Plant species recognition using leaf images and convolutional neural networks

Isabelly dos Santos Pacheco[1], Rosimeri de Oliveira Fragoso[1], Lilian Tais de Gouveia[1], Luciano Jose Senger[1]

[1]Universidade Estadual de Ponta Grossa (UEPG) – Ponta Grossa – PR – Brasil

21001026@uepg.br, meri_ol@yahoo.com.br, ltgouveia@uepg.br, ljsenger@uepg.br

Abstract. *The identification of plant species is essential in botany and has attracted the interest of researchers in the field of computer science. Such identification requires the assistance of botany experts and, due to the substantial number of species and the similarity among them, it can be time-consuming and subjective. To automate the process of plant identification, computer systems that capture and process plant images have been considered. These systems use machine learning and therefore require image samples for training and model construction. Among the techniques that can be used for machine learning, convolutional neural networks have shown promise due to their ability to use images without prior preprocessing and background information. This work investigates the use of machine learning through convolutional neural networks to identifying plant species. For this, a new dataset of images from 35 plant species were created, collecting images from an arboreal collection, and, using data augmentation, this dataset was expanded. This dataset was used to evaluate the accuracies of four convolutional neural network models. The better accuracy value was equal to 89%, when using the MobileNetV2 model.*

Keywords: plant species identification; machine learning, image processing.

1. Introduction

Plants releases oxygen into the atmosphere, captures carbon dioxide, prevents erosion, and contributes to the regulation of relative humidity in the air. Furthermore, living organisms obtain a source of oxygen and nutrients through plants (Antonelli, Smith, & Simmonds, 2019). Knowledge of plants is essential for improving ecosystems, agriculture, and sustainability (Lee, Lim, Song, & Alqahtani, 2023). Due to this fact, researchers have shown interest in new techniques to identify plant species.

To identify plant species, there is a meticulous process involving experts. A plant is commonly identified through visual analysis of its parts, such as stems, fruits, and, primarily, its leaves (Cope, Corney, Clark, Remagnino, & Wilkin, 2012; Zhuang, et al., 2019). The identification of plants species is considered a major challenge due to the following factors: inter-species similarity, high intra-species variability, imbalanced data, and the number of species (Britto, Pacífico, & Ludermir, 2019; Moresco, De S. Britto, Costa, Senger, & Hochuli, 2022).

Computer vision and machine learning have been considered as a solution for plant identification. Computer vision systems use images of plant parts to automatically identify species, allowing non-experts to obtain species identification through image capture. The literature demonstrates that there are two main approaches to plant species identification: a) image processing and non-automatic feature extraction through techniques like histogram surveys, filters, and texture analysis; and b) automatic feature extraction through deep learning. Approaches using automatic feature extraction and convolutional neural networks have demonstrated better classification accuracy compared to non-automatic feature extraction approaches (Zhuang, et al., 2019).

One issue in machine learning tasks is the lack of plants datasets, which are insufficient in diversity and size (García-Ordás, Benítez-Andrades, García-Rodríguez, Benavides, & Alaiz-Moretón, 2020). Thus, regardless of the computational approach to be employed, it is necessary to collect images and label them through consultation with experts in the field of botany. This paper investigates the use of plant leaf images for species classification through convolutional neural networks and transfer learning. The contribution of this paper is: a) a new plant species image dataset; b) an evaluation of machine learning, using convolutional neural networks and transfer learning, on the problem of plant species identification.

2. Background

Due to problems in distinguishing noise or to filter background information from images, deep learning and convolutional neural networks (CNNs) have been used in plants species classification (Moresco *et al.*, 2022). A CNN is a hierarchical and multi-layer feature extractor: each convolutional layer performs a convolution operation on the image input and passes the extracted features to the next layer (Figure 1). Batch normalization is performed on the output of the convolutional layers, whereby the extracted features are normalized by adjusting and scaling the artificial neuron activations. Pooling layers take each output from the convolutional layer's feature map and prepares a condensed feature map. Before passing this information to fully connected layers for the final task, such as classification, it is necessary to convert this three-dimensional representation into a one-dimensional vector. The Flatten layer precisely accomplishes this task by reshaping the data, enabling it to be fed into fully connected layers. Lastly, a fully connected layer (dense layer) and the SoftMax layer computes the score of each class and infers the category of the input image. The number and configuration of the layers defines a convolutional network architecture. A CNN model is a specific instance of a CNN architecture with weights trained for a particular task.

