Computerized System for Identification of Pests Using Machine Vision

S.K.PRAVIN KUMAR[1] Dr.SUMITHRA.M.G[2], N.SARANYA[3]

Assistant Professor, Dept. of ECE, United Institute of Technology, Coimbatore [1].
Professor, Dept. of ECE, BannariAmman Institute of Technology, Coimbatore [2].
Assistant Professor, Dept. of IT, Sri Ramakrishna Engineering College, Coimbatore [3].
Email:skpravinece@gmail.com[1]

Abstract

"Agriculture not only gives riches to a nation, but the only riches a Nation can call as its own". Invasion of pests over crops is one of the most challenging tasks for the crop technicians and farmers in the field of agriculture. This result in damage of crops, leading to low yield and thereby incur serious loss to the farmers and the Nation as a whole. Hence early detection of pests is a key factor for crop management. A detailed study has been carried out to investigate the use of computer vision and image processing techniques in agricultural applications. A framework to identify and classify pests is presented in this paper. Image database of the pests is taken for consideration. The geometrical features that include the characteristics of the mean and variance and the morphological features that include the size and shape of the pest image are extracted. The extracted features of a set of training images are compared with the extracted features of the testing images to enable effective classification. Probability and likelihood of the pests are taken for consideration. Naïve bates classification approach is adopted to identify the class of Pests. The analyses prove that such methods can be used for agricultural applications in areas such as precision agriculture.

Keywords- Pests, image pre-processing, feature extraction, Naïve Bayes approach

I.Introduction

Agriculture is the foundation of civilization and a stable economy of a country. The discovery of agriculture is the first big step in a civilized life. In agriculture, invasion of pests is considered as the most challenging tasks for the crop technicians and farmers. It damages the crops and incur a serious loss to the farmers and thereby a decline in the economy of the Nation. Early detection of the presence of pests is a key point for the management of the crops. Integrated Pests Management (IPM) is one of the most important and accurate methods for pest control and it is a method that minimizes the environmental impacts. Integrated Pests Management involves three important strategies namely detection, identification of pests and application of the correct management. The challenge involved in detection and identification of the pests can be overcome by the development of an automatic pest identification system. Even though the morphological features of the pests such as the size, shape and color vary, it is difficult to distinguish those pests visually. Manual classification of pests in the agricultural field is time consuming and difficult too, because it requires some technical expertise.

The main purpose of the paper is to develop an automated system to classify the field pests based on their class and family. An efficient algorithm is developed for the same. The algorithm initially extracts the still images of the pest by means of image processing techniques. Based on the features extracted, classification of pest is performed using Naïve Bayes classifier technique. The entire process is developed on a MATLAB platform.

The rest of the paper is organized as the system overview in section (II), image pre-processing techniques in section (III), feature extraction mechanism in section (IV), classification technique in section (V) and finally the conclusion is provided in section (V) with additional information about how it can be extended in the future.

II. System Overview

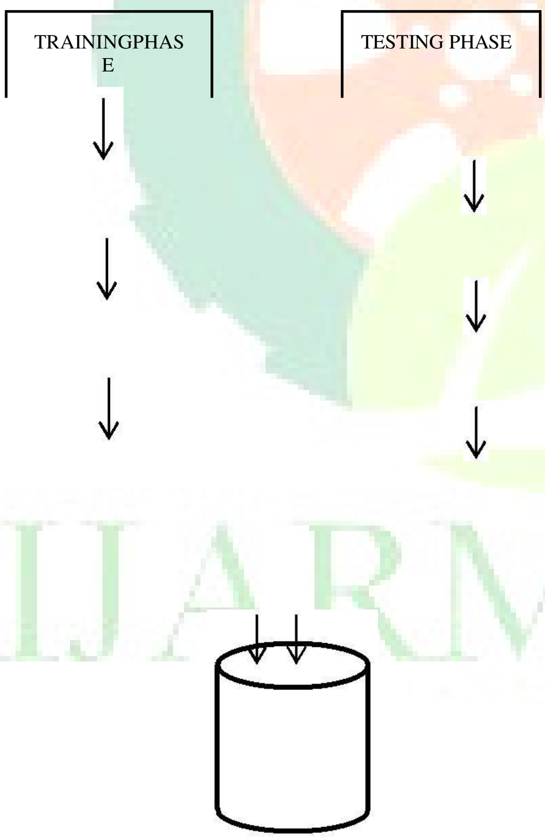
The entire system"s architecture relies on machine vision technique which gives a detailed description about image acquisition and preprocessing, feature extraction and classification of pests based on the above analysis.

