Data Augmentation and Transfer Learning Applied to Charcoal Image Classification

Luciana T. Menon[1], Israel A. Laurensi[1], Manoel C. Penna[1], Luiz E. S. Oliveira[2], Alceu S. Britto Jr[1,3]
[1]Pontifical Catholic University of Parana (PUCPR), Curitiba, PR, Brazil
[2]Federal University of Parana (UFPR), Curitiba, PR, Brazil
[3]State University of Ponta Grossa (UEPG), Ponta Grossa, PR, Brazil
alceu@ppgia.pucpr.br

***Abstract*—This paper proposes the evaluation of data augmentation impact in the process of microscopic charcoal image classification. Two data augmentation approaches were explored, namely morphological transformations and sub-images. From the augmented data, a pre-trained Inception-v3 network was used to train a classifier of charcoal species. The best result was found through the technique of sub-images, with an average accuracy of 99.36%.**

***Keywords*—data augmentation, microscopic charcoal image classification, transfer learning**

I. Introduction

Currently, Brazil is the world's largest producer and consumer of wood charcoal, but much of this production is still coming from illegal logging. The production of illegal charcoal is responsible for the annual emission of over 5 tons of carbon dioxide caused by the inadequate removal of trees and the release of carbonization gases [1].

In Brazil, governmental surveillance activities are performed based on the Document of Forestry Origin (DOF), an official document containing information on the origin of the charcoal and the species used to produce it [2]. The government agents must verify the compliance between the loading of charcoal being transported and its DOF. However, the distinction of charcoal species is not a trivial task, making it difficult for environmental control agencies to identify the source of charcoal that is produced, transported and sold illegally [3].

The charcoal classification is based on its anatomical characteristics and thus demands the knowledge of a specialist in wood anatomy [4]. Moreover, the possibility of biases and mistakes by humans has to be considered. Besides that, it is impractical for a human to analyze and identify a large number of species, given that just in Brazil approximately 8000 species of trees are known [2].

In order to improve the monitoring of charcoal source and ease the work of those involved in the control and monitoring of charcoal production, it would be useful to develop fast and efficient techniques for charcoal classification. However, there is a lack of data to conduct experiments in this field due to the difficulties of charcoal image extraction and labeling, and the requirement of specific carbonization equipment and expertise in the area.

A possible approach for charcoal classification is based on its microscopic images, from which some characteristics are extracted and analyzed to discriminate between species. Given the importance of creating a robust classifier for the identification of charcoal species and the problem of lack of data in this context, together with the difficulty in using deep learning techniques with few data, this paper proposes the use of transfer learning in the CNN Inception-v3 [5] and data augmentation techniques, sub-images and morphological transformations, to perform the automatic classification of microscopic images of charcoal.

The main contributions of this paper are three. Direct, the creation of a charcoal species classifier through its microscopic images. Social, in the sense that it facilitates the work of those involved in the control and monitoring of the production of charcoal. And scientific, which is the evaluation of data augmentation techniques in the deep learning classification process.

The remainder of the paper is structured as follows: In section II, the theoretical background of charcoal and wood classification is given. The proposed classification method is presented in section III. Results and concluding remarks are given in section IV and V.

II. Related Works

Successful CNNs, such as "AlexNet" [6], Inception [7], VGG [8] and ResNet [9], have already proven to be very powerful resources in image classification, without the need of handcrafting the features to be classified. Automatic feature extraction and classification can be used in different domains, given that sometimes it is difficult to establish a well-rounded feature extractor that represents a subset to be labeled, especially when a generalization is important.

One of the most important projects when it comes to image classification is the ImageNet [10] competition, entitled ILSVRC (ImageNet Large Scale Visual Recognition Challenge). Since 2010, participants are given the task to classify images of 1000 different classes. This kind of challenge brings to light, every year, a lot of discussion and new ideas to image classification and object recognition. The architectures VGG [8] and Inception [7] yielded one of the highest performances in the ILSVRC of 2014.

