

PlantAI: Aplicativo de classificação e localização de plantas em risco de extinção na Mata Atlântica

Felipe C. de Souza¹, Giuliano A. Bertoti², Rafael D. C. dos Santos¹,
Amita Muralikrishna¹, Felipe M. Carlos¹, Adriano P. Almeida¹, Helvecio B. L. Neto¹

¹Instituto Nacional de Pesquisas Espaciais (INPE)
CEP 12.227-010 – São José dos Campos – SP – Brasil

²Faculdade de Tecnologia de São Paulo (FATEC)
CEP 12.247-014 – São José dos Campos – SP – Brasil

{felipe.carvalho,rafael.santos,amita.muralikrishna,felipe.carlos,
helvecio.neto,adriano.almeida}@inpe.br,giuliano.bertoti@fatec.sp.gov.br

Abstract. *Identification of species of plants in extinction is an important but complex task. The Lista Vermelha project exists to catalog list of plants classified into different threat levels, and has a team composed of professionals responsible for assessing the risk of extinction of species of brazilian flora. However, due to the diversity and similarity contained in Brazilian biomes, the correct identification of plants is not a trivial task. Technological advances have contributed to this goal, making it possible to obtain a large amount of data from different sources, which motivated this work to present a prototype smartphone application instructed to classify plants with different threat levels of the Mata Atlântica, and map them collaboratively. A Convolutional Neural Network was trained with augmented samples of different plant species, using the transfer learning technique in different convolution blocks of the MobileNet model, and is used for online classification of image collected with the application. The experiments were conducted in the city of Jacareí-SP. For the two species evaluated, Araucaria and Pitanga, it was obtained an accuracy of more than 90%.*

Resumo. *A identificação de espécies de plantas em extinção é uma tarefa importante, porém complexa. O projeto Lista Vermelha existe para catalogar a lista de plantas classificadas em diferentes níveis de ameaça e conta com uma equipe composta por profissionais responsáveis em avaliar o risco de extinção de espécies da flora brasileira. Ainda assim, devido à diversidade e à similaridade contida nos biomas brasileiros, a identificação correta de plantas não é uma tarefa trivial. Os avanços tecnológicos vieram contribuir neste sentido, tornando possível a obtenção de um grande volume de dados de diferentes fontes, o que motivou este trabalho a apresentar um aplicativo protótipo instruído para classificar plantas com diferentes níveis de ameaça da Mata Atlântica, e mapeá-las de forma colaborativa. Uma Rede Neural Convolutiva foi treinada com amostras aumentadas de diferentes espécies de plantas, utilizando a técnica de transferência de aprendizado em diferentes blocos de convolução do modelo MobileNet, e é utilizada para a classificação online de imagens coletadas com o aplicativo. Os experimentos foram realizados no município de Jacareí-SP. Para as duas espécies avaliadas, Araucária e Pitanga, obteve-se a acurácia superior a 90%.*

1. Introdução

No século XX, a flora brasileira era estimada em 55.000 espécies de plantas, dado que, em um único hectare da Mata Atlântica do estado da Bahia foram identificadas 476 espécies distintas [MMA 1998]. Atualmente, são reconhecidas 46.848 espécies de plantas no Brasil, sendo elas: Algas, Angiospermas, Briófitas, Fungos, Gimnospermas, Samambaias e Licófitas [Flora do Brasil 2020 em construção 2020].

A Mata Atlântica é o bioma que mais sofre com o desmatamento, visto que conta com aproximadamente 35% de todas as espécies vegetais existentes no Brasil e é responsável pela produção de alimentos, madeiras, fibras, óleos e remédios [MMA 2018]. A análise feita pelo [MMA 2016] mostra que a principal ameaça que aflige a Mata Atlântica são projetos de infraestrutura e desenvolvimento urbano, considerando que o bioma está presente em 3.429 municípios, em que quase 90% da população vive em áreas urbanas, onde concentra-se uma parte significativa do que restou da Mata Atlântica [Hirota 2016].

Com a finalidade de aprofundar o conhecimento sobre o bioma e melhorar a qualidade de vida da população em geral, no município de São Paulo, criou-se o Mapeamento dos Remanescentes do Bioma Mata Atlântica (PMMA)¹. Entretanto, é interessante que todos os municípios que englobam a Mata Atlântica detenham conhecimento sobre sua flora local, para possibilitar que a população contribua com o mapeamento e a identificação de espécies ameaçadas.

