Leveraging Smart-Phone Cameras and Image Processing Techniques to Classify Mosquito Species

Mona Minakshi
University of South Florida
Tampa, Florida, USA.
mona6@mail.usf.edu

Pratool Bharti
University of South Florida
Tampa, Florida, USA.
pratool@mail.usf.edu

Sriram Chellappan
University of South Florida
Tampa, Florida, USA.
sriramc@usf.edu

ABSTRACT

Mosquito borne diseases continue to pose grave dangers to global health. An important step in combating their spread in any area is to identify the type of species prevalent there. To do so, trained personnel lay local mosquito traps, and after collecting trapped specimens, they visually inspect each specimen to identify the species type and log their counts. This process takes hours and is cognitively very demanding. In this paper, we design a smart-phone based system that allows anyone to take images of a still mosquito that is either alive or dead (but still retaining its physical form) and automatically classifies the species type. Our system integrates image processing, feature selection, unsupervised clustering, and an SVM based machine learning algorithm for classification. Results with a total of 101 mosquito specimens spread across nine different vector carrying species (that were captured from a real outdoor trap in Tampa, Florida) demonstrate high accuracy in species identification. When implemented as a smart-phone application, the latency and energy consumption were minimal. With our system, the currently manual process of species identification and recording can be sped up, while also minimizing the ensuing cognitive workload of personnel. Secondly, ordinary citizens can use our system in their own homes for self-awareness and information sharing with public health agencies.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning**; *Computer vision*; • **Human-centered computing** → *Smartphones*; • **Applied computing** → *Health care information systems*;

KEYWORDS

Machine Learning, Image Processing, Public Health, Mosquitoes, Smart-phones, Computer Human Interaction

1 INTRODUCTION

Mosquito borne diseases (e.g., Malaria, Dengue, West Nile Fever, and most recently Zika Fever) are amongst the biggest health care concerns across the globe today. To mitigate the spread of mosquito-borne diseases, it is vital to combat the spread of mosquitoes. Of critical importance in this mission is the identification of species prevalent in an area of interest. This is important, because there are close to 3, 500 different species of mosquitoes present in the world today [1], and with increasing globalization and warming, they are spreading to newer locations, with most of them acting as vectors for several diseases. In any given location, multiple species are usually found at the same time. However, the process of species identification is not at all easy. We present details below.

1.1 Current Trends in Species Identification

As of today, to derive populations of mosquitoes in any area, trained professionals lay traps, and pick them soon after to sort trapped specimens. Sometimes, hundreds of mosquitoes can be trapped in a single day. Then, to identify each specimen trapped, it is placed under a microscope, and visually identified, which takes hours each day for all specimens. Depending on location and time of year, this process can repeat multiple times in a single week, and is cognitively demanding. We also point out that such kinds of mosquito control facilities are expensive to manage, and they are very few even in advanced countries. In low economy countries, where mosquitoes pose a greater danger, such facilities are even more scarce. With rising temperatures and population migrations, mosquitoes are believed to be invading newer areas across the world, and detecting them early is a huge challenge today.

1.2 The Lack of a Citizen-Science Approach

Experts at mosquito control facilities acknowledge that, depending on location and time of the year, they can receive hundreds of calls each day from concerned citizens about mosquitoes in their neighborhoods. Due to limited resources, knowledge of mosquito species types can play a vital role in prioritizing schedules for trap placement and spraying repellents during peak times, since different mosquito species are vectors for different diseases. Sadly, despite citizens willing to assist in this process, there is no way to enable that now. One practice recommended by experts is to ask citizens to collect a few mosquitoes (after spraying on them), and store them in a transparent bag for the experts to identify them later. But this process is cumbersome, and the need for technology based solutions to empower citizens in this effort became clear.

1.3 Our Contributions

Based on the facts mentioned above, and coupled with the increasing global spread of mosquito-borne diseases, public health experts

MobiQuitous 2018, November 5-7, 2018, New York City, United States

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00

we spoke to during this study were highly receptive to any technology based solution for mosquito species identification and recording that is accurate, comfortable and fast, so that a) human resources in public health can utilized more effectively, and b) citizens can be better informed and hence better served. Towards this extent, in this paper, we design a smart-phone based system that enables anyone to take images of a still mosquito that is alive or dead (after possibly spraying or trapping), but still retaining its physical form, and then processes the captured images for species identification. Our specific contributions are listed below.

a). A Database of 303 **mosquito images spread across nine different species:** In Fall 2016 and Spring 2017, we visited the mosquito control board in Hillsborough County, Florida to collect specimens of female vector carrying mosquitoes that were captured in outdoor traps. When a trap is set, specimens of dead mosquitoes are collected the next day to prevent their decaying (that will complicate visual identification). Then, the personnel there helped us visually identify the species type of 101 mosquito specimens that were evenly distributed across nine different species, details of which are presented in Table 1. We immediately took one image of each specimen using a Samsung Galaxy S5 Smart-phone in three visually distinct backgrounds (with a different camera orientation for each background). As such, we obtained a total of 303 mosquito image samples for model development. To the best of our knowledge, this is the first instance of a dataset containing tagged images of mosquito species taken via a smart-phone camera (which we will publicly release soon, and is also available upon request). In Figure 1, we present one representative smart-phone image for each of the nine species we attempt to classify in this paper.

b). Feature Extraction, Dimensionality Reduction, Clustering and Classification: Our proposed technique for species identification incurs multiple steps. First, we reduce the image size from around 2988×5322 pixels per image to 256×256 pixels for faster processing, followed by employing median filters to remove noise. Second, we employ contour segmentation and Gaussian mixtures to carefully extract only the mosquito portion from each image and segmenting out the background. Third, we then convert each pixel into *Lab* color space, and extract 39 features based on Local Binary Patterns [12], and Haralick Texture Features [15], both of which preserve textural patterns (which are distinct across mosquito species) better. Fourth, we apply Linear Discriminant Analysis (LDA) to further reduce the number of features to just 8, that are subsequently used for classification. Towards the end, we design a 2-step process that involves unsupervised clustering, and a support vector machine algorithm for species identification [1].

c). Results: Our performance evaluations yielded an overall accuracy of around 80% in species identification with good precision and recall. The latency consumed during classification when the entire system is implemented as a smart-phone app on a Samsung Galaxy S5 phone was less than 2 seconds. Towards the end of the paper, we also present important practical impacts of our system.

