Abstract: Agriculture provides food for human existence, where insects damage the crops. The identification of the insect is a difficult process and subjected to expert opinion. In recent years, researches using deep learning in fields of object detection have been widespread and show accuracy as a result. This study shows the comparison of three widely used deep learning meta-architectures (Faster R-CNN, SSD Inception and SSD Mobilenet) as object detection for selected flying insects namely Phyllophaga spp., Helicoverpa armigera and Spodoptera litura. The proposed study is focused on accuracy performance of selected meta-architectures using small dataset of insects. The meta-architecture was tested with same environment for all three architectures and Faster RCNN meta-architecture performs outstanding with an accuracy of 95.33%.

Keywords: CNN, Deep Learning, Insect Detection, Object Detection, Pest Classification, TensorFlow.

I. INTRODUCTION

Insect detection cause significant reduction in both quality and quantity of crops in the field of Agriculture [1] and hence it become very crucial point for agriculture commodity. There are several species of insects in agricultural field, which requires a lot of time for manual classification by insect experts and different species of insects might have similar phenotypes except complicated phenotypes due to different environments and growth periods of insect. Currently, Insect detection process is carried out by manually or by using different trapping tools. One of the most used approaches is bare eye inspection, but for this method needs the specialist entomologist which is expensive [2]. Many times availability of expert and their services is time-consuming. Hence it is contemporary need to develop more rapid, precise and effective remedial approaches to grab this problem. Deep Convolutional Neural Network (CNN) is become widespread and alternate tool for better execution and outperforms than other mainstream computer vision techniques for object detection and identification [3]. Hence, in present study, the widespread CNN meta-architecture was used to identify the different species of insects viz. Phyllophaga spp., Helicoverpa armigera and Spodoptera litura insects with aimed to detect and compare the different meta-architectures of deep learning for identification of insect.

II. RELATED WORK

In the field of agriculture, object detection is raising interest for research work. We have studied similar work like moth detection from trap [4], tomato plant diseases and insects’ recognition [5], fruit detection [6], automatic insect detection on bean and potato crops [7] and insect image detection and recognition based on bio-inspired methods [8]. We found that deep learning object detection and identification technique offers better performance than other popular image processing techniques [9]. Various deep neural network architectures have been proposed in past [10], however, our work aimed to use the latest architectures and measure accuracy performance of the networks using small dataset of insects. CNN meta-architectures, namely-Faster-RCNN, SSD Inception and SSD Mobilenet which is obtain very high detection accuracy [11] used in this study.

III. MATERIALS AND METHODS

There are several different class of insects found in major crops of Saurashtra region of Gujarat, India. Out of them, the three main insects namely Phyllophaga spp., Helicoverpa armigera and Spodoptera litura were selected in the present study. All the three insects were collected from agriculture field nearby area of Junagadh district, Gujarat, India and were subjected to capture the training input images of insects in a controlled environment by using digital mobile camera. The few of captured images are presented in Fig. 1 and the link for whole dataset is provided in Appendix.

Each captured training input image contained individual insect or in combination of all two to three insect class and were given identical code to each class of insect which is as per mentioned in Table- I.

Table- I: Identical code for Insects which are used during the experiment

<table>
<thead>
<tr>
<th>Sr.</th>
<th>Insects</th>
<th>Identical code</th>
<th>No. of Insects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Phyllophaga Spp.</td>
<td>INS-A</td>
<td>319</td>
</tr>
<tr>
<td>2</td>
<td>Spodoptera litura</td>
<td>INS-B</td>
<td>131</td>
</tr>
<tr>
<td>3</td>
<td>Helicoverpa armigera</td>
<td>INS-C</td>
<td>143</td>
</tr>
</tbody>
</table>
Insect Identification Among Deep Learning’s Meta-architectures Using TensorFlow

Fig. 1. Few of captured images of flying insects (Phyllophaga spp, Helicoverpa armigera and Spodoptera litura)

All the captured images were annotated the label by using Labeling tool and generated CSV file (Table- II). During the study, the libraries of TensorFlow and Keras were utilized to perform insect detection and identification.

