

Original papers

Estimating soybean leaf defoliation using convolutional neural networks and synthetic images



Lucas Abreu da Silva^a, Patrik Olá Bressan^b, Diogo Nunes Gonçalves^a, Daniel Matte Freitas^a, Bruno Brandoli Machado^a, Wesley Nunes Gonçalves^{a,*}

^a Federal University of Mato Grosso do Sul, Rua Itibiré Vieira, s/n, 79907-414 Ponta Porã, MS, Brazil

^b Federal Institute of Mato Grosso Sul, Rua Pedro de Medeiros, s/n, 79310-110 Corumbá, MS, Brazil

ARTICLE INFO

ABSTRACT

Keywords:

Soybean defoliation
Convolutional neural networks
Synthetic images
Regression

Agriculture plays an important role in the economy of several countries, contributing from the production of food and income to the generation of jobs. To improve productivity in agriculture, proper crop management should be accomplished through pest control. One approach is to monitor the defoliation level, that is, the percentage of leaf damaged by insects. Despite the importance, this monitoring is performed manually most of the time, which affects reliability as well as it is considered to be a time-consuming task. In this paper, we propose a new fully-automatic method to estimate the defoliation level based on convolutional neural networks (CNN). The main CNN architectures (AlexNet, VGGNet and ResNet) were adapted from classification to regression by replacing the softmax layer with a fully-connected layer with one neuron and a sigmoid activation function. Since CNNs require a large number of training examples, this paper also proposes approaches for generating synthetic defoliation images. In this way, our method is trained only with synthetic images and evaluated using real images. In the experiments, we obtained a root mean square error of only 4.57 even for images with severe defoliation. Additionally, we presented experimental evidence that the proposed method reconstructs the damaged leaf parts to then estimate the defoliation level.

1. Introduction

Agriculture is one of the most important activities for developing countries, contributing to the production of food and raw material. In addition, agriculture provides employment opportunities for a large portion of a country population and represents a large part of the national income (FAO, 2017). Despite the importance and productivity growth in recent years, agriculture in 2050 will have to produce almost 50 percent more than it did in 2012 to meet the demand of the world population (FAO, 2017).

To increase productivity, proper management of a crop, including pest control, is crucial. Annually, plant pests cause crop losses of 20 to 40 percent of production (FAO, 2017). The losses caused by invasive insects cost the global economy around US\$70 billion annually (Bradshaw et al., 2016). The main consequence of invasive insects is the herbivory and injury that result in a functional reduction of the total leaf surface of the plant. The loss of leaf area, namely defoliation, affects the photosynthesis and energy production of the plant, which decreases its nutrition and consequently the final production of the crop. In this way, it is important to monitor the defoliation level to take preventive actions. For soybeans in the vegetative stage, it is

recommended to start treatments when the defoliation level reaches 30% (Kogan et al., 1977).

In most cases, defoliation level is estimated visually by experts using a guide (examples of leaves and their respective defoliation level) (Kogan et al., 1977). Another common way is to use the grid counting method (Kvet and Marshall, 1971), in which the expert positions the leaf in a grid and counts the number of squares that overlap the leaf in order to estimate its area. Therefore, defoliation level estimation using the above techniques is a time-consuming and subjective task. In addition, the defoliation level is generally overestimated by these techniques, which leads to unnecessary insecticide applications (Wilhelm et al., 2000). Also, leaf area meters devices, such as the LI-3000A and LI-3100 (Barclay et al., 2000; dos Santos et al., 2016), are currently available to automatically estimate the leaf area. To estimate the defoliation level, these devices need to calculate the leaf area before and after herbivory. Although the leaf area after herbivory can be correlated to the defoliation level, it does not estimate the damaged area, especially when it occurs at the edges.

To avoid subjectivity and speed up the measurement process, computational methods have been proposed. Mobile applications (Gong et al.,

* Corresponding author.

E-mail addresses: lucasabreudasilva@gmail.com (L.A. da Silva), patrik.bressan@ifms.edu.br (P.O. Bressan), daniel.freitas@ufms.br (D.M. Freitas), brunobrandoli@gmail.com (B.B. Machado), wesley.goncalves@ufms.br (W.N. Gonçalves).

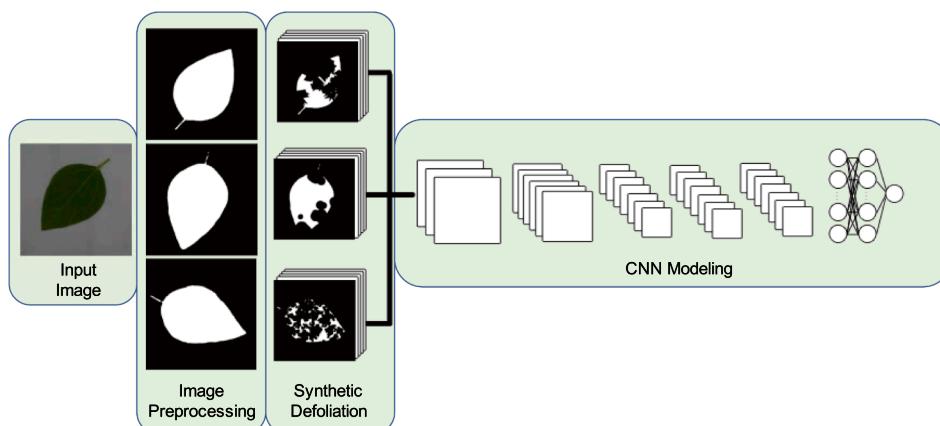


Fig. 1. The proposed method can be described by the steps illustrated in this figure. Initially, images are preprocessed and then synthetic defoliation is generated using the proposed approaches. Then, a CNN is trained with regression to estimate the defoliation level.

2013b; Gong et al., 2013a) and desktop software (Karatassiu et al., 2015; Parmar et al., 2016; Houborg and McCabe, 2018; Fan et al., 2018) are available to estimate the leaf area, although they do not estimate defoliation level. Similar works to estimate the defoliation level were proposed in Machado et al. (2016) and Nazaré-Jr et al. (2010). Machado et al. (2016) presented a mobile application that estimates the defoliation level using Bezier curves to restore the original leaf border. Although this method increase reliability and speed, it still requires the intervention of an expert, for example, to draw the leaf border using the Bezier curve. On the other hand, Nazaré-Jr et al. (2010) proposed an automatic method to estimate the defoliation level. In this method, corners of damaged areas are detected and then line segments are drawn between each corner. However, the method is not suitable for severe defoliation, where the corners required by the method are not present in the image.