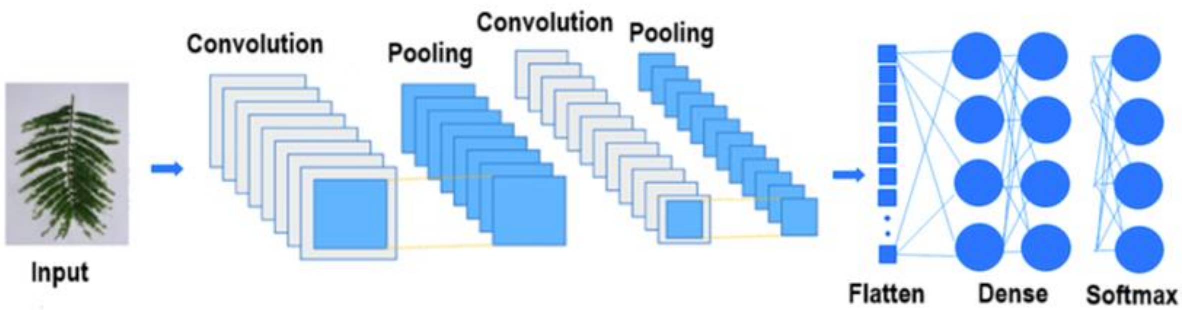

Figure 1 – A convolutional neural Network with two convolution and two max-polling layers. These layers are trained to create a feature representation (Flatten layer). The two dense layers and the SoftMax layer are trained to build a classifier.

One major issue when using Convolutional neural networks to build classification models is the requirement of many samples in the training dataset. The accuracy and generalization capabilities of CNNs is directly related to number of see samples. When many samples are not available, pre-trained networks can be used, using the concept of transfer learning. Transfer learning allows the knowledge available from a large dataset be transferred to other classification task. Instead of training a CNN from scratch for a new task, transfer learning starts with a pre-trained model on a large dataset, often on a diverse task like image recognition. CNN models using the ImageNet dataset are frequently used transfer learning. The knowledge captured by the model in its convolutional layers using the ImageNet dataset, which learn hierarchical features, can be transferred to a new task with a smaller dataset. By using pre-trained features, the model can generalize better and require less training data.

Convolutional neural networks have been evaluated for plant species identification, using leaf images as input to train and evaluate classification models. Lee *et al.* (2023) conducted a study using convolutional neural networks and the MalayaKew Leaf dataset . CNNs were used to extract features from images. The extracted features, acquired from the flatten convolutional network layer, were classified using multilayer perceptron (MLP) and support vector machines (SVM). The classifier achieved an accuracy of 99.5%.

Liu *et al*. (2018) used a modified version of the LeNet model. The model was used in both feature extraction and classification. Data augmentation techniques were used to expand the size of the image's dataset. In the study, horizontal flip, vertical flip, noise, color jittering, and rotation were applied in the original dataset to generate new images. In this study, an accuracy of 87% was achieved.

A convolutional neural network system has been proposed in (Le Huy Hien & Van Hieu, 2020). The PlantCLEF2003 dataset, which consists of 51,273 images from 609 plant species, were used to train classification models. The paper evaluates the Resnet50V2, Inception, Resnet V2, MobilenetV2 and VGG16 models when used to extract features from images. Support Vector Machine and k-nearest neighbor (KNN) classifiers were also evaluated, and the highest accuracy of 95.6% were observed when using the MobilenetV2 model.

