K-means Based Automatic Pests Detection and Classification for Pesticides Spraying

Muhammad Hafeez Javed
Foundation University Islamabad

M Humair Noor
Foundation University Islamabad

Babar Yaqoob Khan
Foundation University Islamabad

Nazish Noor
Foundation University Islamabad

Tayyaba Arshad
Foundation University Islamabad

Abstract—**Agriculture is the backbone to the living being that plays a vital role to country's economy. Agriculture production is inversely affected by pest infestation and plant diseases. Plants vitality is directly affected by the pests as poor or abnormal. Automatic pest detection and classification is an essential research phenomenon, as early detection and classification of pests as they appear on the plants may lead to minimizing the loss of production. This study puts forth a comprehensive model that would facilitate the detection and classification of the pests by using Artificial Neural Network (ANN). In this approach, the image has been segmented from the fields by using enhanced K-Mean segmentation technique that identifies the pests or any object from the image. Subsequently, features will be extracted by using Discrete Cosine Transform (DCT) and classified using ANN to classify pests. The proposed approach is verified for five pests that exhibited 94% effectiveness while classifying the pests.**

Keywords—*Automatic plant pest detection; pest classification; Delta-E; discrete wavelet transform; support vector machine*

I. INTRODUCTION

Enhancements and improvements in the modern computer science has helped us in my unique ways. This is helping us in many ways making our daily tasks easy and convenient for us. It helps us to reduce the time consumption of different activities. Agriculture is also a fundamental part of any country. While technology is also taking part to improve it and help us to make it more easier and easy for the farmers. However it require the continues observation of expert, which was done manually, so it is time consuming process and not much efficient and expensive in large farms. As it requires experts having a precise knowledge of pest management through different techniques like black light traps and sticky traps, and framers spray the pesticides which may also harmful for the crops as well as for the environment. To reduce the effect as technology is playing role in cropping, vigilance now it is being used in the field of pest control.

The main purpose of this work is to establish such a system which may help us to detect pests as well as classify them. This will help us in future in recognizing that whether a pest is a friend one or enemy, so we can decide whether it required a spray or not. As the previous researches has shown that image processing may also help to detect plant diseases by the presences of pest. On the other hand, there are some plants with no pests, but in past they were used by some pests. During their stay many pests leave their larva over the leaves, which are more dangerous.The injuries effects of pest in plant affects the agriculture which shows the negative influence on economy of a country it.There are many pests,which badly effects the plants some are dangerous like snails,flies,caterpillar,fungal infections, while some are friendly for the crops like lady-birds,earwig and many more,so we need to identify them, because spraying blindly is very dangerous [1]. Prevention of plant from pest is more viable then to only detect and cure them after being effected by pests. For effective prevention precise knowledge about the pests management must be required before demonstration of maintaining pests.With a specific end goal to spare the work and ensure humans wellbeing, local and remote specialists have been inquired about on misusing the robot utilized for reaping products of the soil pesticides to plants in nursery diffusely [2]. If pest detection is not done, it will proceeds towards the reduction of crops,which results in increase of poverty. It is required to use a proper well organized pest detection methodology which may help us to reduce these pesticides and increase a healthy yields [3]. We have seen a typical thing that the vast majority of those models were proposed for some particular condition. Here and there these models are proposed for particularly for a solitary sort of pests in view of the colors or shape. While our condition incorporates creepy crawlies, mold, rodents, microscopic organisms and other such creatures. While a couple of these are valuable to individuals from numerous points of view, others are hurtful in nature [4]. Automatic detection of pest and determine the pests on plant through image processing and classify them on different properties of image is the best technique for accruing pest detection. To have the feasible pest control by using automatic pest detection through k-mean clustering which help to determine the pest invasion on the regular basis.

II. RELATED WORK

Earlier in this field some work has already been done. [5] has proposed a model to detect 3D position of the detected object, they have discussed that this model was proposed only for red colored pest and tested on red colored models. In [6] has explained a proper way to identify an infected area on the leaves. This approach shows that if there are some effected leaves there must be some pests. It involves some color

transformations from the RGB to l3a, l3b, H space. After this some thresholding functions were used to identify the affected area. In [4], author focused on a surveillance system, data was collected by using a binocular. This was a stereo system, which is using to control the sprays by robotic device. Some researchers used existing old traditions, like sticky materials, or pheromones, or attractive fragrances to attract the pesticides towards traps.

