Detecting Pine Trees Damaged by Wilt Disease using Deep Learning Techniques Applied to Multi-Spectral Images

RUIRUI ZHANG¹,², JOONWHOAN LEE², and JIE YOU²
¹School of Computer Science and Engineering, Cangzhou Normal University, Cangzhou, 061001, Hebei, China
²Department of Computer Science and Engineering, Artificial Intelligence Lab, Jeonbuk National University, Jeonju, 54896, Jeollabuk-do, South Korea
Corresponding author: JOONWHOAN LEE (e-mail: chlee@jbnu.ac.kr).

ABSTRACT Pine wilt disease (PWD) is responsible for significant damage to East Asia’s pine forests, including those in Korea, Japan, and China. Preventing the spread of wilt disease requires early detection and removal of damaged trees. This paper proposes a method of detecting disease-damaged pines using ortho-images corrected from 5-band multi-spectral images captured by unmanned aviation vehicles. The proposed method relies on a ResNet18 backbone network connected to a modified DenseNet module, classifies the 5-band multispectral (RGB, NIR, Red_Edge) ortho-image patches, and visualizes the results as a heat map. The patch-based classifier was retrained with hard negative examples, after which it achieved 98.66% accuracy, an improvement over the 96.0% accuracy associated with the same method applied to RGB images. The resulting heat map reflects the approximate distribution, and movement of the disease. Disease locations are also predicted by local maximums in the heat map. When the distance between a ground truth and the predicted location is less than visible distance, e.g. about 5m, it is counted as a correct detection. The proposed detection which consists of heat map generation followed by localization achieves Recall of 93.39%, Precision of 88.26%, and F1-score of 90.75%.

INDEX TERMS Pine wilt disease, multispectral ortho-image, heat map, deep learning classification

I. INTRODUCTION

Pine wilt disease (PWD), known as the “cancer of pine trees”, is devastating to pine forests. The pine wood nematodes (Bursaphelenchus xylophilus) that cause the disease are native to North America [1], though today the disease has spread to East Asia’s pine forests, including those of South Korea, Japan, and China. Early and accurate detection and removal of damaged trees are critical to curbing the spread of pine wilt disease [2], [3].

At present, two preventative and controlling measures can be taken in response to PWD; first, damaged trees may be cut down and burnt; second, use chemical or biological methods for control and prevention such as killing the activity of pine wood nematodes with bacteria to kill pine wood nematodes [4]. In both cases, accurate detection of damaged trees is a prerequisite to control the disease. Ground surveys are a typical means of monitoring dead pine trees, and are highly accurate and reliable. However, mountain topography and road conditions may increase the labor costs of the survey while lowering its efficiency, impairing the deployment of timely and comprehensive responses [5].

Unmanned Aerial Vehicles (UAVs) with multiple sensors are widely used for low-altitude remote sensing purposes in commercial agriculture because of their low cost and high efficiency [6] Remote sensing satellites are also used in forestry and agriculture. In 2006, [7] used satellite imagery to locate forest areas damaged by PWD.

Improvements in deep learning models have led to increased utilization of machine learning and computer vision in many areas [8], [9]. Accurate deep learning-based aerial image analysis has brought about a paradigm shift in the fields of image classification, segmentation, and object detection, and led to numerous successful cases. A deep neural network makes it possible to automatically and efficiently learn local features to construct an accurate
classifier through supervised training [10]. For example, classified land (crop, building, etc.) by using a Convolutional Neural Network (CNN) on images acquired by aerial hyperspectral sensors, ultimately achieving an accuracy of 96%. In many countries, UAVs are often used as land surveying tools for agricultural purposes. A German research team has successfully used deep learning techniques to segment crops from weeds using drone images [11].

This paper proposes a method for detecting damaged pine trees using multispectral ortho-images which includes the bands of Red, Green, Blue (RGB), Near InfraRed (NIR), and Red Edge taken by UAVs. Because ortho-images obtained from a UAV-equipped multisensor are expensive to obtain, a sufficient number of data samples is hard to be expected. Usually one can depend on conventional computer vision approach with the small amount data rather than deep neural networks which have lots of parameters to be adjusted by training [12], [13].

To overcome the lack of sufficient data problem, this paper proposes a patch-based deep classification, in which the ortho-images are discriminated into damaged or normal tree patches. For the input of ortho-image patches, the classifier determines the degree of damage, which can be visualized as a heat map. This heat map is then used to construct another ortho-image, which indicates the spread and trend of the disease and identifies the location of trees damaged by PWD.

