

Reconhecimento de plantas utilizando Redes Neurais Convolucionais em dispositivos móveis

Plants recognition using embedded Convolutional Neural Networks on Mobile devices

RESUMO

Neste trabalho propomos um aplicativo móvel capaz de reconhecer plantas medicinais brasileiras para uso de estudantes universitários que não têm contato prévio com as espécies e também profissionais que atuam em centros de saúde. Descrevemos a geração do banco de dados com base na lista de plantas medicinais e tóxicas comuns do Ministério da Saúde. Também implementamos técnicas de inteligência artificial para realizar a tarefa de reconhecimento, usando uma classe de Redes Neurais Convolucionais (CNN), com foco em diminuir o recurso de computação necessário para executar tarefas de *deep learning* e também otimizar a execução das arquiteturas em dispositivos embarcados e móveis.

PALAVRAS-CHAVE: Classificação de imagens. Redes Neurais Convolucionais. Reconhecimento de plantas.

Sthefanie Monica Premebida

sthefanie@alunos.utfpr.edu.br

Universidade Tecnológica Federal do Paraná,

Ponta Grossa, Paraná, Brasil

Marcella Scoczynski Ribeiro Martins

marcella@utfpr.edu.br

Universidade Tecnológica Federal do Paraná,

Ponta Grossa, Paraná, Brasil

Denise Pechebovicz

pechebovicz@alunos.utfpr.edu.br

Universidade Tecnológica Federal do Paraná,

Ponta Grossa, Paraná, Brasil

Jakson L. Bittencourt

jakson@alunos.utfpr.edu.br

Universidade Tecnológica Federal do Paraná,

Ponta Grossa, Paraná, Brasil

Recebido:

Aprovado:

Direito autoral: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



ABSTRACT

In this work we propose a mobile application capable of recognizing Brazilian medicinal plants to be used by universities, students that have not previous contact with the species and professionals working on health centers. We describe the database generation based on the Brazilian Ministry of Health list of medicinal and common toxic plants. We also implement artificial intelligence techniques to perform the recognition task using a class of Convolutional Neural Networks (CNN) focused on lowering the computation resource necessary to run deep learning tasks and also optimizing the execution of the architectures on embedded and mobile devices.

KEYWORDS: Image classification. Convolutional Neural Networks. Plants Recognition.

3 SAÚDE E BEM-ESTAR



15 VIDA TERRESTRE



1 INTRODUÇÃO

O reconhecimento de plantas tem demonstrado um amplo uso na área científica em campos como biologia, agricultura, medicina e farmácia (ANTONIO; TESSER; MORETTI-PIRES, 2013). Normalmente, aplicativos com essa finalidade de reconhecimento de plantas fornecem informações medicinais. A informação medicinal é um conhecimento acumulado que tem sido passado de geração em geração e também foi provado cientificamente os seus benefícios (SCHULZ; HÄNSEL; TYLER, 2001).

A figura dos curandeiros resistiu ao tempo e às inovações da medicina científica, que perduram até os dias atuais com grande demanda da população. Essa demanda está relacionada à compra de ervas medicinais e produtos feitos para tratamento médico alternativo para vários problemas de saúde, como por exemplo, chás medicinais responsáveis por efeitos terapêuticos (PEDRO et al., 2016).

Uma das coisas que se faz necessário aprender sobre esse tipo de tratamento são os ingredientes ativos das ervas, como indicações terapêuticas, contra-indicações, efeitos de overdose e reações alérgicas, para orientar os usuários sobre possíveis interações químicas ou envenenamento por medicamentos, além dos procedimentos de limpeza, armazenamento e riscos envolvidos com a identificação da planta errada.

Além do uso caseiro das plantas medicinais, em 2009 o Sistema Único de Saúde (SUS) ampliou a lista de fitoterápicos disponíveis em assistência farmacêutica básica em todo o país. O Ministério da Saúde também espera que com o programa os estados sejam encorajados a fornecer tratamentos com esses medicamentos. Atualmente, 12 estados brasileiros ofertam o tratamento e o programa considera 71 plantas como medicinais.

