JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Insects Classification in Agriculture Field using Deep learning Technologies

Virender Kumar
Electronics and Communication Engineering Department,
Chandigarh University, Gharuan, Punjab, India

Abstract— ***The Indian economy heavily depends on agriculture, crucial for addressing food shortages and ensuring nutritious food supply. However, farmers currently encounter challenges in identifying crop-damaging pests and selecting appropriate pesticides, often leading to the use of ineffective chemicals that reduce yields and cause financial losses. Traditional pest identification relies on skilled taxonomists assessing physical characteristics, posing limitations. To address this issue, experiments utilized shape features and machine learning methods such as support vector machines (SVM), k-nearest neighbors (KNN), naive Bayes (NB), and convolutional neural networks (CNN) to classify 15 insect classes in the IP102 dataset.***

I . INTRODUCTION

Agriculture plays an important role in the Indian economy. Nowadays, farmers are facing a severe crisisin yield of crops. Crop production is decreased owing to the infection on crops caused by pests. The occurrenceof pests on fields has certain negative effects on crops, it also affects the quantity and quality of agricultural products. Pests have caused a severe threat to the growthof crops thus causing economic loss to the farmers. Every year the causes due to pests and their impacts is considerably increasing. If the pests on crops are not identified in time, it results in food insecurity. Mostly thepests are detected manually, but the manual methods won't help in meeting the actual needs. The crops shouldbe monitored continuously and labor intensive whichcauses high labor costs and it also pushes the farmer intopoverty. In January 2022, a ground survey on insect infestation that destroyed the chilli crops was taken by the Human Rights Forum and Rythu Swarajya Vedika was claimed. This is because of an invasive pest from Indonesia which was rapidly spreading in Telangana andAndhra pradesh. Due to this 20 farmers committed suicide. The pest was identified as Thrips parvispinus byHRF. The loss of precious lives of farmers and the miseries of their families would have been avoided if thepests had been identified in earlier stages so that causes would have been controlled. According to a report by University of California(UC) Agriculture and Natural Resources, scientists and other members of the International Society for Plant Pathology, pathogens and pests are reducing crop yields for five major food crops by 10% to 40%. In India, the crop yield has decreased on an average of 15% due to the effects caused by pests on crops. Countries under development have faced a lot of problems due to scarcityof food grains which leads to increase in poverty and fallin the GDP of the country. Economic crisis can be overcomed but death can't be brought back to life. Various solutions for this global problem have been sought for in the past decade by various researchers and scientists. Detecting the pest in the earlier stage will helpin reducing the said consequences. In order to do so a YOLO based detection and classification of insects. YOLO is an algorithm that uses neural networks to provide real-time object detection. This algorithm is popular because of its speed and accuracy. Object detection is a phenomenon in computer vision thatinvolves the detection of various objects in digital images or videos. Some of the objects detected include people, cars, chairs, stones, buildings, and animals and are used to detect insects. YOLO algorithm works using the following three techniques namely residual blocks, Bounding box, regression, intersection Over Union (IOU). Detection and Classification of insects has been performed on IP102 dataset. The dataset actually has 102classes of insect species and contains more than 75000 images .This paper focuses on detecting and classifying 15 most frequently affecting pests.