To overcome these issues, this paper proposes a method to estimate the defoliation level of soybean leaves using deep learning, which is, to the best of our knowledge, the first fully automatic method with low mean error for severe defoliation. The proposed method adapts the traditional convolutional neural networks (CNN) for regression by replacing the last layer with a fully-connected layer composed by one neuron and a sigmoid activation function. However, it is well known that deep CNN training requires a large number of examples (Oquab et al., 2014), in our case, a large set of images with defoliation and their estimated level by experts. To deal with the number of training images, this paper also proposes methods for generating images with synthetic defoliation. Through these methods, CNN training can be done with the desired number of images at no cost to obtain them.

After the training using only synthetic images, the proposed method was evaluated in real images containing defoliation caused by *S. frugiperda* caterpillars. Experimental results show that the proposed method obtained a mean squared error of only 4.57, even for leaves with severe defoliation in

which a large part of the leaf is damaged. During the analysis of the proposed method, we found that the CNN layers learned to reconstruct parts of the leaf to then estimate the defoliation level. This finding and the quantitative results show that the proposed method is robust and can be used to estimate the defoliation level properly.

This paper is organized as follows. A new method for estimating the defoliation level is presented in Section 2. In Section 3 we present the experimental setup, which includes the image dataset, training configurations and evaluation metrics. In Section 4 experimental results and discussion are provided. Finally, Section 5 concludes the paper and provides the future works.

2. Proposed method

In this section, we describe the proposed method for estimating defoliation in soybean leaves using convolutional neural networks (CNN). Basically, the proposed method can be described in three steps: (i) image preprocessing, (ii) generation of synthetic defoliation, (iii) CNN modeling and training. Fig. 1 illustrates the steps of the proposed method that are detailed in the sections below.

2.1. Image preprocessing

First the input image is resized to a fixed size of 256×256 pixels that corresponds to the CNN input size. This resizing aids in the image standardization, besides contributing to the computational cost. The colored image is then binarized by applying the Otsu thresholding (Bangare et al., 2015) so that the image contains only two regions corresponding to the leaf and the background. Finally, the image is rotated at a random angle in order to make our method invariant to rotation. Fig. 2 shows the preprocessing of an input image.

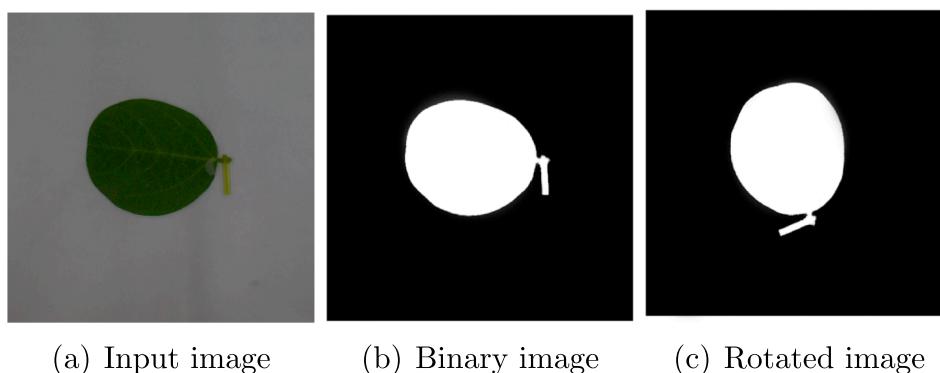


Fig. 2. Example of preprocessing of an input image. The image is resized to 256×256 pixels, binarized and rotated at a random angle.

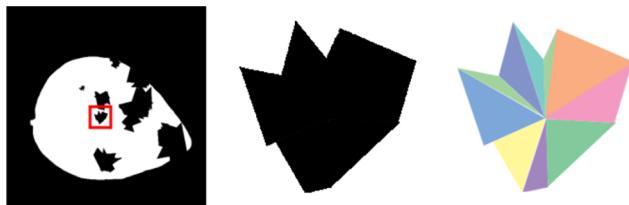


Fig. 3. Example of synthetic defoliation generated by the Polygonal method.



Fig. 4. Example of synthetic defoliation generated by the Circular method.

2.2. Synthetic defoliation methods

To achieve good accuracy, training of most methods relies on huge amounts of labeled data. However, acquiring labeled images for defoliation requires the manual labor of a specialist, making it a time-consuming and expensive process. To overcome this problem, this work proposes three methods for synthetic defoliation. Through these methods, algorithms can be trained with thousands of images at no cost of development.

The synthetic defoliation methods receive a pre-processed image and return a new image with defoliation and its defoliation level. Basically, the proposed methods remove leaf-belonging pixels in different ways to simulate an actual defoliation. The defoliation level is then estimated by the number of pixels removed n relative to the initial

leaf area a , according to Eq. (1). The synthetic defoliation methods, called Polygonal, Circular 1 and Circular 2, are described below.

$$d = 100 * \frac{n}{a} \quad (1)$$

- **Polygonal Defoliation:** the first method simulates defoliation using polygons formed in the leaf area. Given a random pixel p belonging to the leaf, triangles of random sizes are generated with one end in p , forming a polygon as shown in the Fig. 3. The number of polygons is chosen according to the desired defoliation level. Fig. 5(a) illustrates examples of synthetic defoliation generated by the polygonal method.
- **Circular Defoliation 1:** this method consists of making circles in the leaf region. Given a random pixel p of the leaf, we generate a main circle with random radius and center in p . Then, secondary circles with different radii are generated in the circumference of the main circle, as shown in Fig. 4. Examples of synthetic defoliation using the Circular method are shown in Fig. 5(b).
- **Circular Defoliation 2:** the latter method is similar to the previous one, in which a main circle with center in a random pixel of the leaf is generated. However, secondary circles with random radius are generated within the main circle. Fig. 5(c) shows examples of synthetic defoliation using the circular method.