[1]Perspectives on difficulties in applying Deep and Transfer Learning Techniques for our classification problem are presented in Section 5 towards the end of the paper.

2 RELATED WORK

We briefly present important related work on technology based solutions (image and others) to identify mosquitoes, while also clarifying the significance of our system proposed in this paper.

a). Image based Techniques using Digital Cameras: In [10], a solution is proposed to detect *Aedes aegypti* species using images taken from a $500x$ optical zoom camera, and a support vector machine classification algorithm. Using a sample of 40 images, seven textural features, and a support vector machine classification algorithm, an accuracy of 92.5% was demonstrated in classifying *Aedes aegypti* species from others. This solution though is expensive, and addresses a binary classification problem only.

Work in [14] and [13] discuss machine learning techniques to classify mosquitoes from insects like flies and bees using images taken from digital cameras. The problem addressed in these papers is too generic though. In a recent paper [32], the authors address a problem similar to ours, but sufficiently different. Specifically, 12 adult mosquito specimens from 3 genera (*Aedes*, *Anopheles* and *Culex*) were collected, and the right wing of each specimen was photographed using a sophisticated digital camera coupled with a microscope. Then, using coordinates at intersections of wing veins as a feature, followed by a Neighbor Joining Tree classification method, the accuracy in genus identification (among three) was 90%. This technique again is expensive and requires expertise.

b). Using Techniques other than Imaging: In [8], the authors attempt to use optical (rather than acoustic) sensors to record the "sound" of insect flight from a small distance, and then design a Bayesian classifier to identify four species of mosquitoes (*Aedes aegypti*, *Culex quinquefasciatus*, *Culex stigmatosoma*, and *Culex tarsalis*), and achieve an accuracy of 96%. Similarly, the work in [23] also leverages smart-phone microphones to capture and process acoustics of mosquito flight, along with location and time of observation. The claim is that these features are unique to classify mosquito species. More innovative techniques like hydrogel-based low-cost microfluidic chips, baited with odorants to capture saliva droplets of mosquitoes are being designed by Dr. Manu Prakash at Stanford University in order to serve as a test for vector species and pathogens. All of these techniques require "live" and "mobile" mosquitoes, with sensing devices placed close to them. They are not suited for ubiquitous and in-home use by common citizens.

c). Other Related Work: A survey on imaging techniques to classify insects is presented in [19]. However, mosquitoes are not classified there. In [26], the authors ask citizens to use smart-phones for imaging and reporting about mosquitoes they encounter, but species classification is not discussed. In [24], Munoz et.al. propose a deep learning framework to classify larvae of mosquitoes from larvae of other insects, with smart-phone images. In [5], intensity of red blood cells computed from thin blood smear images were used to identify the presence of malarial (plasmodium) parasites in blood samples. Microsoft's "Project Premonition" is an ambitious effort to use drones and DNA sequencing techniques to identify mosquito species in hot-spots [4]. Like we do in this paper, these recent works highlight important, but orthogonal tech-based solutions to combat mosquito-borne diseases, but ubiquitous and easy to use solutions for identifying mosquitoes species are not yet there.

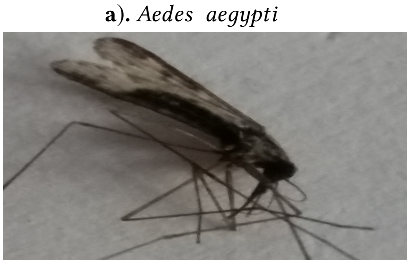
Table 1: Relevant Details on our Dataset of Mosquito Species

Species	No. of Specimens	No. of Image Samples (3 per Specimen)	Disease Spread	Geographical Location
Aedes aegypti	11	33	Zika fever, Dengue, Chikungunya	South America , North America, Asia and Africa
Aedes infirmatus	10	30	Eastern equine encephalitis (EEE)	South America and North America
Aedes taeniorhynchus	8	24	West Nile Virus	South America and North America
Anopheles crucians	15	45	Malaria	South America , North America and Africa
Coquillettidia perturbans	14	42	West Nile Virus	South America and North America
Culex nigripalpus	10	30	West Nile virus	South America , North America and Africa
Mansonia titillans	11	33	Venezuelan equine encephalitis (VEE)	South America , North America and Africa
Psorophora columbiae	11	33	Venezuelan equine encephalitis (VEE)	South America , North America and Africa
Psorophora ferox	11	33	West Nile Virus	South America , North America and Africa