Um avanço notável foi a criação do Projeto Lista Vermelha, o qual avalia o risco de extinção de espécies da flora do Brasil, a fim de auxiliar na conservação do meio ambiente [CNCFlora 2020]. Porém, devido ao grande volume e a complexidade dos dados, o uso desta lista como referência por parte dos municípios ainda é uma tarefa difícil.

Com o intuito de auxiliar a população na tarefa de identificação de plantas em centros urbanos, diversos aplicativos foram desenvolvidos para dispositivos móveis, como o *Leafsnap*, sendo o primeiro a reconhecer flores e árvores por meio de fotos capturadas, utilizando como característica o formato da folha [Kumar et al. 2012]. Já o aplicativo *Fruit Map*² em vez de um sistema de classificação, fornece aos usuários um mapa colaborativo para que possam adicionar as plantas identificadas. Os aplicativos PlantSnap³ e PlantNet⁴ dispõem de ambas funções mencionados anteriormente, porém, como os anteriores, cobrem uma baixa variedade de plantas da Mata Atlântica e não possuem informações sobre o atual estado da planta em relação ao seu nível de ameaça. Partindo desta motivação, este trabalho apresenta um protótipo de um aplicativo que tem por objetivo classificar plantas, usando Redes Neurais Artificiais (RNA), e mapeá-las de forma colaborativa, informando o risco de extinção de acordo com a Lista Vermelha (ameaçada de extinção, quase ameaçada e espécie não avaliada quanto à ameaça) de cada planta classificada.

O presente trabalho é composto pelas seguintes seções. A Seção 2 aborda os conceitos fundamentais para o entendimento deste trabalho. A Seção 3 descreve a metodologia aplicada neste trabalho. A Seção 4 apresenta os resultados e os experimentos realizados. Por fim, a Seção 5 apresenta as conclusões e sugestões para trabalhos futuros.

¹<https://bit.ly/33E6hru>

²https://play.google.com/store/apps/details?id=com.adarley1.fruitmaphl=pt_BR

³<https://www.plantsnap.com/explore-map>

⁴<https://identify.plantnet.org/>

2. Fundamentação Teórica

Nesta seção, aborda-se o conceito de mapeamento colaborativo e uma breve explicação sobre as Redes Neurais Convolucionais.

2.1. Mapeamento Colaborativo

O engajamento generalizado de cidadãos, por vezes com poucas qualificações formais, na produção de informações geográficas, é denominado de Informação Voluntária Geográfica (VGI, do inglês, *Volunteered Geographic Information*) [Goodchild 2007]. No passado as informações geográficas eram produzidas somente por agências oficiais; atualmente, com os recentes avanços tecnológicos e da Internet, tais informações não só passaram a ser disponibilizadas online, mas também começaram a ser fornecidas por voluntários [Degrossi 2015].

Entre as diversas aplicações que usam VGI, destaca-se a plataforma OpenStreet-Map⁵, que usa informações providas de voluntários para criar um mapa global. Outro exemplo de sucesso é a Wikimapia⁶, onde qualquer pessoa com acesso à Internet pode adicionar descrições, incluindo *links* e fotos, referentes a qualquer localidade no mapa.

Os sistemas que usam VGI têm obtido êxito em adquirir informações geográficas detalhadas e oportunas a um custo baixo, mas sofre de algumas deficiências [Goodchild and Li 2012]. A principal questão sobre uso de VGI refere-se à qualidade dos dados, pois grande parte dos voluntários não são especialistas. Com o objetivo de produzir classificações mais fidedignas e reduzir o viés humano na identificação da planta, este trabalho optou por realizar a classificação das plantas por meio de fotos capturadas com celulares, de forma automática usando uma RNA, especificamente, uma Rede Neural Convolucional (CNN, do inglês, *Convolutional Neural Network*).

2.2. Redes Neurais Convolucionais

Redes Neurais Convolucionais ou Redes Convolutivas são um tipo especializado de Rede Neural Artificial. Seu principal uso é em problemas de visão computacional, como classificação de imagens [LeCun et al. 2015]. As CNNs são o atual estado-da-arte para classificação de imagens, em que o modelo proposto por [Krizhevsky et al. 2012] foi capaz de rotular 1000 objetos com mais de 1.2M de imagens do conjunto ImageNet.