2.3. CNN modeling and training

Given a set of images generated by the synthetic defoliation methods, we trained a convolutional neural network. Due to the recent results, three architectures have been evaluated: AlexNet (Krizhevsky et al., 2012), VGGNet (Simonyan and Zisserman, 2014) and ResNet (He et al., 2016). These architectures were proposed for classification problems, that is, given an input image the objective is to classify it into one of the known classes. To adapt an architecture to a regression problem, we have removed the softmax layer that returns the class probabilities and added a fully-connected layer with one neuron and a

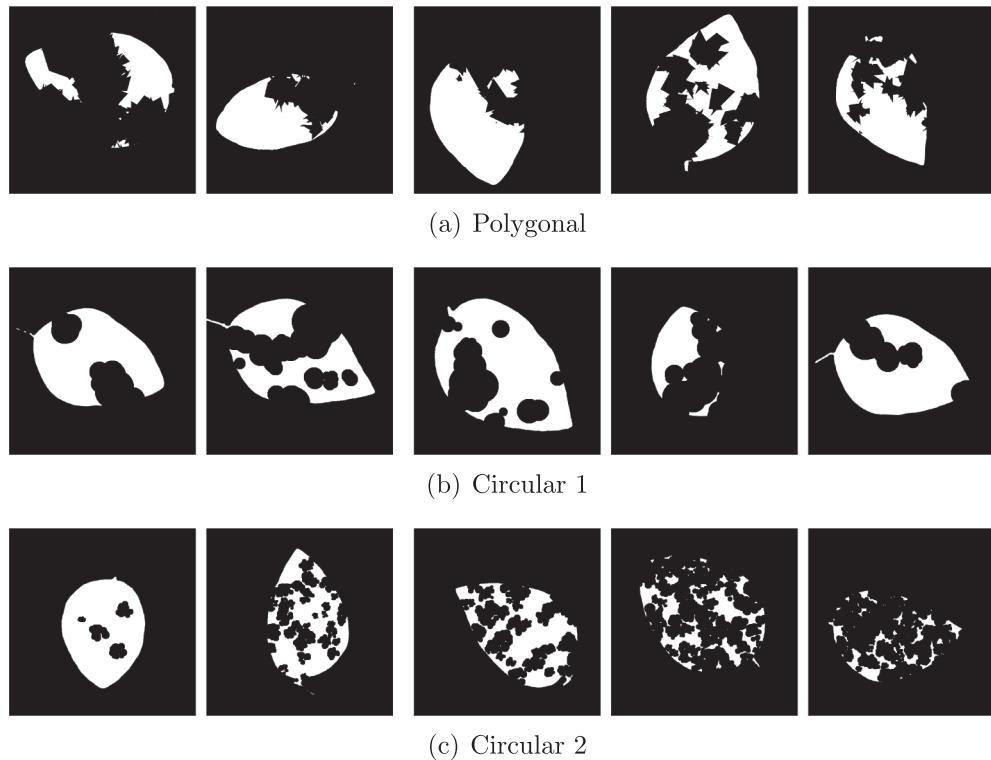


Fig. 5. Example of five images generated by the three methods: polygonal, circular 1 and circular 2.

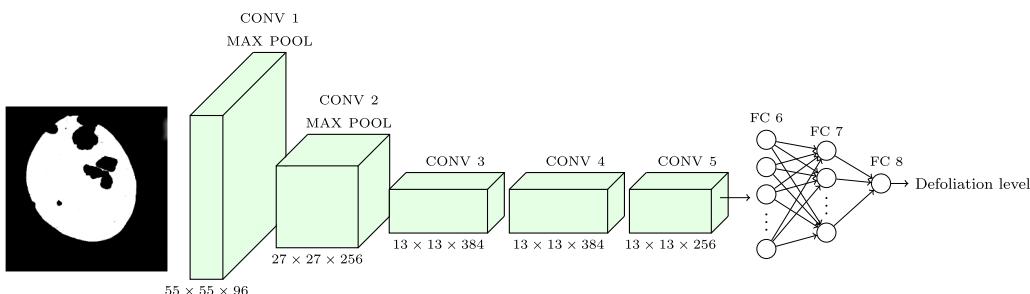


Fig. 6. Example of AlexNet Activation Maps. This architecture consists of convolutional, max-pooling and fully-connected layers, and the last layer contains only one neuron to estimate the defoliation level.

sigmoid activation function to estimate the defoliation level. Also, we have trained the architectures using Root Mean Square Error (RMSE) loss function instead of cross-entropy.

2.3.1. AlexNet

AlexNet (Krizhevsky et al., 2012) is the architecture responsible for the recent success of convolutional neural networks, mainly because of its results in the ILSVRC ImageNet competition. This architecture consists of five convolutional layers, three max-pooling layers with filters of size 3×3 applied with a stride of 2, and three fully-connected layers, the first two with 4096 neurons and the last one with 1000 neurons. The convolutional and fully-connected layers are followed by the ReLU activation function, except the last fully-connected layer which is followed by softmax action function. Example of AlexNet activation maps after the proposed modification can be seen in Fig. 6. We can observe that the last layer is composed of one neuron responsible for estimating the defoliation level. In addition, the last layer after the modification is followed by a sigmoid action function in order to return values between 0 and 1 (defoliation level).

2.3.2. VGGNet

VGGNet (Simonyan and Zisserman, 2014) proposed a significant increase in the number of layers when compared to AlexNet, enabling an improvement in image recognition. This architecture consists of 13 convolutional layers, 5 max-pooling layers and 3 fully-connected layers.

Each convolutional layer has filters with a small receptive field of 3×3 , stride of 1 and padding that preserves the spatial resolution (e.g., the padding is 1 for 3×3 convolutional layers). The number of filters varies from 64 to 512 depending on the layer depth. The five max-pooling layers are performed over a 2×2 pixel window using stride of 2. All convolutional layers are followed by the ReLU activation function.

The final part of the VGGNet is composed of three fully-connected layers. The first two fully-connected layers have 4096 neurons and ReLU activation function. In this architecture after the modifications, we also have only one neuron in the last fully-connected layer with a sigmoid activation function.

2.3.3. ResNet

To train deeper convolutional neural networks, He et al. (2016) proposed a residual learning framework. The core idea of ResNet is to create residual blocks composed of convolutional layers. The main difference is that the input of the block is added to its output obtained after the block convolutions. These special skip connections simplify the optimization of convolutional neural networks.

In this work, we modify the ResNet with 50 layers so that the last layer has one neuron with a sigmoid activation function. In this architecture, the first convolutional layer has 7×7 filters with a stride of 2. A 3×3 max-pooling layer with stride of 2 is performed after the convolutional layer. Then a stack of residual blocks are applied, which are composed of convolutional layers with 1×1 and 3×3 filters. Further details on all parameters can be found in He et al. (2016).