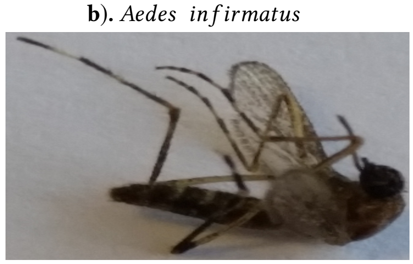
a). *Aedes aegypti*

b). *Aedes infirmatus*

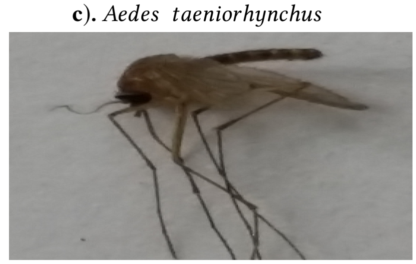
c). *Aedes taeniorhynchus*



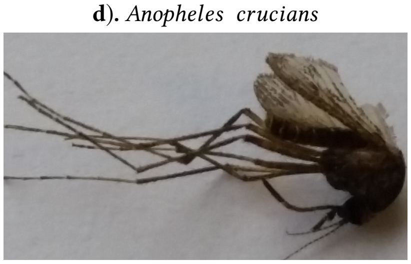
d). *Anopheles crucians*



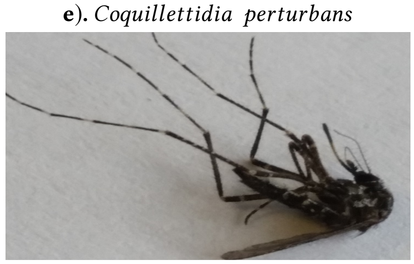
e). *Coquillettidia perturbans*



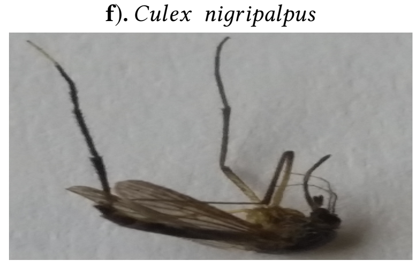
f). *Culex nigripalpus*



g). *Mansonia titillans*



h). *Psorophora columbiae*



i). *Psorophora ferox*

Figure 1: One Representative Sample in our Dataset for Each Species Classified. This Figure is best viewed in Color.

d). Our Prior Work on Identifying Mosquito Species from Smart-phone Images: In our prior work in [22], we leverage smart-phone images to identify a total of 7 mosquito species. However the technique in [22] had limitations stemming from poorer accuracy, inability to handle images taken in different backgrounds, and is also computationally very expensive to process on a smart-phone (due to the processing of many features). In our improved system proposed here, the number of species identified is 9; our improved system includes background segmentation that compensates for images taken in differing backgrounds; and is computationally much more efficient to enable processing on a smart-phone.

e). Summarizing Related Work: To summarize, tech-based solutions to combat the spread of mosquito-borne diseases is an important need of the hour. However, there is no system yet that enables common citizens participate in mosquito identification. This paper fills the gap by designing a smart-phone based system that enables anyone to take images of a still mosquito that is alive or dead (after possibly spraying or trapping), but still retaining its physical form, and then processes the images for species identification. Our system is cheap, ubiquitous, and easily expandable to include more mosquito species beyond the current *nine* classified in this paper.

On a related note, we are aware of smart-phone apps to identify types of plants, flowers and certain types of insects as well. However, algorithms used in these apps are not publicly available. We believe though that the problem of identifying mosquito species from images is much harder than the ones above, since there are no obvious (and un-aided) visually discernible markers across species types to the naked eye. We witnessed public health workers with decades of experience still needing a microscope and careful analysis to identify the species type of a mosquito specimen, hence demonstrating the complexity of our problem attempted here.

3 DATA COLLECTION

In the Hillsborough County where we collected our specimens from, there is a dedicated mosquito control board for trapping, collecting, and manually identifying mosquito species. In this county alone, up to 40 species of mosquitoes are prevalent, not all of them at the same time though. Every week, personnel lay traps for mosquitoes in selected areas, and dead specimens are collected the next day, brought to the lab, and each specimen is visually identified using a microscope, and population results are logged. The early collection of specimens is important because, once dead, they decay fast, making visual identification harder if delayed.

During a couple of months between Fall 2016 and Spring 2017, we participated in multiple such efforts and were given a total of 101 female mosquito specimens from a total of nine different mosquito species, which were the ones most prevalent that time of the year in that county. Each specimen was carefully identified and labeled by experts in the board for us to get the ground truth data. Table 1 presents details on our data set. A Samsung Galaxy S5 phone was then used to capture an image of each specimen under the same indoor light conditions, with the camera located one feet above each specimen without flash. Three images of each specimen were captured in a different phone orientation, on top of one of three backgrounds: a relatively white background, a yellow background and a pink background. In total, 303 images were captured. Figures 1 (a) to (i) present one representative smart-phone image of each of the nine species which we attempt to classify in this paper, when captured in a relatively white background. Features of the smart-phone camera used, are presented in Table 2.

a). Utility of Images Captured: Upon seeing the images generated, our colleagues at the Mosquito Control Board indicated that they were sufficiently rich for a trained expert to visually identify the species from the images. We were thus motivated to achieve the same via learning techniques, that could be implemented on a smart-phone so that common citizens can do the same.

b). A Note on Gender of Specimens in our Dataset: Note here that all the 101 mosquito specimens we collected were *female*. Among mosquitoes, only females engage in a blood meal (to provide nutrients for egg production), while males only feed on plant nectar. As such, only female species can carry disease vectors. In the traps that were laid for our experiments, CO_2 was used as a bait, which is typical. The presence of CO_2 tricks a female mosquito into believing that there is a blood meal present, and hence gets trapped [20]. Capturing male mosquitoes would have require separate traps with 'nectar' baits, that was beyond our scope. Nevertheless, it is generally true that external morphological characteristics of both males and females for any particular mosquito species are visually similar (with males consistently having a feather like proboscis [18]), and hence we are confident that our proposed techniques can be easily adapted to detect both species and genders, and is part of our future efforts, with more experiments.

Table 2: Samsung Galaxy S5 Camera Features

Camera Details	Specification
Sensor Resolution	16 MP
Aperture size	F2.2
Focal length	31mm
Shooting Mode	High Dynamic Range mode
Camera Light Source	Daylight
Background	White, Yellow & Pink

4 OUR TECHNICAL APPROACH

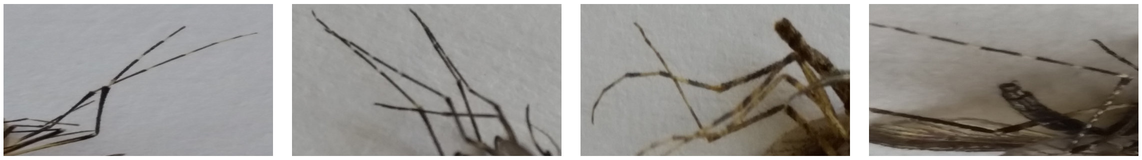
In this section, we present our technical approach to classify mosquito species from smart-phone images. There are a sequence of steps in our approach - image resizing, noise removal, background segmentation, feature extraction, dimensionality reduction, unsupervised clustering and classification.

