



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

XXV SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS **SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2021**

Identificação de Espécies de Plantas em Imagens com Aprendizado Profundo **Baseado em Taxonomia**

Luciano A. Dourado Filho¹; Angelo Conrado Loula²; Rodrigo Tripodi Calumby³

1. Bolsista PIBIC/CNPq, Graduando em Engenharia da Computação, UEFS, e-mail: lucianoadfilho@ecomp.uefs.br

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

3. Coorientador, Departamento de Ciências Exatas, UEFS, e-mail: rtcalumby@uefs.br

PALAVRAS-CHAVE: Deep Learning; Plantas; Taxonomia.

INTRODUÇÃO

Considerando a flora mundial, estimativas atuais apontam para a existência de cerca de 420.000 espécies de angiospermas. Devido às especificidades das espécies e variações fenotípicas, sua identificação é considerada uma tarefa muito difícil para o cidadão comum e até mesmo para especialistas em botânica [1,2]. Diante disso, diversas iniciativas têm fomentado o desenvolvimento de trabalhos envolvendo a classificação automática de plantas no nível de espécies [2,3] através de redes neurais convolucionais profundas (RNCP), a ponto de possibilitarem o desenvolvimento de modelos de reconhecimento capazes de superar especialistas humanos. Apesar disso, esses estudos ainda estão longe de atingir representatividade significativa em relação à quantidade de espécies estimada ao redor do mundo. Nesse sentido, além de promover a aquisição de bases de dados com maior cobertura de espécies e espécimes, estudar novas estratégias para o desenvolvimento de modelos de RNCP para identificação de espécies de plantas é de fundamental importância para a obtenção de sistemas de reconhecimento mais eficazes para auxiliar em tarefas que envolvem identificação de plantas. Nesse sentido, apesar de na literatura os trabalhos serem majoritariamente voltados para identificação a nível de espécie, trabalhos recentes [4,5] demonstraram que a associação de outras informações taxonômicas, como gênero e família podem auxiliar na potencialização da eficácia de identificação de espécies com RNCP.

Em [4], os autores avaliaram arquiteturas de RNCP para classificação de em diferentes níveis taxonômicos utilizando imagens de herbário. Para isso, eles propuseram uma arquitetura de RNCP (denominada *TaxonNet*) que realiza classificação multirrotulo de cada nível taxonômico (família, gênero e espécie) de maneira hierárquica e compararam sua eficácia em relação à outras duas arquiteturas: *Multi-Task Classification Model* (MCM) [6], que corresponde à uma RNCP adaptada para realizar classificação multirrotulo (isto é, com múltiplos classificadores) e *Flat Model*: RNCP convencional para classificação multiclasse. A partir disso, além de demonstrarem os modelos mais

eficazes a respeito dos níveis taxonômicos, os autores puderam constatar que informações de níveis taxonômicos superiores podem corroborar no processo de identificação de espécies.

Apesar de alcançarem resultados promissores através de imagens de herbário, esses modelos apresentam aplicabilidade restrita ao contexto de herbário, ao passo que a utilização de imagens *in natura* possibilita um conjunto mais amplo de aplicações em diversas áreas além da botânica, como farmacologia, agricultura, engenharia ambiental, etc. Por conta disso, neste trabalho comparamos a performance das estratégias MCM e *Flat Model* para realizar classificação à nível de espécie, porém utilizando um conjunto de dados com imagens *in natura*. Além disso, propomos uma arquitetura de RNCP para realizar a fusão de características da hierarquia taxonômica (gênero e família).

