

## Relatório Projeto Final

**Nome:** Lucas Bragança da Silva

**Disciplina:** Aprendizado em Redes Neurais Profundas - INF 721

**Professor:** Levi Henrique Santana de Lelis

### 1. Introdução

Recentes avanços em técnicas de detecção de objetos em imagens por meio de redes neurais convolucionais profundas (CNN) têm permitido a construção de várias tipos aplicações. Grandes empresas como Google e Pinterest usam esses tipos de redes em seus produtos como o Google Photos e Pinterest Visual Search [1]. Projetos como Microsoft COCO (Common Objects in Context) e Google Imagenet ajudam avançar o estado da arte de detectores de objetos por meio de CNNs, onde grandes conjuntos de dados de treinamento já rotulados são disponibilizados para pesquisadores utilizarem em competições [2].

Detectores de objetos podem ser usados em inúmeras aplicações tanto na área acadêmica quanto na indústria. O poder de generalização de modelos de CNNs para detecção de objetos permitem usar uma mesma modelagem de rede para solucionar variados tipos de problemas. Neste trabalho foi realizado um estudo acerca do uso de um modelo CNN da literatura para auxiliar a realizar medidas de raízes e caules de plântulas<sup>1</sup>. Para isso um modelo pronto desenvolvido utilizando *Tensor Flow* foi utilizado para treinar uma rede.

O modelo adotado neste trabalho foi vencedor do desafio COCO 2018 e possui inúmeros tutoriais de sua utilização na internet. Para entender o funcionamento do modelo, primeiramente foi seguido o tutorial disponível no Github do Edge Electronics<sup>2</sup>, onde o mesmo modelo foi usado para aprender a detectar cartas de baralho como mostrado na Figura 1. Embora o tutorial tenha

---

<sup>1</sup> A plântula é o embrião vegetal já desenvolvido e ainda encerrado na semente ou também, planta recém-nascida. Fonte: <https://pt.wikipedia.org/wiki/Pl%C3%A2ntula>

<sup>2</sup>Disponível em: <https://github.com/EdgeElectronics/>

sido feito para ambiente Windows foi facilmente seguido de um ambiente Linux o que ajudou no restante do desenvolvimento deste trabalho.

Este trabalho está dividido da seguinte maneira: Na Seção 2 é dada a motivação envolvida no desenvolvimento do trabalho, na Seção 3 é apresentada



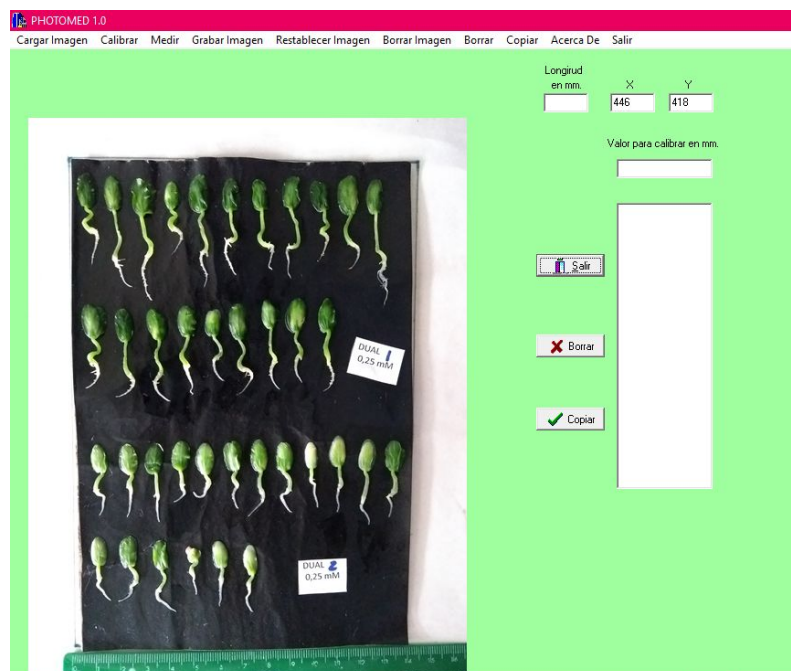
**Figura 1:** Exemplo de detecção de cartas de baralho.

a metodologia adotada no trabalho, na Seção 4 são apresentados os resultados obtidos e na Seção 5 é dada a conclusão do trabalho assim como uma discussão sobre trabalhos futuros.