Finally, an average-pooling and a fully-connected layer

3. Experiments

3.1. Image dataset

To evaluate the proposed method, we have used the image dataset presented by Sarath et al. (2015). To build this image dataset, soybean cultivars (DM 6563 RSF IPRO and BRS 284) were grown in pots in a greenhouse. After 40 days of germination, 150 leaves were removed and photographed on a white paper background using a Sony Alpha DSLR-A350 camera. Example of five leaves can be seen in Fig. 7(a).

Then, 24-h fasting *Spodoptera Frugiperda* caterpillars were placed for herbivory. All caterpillars were 10 days old. Twenty-four hours after the placement of the caterpillars, image capture of the leaves was performed again. Example of five leaves after the herbivory can be seen in Fig. 7(b). Thus, two images (before and after) were obtained for each of the 150 leaves.

The defoliation level of each leaf was estimated using the software proposed in Sarath et al. (2015) and corroborated by the grid count method. Basically, the software calculates the leaf area before and after the herbivory to estimate the percentage of defoliation. It is important to mention that the objective of the proposed method is to estimate the defoliation level using only the leaf image after herbivory.

3.2. CNN training

CNN training was performed using only images generated by synthetic defoliation methods. In this way, three image datasets were created, one for each synthetic defoliation method (Polygonal, Circular 1 and Circular 2 Defoliation). For each dataset, we generated 10,000 images that are equally spaced in each 10% range of defoliation. Therefore, 1000 images have defoliation between 0% and 10% and so on. In addition, a fourth dataset composed by all images of the datasets was created to evaluate the influence of the three datasets together in the training. The objective is to show that CNNs trained only with synthetic images are capable of estimating the defoliation of an actual image.

CNNs were trained using Stochastic Gradient Descent (SGD) for 30 epochs with learning rate of 0.01. For AlexNet and VGGNet, it was used a batch size of 100, in contrast, ResNet was trained with batch size of 10. All CNNs were trained in a computer with i7-5820 3.3GHZ CPU, 32 GB RAM, and NVidia Titan XP GPU. Our implementation was written in python using the Tensorflow library.¹

Table 1 shows the total training time using 10000 images during 30 epochs. AlexNet took about 10 min to complete the training while ResNet and VGGNet took approximately 46 min and 11 h respectively. The longer time of ResNet and VGGNet is due to the greater number of layers compared to the number of AlexNet layers.

¹ <http://tensorflow.org/>.

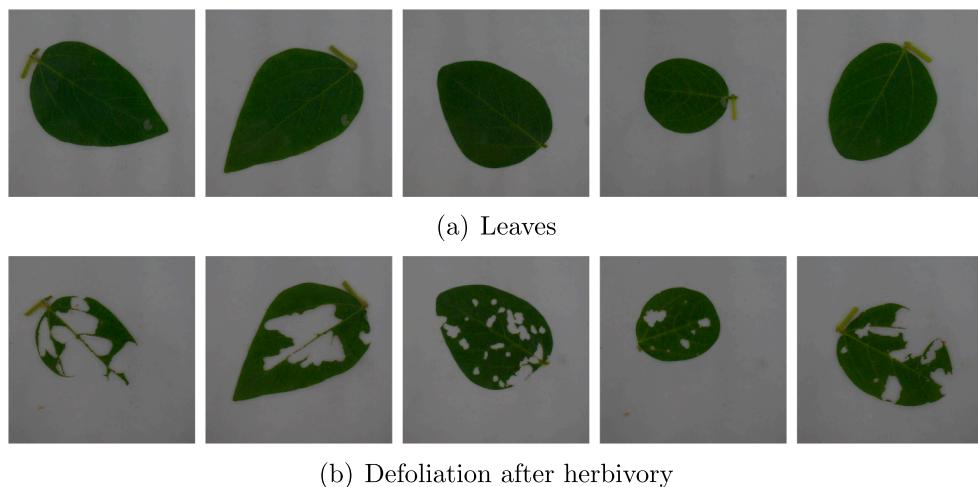


Fig. 7. Images of leaves before and after herbivory. Defoliation level can be estimated by the ratio of the leaf area of both images.

Table 1
Total training time in a set of 10,000 images during 30 epochs.

Architect.	Training time	Time per epoch
AlexNet	10 m 28 s	21 s
ResNet	46 m 58 s	1 m 34 s
VGGNet	11 h 13 m 45 s	22 m 27 s

According to the loss function of Fig. 8, it is shown that the training occurred properly, obtaining a low value after 30 epochs. We can also observe from this figure that 30 epochs are sufficient, since the loss function has stabilized.

3.3. Evaluation metric

After the training, the proposed method for defoliation estimation was evaluated using the dataset of real images (Section 3.1). Therefore, it is important to mention that the proposed method is trained only with synthetic images and evaluated using real images.

Given the defoliation level d_i for image i and the estimate of defoliation \hat{d}_i obtained by the proposed method, Root Mean Square Error (RMSE) can be calculated by Eq. (2).

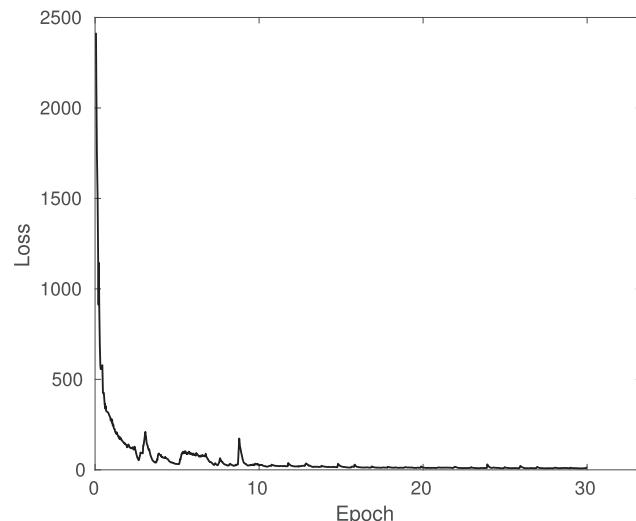


Fig. 8. Loss function obtained during the training of the proposed method.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i - \hat{d}_i)^2} \quad (2)$$

where n is the number of images.

4. Results and discussions

In this section, we present the results and discussions in three experiments: quantitative evaluation using root mean square error, qualitative evaluation, and interpretability.