II. RELATED WORKS

Insect classification and detection in field crops using modern machine learning techniques by Thenmozhi Kasinathan, Dakshayani Singaraju, Srinivasulu Reddy Uyyala (2021) has made a comparative study on various machine learning techniques such as ANN, SVM, KNN, NB and CNN. They have used the Grabcut algorithm for the detection of insects. Fruits and vegetables quality evaluation using computer vision: A review by Anuja Bhargava and Atul Bansal (2017) proposed a paper for inspecting the quality of fruits and vegetables using computer vision. They have implemented deep residual learning for identifying the pests and achieved an accuracy of 98.67%. Detection of stored-grain insects using deep learning by Yufeng Shen, Huiling Zhou, Jiangtao Li, Fuji Jian and Digvir S. Jayas (2018) have applied the object detection algorithm, which was based on Faster R-CNN, to detect stored-grain insects under field conditions with impurities. This could detect the insects with slight adherence. An improved inception network was also developed to enhance the accuracy of small insect detection through the deep convolutional neural networks.The improved inception model with SVD operation trained on images of 600 × 800 pixels resolution with augmentation was selected as the final model. Research on insect pest image detection and recognition based on bio-inspired methods by Limiao Deng, Yanjiang Wang, Zhongzhi Han and Renshi Yu (2018) made their work based on human visual systems. The features were extracted using the SIFT-HMAX model and Local Configuration Pattern (LCP) algorithm and fed to SVM for performing classification. Pest Detection on Leaf using Image Processing by Harshita Nagar and R.S. Sharma (2021) has applied a Wavelet transformation algorithm for extracting the area of interest. Then by using FDWT, the detailed image is constructed and the input is given to the ORG (Oriented FAST and rotated BRIEF) algorithm. Then classification is done using the Dynamic Time Warping Algorithm. Pest identification via deep residual learning in complex background by Xi Cheng, Youhua Zhang, Yiqiong Chen, Yunzhi Wu, Yi Yue (2017) implemented CNN for classification of insects under complex background by using deep residual learning and have achieved an accuracy of 98.67% for 10 classes of insects. Precise Agriculture: Effective Deep Learning Strategies to Detect Pest Insects by Luca Butera, Alberto Ferrante, Mauro Jermini, Mauro Prevostini Cesare Alippi (2022) implements FasterRCNN, SSD, and RetinaNet models for object detection and VGG, ResNet, DenseNet, MobileNet and Feature Pyramid Network for objectclassification. They have used Stochastic Gradient Descent and Adam optimizer as optimization algorithms. Multi-level learning features for automatic classification of field crop pests by Chengjun Xi , Rujing Wang , Jie Zhang, Peng Chen, Wei Dong, Rui Li, Tianjiao Chen andHongbo Chen proposed an unsupervised method with the use of multilevel learning features for the automatic classification for classification of insects. The boosting method is applied by having SVMs with different kernels as weak classifiers, which can be formed as a strong classifier. Automated Pest Detection With DNN on the Edge for Precision Agriculture by Andrea Albanese , Matteo Nardello, and Davide Brunelli (2021) implemented deep neural network architectures such as LeNet-5, VGG16 and MobileNetV2. Automatic detection of insect predation through the segmentation of damaged leaves by Gabrielda Silva, Vieira Bruno Moraes Rocha, Afonso Ueslei Fonseca, Naiane Mariade Sousa, Julio Cesar Ferreira, Christian Dias Cabacinha, Fabrizzio Soares (2022) proposed an automatic method for detecting injured leaf areas caused by insect predation. They have implemented the K-fold cross-validation method.

III. PROPOSED WORK

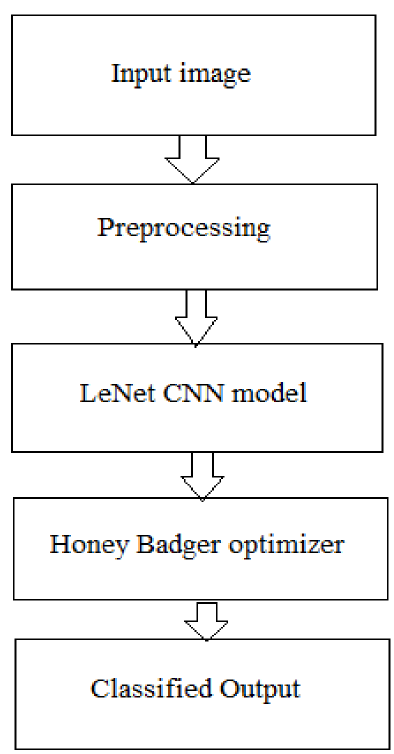


Fig. 1. Framework for insect classification and detection

TABLE 1 - INSECT DETAILS USED FROM IP102 DATASET

Insect class	Number of insects
Mole cricket	100
Potosi Brevitarsis	100
Flea beetle	100
Oides decempunctata	100
Xylotrechus	100
Icerya purchasi Maskell	100
Salurnis marginella Guerra	100
Cicadellidae	100

Insect dataset for detection and classification

Insect classification and insect detection were performed for IP102 dataset for different field crops. 15 insects were taken into consideration to detect and classify pests.

