



Mosquito Larvae Image Classification Based on DenseNet and Guided Grad-CAM

Zaira García¹ , Keiji Yanai² , Mariko Nakano¹  ,
Antonio Arista¹ , Laura Cleofas Sanchez¹ , and Hector Perez¹ 

¹ Instituto Politecnico Nacional, Mexico City, Mexico
zgn_1607@hotmail.com, mnakano@ipn.mx

² The University-of-Electrocommunications, Tokyo, Japan

Abstract. The surveillance of *Aedes aegypti* and *Aedes albopictus* mosquito to avoid the spreading of arboviruses that cause Dengue, Zika and Chikungunya becomes more important, because these diseases have greatest repercussions in public health in the significant extension of the world. Mosquito larvae identification methods require special equipment, skillful entomologists and tedious work with considerable consuming time. In comparison with the short mosquito lifecycle, which is less than 2 weeks, the time required for all surveillance process is too long. In this paper, we proposed a novel technological approach based on Deep Neural Networks (DNNs) and visualization techniques to classify mosquito larvae images using the comb-like figure appeared in the eighth segment of the larva's abdomen. We present the DNN and the visualization technique employed in this work, and the results achieved after training the DNN to classify an input image into two classes: *Aedes* and Non-*Aedes* mosquito. Based on the proposed scheme, we obtain the accuracy, sensitivity and specificity, and compare this performance with existing technological approaches to demonstrate that the automatic identification process offered by the proposed scheme provides a better identification performance.

Keywords: Mosquito larvae · Classification · Deep Neural Network · Mosquito control · Mosquito surveillance

1 Introduction

Worldwide viral diseases transmitted by mosquitoes have the most significant repercussions in public health. The principal viruses spread by the bite of the female *Aedes* mosquitoes (*Aedes aegypti* and *Aedes albopictus*) are arboviruses, which cause Dengue, Zika, Chikungunya and Yellow fever. When an *Aedes* mosquito bites a person, who has been infected previously with an arbovirus, the mosquito can become a carrier of the virus. If this mosquito bites another person, then the person also can be infected with arbovirus [1]. In the Dengue fever, the World Health Organization (WHO) estimates that the four serotypes of this virus are a menace for 40% of the world population that lives in subtropical and tropical areas [2]. In 2016, the WHO declared Zika virus a public health emergency and advised that pregnant women need to be protected from mosquitoes, because Zika fever can cause serious birth defects of her unborn baby [3].

Aedes aegypti and *Aedes albopictus* use natural and artificial water-holding containers to lay their eggs. After hatching, larvae grow and develop into pupae and subsequently into a terrestrial flying adult mosquito. Mosquito surveillance is a key component of any local integrated vector control program. Collecting data of the mosquito population in many geographic areas to identify the presence or absence of the *Aedes aegypti* or *Aedes albopictus*, support the creation of maps to determine the specific areas to fumigate. The specimen collection can be divided into three main surveys: ovitraps, immature stage surveys and adult mosquito trapping. Nevertheless, the transportation of specimens to laboratories and the analysis performed by an entomologist with special equipment are very time-consuming tasks in comparison with the mosquito's short lifecycle which takes about ten days from eggs to adult mosquitos under a favorable condition [4].

Nowadays, exist some technological approaches to support the identification of *Aedes* mosquitoes at the larvae stage focused on the eighth segment of its abdomen, denominated comb-like figure region. The first attempt presented in [5] used texture descriptors, such as Co-Occurrence matrix, Local Binary Pattern (LBP) and 2D Gabor filters, to extract characteristics and classify them using a Support Vector Machine (SVM). The better results obtained by 2D Gabor filters together with the SVM achieve an accuracy of the 79%. In [6], AlexNet DNN architecture with transfer learning technique is used, showing high sensitivity but low specificity. The work [7] proposed using VGG-16 in combination with VGG-19 to enhance the accuracy obtained during the classification. However, all before mentioned works use larvae images in grayscale and the comb-like figure area is cropped manually before the classification, so these methods need human intervention to carry out the recognition.