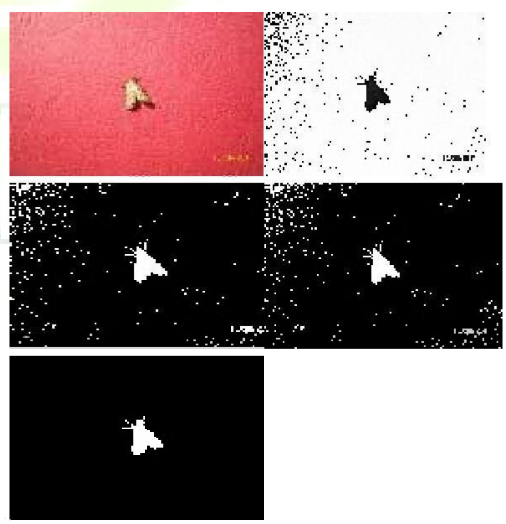
A. Acquisition and Pre-processing

The first phase is the image acquisition phase. In this phase the images of the pests that are to be classified are selected. The images as such are not suitable for classification purposes because of various factors such as noise, lighting variations etc. Hence these images would be pre-processed using certain filters to remove unwanted features in the images.

B. Feature Extraction

Feature extraction is employed so that classification of the pests can be made easier and of higher accuracy. Certain statistical analysis tasks are completed to choose the best features to represent the given image thus minimizing feature redundancy. After the extraction of features, the bayesian technique is used to classify the images according to the specific problem at hand.

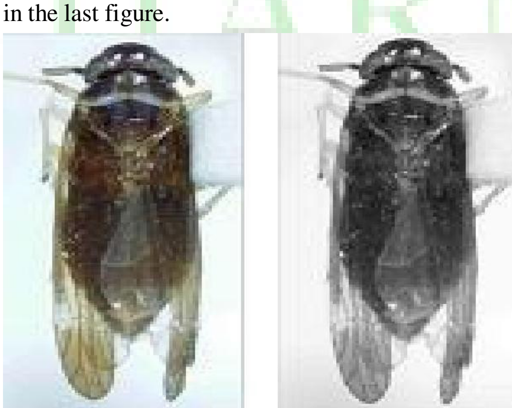


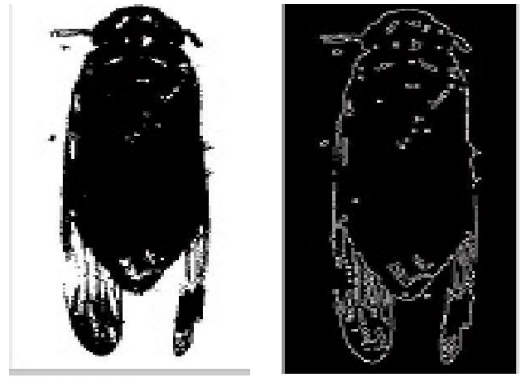


IV. Feature Extraction

The theory involved in feature extraction, which is the first step in the classification process is discussed below. The method followed for extracting the feature set is called the colour co-occurrence method or CCM method in short. It is a method, in which both the colour and texture of an image are taken into account, to arrive at unique features, which represent that image. It is well known in the image processing research community that classification accuracies are highly dependent on the feature set selection. In other words, the classification accuracy is as good as the feature set that is selected to represent the images. Therefore, careful consideration must be given to this particular step.