4.1 Image Resizing

In our case, a single smart-phone image contains 2988×5322 pixels. This is large, and will computationally be prohibitive for the phone during image processing and features extraction, and even more so when there are multiple images. For practicality, we resize each image captured to a size of 256×256 pixels. This reduced the image size from around 3MB to 16KB, making processing much more practical and fast during model development and also run-time execution, without compromising accuracy.

4.2 Noise Removal

In our paper, we implemented a median filter to reduce noise. Median filter [17] is a nonlinear technique, where each pixel value



a). *Aedes aegypti* **b**). *Aedes taeniorhynchus* **c**). *Coquillettidia perturbans* **d**). *Psorophora columbiae*

Figure 2: Edge Contrast in Legs of Different Species. This Figure is best viewed in Color.

a). *Aedes aegypti* **b**). *Aedes taeniorhynchus* **c**). *Coquillettidia perturbans* **d**). *Psorophora columbiae*

Figure 3: Color Contrast in Wings of Different Species. This Figure is best viewed in Color.

in a window of size $n \times n$ pixels is replaced by the median of all pixel values in that window. In our case, we choose $n = 3$. In other filtering techniques like mean filter, pixels are replaced by mean values in a window, and in some cases, the mean value computed is not one that is actually there in the image, resulting in poorer retention of image fidelity, which also compromises edge and color preservation. Median filters avoid this problem, since median values of pixels are computed and retained during noise removal.

For our problem, edge and color preservation are crucial since textural patterns of a mosquito that make up the edges (e.g., legs and wings), and their colors, aid in classification. For example, from Figure 2, we see that the legs of *Aedes aegypti* and *Psorophora columbiae* have a combination of black and white color patterns; and the legs of *Aedes taeniorhynchus* and *Coquillettidia perturbans* have yellowish and black patterns. But the white and black patches in the case of *Psorophora columbiae* are thinner than that of *Aedes aegypti*. Similarly from observation of Figure 3, we see that the wings of *Aedes aegypti* are slightly whiter compared to others; the wings of *Psorophora columbiae* are slightly blacker than others; and those of *Aedes taeniorhynchus* and *Coquillettidia perturbans* are more brown [2]. There are distinct color/ textural patterns even in the scales and shapes of wings of various species, hence demonstrating the importance of edge and color preservation, and the importance for median filters to remove noise.

4.3 Background Segmentation

The next step is background segmentation. Since, we anticipate mosquito images to be captured in a variety of backgrounds, compensating for differing backgrounds is vital. The technical challenge here is automatically segmenting out all of the background information, while retaining only the region of interest (i.e., the mosquito).

In our technique, we employ a 2-step process. The first step is to detect the edges of the mosquito in the image to find contours, that actually encompass a significant part of the image [6]. Following which, we identify portions within the image that need to be categorized as background by comparing images before and after contour detection. To do so, we implemented Sobel edge detection algorithm for our problem, where the algorithm takes the derivative of each pixel intensity (retrieved after converting image to gray scale) with respect to its neighboring pixel [29]. The derivative of the image is discrete as it consists of a $2D$ array and we need to take it in two directions: x-axis and y-axis. For example, the derivative of any arbitrary pixel in the x-axis will be calculated by taking the difference of pixel intensities between its left and right neighbor. The same applies to compute the derivative in y-axis. Whenever there is edge, there is a prominent change in pixel intensity. This will cause significant change in derivative value. This significant change denotes the presence of edge. In order to identify contours, we need to know edge intensity and its direction. Direction of the edge, θ is calculated as $\theta = \tan^{-1} \frac{g_x}{g_y}$, where g_x and g_y are the derivatives of each pixel intensity in x and y axis while edge intensity is calculated as, $Edge\_Intensity = \sqrt{g_x^2 + g_y^2}$.

After retrieving direction and intensity, we get many contours enclosed within the edges. The significant contours encompass the largest number of (x, y) coordinates. Then we compare the locations of each pixel of the significant contours with the locations of pixels in the original image. The pixel intensity at locations which are not in the significant contour are considered as background. While this may look like it solves our problem, there is one issue. For those portions of the background that are enclosed within identified edges (e.g., within mosquito legs), those are not segmented out, and are considered a part of the mosquito still. Such problems don't exist in regular image processing applications like face detection. However, correcting this issue is accomplished in our next step.

Now that certain portions of the background are extracted, the next step is to create a probabilistic model which assumes that

[2]Figures 2 and 3 are best viewed in color.

Figure 4: Results of Background Segmentation: Original Image taken in Pink Background, Segmentation with Significant Contours, Segmentation with Integration of Significant Contours and Gaussian Mixture Model

the background pixels are generated from a Gaussian mixture [3][30][31]. In this step, we create different Gaussian mixtures for known background pixels (*RGB* color space background pixels retrieved from the first step). For accurately segmenting the background from the mosquito image, we introduce a threshold called T. In our set-up, if the probability that the intensity of any pixel belongs to the Gaussian mixture is higher than T, that pixel is considered as background and is segmented out. In case of images with many background portions, only a few of them will be considered as background if T is set too low, while if it is too high, then it will treat portions of the foreground image as background. We initialize T with a random number between 0 to 1, and with repeated trial and error, we identify that setting $T = 0.65$ gives us best results.

In our problem, we expect a relatively uniform background, since the smart-phone needs to be close to the mosquito during imaging, and overall focus area is less. As such, we believe these parameter settings are general across backgrounds. Note that, since the distribution of pixels in the background is known apriori, shadows, and other portions of the background enclosed within edges are also removed in this technique. The effectiveness of our proposed 2-step approach in segmenting the background from an *Aedes aegypti* mosquito image taken in a pink background from our dataset is shown in Figure 4.

4.4 Feature Extraction

The next step in our system is feature extraction. Unfortunately, the standard *RGB* color space did not us give good results since the perceptible color differences across species is minimal there. We then proceeded with the *Lab* color space [27], that also considers lightness as a factor for determining color, and provides superior color perception [2]. This color space has three dimensions where, L represents lightness, and a and b represent the the color opponents ranging from green−red and blue −yellow.