In (Sundara Sobitha Raj & Vajravelu, 2019), the MobileNet and DenseNet models were combined to extract the features from plant leaf images. The extracted features were used to train traditional machine learning classifiers, such as Naïve Bayes and multi-layer perceptron. Four image datasets were considered in the experiments: Folio (Munisami, Ramsurn, Kishnah, & Pudaruth, 2015), Swedish leaf (Söderkvist, 2001), Flavia (Wu *et al.*, 2007), and a dataset built by the authors. The authors reported accuracy values among 96% to 99% and the best results were reported when using a combination of CNN and traditional machine learning classifiers.

Moresco et al. (2022) evaluate different architectures of CNNs in classifying plant species from leaf images and investigate the fusion of their intermediate layers. The authors combined up to 5 convolutional blocks for each model experimented and studied the impact on the CNN accuracy and its capacity to recognize classes unseen during training. A Siamese Neural Network model, which contains two sibling convolutional architectures that share their weights, were combined with the VGG16, MobileNet and DenseNet models and evaluated the Flavia and MalayaKew datasets. The best results show an accuracy of 100% for the Flavia dataset and 94.31% for the MalayaKew dataset.

3. Methods

Leaf images from plants were obtained from the Arboreal Collection at Augusto Ribas Agricultural College (CAAR/UEPG). This plant collection is situated in Augusto Ribas College, located at the Ponta Grossa State University, Ponta Grossa, Paraná, Brasil. The images were taken using a smartphone camera with a resolution of 1659 x 2658 pixels and 24 bits of color depth. For each plant, images samples were collected using a white paper sheet background, varying the leaf orientation. The Figure 2 shows three samples of plants images included in the dataset. The number of plant species used to build the dataset is equal to 35. Data augmentation, using rotation and zoom, were used to increase the data size from 730 to 1986 images in the CAAR dataset (Figure 3).

Convolutional neural networks and transfer learning were experimented with using the MobileNet, MobileNetV2, VGG16, and VGG19 as base models. A percentage of 70% of the images were used for training, and the remaining for validation. The number of training epochs was set to 50 for all models, except for the MobileNetV2 model, which used 100 training epochs. In general, it was experimented number of epochs at which the accuracy and loss function on both the training and validation sets became stable. The Figure 4 shows a plot of the convolutional network layers with a VGG16 based transfer learning model. As seen in Figure 4, additional dense layers have been added to the base model. Only these layers have their weights updated on training, while the layers of the base network have their weights frozen (not updated on training). The pre-trained CNN models were experimented using the Keras API and the Python language (Joseph, Nonsiri, & Monsakul, 2021).

(a) *Parapiptadenia rigida* (b) *Jacaranda mimosifolia* (c) *Tabebuia roseoalba* (d) *Libidibia ferrea*

Figure 2– Four samples of plant images included in the CAAR dataset: *Parapiptadenia rigida* (a), *Jacaranda mimosifolia* (b), *Tabebuia roseoalba* (c) and *Libidibia ferrea* (d)

(a) (b) (c) (d)

Figure 3 – Images of the *Jacaranda mimosifolia* leafs. In (a), the original image. The images from (b) to (d) were created using data augmentation.

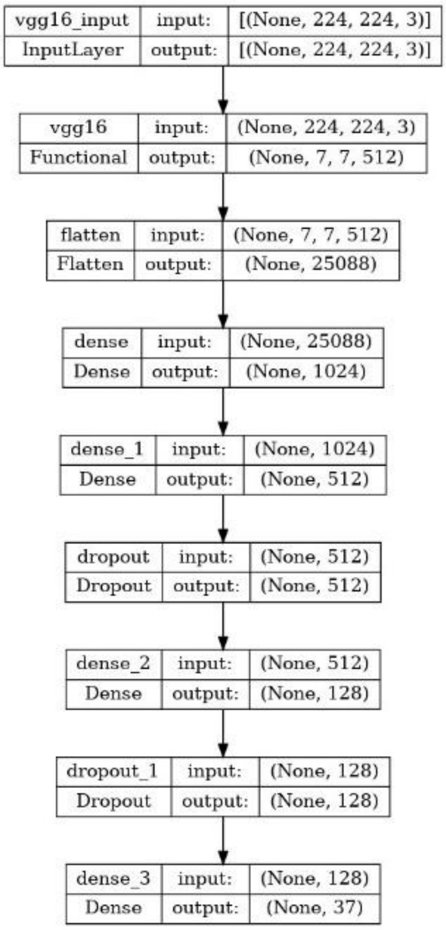

Figure 4 – Number and configuration of the model layers. The plot shows the VGG16 as transfer learning model (three first layers). The additional layers configuration was also fixed when using the other base models experimented (VGG19, MobileNet and MobileNetV2).