Existem algumas plantas com ingredientes ativos que causam intoxicação em humanos e animais, conhecidas como plantas tóxicas. As plantas tóxicas incluem todos os vegetais que o contato, inalação ou ingestão resultam em danos à saúde e podem levar a morte em alguns casos. Muitas das plantas tóxicas são consideradas ornamentais, presentes em diversos locais ao nosso redor, facilitando o risco de intoxicação, além disso, as toxinas presentes nas plantas podem afetar diretamente a produção animal, sendo capaz de prejudicar o agronegócio.

De acordo com os dados do Sistema Nacional de Informações Tóxico-Farmacológicas (SINITOX), em 2010 o Brasil registrou 1123 casos de intoxicações por plantas, 330 deles no Sul do país (CAMPOS et al., 2016). Portanto ao otimizar o reconhecimento das plantas, incentiva-se o estudo das aplicações da planta e conhecimento das suas indicações e contra indicações, e também pode fornecer melhores informações sobre as plantas tóxicas.

O principal objetivo desse trabalho é implementar um modelo de rede neural capaz de reconhecer as plantas medicinais brasileiras que podem ser usadas pelo SUS em tratamentos fitoterápicos. O modelo será incorporado em um dispositivo móvel, tornando o aplicativo portátil, rápido e preciso.

A ferramenta pode ser usada por estudantes que não tem conhecimento prévio das espécies, profissionais que trabalham em Postos de Saúde e fazendeiros fornecendo um serviço rápido e fácil evitando a intoxicação dos animais por plantas tóxicas. Assim, o aplicativo é um meio fácil e acessível de reconhecimento de plantas, contribuindo nas áreas da agricultura, saúde e conhecimento popular.

2 MATERIAL E MÉTODOS

2.1 Pesquisas relacionadas

As redes neurais têm mostrado resultados precisos na classificação de plantas, principalmente na análise de folhas (Wu et al., 2007), classificando 32 tipos de plantas com precisão acima de 90%. Normalmente os principais objetivos da aplicação dessas técnicas de classificação de plantas estão relacionados à agricultura, para construir máquinas autônomas que reconheçam e removam ervas daninhas (ADHIKARI; YANG; KIM, 2019), detecção de doenças de cultivo (MOHANTY; HUGHES; SALATHÉ, 2016) e classificação de grãos (VISEN et al., 2002), mas pode ser estendido para áreas como farmácia, medicina e botânica. Uma técnica interessante foi aplicada para o reconhecimento de plantas, a proposta foi um modelo CNN aprender automaticamente as características de representação para categorias de plantas (LEE et al.,). Esse método consiste no reconhecimento de 44 classes, coletadas no Royal Botanic Gardens e teve uma performance de 99.6%.

2.2 Redes Neurais Convolucionais

Redes Neurais Convolucionais ou em inglês *Convolutional Neural Networks* (CNN) se tornaram muito populares para classificar imagens após a Alexnet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) ganhar o Imagenet Challenge (ILSVRC) em 2012 e a tendência seguida por essa área da inteligência artificial era criar arquiteturas cada vez mais complexas, robustas e precisas, porém essas melhorias tem um custo. Para executar essas poderosas arquiteturas é necessário um computador com processamento gráfico (GPU) de ponta. Nesse trabalho, usamos uma abordagem diferente aplicada à CNN, que em vez de focar em precisão extremamente alta, as arquiteturas usadas aqui estão focadas na implementação eficiente em dispositivos móveis. As redes usadas são chamadas Mobilenets (HOWARD et al., 2017).

Essas redes usam convoluções separáveis em profundidade para tornar a aplicação mais leve, reduzindo a quantidade de esforço computacional para obter a previsão. Esse tipo de convolução separável é especialmente eficiente quando a saída dos canais de uma camada são maiores do que os canais de entrada. A convolução de imagens é semelhante à convolução 2D, mas multiplica o valor da cor do pixel e de seus pixels vizinhos por uma matriz. Essas matrizes são geralmente pequenas (3x3 ou 5x5) e chamadas de kernels (LUDWIG, 2013). Existem dois tipos de convolução de imagens, as normais e as separáveis em profundidade. Os Mobilenets têm dois hiperparâmetros que podem diminuir ainda mais o custo computacional:

- α , multiplicador de largura que pode diluir uma rede uniformemente, escolhendo cada camada. Sendo $\alpha=1$ a versão padrão das redes, redes com valores inferiores são as reduzidas Mobilenets.
- ρ , é um multiplicador de resolução que é aplicado à imagem de entrada e a representação interna de cada camada, geralmente esse parâmetro é definido implicitamente quando a resolução é definida, o valor típico é 224 mas pode ser tão baixo quanto 128.

