An Entire-and-Partial Feature Transfer Learning Approach for Pest Occurrence Frequency

Chen Yuh-Shyan, Lo Chia-Ling and Hung Hsiang-Ching

July 28, 2019

An Entire-and-Partial Feature Transfer Learning Approach for Pest Occurrence Frequency

Yuh-Shyan Chen, Chia-Ling Lo, and Hsiang-Ching Hung
Department of Computer Science and Information Engineering, National Taipei University
New Taipei City, 23741, Taiwan, R.O.C.
E-mail: yschen@gmail.ntpu.edu.tw

Abstract— **The frequency of pest occurrence has always been a task of agricultural time and labor. This paper attempts to solve the above problems through the combination of deep learning and agriculture. We propose an entire-and-partial feature transfer learning architecture to perform pest detection, classification and counting tasks that have the final goal of presenting pest occurrence frequency. In the partial-feature transfer learning, different fine-grained feature map are strengthened to the entire-feature transfer learning use weight scheme. Finally, different fine-grained feature map are strengthened to the entire-feature transfer learning use weight scheme and the cross-layer of the entire-feature network is combined with multi-scale feature map. The entire-feature transfer learning approach enhances the feature map close to the input to the network layer near the output layer, creating a shortcut topology for the input and output layers to reduce the gradient disappearance problem common to deep networks. The experiments result shows that the detection accuracy can be significantly improved and the proposed method can reach 90.2%.**

Index Terms: partial-feature transfer learning, cross-layer, frequency of pest occurrences, entire-and partial feature, multi-task learning.

I. Introduction

Agriculture is indispensable for human life, so increasing crop yield is important. However, crop yields reduce production due to pests, which is slow down the growth of agriculture. The method of pest counting used by traditional farmers cannot achieve immediate results. However, if the information on the number of pests cannot be obtained in time, it is impossible to take immediate measures. Therefore, the relevant information of pests is quite important. For example, the density of pest populations and the number of individual pest species have been quite time consuming and laborious in terms of pest density and quantity statistics. Due to the current deep learning and the rise of machine vision, there are many studies that apply these techniques to agriculture and have some excellent results. Vakilian *et al.*[1] used many different classification methods and extracted multiple features of pests, some research used SVM [2] and artificial neural network.. Deep learning requires an image of the pest to be predicted in order to predict the number and class of pests. Some studies have chosen to use pest samples to obtain pest data set, but this method trains models with limited generalization capabilities, so it is usually only applicable to the environment in which the data set was originally acquired. The collection method which considers the cost effectiveness and close to the actual environment. Finally, the collection method chooses to use the sticky traps as the data sample. This data collection method is used because the pest is easily attracted by color and pheromones. When the pest is stuck on the sticky trap, use the camera to get the image of the pest. The traditional machine learning method is mainly to manually extract, the development of deep learning in recent years does not need to be manually performed. The convolutional neural networks can helps to obtain the main features of objects in an image, but the more deep learning samples, the higher the accuracy performance. This method is mainly based on Haar random forest extraction. The proposed a framework for pest detection and recognition for image pest localization and recognition. In order to balance cost and performance. The proposed method embed the coarse-level and the fine-level feature map to the entire-feature network, in order to make full use of the multi-level feature map of the partial-feature network as the intput, respectively, the two feature map of the coarse level and fine level are embed in the entire-feature network according to use weighting scheme. The feature maps of partial-feature network to enhance the entire-feature network. This deep network structure use to enhance pest detection and classification. The main contributions of the proposed approach are summarized as follows: This paper proposes a partial feature transfer learning to generate multi-level feature maps, which has enhanced the problem that the feature message is lost due to features in the pooling layer. An entire-feature transfer learning is proposed to connect the network layers close to the input and output. This solves the problem that deep networks often encounter gradients disappearing. This paper proposed entire-and-partial feature transfer learning, embedding partial-feature maps into the pooling layer of entire-feature transfer learning, combining the feature maps of the two transfer learning. This mechanism develops a deep learning application for the classification and warning system for pest. Deep learning in pest classification and warning systems, using low-power IoT communication Technology (LPWA) Telecom NB-IoT communication module. The aim is to solve the problem of farmers uses labor to identify pests and calculations, so the first research work plans to use deep learning technology, image recognition technology and edge computing technology. This approach improve traditional agricultural pest management to manually set up sticky insect traps and estimate the number of pests which is time consuming and requires a lot of manpower and may also have a high error rate. The proposed mechanism uses pest classification and warning system, it is expected to play the role of automatic monitoring of pests and through the high-speed computing of the edge computing platform and the advantages of the high-wide coverage of low-power IoT communication technology, combined with deep learning, effective pest management, reduce pest risk and ensure crop quality. Finally, the transmission data uses LPWA communication technology to transmit pest classification results to the cloud server and immediately displays changes in the number of pests on the visual interface, enabling farm staff to take immediate measures of pest changes. The remainder of the thesis is organized as follows. Section 2 describes the related work and motivation. Section 3 describes the system model, problem formulation and basic idea of our proposed scheme. Section 4 describes our proposed an entire-and-partial feature transfer learning approach for pest occurrence frequency scheme. Section 5 provides the experiment result, and the conclusion is finally given in Section 6.