The accuracy, loss function, recall, precision, and f1-score measures were used to evaluate the results. During the training, the accuracy and the loss function values were observed. The overall correctness of the model was evaluated using the accuracy values. The accuracy is the ratio of the number of correct plant species predictions and the total number of predictions. A loss function (also known as a cost function or objective function) quantifies how well the model is performing during training. It represents the discrepancy between the predicted values and the actual ground truth labels. The goal during training is to minimize the value of the loss function.

The Precision metric evaluates how often the model is correct when predicting a target plant species. The precision is computed by dividing the number of correct positive predictions (true positives) for a plant species by the total number of instances the model predicted as positive, considering both true and false positives plant species. High precision (near to 1, in a 0 to 1 scale) indicates that the model makes fewer false positive errors. The recall (also known as sensitivity) indicates the performance of the model when finding all objects of a target plant species. High recall indicates that the model captures most of the positive instances. The F1-score is the harmonic mean of precision and recall. It balances precision and recall, providing a single metric that considers both false positives and false negatives.

Also, a confusion matrix was used to describe the classification results. The confusion matrix provides a comprehensive summary of the performance of a classification model by illustrating the number of correct and incorrect predictions for each plant species. The rows represent the true classes, while the columns represent the predicted species. Each cell in the matrix indicates the number of instances where a given class was predicted correctly or incorrectly. This allows for an analysis of the model's performance.

4. Results

The Figures 5, 6, 7 and 8 show the accuracy and loss values for each experimented CNN model according to each epoch of the training and validation. The accuracy values direct vary with number of epochs. The best results expressed into a 0 to 1 scale, where 1 represents an accuracy of 100%, are presented in Table 1. The MobileNetV2 model obtained the better accuracy value (93%) than the accuracy values from the other models. For this model, the Table 2 presents the classification report when using the validation data and the Figure 9 depicts the confusion matrix.

The precision observed was equal to 100% for 14 plant species. Among all plant species, the worst precision values were observed for the Psidium guajava L. (65%), Prunus myrtifolia (L.) Urb. (57%) and Myrcianthes pungens (O.Berg) D.Legrand (67%) plant species. The Prunus myrtifolia (L.) also presented one of the worst recall values (33%), followed by the Prunus cerasifera Ehrh. and the Croton floribundus Spreng species, with recall values of 42% and 52% respectively. The confusion matrix indicates (Figure 9) , when predicting the Prunus myrtifolia (L.), that the model did not distinguish among mainly the Ocotea catharinensis Mez, the Eugenia involucrata DC. and the Psidium guajava L. species.

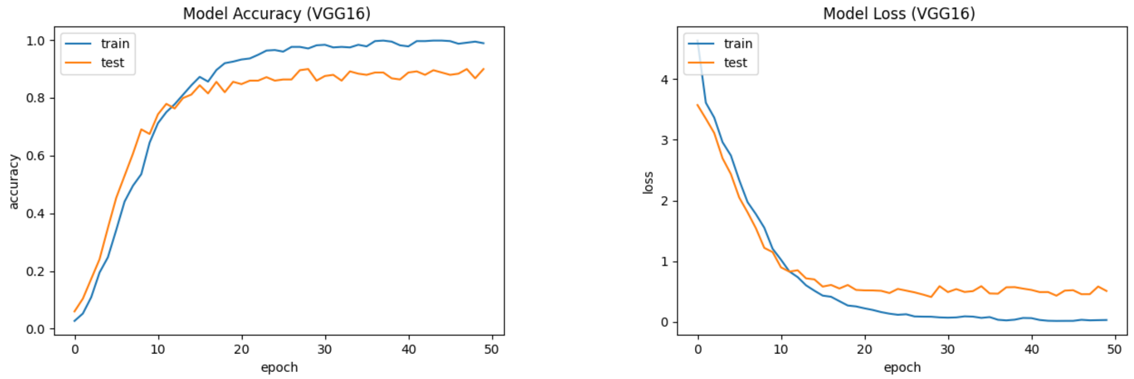

Figure 5. VGG16 model results.