978-1-7281-3227-3/19/$31.00 ©2019 IEEE

These CNN architectures can be applied to a variety of domains. Studies around automatically feature extraction and image classification has been increasing in different areas. Given that, from the transfer learning technique, it is feasible to recognize and apply knowledge learned in previous tasks to novel tasks, using the weights of a network already trained in a domain to initialize and retrain a neural network in another domain [11].

It is important to note, that even though CNNs architectures are powerful in the image classification task, they require a large amount of data to be trained with, given the number of parameters and deep layers to convey generalization.

Transformations techniques such as rotation, cropping, mirroring, zooming, and distortion can be used to increase the amount of data used to train these networks. A more accurate classifier can be achieved with data augmentation since with a greater amount of data it is possible to reduce the sensitivity of the model and avoid learning incorrect information [12]. The main idea of data augmentation is that the transformations applied to the data do not alter the semantic meaning of the image, thus allowing the generation of new samples.

Image processing and pattern recognition have already been successfully applied in the classification of charcoal and wood species by some authors.

As proposed in [13], the use of techniques based on dividing the image into several sub-images, and the use of handcrafted features based on Local Binary Patterns (LBP) were applied to the problem of wood charcoal image classification. It was shown that, along with machine learning classifiers, such as Support Vector Machine (SVM) and Random Forest (RF), these techniques demonstrated good results in the observed accuracy of the classifiers.

Sub-images combined with different classifiers trained from feature sets based on color, orientation and texture were used to classify forest wood species [14]. Hardwood and softwood image classification represented through structural features, gray level co-occurrence matrix and texture patterns also have been evaluated in different machine learning classifiers, k-Nearest Neighbors (k-NN), Linear Discriminant Analysis (LDA), and SVM [15].

Deep learning models have been receiving increased attention in recent years and also have been applied to this problem. The usage of CNNs as feature extractors followed by the use of a traditional classifier, such as SVM or RF, was proposed in [13]. In [16], a fuzzy logic pre-classifier was proposed to improve the classification accuracy of wood species. Pore size and distribution in the image were used to cluster the wood image database into smaller databases which were then classified by a neural network. Convolutional Neural Networks were also explored for the classification of macroscopic and microscopic wood images [17].

III. Experimental Details

A. Database

In this work, an open-access database of microscopic images of wood charcoal species, comprising 44 species, was used (see Fig. 1 for some examples). Twelve samples of each species produce a database of 528 images [13]. The samples have 1280 x 1040 pixel resolution, are available in grayscale Tagged Image File Format (TIFF) and were cataloged by the Laboratory of Wood Anatomy, Federal University of Parana, Brazil.

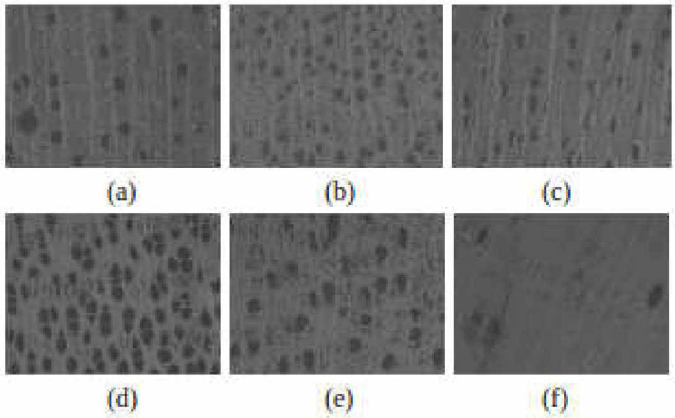

Fig. 1. Database image samples. (a) Apuleia molaris. (b) Aspidosperma populifolium. (c) Astronium gracile. (d) Poeppigia procera. (e) Vochysia densiflora. (f) Hymenaea courbaril.

B. Proposed Method

In the proposed method, two techniques of data augmentation were used: morphological transformations and random cropping, as can be seen in Fig. 2.

The results reported here were compared to those of the database authors [13], using the experimental protocol proposed by them. The main difference between our proposed method and theirs is that here the CNN is used as an end-to-end solution, while they use it just as a feature extractor.