II. Related Works

In recent years, the field of deep learning has become a hot and exciting new field of research. Deep learning currently has some practical examples for agricultural applications. Controlling the number of pests on the farm is very important for farmers. Early farmers spent a lot of human resources and time on counting pests. Some results support farmers to accelerate the statistics of various pests on the farm. A number of studied have been carried out [3], [4], [5], [6] to address the management of pest populations in traditional farms, but there are still face a huge challenges in integrating cross-domain knowledge. Yuan Hong *et al*. [7] counting insect used two of method, which are YOLO and SVM classifier. Limiao *et al*. [8] proposed to generate a saliency map using the Natural Statistical Model (SUN) method and detect a region of interest (ROI) in the pest image. Redmon *et al*. [9] proposed a new end-to-end object detection method called YOLO. This method constructs object detection into a regression problem and separates the bounding box and related class probability according to space. Peikari *et al*. [10] proposed the idea of introducing high-density regions in the data space for clustering, which is used to enhance the SVM classifier efficacy. Daniel *et al*. [11] proposed to extract the characteristics of mosquitoes from images and realize a convolutional neural network model for identifying adult mosquitoes. Sangheon *et al*. [12] proposed adds only highly confidence predictive data to the training corpus in each training to make the classifier stronger. Ebrahimi *et al*. [13] used the new image processing technology SVM method, which has a differential kernel function to detect and identify pests that may be found on strawberry plants.

A. Basic Ideas and Challenges

The model uses a complete convolutional neural network and a supervised network to extract features from the image

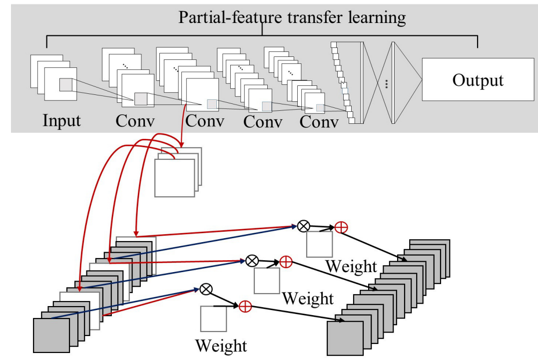

Fig. 1. The basic idea of partial-feature learning.

and then predict the image through a deep learning model. The partial-feature transfer learning structure includes a feature stream deep network having a plurality of m outputs. It generates different levels of image feature maps for each layer. Since the hierarchical output multi-level feature map, the feature map in the extracted background can be changed by changing the hyper-parameter ϕ. Finally, the partial-feature network output feature map P_c and the corresponding weight $\hat{W}$ perform feedforward work to generate a multi-level feature map feed image. The coarse and fine level of layers are selected as feature stream outputs $P_{a,b,c}$ to be merged by the max pooling layer. The deconvolutional layer adjusts the output to the same size and performs an average blending operation. After the feature map is merged, the outputs fed to the next layer. The feature map E_f and E_s is combined with entire-and-partial feature network, which are inherited by the next layer of the network layer and produces feature output.

B. Problem Formulation

In the classification of pests, due to category imbalances, losses caused by frequent categories can control all losses and lead to instability in early training. The approach of pest detection and classification work is included, as well as three objective function loss: pest classification loss, pest location loss and boundary classification loss.