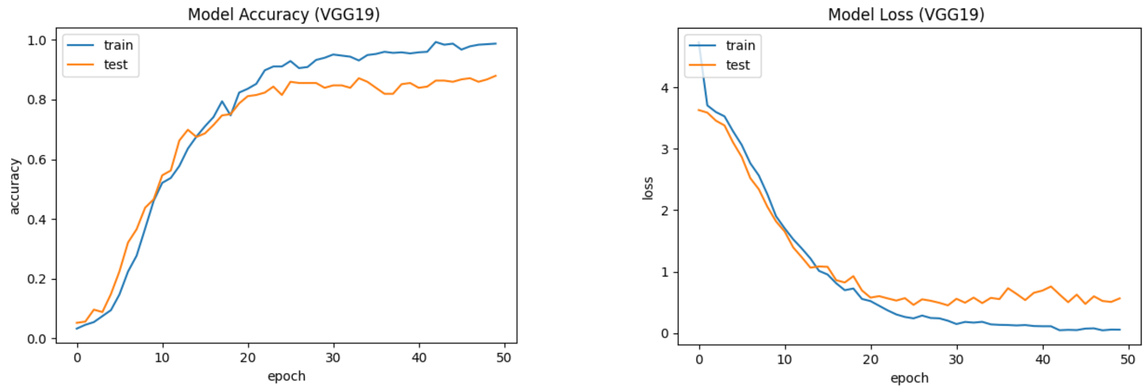

Figure 6. VGG19 model results.

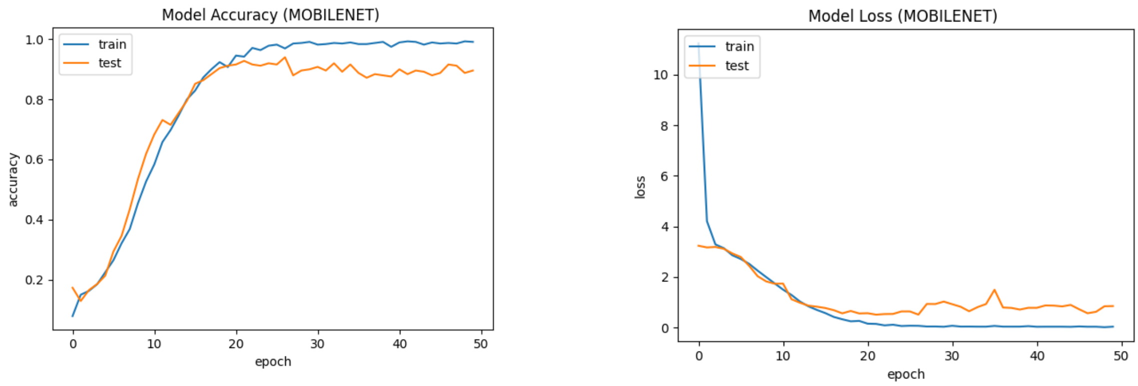

Figure 7. MobileNet model results.