The first step to be followed to perform feature extraction is to obtain gray image from that of the original image. Thus it is necessary to obtain gray image with average, gray image with lightness and gray image with luminosity. The original input image of a planthopper that is obtained by programming using MATLAB is shown in first figure 3.a. The gray image that was processed from original image by means of image processing tool box using MATLAB is shown in the second figure. In order to obtain the gray image from the original image certain comments are to be given in the command window. The original image containing RGB space is converted into a gray image. The Binarized image of the planthopper that contains only the values of zeroes and ones, so called the black and white image is shown in the next figure. The border of the input image of the planthopper is shown in the next figure. The histogram of the gray image that is processed using MATLAB is shown in the last figure.





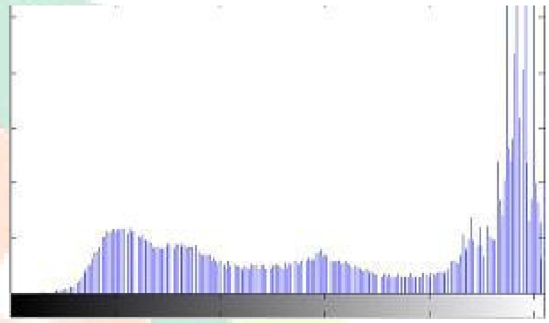


Figure3: a) original image, b)gray image, c)Binarized image, d) border of the image, e) HOG of the image

V. Classification Technique

In machine learning, a large number of classifiers are used for classification of respective applications. Some classifiers include supervised classifiers like fisher Fisher's linear discriminant, Quadratic discriminant analysis, nearest neighbours, Parzen window methods Support Vector Machines and Ensemble methods. Unsupervised classifiers include Gaussian mixture models.

A. Naïve Bayes Classifier

Naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features. A naive Bayes classifier assumes that the value of a particular feature is unrelated to the presence or absence of any other feature, given the class variable. For example, a pest may be considered to be a plant hopper if it is brown, has antennae and about 1"" in length. A naive Bayes classifier considers each of these features to contribute independently to the probability that this pest is a plant hopper, regardless of the presence or absence of the other features. For some types of probability models, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood; in other words, one can work with the naive Bayes model without accepting Bayesian probability or using any Bayesian methods.

B. Probabilistic model

Abstractly, the probability model for a classifier is a conditional model $p(C|F1\,F2 \ldots Fn)$ over a dependent class variable „C" with a small number of outcomes or classes, conditional on several feature variables F1 through Fn. The problem is that if the number of features is large or if a feature can take on a large number of values, then basing such a model on probability tables is infeasible. We therefore reformulate the model to make it more tractable.

Using Bayes' theorem, this can be written as

$$p(C|F1, F2 \ldots Fn) = \frac{p(C)p(F1, F2, \ldots Fn)}{p(F1, F2 \ldots Fn)}$$

The above equation can be thus mentioned as

$$posterior = \frac{prior * likelihood}{evidence}$$

In practice, there is interest only in the numerator of that fraction, because the denominator does not depend on C and the values of the features F are given, so that the denominator is effectively constant. The numerator is equivalent to the joint probability model $p(C|F1\,F2 \ldots Fn$ which can be rewritten as follows, using the chain rule for repeated applications of the definition of conditional probability

$$p(C|F1, F2 \ldots Fn) = p(C)p(F1, F2, \ldots Fn)$$

$$= p(C)p(F1|C)p(F2, \ldots Fn|C, F1)$$

$$= p(C)p(F1|C)p(F2|C, F1)p(F3, \ldots Fn|C, F1, F2)$$

Now the "naive" conditional independence assumptions come into play: assume that each feature Fi is conditionally independent of every other feature Fj for j≠i, given the category „C". This means that

$$p(Fi|C, Fj) = p(Fi|C)$$

$$p(Fi|C, Fj, Fk) = p(Fi\ C) \; |$$

$$p(Fi|C, Fj, Fk) = p(Fi, C)$$

To demonstrate the concept of Naïve Bayes Classification, consider the example given in Fig 4.