The variable $p_i(c)$ defined as the i-th pest classification loss and the variable defined as $\hat{p_i}(c)$ the i-th ground truth box of classification loss. The prediction error of the target pest, sum of squared prediction probability errors shows in equation 1.

$$\begin{aligned} &arg\ min \left\{ \textstyle\sum_{i=0}^{s^2} I_i^p \ \sum_c \left(p_i(c) - \hat{p_i}(c)\right) \right\} \\ &subject\ to \begin{cases} I_i^p \ = 1, otherwise = 0 \\ \hat{p_i}(c) \epsilon \left\{0,1\right\}, \forall i \\ 0 \leq p_i(c) - \hat{p_i}(c) \leq 1 \end{cases} \end{aligned} \tag{1}$$

The boundary prediction error of the target object represent to equation 2 and the variable B defined as the prediction of bounding box.

$$\begin{aligned}&arg\ min\{\lambda_a \sum_{i=0}^{s^2} \sum_{j=0}^{B} I_{ij}^{p} \left[(x_i-\hat{x}_i)^2\right.\\&+(y_i-\hat{y}_i)^2]+\lambda_a \sum_{i=0}^{s^2} \sum_{j=0}^{B} I_{ij}^{p}[(\sqrt{w_i}-\sqrt{\hat{w}_i})^2\\&+(\sqrt{h_i}-\sqrt{\hat{h}_i})^2]\}\\&\ subject\ to\left\{\ I_i^p\ =1, otherwise=0\right.\end{aligned} \tag{2}$$

The variable C_i defined as the i-th pest of confidence score as shown in equation 3, and $\hat{C}_i$ defined as the i-th ground truth box of the pest. The pest is detected the confidence loss of box.

$$\begin{aligned}&arg\ min\left\{\sum_{i=0}^{s^2} \sum_{j=0}^{B} I_{ij}^{p}[(C_i-\hat{C}_i)^2]+\right.\\&\lambda_n \sum_{i=0}^{s^2} \sum_{j=0}^{B} I_{ij}^{p}[C_i-\hat{C}_i)^2]\\&subject\ to\left\{\begin{array}{l} I_i^p\ =1, otherwise=0\\ 0\le \hat{c}_i\le 1, \forall i,j\end{array}\right.\end{aligned} \tag{3}$$

III. An Entire-and-Partial Feature Transfer Learning Approach for Pest Occurrence Frequency

The main deep learning model of this paper combined with multi-task learning, entire-feature transfer learning and partial-feature transfer learning as shown as Fig. 2.

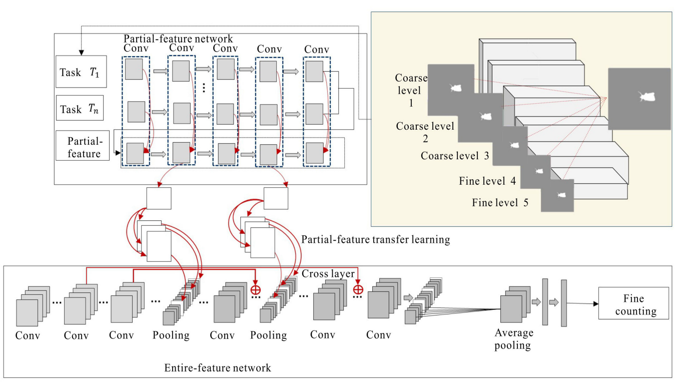

Fig. 2. The entire-and-partial feature transfer learning architecture.

(1) **Entire-and-Partial feature learning** phase: The phase put in multiple feature to partial-feature network, this idea is use multi-task learning with convolutional network that are shared parameter of across tasks. The partial-feature network generation multi-level feature map embed in entire-feature network.

(2) **Partial-feature transferring** phase: The phase includes the entire-feature network and the feature maps of the partial-feature network inputs are embedded in the max pooling layer of the entire-feature network. The feature map embedding method uses a weighting scheme for giving different weight values according to different fine-grained features and selecting appropriate channels for them to join the entire-feature network.

(3) **Entire-feature transferring** phase: The phase more feature information from upsampling features and feature maps from earlier. The entire-feature network yields more meaningful feature information from upsampling features and feature maps from earlier layer. The function of this layer is to connect the feature map of the previous layer with the feature map of the layer. This operation is also to enhance the accuracy of small target detection.

(4) **Pest detection and classification** phase: The phase emphasizes the accurate location of the average detection accuracy of the detection and classification of multiple overlapping thresholds of the network and the measurement of current test evaluation criteria.

The detailed operations of the An Entire-and-Partial Feature Transfer Learning Approach for Pest Occurrence Frequency phases are described as follows.

A. Phase I: Entire-and-Partial feature learning phase

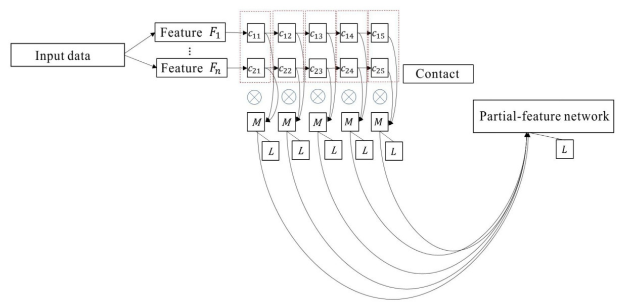

Fig. 3. The multi-feature used multi-task learning to shared parameter.