4.1. Quantitative evaluation

Table 2 presents RMSE for the proposed method using different CNN architectures and synthetic defoliation methods for training. As we can see, AlexNet achieved the best result for Polygonal and Circular 1 with RMSE of $4.95(\pm 6.82)$ and $5.02(\pm 6.61)$, respectively. On the other hand, VGGNet obtained the best result of $6.05(\pm 8.95)$ for Circular 2. ResNet presented inferior results when compared to AlexNet and VGGNet in the three synthetic defoliation methods. This indicates that architectures with few convolutional layers already obtain good results, since the input images are binary.

Comparing the synthetic defoliation methods, Polygonal obtained the best results for all architectures, followed by Circular 2 and Circular 1. However, the best result is obtained when AlexNet is trained with images generated by all three synthetic defoliation methods. This result of only $4.57(\pm 5.80)$ is very relevant, since the proposed method was trained only with synthetic images and validated with real ones.

Fig. 9 shows a histogram of absolute errors obtained by the proposed method using AlexNet. For an image i , the absolute error e_i can be calculated by Eq. (3), where d_i is the ground truth defoliation level and \hat{d}_i is the defoliation level estimated by our method. It is observed that most of the absolute errors were between [0, 1], that is, 32 of the 125 images had an absolute error smaller than 1. We can also observe that only five images had an absolute error greater than 10, which is an

Table 2

Root mean square error (RMSE) for architectures and synthetic defoliation approaches used by the proposed method. The proposed method was trained only with synthetic images and the RMSE was calculated only for real images.

Architect.	Polygonal	Circular 1	Circular 2	All
AlexNet	$4.95 (\pm 6.8)$	$5.02 (\pm 6.6)$	$7.63 (\pm 13.2)$	$4.57 (\pm 5.8)$
VGGNet	$5.29 (\pm 7.6)$	$6.48 (\pm 8.1)$	$6.05 (\pm 8.9)$	$4.65 (\pm 6.4)$
ResNet	$12.58 (\pm 17.1)$	$15.04 (\pm 18.5)$	$13.35 (\pm 17.7)$	$14.60 (\pm 18.8)$

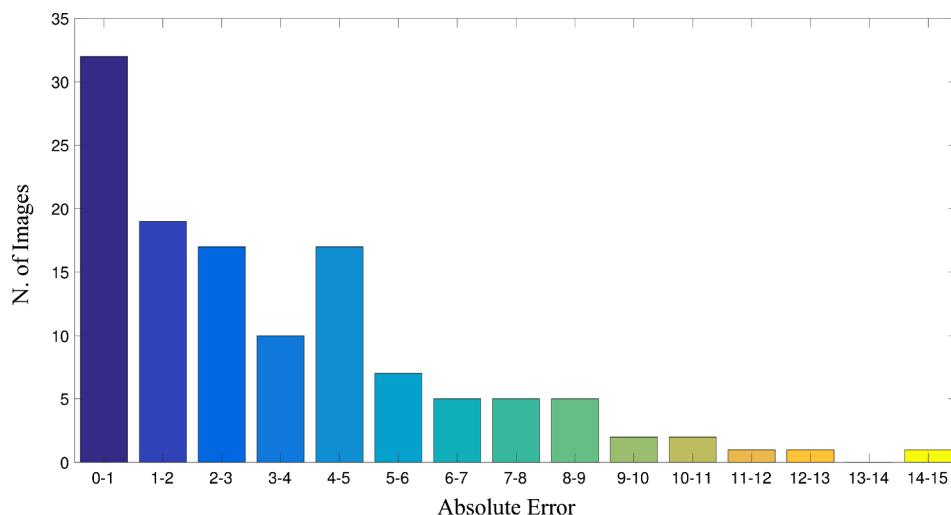


Fig. 9. Histogram of absolute errors obtained by the proposed method. The x-axis represents error range while the y-axis represents the number of images in each range.

expressive result for a fully automatic method. The maximal defoliation level of the test set is 64%. In this sample, our method estimated the defoliation level as 55.13%, an absolute error of 8.87.

$$e_i = |d_i - \hat{d}_i| \quad (3)$$

To visualize the distribution of estimated and ground truth defoliation levels, Fig. 10 shows two boxplots. As we can see, the two distributions are similar with mean and standard deviation close, which corroborates the defoliation estimation performed by the proposed method.

4.2. Qualitative evaluation

For qualitative evaluation of the proposed method, Fig. 11 shows examples of images and their respective defoliation level. Fig. 11(a)-(d) present examples of good estimates obtained by the proposed method, which achieved errors less than 1% in the defoliation level. On the other hand, Fig. 11(e)-(h) present examples of poor estimates of the defoliation level performed by the proposed method. As expected, the most difficult cases occur when the defoliation level is high and at the edges, where the method would have to rebuild the leaf shape. Despite

the challenges, the highest error of these examples was approximately 12%, which is similar to human perception (Shen, 2003).

On the other hand, Fig. 11(i)-(l) present examples of severe defoliation whose level was properly estimated by the proposed method. These results show that, even in leaves with severe defoliation, the proposed method can lead to a good estimate in most cases. Hence, the approach described here has proven to be effective for defoliation estimation.

4.3. Interpretability of the proposed method

To visualize the learned features of the proposed method, we select the activation maps of the first, second and third layers, which are shown in Fig. 12. It is possible to notice that after each layer, the proposed method reconstructs parts of the leaf, which would facilitate the defoliation estimation. After the first layer, the activation maps show that the inner regions of the leaves have been reconstructed. It was possible to visualize that the reconstruction continues in the second and third layers according to Fig. 12.

The proposed method can be interpreted as gradually transforming the leaf image into a representation in which the defoliation level can be estimated by regression. To visualize the space of this representation, we fed our method with leaf images to obtain high-dimensional vectors of the last layer. Then, we embed these vectors in two dimensions using t-Distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten, 2014). The result can be seen in Fig. 13, where each point corresponds to an image and its circumference is proportional to the defoliation level. We can observe that images with high defoliation level were grouped in the upper right corner while images with low defoliation level were grouped in the lower left corner. Thus, it is possible to validate the hypothesis that the information extracted by the proposed method is relevant for estimating defoliation level.