METODOLOGIA

Para realizar extração de características e a identificação da espécie (multiclasse), utilizamos a arquitetura ResNet50 [7] (*Flat model*). Para implementar o modelo MCM foram acoplados à ResNet50, 3 classificadores neurais com ativação *Softmax* que permitiram realizar o treinamento para identificação simultânea (multirrótulo) de espécie, gênero e família. Já para desenvolver a arquitetura proposta, denominada *MCM_Fusion*, treinamos um submodelo MCM para realizar classificação simultânea nos níveis de gênero e família. Em seguida, acoplou-se as respectivas camadas pré-treinadas (com pesos congelados) como extratores de características ao modelo proposto (*MCM_Fusion*). Isso possibilita concatenar as características extraídas para identificação de gênero e família, às características que aprenderá a extrair para identificar espécies. A intuição deste processo é que no *MCM_Fusion* o modelo possa utilizar o conhecimento prévio (características discriminativas de gênero e família) durante o aprendizado para reconhecimento de espécies. Para isso utilizamos o conjunto de dados do PlantCLEF 2015¹, que dispõe de aproximadamente 115 mil imagens distribuídas entre 1000 espécies, 516 gêneros e 124 famílias, das quais cerca de 21 mil imagens são destinadas exclusivamente para avaliação de eficácia.

Considerando mitigar o desbalanceamento de classes (espécies), introdução de invariâncias e outros benefícios da utilização de aumento de dados, neste trabalho utilizou-se um esquema robusto de aumento de dados. Para isso, cada imagem do conjunto de dados original foi submetida a um processo que consistiu na seguinte heurística: divisão em 9 quadrantes, dos quais se escolhem aleatoriamente 5 recortes (*random crop*), aplicação de um recorte central compreendendo 50% das dimensões da imagem original (*central crop*) e aplicação de rotação (*random rotate*) entre $\pm 90^\circ$. A partir disso é possível obter um conjunto de 8 imagens (imagem original + 5 recortes + recorte central + original rotacionada) sobre o qual aplica-se individualmente uma translação aleatória (*random translate*) em cada uma das imagens. Por fim, com a obtenção de um conjunto de 16 imagens a partir da imagem original, aplica-se um

¹ Mais informações: <https://www.imageclef.org/lifeclef/2015/plant>

espelhamento vertical sobre cada uma das imagens do conjunto, permitindo obter 32 imagens para cada uma das imagens do conjunto de dados.

Dessa forma os experimentos foram conduzidos considerando o protocolo de avaliação experimental de amostragem aleatória estratificada com uma divisão do conjunto de treino em 70% para treinamento e 30% para validação. Com isso, para fins de comparabilidade, as redes foram treinadas sob as mesmas configurações paramétricas (n. épocas, *batch*, otimizador, número de passos por época e função de custo). Dessa forma, o treinamento foi realizado até a convergência, com aplicação de aumento de dados apenas sobre o conjunto de treinamento e utilizando o conjunto testes independente. Durante o treinamento, a função de custo utilizada foi a *Categorical Crossentropy*, o otimizador *Adam* e medida acurácia. Já na fase de testes, computou-se a *Micro-F1* (média harmônica entre *precision* e *recall*) que considera pesos iguais para as diferentes classes.

RESULTADOS E DISCUSSÃO

No processo de avaliação experimental, as redes foram treinadas para realizar a identificação de espécies de plantas a partir do conjunto de testes. Como apresentado na Figura 1, de maneira similar ao apresentado em [4], o modelo MCM possibilitou atingir maior eficácia em termos de Micro-F1 (0,4566), em relação aos outros modelos avaliados. Apesar disso, o modelo *MCM_Fusion*, que incorpora as características extraídas (dos níveis gênero e família) para classificação de espécies e o modelo *Flat* (ResNet50), apresentaram resultados próximos ao MCM, com Micro-F1 correspondente a 0,4464 e 0,4370, respectivamente.