A arquitetura de uma CNN é composta por três camadas fundamentais: convolução, *pooling* e a camada totalmente conectada (FC, do inglês, *Fully Connected*). A camada de convolução é formada por duas ideias essenciais: campo receptor local e o compartilhamento de pesos sinápticos. No campo receptor local, de maneira oposta às RNAs tradicionais, nas quais todo o volume da camada de entrada é conectado na camada oculta, apenas uma região específica é retirada da camada de entrada [Nielsen 2015]. A operação de convolução é feita pelo deslocamento de um filtro sobre a imagem de entrada; em seguida, é calculada a soma dos produtos, e o resultado é anexado a um mapa de características. O objetivo da camada de *pooling* é reduzir a resolução espacial do mapa de características. Há duas formas de utilização do *pooling*, a agregação pela média ou pelo valor máximo [Goodfellow et al. 2016]. A camada totalmente conectada é composta pela conexão de todos os valores de entrada por todos os valores de saída.

⁵<http://www.openstreetmap.org>

⁶<http://www.wikimapia.org>

Para que uma CNN seja treinada do início ao fim, exige-se um grande volume de imagens e um alto poder computacional. Como opção, usa-se a técnica de transferência de aprendizado [Yosinski et al. 2014], na qual uma rede previamente treinada é utilizada como base para uma nova aplicação, de forma que os pesos se adaptem ao novo problema. Neste trabalho, utilizou-se o modelo MobileNet, criado para aplicações de dispositivos móveis e visão embarcada. Sua arquitetura é apresentada na Figura 1, composta por 13 blocos de convolução, os quais utilizam a técnica de convolução separável por profundidade (do inglês, *Depthwise Separable Convolution*), que consiste na aplicação de apenas um filtro para cada camada de cores [Howard et al. 2017].

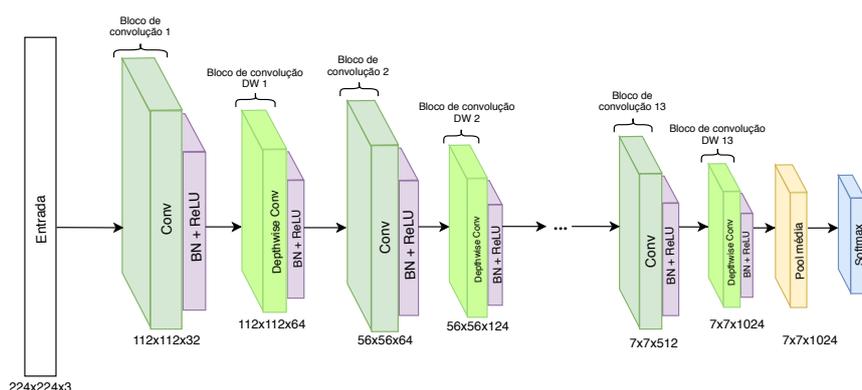


Figura 1. Arquitetura do modelo MobileNet

3. Metodologia e Métodos

Nesta Seção, apresentam-se os procedimentos metodológicos adotados neste trabalho. As subseções a seguir baseiam-se no fluxo descrito pela [Schutt and O’Neil 2013], o qual descreve as etapas necessárias para a criação de um projeto de ciência de dados. As etapas são: coleta de dados, pré-processamento das imagens e configurações da CNN. Para o pré-processamento das imagens utilizou-se a biblioteca OpenCV⁷. Na etapa de treino do modelo neural, utilizou-se o *framework* Keras⁸. Por fim, criou-se o mapa colaborativo com o banco de dados *Firestore*⁹, o qual possui sincronização em tempo real, destinado para aplicações de mapeamento colaborativo.

3.1. Coleta de dados

As imagens foram obtidas através da base de dados *Encyclopedia of Life* (EOL)¹⁰, um *hub* de imagens de diversos seres vivos contidos na natureza, incluindo plantas. O EOL é uma plataforma colaborativa, em que qualquer usuário pode, além de adicionar suas próprias imagens, avaliar imagens de outros usuários em uma escala de zero a cinco estrelas. Às imagens avaliadas com cinco estrelas é atribuído um selo de confiança. Neste trabalho, utilizou-se apenas imagens com selo de confiança e para aproximar-se de um cenário de uso real, optou-se por incluir imagens com elementos adicionais, como nuvens, folhas de outras plantas e pessoas.

