



UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Autorizada pelo Decreto Federal nº 77.496 de 27/04/76

Recredenciamento pelo Decreto nº 17.228 de 25/11/2016



PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COORDENAÇÃO DE INICIAÇÃO CIENTÍFICA

XXIV SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS **SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2020**

Técnicas de Aumento de Dados Para Reconhecimento de Espécies de Plantas baseado em Deep Learning

Luciano Araújo Dourado Filho¹ e Rodrigo Tripodi Calumby²

¹Estagiário PEVIC, Graduando em Engenharia Da Computação, UEFS, e-mail: lucianoadfilho@ecomp.uefs.br

²Orientador, Departamento de Ciências Exatas, UEFS, e-mail: rtcalumby@uefs.br

PALAVRAS-CHAVE: Data Augmentation; Reconhecimento de Plantas; Deep Learning.

INTRODUÇÃO

O esforço recente de milhares de botanistas e pesquisadores ao redor do mundo de catalogar e digitalizar milhares de espécies de plantas tem facilitado o acesso a grandes volumes de dados e crescentes em número de espécies (Soltis, 2017; Beech et al, 2017). Paralelamente, a exploração de arquiteturas paralelas por meio de técnicas de alto desempenho tem viabilizado aplicações de *Deep Learning* (DL), como Redes Neurais Convolucionais (CNN), que por muito tempo foram consideradas inviáveis, mas se tornaram amplamente populares devido à sua capacidade de aprender a extrair características a partir dos próprios dados, (Lee et al. 2017). Isso possibilitou ultrapassar obstáculos previamente associados a algoritmos de extração de características feitos à mão para classificação baseada em um órgãos individuais das plantas (Pawara et al, 2017; Ghazi et al, 2017), a até mesmo superar especialistas humanos em desafios de classificação abrangendo grandes quantidades de espécies, (Goëau et al, 2018; Bonnet et al, 2018). Apesar disso, grande parte dessas coleções apresentam uma distribuição irregular de imagens por espécie (Picek et al, 2019), o que motiva os pesquisadores a utilizarem técnicas de aumento de dados para contornar estes obstáculos através da ampliação destes conjuntos por meio da aplicação de transformações sobre os conjuntos de dados originais. Isso tem se mostrado eficaz não apenas em contornar desbalanceamento, mas para aumentar a capacidade de generalização de modelos de aprendizado profundo e introduzir invariância (Ghazi et al, 2017). Assim, esse trabalho buscou avaliar o impacto da aplicação de múltiplas técnicas de aumento sobre modelos de aprendizado profundo com objetivo de realizar classificação de imagens de plantas, a fim de tentar estabelecer relações entre as estratégias de aumento empregadas e as consequências sobre a eficácia desses modelos. Acredita-se que avaliações como essas podem ajudar a endereçar o desenvolvimento de sistemas para reconhecimento de espécies de plantas e estudos de DL baseados em modelos treinados com conjuntos desbalanceados.

METODOLOGIA

Para realizar uma análise comparativa dentre as principais técnicas de aumentos de dados foi realizado um levantamento bibliográfico que permitiu identificar as técnicas mais apropriadas para aplicação no contexto de reconhecimento de espécies de plantas. Dentre esses, os métodos de aumento de dados mais comuns podem ser classificados de acordo com dois tipos de transformações: geométricas, em que a constituição geométrica das imagem são alteradas; e fotométricas, nas quais a geometria das imagens são preservadas e seus espaços de cores são modificados. Exemplos de técnicas geométricas comumente utilizadas na literatura incluem, *Crop* (recorte), *Flip* (espelhamento), *Translate* (translação) e *Rotate* (rotacionar) e as múltiplas combinações destas técnicas, (Ghazi et al, 2017; Picek et al, 2019). Já as transformações fotométricas incluem, principalmente, mudanças de brilho, ajuste de foco, mudança para escala de cinza, entre outras.

Este trabalho considerou a aplicação de transformações geométricas e algumas de suas combinações. Para isso, utilizamos a base de dados do Plant CLEF 2013 (Goëau et al, 2013), que dispõe de um total de 26.077 imagens de 250 espécies de plantas da flora Francesa, divididas em 20.985 imagens de treino e 5.092 imagens exclusivas para teste. As imagens possuem diferentes tipos de fundos e ângulos de captura de diferentes órgãos das plantas, como folhas, flores, frutos, galhos e troncos. Para a extração de características e classificação, utilizamos a arquitetura de rede de *Deep Learning* ResNet50 (He et al, 2013). A Tabela 1 apresenta as transformações aplicadas, suas combinações, o fator de aumento e o número de imagens resultante.