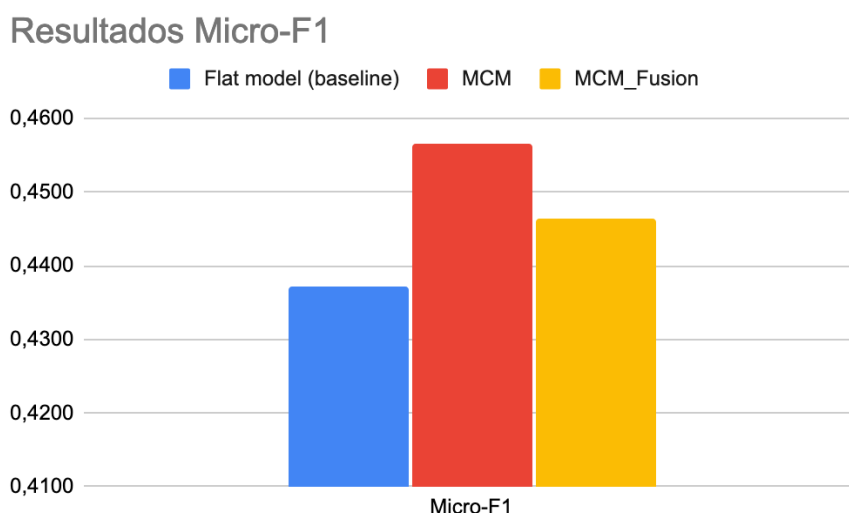


Figura 1. Resultados obtidos de Micro-F1 sobre o conjunto de testes para cada arquitetura avaliada.

Tendo em vista as dificuldades por trás do processo de otimização de redes neurais convolucionais em termos de custo computacional, avaliar diversas configurações paramétricas, principalmente envolvendo grandes quantidades de dados é bastante

custoso. Diante disso, poucos experimentos puderam ser conduzidos no sentido de buscar melhores configurações de hiperparâmetros para as redes avaliadas. Apesar disso, levando em consideração o cenário de classificação, em termos de quantidade de classes (espécies) avaliadas, heterogeneidade das imagens e desbalanceamento entre as classes, os resultados obtidos foram promissores. Além de ajustes de configurações de hiperparâmetros, outro aspecto que pode estar sujeito a modificação para obtenção de melhores resultados refere-se aos métodos de aumento de dados. Os métodos aplicados, com base em [4], possibilitaram ampliar a parte do conjunto de dados relativa à fase de treinamento em 32 vezes, totalizando em aproximadamente 2.055.000 imagens. Neste sentido, diversas manipulações paramétricas e até mesmo heurísticas podem ser aplicadas a fim de obter configurações que permitam obtenção de maior eficácia com maior eficiência possível. Dessa forma, acreditamos que a realização de ajustes finos como os mencionados anteriormente possam evidenciar melhor os contrastes entre essas arquiteturas e assim possibilitar a condução de análises mais aprofundadas a fim de obter novas conclusões a respeito dos modelos avaliados.

CONSIDERAÇÕES FINAIS

Os resultados obtidos nos permitiram inferir qual arquitetura apresentou o melhor desempenho dentro das circunstâncias avaliadas. No entanto, acreditamos que ao avaliar as redes sobre cenários mais amplos de configurações de hiperparâmetros, as otimizações individualizadas podem trazer à tona diferenças de performance mais significativas entre elas, permitindo uma análise mais aprofundada.

REFERÊNCIAS

- [1] GOËAU, H. et al. Multi-organ plant identification. In: Proceedings of the ACM IWMAED. New York, NY, USA, 2012, p. 41–44.
- [2] BONNET, P. et al. Plant identification: Experts vs. machines in the era of deep learning - deep learning techniques challenge flora experts. In: Proceedings of the MTAEBI, 2018. p.131–149.
- [3] GHAZI, Mostafa Mehdipour; YANIKOGLU, Berrin; APTOULA, Erchan. Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing*, v. 235, p. 228-235, 2017.
- [4] CARRANZA-ROJAS, Jose et al. Automated identification of herbarium specimens at different taxonomic levels. In: *MTAP for environmental & biodiversity informatics*. Springer, Cham, 2018. p. 151-167.
- [5] SEELAND, Marco et al. Image-based classification of plant genus and family for trained and untrained plant species. *BMC bioinformatics*, v. 20, n. 1, p. 1-13, 2019.
- [6] GOODFELLOW, Ian J. et al. Multi-digit number recognition from street view imagery using deep convolutional neural networks. arXiv preprint:1312.6082, 2013.
- [7] HE, Kaiming et al. Identity mappings in deep residual networks. In: *ECCV*. Springer, Cham, 2016. p. 630-645.