ResNet18 is used for a backbone of the classifier. In addition, the front-end DenseNet-based universal function approximation block (UFAB) is integrated [14] with ResNet18 to enhance pool of multispectral bands. The patch-based classifier is retrained with the hard negative examples to get a more elaborated decision boundary, which is finally used to build the heat map [15] to visualize the approximate distribution of diseased trees. Our system shows 98.66% accuracy, which is improvement over 96.0% achieved by RGB image. Disease locations are also predicted by local maximums in the heat map. When the distance between a ground truth and the predicted location is less than visual distance, e.g. about 5m, it is treated as a correct detection. In this way, the proposed detection which consists of heat map generation followed by localization, achieves Recall of 93.39%, Precision of 88.26%, and F1-score of 90.75%. The contributions of the paper can be summarized as follows:

- A patch-based deep learning approach to detect PWD damaged trees in multispectral images is developed, which efficiently uses a small set of training images. The proposed method could be effectively applied for detection tasks in any type of expensive multi band images including hyperspectral images.
- In order to boost the classifier performance, the front-end UFAB was integrated with the ResNet18-based patch classifier, and retraining with hard negative examples was performed.
- The heat map produced from the patch-based classifiers shows the approximate distribution, and its local maximums can be treated as the locations of damaged tree. The experiment shows the effectiveness.

II. RELATED WORK

A. PWD DETECTION

PWD poses a significant threat to the global forest ecosystem, and is capable of disrupting the timber industry and global warming mitigation efforts. Good management of PWD requires early identification and removal of affected trees. The disease symptoms typically appear first on a tree’s top branches, which are difficult to observe at ground level and necessitates the identification of damaged trees from a distance by scanning a relatively large area [16]. UAVs and remote sensing satellite images have emerged as viable alternatives to human-based scans to perform this task. [17] conducted a study in 2017 in Sejong as part of the city’s forest conservation effort and used drones to successfully identify 23 trees (out of the 231 trees total) suspected to be damaged with PWD. In this study, RGB ortho-images were observed by human eyes to detect the damaged tree.

Multispectral images are specially over-determined, and accordingly provide information sufficient to identify and distinguish specially unique materials [18]. RGB, in comparison, only carries three bands of information. Multi-spectral images potentially offer more accurate and detailed information than RGB data alone, and can be used to discriminate the objects that may not be easy in RGB images. Some studies have utilized the spectral characteristics of foliage derived from multispectral images to detect yellow rust, a disease that affects wheat. [19] evaluated the performance of ten spectral vegetation indices at identifying rust infection in individual wheat leaves. [20] studied the effects of different wheat rust disease symptoms on vegetation indices extracted from hyperspectral measurements.

In 2001, [21] analyzed IKONOS high-resolution satellite imagery and detected 15 damaged pine trees out of a total of 22 samples. In this study, the spectral histogram analysis of proved to be the most effective for disease detection. These authors also used the R and NIR band along with IKONOS high-resolution 1m images and various spectral bands of 4m images.

B. DEEP LEARNING TECHNIQUES

Models for early-stage PWD detection used regular machine learning classifiers, including k-nearest neighbors (k-NN) [22], support vector machine (SVM) linear [23] and Gaussian kernel [24], random forest [25], and Gaussian Naive Bayes [26]. The advantage of a convolutional neural network (CNN) is that its multilayer structure allows it to automatically learn advanced features that improve classification and recognition. With CNNs having made tremendous progress in image detection and classification, today’s researchers routinely use it in remote sensing applications, achieving reliable results.

[27] used a Faster-RCNN (Faster Region Convolutional Neural Networks) deep learning framework based on an
A. DATA DESCRIPTION

In this study, the damaged pine trees involved were mainly Pinus armandii, Pinus koraiensis and Larix gmelinii. The training dataset comprised of multi-spectral ortho-images of the Pohang, Jinju and Daegu areas taken by Smart Geo Co. Ltd (Fig. 1) in 2018-2020. RedEdge-Mx was equipped with UAV to capture the multi-spectral images. The maximum flying altitude of each drone was set at 150m by the Korean Forestry Promotion Institute. The resolution of the 5-band multi-spectral ortho-images used in this study was about 20cm per pixel.