Table- II: Few data of the generated CSV File which were used during generating TFRecords

<table>
<thead>
<tr>
<th>Filename</th>
<th>Width</th>
<th>Height</th>
<th>Class</th>
<th>X-min</th>
<th>Y-min</th>
<th>X-max</th>
<th>Y-max</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMG</td>
<td>4160</td>
<td>3120</td>
<td>INS-A</td>
<td>892</td>
<td>797</td>
<td>1178</td>
<td>1334</td>
</tr>
<tr>
<td>105659.jpg</td>
<td>4160</td>
<td>3120</td>
<td>INS-A</td>
<td>1368</td>
<td>707</td>
<td>1815</td>
<td>1112</td>
</tr>
<tr>
<td>105659.jpg</td>
<td>4160</td>
<td>3120</td>
<td>INS-B</td>
<td>2192</td>
<td>1645</td>
<td>2401</td>
<td>2150</td>
</tr>
<tr>
<td>162650.jpg</td>
<td>720</td>
<td>480</td>
<td>INS-A</td>
<td>311</td>
<td>201</td>
<td>403</td>
<td>300</td>
</tr>
<tr>
<td>163319.jpg</td>
<td>720</td>
<td>480</td>
<td>INS-A</td>
<td>304</td>
<td>166</td>
<td>415</td>
<td>283</td>
</tr>
<tr>
<td>163702.jpg</td>
<td>720</td>
<td>480</td>
<td>INS-A</td>
<td>269</td>
<td>116</td>
<td>423</td>
<td>272</td>
</tr>
<tr>
<td>120445.jpg</td>
<td>1920</td>
<td>1080</td>
<td>INS-A</td>
<td>1153</td>
<td>990</td>
<td>1246</td>
<td>1080</td>
</tr>
<tr>
<td>120450.jpg</td>
<td>1920</td>
<td>1080</td>
<td>INS-A</td>
<td>162</td>
<td>586</td>
<td>326</td>
<td>693</td>
</tr>
<tr>
<td>120450.jpg</td>
<td>1920</td>
<td>1080</td>
<td>INS-A</td>
<td>835</td>
<td>205</td>
<td>1010</td>
<td>388</td>
</tr>
<tr>
<td>120450.jpg</td>
<td>1920</td>
<td>1080</td>
<td>INS-A</td>
<td>1251</td>
<td>15</td>
<td>1364</td>
<td>115</td>
</tr>
<tr>
<td>120450.jpg</td>
<td>1920</td>
<td>1080</td>
<td>INS-A</td>
<td>1314</td>
<td>666</td>
<td>1522</td>
<td>866</td>
</tr>
<tr>
<td>120456.jpg</td>
<td>1920</td>
<td>1080</td>
<td>INS-A</td>
<td>481</td>
<td>341</td>
<td>576</td>
<td>418</td>
</tr>
<tr>
<td>120456.jpg</td>
<td>1920</td>
<td>1080</td>
<td>INS-A</td>
<td>856</td>
<td>53</td>
<td>981</td>
<td>170</td>
</tr>
<tr>
<td>120456.jpg</td>
<td>1920</td>
<td>1080</td>
<td>INS-A</td>
<td>553</td>
<td>959</td>
<td>658</td>
<td>1049</td>
</tr>
</tbody>
</table>

*Use Appendix for more detail

The process flow for training the model of selected insect is illustrated in Fig. 2 for each architecture. As per the process flow, all the generated CSV files were further processed to make TFRecord formats by using python code which is simple format for storing a sequence of binary records as protocol buffers. All the TFRecord formats were processed for training meta-architectures namely SSD Inception, SSD Mobilenet and Faster R-CNN. All experiment procedures were carried out on Intel Xeon Gold 6136 Processor, 128GB RAM with Nvidia Quadro P5000 GPU workstation.

A. SSD Inception

Single Shot MultiBox Detector (SSD) is a fast single-shot object detector for multiple categories. The main feature of this model is the use of multi-scale convolutional bounding box for efficiency [12]. All the TFRecords were p have used combination of SSD with inception feature extractor which is high performance vision networks that have a relatively modest computation cost, simpler and monolithic architectures [13].

B. SSD Mobilenet

We have used another combination of SSD with Mobilenet feature extractor which is based on a streamlined architecture that uses depth wise separable convolutions to build light weight deep neural networks. It can be deployed as an effective base network in modern object detection systems [14].