## **2. Motivação**

A pesquisa por novos herbicidas que são produzidos por meio de compostos naturais é suma importância para produção agrícola não nociva ao meio ambiente. O fato desses herbicidas serem capazes de matar plantas daninhas e fungos com efeitos menos maléficos que os defensivos agrícolas comerciais contribui para preservação do meio ambiente e para a produção de alimentos mais saudáveis. Para a descoberta de novos herbicidas naturais pesquisadores da área agroquímica precisam testar a efetividade de novos compostos durante o processo da pesquisa. Para isso são realizados testes laboratoriais com plântulas de algumas espécies de plantas, onde elas são submetidas a um novo composto e posteriormente são realizadas comparações

e testes estatísticos com outras plântulas que não foram submetidas ao produto. Esses testes são realizados por meio das medidas de cada plântula, onde são mensurados o caule e raiz de mais 3000 amostras dependendo de quantos compostos serão testados. A tarefa de medir os caules e as raízes das plântulas são realizadas atualmente com o auxílio de um *software*. Para isso são tiradas fotos de todas as plântulas com um fundo preto e com uma régua como mostrado na Figura 2. Cada foto possui em média 30 amostras, e com a auxílio do *software* PhotoMed são realizadas as medidas de forma manual dessas duas características das plântulas. A Figura 2 mostra a tela do software PhotoMed com uma imagem aberta. A régua nas fotos serve para calibrar as medidas, assim, primeiramente o usuário deve marcar uma medida de acordo com a régua e então entrar com essa medida em uma caixa de entrada para realizar a calibração. Uma vez calibrado o usuário deve marcar na foto todas as médias que deseja realizar. Esse processo é demorado e pode levar dias para realizar todas medidas.



**Figura 2:** Tela do software de medição PhotoMed.

### 3. Metodologia

Para o desenvolvimento do trabalho primeiramente foi realizada uma pesquisa de modelos de CNNs para detecção de objetos que poderiam ser utilizados para identificar o caule e a raiz das plântulas. O modelo Fast R-CNN v2 é capaz de capturar até pequenos detalhes de objetos [3]. A capacidade desse modelo torna possível sua utilização no problema de medição, que primeiramente passa pela etapa de detectar o que é caule e o que é raiz em uma plântula, tarefa essa que pode ser difícil até mesmo para humanos em algumas situações e uma vez detectado essas partes realizar a medição de alguma forma.

Após a escolha do modelo foi preciso criar o conjunto de dados de treinamento, para isso foi utilizado o *software* LabelImg<sup>3</sup>. O LabelImg permite a criação de rótulos em imagens por meio de criação de retângulos sobre o objeto que se deseja rotular. Ao fim da marcação de todos os rótulos na imagem é possível salvar um arquivo XML no formato PASCAL VOC, que pode ser utilizado para criação dos dados treinamento. A Figura 3 apresenta a tela do aplicativo LabelImg com os retângulos marcando as três classes do problema a folha, o caule e a raiz.

Para o treinamento do modelo foram anotadas 179 imagens com aproximadamente 30 plântulas cada, onde 144 foram utilizadas no conjunto de treinamento e 35 foram utilizadas para o conjunto de teste.

Para a execução do treinamento do modelo foi utilizado uma máquina com processador Xeon e 64 GB de memória principal adaptado com uma GPU Nvidia Tesla k40. O sistema operacional utilizado foi Linux Ubuntu 16.04. Para a configuração do ambiente foi utilizado o Anaconda<sup>4</sup> com ambiente virtual Python 3.5. Todo código usado está disponível no Github do *Tensor Flow*<sup>5</sup>, onde é possível baixar e rodar os modelos prontos de *machine learning* codificados de