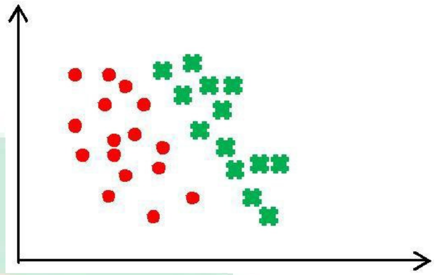


Figure 4: Graph to demonstrate Naïve Bayes approach

As indicated, the objects can be classified as either GREEN or RED. Our task is to classify new cases as they arrive, i.e., decide to which class label they belong, based on the currently exiting objects. Since there are twice as many GREEN objects as RED, it is reasonable to believe that a new case (which hasn't been observed yet) is twice as likely to have membership GREEN rather than RED. In the Bayesian analysis, this belief is known as the prior probability. Prior probabilities are based on previous experience, in this case the percentage of GREEN and RED objects, and often used to predict outcomes before they actually happen.

Prior Probability of GREEN: number of GREEN objects / total number of objects

Prior Probability of RED: number of RED objects / total number of objects

Since there is a total of 30 objects, 16 of which are GREEN and 14 RED, our prior probabilities for class membership are:

Prior Probability for GREEN: 16 / 30

Prior Probability for RED: 14 / 30

Having formulated our prior probability, we are now ready to classify a new object (WHITE circle in the diagram below). Since the objects are well clustered, it is reasonable to assume that the more GREEN (or RED) objects in the vicinity of X, the more likely that the new cases belong to that particular color. To measure this likelihood, we draw a circle around X

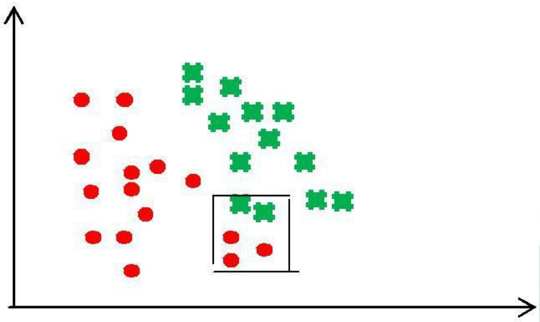


Figure 5: Graph to measure the likelihood

Liklehood of x given green

$$\text{Liklehood of } x \text{ given red} \propto \frac{\text{number of red in vicinity of } x}{\text{total number of red cases}}$$

From the illustration above, it is clear that Likelihood of X given red is smaller than Likelihood of X given green, since the circle encompasses 3 GREEN object and 2 RED ones. In the Bayesian analysis, the final classification is produced by combining both sources of information, i.e., the prior and the likelihood, to form a posterior probability using the so-called Bayes' rule. Finally, we classify X as Green since its class membership achieves the largest posterior probability.

VI. System Implementation

The implementation of the algorithm mainly consists of two main categories namely image processing followed by feature extraction and classification using naïve bayes classifier. Feature extraction involves parameters like conversion of RGB to gray image, calculation of mean, variance and diameter of the trained images and also testing images. Naïve bayes classification involves calculation of the posterior and likelihood of the training and testing images. Combining both the pests are classified based on a statistical approach, so called the probabilistic approach.

A. Training and testing dataset

Three classes of pests are identified in this pest classification approach. The name of the pests under classification is ant, feltia and planthopper. Number of training and testing images may vary based on designer"s convenience. In this pest development algorithm certain number of trained and testing dataset are taken. Number of testing and training images of the three class of pests is summarized in the table 1.