In this paper, we propose *Aedes* mosquito's identification algorithm using DenseNet 121 [8] and Guided Grad-CAM [9] to visualize the area of interest on which the trained DenseNet focuses. The area of interest determined by Dense-Net coincides with the region of the comb-like figure of larva's abdomen, which is efficiently used to distinguish larvae of *Aedes* mosquitos from other genera. The experimental results of the proposed algorithm show higher accuracy, sensitivity and specificity of classification, comparing with the before mentioned previous methods [5–7].

The rest of the paper is organized as follows: Sect. 2 describes the DenseNet 121 that is applied in the proposed method to achieve the classification of raw mosquito larvae images. Section 3 presents the obtained results applying Guided Grad-CAM to identify the area of interest in which the DNN focus to achieve the classification. Finally, in Sect. 4 the conclusions are presented.

2 Proposed Methods

The proposed method uses a database of raw mosquito larvae images, in other words we use larvae images without any manual interventions. As the comb-like figure area is relatively small in the input raw image, we decided to use Dense Convolutional Network (DenseNet) [8], in which a different connectivity of pattern between layers is

processed in feed-forward way. Whereas other DNNs have L connections if the number of layers is L , DenseNet has N direct connections, which is given by

$$N = \frac{L(L+1)}{2} \tag{1}$$

In Fig. 1 an example of DenseNet architecture is presented. Dense Blocks are made up of 1×1 convolution and 3×3 convolution layers.

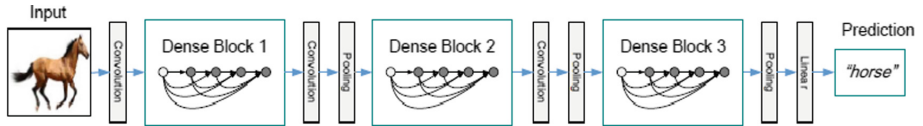


Fig. 1. DenseNet architecture example.

This configuration gives DenseNet many advantages that let the DNN improve the information flow between layers, strengthen feature propagation and feature reuse although the region of interest is relatively small. Also, this configuration reduces overfitting on tasks with smaller training set sizes [8]. Thus, the feature maps of preceding layers to be used as inputs into all subsequent layers are concatenated, supporting the final classifier decides based on all of them.

$$x_\ell = H_\ell([x_0, x_1, \dots, x_{\ell-1}]) \tag{2}$$

In (2), x_ℓ denotes the output of the ℓ^{th} layer, $H_\ell(\cdot)$ a non-linear transformation and $[x_0, x_1, \dots, x_{\ell-1}]$ is the feature maps concatenation of the layers $0, \dots, \ell-1$.

The Guided Grad-CAM is a visual explanation technique that allows us visualizing the regions learned by the DNNs. The Guided Grad-CAM does not require the architectural changes or re-training, also it can highlight fine-grained details in the image, and it provides class discriminative capability [9]. This technique can detect the importance of the neurons into the decisions of interest, using the gradient information into the last convolutional layer. The neuron importance of the feature map k for the target class c , α_k^c is given by

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \underbrace{\frac{\partial y^c}{\partial A_{ij}^k}}_{\text{gradients via backprop}} \tag{3}$$

global average pooling

First the gradient of the score for class c , y^c (before softmax) are computed with respect to A^k feature maps of a convolutional layer $\frac{\partial y^c}{\partial A^k}$. The gradients flow back and average-pooled to obtain (3). Then, ReLU is applied to the linear combination of forward activation maps (4).

$$I_{Grad-CAM}^c = ReLU \left(\overbrace{\sum_k \alpha_k^c A^k}^{Linear\ combination} \right) \tag{4}$$

The reason of the use of ReLU in (4) is because only the features that have positive influence on the area of interest are maintained. Figure 2 shows how Grad-CAM and Guided Backpropagation are obtained and the combination of them using point wise multiplication to generate Guided Grad-CAM.