C. Faster R-CNN

Faster Region-based Convolutional Network method (Faster R-CNN) takes as input an entire image and a set of object proposals. The network first processes the whole image with several convolutional networks to produce a feature map. Then, it extracts a fixed-length feature vector from the feature map. This model employs to improve training and testing speed with increasing detection accuracy [15].
To compare the performance of the Deep CNNs architectures for insect classification task, various evaluation metrics such as specificity, sensitivity and accuracy were used in the present study. According to Equations (1)-(4), True Positives (TP) is the numbers of specific insects that predicted as specific insects; False Negatives (FN) is the numbers of insects not predicted as specific insects. True Negatives (TN) is the numbers of non-insects that not predicted as insects, while False Positives (FP) is the numbers of non-insects predicted as insects [16].

Sensitivity is the measure of insect detection that correctly classified and is expressed as Sensitivity or True Positive Rate (TPR).

\[ TPR = \frac{TP}{(TP + FN)} \]  \hspace{1cm} (1)

Specificity is the measure of non-insect that not classified and is expressed as Specificity or True Negative Rate (TNR).

\[ TNR = \frac{TN}{(TN + FP)} \]  \hspace{1cm} (2)

Precision or positive predictive value measures what percentage of correctly classified labels is truly positive and is given as Positive Predictive Value (PPV).

\[ PPV = \frac{TP}{(TP + FP)} \]  \hspace{1cm} (3)

Accuracy (ACC) is used to show the number of correctly classified insect or non-insect divided by the total number of insects and is defined as ACC(%):

\[ ACC(\%) = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100 \]  \hspace{1cm} (4)

IV. RESULTS

ULTS AND DISCUSSION

Insect identification is most important factor to control the damage earlier in crop and the manually identification of insect without help of specialized entomolog is very cumbersome process as well as there are lot of chances of error. With advancement of deep learning and computer vision, it is become possible to detect the insect at earlier. The present study focused to compare the all well-known three meta-architecture of deep learning as well as to find out the test accuracy performance of the networks using small dataset of 128 source insect images. Loss functions serve as ways to measure the distance or difference between a model’s predicted output and the ground truth of trained model. The result of loss function with lower optimal value of training images stabilized at 20k, 60k and 20k epoch iterations respectively, for SSD Inception, SSD MobileNet and Faster R-CNN architectures. As per the training result aspect, the Faster R-CNN looks better performance than other architectures which is depicted in Fig. 3.

As a previous work on automatic insect detection, the methods introduced in this paper have many possible future extensions besides those have been mentioned based on accuracy result. Faster R-CNN performed best as result of real-time tomato plant diseases and pest recognition study [5]. The Random Subspace Classifier (RSC) and Limited Receptive Area (LIRA) recognition rate of 89% and 88% respectively during the study of automatic pest detection on bean and potato crops [7]. There was 75.46% mean average precision (mAP) recorded during similar work of PestNet with large dataset of 80k images with over 580k pests categorized in 16 classes [17].

![Fig. 3. Total Loss Function during transfer learning of (A) SSD Inception, (B) SSD Mobilenet, (C) Faster R-CNN](image-url)
Our study was based on accuracy performance of used meta-architecture of deep learning with small dataset of insects. All above parameters state that the Faster-RCNN approach performs better in all aspects of pest detection. The results also show that SSD Mobilenet meta-architecture perform very well with small dataset of insects.

Fig. 4. Detection result of trained model (A) SSD Mobilenet (B) SSD Inception and (C) Faster R-CNN
V. CONCLUSION

Our experiments show that state of the art CNN meta-architecture Faster-RCNN is better performer with identification accuracy as high as 95.33% and sensitivity equal to 91.06% for selected pest detection operation. We found that SSD Mobilenet also obtained high detection accuracy. Even though we used small sample size, we got good accuracy with Faster-RCNN and SSD Mobilenet architecture in detecting selected pest. To validate our findings, future studies can be done with more classes.

APPENDIX

During the experiment, used training dataset and generated result sets are available on https://github.com/deven1983/pestdetection/tree/master/work.

ACKNOWLEDGMENT

We gratefully acknowledge the support and generosity of Department of Entomology, Junagadh Agricultural University for providing insect data, without which the present study could not have been completed.

REFERENCES