The convolutional layers are used in an entire-feature network, which have different kernel sizes because scale of the pests is different and feature map are extracted from each convolution layer by applying different filter types $\widetilde{W}$. The main function of using a convolution filter is to extract different features from the image and then generate a feature map. The variable is defined as the $\hat{E}_l^k$, feature map k-th of l layer and b_l^k is defined as the k-th bias of layer l .The convolution operation represents the feature map of the layer calculated as follows in equation 4:

$$\hat{E}_l^k=\phi\left(\widetilde{W}_l^k * E_{l-1}^{\hat{k}}+b_l^k\right) \tag{4}$$

The partial-feature network cover the optimal local sparse structure in a convolutional neural network architecture through locally available dense components . The fixed-size filter kernel of the convolutional neural network will not be able to capture this attribute because the spatial correlation of the feature map is multi-level space, demonstrated Fig. 3. For pest detection, the size of the pests in the image is different. Therefore, if a single fixed size filter element is used, the appearance of the pest will be ignored. Therefore, using multiple filter kernels to increase the width of the convolutional neural network is the main feature of the partial-feature network. The feature map generated by the partial-feature network corresponding to the multi-level space of the image uses a plurality of filters. These feature maps are combined into one output exhibited in the Fig. 4.

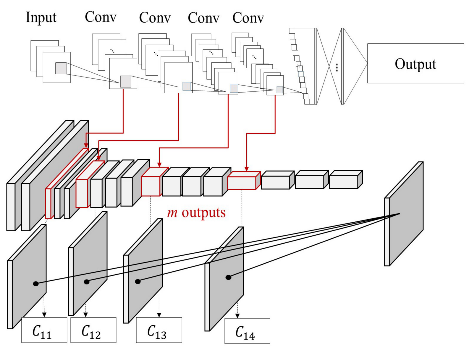

Fig. 4. The entire-and-partial feature network architecture.

S1. The partial-feature learning phase use a multi-level convolutional network to extract features from pest images, thereby enabling image prediction through deep learning models. The partial-feature network utilizes a fully convolutional neural network and supervised learning. The equation 5 trained data is defined as X_n is:

$$T = (X_n, Y_n), n = 1, ..., N \tag{5}$$

the original input and Y_n is the feature map of the image with the label X_n and the partial-feature network outputs is defined as m.

S2. The correspond weight of partial-feature network suppose there are $w = \{w^1, ..., w^m\}$, and W is corresponding parameter with network, two levels of feature aggregation are proposed to integrate feature map of different levels into the coarse and fine levels of the classification network, present to equation 6.

$$L_o(W, w) = \sum_{m=1}^{m} \alpha_m l_o^m (W, w^m) \tag{6}$$

S3. The m-th loss function for output of feature map (m) function as shown in equation 7:

$$\begin{aligned} arg\ min\{-\gamma \sum^{d_1} \log Pr\,(Y_a = 1|\, X : W, w^m) \\ -(1-\gamma) \sum^{d_2} \log Pr\,(Y_b = 0|X : W, w^m)\} \end{aligned} \tag{7}$$

B. Phase II: Partial-feature transferring phase

The output layer feature map generated by filtering the output caused by the partial-feature network, the feature map generated by the partial-feature network is the two inputs of the detection network, as shown in Fig. 5. The detection network obtains the feature map from the mixture network and enhance the feature map corresponding to the output layer uses a weighting scheme.

S1. The weight matrix constrained by the L1 norm is enhanced by the induced channel selection. By choosing superior invariant features, generalization performance and faster convergence speed can be improved. The feature map corresponding to the output layer can be enhanced.

S2. According to equation 8 The method introduced the feature map from the partial-feature network to the convolution layer of the entire-feature network.

$$P_{a,b,c} = \begin{Bmatrix} \max\left\{\hat{F} \otimes \left(1+\hat{W}\right) \otimes P_c\right\} \\ \max\left\{\hat{F} \otimes (1+\beta) \otimes P_c\right\} \end{Bmatrix} \tag{8}$$

Where $\hat{F}$ is feature map of entire-feature map, P_c is the c channel feature map and $\otimes$ indicates that the two matrices have corresponding integrated positions for the flight. Two positive square product combinations are combined use weights.