Image pre-processing

The insect photos were resized to 150* 150 pixels in size. Techniques for enhancing image data, such as rotation, flipping, and cropping operators, are used to expand the training set in order to improve accuracy and get rid of overtraining issues.

Feature extraction

Shape features are crucial characteristics that are used in computer vision and automatic object recognition systems since they are unaffected by scaling, rotation, and translation. Based on the finite shape features that were derived from the bug photos, classification of insects is done.

The RGB insect photos were changed to grayscale images for additional feature extraction. The classifier model is then given the feature extracted by LDA technique.

Linear Discriminant Analysis:

One of the most widely utilized dimensionality reduction methods for supervised classification issues in machine learning is linear discriminant analysis. Itis also regarded as a pre-processing stage for applications of pattern categorization and modelingvariations in Machine Learning.
An automatic dimensionality reduction procedure can be carried out using the multi-class predictive modeling technique linear discriminant analysis. Reducing the number of columns or input variables in a data model is referred to as dimensionality reduction. In situations involving pattern categorization, LDA is most frequently employed for feature extraction. The LDA approach has the benefit of requiring no prior knowledge of the topics to be covered.

Insect classification was performed using several classifiers such as ANN, SVM, KNN, CNN and NB.

Naive Bayes :

This is based on the Bayes Theorem of Probability and presupposes that each predictor is independent of the others.The probability is calculated by the formulastated below.

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

This classifier only operates under the presumption that the existence of one feature does not indicate the existence of any other feature in the insect class. The likelihood that a given tuple from the soybean cropbug dataset belongs to a specific insect class was predicted using naive Bayesian classification. This has yielded an accuracy score of 32.4%.

TABLE - 2 CLASSIFICATION REPORT

	precision	recall	f1 – score	support
2	0.10	0.32	0.15	79
6	0.40	0.15	0.22	111
9	0.12	0.01	0.01	166
18	0.56	0.09	0.15	257
22	0.24	0.06	0.10	170
28	0.11	0.02	0.03	98
33	0.10	0.42	0.17	88
34	0.30	0.08	0.12	102
40	0.33	0.04	0.08	90
41	0.10	0.42	0.16	81
42	0.14	0.02	0.04	93
52	0.00	0.00	0.00	78
62	0.10	0.51	0.17	85
78	0.15	0.06	0.09	78
101	0.08	0.05	0.06	150

KNN Classifier :

It is an inefficient learning algorithm that makes no use of parameters. In this case, the neighbors' contributions are given weights so that the closer neighbors contribute more to the average than the farther neighbors.

$$d(x, y) = (\Sigma \sum_{i=1}^{m} |x_i - y_i|)$$

The KNN method performs better when handling classification problems with high dimensions and little samples. Using this classifier we have achieved an accuracy of 24.4%.