To conduct a 10-fold cross validation, from the original database, 10 new databases were randomly generated maintaining the class distribution. The images were divided into training and testing set, considering 50% of the images for each one. To avoid overfitting, the test images are kept aside and do not interfere in any way in the training. It is important to remember that during the training process, the training images were subdivided into training (80%), validation (10%), and test (10%).

Morphological transformations are a collection of operations that modify the shape of images (see Fig. 3). To evaluate the morphological transformations, we considered the following empirically selected techniques: vertical mirroring; horizontal mirroring; rotation between -10° and 10°; zoom with a factor between 1 and 1.5; and distortion between -5° and 5°.

The protocol for applying the data augmentation techniques is defined as a two-step process: first, both vertical and horizontal mirroring were applied to all images; second, each of the remaining morphological transformations had a 50% chance of being applied to each image. Given this protocol, each augmented base had, after this process, 4400 images with an average of 100 training images per species. The test datasets were not submitted to morphological transformations.

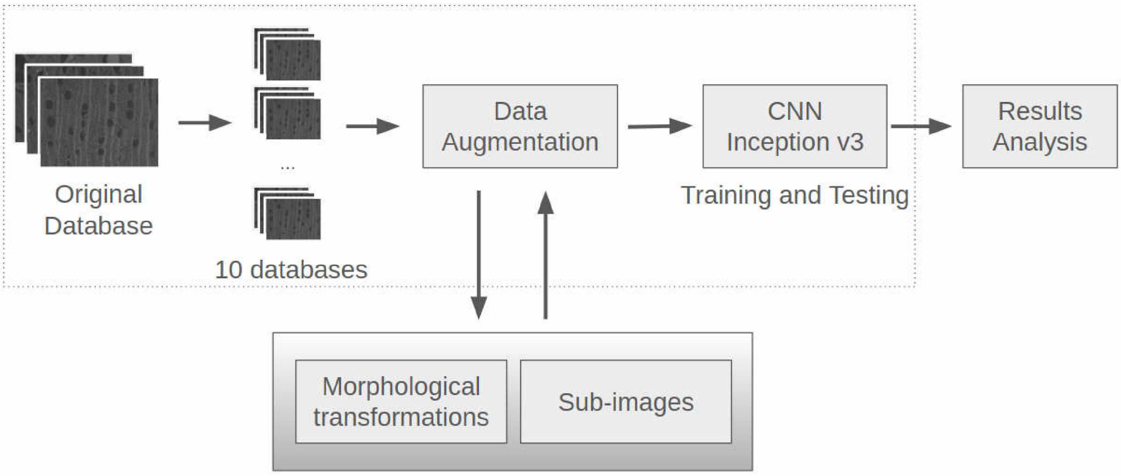

Fig. 2. Experimental protocol pipeline, considering the data augmentation techniques.

To evaluate the random cropping (defined from now on as sub-images) protocol, both training and test datasets were submitted to the process. The sub-images were obtained from each image of the original database. Fig. 4 shows examples of sub-images. Given an image, 25 sub-images were generated with random patches of 256x256 pixels. The number of sub-images as well as their size were empirically selected. This process was applied to all 10 bases. Therefore, each database of sub-images had 6600 images, 150 training images per species.

The accuracy in the sub-images databases was evaluated according to the majority vote, as shown in Fig. 5. For each image, 25 patches were submitted to the CNN and labeled by the classifier. In the end, the label with the highest occurrence was the one chosen to be the label of that image.

IV. Results and Discussion

A. Experimental results

The performance of the proposed method was evaluated for the two data augmentation approaches discussed before. A convolutional neural network, namely the Inception-v3 architecture [5], was used with a learning rate of 0.01 and a training batch size of 100. Tests were conducted with steps ranging from 2000 to 20000 (a step is a forward pass of the batch on the neural network), increasing by 2000 each time. The best results were yielded with the Inception-v3 network fine-tuned with 20000 steps.

The use of data augmentation represented a gain in accuracy using the Inception-v3 network, with an average accuracy of 93.14% when using the morphological transformation dataset and 99.36% when using the sub-images dataset, compared to 89.62% with the original dataset. These results can be observed in Table I.

