MosquitoNet: A Deep Learning-Based Classifier for Mosquito Identification

Jay Gupta & Aliza Kajani
Department of Biomedical Engineering
Duke University
Durham, NC 27705
jg329@duke.edu
ak398@duke.edu

April 29, 2020

Abstract

High-throughput mosquito surveillance technologies are vital for epidemiological studies concerned with investigating malaria incidence in remote areas. Mosquitoes are known to transmit malaria parasites during only certain stages of their gonotrophic (reproductive) cycle. Thus, the present work explores the use of deep convolutional neural network models to determine the gonotrophic state of a mosquito through images of its head, legs, wings, and palps. Three major approaches were considered in this study: (a) training AlexNet with the Duke Global Health Institute (DGHI) mosquito dataset, (b) pre-training the InceptionV3 model with ImageNet and fine-tuning with the DGHI dataset, and (c) pre-training AlexNet with the IP102 dataset with fine-tuning with the DGHI dataset. In order to reduce overfitting and class imbalances, data augmentation and random example oversampling strategies were employed. In all cases, classifier performance was weak and test accuracies above ~45% were not observed. Taken together, this project lays the groundwork and proposes strategies for automated identification of gonotrophic state by inspection of mosquito images.

1 Introduction

While the global incidence of malaria has been steadily decreasing over the past decade, the World Health Organization estimates that over 200 million new cases still occur each year [1]. Malaria infection is caused by the *Plasmodium* parasite, which is carried by certain species of mosquitoes. Each of these species may be endemic to a unique geographical region, making high-throughput mosquito surveillance methods valuable for public health researchers.

Parasite-carrying mosquitoes belong to the *Anopheles* genus and are often one of three species: *gambiae*, *funestus*, or *demeilloni*. Furthermore, mosquitoes can also be categorized into the following states: unfed, blood-fed, gravid, or half-gravid. These states correspond to stages in the mosquito gonotrophic (reproductive) cycle and can be determined by visual inspection of a mosquito's abdomen [2]. Furthermore, a mosquito's gonotrophic stage can be indicative of its ability to transmit Plasmodium parasites and cause malarial infections [3]. Thus, we describe herein the development of a deep learning-based classifier for identification of a mosquito's gonotrophic state through images collected by researchers in remote locations.

2 Related work

There are several studies which propose using convolutional neural networks (CNNs) similar to AlexNet in order to classify mosquitoes; however, the majority of them focus on classifying mosquitoes by their age and/or species, as these are key factors in predicting probability of disease transmission. This work, instead, will focus on attempting to classify mosquitoes by their gonotrophic state.

Motta et al. makes use of three CNNs, including AlexNet, for the automatic morphological classification of adult mosquitoes of the genus *Aedes* [4]. Certain mosquitoes of this genus are well-known carriers of arboviruses such as Zika, dengue, and chikungunya. This work demonstrated the capacity of CNNs, such as AlexNet, to classify mosquitoes by species and gender. Similarly, Park et al. makes use of three different CNNs, including ResNet-50 and VGG-16 to classify mosquitoes by species [5]. This work focused on classifying mosquitoes with high inter-species similarity and intra-species variations of the genera *Anopheles*, *Aedes*, and *Culex*. Park et al. was able to achieve 97% accuracy with the use of data augmentation and pre-training, both of which will be explored in the study presented here.

3 Methods

3.1 Dataset

The deep learning model was developed, trained, and evaluated using a dataset recently collected and annotated by Duke Global Health Institute (DGHI) researchers, led by Dr. Wendy Prudhomme-O'Meara. The dataset consists of images of 1,327 unique mosquitoes. Each mosquito was imaged three times (*i.e.* from three different directions) giving a total of 3,981 images in the dataset. Mosquitoes in the example images belong to the *Anopheles* genus and are one of three species: *gambiae, funestus,* or *demeilloni*. Furthermore, each example mosquito is in one of the following gonotrophic states: unfed, blood-fed, gravid, or half-gravid.