4.4. Comparison with BioLeaf - Foliar Analysis

In this section, we compare the proposed method with BioLeaf - Foliar Analysis (Machado et al., 2016), a mobile application that estimates the defoliation level by counting the leaf holes. Given an RGB image, it is converted to CIE La*b* color space, binarized by Otsu thresholding and smoothed out for noise suppression. When the leaf has defoliation at the borders, the user must draw the missing parts manually with the help of Bezier curves. Finally, the defoliation level is

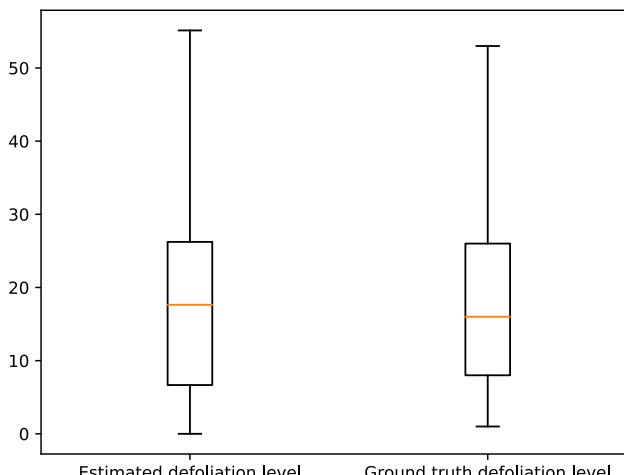


Fig. 10. Distribution of estimated and ground truth defoliation levels.

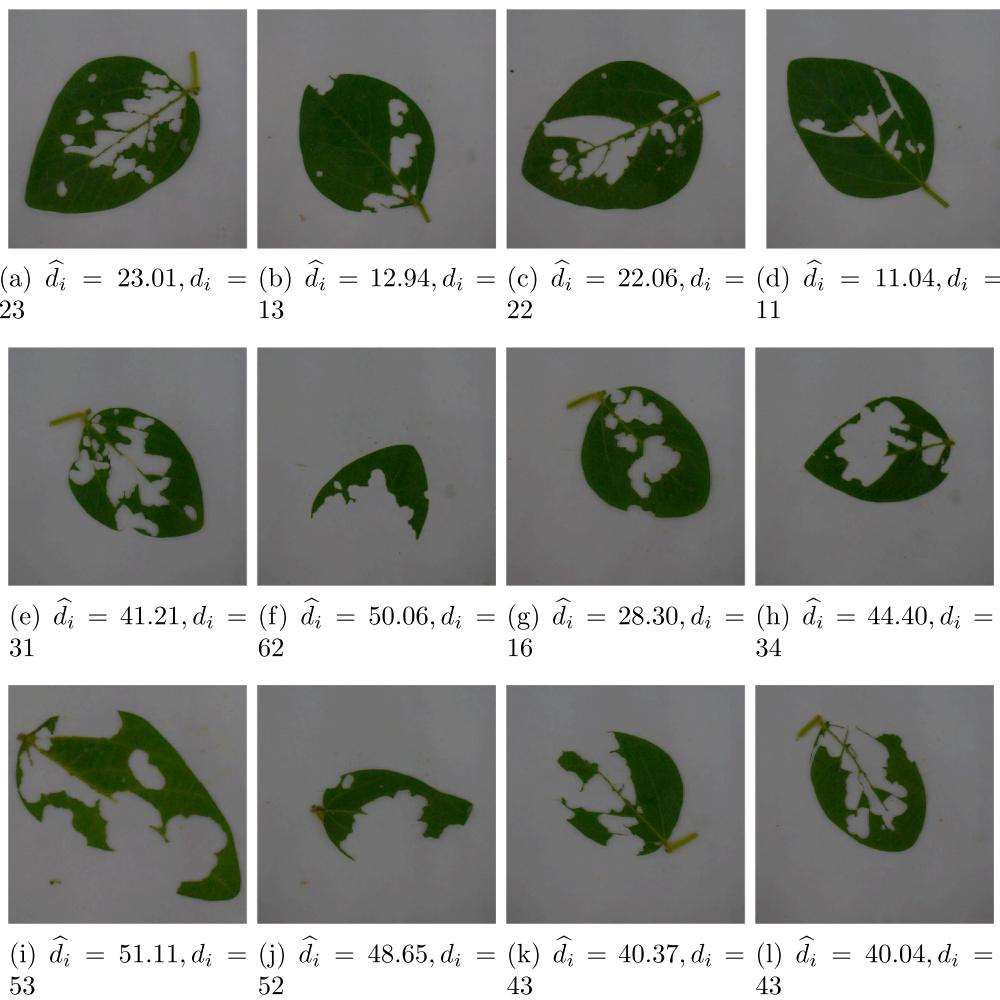


Fig. 11. Examples of images with the defoliation level d_i and its estimation \hat{d}_i obtained by the proposed method. The first row corresponds to the examples with good estimates while the second row presents unsuitable estimates. The third row presents challenging examples whose proposed method has reached a good estimate.

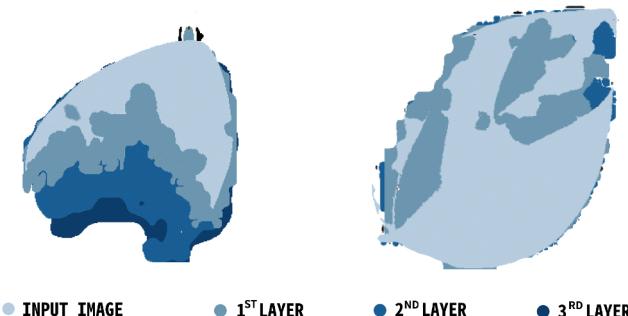


Fig. 12. Activation of the first, second and third layers of the proposed method. We can observe that the hidden layers reconstruct the damaged parts of the leaf.

estimated by counting the number of pixels belonging to leaf holes in relation to the total leaf area.

For comparison with BioLeaf, three users estimated the defoliation level and drew the edges when necessary. Thus, the BioLeaf results correspond to the average of the three estimates. Table 3 presents the estimation of the defoliation level obtained by the proposed method and BioLeaf. In the first row, we show the leaves in which the proposed method was superior to BioLeaf. As we can see, the estimation of the proposed method is better when defoliation occurs at the leaf borders. On the other hand, BioLeaf presents good precision when the

defoliation occurs in the inner regions or in small parts of the border, as presented in the leaves of the second row of the figure.

Fig. 14 shows the defoliation estimated by the proposed method (blue circles) and BioLeaf (green circles) in 58 randomly chosen leaves from the image dataset of Section 3.1. The x-axis represents the defoliation while the y-axis represents the estimation obtained by the methods. The radius of the circles is proportional to the error.