In order to extract features after transforming images to *Lab* color space, we focused on textures. Recall from Figures 2 and 3 the importance of textures (patterns and colors in legs and wings) in aiding species identification. Furthermore, textural patterns do not change much as the mosquito grows, and interacts with nature in the wild. Essentially, in texture analysis, we derive the dependency of intensity or variance across pixels in the image. This can be done in two ways. One is structural that captures dependencies among neighboring pixels, that enables superior perception of textures as primitives (spots, edges, curves and edge ends). The other is statistical, that computes local features by analyzing the spatial distribution of gray values of an image [16].

Local Binary Patterns [12] is a popular approach that extracts a combination of structural and statistical properties of an image. In this technique, textures are extracted on the basis of local patterns formed by each pixel. To do so, each pixel is labeled by thresholding the 3×3 neighborhood of each pixel with the center pixel value. In other words, for each pixel of an image, we compare the pixel value of their 8 neighbors either clockwise or counter-clockwise. If the neighbor pixel value is greater than center's pixel value, we replace it with 1, otherwise with 0. This will give 8 binary digits, which are converted to decimal values, which will replace the value in the center pixel. The process repeats for all pixels in the image. The range of decimal values lies from 0 to 255. In Figure 5, we show a representative instance of determining Local Binary Patterns.

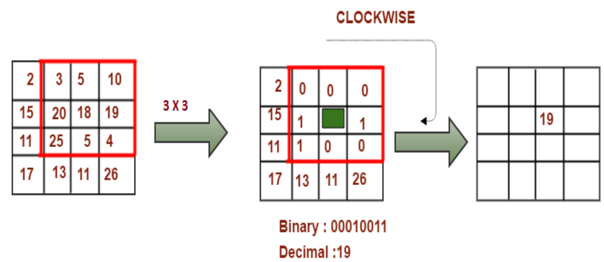


Figure 5: Local Binary Pattern Calculation for a Single Pixel

We then derive a histogram with 26 bins for the number of decimal values in each pixel in the range of 0 to 9; 10 to 19 and so on, up to 250 to 255. The number of values in each of the 26 bins is a feature. Essentially, when the number of bins with non-zero entries is less, it indicates fewer textural patterns, and when it is more, it is an indicator of more textural patterns.

While Local Binary Patterns do yield structural and statistical information on local textures, they cannot capture spatial dependencies among textures, which contrast mosquito species (e.g., alternating black and white patches in legs, variations in thickness of patches etc.). To capture these on a global scale, we derive Haralick textural features, which employ higher order statistics to capture neighborhood properties of textures.

The basis of Haralick features [15] is a gray-level co-occurrence matrix, where gray-level indicates the intensity of a pixel in two dimensions. At the start, a square matrix of dimensions $G = N_g \times N_g$ is constructed, where N_g denotes the number of gray levels in an image. An Element $[i,j]$ in the matrix is generated by counting the number of times a pixel with value i is adjacent to a pixel with value j, and then dividing the entire matrix by the total number of such comparisons made. Each entry in the matrix is thus the probability that a pixel with value i will be found adjacent to a pixel of value j. Subsequently, using the pixel intensity dependencies identified in Matrix G, we compute 13 Haralick features to capture spatial dependencies across textural patterns in the image. Table 3 presents these features, and how to compute them from the Matrix G below, where $p(i, j)$ is defined as the probability that a pixel with value i will be found adjacent to a pixel of value j.

$$G = \begin{bmatrix} p(1,1) & p(1,2) & p(1,3) & \dots & p(1,N_g) \\ p(2,1) & p(2,2) & p(2,3) & \dots & p(2,N_g) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p(N_g,1) & p(N_g,2) & p(N_g,3) & \dots & p(N_g,N_g) \end{bmatrix}.$$

4.5 Dimensionality Reduction

Recall now that we have extracted 39 features from each mosquito image: 26 LBP and 13 Haralick Features. To make our solution computationally efficient, we employed Linear Discriminant analysis [21] for dimensionality reduction, where the aim is to find a linear combination of the 39 features by projecting them into a lower dimensional sub-space to avoid computational cost and over fitting, while the identified subspace maintains class variability and reduced correlation among features. To do so, let us assume, we have K classes and each having mean μ_i, and covariance $\sum$, where $i = 1, 2, 3, \dots .K$. Then, the scatter between class variability is defined using sample covariance of the class means as:

$$\sum_b = \frac{1}{K}\sum_{i=1}^{K}(\mu_i - \mu)(\mu_i - \mu)^T, \tag{1}$$

where μ is the mean of the all class means. The separation of class in a direction $\vec{w}$, which is an eigenvector of $\sum^{-1}\sum_b$, is computed as,

$$S = \frac{\vec{w}^T \sum_b \vec{w}}{\vec{w}^T \sum \vec{w}}. \tag{2}$$

If $\sum^{-1}\sum_b$ is diagonalizable, the variability between features will be contained in the subspace spanned by the eigenvectors corresponding to the $K-1$ largest eigenvalues (since $\sum_b$ is of rank $K-1$ at most). These $K-1$ values will be our features for classification. In our case, since we have nine classes of mosquito species, eight final features are returned after LDA, that will be used for model development.

4.6 Unsupervised Clustering

Our first attempt to classify mosquito species is to investigate the efficacy of our eight features extracted as above, by checking to see if an unsupervised learning algorithm can by itself cluster image samples. To do so, we designed as Expectation-Maximization (EM) algorithm [7] for clustering unlabeled mosquito images, where the idea is to estimate the Maximum Likelihood (ML) parameters from the observed samples. Assuming that each image is sampled from a mixture of Gaussian distributions, the EM algorithm attempts to find the model parameters of each Gaussian distribution from which the sample most likely is observed, while increasing the likelihood of the parameters in each iteration. It comprises of two steps in each iteration. In the expectation, or E-step, model parameters are estimated based on observed samples. This is achieved using the conditional expectation. In the M-step, the likelihood function of model parameters is maximized under assumption that the observed sample is sampled from the estimated parameter. The iteration goes until convergence. Convergence is guaranteed since the algorithm is bound to increase the likelihood function at each iteration.

