Journal of Entomology and Zoology Studies
Available online at www.entomoljournal.com

E-ISSN: 2320-7078
P-ISSN: 2349-6800
JEZS 2017; 5(6): 1079-1082

Received: 14-09-2017
Accepted: 17-10-2017

Hafiz Gulfam Ahmad Umar
Department of Computer Science and Information Technology, Ghazi University, Dera Ghazi Khan (32200), Pakistan

Qaisar Abbas
Department of Economics, Ghazi University, Dera Ghazi Khan (32200), Pakistan

Fatima Gulzar
Department of Economics, Ghazi University, Dera Ghazi Khan (32200), Pakistan

Correspondence
Hafiz Gulfam Ahmad Umar
Department of Computer Science and Information Technology, Ghazi University, Dera Ghazi Khan (32200), Pakistan

Insect classification using image processing and bayesian network

Hafiz Gulfam Ahmad Umar, Qaisar Abbas and Fatima Gulzar

Abstract
In every discipline classification is a difficult job especially in the case of insect species. Outstanding high degree of resemblance of the appearance between distinct species classification becomes a problematic challenge. The present study discussed different techniques to classify insects especially using Bayesian network and proposed an insect identification scheme that can identify insect colored images. Four different classes of insects were selected for an evidence of concept. Classification was accomplished by manipulating insects' shape feature and their histogram and color. Since each insect has different color and distinctive body shapes, so the current study also intend to propose a Bayesian classification approach that allows proficiently development of many valuable applications in vector control for both medical and agricultural entomology. Precisely classify insects by using proposed framework would have significant implications for entomological research. The present study classification model is very robust and fit for a general classification framework that permits to easily integrate arbitrary number of features.

Keywords: insect classification, image processing, bayesian network

Introduction
Nowadays in the world 0.9 million different varieties of living insects are known [1]. From Smithsonian Institution's Department of Entomology in Latin American forest canopies studies conducted by Terry Erwin, the number of living species of insects has been estimated to be 30 million [1]. Most authorities agree that there are more insect species that have not been described: there are insect species that have been previously named. In the last decade, much attention has been given to the entomofauna that exists in the canopies of tropical forests of the world [2]. From the last sixty years to date there have been many research efforts to accomplish this task, however, none of the research has had a lasting impact. We feel that the lack of progress in this quest can be attributed to several related factors that make data collection and data classification difficult, resulting in poor-quality and limited data make this process complicated for classifiers to classify unknown insects [3]. Computerized system of image processing and object discovery is a perfect method to improve the deficiencies of this traditional technique besides ornamental accuracy. There have been many successful attempts of using machine learning in automation of laborious intensive tasks [4, 5].

After succeeding on detection, researches become interested in doing classification of an image that come from similar family tree. The challenging part in classification is the object shares almost similar features and the variation points are limited. Another difficulty of classifying an object is that some species are difficult to distinguish visually, since they differ in term of biological characteristics. One hundred species of insects are selected as samples and the image samples are captured from the right wing of the insects. A part of them are back wing with forewing, the rest are forewing [6]. The image is resized to speed up the processing speed and then filtered using mean shift algorithm [7]. Then, it is converted to gray scale image and binarized with certain threshold. A general-purpose knowledge integration framework that employs BN in integrating both low-level and semantic features. The efficacy of this framework is demonstrated via three applications involving semantic understanding of pictorial images. The first application aims at detecting main photographic subjects in an image. As Insect classification delivers a unique prospect to apply emergent techniques used in Hierarchical Deep CNNs [8] (HD-CNNs). Inspired by the success of recent work on HD-CNNs, pipeline for training and designing testing a hierarchical architecture achieves talented results considering the predominantly messy nature of our data set.