S3. The pooling is controlled by conventional specifications. This pooling of weighted matrices is defined as a regularization pool. In order to achieve sparsity, the target functions of network training are as follows equation 9:

$$arg\ min\left\{L = |X_i - Y_i| + \Phi|\hat{W}_F|_2 + \Phi|\hat{W}_P|_1\right\} \tag{9}$$

The symbol Φ is hyper parameter, X_i is the feature map of the convolutional network from input to output, Y_i is the labeled data with real, $\hat{W_F}$ is the weight of feature map in entire-feature network, $\hat{W_P}$ is the weight of feature map in partial-feature network, $|\hat{W_F}|$ used L2 regularization and $|\hat{W_P}|$ used non-smooth regularization, can not only prevent overfitting but also enhance the generalization of the model. First, regularize feature map of weights and $\hat{W_F}$ refers to weights in other layers. The Weight of feature map Selection the appropriate pooling layer uses the addition L1 norm to make the weight $\hat{W_P}$ sparse.

C. Phase III: Entire-feature transferring phase

The phase III use the cross layer method in entire-feature network because deep is very important for CNN recognition results, the deeper the network, respectively the training complexity will be greatly improved (the more common reason is because the deeper the network, the more difficult the gradient descent is to update to the front layer, resulting in parameter updates. Extremely slow), compared to the general neural network activation function is to make the input nonlinear transformation and the unit in the cross layer method is also transparent to this concept, every two layers convolution layer will have a shortcut to directly input the input To the back of the network, because the traditional convolution layer will inevitably have data loss when transferring data and this unit can preserve the integrity of the data to some extent, the entire network only needs to learn input and output.

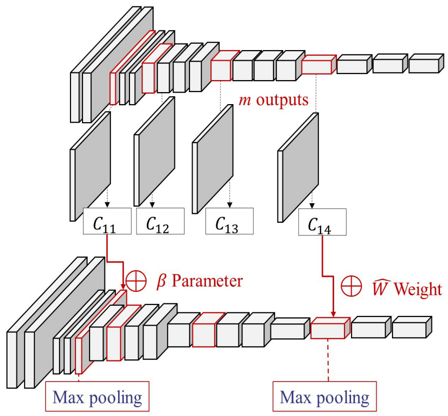

Fig. 5. The partial-feature transfer fine-grained feature map to entire-feature transfer learning architecture.

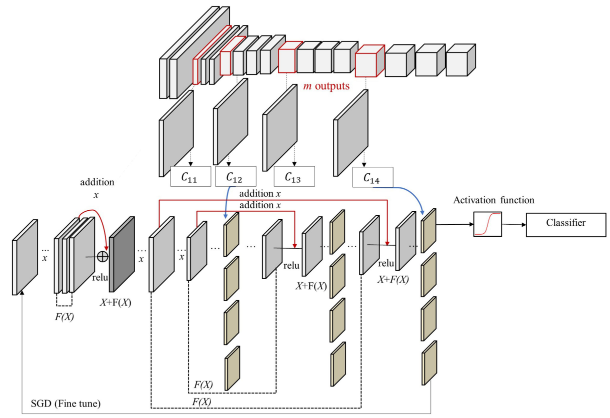

Fig. 6. The entire-feature transferring phase used the cross layer methods.

S1. The difference can simplify the training. Considering the low capacity of the backbone network in feature extraction and the disappearance gradient in back propagation, a convolutional dense connection structure is adopted. The network Improve accuracy by enhancing feature extraction while ensuring maximum information flow in the network.

The connection structure of the entire-feature network present to equation 10, in which the feature map of the first l-1 layer are connected together and used as input to the layer 1. The variable H_l defined as the transformation of the l-th building block. The desired output is $H_l(x)$.

$$F_l(x) = H_l(x) - x \tag{10}$$

The traditional CNN training transformation according to equation 11 of the l-th building block is:

$$H_l(x) = F_l(x) + x \tag{11}$$

The cross-layer mode connection in equation 13 introduces neither additional parameters nor computational complexity. Between ordinary and shortcut topology. The dimensions must be equal to x and F and the linear projection performed by W_s is used to match the dimension's shortcut connection:

$$y = F(x, W_i) + x \tag{12}$$

$$y = F_{(x,Wi)} + W_s x \tag{13}$$

D. Phase IV: Pest detection and classification phase

Accurate positioning of the average detection accuracy of multiple overlapping thresholds and measurement of current test evaluation criteria are emphasized. Apply an NMS with a low threshold when the overlap criterion in the average accuracy.