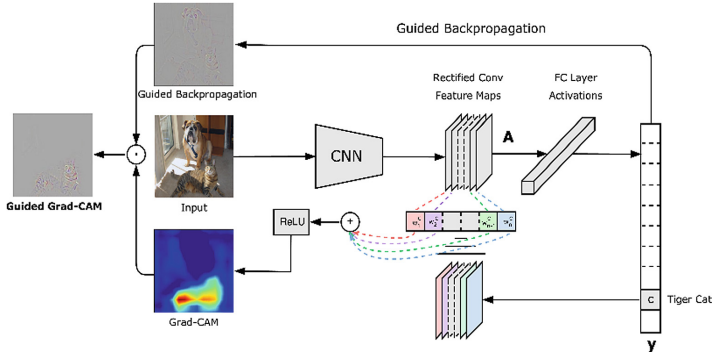


Fig. 2. Obtaining guided Grad-CAM.

3 Experimental Results

In this section, we show and explain the results achieved after training DenseNet 121. We modify the classification layer of the network to be able to discriminate into two classes: Aedes and Non-Aedes mosquito. Figure 3 illustrates an example of the images used during training, validation and test phases. The principal feature of larvae of Aedes

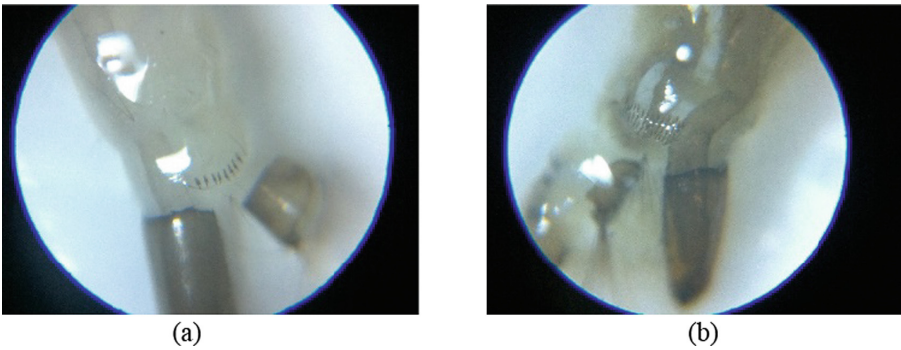


Fig. 3. 8th segment of larva’s abdomen. (a) Aedes sample. (b) Non-Aedes sample.

mosquito is single line well-defined comb-like figure as shown by Fig. 3(a), while other genera of mosquitoes shows disordered comb-like figure as shown by Fig. 3(b).

Our mosquito database is made up of 760 images divided in three sets. 600 images as training set, 60 as validation set and 100 for test. We use data augmentation technique to generate more data with random transformation such as rotation, width and height shift, flip, shear and zoom. It is important to mention that this procedure is only applicable to the training data and the final size of this data is 4,200 images. We use different values of hyperparameters to get different models to obtain an optimum model. Finally, we employed the Grid Search parameter tuning technique to determine one from the accuracy point of view, which is given by Table 1.

Table 1. Hyperparameters used to train DenseNet 121.

Hyperparameter	Values
Epoch	150
Batch size	10
Learning rate	0.01
Momentum	0.2
Decay	$10e^{-5}$

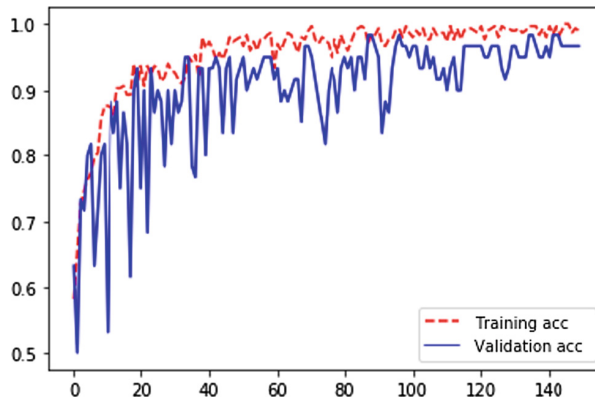


Fig. 4. Training and validation accuracy.