⁷<https://firebase.google.com/>

⁸<https://keras.rstudio.com/>

⁹<https://firebase.google.com/>

¹⁰<https://eol.org/>

O conjunto de dados coletado abrange sete espécies de plantas: Araucária (a), Guapeba (b), Guapuruvu (c), Guatambu (d), Ipê-roxo (e), Jequitibá-vermelho (f) e Pitanga (g), apresentados na Figura 2. De acordo com a lista vermelha, às espécies Araucária (a), Guapeba (b) e Guatambu (c) possuem risco elevado de extinção na natureza, classificadas “em perigo”. O Ipê roxo (d) se encontra no estágio de “quase ameaçada”, correndo risco de ser qualificado em uma categoria de ameaça num futuro próximo. As outras plantas: Pitanga (g), Jequitibá-vermelho (f) e Guapuruvu (c) ainda não foram avaliadas quanto a estarem ameaçadas.

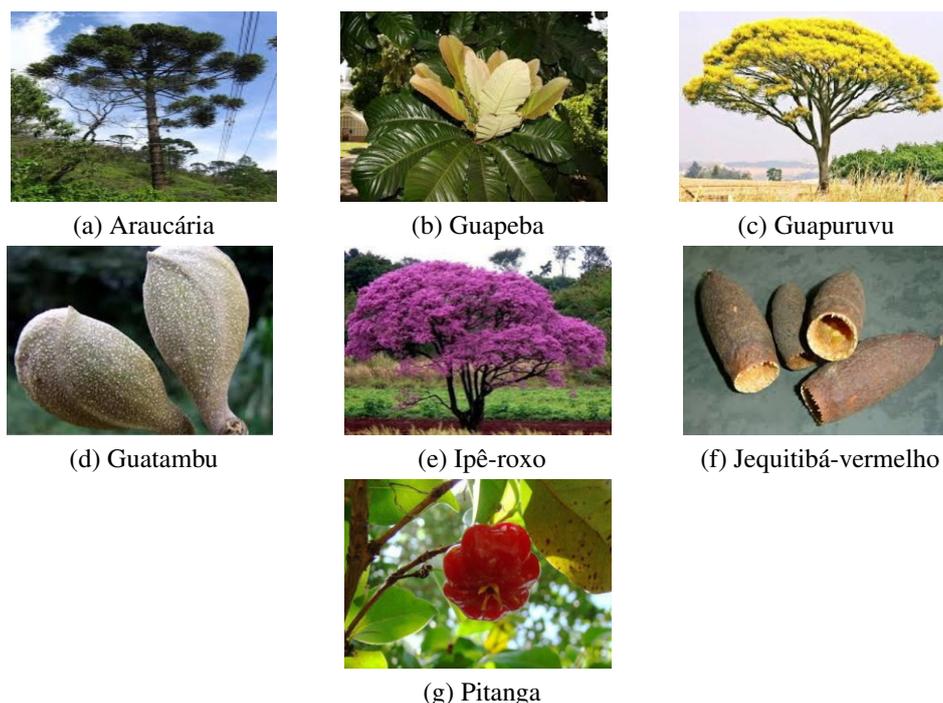


Figura 2. Amostra de cada espécie do conjunto de imagens

3.2. Pré-processamento

Com o intuito de aumentar a quantidade de imagens do conjunto de dados e garantir que a rede tenha uma boa capacidade generalização, utilizou-se a técnica de criação de imagens artificiais, chamada *data augmentation* [Perez and Wang 2017]. *Data augmentation*, ou aumento de dados, consiste na criação de imagens artificiais a partir da original, adicionando filtros sobre a imagem. Assim, para cada imagem do conjunto de dados, adicionou-se seis filtros: Horizontal, Vertical, Gaussiano, Rotativo 45°, Rotativo -45°, Rotativo 80°.

A Figura 3 apresenta a quantidade total de imagens, antes e após a utilização da técnica. Em seguida, as imagens foram redimensionadas para 224x224 pixels. Com isso, destinou-se 70% das imagens para o treinamento, 20% para validação e 10% para testes.