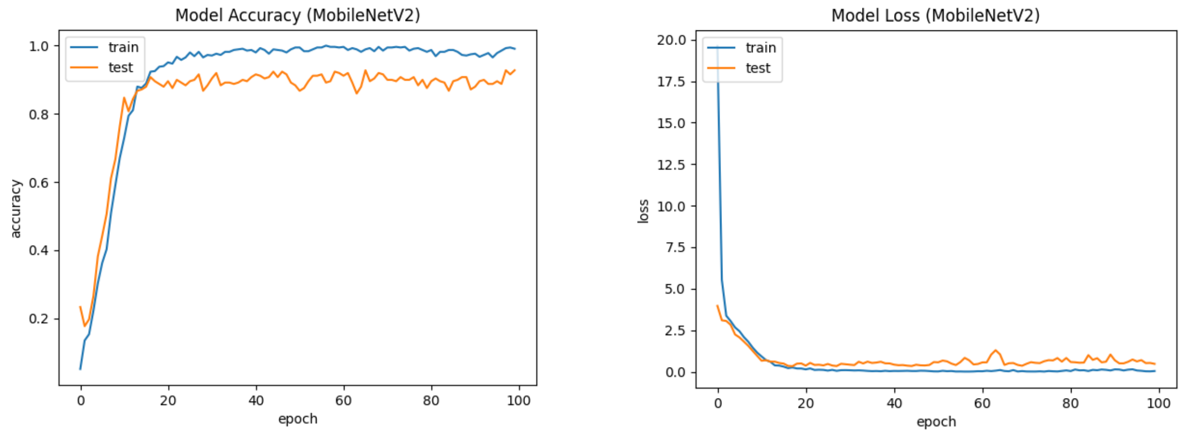

Figure 8. MobileNetV2 model results.

Table 1. VGG16, VGG19, MOBILENET, and MOBILENETV2 best accuracy values.

Performance Metrics/ Model	MobileNet	MobileNetV2	Vgg16	Vgg19
Accuracy	0.88	0.89	0.84	0.84

5. Conclusion

This paper presented the CAAR plant image dataset and an evaluation of its usage to train convolutional neural models aiming to identify plant species. The dataset comprises 35 plant species and a number of 1986 images. The results indicated that the MobileNetV2 model was better than the other models, achieving an accuracy of 89%.

Improvements can be done in the CAAR dataset, like expand its size and variety, aiming to achieve better accuracy values in computational classification tasks. In future work, extensions in the classification models can be evaluated. For instance, the inclusion of new layers in the network topology and the use of fine tuning can be studied. Furthermore, other network models and other traditional machine learning algorithms can also be evaluated, in addition to those presented in this paper.

Table 2. MobileNetV2 model precision, recall, f1-score values.

Plant Species	**precision**	**recall**	**f1-score**	**support**
Parapiptadenia rigida (Benth.) Brenan	0,88	1,00	0,93	14
Psidium longipetiolatum D.Legrand	0,91	0,83	0,87	24
Gymnanthes klotzschiana Müll.Arg.	0,88	1,00	0,93	14
Campomanesia xanthocarpa (Mart.) O.Berg	0,69	0,92	0,79	12
Cinnamomum verum J.Presl	0,73	0,92	0,81	12
Ocotea catharinensis Mez	1,00	0,71	0,83	14
Ocotea odorífera (Vell.) Rohwer	0,67	1,00	0,80	14
Croton floribundus Spreng.	0,79	0,52	0,63	21
Roupala montana var. *brasiliensis* (Klotzsch) K.S.Edwards	1,00	0,93	0,96	28
Cassia leptophylla Vogel	1,00	1,00	1,00	12
Cedrela fissilis Vell.	0,77	1,00	0,87	10
Eugenia involucrata DC.	0,71	1,00	0,83	10
Prunus serrulata Lindl.	0,85	0,92	0,88	12
Cupania vernalis Cambess.	1,00	1,00	1,00	10
Eugenia pyriformis Cambess.	1,00	0,79	0,88	14
Psidium guajava L.	0,65	0,93	0,76	14
Inga vulpina Mart. ex Benth.	0,68	0,81	0,74	21
Inga sessilis (Vell.) Mart.	1,00	0,64	0,78	14
Tabebuia roseoalba (Ridl.) Sandwith	0,93	0,88	0,90	16
Jacaranda mimosifolia D. Don	1,00	1,00	1,00	8
Laurus nobilis L.	0,83	0,71	0,77	14
Luehea divaricata Mart. & Zucc.	0,59	1,00	0,74	10
Senna macranthera (DC. ex Collad.) H.S.Irwin & Barneby	1,00	1,00	1,00	14
Curitiba prismatica (D.Legrand) Salywon & Landrum	0,76	0,93	0,84	14
Myrcianthes pungens (O.Berg) D.Legrand	0,67	0,60	0,63	10
Bauhinia variegata L.	1,00	0,75	0,86	12
Libidibia ferrea (Mart. ex Tul.) L.P.Queiroz	0,92	1,00	0,96	12
Paulownia fortunei var. *mikado*	1,00	0,93	0,96	14
Prunus myrtifolia (L.) Urb.	0,57	0,33	0,42	12
Eugenia uniflora L.	0,92	0,75	0,83	16
Prunus cerasifera Ehrh.	0,71	0,42	0,53	12
Schinus molle L.	1,00	1,00	1,00	14
Sequoia sempervirens (D.Don.) Endl.	1,00	0,92	0,96	12
Poincianella pluviosa (DC.) L.P.Queiroz	1,00	0,86	0,92	14
Tipuana tipu (Benth.) Kuntze	1,00	1,00	1,00	14

