentation**. Elsevier B.V., 2018. v. 70. 41–65 p. ISSN 15684946. ISBN 9781467345873. Disponível em: <https://doi.org/10.1016/j.asoc.2018.05.018>.

JORDAN, J. **An overview of semantic image segmentation**. 2018. Disponível em: <http://www.jeremyjordan.me/semantic-segmentation/>.

JUNIOR, L. C. M.; COVOLAN, J. A. U. Detecção e classificação de plantas daninhas e seu estado da arte. v. 11, p. 391–403, 2018.

LARESE, M. G. et al. Automatic classification of legumes using leaf vein image features. **Pattern Recognition**, Elsevier, v. 47, n. 1, p. 158–168, 2014. ISSN 00313203. Disponível em: <http://dx.doi.org/10.1016/j.patcog.2013.06.012>.

RIOS, M.; NETTO, J. F. Uma Abordagem Utilizando Visão Computacional para Monitoramento de Robôs Móveis em Ambientes de Tarefas na Robótica Educacional. **Anais do XXVII Simpósio Brasileiro de Informática na Educação (SBIE 2016)**, v. 1, n. Cbie, p. 480, 2016.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 9351, p. 234–241, 2015. ISSN 16113349.

Sena Júnior, D. G. de 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**, v. 5, n. 3, p. 502–509, 2001.

SPOLTI, A. C. Classificação de vias através de imagens aéreas usando Deep Learning. 2018.

VIDAL, W. N.; VIDAL, M. R. R. **Vidal&Vidal\_2006.pdf**. 4. ed. Viçosa: Universidade Federal de Viçosa, 2006. 124 p.

VIEIRA, J. P.; PARREIRA, F. J.; SILVEIRA, S. R. Desenvolvimento De Um Sistema Para Auxílio Ao Diagnóstico De Diabetes Empregando Redes Neurais Artificiais (Sadd). **Revista de Sistemas e Computação**, v. 6, n. 1, p. 13, 2016.

VIJAYALAKSHMI, B.; MOHAN, V. Kernel-based PSO and FRVM: An automatic plant leaf type detection using texture, shape, and color features. **Computers and Electronics in Agriculture**, Elsevier B.V., v. 125, p. 99–112, 2016. ISSN 01681699. Disponível em: <http://dx.doi.org/10.1016/j.compag.2016.04.033>.

YADAV, A. et al. Automatic Identification of Botanical Samples of leaves using Computer Vision. **2017 International Work Conference on Bio-Inspired Intelligence: Intelligent Systems for Biodiversity Conservation, IWOBI 2017 - Proceedings**, 2017.