3.2 Convolutional neural network

Deep convolutional neural network (CNN) architecture was based on AlexNet, which was popularized in 2012 when it achieved a Top-5 error ranking in the ImageNet Large Scale Visual Recognition Challenge. The network employs "five convolutional layers, some of which are followed by max-pooling layers, and two globally connected layers with a final 1000-way softmax" classifier [6]. AlexNet has been applied in previous work on insect and pest detection, and was therefore hypothesized to be a good choice for the mosquito classification task described here [7].

3.3 Physical layers

Pixel size and camera angle were considered physical dimensions of the image acquisition system. The DGHI mosquito dataset consisted of images that were captured at high resolution. Images were downsampled to 224 x 224 pixels prior to training in order to reduce computational complexity and training time. Further, each example mosquito in the dataset was imaged from three directions (*i.e.* head-legs, wings, and palps). The network was trained multiple times, using data from a single camera angle during each of these training cycles, in order to determine whether there was an optimal camera angle for the classification task. Model accuracy after training with each specific camera angle were compared in order to determine whether one view outperformed the others.

3.4 Pre-training

Due to the limited size of the DGHI mosquito dataset, it was hypothesized that pre-training on a very large, general object detection dataset may further improve classification accuracy on the DGHI mosquito dataset. This was tested by using the DGHI dataset to fine-tune an instance of the InceptionV3 model that was pre-trained on the ImageNet dataset and was available through the Keras Applications API.

Additionally, it was hypothesized that pre-training on a dataset of example insect images would further improve accuracies by allowing the deep CNN to identify visual features that are uniquely characteristic of mosquitoes and other pests. Thus, the AlexNet model was pre-trained on the IP102 dataset, a large benchmark dataset recently developed for insect pest recognition [8]. IP102 consists of 75,222 images of insects belonging to 102 classes. These insects may share fundamental morphological features with mosquitoes, thus the effect of pre-training with IP102 on model convergence and accuracy with the DGHI mosquito dataset was tested.

4 Results

4.1 Training each camera angle separately, with and without data augmentation

An implementation of AlexNet was written using Tensorflow and the high-level Keras API. The model was trained four times. Specifically, it was trained on data from each of the three camera angles separately, as well as the full dataset consisting of images from all three camera angles together. Test accuracies, shown in Table 1, for training separately on each camera angle were very similar (around 40% accuracy in each case). However, mild overfitting was observed as the training accuracies were often higher for the models. In order to reduce the likelihood of overfitting, the models were retrained after data augmentation strategies were applied. Images in the dataset were rotated, flipped, and shifted horizontally and vertically prior to model training. Test accuracies were not significantly improved with data augmentation, but it was noticed that there was much less overfitting in cases where augmentation was performed. Additionally, about a 5% increase in accuracy was noticed when the model was trained on the entire dataset (containing images from all three camera angles).

Table 1: Test accuracies after training the AlexNet model on images captured from unique camera angles, before and after data augmentation strategies were applied.

Camera Angle	Data Augmentation	Final Test Accuracy
Head-legs	No	40.23%
	Yes	42.11%
Wings	No	34.59%
	Yes	40.23%
Palps	No	40.98%
	Yes	39.85%
All	No	44.04%
	Yes	45.67%

4.2 InceptionV3 model

To determine whether pre-training would improve classification accuracy, the DGHI dataset was used to fine-tune an instance of an InceptionV3 model that was pre-trained on the ImageNet dataset and was available through the Keras Applications API. First, all InceptionV3 model weights were frozen and a trainable, softmax classifier was added at the top of the network. This classifier was trained for five epochs, after which point the weights of the last 2 “inception modules” were unfrozen. The model was trained again in order to fine-tune the parameters of these final inception modules and the classifier stage as well. This process was repeated four times, first using data from each of the three camera angles separately, then the full dataset. Accuracy was monitored at the end of each epoch and learning curves are presented in Figure 1. The results suggest that, after fine-tuning weights of the last 2 inception models, the network was largely overfitting and was unable to produce accurate results (as suggested by the large gaps between the training and testing accuracy curves).