As shown in Fig. 6, the use of data augmentation techniques provides an improvement on the average accuracy, with respect to the original method.

In order to verify the significant difference in the average accuracy of the classifier trained from different methods of

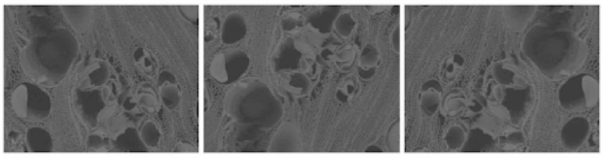

Fig. 3. Examples of images after the morphological transformations.

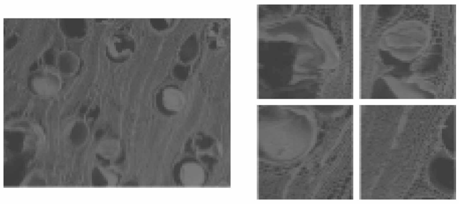

Fig. 4. Examples of sub-images. On the left, the original image and on the right, four sub-images generated from the original.

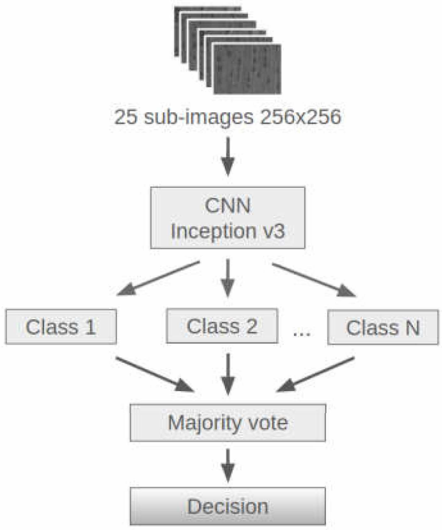

Fig. 5. Sub-imaging process and decision based on majority vote.

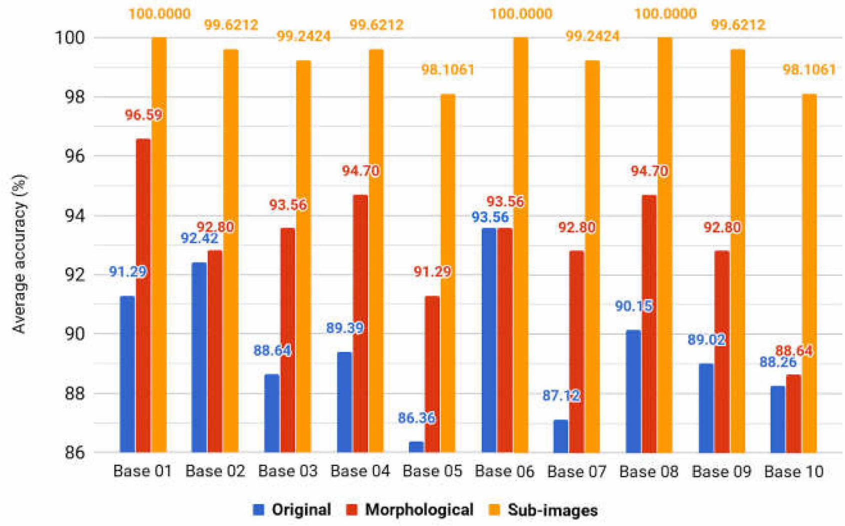

Fig. 6. Results obtained with the CNN for the original dataset and using the morphological transformation and sub-images data augmentation technique.

data augmentation, the Friedman [18] statistical test was applied with a significance level of 0.05. It was concluded that the use of data augmentation had an impact on the observed accuracy. It is also noticed that the rank order reported by the test indicates that the technique of sub-images obtained a better result than the others, as shown in Fig. 7.

Given that there is a significant difference in applying at least one of the data augmentation techniques, the Nemenyi [19] post hoc test was used to identify the internal differences between each of the observed bases. The Nemenyi test compares all observed variables, one against the other, identifying if one variable is significantly different from the other when the observed rank average is greater than the calculated critical distance (CD). In this case, $CD = 1.04$.