Although we determine the best hyperparameters, the several models obtained before were also used as inputs for Guided Grad-CAM to determine the area of the image that the network is focused on, because we are interested in making the DenseNet 121 identify the area of the comb-like figure. The model obtained with the hyperparameters in Table 1 achieves 97% accuracy in the classification of the samples as it shows in Fig. 4.

Table 2 presents the confusion matrix obtained using the test set composed by 50 larvae of *Aedes* mosquitoes and 50 of other genera, after the DNN with optimum hyper-parameter is trained.

Table 2. Confusion matrix results with 97% of accuracy.

	Aedes	Non-Aedes
Aedes	48	2
Non-Aedes	1	49

Table 3 shows a comparison of the accuracy, sensitivity and specificity values between the before mentioned methods and our method.

Table 3. Comparison of our method with previous methods.

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)
Co-Occurrence Matrix and SVM [5]	67.00	64.51	58.06
LBP and SVM [5]	72.00	67.74	77.41
2D Gabor filters and SVM [5]	79.00	77.41	80.64
AlexNet DNN [6]	80.92	100	70
VGG-16 with VGG-19 [7]	91.28	94.18	88.37
<i>Proposed scheme</i>			
DenseNet 121 and Guided Grad-CAM	97.00	97.00	96.00

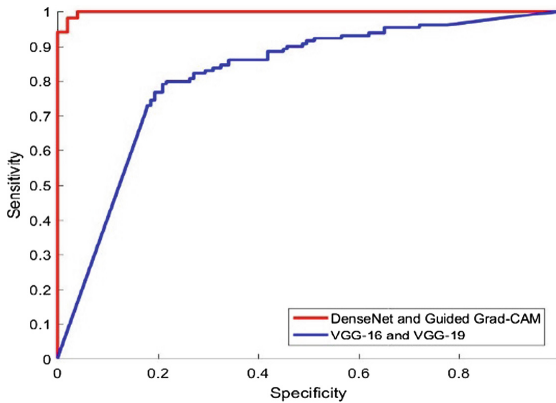


Fig. 5. Comparison of ROC curves.

Figure 5 shows the receiver operation characteristics (ROC) curve of the proposed method using DenseNet and Guided Grad-CAM and the method which employs VGG-16 and VGG-19 [7]. The ROC curves demonstrate that our proposed method draws a curve nearer to the optimum value (that is one) than the VGG-16 and VGG-19 method [7] which is the recent one with high values of accuracy, sensitivity and specificity in the state of arts.

As a verification phase, the Guided Grad-CAM is applied to the obtained model to visualize the region of interest. In Fig. 6 the area of interest determined by the network

for each class: *Aedes* and Non-*Aedes* mosquitoes are shown for one training data and one test data. From Fig. 6, we can observe clearly that the interest region of the *Aedes* larva indicated by the DenseNet is the comb-like figure, and although this region for Non-*Aedes* sample is not as clear as the *Aedes* one in both training and testing stages, these regions are highlighted as we expected. The visualization of the interest region where the trained DNN focused on allows us an efficient verification of the training process of the DNN. Additionally, the visualization of interest region helps us to select not only an appropriate architecture of the DNN but also hyper-parameters used in it.