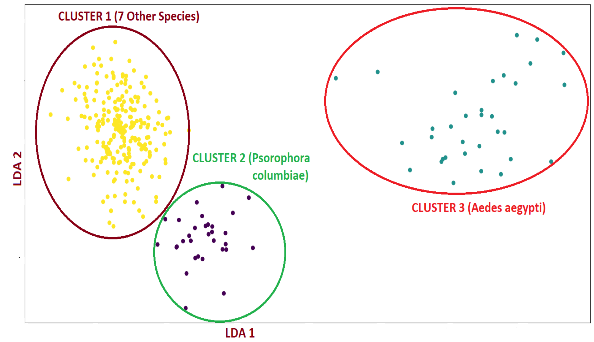


Figure 6: Three Clusters Identified after EM Clustering

With this clustering technique, we found very good performance when the number of clusters selected were 3, and with top 2 LDA features having highest variance. Figure 6 presents results, where all samples belonging to *Aedes aegypti* and *Psorophora columbiae* were each clustered separately using just 2 features. This is a very interesting result from unsupervised clustering that justifies our selection of features as representative. However, all samples in 7 other species were clustered separately. These species are identified in Table 4.

4.7 Classification Method

With two of the three species already identified via clustering, we present the final step of classifying the remaining 7 species. To do so, we use Support Vector Machines [9], which is an established supervised classification and regression machine learning algorithm, and requires minimal overhead to train and test. It gives fast and high performance with very little tuning of parameters. The main aim in SVM is to maximize the margin between classes to be identified by determining training instances that are called as support vectors which are used to define class boundaries. The middle of the margin is the optimal separating hyperplane between two classes. While testing, we calculate the probability of each sample belonging to particular species and output the one that has highest probability.

Recall that, we are taking three smart-phone images of each mosquito specimen in different orientations. As such, three images will be given for classification in each instance. Since the number of species to be identified is only seven (after Clustering), for features from these samples alone, we reapply LDA to identify six features for classification. When implementing the SVM algorithm for this set (of 3 images each per specimen to be identified), we compute the average probabilities of each species as identified from the SVM algorithm for each of the 3 images, and output the one with the highest average probability among all species classified.

Table 3: Formulas for Haralick's 13 features

Features	Formula
Angular Second Moment	$\sum_i \sum_j p(i,j)^2$, where $p(i,j)$ is defined as the probability that a pixel with value i will be found adjacent to a pixel of value j
Contrast	$\sum_{n=0}^{N_g-1} n^2 \{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)\}, \lvert i-j \rvert = n$
Correlation	$\frac{\sum_i \sum_j (i,j) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$, where x and y are the row and column of an entry in co-occurrence matrix G, and $\mu_x, \mu_y, \sigma_x, \sigma_y$ are the means and std. deviations of p_x, p_y which is partial probability density functions of pixel x and y respectively
Sum of Squares: Variance	$\sum_i \sum_j (i-\mu)^2 p(i,j)$
Inverse Difference Moment	$\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
Sum Average	$\sum_{i=2}^{2N_g} i p_{x+y}(i)$, where $p_{x+y}(i)$ is the probability of the co-occurrence matrix coordinates summing to $x+y$
Sum Entropy	$\sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\} = f_s$
Sum Variance	$\sum_{i=2}^{2N_g} (i-f_s)^2 p_{x+y}(i)$
Entropy	$-\sum_i \sum_j p(i,j) \log(p(i,j))$
Difference Variance	$\sum_{i=0}^{N_g-1} i^2 p_{x-y}(i)$
Difference Entropy	$\sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\}$
Information Measure of Correlation 1	$\frac{HXY - HXY1}{max\{HX, HY\}}$, where HXY = $-\sum_i \sum_j p(i,j), HX, HY$ are the entropies of $p_x, p_y, HXY1 = -\sum_i \sum_j p(i,j) \log\{p_x(i) p_y(j)\}$
Information Measure of Correlation 2	$(1 - exp[-2(HXY2 - HXY)])^{1/2}$, where HXY2 = $\sum_i \sum_j p_y(j) \log\{p_x(i) p_y(j)\}$

Table 4: Cluster Results

Cluster	Species
1	*Aedes infirmatus*, *Aedes taeniorhynchus*, *Anopheles crucians*, *Coquillettidia perturbans*, *Culex nigripalpus*, *Mansonia titillans*, and *Psorophora ferox*
2	*Psorophora columbiae*
3	*Aedes aegypti*

5 RESULTS

a). Overview of Evaluation Methods: Recall that for two species, namely *Aedes aegypti* and *Psorophora columbiae*, the classification accuracy was 100% with Clustering alone. For the other seven species, we evaluate the ability of our SVM algorithm for classification under 10-fold cross validation technique, which is standard for our problem scope.

b). Results and Interpretations: Figure 7 presents results in terms of Precision, Recall and F1-Measure for seven species, wherein for each specimen, the average classification probability for all 3 images of that specimen are computed, and the highest one is returned. The accuracy in this case for these seven species is 71.07%. Combined with 100% accuracy for two other species, the overall accuracy of our system for all nine species is 77.5%.

For curiosity, we attempt to output two species which have the top two highest classification probabilities from SVM, instead of only the top most (as shown above in Figure 7). In other words, we will consider our system accurate if the actual species is among the top two species outputted from our algorithm. Figure 8 presents results, and the accuracy naturally improves to 87.15% for the 7 species, resulting in an overall accuracy for nine species as 90.03%.

Interestingly, if we aim to identify each image of each specimen separately (without considering them as part of a set), the accuracy is only 47.16%. We do not present figures in this paper due to space limitations, but it reveals the importance of capturing images in multiple orientations for enhanced accuracy to identify mosquito species, which as we hope readers agree is quite practical for our application scenario, where citizens engage in the imaging/ species identification process. In fact, for visual identification under a microscope, usually one orientation is not sufficient, and multiple orientations are needed for species identification even for experts.

c). Complexity of Execution: Training our EM Clustering, and Support Vector machine classification model were implemented on a machine with Intel Core *i*7 CPU @2.6 GHz with 16 GB RAM configuration. Training the model took less than a few minutes.