As can be seen in Fig. 7, the original base and the dataset augmented with morphological transformations did not present a significant difference ($1.95 - 1.1 = 0.9 < CD$). It is also shown in Fig. 7, that the classifier trained from the dataset augmented with sub-images has a significantly higher accuracy than the others.

It was concluded that the use of sub-images performs better than morphological transformations and is, therefore, the method proposed in this paper to solve the wood charcoal species classification problem.

The results reported in Table II were compared to those of the database authors [13]. In the approach reported by them, the features were extracted from "pool_3" layer of the Inception-v3 and then used to train Linear SVM and RF classifiers. All the results described in [13] refer to 80 sub-images. However, the authors state that similar results can be achieved using a smaller number of patches. For Linear SVM the best performance was achieved with 9 patches and for RF, 40 patches were required to reach the performance shown in Table II. For our proposed model, the reported result refer to 25 patches of the original image.

TABLE I
AVERAGE CLASSIFICATION ACCURACY FOR EACH DATASET USED.

Dataset	Accuracy (%)
Original	89.62 ± 2.27
Morphological	93.14 ± 2.14
Sub-images	99.36 ± 0.72

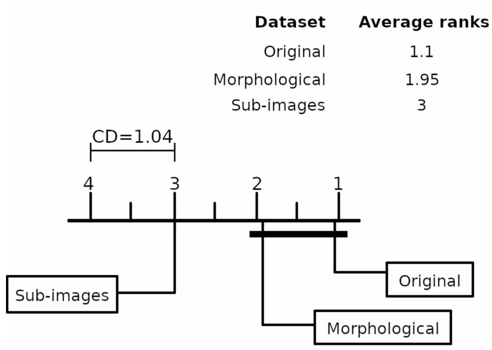

Fig. 7. Nemenyi test from the average ranks obtained through the Friedman test.

TABLE II
RECOGNITION RATES FOR CLASSIFIERS TRAINED FROM INCEPTION-V3 FEATURES.

Method	Accuracy (%)
Random Forest [13]	93.9 ± 4.0
Linear SVM [13]	95.7 ± 4.7
Proposed Model	99.36 ± 0.72

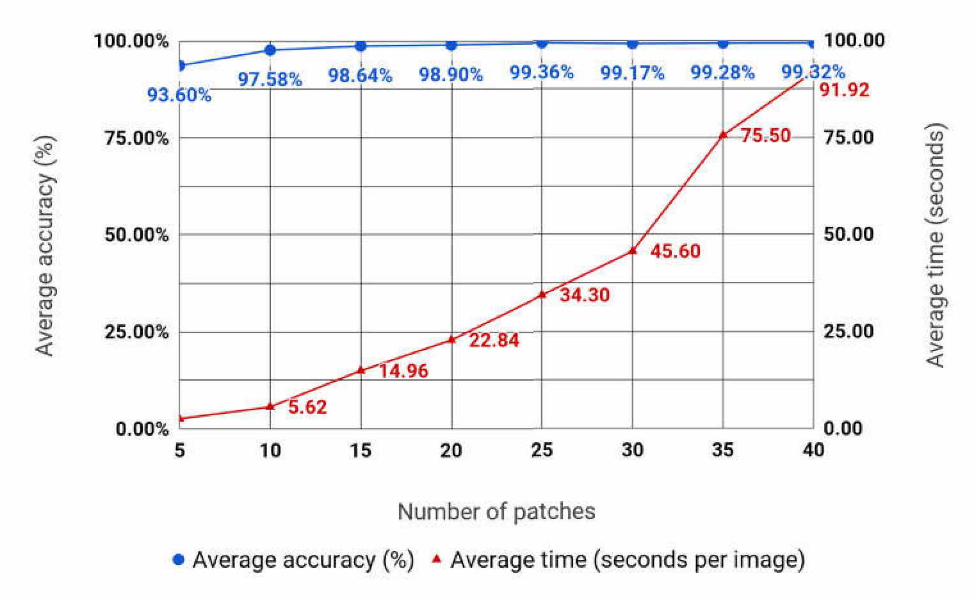

Fig. 8. Average accuracy and the average time taken (in seconds) to execute the test with the Inception-v3 network for each sub-image size analyzed. Results obtained with the CNN for the original dataset and using the morphological transformation and sub-images data augmentation techniques.