In [7], author has proposed a method, in which he has used some traps with sticky materials. Pests with the help of some smelly materials get attracted towards these traps. Where they get stuck, but this is a special way to reduce pests, it will not help about those pests which cannot fly, or which are at initial level. There are some common techniques to detect pests where some used clustering. Clusters basically divide an image on behalf of multiple factors, like colors, geometries, distances etc. A proposed method to detect objects from an image, where they have transformed an RGB image to another color space. Than by applying C-Mean Clustering they identify the different objects. The said technique have been used to detect effected areas in human body even from different diseases.

In [8], [9], authors proposed a novel technique to detect the pests at some initial stages. This idea was an excellent approach towards the detection of pests, but with some limitations. It only detects the White flies a specific pest. It will not work for the others. Similarly it is a vision dependent, system which depends upon the angle of vision. If this angle changes results of the said system will get damaged. In [10], author used the MATLAB neural network system toolbox to develop the forecast framework about natural product tree infections and bug bothers in light of the Backpropagation neural network as per the climate data and the event status of organic product tree illnesses and creepy crawly bugs in plantations. The history record information were used to prepare the information. Toward the end they think about both the anticipated esteems and unique information in the preparation set for the bug, or plant infection. From the outcomes we can check climate framework is performing precisely or suddenly. The applications that predicts organic product tree infections and pests irritations can use the framework.

In [3], author presented a crop based detection technique. Authors has developed a cotton crop pest's dataset. Further they have applied a thresholding function on this raw data. They have used multilayer ANN for the classification of pests.An other vision based technique used for the pest detection by [11], which was an SVM based technique used for the indoor green house. It was a yield dependent technique which is based upon the color variation. It uses very high definition images which results into more high processing and calculations. These images may not be suitable in real-time environments. In [12], author explained the uses of ANN and Bayesian, for correctly identifying a rice fields pest named as Yellow Stem Borer. They have used a dataset specifically designed for the rice field which was compiled by collecting the data over the duration of 12 years, for different climates. Earlier they have collected all the data and classify it using ANN, while after this the results collected by the ANN were further predicted using Bayesian. This methodology was well organized, but was dependent over a specific dataset [13]. There are several systems which are available among them some are color based or pest based which are presented by different researchers, among them some are [14], [15].

Reviewing the existing work it was observed that most of the models were color, yield or pest dependent. It was necessary to design a generic detection and classification technique which should not be dependent on colors or yields [1].

III. MATERIAL AND METHODOLOGY

This methodology is based upon three different steps, on initial stage pest will be detected. Pest detection consist of pre-processing to reduce noise and to prepare the image for further processing. After pre-processing this image will be segmented using k-mean clustering. Second step is to extract the features of detected pest. Third step is classification of the pest on the behalf of extracted features. A brief flow and working of the system shown in Fig. 1. At first step system will acquire the image using different technique or devices. These earlier level images are not a suitable thing for the classification. It will be preprocessed. To enhance the appearance base intensity variations we have used the Gaussian filter described in (1).

A. Image Pre-processing

To prepare the image to for the system for main operation image was preprocessed. Pre-processing consist of the steps like: I) Smoothening filter II) Color transformation.

B. Smoothening Filter

Due to different factors during capturing the image it is possible that some sharp edges appears. These sharp edges and lightening effects may cause the local maxima during our segmentation process. This could affect our system's performance. That's why this was essential to remove these edge and smooth the image. For this purpose we will use Gaussian filter. It will consider a point on the image as a center and smooth the image as per given equation.

$$G(x,y) = e^{-(x^2+y^2)/\alpha} \tag{1}$$

Gaussian filter required a Gaussian center, which is taken at (0,0), while is showing the Standard deviation.

C. Color Transformation

After smoothening images will be transformed to from RGB to L*, a*, b* color space. This process was taken to take the maximum values of L*, a*, b*, which were further be used as threshold. After pre-processing images will be as shown in