We implemented the entire process of classification (image preprocessing, feature extraction, LDA, Clustering and Classification algorithm) as an application on a Samsung Galaxy S5 Smart-phone. The average time it took to classify a species was less than 2 seconds, with negligible energy consumption. Total memory consumed by the application in the phone was 23MB.

d). Difficulties in Designing Deep and Transfer Learning Techniques to Identify Mosquito Species: We understand that deep-learning is state-of-art in object recognition. However, for effective model development using deep learning, tens of thousands of images are needed, since deep learning enables automatic feature extraction from the dataset. Generating 303 images in this paper was itself a challenge. Generating tens of thousands of mosquito images requires much more resources. Data Augmentation in one approach to create larger datasets via flipping, blurring, zooming and rotating images [25]. But this was not effective for us, because these are regularization techniques, that have applicability when images classes are more diverse. But since there is minimal diversity in the physical appearance (and hence images) among various species of mosquitoes, this approach will likely introduce more noise, resulting in poorer accuracies. Our attempt in generating a dataset of 2000 mosquito images from the original 303, using augmentation, followed by species classification yielded an accuracy of only 55%. Enhancing our dataset size using open source images (e.g., Google Images) are not possible because there were not enough images tagged with the name of species, and even then we cannot guarantee that they were correctly tagged.

Another more recent technique is Transfer Learning, where the idea is to extend an existing model already trained to identify certain classes, in order to identify newer classes. Unfortunately, even the most popular VGGNet model [28] trained to recognize 1000 classes of images using the ImageNet database [11] fetched us only 47% accuracy. Primarily, no class among the 1000 in ImageNet were even remotely representative of mosquitoes, hence explaining low accuracy in species classification using Transfer Learning.

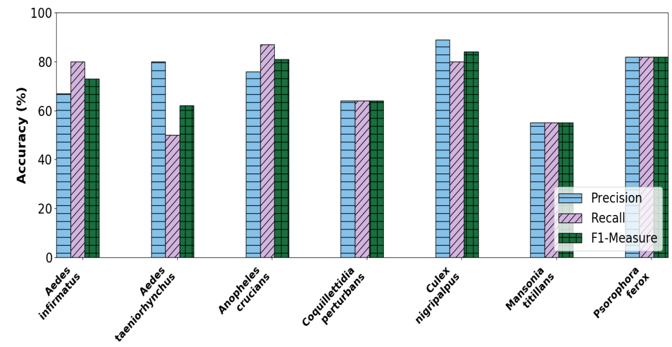


Figure 7: Precision, Recall and F1-Measure for 10-fold Cross-Validation Method for Seven Species

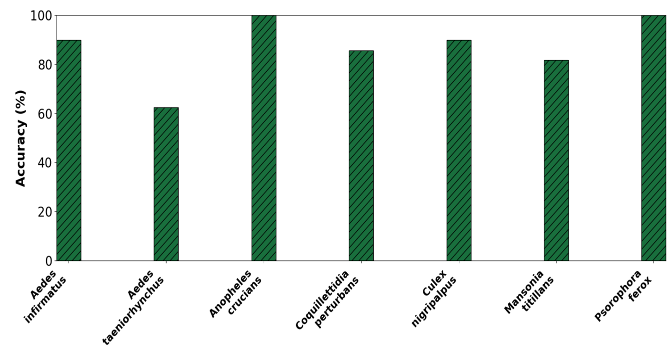


Figure 8: Accuracy of Top 2 Results for 10-fold Cross-Validation Method for Seven Species

6 PRACTICAL IMPACT AND FUTURE WORK

In this paper, we design a system that allows any citizen to take image(s) of a still mosquito that is either alive or dead (via spraying or trapping), but still retaining its physical form, and subsequently processes the image(s) to identify the species type in real time.

a). Practical Impact: At peak times, hundreds of requests come daily from people complaining of mosquitoes in their neighborhoods. Deciding where to divert resources for trap laying and spraying is a constant problem for public health workers. In fact, in Florida, during the Zika Virus scare in 2016, the lack of information about species type during calls from concerned citizens was a huge problem for public health workers we spoke to. With knowledge on species type and density, reported by citizens themselves using our system, urgent needs can be better prioritized. Furthermore, with a system like ours in place available at mosquito control facilities, the process of species identification and logging is much faster. Expertise of public health workers can hence be shifted from the cognitively demanding task of species identification via a microscope, to more useful tasks in combating mosquitoes spread.

b). Future Work: We are now generating images of more mosquito specimens (male and female) in the Hillsborough County. With more species and specimens, and using more smart-phones for imaging, we hope to demonstrate superior validity of our system. The process of data collection though is very laborious, requiring months of laying traps, and tagging/ imaging specimens. We are now working with public health experts to design a user-friendly smart-phone app that citizens can use for imaging, classification and reporting of mosquitoes. After testing, we will release it for public use in the Hillsborough county, and evaluate it. Images collected and tagged in this manner will also be publicly shared. Expanding our results to beyond Florida, and possibly beyond the US is also on our agenda, but is very challenging - technically and logistically.

ACKNOWLEDGMENTS

We are grateful to the staff at Hillsborough County Mosquito and Aquatic Weed Control Unit for helping us with mosquito species collection and identification. This work was supported in part by US National Science Foundation under Grant CBET #1743985. Any opinions and findings are those of the authors alone and do not necessarily reflect the views of the funding agency.

REFERENCES

[1] 2018. CDC - Malaria - About Malaria - Biology - Mosquitoes - Anopheles Mosquitoes. https://www.cdc.gov/malaria/about/biology/mosquitoes/. (Accessed on 03/02/2018).

[2] 2018. Explanation of the LAB Color Space. https://www.aces.edu/dept/fisheries/education/pond_to_plate/documents/ExplanationoftheLABColorSpace.pdf:. (Accessed on 02/24/2018).