SVM Classifier :

This algorithm uses supervised machine learning. The quantity of features that we have determines how many dimensions this space has. The data items are represented as a point in a k-dimensional space with "k" features, where the value of a given coordinate corresponds to a particular feature value. Identificationof the hyper-plane, which clearly distinguishes the classes, is used to classify insects. In order to reduce the total interspecific error rates for classifying mosquitoes and fruit flies, SVM classifiers were implemented with 9 fold cross validation. The performance of the classification models for a certain set of test data is evaluated using a matrix called the confusion matrix. Only after the true values of the test data are known can it be determined. Although the matrix itself is simple to understand, some of the terminology used in connection with it might be.

```
[[ 17   6   0   3   5   4   2   4   4   1   3   5   1   0  24]
 [  1  60   1   1   2   4   1   1  10   1   6   1   2   2  18]
 [  2   6  53   0   5  13   6   6   4   2   6   4   6   2  51]
 [  5  40   4  57   6   1   6  10   3   4   8   6   2  13  92]
 [  6  19  12   2  25   4  11  12  11   8   7   6   1   8  38]
 [  1  11   5   1   2  17  11   6   3   0   1   3   1   4  32]
 [  0   3   1   0   0   4  36   2   3   0   2   5   2   0  30]
 [  1   4   9   0   3   2   1  39   3   2   2   5   2   0  29]
 [  3   6   1   4   1   1   4   7  36   0   1   0   4   1  21]
 [  1   3   2   2   4   2   4   1   1  37   0   4   0   3  17]
 [  4   9   5   1   2   3   4   2   2   2  16   2   1   1  39]
 [  4   6   4   3   8   2   8   4   2   2   2   4   1   1  27]
 [  5   5   7   2   3   3   3   2   0   0   1   2  25   1  26]
 [  2   9   4   2   3   0   4   6   2   1   2   4   0  22  17]
 [  1   9   8   1   3   4   4   6   1   0   5   2   2   3 101]]
```

Fig. 3. Confusion matrix

It is also referred to as an error matrix since it displays the errors in the model performance as a matrix.

$$classification\ accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

The true positive, false positive, false negative, and true negative are denoted by TP, FP, FN, and TN. If the bug in the image is accurately identified, it istreated as TP; if not, it is treated as FN. If the classification is done incorrectly, the insect that is not visible in the image is referred to as TN; otherwise, it is referred to as FP.

Rose Chafer *Blue-Striped Leafhopper*

Salurnis marginella *Cyrtoclytus capra*

Fig. 3. Sample images from IP102 dataset

ANN Classifier

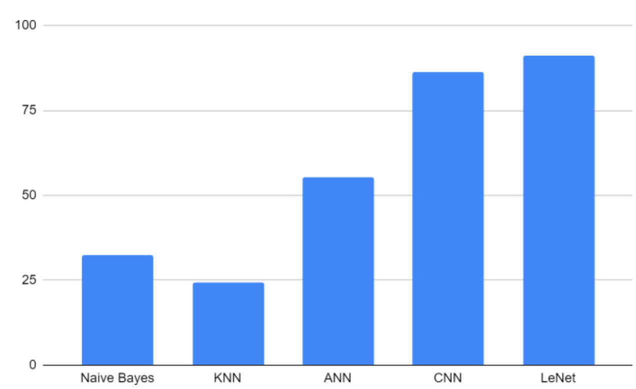
Artificial Neural Networks are brain-inspired algorithms that are used to forecast issues and model complex patterns. The internal structure of an artificial neural network (ANN) can alter depending on the information passing through it. The idea of biological neural networks in the human brain gave rise to the Artificial Neural Network , a deep learning technique.Using this classifier we have achieved an accuracy of 55.4%.

CNN Classifier

A CNN is a particular type of network design for deep learning algorithms that is utilized for tasks like image recognition and pixel data processing. Althoughthere are different kinds of neural networks in deep learning, CNNs are the preferred network architecture for identifying and recognizing objects.With no loss ofinformation, its integrated convolutional layer lowers the high dimensionality of images. Here we have achieved an accuracy of 86.2% using this classifier.