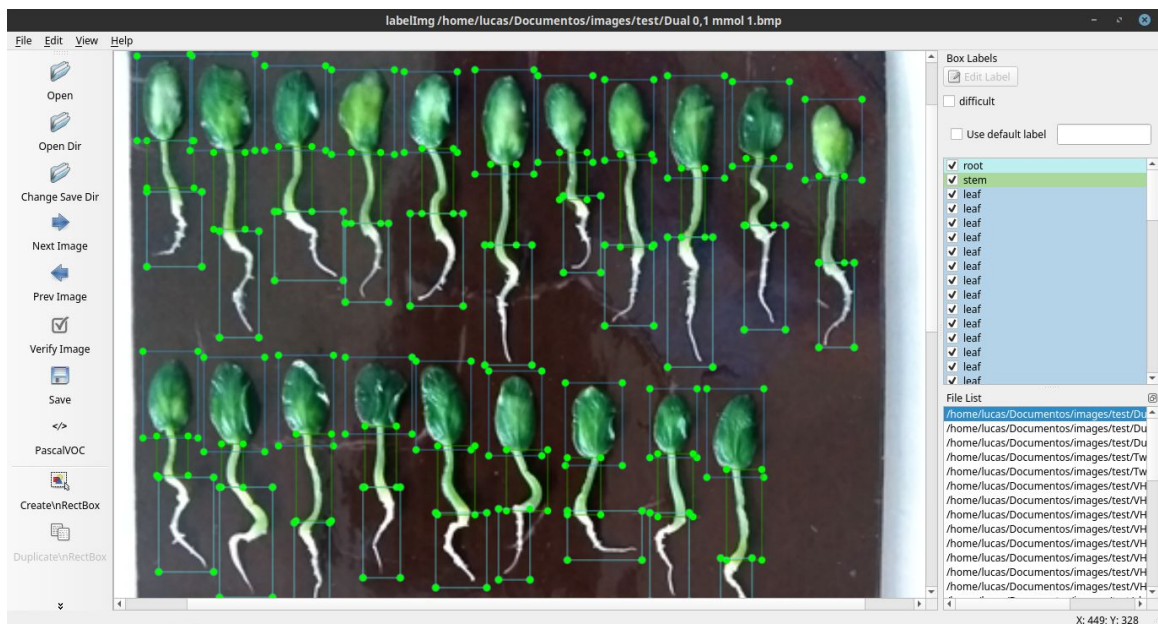
---

<sup>3</sup> Disponível em: <https://github.com/tzutalin/labelImg>

<sup>4</sup> Disponível em: <https://www.anaconda.com/distribution/>

<sup>5</sup> Disponível em: [https://github.com/tensorflow/models/tree/master/research/object\\_detection](https://github.com/tensorflow/models/tree/master/research/object_detection)

forma eficiente ficando a cargo do usuário apenas criar o conjunto de treinamento para o seu problema.