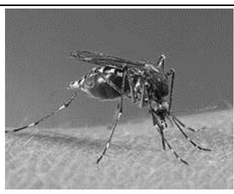
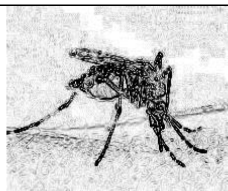
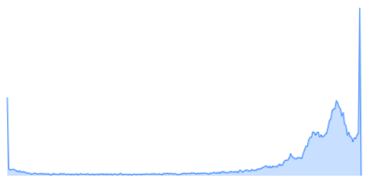
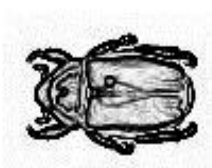
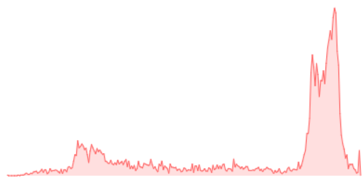
Bayesian networks are a kind of Probabilistic Graphical Model that can be used to make models from data for expert opinion [19]. These models can be practice for an extensive range of responsibilities including anomaly detection predictions diagnostics, reasoning, time series prediction and automated insight, decision making under ambiguity [9]. Each class of insect has different body shape and color, so classification is attained by manipulating insects' color and their shape feature.The proposed insect identification process starts by extracting features from and splitting them into training sets. Cataloging of insect species is a particularly difficult contest because of the high degree of resemblance of the appearance between distinct species [8]. We can easily collect large amounts insect data that a Bayesian classification approach allows to efficiently learn classification models that are very robust to over-fitting and easily incorporate arbitrary number of features. The paramount studies to use arthropod and arachnid organisms as detectors were steered by the United States Army in 1963 [10]. Image processing is the procedure of translating "pixels to predicates", i.e., iconic image illustrations to emblematic form of knowledge [11]. Image understanding is the highest processing level in computer visualization [12], as conflicting to image processing, which converts one image demonstration to another for instance converting raw pixels to an edge map. Considerable initial successes in image understanding have been made in unnatural environments such as automatic military target recognition [13] and document [14] and medical [15] image understanding but still image processing in unconstrained environments is an open problem [16]. One foremost methodology to semantic image understanding is based on the above premise. Examples or training data are collected. These patterns are fall into clusters in the feature space, and then used to train an appropriate classifier to classify novel test images. In essence, prototype based approaches apply pattern recognition techniques. Zhu Le Qing and Zhang Zhen present an investigation by using color histogram and Gray Level Co-occurrence Matrix (GLCM) for insect recognition. A survey on this specific area of research in 2012 work on optimizing accuracy, specificity trade-offs in large scale visual recognition. In this research discovered the concept of making predictions down a hierarchical class tree [17]. Here we review the past literature in vision-based insect recognition and Classifications that will helpful to enhance awareness of manipulating for image classification problems [18].

2. Materials and methods

The present study selected four kinds of arthropods for classifications: *Aedes aegypti*, scrab beetle, black window spider and Termite which have distinct color, shapes and size. In our work, we are classifying four types of insects which are stated in Fig. 1, we obtained 10 samples of data for each species and the total numbers of samples were 40 images for the data training. During data testing the present study select 4 images for each species to be tested later on the system. Most of the dataset is downloaded from internet images. We can collect data set from other sources. We chose four type of insect to be tested due to their wide range of figures and easy to attain certain characteristic from feature extraction.