LeNet 5 architecture

Convolutional neural networks are a type of feed-forward neural network that excels at processing large-scale images because their artificial neurons mayrespond to a portion of the surrounding cells in the coverage range. The LeNet-5 CNN architecture has seven layers. Three convolutional layers, two subsampling layers, and two fully linked layers make up the layer composition.Here , we have achieved an accuracy of 91.01%.



```
Set parameters t_max, N, β, C.
Initialize population with random positions.
Evaluate the fitness of each honey badger position x; using objective function and assign to f_i, i ∈ [1,2,..., N].
Save best position X_prey and assign fitness to f_prey.
while t ≤ t_max do
    Update the decreasing factor α using α = C × exp (−t/t_max), t_max = maximum number of iterations.
    for i = 1 to N do
        Calculate the intensity I; using I_i = r2 × S/4πd_i^2, r2 is a random number between 0 and 1.
        if r < 0.5 then
            Update the position X_new using x_new = x_prey + F × β × I × x_prey + F × r3 × α × d_i ×
            |cos(2πr4) × [1 − cos(2πr5)] |
        else
            Update the position X_new using x_new = x_prey + F × r7 × α × d_i, r7 is a random number between 0 and 1.
        end if
        Evaluate new position and assign to fnew:
        if f_new ≤ f_i then
            Set x_i = X_new and f_i = f_new
        end if
        if f_new ≤ f prey then
            Set X_prey = X_new and f_prey = f_new
        end if
    end for
end while Stop criteria satisfied.
Return X_prey
```

Fig. 5. Honey Badger Algorithm

Grabcut detection:

GrabCut is a graph-cut-based technique for segmenting images. The algorithm estimates the colour distribution of the target object and the background using a Gaussian mixture model, starting with a user-specified bounding box around the object to be segmented. The color image is read by OpenCV in BGR (Blue, Green, and Red) format order. An array of zeros with the same size as the input insect image is used to construct the mask image with the name "mask." Zero-filled arrays are used to build both the foreground and background array models. When the foreground insect is separated from the background, these two arrays are used internally. The bounding box's coordinates are specified. It includes the area of the input image where the insect should be segmented. By applying a mask image "mask" to the input image, the GrabCut algorithm is modified to separate the foreground insect from the coloured background image. The algorithm executes for five iterations while the bounding box initialization mode is chosen. All of the definite and probable background pixels are set to zeroin the output mask image 'mask output', while all ofthe definite and probable foreground pixels are set to one in the mask image 'mask'. To create the segmentedinsect picture, "image seg," the BGR channels of the input image were multiplied by the output mask image, "mask output." Another mask called "background img" is produced with an array made up entirely of zeros, matching the size of the input image.The 'background img' mask's pixels are then all converted to white. The 'image segmented' image is created by combining the 'background img' mask with the 'image seg'. The segmented image "image segmented" is further processed by using the Gaussian blur to eliminate the Gaussian noise and converting it to the HSV color model. Histogram equalization increases the image's contrast. This study adapts the inverted binary thresholding. The contours in the binary image are found using OpenCV's findContours function. The contours are located, the largest counter is chosen, andthe remaining pointless contours are removed. Therotating rectangle containing the bug in the input image is drawn using the minAreaRect function's computed minimum-area bounding rectangle for the largest contour and its coordinates. Fig. 4. is Grabcut insect foreground mask.

```
Input: Np    : population size;
       Dim  : problem dimension;
       Kmax: The upper limit of the number of iterations.
Output: Global optimal individual position in the population;
        Fitness value.
1.Initialize the position of each Gannet in the population according to
  xid = r0(ubd − lbd) + lbd, i = 1, 2, . . ., N, d = 1, 2, . . ., Dim
2.Generate a position matrix MXi based on each initialized Gannet position and calculate
  the fitness value for each Gannet
3. for Itk< Kmax do
4.    if rand ≥ 0.5 then
5.      for MXi do
6.        if rand  ≥ 0.5 then
7.            MXi(t + 1) = u1 + u2 + Xi(t)
8.        else
9.            MXi(t + 1) =  v1 + v2 + Xi(t)
11.       end if
12.     end for
13.  else
14.     for MXi do
15.       if c ≥ 0.2 then
16.            MXi (It + 1) = Xi (It) + t × Delt × (Xi (It) – Xbest (It))
17.         else
18.             MXi (It + 1) = Xbest (It) − (Xi (It) – Xbest (It)) × t × Lv
19.         end if
20.      end for
21.   end if
22.   for MXi do
23.      Calculate fitness value of each Gannet Xi in MXi;
24.      Update MXi based on Xi fitness;
25.   end for
26.end for
```