**Figura 3:** Tela software LabelImg.

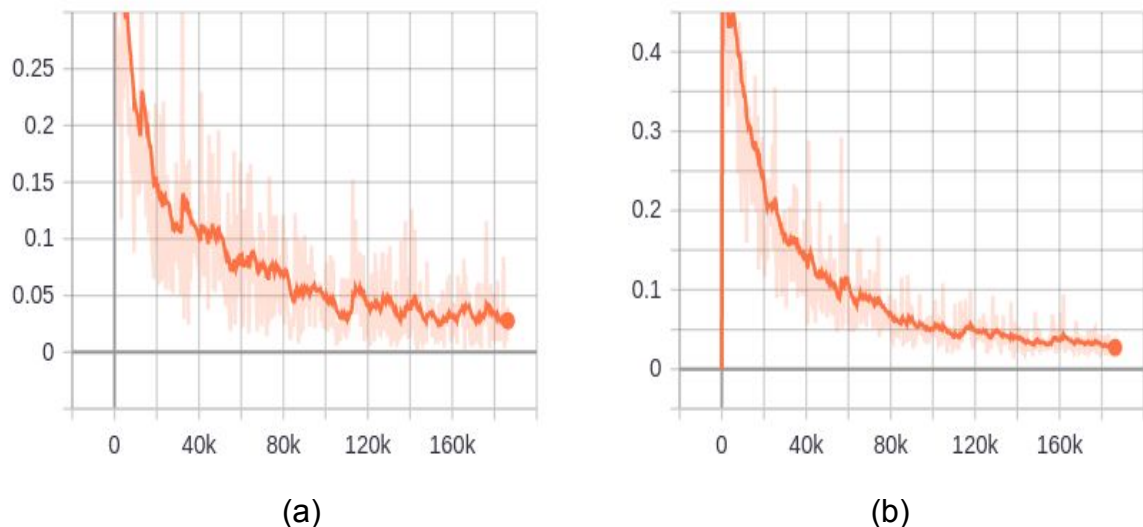
#### 4. Resultados

Nesta seção são apresentados os resultados do uso do modelo *Fast R-CNN* para detectar folhas, caules e raízes em plântulas. Devido o pouco tempo para rotular os dados de treinamento, foram utilizados apenas plântulas da espécie Pepino, sendo possível o uso do modelo treinado apenas nesta espécie. Após a criação de todos os rótulos como mostrado na Seção 3, foi utilizado um script fornecido na API do *detection object* do *Tensor Flow*<sup>4</sup> para gerar os arquivos *.recorder* do *Tensor Flow* que são usados como entrada da rede durante o processo treinamento.

Embora apenas uma pequena quantidade de imagens de treinamento terem sido rotuladas, cada imagem contém várias amostras de folhas, raízes e caules, o que torna o treinamento lento mesmo usando uma GPU como acelerador. O modelo foi treinado durante 15 horas , os gráficos da Figura 4

foram obtidos da aplicação *Tensor Board*<sup>6</sup> onde é possível acompanhar o processo de treinamento por meio de gráficos em tempo real. Podemos notar que a função de perda do modelo cai gradualmente durante as épocas de treinamento, o que significa que o modelo está aprendendo a detectar e classificar cada classe do problema.

Com o modelo devidamente treinado é possível exportar um grafo de inferência, onde é possível realizar a inferência fornecendo uma imagem como entrada e obtendo como saída um conjunto de retângulos compostos por 4 pontos que indicam a localização na imagem em pixels assim como a classe pertencente aquele retângulo. A Figura 5 mostra os retângulos desenhados sobre uma imagem de entrada com as plântulas onde é possível verificar a detecção das raízes, caules e folhas com porcentagem do modelo ter acertado a classificação escrita dentro da retângulo.



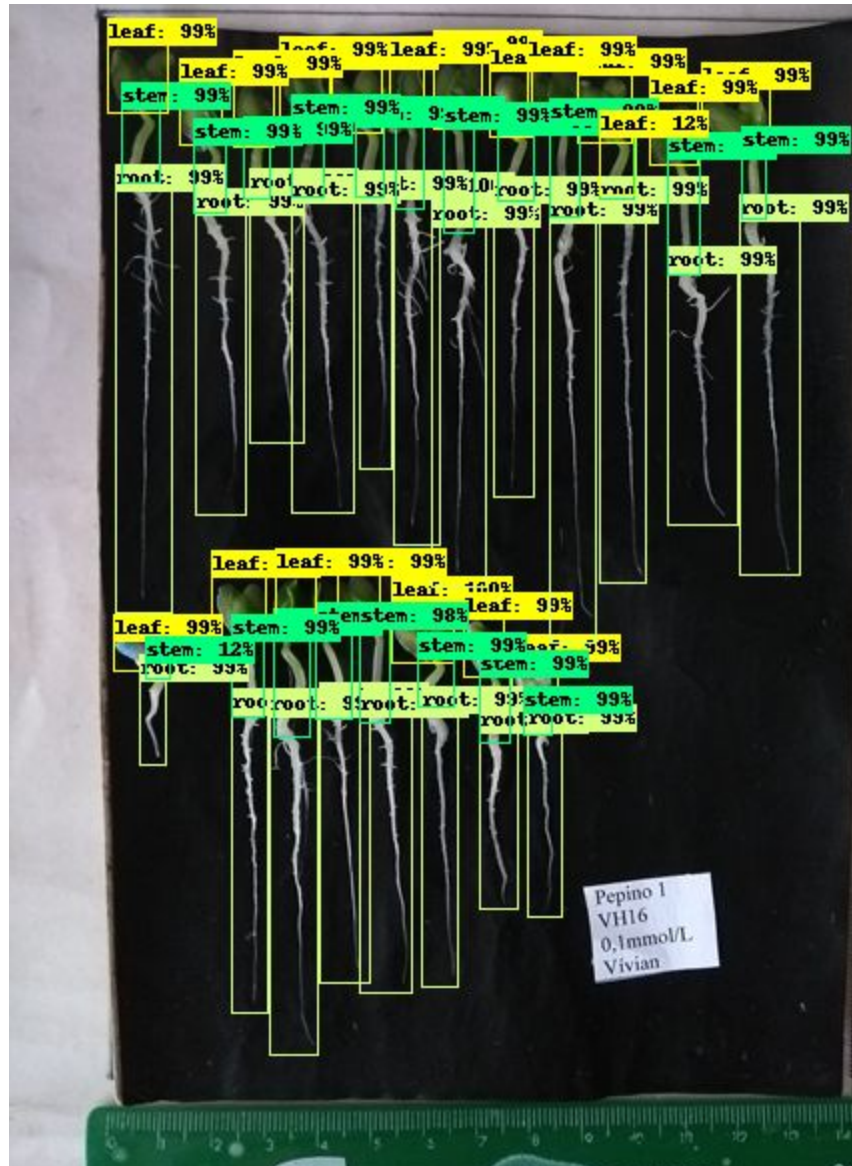
**Figura 4:** Gráficos de perda de erro de (a) classificação e de (b) localização.