3.3. Hiperparâmetros da CNN

Como mencionado na Seção anterior, utilizou-se o modelo MobileNet (Figura 1) pré-treinado com os parâmetros do conjunto de dados ImageNet. A Tabela 1 apresenta os hiperparâmetros avaliados para o treino da rede. Escolheu-se a função de ativação ReLu

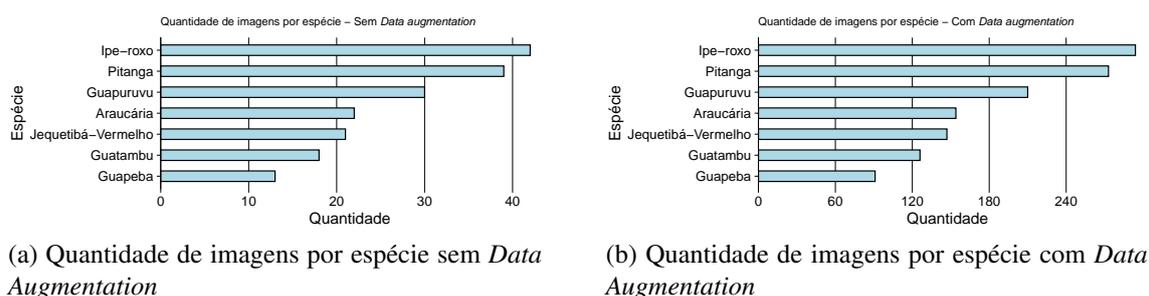


Figura 3. Quantidade total de imagens por espécies

(*Rectified Linear Units*) pelo fato da eficiência no treino quando comparada com a tangente hiperbólica [Rawat and Wang 2017]. Por padrão, utilizou-se a mesma função de otimização em que foi treinado o modelo MobileNet. Por fim, três blocos de convolução foram avaliados, com o objetivo de identificar o modelo com a maior capacidade de generalização. Optou-se pelo uso dos blocos de convolução localizados ao final do modelo, por apresentarem características específicas do conjunto de dados em que foram treinados.

Modelo	Função de Ativação	Taxa de Aprendizado (η)	Função de Otimização	Épocas	Blocos de Convolução
MobileNet	ReLU	0.00001	RMSprop	23	10, 11 e 12

Tabela 1. Hiperparâmetros utilizados para o treino da CNN

4. Resultados e Discussões

Nesta Seção, apresentam-se os resultados do treino da CNN, a interface gráfica do aplicativo e os experimentos realizados.

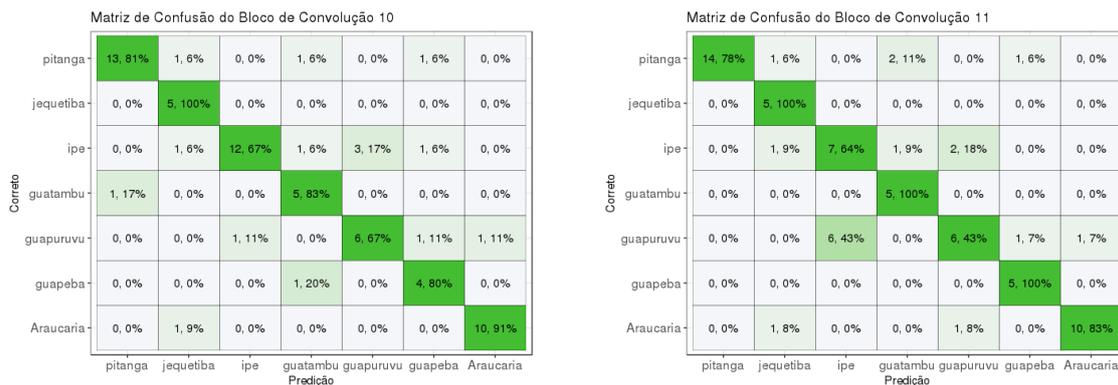
4.1. Resultados do treinamento da CNN

A Figura 4 apresenta as matrizes de confusão geradas pelo conjunto de teste, para cada bloco de convolução avaliado. Observa-se que o bloco de convolução 10 (Figura 4 (a)) obteve mais verdadeiros positivos nas classes Pitanga, Guapuruvu e Araucária, porém, obteve mais falsos positivos na classe Guapeba.