The methods presented good estimates for both low and high defoliation since they have high correlation with the ground truth. Pearson correlations for the proposed method and BioLeaf are respectively 0.987 and 0.986. Despite the similar results, it is important to emphasize that the proposed method is completely automatic, making the estimation process much faster.

4.5. Leaf shape influence

The performance of the proposed method was also evaluated in leaves with different shapes. We have obtained images of two species, *Acer Campestre* and *Juglans Regia*, from the Middle European Woody Plants (MEW) dataset (Novotný and Suk, 2013). The images were submitted to synthetic defoliation methods in order to obtain 10000 training images and 1000 test images. Both training and test sets have different defoliation levels that can be seen in Fig. 15.

Initially, we evaluated the ability of the proposed method to predict defoliation of other species when trained only on soybean leaves. In this scenario, the first row of Table 4 shows that the proposed method was

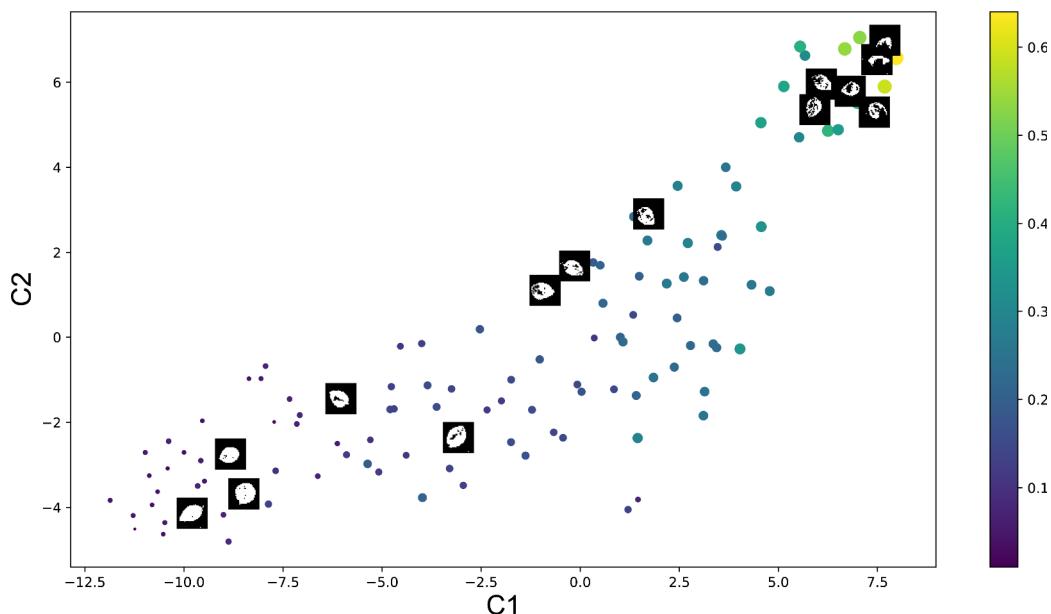


Fig. 13. Projection of the last layer in two dimensions for different images using t-SNE. It is observed that the images were grouped according to the defoliation level.

Table 3

Comparison of the defoliation level estimated by the proposed method and BioLeaf. The ground truth defoliation level was obtained as described in Section 3.1.

Methods				
Ground Truth	24	15	41	53
Proposed Method	26.52	14.56	42.64	51.11
BioLeaf	29.66	13.28	44.47	50.02
Methods				
Ground Truth	7	16	17	25
Proposed Method	5.05	14.63	14.97	23.02
BioLeaf	6.1	15.13	15.14	25.49

able to predict the defoliation level with good RMSE (9.82 and 9.62) even when trained in images of another species. We then trained our method using images with the new species and the RMSE is presented in the second row of Table 4. As expected, the RMSE decreased for both species presenting a very promising value (e.g., from 9.82 to 5.17).

5. Conclusion

Estimating defoliation is a crucial step in the adoption of insect-pest control strategies. To automate this process, this paper proposed a completely automatic method to estimate the defoliation level using convolutional neural networks (CNNs). In the proposed method, we replace the last layer of the CNNs (responsible for classification) by a layer to perform the regression of the defoliation level. Since CNNs require many examples for training, we also proposed approaches for

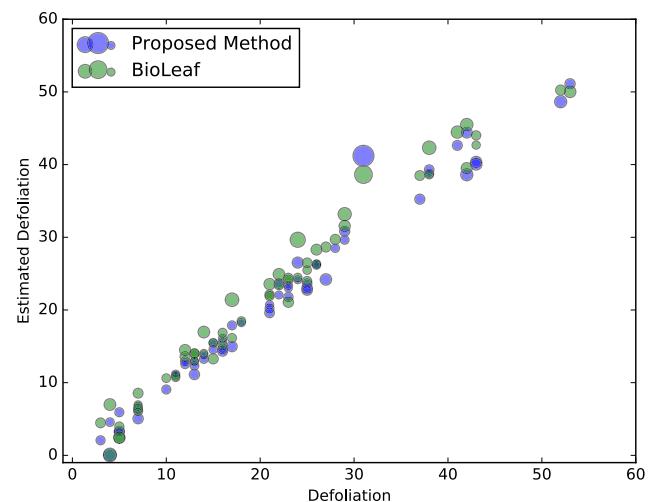


Fig. 14. Comparison of the proposed method and BioLeaf for 58 randomly leaves. Pearson correlation for the proposed method and BioLeaf are 0.987 and 0.986, respectively.

generating images with synthetic defoliation to train our method.

In the experiments, we have evaluated three CNN architectures (AlexNet, VGGNet and ResNet) that were trained only with synthetic images generated by the proposed approaches. We found that AlexNet estimated defoliation adequately with root mean square error of only 4.57, which is an expressive result for leaves with severe defoliation. In the proposed method analysis, it was observed that the internal layers reconstructed the missing parts of the leaf to perform the defoliation estimation. In addition, visualization of the last layer projected for two dimensions showed that the proposed method extracted information directly related to the defoliation level. As part of the future works, we intend to extend the method to color images and include new methods to generate synthetic defoliation. We also intend to evaluate the influence of including real images in training.