Figure 1: Learning curves for fine-tuning the InceptionV3 model.

To identify the root cause of this overfitting, confusion matrices for each training case were generated (Figure 2). The confusion matrices suggest that, in all cases, the InceptionV3 model was almost exclusively predicting that test examples belonged to either the blood fed or half gravid classes. This result is intuitively warranted by the fact that the half gravid and blood fed classes are disproportionately represented in the DGHI mosquito dataset (Figure 3). Over 60% of the dataset consists of examples from the blood fed or half gravid classes, resulting in the imbalanced predictions.

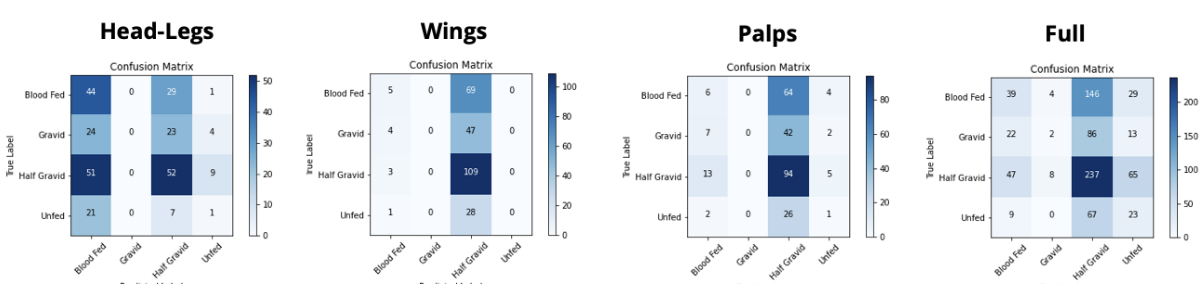

Figure 2: Confusion matrices for the fine-tuned InceptionV3 model.

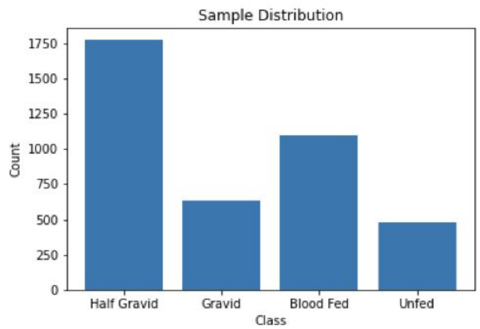

Figure 3: Distribution of samples across classes in the DGHI mosquito dataset.

Two strategies were employed in order to reduce overfitting during the fine-tuning process: (a) data augmentation, and (b) random oversampling of the minority class. First, the softmax classifier and top 2 inception modules were fine-tuned using the full DGHI dataset (*i.e.* images from all 3 camera angles) after augmentation through rotations, flips, and horizontal and vertical shifts. As shown in Figure 4, overfitting was no longer present (indicated by similar train and test accuracies after the final epoch). However, the accuracy did not improve above ~45%, like AlexNet. Furthermore, in every case except for 1, the model was predicting that examples belonged to the half gravid class explaining the low accuracy. The model was then trained after minority classes were randomly oversampled in the training dataset to minimize the effect of class imbalance. Results indicated that this manipulation did not eliminate overfitting, however, it did ensure that model predictions were distributed somewhat more evenly across all four classes (Figure 4).