A Tabela 2 apresenta a acurácia e a taxa de erro dos blocos de convolução avaliados. O bloco de convolução de 10 apresentou maior acurácia, pelo fato de sua profundidade, sendo o mais próximo da camada de entrada, no entanto, levou mais tempo no treinamento. Desta forma, escolheu-se a rede do bloco de convolução 10 para utilizar no aplicativo.

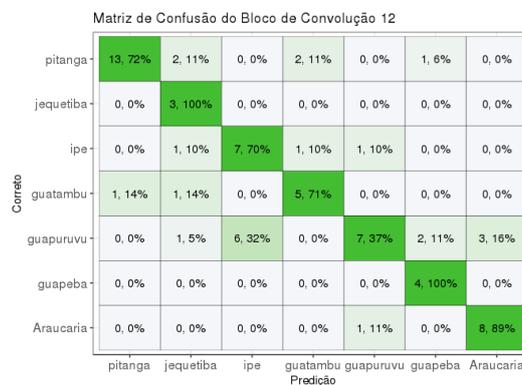
Bloco Convolução	Acurácia	Taxa de Erro
10	79.8%	0.66
11	74.7%	0.91
12	65.5%	0.94

Tabela 2. Acurácia do blocos de convolução



(a) Matriz de confusão do modelo treinado a partir do bloco de convolução 10

(b) Matriz de confusão do modelo treinado a partir do bloco de convolução 11



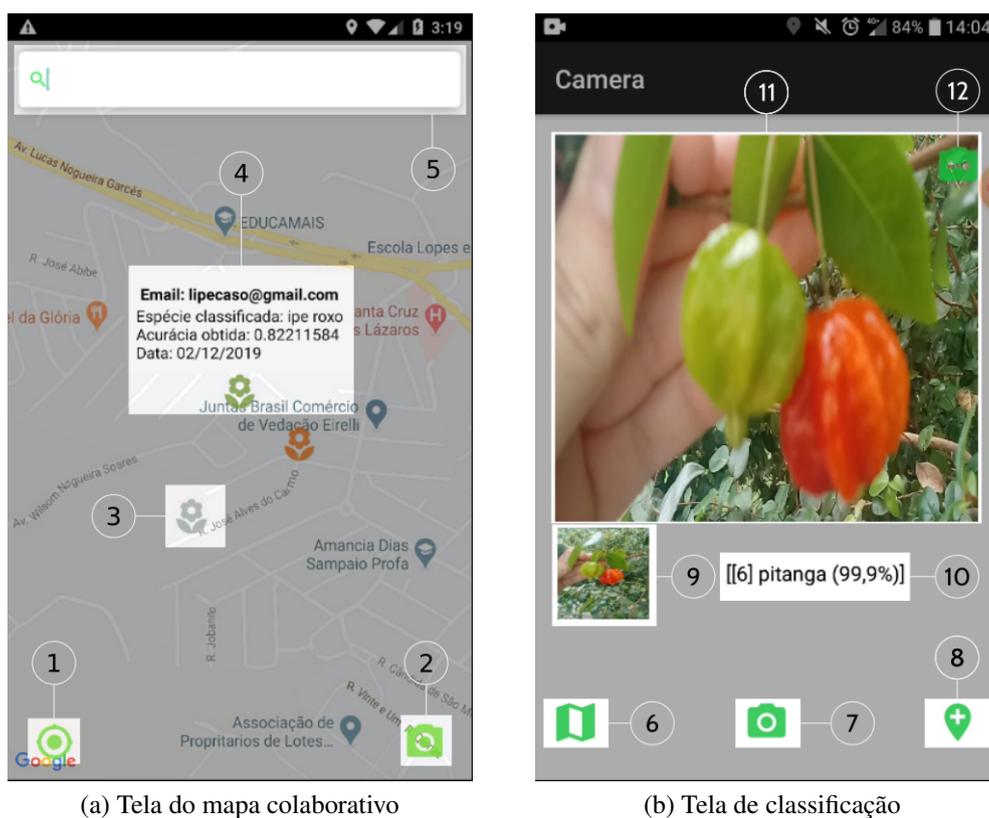
(c) Matriz de confusão do modelo treinado a partir do bloco de convolução 12