S1. This is detection frame b_i may be very close to the object, so the detection threshold O_t is called here (in 0.7 overlap range), when b_i low N_t suppression occurs because the score is slightly lower than M (M is not covered). The possibility of this case increases the overlap threshold increase criterion.Therefore, suppress all nearby low N_t detection the box will increase the miss rate. In addition, when O_t is low, using a high N_t will increase the false positive, thereby reducing multiple thresholds average accuracy. In this case, the true positive increase will be much lower. The increase in false positive increases the true positive number because the number of objects is usually much smaller than the number of RoIs produced by the detector. The scoring function as follows equation 14:

$$s_i = \left\{ \begin{matrix} s_i, iou\left(M, b_i < N_t\right) \\ S_i\left(1 - iou\left(M, b_i \geq N_t\right)\right), \\ iou\left(M, b_i \geq N_t\right) \end{matrix} \right\} \tag{14}$$

The detection score above the threshold N_t will be attenuated by the above function to a linear function that overlaps with M. The detection frames far from M will not be affected and larger penalties will be assigned to the closer detection frames.

S2. The labelled data need to pass the model in order to training. Suppose the image have divided into a grid of size 3 X 3 and there are a total of 6 classes which sets to be classified the objects. So, for each grid cell, the label will be an eleven dimensional vector: Here, $\hat{P}_i(c)$ defines whether an object is present in the grid or not (it is the probability) $\hat{x}_i$, $\hat{y}_i$, $\hat{w}_i$, $\hat{h}_i$ specify the bounding box if there is an object $\hat{c}_1$, $\hat{c}_2$, $\hat{c}_3$, $\hat{c}_4$, $\hat{c}_5$, $\hat{c}_6$ represent the classes. So, if the object is a $\hat{c}_1$, $\hat{c}_1$ vector element will be 1 and $\hat{c}_2$ $\hat{c}_3$ $\hat{c}_4$ $\hat{c}_5$ $\hat{c}_6$ vector element will be 0, and so on. The classification approach select the first grid from the above example: Since there is no object in this grid, $\hat{P}_i(c)$ will be zero that means that it doesnt matter what $\hat{x}_i$, $\hat{y}_i$, $\hat{w}_i$, $\hat{h}_i$, $\hat{c}_1$, $\hat{c}_2$, $\hat{c}_3$,$\hat{c}_4$, $\hat{c}_5$, $\hat{c}_6$ contain as there is no object in the grid.

IV. Performance Analysis

The training and fine-tuning the execution algorithm, sample data set 14000 data divides the 8000 training set and 6000 data into effective sets in the system, the learning rate is set to 0.001 and the iteration is set to 70,000. The system uses two algorithms, namely partial-feature transfer learning and the entire-feature transfer learning. In the data graph, it belongs to the red line, the baseline phase belongs to the green line and detection based on YOLO model and classification with SVM phase elongs to the blue line.

(1) **Average precision (AP)**: This evaluation index calculates the accuracy of each category and average precision. The accuracy here is to use true positive in class false positive in class c as the criterion. The predicted proposal and ground truth are the same (the correct type and the overlap is high enough). The predicted proposal and ground truth do not match (the type is wrong or the overlap is not high enough).

(2) **Recall**: This evaluation index calculates the accuracy and average of each category. The accuracy here is the use of true positive in class, false negatives) in class c as the criterion and the predicted proposal and ground truth match (the correct type and the overlap is high enough) and the predicted proposal and ground truth do not match (the correct type but not predicted or undetected objects).

(3) **Training loss**: The Training losses are used to assess pest bonding box and pest classification.

(4) **Class accuracy**: After the pest is detected, the pest bounding box is obtained and then the bounding box is used to obtain the classification result of the pest. In this analysis, only the classification results are presented without the accuracy of pest location.

A. Average precision (AP)

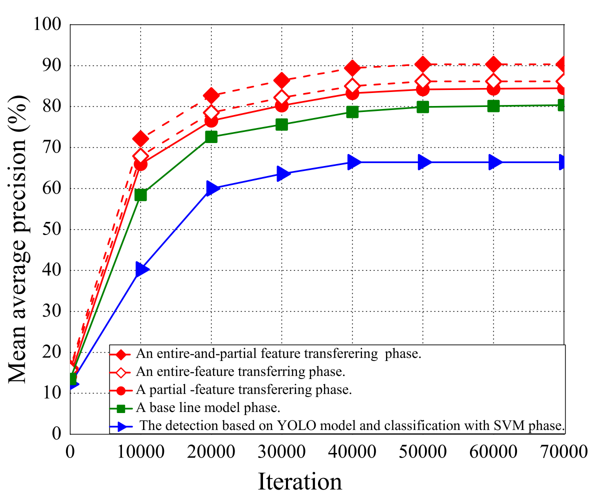

Fig. 7. The average precision of iterations.