Aedes Aegypti | Scrab | Black window spider | Termite

Fig 1: Four Species under study

 1a	 1b	 1c	
 2a	 2b	 3c	
 3a	 3b	 3c	

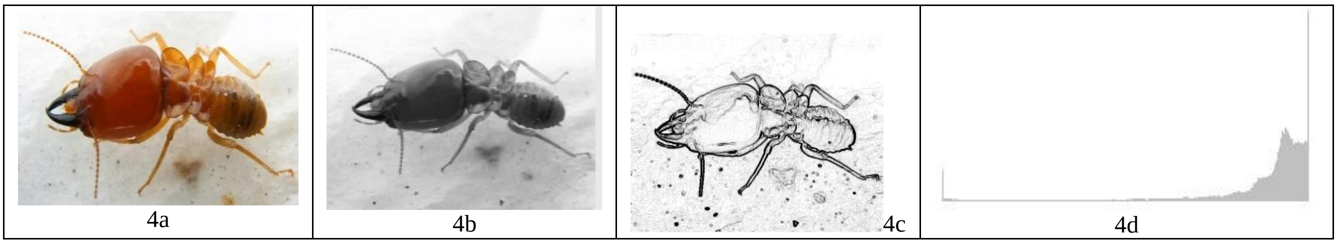


Fig 2: Feature extraction by edge detection and histogram

The next step in our experiment is abstraction of features in an image. There are several types of feature extraction techniques which are used nowadays. We chose to use feature extraction based on color and histogram. Proposed system is implemented and concentrated on visual contents of an image such as color, shape, texture and spatial layouts, histogram. From layman's bed, SIFT starts by detecting edges and corners in the image. On the resulted image, SIFT tries to find interesting a Region of Interest points that are distinguishing that image from the others. Then, out of each ROI, it extracts a histogram where each of the vats is count of specific edge or corner orientation. These histograms can be concatenated or quantized into some smaller number of groups with a clustering method like K-means. Image retrieval is implemented in two different phases. One is new image insertion with features in database and other one is searching a new image in available database. After storing process accomplished successful we are using Aedes Aegypti and Scarb pictures these two items establish two label namely ‘1a’for butterfly and ‘2a’ and 3a,4a for black window spider, lady bug respectively. Using MATLAB all image features in R, G, B color projection values are extracted and stored in database using specified programming methods. In third step Threshold calculation is taken for categorizing the images into a similar feature groups. In this step, threshold value is computed based on the histogram calculation.

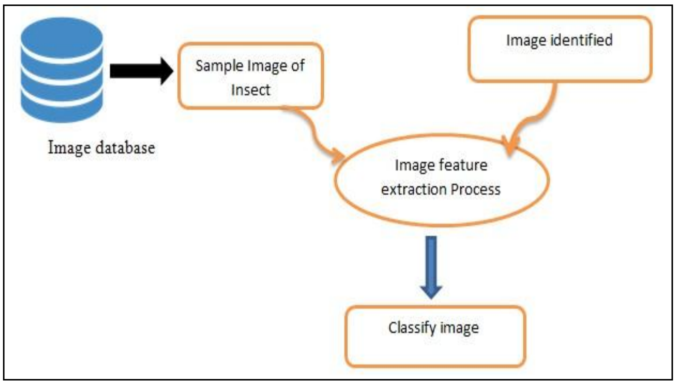


Fig 3: Graphical representation of proposed system.

They have been exercising a seven steps for classifying an insect. The seven procedures are: First, the image is loaded from a file. Then, image is converted to grayscale. Third, the edged is extracted using Canny edge algorithm. Close operator is used on the edges image.

4. Using Bayesian Network for Classification

Bayesian inference is a used for find the probability for a hypothesis as more evidence or information becomes available when there are numerous features used for classification, we require to consider the possibility of missing values, which happens when some remaining features are not observed. Bayes theorem provides a way of calculating posterior probability P(c|x) from P(c), P(x) and P(x|c). Look at the equation below:

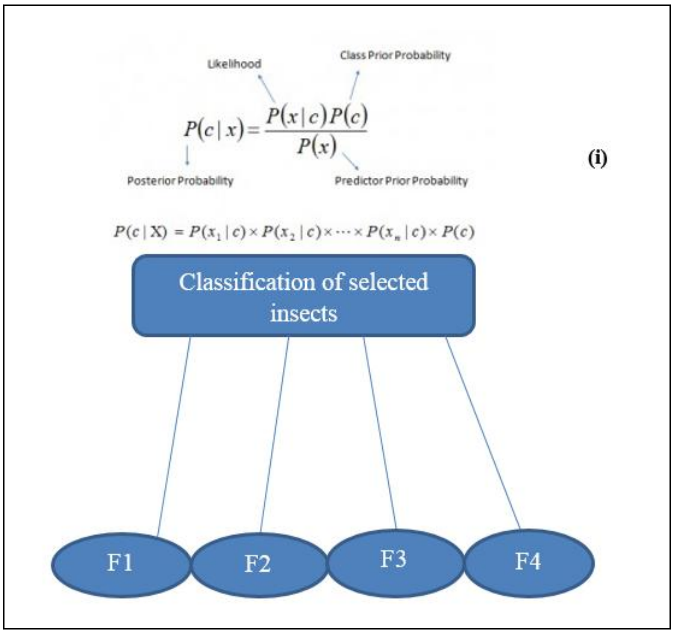


Fig 3: Nominated features of Classified insects.

Four types of insects are selected that are given above. F1…..F4 are nominated features of selected classified insects. Bayesian inference is used to find posterior probability as the consequences of selected classified insects of prior probability. These probability values computed from observed data.