A Figura 1 ilustra a aplicação de duas técnicas de aumento. Do lado esquerdo, a imagem **a** é transformada através da técnica de rotação de forma isolada, originando como produto duas imagens **b** e **c**, rotacionadas R graus nos sentidos horário e anti-horário. Do outro lado, a combinação entre as técnicas de espelhamento vertical e rotação dão origem a duas imagens **b** e **c**, verticalmente invertidas e rotacionadas R graus no sentido horário e anti-horário.

Para realizar a avaliação sobre os modelos de aprendizado foi adotado o protocolo de amostragem aleatória estratificada, com uma divisão de 80% dos dados para treino e 20% para validação, seguida por uma etapa extra de avaliação com o conjunto de testes. As medidas de avaliação utilizadas foram acurácia e a função de custo ou perda (*Loss*), neste caso a Entropia Cruzada Categórica. A política de atualização de pesos para treino da rede foi o Gradiente Descendente (*mini batch*) e o otimizador utilizado foi o *Adam*.

As configurações paramétricas das técnicas de aumento foram: *Rotate*, com 30 graus nos

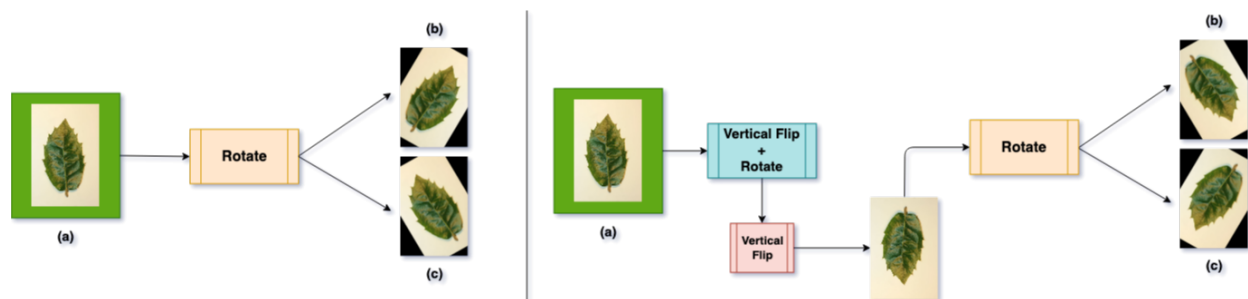


Figura 1: Exemplo de transformações geométricas.

sentidos horário e anti-horário; *Crop*: divisão da imagem original em quatro partes dos seus cantos mais um corte central compreendendo 50% da sua cópia original; *Translate*: quatro translações diferentes, a partir de deslocamentos equivalentes à 20% da largura da imagem no eixo horizontal, e 20% da altura da imagem no eixo vertical; e *Vertical Flip* (espelhamento vertical). Os experimentos envolvendo as combinações das técnicas anteriores basearam-se nos mesmos parâmetros.

Utilizando a rede ResNet50, pré-treinada sobre o conjunto de dados da ImageNet, realizou-se o ajuste fino sobre a última camada de pesos e o último bloco de convoluções, que incluiu o ajuste de 5.000.000 de parâmetros de um total de aproximadamente 24.000.000. Foram avaliados de forma comparativa, o treinamento inicial (*Baseline*) com o conjunto de dados original sem transformações e os treinamentos envolvendo o conjunto original embaralhado com os conjuntos transformados de acordo com os experimentos citados na Tabela 1. Assim, os modelos foram treinados por 100.000 iterações, com uma taxa de aprendizado de 2×10^{-6} e *batch* de tamanho 64. Em seguida, a partir dos conjuntos de pesos obtidos do treinamento realizou-se uma etapa extra de avaliação com aumento na fase de testes a partir do critério de decisão de média das probabilidades de classificação.

RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta os resultados obtidos a partir das etapas de validação e testes com as configurações de treinamento propostas. Pode-se observar que os experimentos envolvendo a aplicação de um único tipo de transformação em sua maioria apresentaram desempenho superior ao *Baseline* em relação aos experimentos que envolveram a aplicação de múltiplas transformações de forma combinada durante a etapa de testes, com exceção da técnica *Vertical Flip + Crop*, que apresentou desempenho superior à aplicação exclusiva de *Rotate*.