B. DATA PREPROCESSING AND DATASETS CONSTRUCTION

1) Data preprocessing

In Fig. 1, the information in parentheses represents spatial resolution. When training a learner, in general, the training/validation/test datasets should all come from the same data distribution. Unfortunately, however, only a limited number of ortho-images were taken at different time and places in this study. It was difficult to separate training images from test images because of the uneven imaging conditions associated with the shooting and production of the multispectral ortho-images. In order to avoid the covariate shifts due to different data distributions, training and test sets were taken from different areas in the same image. Accordingly, each image was divided into two parts, the one used for training and verification (the blue ellipse), and the other used for testing (the red ellipse).

FIGURE 1. Multi-spectral ortho-images and separation of positive examples.

The pixel values of training image were normalized to z-score for each band. While the genuine status of damaged trees would ideally have been determined through a direct human observation, it is impractical to obtain a sufficient amount of ground truth data. Therefore, the ground truth of positive examples was based on either Normalized Difference Vegetation Index (NDVI) for Jinju and Daegu areas, or on expert’s observation for Pohang area. The

III. MATERIALS AND METHODS

A. DATA DESCRIPTION

In this study, the damaged pine trees involved were mainly Pinus armandii, Pinus koraiensis and Larix gmelinii. The training dataset comprised of multi-spectral ortho-images of the Pohang, Jinju and Daegu areas taken by Smart Geo Co. Ltd (Fig. 1) in 2018-2020. RedEdge-Mx was equipped with UAV to capture the multi-spectral images. The maximum flying altitude of each drone was set at 150m by the Korean Forestry Promotion Institute. The resolution of the 5-band multi-spectral ortho-images used in this study was about 20cm per pixel.
ground truth information was confirmed by the Korean Forestry Promotion Institute, and the data preprocessing was performed on ArcGIS 10.5 software. Table 1 depicts the distribution of ground truth values in the dataset.

**TABLE 1.** Positive examples of training, validation and test.

| Number of positive samples (year) | Train/Valid | Test | Total |
|----------------------------------|-------------|------|-------|
| Pohang(2018)                     | 128/35      | 50   | 213   |
| Pohang(2019)                     | 50/0        | 2    | 7     |
| Jinju(2020)                      | 385/80      | 85   | 550   |
| Daegu(2020)                      | 410/153     | 362  | 925   |
| Total                            | 928/268     | 499  | 1695  |

2) Dataset construction

Because only 928 true data points were available for training, which is a manifestly insufficient number, image rotation, image flipping, and other geometric transformation methods were performed to augment the dataset. For data augmentation, the damaged area was first placed at the center of a 400x400 patch image, which was transformed through 15-degree rotations, as well as left-right and up-down flips. To ensure that the location of the PWD damaged trees changed sufficiently inside a patch for robust functioning of our system, the 32x32 images were overlapped and five were taken to augment the data in the center of 400x400 image. Fig. 2 and Fig. 3 provide an example that was extracted through this process. Ultimately, 115,000 patch image were used as positive examples after augmentation, among which 90,000 and 25,000 were used for training and validation data, respectively.

**FIGURE 2.** Augmentation of positive samples by geometric transformation and displacement.

**FIGURE 3.** Extract positive samples by overlapping.

**FIGURE 4.** Patch images for positive and negative samples.

C. NETWORK STRUCTURE OF PATCH CLASSIFIER

Feature extraction and classification were performed with the deep network architecture on each 32x32 patch image extracted in the data preprocessing step. Fig. 5 shows the architecture of the proposed model. According to the Universal Approximation Theorem, neural networks can approximate various non-linear functions; accordingly, a dense network structure with five input bands called UFAB was placed on the front-end of the ResNet18 classifier. The role of UFAB implemented by dense network blocks is to non-linearly combine 5 bands in adaptive way by training [14].

The network included two well-known architectures (DenseNet and ResNet) with some modifications for deep feature extraction. The classifier discriminated the areas damaged by PWD within each 32x32 patch image. ResNet18 was pre-trained on the ImageNet dataset (with the exception of the first stage, described later), and the probability of PWD was calculated by SoftMax.

The rationale behind our model combining these two architectures is as follows:

1) DenseNet learns potential input from each input channel, automatically learns features based on potential bands through network training, and provides a rich feature (or band) pool to support the accurate classification of diseased trees. DenseNet can be placed in front of ResNet backbone, where it alleviates
the vanishing-gradient problem, strengthens feature propagation, and encourages feature reuse. It can also reduce the number of bands in multispectral images without performance reduction, once the essential bands which produce sufficient learned features are identified.