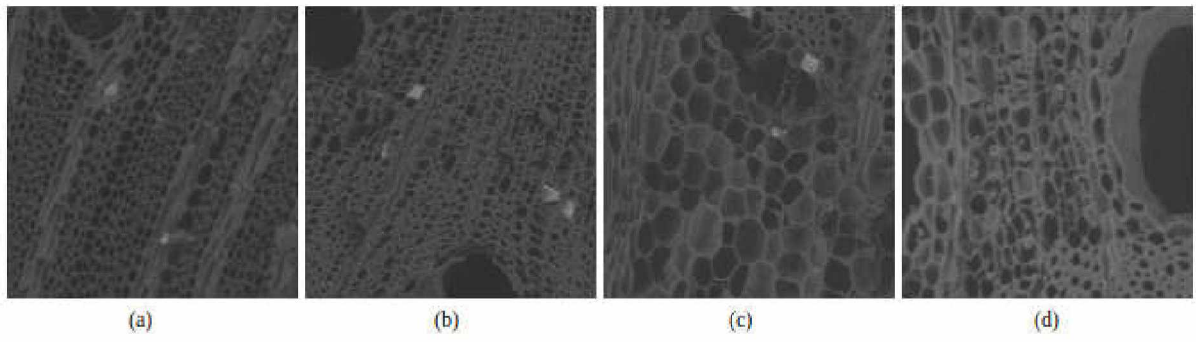

Fig. 9. Examples of misclassified charcoal images: (a) Copaifera langsdorfii misclassified as (b) Hymenaea courbaril, (c) Vatairea guianensis misclassified as (d) Vatairea paraensis.

Accuracy tends to increase with more patches, although the time consumed to test every image also increases. Fig. 8 shows the average accuracy obtained for each number of patches tested. In conjunction with the average accuracy, the time spent for the network (seconds per image) to analyze all patches for a given image was also observed and is reported on the right vertical axis.

Since each training base has 264 images and tests were performed on 10 randomly selected sets of images, 2640 images of charcoal were classified in each of the experiments. Considering our best result of 25 sub-images only 17 were misclassified. From the confusion matrix, it was observed that the most confusing classes were the following: Copaifera langsdorfii and Vatairea guianensis, both with 4 misclassified images. These classes were commonly classified as Hymenaea courbaril and Vatairea paraensis, respectively. Fig. 9 shows two examples of misclassified species.

B. Discussion

The protocol of morphological transformations achieved a slight improvement over the original database, given that a greater number of training examples allows the neural network to better adjust its parameters and thus obtain a better performance.

The high accuracy observed with the sub-images protocol is a reflex of the patterns present in the images. The results show that smaller patches of the charcoal images are sufficient to accurately distinguish between different classes, given that these 256x256 patches still represent the main pattern of each class. Smaller patches could be used to speed up training and prediction time, although the first layer on the Inception-v3 network is a convolution filter of 299x299x3, which could lead to loss of information due to resizing.

Another reason for the high accuracy is the voting process, which gives the neural network more examples and chances for the classification task, while still maintaining a significant difference between classes when working with sub-images.

In regards to the time consumed on the sub-images, 60 patches were also tested, leading to an average of 16 hours (228 seconds for each image) of prediction time across all 10 databases. Although this yielded a higher accuracy (99.47%), the time is exponentially larger (for 25 patches, the average time was 2.5 hours, with 34 seconds for each image) and is impractical for real-time predictions.

As the patches are randomly selected, it was observed that some of them overlapped during the experiments. However, it has been verified that this is very rare, less than 1% of images overlap.

The tests were performed with the Inception-v3 [5], which is a well-known network in the literature for obtaining good results in the image classification task, with a reasonable number of parameters and training time. However, smaller networks could also be tested in order to compare the time and accuracy.