De forma geral, observou-se nos resultados obtidos o potencial de melhoria de eficácia a partir do aumento de dados com métodos específicos, em especial com aqueles aplicados individualmente. Contudo, acredita-se que o treinamento baseado em uma quantidade fixa de iterações, possa ter desfavorecido os métodos treinados sobre os conjuntos de dados maiores (com diferentes tipos de aumento), já que o modelo acabou sendo exposto a todo o conjunto de dados menos vezes em relação ao aumento com apenas um único método. Isso sugere que a abordagem de treinamento com limite de iterações possa não ser adequada para comparação de métodos de aumento, gerando, assim, a necessidade de uma mudança de abordagem, por exemplo, com treinamento baseado em uma determinada quantidade de vezes que o modelo é exposto a todo o conjunto de dados. Assim,

Experimento	Fator de Aumento	Quantidade de Imagens	Acurácia de Validação	Acurácia de Teste
Baseline	-	20985	0,7756	0,4372
Crop	5X	125910	0,7662	0,4595
Vertical Flip	1X	41970	0,8841	0,4570
Rotate	2X	62955	0,8945	0,4264
Translate	4X	104925	0,9591	0,4519
Vertical Flip + Crop	5X	125910	0,7577	0,4295
Vertical Flip + Rotate	2X	62955	0,8169	0,4089
Translate + Crop	20X	440685	0,4029	0,3291
Translate + Rotate	8X	188865	0,4746	0,2188

Tabela 1: Configurações experimentais e desempenho na etapa de testes.

novos experimentos nesse sentido são necessários para validação. Observou-se também grande variação de eficácia entre a etapa de validação e teste. Considerando o processo de treinamento de modelos preditivos, este fenômeno pode ter diversas origens e possibilidades de tratamento e que poderão ser alvo de novos trabalhos.

CONSIDERAÇÕES FINAIS

Os resultados obtidos possibilitaram demonstrar a eficácia das técnicas de aumento de dados no que diz respeito à melhoria do acerto de classificação de modelos de aprendizado baseados em Redes Neurais Profundas. Além disso, permitiram inferir dentre as abordagens utilizadas e suas combinações, as que apresentaram melhor desempenho, além de fornecer intuições sobre formas mais eficazes de treinamento e comparação de métodos avaliados sobre essas técnicas. Dessa forma, destaca-se a necessidade de realização de trabalhos futuros compreendendo otimizações do processo de treinamento e abordagens de comparações mais adequadas para combinações de múltiplos métodos de aumento, possivelmente incluindo também o emprego de transformações fotométricas.

REFERÊNCIAS

- SOLTIS, P.S. 2017. Digitization of herbaria enables novel research. *AMJBOT*, 104(9), pp.1281-1284.
- BEECH, E; RIVERS, M; OLDFIELD, S; SMITH, P.P. 2017. GlobalTreeSearch: The first complete global database of tree species and country distributions. *J SUSTAIN FOREST*, 36(5), pp.454-489.
- LEE, S,H; CHAN, C,S; MAYO, S.J; REMAGNINO, P. 2017. How deep learning extracts and learns leaf features for plant classification. *Pattern Recognition*, 71, pp.1-13.
- PAWARA, P; OKAFOR, E; SURINTA, O.; SCHOMAKER, L; Wiering, M. 2017, February. Comparing local descriptors and bags of visual words to deep convolutional neural networks for plant recognition. *INT C PATT RECOG (Vol. 2, pp. 479-486)*. SCITEPRESS.
- Ghazi, M.M., Yanikoglu, B. and Aptoula, E., 2017. Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing*, 235, pp.228-235.
- PICEK, L; SULC, M; MATAS, J. 2019. Recognition of the Amazonian flora by InceptionNetworks with Test-time Class Prior Estimation. In *CLEF (Working Notes)*.
- GOËAU, H; BONNET, P; Joly, A. 2018. Overview of ExpertLifeCLEF 2018: how far automated identification systems are from the best experts?. In *CLEF (Working Notes)*.
- BONNET, P.; GOËAU, H; HANG, S.T.; LASSECK, M; SULC, M; MALÉCOT, V; JAUZEIN, P; MELET, J.C; YOU, C; JOLY, A., 2018. Plant identification: experts vs. machines in the era of deep learning. In *MULTIMED TOOLS APPL for Env & Bio Info* (pp. 131-149). Springer, Cham.
- GOËAU, H; JOLY, A; BONNET, P; BAKIC, V; BARTHÉLÉMY, D.; BOUJEMAA, N; Molino, J.F. 2013, October. The imageclef plant identification task 2013. In *Proceedings of the 2nd ACM international workshop on MAED* (pp. 23-28).
- HE, K; ZHANG, X; REN, S; SUN, J. 2016. Deep residual learning for image recognition. *PROC CVPR IEEE* (pp. 770-778).