Fig. 6. Gannet Optimization Algorithm

4. Optimization Techniques

4.1. Honey Badger Algorithm

The recently developed optimization method known as the Honey Badger Algorithm (HBA) was influenced by the clever foraging of honey badgers. The exploration and exploitation phases of HBA express the honey badger's digging-based search techniques. A new method and a variety of tools are developed by the new optimizer to balance exploitation and exploration. It is simple to understandand has minimal control settings. Fig. 5. explains Honey Badger Algorithm.

4.2 Gannet Optimization Algorithm

Exploration and exploitation are made possible by theGOA, which mathematically represents the different distinctive actions of gannets when foraging. The ideal region within the search space is explored using GOA's U- and V-shaped dive patterns, with abrupt turns and random walks ensuring better answers are identified there.In high dimensions, GOA runs faster and can offer a more effective solution.

5. CONCLUSION

The outcomes of classifying and detecting various insects using IP102 dataset using yolo and inception based model was illustrated in this research. To expand the dataset and increase the accuracy, all of theinsect photos were rescaled, pre-processed, and enhanced. In the real-time field, it can be difficult to classify insects with greater precision because of shadows, leaves, mud, branches, flower buds, etc. in the main agricultural field crops. Naive bayes algorithm produced the accuracy of 14.23 and KNN produced the accuracy of 37%.

References

[1] Kasinathan, T., Singaraju, D. and Uyyala, S.R., 2021. Insect classification and detection in fieldcrops using modern machine learning techniques. Information Processing in Agriculture, 8(3), pp.446-457.
[2] Bhargava, A. and Bansal, A., 2021. Fruits and vegetables quality evaluation using computer vision: A review. Journal of King Saud University-Computer and Information Sciences, 33(3), pp.243-257.
[3] Shen, Y., Zhou, H., Li, J., Jian, F. andJayas,D.S., 2018. Detection of stored-grain insects using deep learning. Computers and Electronics in Agriculture, 145, pp.319-325.
[4] Deng, L., Wang, Y., Han, Z. and Yu, R., 2018. Research on insect pest image detection and recognition based on bio-inspired methods. Biosystems Engineering, 169, pp.139-148.
[5] Nagar, H. and Sharma, R.S., 2021, January. Pest Detection on Leaf using Image Processing. In 2021 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-5). IEEE.
[6] Cheng, X., Zhang, Y., Chen, Y., Wu, Y. and Yue, Y., 2017. Pest identification viadeepresiduallearning in complexbackground. ComputersandElectronicsin Agriculture, 141, pp.351-356.
[7] Butera, L., Ferrante, A., Jermini, M., Prevostini, M. and Alippi, C., 2021. Precise Agriculture: Effective Deep Learning Strategies to Detect Pest Insects. IEEE/CAA Journal of Automatica Sinica, 9(2), pp.246-258.
[8] Xie, C., Wang, R., Zhang, J., Chen, P., Dong, W., Li, R., Chen, T. and Chen, H., 2018. Multi-level learning features for automatic classification of field crop pests. Computers and Electronics in Agriculture, 152, pp.233-241.
[9] Albanese, A., Nardello, M.andBrunelli,D.,2021. Automated pest detection with DNN on the edge for precision agriculture. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 11(3), pp.458-467.
[10] da Silva Vieira, G., Rocha, B.M., Fonseca, A.U., de Sousa, N.M., Ferreira, J.C., Cabacinha, C.D. and Soares, F., 2022. Automatic detection of insect predation through the segmentation of damaged leaves. Smart Agricultural Technology.