Figura 4. Matrizes de confusão de cada bloco de convolução avaliado

4.2. Interface do protótipo

A Figura 5 (a) apresenta a tela do mapa colaborativo, a legenda numérica apresenta as seguintes funcionalidades: o ícone de localização (1) possui a função de localizar a atual posição do usuário; o ícone da câmera (2) inicia a tela de classificação; os ícones de flores (3) recebem diferentes cores de acordo com a atual situação da planta classificada, em que plantas “em perigo” recebem o marcador laranja, plantas “quase ameaçadas” recebem o marcador verde claro e as plantas que não se encontram em nenhuma das situações anteriores recebem a cor cinza. A paleta de cores utilizada baseia-se no Projeto da Lista Vermelha. Por fim, a barra de busca (5), integrada com a API do *Google Places*, facilita a busca do usuário por localizações com sugestões de cidades, bairros e ruas. No marcador verde (4), observam-se algumas informações relevantes, como o e-mail do usuário que identificou a planta, a espécie identificada, a acurácia obtida e a data de classificação.

A tela de classificação, apresentada na Figura 5 (b), assim como a tela do mapa colaborativo, é composta por números para facilitar o entendimento de cada funcionalidade. Os ícones da horizontal inferior (6,7,8), representam os ícones de mapa, câmera e adição, respectivamente. O mapa (6) possui a função de voltar para a tela do mapa colaborativo, a câmera (7) tem por objetivo capturar imagens, e conseqüentemente, classificá-las (10). Logo após a classificação, é possível visualizar a imagem capturada (9). Por fim, o usuário tem a opção de adicionar a classificação feita no mapa colaborativo (8).



(a) Tela do mapa colaborativo

(b) Tela de classificação

Figura 5. Telas do aplicativo protótipo desenvolvido neste trabalho

4.3. Experimentos

Os experimentos foram realizados na cidade de Jacareí, município localizado no interior do estado de São Paulo. Neste experimento, objetivou-se avaliar duas espécies de plantas, Araucária e Pitanga.

Para classificar a Araucária, diversas fotos foram capturadas, alternando entre o caule e as folhas da árvore. Notou-se que a classificação pelas folhas da Araucária atingiu maior acurácia do que pelo caule. A Figura 6 (a) apresenta a classificação da Araucária, na qual obteve-se 91,4% de acurácia. É interessante observar que a imagem classificada (Figura 6 (a)) possui diversos elementos, como folhas de outras árvores e nuvens.

A Figura 6 (b) mostra a classificação de uma Pitangueira através da folha, diferentemente da Figura 5 (b), que identificou a planta pelo fruto. Em ambas classificações atingiu-se aproximadamente 100% de acurácia, devido a dissimilaridade dos atributos da Pitangueira em comparação às outras plantas. No entanto, plantas com frutos semelhantes aos da Pitangueira, como a Acerola, são identificadas como Pitangueiras.

5. Conclusão

Neste trabalho, apresentou-se um protótipo de aplicativo desenvolvido para classificar plantas e mapeá-las de forma colaborativa, informando o risco de extinção de acordo com a Lista Vermelha. Para realizar a classificação das plantas treinou-se um modelo de CNN baseado na arquitetura MobileNet. O uso da base de imagens voluntárias da plataforma EOL para o treinamento da CNN mostrou-se eficaz durante a classificação, especialmente

ao lidar com casos em que as imagens apresentavam também outros elementos de um cenário real.

Este trabalho pode possuir numerosas aplicações na sociedade, tais como: criação de áreas de preservação ambiental; mecanismo didático para aulas de geoprocessamento; auxílio aos munícipes na localização de plantas que se encontram ameaçadas em suas proximidades. Para trabalhos futuros, pretende-se avaliar outros modelos de CNN e expandir a quantidade de espécies de plantas a serem classificadas. Outra funcionalidade interessante é a adição de um sistema de *feedback* de classificações, no qual cada planta localizada no mapa recebe uma avaliação, de zero a cinco, para que assim, plantas mal avaliadas sejam removidas do aplicativo.