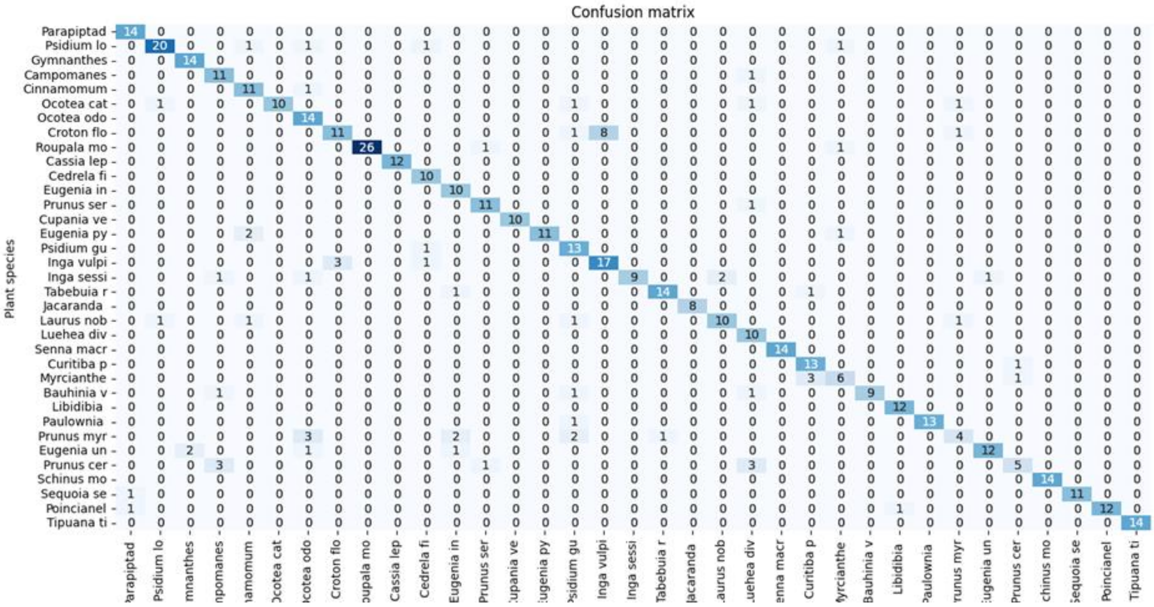

Figure 9. Confusion matrix.

ACKNOWLEDGEMENTS

This work was supported by Fundação Araucária and CNPQ.