The experiment results of the average precision under the each iterations, pest bounding box localization and classification are shown in Fig. 7, The average precision is the evaluation index calculates the accuracy of each category and averages. The accuracy here is to use true positive in class false positive in class c as the criterion. The predicted proposal and ground truth are the same (the correct type and the overlap is high enough). The predicted proposal and ground truth do not match (the type is wrong or the overlap is not high enough).

B. Recall

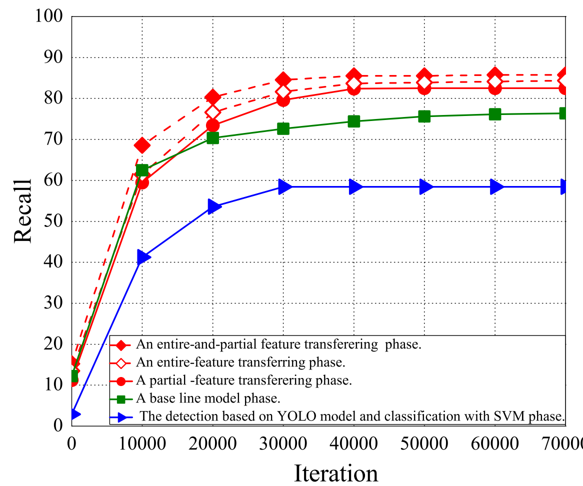

Fig. 8. The Recall of iterations.

The experiment results of Recall under the each iterations, pest bounding box localization and classification are shown in Fig. 8, The Recall is the evaluation index calculates the accuracy of each category and averages. The accuracy here is the use of true positive in class, false negatives) in class c as the criterion and the predicted proposal and ground truth match (the correct type and the overlap is high enough) and the predicted proposal and ground truth do not match (the correct type but not predicted or undetected objects).

C. Training loss

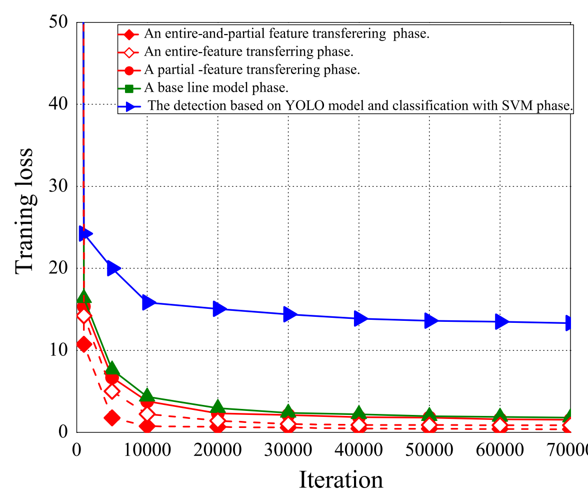

Fig. 9. The training loss of iterations.

The experimental results of training losses at each iteration, location and classification loss of pest bounding box as shown in Fig. 9, first, the each predicted frame with IOU and all ground truth losses calculated and the loss is the confidence error of the calculated background. If a threshold is less than this value, the predicted frame is to be marked as a background and needs to be calculated an object confidence error. And then is the coordinate error is calculated and calculate the difference between the ground truth box and the predicted width. Finally, calculates each part with the loss value that matches its ground truth, the error including coordinate, confidence and classification. Matching principle For the basic fact, first determine which cell the center point falls on and then calculate the the previous box with IOU value and the base fact of the cell. When calculating the IOU value, the coordinates are not considered, only the shape is considered, so the previous frame and the ground truth point are offset to the same position (origin) and then the corresponding IOU value is calculated. The previous box with the largest IOU value matches the ground truth and the corresponding prediction box is used to predict the ground.

D. Class accuracy

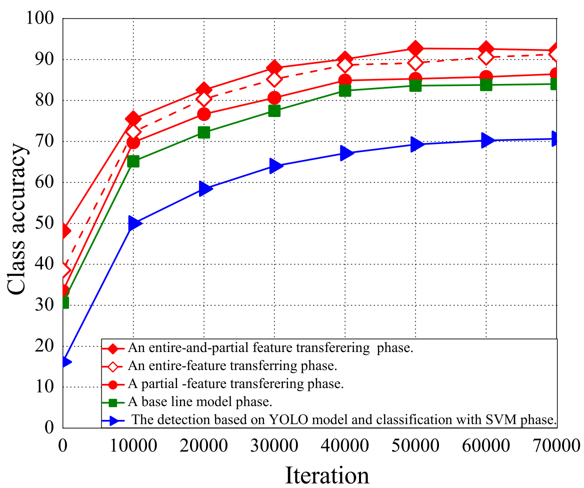

Fig. 10. The class accuracy of iterations.