2.3 Base de dados

Em nossa pesquisa consideramos um banco de dados com 10.162 imagens, incluindo diferentes estados de vida da planta, imagens com origens e ambientes diferentes. Para gerar essa base de dados, usamos o programa de download do Google Imagens (VASA,), permitindo o download de imagens com base em palavras-chave ou frases-chave.

Os rótulos usados para a pesquisa são baseados no Ministério da Saúde¹, os principais são nomes científicos de gêneros/espécies de plantas. Para a categoria de tóxicos consideramos uma lista da Fundação Oswaldo Cruz (FIOCRUZ)², uma instituição científica de pesquisa e desenvolvimento em ciências biológicas localizadas no Rio de Janeiro e considerada uma das principais instituições de pesquisa em saúde pública do mundo.

O conjunto de dados é composto por 72 categorias, sendo as 71 originais do Ministério da Saúde mais a categoria de plantas tóxicas.

2.4 Aumento de dados

Nesse trabalho o número de imagens de treinamento representa 80% do conjunto de dados originais, para cada categoria. A categoria com maior número de amostras é a rotulada como "tóxicas", que agrega todas as amostras tóxicas do banco de dados, tendo 1.018 amostras no treinamento conjunto, mas outras categorias tem menos amostras, sendo o caso da *Lippia Sidoides* com apenas 25.

Apesar do aumento no número de amostras, o objetivo principal é melhorar a generalização do modelo, gerando "novos" dados de treinamento. Com mais dados, o modelo pode treinar com entradas ligeiramente modificadas, podendo treinar com recursos mais robustos. Normalmente a aplicação desse método melhora os valores de precisão da rede.

A tarefa de aumento de dados foi implementada usando a biblioteca *Python Augmentor* (BLOICE; STOCKER; HOLZINGER, 2017) que inclui alguns métodos de aumento em nosso conjunto de dados, como rotação, recortes, espelhamento, zooms aleatórios, distorções aleatórias e enviesamento. Após o aumento de dados, o número total de imagens aumentou para 151.128 com 5230 na categoria tóxicos e um número médio de 2.000 nas outras categorias.

2.5 Arquitetura de Rede

Uma versão mais recente, Mobilenet v2 (SANDLER et al., 2018), será usada no desenvolvimento da pesquisa. A versão usada melhora a tecnologia *state-of-the-art*, obtendo melhores e mais rápidos resultados usando resíduos e gargalos lineares. Usaremos $\alpha=1$ e $\rho=224$ como pesos iniciais de treinamento. Um classificador foi adicionado ao modelo e as camadas convolucionais foram congeladas, utilizando apenas os pesos do classificador. O modelo completo possui 155 camadas e estão congeladas, mas para que o desempenho aumente, foi feito um ajuste fino para treinar as camadas superiores, descongelando as camadas 101 a 155 e retreinando a rede. Diferentes tamanhos de épocas são usados no treinamento e no ajuste fino (5, 10, 20, 30) e a comparação pode ser vista nos resultados. O resultado final é treinado para 20 épocas e têm o ajuste fino com 20 épocas.

2.6 Estrutura

O objetivo do trabalho é utilizar esse modelo de rede em um dispositivo móvel, para que isso seja feito, utilizamos a estrutura TensorFlow Lite, cujas ferramentas ajudam a desenvolver redes neurais para serem aplicadas em dispositivos móveis, incorporados e IoT. Portanto, a rede convertida pode ser implantada em um celular, e podemos verificar a latência e a precisão da aplicação.

¹ Maiores informações em <http://bit.do/fa2av>

² A lista completa está disponível em <http://www.fiocruz.br/biosseguranca/Bis/virtual/%20tour/hipertextos/up2/plantas-toxicas.htm>

3 RESULTADOS E DISCUSSÃO

Nosso conjunto de dados é dividido em treinamento e validação, usando 80% e 20% das imagens, respectivamente. O processo de treinamento com os dados originais, sem aumento, e Mobilenet v1 apresenta resultados ruins. A perda é sempre maior que 1 e a precisão está entre 0,4 e 0,5. A perda representa como muitas amostras (porcentagem) que o algoritmo estimou errado, e é um número entre 0 e 1. Por exemplo, uma perda de 0,25 representa que a técnica acerta 3 em 4 vezes e a precisão representa o quão certo o algoritmo está.