6. References

Antonelli, A., Smith, R. J., & Simmonds, M. S. (2019). Unlocking the properties of plants and fungi for sustainable development. Nature Plants, 5, 1100–1102. doi:10.1038/s41477-019-0554-1

Britto, L., Pacífico, L., & Ludermir, T. (2019). Um Sistema Automático de Classificação de Folhas Baseado em Características da Forma. Anais do XVI Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2019). Sociedade Brasileira de Computação - SBC. doi:10.5753/eniac.2019.9353

Cope, J. S., Corney, D., Clark, J. Y., Remagnino, P., & Wilkin, P. (2012). Plant species identification using digital morphometrics: A review. Expert Systems with Applications, 39, 7562–7573. doi:10.1016/j.eswa.2012.01.073

García-Ordás, M. T., Benítez-Andrades, J. A., García-Rodríguez, I., Benavides, C., & Alaiz-Moretón, H. (2020). Detecting Respiratory Pathologies Using Convolutional Neural Networks and Variational Autoencoders for Unbalancing Data. Sensors, 20, 1214. doi:10.3390/s20041214

Hu, J., Chen, Z., Yang, M., Zhang, R., & Cui, Y. (2018). A Multiscale Fusion Convolutional Neural Network for Plant Leaf Recognition. IEEE Signal Processing Letters, 25(6), 853–857. doi:10.1109/LSP.2018.2809688

Joseph, F. J., Nonsiri, S., & Monsakul, A. (2021). Keras and TensorFlow: A Hands-On Experience. Em EAI/Springer Innovations in Communication and Computing (pp. 85–111). Springer International Publishing. doi:10.1007/978-3-030-66519-7_4

Le Huy Hien, N., & Van Hieu, N. (2020). Recognition of Plant Species using Deep Convolutional Feature Extraction. 11, 904-910.

Lee, C. P., Lim, K. M., Song, Y. X., & Alqahtani, A. (2023). Plant-CNN-ViT: Plant Classification with Ensemble of Convolutional Neural Networks and Vision Transformer. Plants, 12, 2642. doi:10.3390/plants12142642

Lee, S. H., Chan, C. S., Wilkin, P., & Remagnino, P. (2015). Deep-plant: Plant identification with convolutional neural networks. 2015 IEEE International Conference on Image Processing (ICIP), (pp. 452-456). doi:10.1109/ICIP.2015.7350839

Liu, J., Yang, S., Cheng, Y., & Song, Z. (2018). Plant Leaf Classification Based on Deep Learning. 2018 Chinese Automation Congress (CAC), (pp. 3165-3169). doi:10.1109/CAC.2018.8623427

Moresco, M., De S. Britto, A., Costa, Y. M., Senger, L. J., & Hochuli, A. G. (2022). Combining Muti-Layer Features For Plant Species Classification in a Siamese Network. 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE. doi:10.1109/smc53654.2022.9945552

Munisami, T., Ramsurn, M., Kishnah, S., & Pudaruth, S. (2015). Plant Leaf Recognition Using Shape Features and Colour Histogram with K-nearest Neighbour Classifiers. Procedia Computer Science, 58, 740–747. doi:10.1016/j.procs.2015.08.095

Pérez-Pérez, B. D., García Vázquez, J. P., & Salomón-Torres, R. (2021). Evaluation of Convolutional Neural Networks' Hyperparameters with Transfer Learning to Determine Sorting of Ripe Medjool Dates. Agriculture, 11, 115. doi:10.3390/agriculture11020115

Söderkvist, O. (2001). Computer Vision Classification of Leaves from Swedish Trees. Fonte: https://api.semanticscholar.org/CorpusID:60245555

Sundara Sobitha Raj, A. P., & Vajravelu, S. K. (2019). DDLA: dual deep learning architecture for classification of plant species. IET Image Processing, 13, 2176–2182. doi:10.1049/iet-ipr.2019.0346

Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y.-X., Chang, Y.-F., & Xiang, Q.-L. (2007). A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. 2007 IEEE International Symposium on Signal Processing and Information Technology. IEEE. doi:10.1109/isspit.2007.4458016

Zhang, S., Huang, W., Huang, Y.-a., & Zhang, C. (2020). Plant species recognition methods using leaf image: Overview. Neurocomputing, 408, 246–272. doi:10.1016/j.neucom.2019.09.113

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., . . . He, Q. (2019). A Comprehensive Survey on Transfer Learning. A Comprehensive Survey on Transfer Learning. arXiv. doi:10.48550/ARXIV.1911.02685