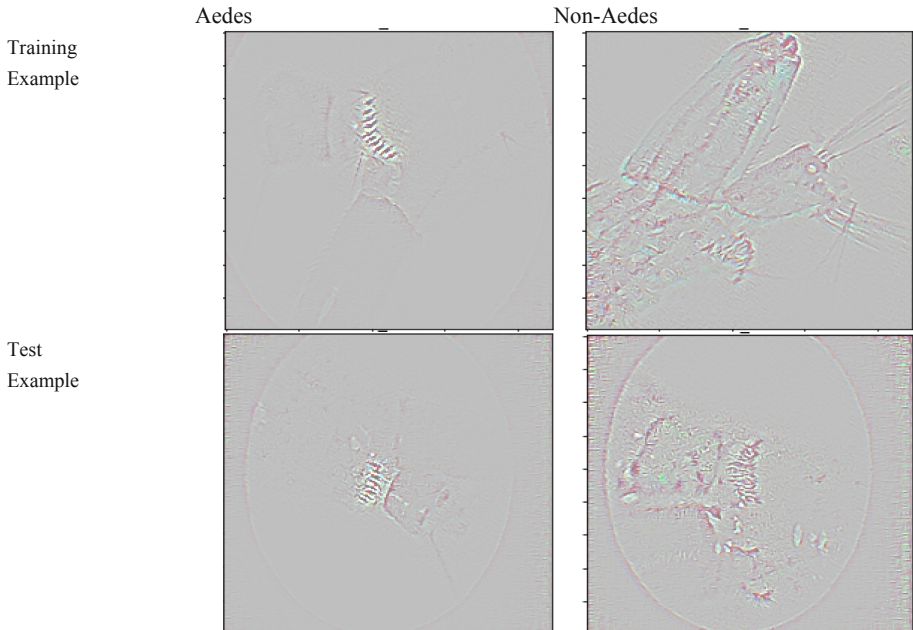


Fig. 6. Guided Grad-CAM example results for training and test data.

4 Conclusions

In this work we proposed a novel method to classify mosquito genera from the mosquito larvae images using the DenseNet 121 as the DNN architecture and the Guided Grad-CAM as a visualization technique of the interest region on where the trained DenseNet focused. The experimental results show that the proposed method provides the accuracy of 97%, which is superior to the performance of the previously reported methods [5–7]. The proposed scheme does not need any special pre-processing for the images such as manual segmentation of the area where the comb-like figure is presented or convert a color image to a grayscale one. This characteristic allows automatic recognition of the larvae without any human intervention, once larva's image is captured and introduced into the system. The future work includes building a bigger dataset in order to reinforce the training and obtain a higher classification accuracy.

References

1. Salud Documentos Técnicos. <https://drive.google.com/file/d/0B0K9c-Z-JA2nSWRrYWRFVzhXa0E/view>. Accessed 31 Jan 2019
2. KidsHealth. <https://kidshealth.org/en/parents/dengue.html?WT.ac=pairedLink>. Accessed 05 Feb 2019
3. CDC. <https://www.cdc.gov/zika/pdfs/Draft-Environmental-Assessment-Mosquito-Control-for-publication.pdf>. Accessed 05 Feb 2019
4. CDC. <https://www.cdc.gov/chikungunya/pdfs/Surveillance-and-Control-of-Aedes-aegypti-and-Aedes-albopictus-US.pdf>. Accessed 07 Feb 2019
5. Garcia-Nonoal, Z., Sanchez-Ortiz, A., Arista-Jalife, A., Nakano, M.: Comparison of image descriptors to classify mosquito larvae. In: Proceeding of CAIP, pp. 271–278 (2017). (in Spanish)
6. Sanchez-Ortiz, A., et al.: Mosquito larva classification method based on convolutional neural networks. In: International Conference of Electronics, Communications and Computers CONIELECOMP, pp. 155–160 (2017)
7. Arista-Jalife, A., Sanchez, A., Nakano, M., Tunnermann, H., Perez-Meana, H., Shouno, H.: Deep Learning employed in the recognition of the vector that spreads Dengue, Chikungunya and Zika viruses. In: International Conference on Intelligent Software Methodologies SoMeT, vol. 17, no. 1 (2018)
8. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: CVPR, vol. 1, p. 3 (2017)
9. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-CAM: visual explanations from deep networks via gradient-based localization, vol. 7, no. 8 (2016). <https://arxiv.org/abs/1610.02391v3>