D. 2 K-mean clustering

For detection purpose here it was used k-mean clustering algorithm. It was earlier used by some of the researchers but existing techniques were facing some problems with green colored pests. In our proposed model the k-mean clustering is applied sequentially. Initially on the image was clustered using k-mean clustering. K-mean clustering technique decide on behalf of distance between two instances. We have used Euclidean distance to decide about any point on the picture.
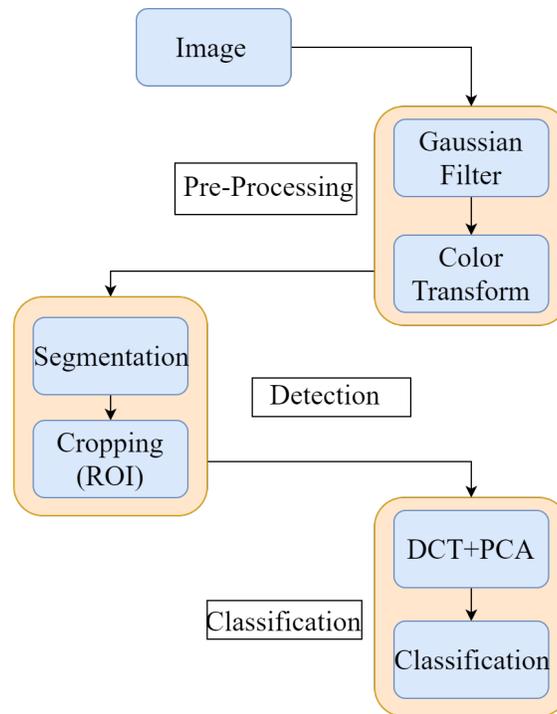
Fig. 1. A complete flow of proposed model.
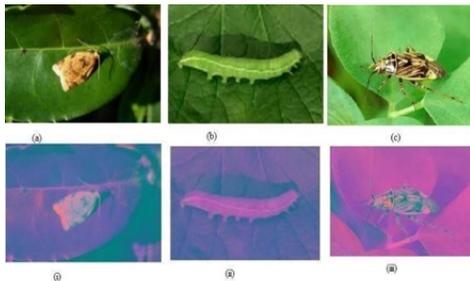
Fig. 2. Input images and their corresponding preprocessed images.
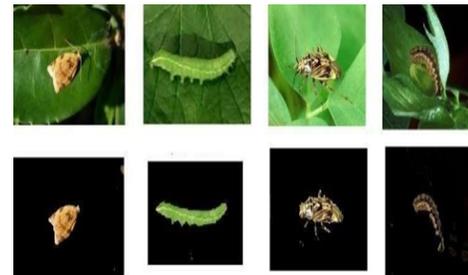
Fig. 3. Input images and their cropped images.

To overcome the colored problem we have improved the working of this detection algorithm. Our algorithm make the segments on the behalf of initially on the number of colors than remaining will be clustered in the behalf of distance. Due to this detection results were based upon the number of colors we used to classify. We selected 5 number of colors in which image has to be segmented. This threshold value was selected on the experimental results observed by us. K-means clustering is a technique from vector quantization. Basically, it is a Signal processing technique, which was utilized in Data Mining for the data clustering. K-mean Clustering normally measure 'n' Observations, then categorized in k Clusters. This decision will be made on the behalf of nearest matched clusters. It normally takes a set of observations X1, X2, X3, ... Xn, where each observation from this set is a p-dimensional real vector. The said techniques actually categorize the sample in k number of classes as S1, S2,S3, ... Sk where (k¡=n). Actually the combinations are combined using Euclidean distance, the objects having the minimum distance will be clustered in one class. While the distance of the clusters will be calculated from the center which were selected earlier.

E. Classification

Segmented images will be cropped on the next step. At this step the area of interest will be cropped as per shown in Fig. 2.

Fig. 3 shows Input Images and their cropped images. These are the results when number of colors are taken 3. After this cropping we have extracted the features using DCT. Feature set was too much large almost of 50x 50. As such high dimensions need a complex ANN. While on the behalf of results it was seen that some of the features were not effecting the results in both ways. So we had decided to reduce the dimensionality using PCA (Principal Component Analysis). PCA returned a set of 1x 32 set of features. On behalf of this a neural network of 32 input neurons was constructed. This ANN was consisting of one hidden layer. ANN was using a parallel classifier as we were working one out of 10 classes. To improve the results and

TABLE I.   SEGMENTATION RESULTS OF DIFFERENT NUMBER OF CLUSTERS

| Color (n)	Correctly segmented		Incorrectly segmented	
	Images	*%age*	*Images*	*%age*
1	0	0	64	100
2	23	35.9375	36	56.25
3	30	46.875	31	48.4375
4	42	65.625	17	26.5625
5	58	90.625	1	1.562
6	31	48.4375	21	32.8125
7	20	31.25	35	54.6875
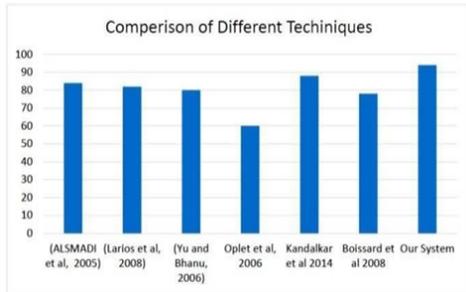
Fig. 4.   Comparison of different techniques.

validate them we have used a 10 cross fold validation method.