[3] 2018. Image Derivative · Chris McCormick. http://mccormickml.com/2013/02/26/image-derivative/. (Accessed on 03/01/2018).

[4] 2018. Project Premonition - Microsoft Research. https://www.microsoft.com/en-us/research/project/project-premonition/. (Accessed on 02/23/2018).

[5] Ahmedelmubarak Bashir, Zeinab A Mustafa, Islah Abdelhameid, and Rimaz Ibrahem. 2017. Detection of malaria parasites using digital image processing. In *Communication, Control, Computing and Electronics Engineering (ICCCCEE), 2017 International Conference on*. IEEE, 1–5.

[6] D Baswaraj, A Govardhan, and P Premchand. 2012. Active contours and image segmentation: The current state of the art. *Global Journal of Computer Science and Technology* (2012).

[7] Sean Borman. 2004. The expectation maximization algorithm-a short tutorial. *Submitted for publication* (2004), 1–9.

[8] Yanping Chen, Adena Why, Gustavo Batista, Agenor Mafra-Neto, and Eamonn Keogh. 2014. Flying insect detection and classification with inexpensive sensors. *Journal of visualized experiments: JoVE* 92 (2014).

[9] Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine learning* 20, 3 (1995), 273–297.

[10] Anna Monica M De Los Reyes, Anna Camille A Reyes, Jumelyn L Torres, Dionis A Padilla, and Jocelyn Villaverde. 2016. Detection of Aedes Aegypti mosquito by digital image processing techniques and support vector machine. In *Region 10 Conference (TENCON), 2016 IEEE*. IEEE, 2342–2345.

[11] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 248–255.

[12] Anita Dixit and Nagaratna P Hegde. 2013. Image texture analysis-survey. In *Advanced Computing and Communication Technologies (ACCT), 2013 Third International Conference on*. IEEE, 69–76.

[13] Colin Favret and Jeffrey M Sieracki. 2016. Machine vision automated species identification scaled towards production levels. *Systematic Entomology* 41, 1 (2016), 133–143.

[14] Masataka Fuchida, Thejus Pathmakumar, Rajesh Elara Mohan, Ning Tan, and Akio Nakamura. 2017. Vision-based perception and classification of mosquitoes using support vector machine. *Applied Sciences* 7, 1 (2017), 51.

[15] Robert M Haralick, Karthikeyan Shanmugam, et al. 1973. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics* 6 (1973), 610–621.

[16] PS Hiremath and Rohini A Bhusnurmath. 2013. Texture Image Classification Using Nonsubsampled Contourlet Transform and Local Directional Binary Patterns. *International Journal* 3, 7 (2013).

[17] T Huang, GJTGY Yang, and G Tang. 1979. A fast two-dimensional median filtering algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 27, 1 (1979), 13–18.

[18] PG Jupp. 1996. Mosquitoes of Southern Africa: Culicinae and Toxorhynchitinae. Hartebeespoort.

[19] Maxime Martineau, Donatello Conte, Romain Raveaux, Ingrid Arnault, Damien Munier, and Gilles Venturini. 2017. A survey on image-based insect classification. *Pattern Recognition* 65 (2017), 273–284.

[20] Conor J McMeniman, Román A Corfas, Benjamin J Matthews, Scott A Ritchie, and Leslie B Vosshall. 2014. Multimodal integration of carbon dioxide and other sensory cues drives mosquito attraction to humans. *Cell* 156, 5 (2014), 1060–1071.

[21] Sebastian Mika, Gunnar Ratsch, Jason Weston, Bernhard Scholkopf, and Klaus-Robert Mullers. 1999. Fisher discriminant analysis with kernels. In *Neural Networks for Signal Processing IX, 1999. Proceedings of the 1999 IEEE Signal Processing Society Workshop*. IEEE, 41–48.

[22] Mona Minakshi, Pratool Bharti, and Sriram Chellappan. 2017. Identifying mosquito species using smart-phone cameras. In *Networks and Communications (EuCNC), 2017 European Conference on*. IEEE, 1–6.

[23] Haripriya Mukundarajan, Felix JH Hol, Erica A Castillo, Cooper Newby, and Manu Prakash. 2016. Using Mobile Phones as Acoustic Sensors for the Surveillance of Spatio-temporal Mosquito Ecology. (2016).

[24] J. P. Munoz, R. Boger, S. Dexter, R. Low, and J. Li. 2018. Image Recognition of Disease-Carrying Insects: A System for Combating Infectious Diseases Using Image Classification Techniques and Citizen Science. (2018).

[25] Luis Perez and Jason Wang. 2017. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621* (2017).

[26] Antonio Rodríguez García, Frederic Bartumeus, and Ricard Gavaldà Mestre. 2016. Machine learning assists the classification of reports by citizens on disease-carrying mosquitoes. In *SoGood 2016: Data Science for Social Good: Proceedings of the First Workshop on Data Science for Social Good co-located with European Conference on Machine Learning and Principles and Practice of Knowledge Dicovery in Databases (ECML-PKDD 2016): Riva del Garda, Italy, September 19, 2016*. CEUR-WS. org, 1–11.

[27] Michael W Schwarz, William B Cowan, and John C Beatty. 1987. An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models. *ACM Transactions on Graphics (TOG)* 6, 2 (1987), 123–158.

[28] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[29] Irwin Sobel. 2014. An Isotropic 3x3 Image Gradient Operator. (02 2014).

[30] Chris Stauffer and W Eric L Grimson. 1999. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, Vol. 2. IEEE, 246–252.

[31] Brandyn White and Mubarak Shah. 2007. Automatically tuning background subtraction parameters using particle swarm optimization. In *Multimedia and Expo, 2007 IEEE International Conference on*. IEEE, 1826–1829.

[32] André Barretto Bruno Wilke, Rafael de Oliveira Christe, Laura Cristina Multini, Paloma Oliveira Vidal, Ramon Wilk-da Silva, Gabriela Cristina de Carvalho, and Mauro Toledo Marrelli. 2016. Morphometric wing characters as a tool for mosquito identification. *PloS one* 11, 8 (2016), e0161643.