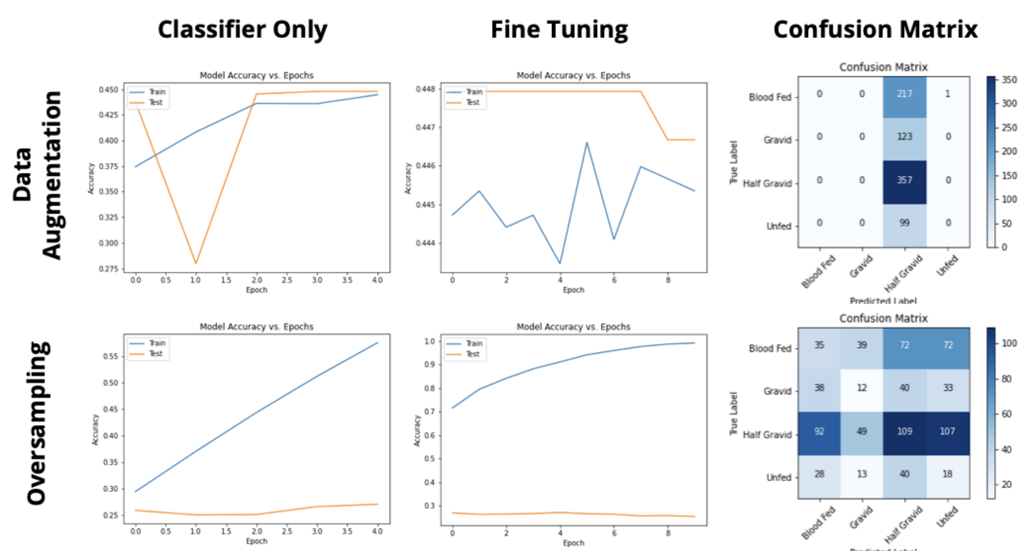

Figure 4: Learning curves and confusion matrices for fine-tuning the InceptionV3 model after data augmentation and oversampling strategies were applied.

4.3 Pre-training on IP102 dataset

Since the InceptionV3 model was largely overfitting the training data, we next applied pre-training strategies to AlexNet, which is a less computationally complex model and should thus be less inherently prone to overfitting. In an attempt to improve AlexNet performance (compared to the previous experiments), the model was pre-trained using the IP102 dataset. The IP102 dataset consists of images of insects, which may share fundamental morphological features with mosquitoes, making this dataset a potentially strategic choice for pre-training. Significantly, only the four largest classes from the IP102 dataset were used for pre-training, since the remaining 98 classes had less than 500 examples each. Pre-training learning curves for both loss and accuracy, shown in Figure 5, indicate that AlexNet was able to classify the four largest IP102 classes with ~67% accuracy. However, after fine-tuning model parameters by re-training weights with the DGHI mosquito dataset, AlexNet accuracy returned to ~44%, which was comparable to the results of previous experiments. Taken together, these results suggest that IP102 pre-training did not improve AlexNet model performance for this classification task.

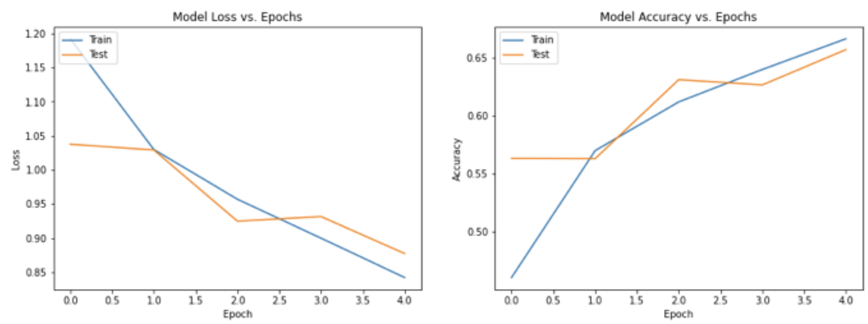

Figure 5: Loss and accuracy learning curves for pre-training AlexNet with the IP102 dataset.