V. Conclusions

We have presented a method to use data augmentation applied to transfer learning for charcoal image classification. Two different techniques were explored, which were morphological transformation and sub-images. Datasets for each technique were generated, using a 50%/50% rule to train and test sets. The pre-trained convolutional neural network Inception-v3 was used for transfer learning on each dataset, including the original, for comparison. The best results were achieved when using 25 patches from the original image, with an average accuracy of 99.36%.

Although higher accuracy could be obtained with more patches, we show that the time taken for the network to process higher numbers (greater than 25) of patches is not viable given the small gain in accuracy. We have compared the results with previous protocols of microscopical charcoal image classification approached in the literature and shown that the use of sub-images using transfer learning with a pre-trained CNN represented a successful gain in accuracy.

Further work can be done to explore other transformations and methods of feature extraction. The use of smaller networks could also be analyzed and even the combination of multiple classifiers. It would also be interesting to apply the same protocol proposed here with other charcoal image datasets.

Acknowledgment

This work was supported by UNIVISION INFORMÁTICA LTDA and CNPq (National Council for Scientific and Technological Development) grant 306688/2018-2.

References

[1] T. A. P. Gonçalves and R. Scheel-Ybert, "Charcoal anatomy of Brazilian species. I. Anacardiaceae," Anais da Academia Brasileira de Ciências, vol. 88, no. 3 suppl, pp. 1711-1725, Oct. 2016.

[2] T. A. P. Gonçalves, J. Sonsin-Oliveira, S. Nisgoski, C. R. Marcati, A. W. Ballarin, and G. I. B. Muñiz, "A contribution to the identification of charcoal origin in Brazil III: microscopic identification of 10 Cerrado species," Australian Journal of Botany, vol. 66, no. 3, p. 255, 2018.

[3] F. M. G. Ramalho, P. R. G. Hein, J. M. Andrade, and A. Napoli, "Potential of near-infrared spectroscopy for distinguishing charcoal produced from planted and native wood for energy purpose," Energy & Fuels, vol. 31, no. 2, pp. 1593-1599, Jan. 2017.

[4] F. Davrieux, P. L. A. Rousset, T. C. M. Pastore, L. A. de Macedo, and W. F. Quirino, "Discrimination of native wood charcoal by infrared spectroscopy," Química Nova, vol. 33, no. 5, pp. 1093-1097, 2010.

[5] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception architecture for computer vision," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84-90, May 2017.

[7] C. Szegedy et al., "Going deeper with convolutions," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[8] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," in 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), 2015.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[10] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009.

[11] S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, Oct. 2010.

[12] A. Heikkinen and M. Barai, "Impact of data augmentations when training the Inception model for image classification," 2017.

[13] T. M. Maruyama, L. S. Oliveira, A. S. Britto Jr, and S. Nisgoski, "Automatic classification of native wood charcoal," Ecological Informatics, vol. 46, pp. 1-7, Jul. 2018.

[14] P. L. P. Filho, L. S. Oliveira, S. Nisgoski, and A. S. Britto, "Forest species recognition using macroscopic images," Machine Vision and Applications, vol. 25, no. 4, pp. 1019–1031, Jan. 2014.

[15] J. Martins, L. S. Oliveira, S. Nisgoski, and R. Sabourin, "A database for automatic classification of forest species," Machine Vision and Applications, vol. 24, no. 3, pp. 567-578, Mar. 2012.

[16] R. Yusof, M. Khalid, and A. S. Mohd Khairuddin, "Fuzzy logic-based pre-classifier for tropical wood species recognition system," Machine Vision and Applications, vol. 24, no. 8, pp. 1589–1604, Jul. 2013.

[17] L. G. Hafemann, L. S. Oliveira, and P. Cavalin, "Forest species recognition using deep convolutional neural networks," in 2014 22nd International Conference on Pattern Recognition, 2014.

[18] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance,". Journal of the American Statistical Association, vol. 32, pp. 674–701, 1937.

[19] P. B. Nemenyi, "Distribution-free multiple comparisons." PhD thesis, Princeton University, 1963.