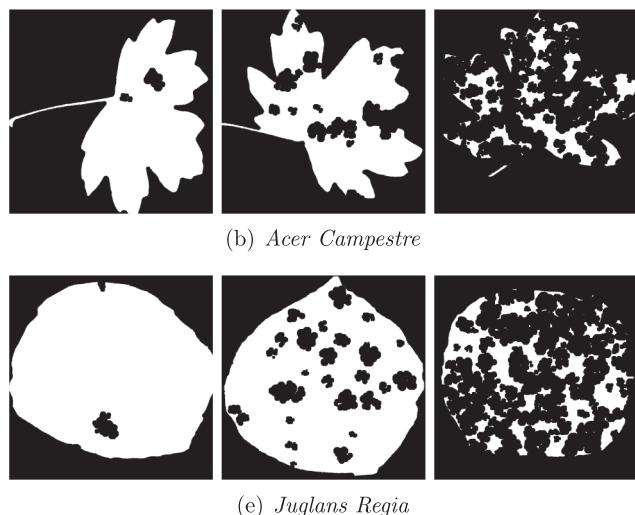


Fig. 15. Image examples with defoliation of two species: *Acer Campestre* and *Juglans Regia*.

Table 4
Root mean square error (RMSE) for two species using the proposed method.

	<i>Acer Campestre</i>	<i>Juglans Regia</i>
No training	9.82 (± 11.9)	9.62 (± 6.8)
Finetuning	5.17 (± 4.4)	4.96 (± 4.6)

Acknowledgments

This work was supported by the FUNDECT - State of Mato Grosso do Sul Foundation to Support Education, Science and Technology, CAPES - Brazilian Federal Agency for Support and Evaluation of Graduate Education, and CNPq - National Council for Scientific and Technological Development. The Titan X Pascal used for this research was donated by the NVIDIA Corporation.

References

- Bangare, S., Dubal, A., Bangare, P., Patil, S., 2015. Reviewing otsu's method for image thresholding. *Int. J. Appl. Eng. Res.* 10 (9), 21777–21783.
- Barclay, H., Trofymow, J., Leach, R., 2000. Assessing bias from boles in calculating leaf area index in immature Douglas-fir with the LI-COR canopy analyzer. *Agric. For. Meteorol.* 100 (2), 255–260.
- Bradshaw, C.J., Leroy, B.A., Bellard, C.A., Roiz, D.A., Albert, C.A., Fournier, A.A., Barbet-
- Massin, M.A., Salles, J.-M.A., Simard, F.A., Courchamp, F.A., 2016. Massive yet grossly underestimated global costs of invasive insects. *Nature Commun.* 7, 12986.
- dos Santos, J.C.C., Costa, R.N., Silva, D.M.R., de Souza, A.A., de Barros Prado Moura, F., da Silva Junior, J.M., Silva, J.V., 2016. Use of allometric models to estimate leaf area in *hymenaea courbaril*. *L. Theoret. Exp. Plant Physiol.* 28 (4), 357–369.
- Fan, X., Kawamura, K., Guo, W., Xuan, T., Lim, J., Yuba, N., Kurokawa, Y., Obitsu, T., Lv, R., Tsuniyama, Y., Yasuda, T., Wang, Z., 2018. A simple visible and near-infrared (V-NIR) camera system for monitoring the leaf area index and growth stage of Italian ryegrass. *Comput. Electron. Agric.* 144, 314–323.
- FAO, 2017. The future of food and agriculture - Trends and challenges. Food and Agriculture Organization of the United Nations, Rome.
- Gong, A., Wu, W., Qiu, Z., He, Y., 2013a. Leaf area measurement using android os mobile phone. *Trans. Chinese Soc. Agric. Mach.* 44 (9), 203–208.
- Gong, A., Wu, X., Qiu, Z., He, Y., 2013b. A handheld device for leaf area measurement. *Comput. Electron. Agric.* 98, 74–80.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 770–778.
- Houborg, R., McCabe, M., 2018. A hybrid training approach for leaf area index estimation via cubist and random forests machine-learning. *ISPRS J. Photogramm. Remote Sens.* 135, 173–188.
- Karatassiou, M., Ragkos, A., Markidis, P., Stavrou, T., 2015. A comparative study of methods for the estimation of the leaf area in forage species. In: *CEUR Workshop Proceedings*. vol. 1498. pp. 326–332.
- Kogan, M., Turnipseed, S., Shepard, B., B. De Oliveira, E., Borgo, A., 1977. Pilot insect pest management program for soybean in southern brazil. *J. Econ. Entomol.* 70 (5), 659–663.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. pp. 1097–1105.
- Kvet, J., Marshall, J., 1971. Assessment of leaf area and other assimilating plant surfaces. *Plant Photosynth. Prod. Manual Methods* 517–555.
- Machado, B.B., Orue, J.P., Arruda, M.S., Santos, C.V., Sarath, D.S., Goncalves, W.N., Silva, G.G., Pistori, H., Roel, A.R., Rodrigues-Jr, J.F., 2016. Bioleaf: a professional mobile application to measure foliar damage caused by insect herbivory. *Comput. Electron. Agric.* 129, 44–55.
- Nazaré-Jr, A.C., Menotti, D., Neves, J.M.R., Sediyama, T., 2010. Automatic detection of the damaged leaf area in digital images of soybean. In: *International Conference on Systems, Signals and Image Processing*. pp. 1–4.
- Novotny, P., Suk, T., 2013. Leaf recognition of woody species in central Europe. *Biosyst. Eng.* 115 (4), 444–452.
- Quab, M., Bottou, L., Laptev, I., Sivic, J., 2014. Learning and transferring mid-level image representations using convolutional neural networks. In: *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, pp. 1717–1724.
- Parmar, D., Ghodasara, Y., Patel, K., Patel, K., Kathiriya, D., Patel, H., 2016. Analysis of plant leaf area using java image processing techniques - scaling and non scaling. *Ecol. Environ. Conservat.* 22 (2), 763–766.
- Sarath, D.S., da Silva, G.G., Peruca, R.D., Machado, B.B., Roel, A.R., Pistori, H., 2015. Automatic quantification of soybean leaf area using color-based image segmentation. In: X SBIAGRO.
- Shen, J., 2003. On the foundations of vision modeling: I. webers law and weberized tv restoration. *Phys. D: Nonlinear Phenomena* 175 (3), 241–251.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- van der Maaten, L., 2014. Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.* 15, 3221–3245.
- Wilhelm, W.W., Ruwe, K., Schlemmer, M.R., 2000. Comparison of three leaf area index meters in a corn canopy. *Crop Sci.* 40 (4), 1179–1183.