The experiment results of the class accuracy under the each iterations, pest classification of pest accuracy are shown in Fig. 10, the class accuracy is the evaluation index calculates the accuracy of each category and averages. The accuracy here is to use true positive in class false positive in class c as the criterion. The predicted proposal and ground truth are the same (the true positive and false positive). This evaluation only consider classification of pest accuracy.

V. Conclusions

This paper attempts to solve the statistics of the frequency of pests that consume a lot of manpower through the combination of deep learning and agriculture. We propose an entire-and-partial feature transfer learning network architecture. In partial-feature networks, different fine grained feature mappings are used to enhance the entire feature network, and the cross layer method of the entire-feature network enhances the input of the network layer near the output layer. The feature mapping creates a shortcut topology for the input and output layers to reduce the gradient disappearance problem common to deep networks. The experimental results show that the detection accuracy can be significantly improved, and the method can improve the accuracy of smaller pests.

VI. Acknowledgments

This research was supported by the National Science Council of the R.O.C. under grants USTP-NTPU-NTOU-108-01 and MOST 107-2622-E-305-002 -CC3. The authors would like to thank the anonymous referees for carefully reading an earlier version of this paper and giving many helpful suggestions.

References

[1] K.A. Vakilian and J Massah. ”Performance Evaluation of a Machine Vision System for Insect Pests Identification of Field Crops using Artificial Neural Networks”. *Archives of Phytopathology and Plant Protection*, 46(11):1262–1269, 2013.

[2] J. Wang, C. Lin, L. Ji, and A Liang. ”A New Automatic Identification System of Insect Images at the Order Level”. *Knowledge-based Systems*, 33(1):102–110, 2012.

[3] Nieuwenhuizen Ard and H.K. Suh J. Hemming. ”Detection and Classification of Insects on Stick-traps in a Tomato Crop Using Faster R-cnn”. In *Proceedings of the Netherlands Conference on Computer Vision*, pages 1–4, Eindhoven, Jan. 2018.

[4] Xiea Chengjun, Wanga Rujing, and Zhanga Jie. ”Multi-level Learning Features for Automatic Classification of Field Crop Pests”. *Computers and Electronics in Agriculture*, 152(7):152–156, May 2018.

[5] Hafeez Javed Muhammad, Noor M Humair, and Yaqoob Khan Babar. ”K-means based Automatic Pests Detection and Classification for Pesticides Spraying”. *International Journal of Advanced Computer Science and Applications*, 8(11):174–176, 2017.

[6] Konstantinos G. Liakos, Patrizia Busato, Dimitrios Moshou, Simon Pearson, and Dionysis Bochtis. ”Machine Learning in Agriculture: A Review”. *Computers and Electronics in Agriculture*, 18(8):2674–2678, 2018.

[7] Yuanhong Zhong, Qilun Lei Junyuan Gao, and Yao Zhou. ”A Vision-based Counting and Recognition System for Flying Insects in Intelligent Agriculture”. *Sensors*, 18(5):234–240, 2018.

[8] Limiao Deng, Yanjiang Wang, and Zhongzhi Hanand Renshi Yu. ”Research on Insect Pest Image Detection and Recognition based on Bio-inspired Methods”. *Biosystems Engineering*, 169(6):139–142, May 2018.

[9] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. ”You Only Look Once: Unified, Real-time Object Detection”. In *Proceedings of Computer Vision and Pattern Recognition*, pages 2–4, Las Vegas, NV, USA, June 2016.

[10] Peikari Mohammad, Salama Sherine, Nofech Mozes Sharon, and Martel Anne L. ”A Cluster-then-label Semi-supervised Learning Approach for Pathology Image Classification”. *Scientific Reports*, 8(1):2–4, May 2018.

[11] Daniel Motta, Alex lisson Bandeira Santos, Ingrid Winkler, Bruna Aparecida Souza Machado, and Daniel Andr Dias Imperial Pereira. ”Application of Convolutional Neural Networks for Classification of Adult Mosquitoes in the Field”. *Public Library of Science*, 14(1):1–18, 2019.

[12] Sangheon Lee and Wooju Kim. ”Sentiment labeling for Extending Initial Labeled Data to Improve Semi-supervised Sentiment Classification”. *Electronic Commerce Research and Applications*, 26(6):35–43, 2017.

[13] M.A. Ebrahimi, M.H. Khoshtaghaza, S. Minaei, and B. Jamshidi. ”Vision-based Pest Detection based on SVM Classification Method”. *Computers and Electronics in Agriculture*, 137:52–58, May 2017.