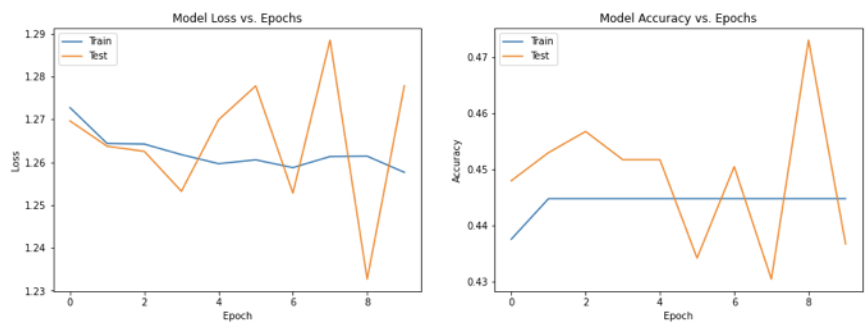

Figure 6: Loss and accuracy learning curves for fine-tuning AlexNet with the DGHI mosquito dataset.

5 Discussion

Three major approaches for classifying images of mosquitoes have been considered in the present work: training AlexNet on the DGHI mosquito dataset with and without data augmentation, pre-training the InceptionV3 model on ImageNet with fine tuning on the DGHI dataset, and pre-training AlexNet on the IP102 dataset with fine tuning on the DGHI dataset. Additionally, the physical layer of camera angle was considered; however, no major differences were identified between the three camera angles (head-legs, wings, and palps) as performance was weak for all.

Over the course of this study, two major challenges were encountered: overfitting and class imbalance. We attempted to mitigate the effects of these challenges by employing data augmentation strategies as well as random oversampling of the minority classes. The data augmentation technique was effective in decreasing the level of overfitting; however, it did not result in a significant increase in model accuracy. Similarly, although random oversampling successfully mitigated the class imbalance effects we observed, the model was still unable to accurately classify the mosquitoes by gonotrophic state.

Although AlexNet and other CNNs have proven useful for classifying mosquitoes by their age, species, and/or gender, we were unable to demonstrate their ability to classify mosquitoes by gonotrophic state. This could be a result of a number of factors, including image resolution and camera angle. Future directions may include optimizing alternative physical layers, such as a downsampling factor for the original, high resolution images. Alternatively, it may be advantageous to use a dataset consisting of images which include the mosquito's full abdomen, since we know it is possible to determine a mosquito's gonotrophic state from visual inspection of the abdomen alone. Future works may also consider using a less complex CNN architecture in order to reduce overfitting, employing more nuanced statistical methods than random oversampling to handle class imbalances, or developing ensemble models that are capable of making classification decisions by simultaneously analyzing images acquired from multiple camera angles.

References

[1] World Health Organization, "Global Health Observatory (GHO) data," WHO, 2019.

[2] M. Service, Medical Entomology for Students: Fifth Edition Mike. 2012.

[3] K. Okuneye, S. E. Eikenberry, and A. B. Gumel, "Weather-driven malaria transmission model with gonotrophic and sporogonic cycles," *J. Biol. Dyn.*, vol. 13, no. sup1, pp. 288–324, 2019, doi: 10.1080/17513758.2019.1570363.

[4] Motta, Daniel et al. "Application of convolutional neural networks for classification of adult mosquitoes in the field." *PloS one* vol. 14,1 e0210829. 14 Jan. 2019, doi:10.1371/journal.pone.0210829

[5] Park, J., Kim, D.I., Choi, B. et al. "Classification and Morphological Analysis of Vector Mosquitoes using Deep Convolutional Neural Networks." *Sci Rep* vol. 10, 1012 (2020). https://doi.org/10.1038/s41598-020-57875-1

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," 2012.

[7] L. Nanni, G. Maguolo, and F. Pancino, "Research on insect pest image detection and recognition based on bio-inspired methods," Oct. 2019.

[8] X. Wu, C. Zhan, Y.-K. Lai, M.-M. Cheng, and J. Yang, "IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition," 2019.