S.NO	NAME OF THE PEST	TRAINED IMAGES	TESTING IMAGES
01	ANT	3	2
02	FELTIA	4	3
03	PLANTHOPPER	5	3

Table1: Number of training and testing images for classification

VII. Conclusion and Future Enhancements

A detailed study was carried out to investigate the use of computer vision and image processing techniques in agricultural applications. The task of pest classification using the above mentioned techniques is successfully implemented. Three different classes of pests: Ant, Feltia, Planthopper was used to develop this algorithm. The image data of the pests is collected and algorithms for feature extraction and classification based on image processing techniques are designed. The manual feeding of the datasets, in the form of digitized RGB colour photographs was implemented for feature extraction and training the Naïve Bayes statistical classifier. The analyses prove that such methods can be used for agricultural applications in areas such as precision agriculture. This project is a feasibility analysis to see whether the techniques investigated in this project can be implemented in outdoor applications. The results that are obtained prove that these methods can indeed be used for such applications. However, it should be noted that all the analyses in this study are done in controlled laboratory conditions. The real world conditions are much more different due to the inherent variability in natural outdoor lighting, temperature and humidity variations. This would be a major challenge to overcome in future implementations so as to make the research portable for real time leaf classification.

The result presented in this paper is promising but several improvements are recommended to reach the requirements of fully automated pest identification system. In the future, the number of classes of pests involved in pest classification may be increased and also other image processing techniques and learning algorithm such as Support Vector Machine may be compared to enable the identification of the detected insect pests to be more efficient and accurate.

Acknowledgement

The authors would like to sincerely thank **Mr Amirthalingam Ramanan**, Senior lecturer Department of Computer Science, University of Jaffna, Sri Lanka and **Dr.K.SenguttuvanPh.D,** Assistant Professor, Department of Agricultural Entomology, Tamilnadu agricultural university, Coimbatore for providing the database of the images of the pests

References

[1] Kanesh Venugoban and Amirthalingam Ramanan, "Image Classification of Paddy Field Insect Pests Using Gradient-Based Features" in International Journal of Machine Learning and Computing, Vol. 4, No. 1, February 2014.

[2] Johnny L. Miranda, Bobby D. Gerardo, Bartolome T. Tanguilig III, "Pest Identification Image Processing Techniques in Detecting Image Pattern through Neural Network", Proc. of the Intl. Conf. on Advances In Computer and Electronics Technology- ACET 2014, ISBN: 978-1-63248-024-8.

[3] SmitaNaikwadi, NiketAmoda, "Advances in image processing for detection of plant diseases", in International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 11, November 2013.

[4] David Knight,JamesPainter,Matthew Potter , "Automatic Plant Leaf Classification for a Mobile Field Guide-An Android Application", Department of Electrical Engineering Stanford, California.

[5] Yudong Zhang and Lenan Wu, "Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine", ISSN 1424-8220, doi:10.3390/s120912489, Sensors 2012.

[6] Ruizhen Han , Yong He and Fei Liu, "Feasibility Study on a Portable Field Pest Classification System Design Based on DSP and 3G Wireless Communication Technology" in ISSN 1424 8220 Sensors 2012.

[7] Abdul Kadir, Lukito Edi Nugroho, AdhiSusanto, Paulus InsapSantosa," Leaf Classification Using Shape, Color, and Texture Features", International Journal of Computer Trends and Technology- July to Aug Issue 2011.

[8] M. J. Sarpola, R. K. Paasch, E. N. Mortensen, T. G. Dietterich, D. A. Lytle, A. R. Moldenke, L. G. Shapiro. "An acquaticimageing system to automate insect classification", Transactions of the ASABE, Copyright 2008, American Society of Agricultural and Biological Engineers, Vol. 51(6): xxx BE 67730001-2351.

[9] Ji-Xiang Du, Xiao-Feng Wang, Guo-Jun Zhang, "Leaf shape based plant species recognition", ScienceDirect-Applied Mathematics and Computation 185 (2007) 883–893.

[10] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, YuXuan Wang, YiFan Chang and QiaoLian Xiang, "A leaf recognition algorithm for plant classification using Probabilistic neural network", IEEE 7th International Symposium on Signal Processing and Information Technology, 2007, Cairo, Egypt.