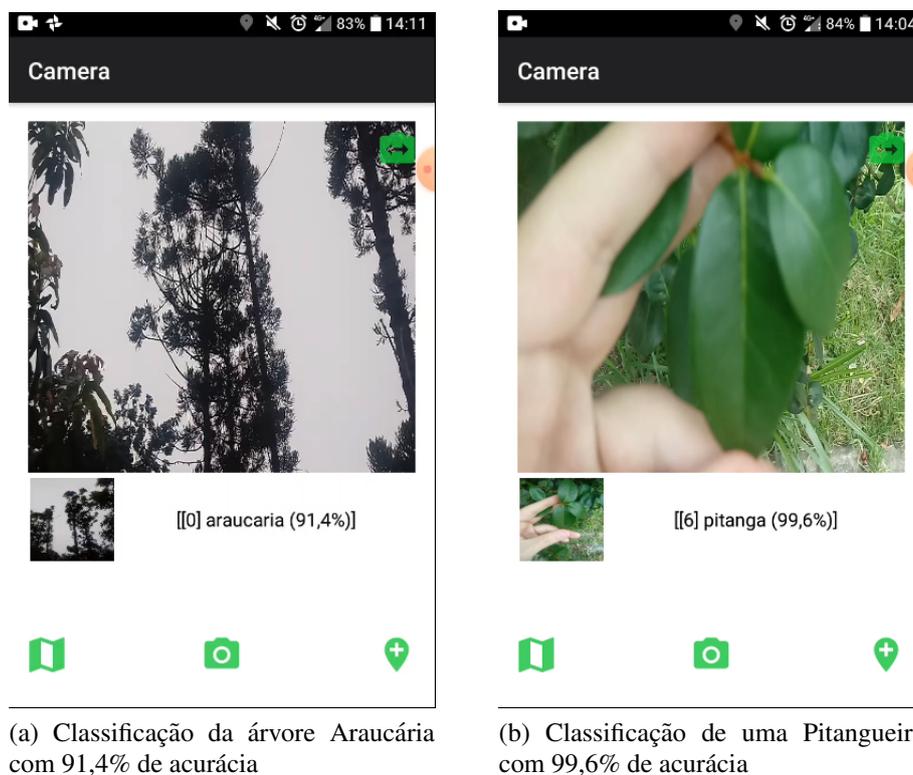


Figura 6. Experimentos realizados na cidade de Jacareí-SP com o uso do aplicativo desenvolvido neste trabalho

Referências

- CNCFlora (2020). Projeto Lista Vermelha. <https://bit.ly/2J8m34J>. Acesso em: 2020-03-20.
- Degrossi, L. C. (2015). *Uma abordagem para obtenção e disponibilização em tempo real de informações geográficas voluntárias no contexto de gestão de risco de inundação*. PhD thesis, Universidade de São Paulo.
- Flora do Brasil 2020 em construção (2020). Jardim botânico do rio de janeiro. Disponível em: <<https://bit.ly/3drneKf>>. Acesso em: 2020-03-15.
- Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *Geo-Journal*, 69(4):211–221.

- Goodchild, M. F. and Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial statistics*, 1:110–120.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- Hirota, M. M. (2016). A mata atlântica invisível nas cidades. Disponível em: < <https://bit.ly/3afkfma> >. Acesso em: 2020-03-13.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.
- Kumar, N., Belhumeur, P. N., Biswas, A., Jacobs, D. W., Kress, W. J., Lopez, I. C., and Soares, J. V. (2012). Leafsnap: A computer vision system for automatic plant species identification. In *Computer vision—ECCV 2012*, pages 502–516. Springer.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.
- MMA (1998). *Primeiro relatório nacional para a convenção sobre diversidade biológica*. <https://bit.ly/2UBftZN>.
- MMA (2016). *Quinto Relatório Nacional para a Convenção sobre Diversidade Biológica*. <https://bit.ly/2UgPR5e>.
- MMA (2018). Mata atlântica. <https://bit.ly/3aePoWS>. Acesso em: 2020-03-15.
- Nielsen, M. A. (2015). *Neural networks and deep learning*, volume 2018. Determination press San Francisco, CA, USA:.
- Perez, L. and Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*.
- Rawat, W. and Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 29(9):2352–2449. PMID: 28599112.
- Schutt, R. and O’Neil, C. (2013). *Doing Data Science: Straight Talk from the Frontline*. O’Reilly Media, Inc.
- Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328.