If we conclude the whole process in a broader way than we can say that this whole methodology depends upon the three major steps: 1) preprocessing: 2) detection; and 3) classification which is shown in Fig. 1. While each Sub category contain more discrete steps for a specific operation. Like preprocessing include the loading of plant pest image which will pass through the Gaussian filter and creating the color transformation or color space changing structure applying the cluster technique further which will passes to the segmentation and create the corresponding 360^0 angular filter using training image known as cropping (ROI) technique to detection while classification contain Feature Extraction, Set the tuning parameter to determine input images and recognized the best matches. Normalized of images for different viewing angles known as Feature Reduction, and Classification phase.

We have tested the crafted features with three different Classifiers, KNN (K-Nearest Neighbor), SVM and ANN. Complete results of the system along with comparison of different techniques have been shown in Table IV. It has been compared with the results of different features combinations along with different classifiers. It also conclude that we have reached to highest results with suitable technique, which was the core objective of our work.

IV.   RESULTS AND DISCUSSION

When we have observed the initial detection results it was observed as seen in Table I. It was seen that the detection phase has shown some good results when total number of clusters have been taken as n=5.

For classification for testing the accuracy of system we have selected four classes. Dataset for training and testing our system was taken by [16]. They have applied some different feature extraction and classification techniques and collected the results shown in Table IV. He has discussed the working

TABLE II.   CLASSIFICATION RESULTS

Pest	Accuracy%age		AverageAccuracy%
	1st Fold	*2nd Fold*	
Assassin Bug	79	89.42	84
T. Plant Bug	66.7	75	70.85
Convergent Lady Beetle	75	92	83.5
White Flies	87.5	100	93.75

TABLE III.   BRIEF COMPARISON OF DIFFERENT TECHNIQUES

Reference	Method	ACC%
ALSMADI et al, 2005 [17]	Grey Level Co-occurrence Matrix+ ANN	84
Larios et al, 2008 [12]	Histogram + Appearance Base	82
Oplet et al, 2006[18]	Scale-invariant feature transform	60
Kandalkar et al, 2014[19]	Discrete Wavelet transform(DWT) + ANN	88
Boissard et al, 2008[7]	Knowledge Base Mappin	78
Our System	DCT + ANN	**94**

of different feature extraction techniques with two classifiers. The results for the classification are as given below. If we make a comparison of different Feature and classifying techniques with our methodology we can see that it has improved results. Result has been collected very precisely, and shown in Table II. We have shown a brief comparison of the Results with some existing benchmark techniques in Table III.

Setup used by [16] has been extended to check the accuracy of the said database with ANN (Artificial Neural Network). It shows a remarkable results on each class of he given dataset, which shows the robustness of the system. This extension contains one more classifier and a features technique. Results of each combination have been shown in Table IV, and graphically illustrated in Fig. 4.

TABLE IV.   COMPARISON OF RESULTS WITH DIFFERENT CLASSIFIERS AND FEATURE TECHNIQUES [16]

Features	Classifier		
	KNN	SVM	ANN
Scale-Invariant Feature Transform (SIFT)	46.5	63.5	70.1
Speeded-Up Robust Features (SURF)	53.0	72.5	60.3
Histograms Of Oriented Gradient (HOG)	73.5	81.0	82
HOG + SIFT	75.0	84.0	85.0
HOG+ SURF	76.5	89.5	90.2
DCT + PCA Our	**92.3**	**90.5**	**94**

V.   CONCLUSION

The result shown above in this paper were impressive and improved than previous techniques. The proposed technique is based on image-processing and k-mean clustering methodology. When a pest will be recognized by using four main phases that which pest is this, after we can classify we can recognize it as a friend or enemy pest. We were very much accurate that we have achieved our results, but still this work could be extended in future, which will be more direction oriented to reduce the pesticide usage. Still by introducing the Deep learning techniques we can improve our results. In future this method can be used to develop an automated spraying system which will spray only those pests which need a spray else it will skip it.