Os resultados usando Mobilenet v2 são semelhantes ao Mobilenet v1, diminuindo ligeiramente a perda e aumentando a precisão. Para melhorar o desempenho da rede, o conjunto de dados foi aumentado e a rede treinada por 10 épocas, em seguida, ajustada para 5 épocas.

Como os resultados parecem estar melhorando à medida que aumenta-se o número de épocas, a rede é retreinada por 30 épocas em seguida, ajustado para 10 épocas. Os resultados podem ser vistos na Figura 1. Para analisar o comportamento do treinamento da rede, o número de épocas de treinamento são reduzidas para 20 e as épocas de ajuste fino aumentaram para 20. Esses resultados podem ser vistos na Figura 2.

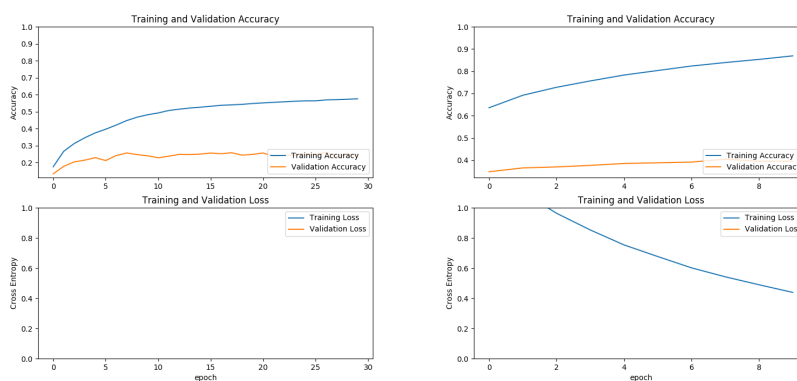


Figura - 1: Treinamento para 30 épocas e o ajuste fino para 10 épocas, respectivamente

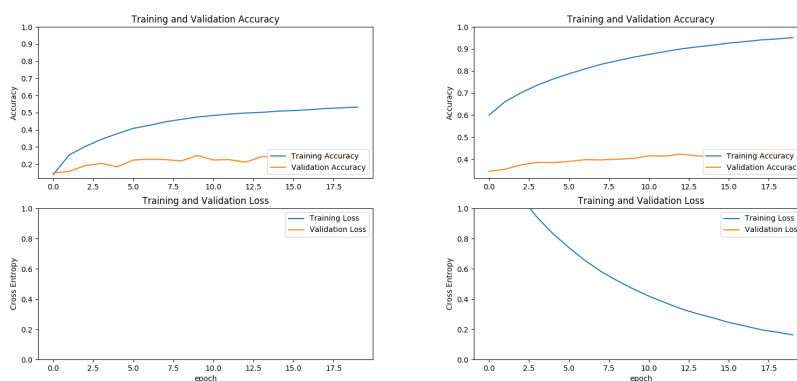


Figura - 2: Treinamento para 20 épocas e o ajuste fino para 20 épocas, respectivamente

A precisão geral da validação é 0,4254 e os 5 primeiros 0,685. Os pesos obtidos nesta última etapa de treinamento e ajuste fino são usados para o modelo embutido em um tele-

fone móvel. Os primeiros testes foram realizados em um Xiaomi Redmi Note7 e o aplicativo rodou sem problemas, em tempo real.

4 CONCLUSÃO

Nesse trabalho implementamos uma CNN para reconhecer as plantas medicinais brasileiras por meio de um dispositivo móvel. A arquitetura da rede foi baseada na Mobilenet v2, sobre um banco de dados composto por 10.162 imagens baixadas do Google Imagens, podendo classificar 71 categorias de plantas de acordo com o Ministério da Saúde e mais uma de plantas tóxicas. Além disso, métodos de aumento de banco de dados foram utilizados para aumentar a generalização do conjunto de treinamento. A estrutura foi desenvolvida usando o TensorFlow e, para ajustar o modelo de rede a um dispositivo móvel, a ferramenta TensorFlow Lite Converter foi aplicada. Nosso modelo incorporado apresentou 0,94 de precisão e 0,18 de perda, valores significativamente maiores quando comparados a outras arquiteturas, mas como esses valores foram obtidos durante o teste, a validação mostrou uma precisão dos 5 principais de 0,685 o que implica que nosso aplicativo ainda precisa ser melhorado antes de ser usado por outros pesquisadores.