Basin theorem; $P(H(i)/E)^n = \frac{P(\frac{E}{H}).P(H)}{P(E)}$ (2)

H (i) = stands for hypothetical probability (insect classification has same feather in image processing)
E = corresponds to new data that were not used in computing the prior probability (classified insects has likelihood feathers in image process).
P (H) = is the estimate of the probability of the hypothesis, **H** before the data **E**, the current evidence that is observed after image processing.
P (H/E) = is the posterior probability, the probability of hypothetical value and current evidence data that given after applying image process.
Classification values designated that hypothetical values are accepted. Probability value of H/E gives the 80 percent accuracy values. It described that image processing in four selected insects and after their classifications are inters link feathers. Each class have 70- 84 percent likelihood features with their cross ponding particular insect group.

5. Result and Discussion

The Proposed system accuracy prediction was suitable as compare to the total number of predictions. The system attained accuracy score of approximately 80 %. Experimental results shows that Aedes due to variation in shape 70% low classification prediction result as compare to Scrab. Black window spider prediction rate 73% while Termite had 71% due to there small size and shape if we compare our results to previous work especially in insect classification our work significant with different type of insects and limited parameters.

6. Conclusion

This paper suggests a spontaneous insect identification framework that can identify insect from colored images. Four types of insects were chosen for evidence. Our system test image to precise calculations over the total satisfactory accuracy score. However, with a record of 80% while insects variations of the image color and shape acquire 100% accuracy. Our classifier Classification values designated that hypothetical values are accepted. Probability value of H/E gives the80 percent accuracy values. It described that image processing in four selected insects and after their classifications are inters link feathers. Each class have 70- 84 percent likelihood features with their cross ponding particular insect group.

6.1 Future Work

The proposed insect classification system produced satisfactory results but there needs a lot of work for improvement. As future work, we want to accumulate more training data set to enhance the system parameters by using optimization techniques such as genetic algorithm. In next step we apply proposed framework to identify a diversity of other species of insects.

7. Reference

1. https://www.si.edu/Encyclopedia_SI/nmnh/buginfo/bugnos.htm
2. Chapman AD. Numbers of living species in Australia and the world, 2009.
3. Htike ZZ, Win SL. Recognition of Promoters in DNA Sequences Using Weightily Averaged One dependence Estimators, Procedia Computer Science. 2013; 23(1):60-67.
4. Htike ZZ, Egerton S, Kuang YC. A Monocular View-Invariant Fall Detection System for the Elderly in Assisted Home Environments.7th International Conference on Intelligent Environments (IE) 2011, 25-28.
5. Htike ZZ. Can the future really be predicted, Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA). 2013, 360-365.
6. Zhu LQ, Zhang Z. Auto-classification of Insect Images Based on Color Histogram and GLCM", Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 2010.
7. Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002; 24(5):603-619.
8. Glen C, Rains, Jeffery K, Tomberlin. Using insect sniffing devices for Detection. Trends in Biotechnology. 2016; 26(6).
9. http://www.bayesserver.com/docs/introduction/bayesian-networks
10. Lubow RE. The War Animals, Doubleday & Company, 1977.
11. Ballard D, Brown C. Computer Vision, Prentice-Hall, Englewood Cliffs, NJ, 1982.
12. Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and Machine Vision, second ed., Brooks & Cole Publishing, Pacific Grove, CA, 1999.
13. Dudgeon DE, Lacoss RT. An overview of automatic target recognition, Lincoln Lab. J. 1993; 6(1):3-10.
14. Schurmann J, Bartneck N, Bayer T, Franke J, Mandler E, Oberlander M. Document analysis—from pixels to contents, Proc. IEEE.1992, 1101-1119.
15. Vailaya A, Figueriredo M, Jain A, Zhang HJ. Content-based hierarchical classification of vacation images, Proceedings of IEEE International Conference on Multimedia Computing and Systems, 1999.
16. Robinson GP, Colchester ACF. Model-based recognition of anatomical objects from medical images, Image Vision Comput. 1994; 8:99-507.
17. Deng J, Krause J, Berg AC, Fei-Fei L. Hedging your bets: Optimizing accuracy-specificity trade-offs in large scale visual recognition. In Computer Vision and Pattern Recognition (CVPR), EEE Conference. 2012, 3450-3457.
18. Zhu LQ, Zhang Z. Auto-classification of Insect Images Based on Color Histogram and GLCM, Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 2010.
19. https://arxiv.org/abs/1407.5656