ACKNOWLEDGMENT

REFERENCES

[1] Z. Tiezhong, C. Libing, and S. Jian, "Study on strawberry harvesting robot: ?. images based identifications of strawberry barycenter and plucking position [j]," *Journal of China Agricultural University*, vol. 1, p. 013, 2005.

[2] X. Zheng, J. Zhao, and M. Liu, "Tomatoes recognition and location from nature background based on binocular stereo vision," *Computer Engineering*, vol. 30, no. 22, pp. 171–173, 2004.

[3] J. Patil and V. Mytri, "A prediction model for population dynamics of cotton pest (thrips tabaci linde) using multilayer-perceptron neural network," *International Journal of Computer Applications*, vol. 67, no. 4, 2013.

[4] L.-Q. Zhu and Z. Zhang, "Auto-classification of insect images based on color histogram and glcm," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on*, vol. 6. IEEE, 2010, pp. 2589–2593.

[5] Y. Li, C. Xia, and J. Lee, "Vision-based pest detection and automatic spray of greenhouse plant," in *Industrial Electronics, 2009. ISIE 2009. IEEE International Symposium on*. IEEE, 2009, pp. 920–925.

[6] A. Camargo and J. Smith, "An image-processing based algorithm to automatically identify plant disease visual symptoms," *Biosystems engineering*, vol. 102, no. 1, pp. 9–21, 2009.

[7] P. Boissard, V. Martin, and S. Moisan, "A cognitive vision approach to early pest detection in greenhouse crops," *computers and electronics in agriculture*, vol. 62, no. 2, pp. 81–93, 2008.

[8] R. K. Priya, C. Thangaraj, C. Kesavadas, and V.-C. Presently, "Fuzzy c-means method for colour image segmentation with l* u* v* colour transformation," *IJCSI International Journal of Computer Science Issues, Special Issue, ICVCI-2011*, vol. 1, no. 1, pp. 1694–0814, 2011.

[9] J. H. Lai, P. C. Yuen, and G. C. Feng, "Face recognition using holistic fourier invariant features," *Pattern Recognition*, vol. 34, no. 1, pp. 95–109, 2001.

[10] G. Liu, H. Shen, X. Yang, and Y. Ge, "Research on prediction about fruit tree diseases and insect pests based on neural network," *Artificial Intelligence Applications and Innovations*, pp. 731–740, 2005.

[11] M. Ebrahimi, M. Khoshtaghaza, S. Minaei, and B. Jamshidi, "Vision-based pest detection based on svm classification method," *Computers and Electronics in Agriculture*, vol. 137, pp. 52–58, 2017.

[12] N. Larios, H. Deng, W. Zhang, M. Sarpola, J. Yuen, R. Paasch, A. Moldenke, D. A. Lytle, S. R. Correa, E. N. Mortensen *et al.*, "Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects," *Machine Vision and Applications*, vol. 19, no. 2, pp. 105–123, 2008.

[13] A. Phinyomark, C. Limsakul, and P. Phukpattaranont, "A novel feature extraction for robust emg pattern recognition," *arXiv preprint arXiv:0912.3973*, 2009.

[14] A. Clément, T. Verfaille, C. Lormel, and B. Jaloux, "A new colour vision system to quantify automatically foliar discolouration caused by insect pests feeding on leaf cells," *Biosystems Engineering*, vol. 133, pp. 128–140, 2015.

[15] Q. Yao, Q. Liu, T. G. Dietterich, S. Todorovic, J. Lin, G. Diao, B. Yang, and J. Tang, "Segmentation of touching insects based on optical flow and ncuts," *biosystems engineering*, vol. 114, no. 2, pp. 67–77, 2013.

[16] K. Venugoban and A. Ramanan, "Image classification of paddy field insect pests using gradient-based features," *International Journal of Machine Learning and Computing*, vol. 4, no. 1, p. 1, 2014.

[17] M. K. Alsmadi, K. B. Omar, S. A. Noah, and I. Almarashdeh, "Fish recognition based on robust features extraction from color texture measurements using back-propagation classifier," *Journal of Theoretical and Applied Information Technology*, vol. 18, no. 1, pp. 11–18, 2010.

[18] A. Opelt, A. Pinz, M. Fussenegger, and P. Auer, "Generic object recognition with boosting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 3, pp. 416–431, 2006.

[19] G. Kandalkar, A. Deorankar, and P. Chatur, "Classification of agricultural pests using dwt and back propagation neural networks."