Uma vez com os retângulos detectados é possível realizar as medidas das partes das plântulas presentes em uma imagem. Para realizar as medidas das raízes e caules presentes em cada retângulo, pode ser utilizada apenas uma aproximação em relação ao tamanho do retângulo, onde a quantidade de *pixels* do maior lado é multiplicado por um fator de calibração obtido com relação

<sup>6</sup> Disponível em: [https://www.tensorflow.org/tensorboard/get\\_started](https://www.tensorflow.org/tensorboard/get_started)



a régua presente na imagem, semelhante do modo de calibração utilizado pelo *software* PhotoMed. Desse modo é possível criar uma aplicação onde dada uma imagem igual da Figura 5, o *script* abre a imagem com todos os retângulos desenhados sobre cada classe das plântulas e pede para o usuário marcar uma reta de 10 cm sobre a régua para realizar a calibração e poder ser realizadas as medidas de forma automática.



**Figura 5:** Detecção de folhas, caules e raízes realizada pelo modelo Fast R-CNN dado como entrada uma imagem do conjunto de teste.

## 5. Conclusão

Neste trabalho foi apresentado o uso de redes neurais convolucionais profundas para realizar a detecção e medidas de folhas, raízes e caules de plântulas de Pepinos. Esse trabalho se concentrou em criar o conjunto de dados de treinamento e realizar o treinamento usando um modelo de detecção de objetos em imagens de alto desempenho. O modelo escolhido neste trabalho foi o modelo vencedor do desafio COCO 2018, conhecido como Fast R-CNN apresentado em [3]. Embora o modelo tenha apresentado bons resultados para as plântulas de Pepino, seria preciso aumentar o conjunto de dados de treinamento, adicionando mais imagens rotuladas de espécies de plântulas diferentes para possibilitar uma melhor generalização do modelo. Com um conjunto de dados mais amplo, seria possível desenvolver uma aplicação mais completa, com mais recursos capaz de ser utilizada por pesquisadores no lugar do software PhotoMed. Os códigos usados para realizar o treinamento estão disponível em um repositório público do Github<sup>7</sup>, este repositório possui os mesmo arquivos que o repositório do *Tensor Flow models* porém apenas com a parte de detecção de objetos o que torna a replicação do trabalho mais simples.

## Referências

- [1] HUANG, Jonathan *et al.* Speed/accuracy trade-offs for modern convolutional object detectors. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2017. p. 7310-7311.
- [2] LIN, Tsung-Yi *et al.* Microsoft coco: Common objects in context. In: **European conference on computer vision**. Springer, Cham, 2014. p. 740-755.
- [3] GIRSHICK, Ross. Fast r-cnn. In: **Proceedings of the IEEE international conference on computer vision**. 2015. p. 1440-1448.

---

<sup>7</sup> [https://github.com/LucasBraganca/projeto\\_final\\_inf721](https://github.com/LucasBraganca/projeto_final_inf721)