Em trabalhos futuros, planejamos expandir e refinar nosso banco de dados usando nomes comuns de plantas como palavras-chave. O banco de dados também pode ser ampliado para imagens e quantidade e qualidade de gêneros / espécies por meio de parcerias com instituições de pesquisa. Além disso, para permitir que o usuário busque mais informações sobre a planta identificada, criaremos uma interface móvel capaz de vincular a classificação às informações do Ministério da Saúde e de outras instituições de pesquisa.

Para melhorar o desempenho da rede neural, outros métodos de aumento podem ser aplicados, bem como outras configurações de treinamento e diferentes otimizadores, incluindo uma análise de desempenho para entender melhor o comportamento da rede.

5 AGRADECIMENTOS

Os autores agradecem à Fundação Araucária (FA), pela bolsa de Iniciação Tecnológica e Inovação (Edital 02/2019 - PROPPG - PROGRAMA INSTITUCIONAL DE INICIAÇÃO CIENTÍFICA DA UTFPR), e também à participação no Programa de Voluntariado em Iniciação Científica e Tecnológica (Edital PROPPG – PIVICT 2019/2020) concedida ao acadêmico .

REFERÊNCIAS

- ADHIKARI, S. P.; YANG, H.; KIM, H. Learning semantic graphics using convolutional encoder-decoder network for autonomous weeding in paddy field. **Frontiers in Plant Science**, Frontiers, v. 10, p. 1404, 2019.
- ANTONIO, G. D.; TESSER, C. D.; MORETTI-PIRES, R. O. Contributions of medicinal plants to care and health promotion in primary healthcare. **Interface-Comunicação, Saúde, Educação**, SciELO Brasil, v. 17, n. 46, p. 615–633, 2013.
- BLOICE, M. D.; STOCKER, C.; HOLZINGER, A. Augmentor: an image augmentation library for machine learning. **arXiv preprint arXiv:1708.04680**, 2017.
- CAMPOS, S. et al. Toxicidade de espécies vegetais. **Revista Brasileira de Plantas Medicinais**, v. 18, n. 1, p. 373–382, 2016.

- HOWARD, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. **CoRR**, abs/1704.04861, 2017.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Commun. ACM**, v. 60, p. 84–90, 2012.
- LEE, S. H. et al. **Deep-Plant: Plant identification with Convolutional Neural Networks**.
- LUDWIG, J. Image convolution. **Portland State University**, 2013.
- MOHANTY, S. P.; HUGHES, D. P.; SALATHÉ, M. Using deep learning for image-based plant disease detection. **Frontiers in plant science**, Frontiers, v. 7, p. 1419, 2016.
- PEDRO, F. et al. Composição centesimal e mineral de plantas medicinais comercializadas no mercado do porto de cuiabá, mato grosso, brasil. **Rev. Bras. Pl. Med**, v. 18, n. 1 supl I, p. 297–306, 2016.
- SANDLER, M. et al. Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation. **CoRR**, abs/1801.04381, 2018.
- SCHULZ, V.; HÄNSEL, R.; TYLER, V. E. **Rational phytotherapy: a physician's guide to herbal medicine**. [S.l.]: Psychology Press, 2001.
- VASA, H. **Google Images Download**. Disponível em: <<https://google-images-download.readthedocs.io/en/latest/index.html>>.
- VISEN, N. et al. Ae—automation and emerging technologies: specialist neural networks for cereal grain classification. **Biosystems Engineering**, Elsevier, v. 82, n. 2, p. 151–159, 2002.
- Wu, S. G. et al. A leaf recognition algorithm for plant classification using probabilistic neural network. In: **2007 IEEE International Symposium on Signal Processing and Information Technology**. [S.l.: s.n.], 2007. p. 11–16.