2) ResNet block was designed to build a deep model as thin as possible in favor of increasing its depth and having fewer parameters for performance enhancement. Prior research [29] has shown that residual learning can ease the problem of vanishing/exploding gradients. In addition, the insufficient number of positive samples, as well as the purpose of this study (i.e., to identify economic disease detection method) prompted us to select ResNet18, a lightweight deep network (compared to ResNet34, ResNet50 and ResNet101).

Due to the front-end dense network, our architecture is not identical to conventional ResNet18. In order to fully use the advantage of transfer learning, all the pre-trained weights were imported except the weights in the first layer because there are more than 3 inputs instead of RGB inputs in conventional ResNet18. The weights of the first layer were randomly initialized. Then, the whole network is fine-tuned in the training process.

The 32x32 RGB, NIR, Red_Edge patch images were upsampled by bicubic interpolation to 224x224 images to fit the size of a ResNet18 input. ResNet18 created a 512-dimensional features through convolution and resolution reduction process, and through a Fully Connected (FC) layer, which is converted into a probability value by final SoftMax block.

As shown in Fig. 5, five bands of RGB, NIR, and Red_Edge are inputs of UFAB, which includes three convolutional layers, each of which consists of 1x1 convolution kernels. The output of the dense network is a 14-dimensional feature (or band). Accordingly, the first convolution filter of ResNet18 should be changed into $7 \times 7 \times 14$.

D. THE ENTIRE PROCESS OF DISEASE DETECTION

Fig. 6 shows the proposed detection process, in which a heat map is produced from the patch-based classifier and the locations of damaged trees are identified. The 32x32 (width x height) patch images were sequentially extracted from the multispectral ortho-image by overlapping them in raster scan order, and feed them into the classifier. The classifier then produced the probability of each patch, which represents the degree of a damaged tree. The result of probability is visualized from red to blue as a heat map as in Fig. 6.

If the processing stride is 32 without overlapping parts, the resolution of heat map is too low to make a heat map looking discontinuous. Taking large overlapping parts such as 16 or 24, the resolution can be increased to 16 or 8, and their resulted heat map looks continuous. Of course, there must be a strategy for determining the probability within an overlapped area. For the purposes, the probability value of each pixel was set to the average of the probability values of the overlapped patches. This may reduce the discrimination error due to the ensemble effect of averaging operation on the overlapped area.

To predict the location of damaged trees, the local maximums [31] around 5x5 neighbors were found in 2-
dimensional image. To reduce the occurrence of spurious peaks, a set of potential damaged positions were taken among the local maximums whose heat value were larger than empirically determined threshold.

In general, diseased trees should be removed by human workers, and it is desirable to be easily found within a visual distance. The detection results must guide the location of the diseased tree. When detection accuracy was calculated, the result was assumed to be correct if there was at least one damaged pine tree within the visual distance, approximately 5 meters, which corresponds to 25 pixels at a resolution of 20cm/pixel. Detection performance was evaluated by counting the number of true positives (TP), false positives (FP), and false negatives (FN).

E. RETRAINING WITH HARD NEGATIVE EXAMPLES

After initial training, the trained classifier frequently tended to judge some healthy areas (e.g., brown ground, healthy vegetation with color similar to that of damaged trees, etc.) as diseased trees. This produced a large number of FPs. To reduce the false detections, hard negative example mining [32], [33] was adopted to obtain the examples for retraining the network. Fig. 7 depicts the process of hard negative mining and retraining.

The trained network performed a binary classification, in which probability values above 0.5 were identified as diseased. In general, the output probability values close to 1 is correlated with obvious disease, and those of the samples cropped from healthy areas should be close to 0. However, the probability of some negative samples with similar disease features (hard negative examples) would be close to or greater than 0.5. Accordingly, the definitive decision criterion 0.5 became inaccurate and produced a large number of FPs in the first round of training. This implies the decision boundary is not properly constructed in the first training due to the lack of hard negative examples. Cropped images in the healthy area were classified using the initial trained model. Then a large of hard negative samples could be obtained by setting the decision threshold between 0.35 and 0.65. Note that these samples for mining were drawn from a healthy area, and there could be a sample that produces a large probability close to 1. Those samples may be caused by wrong annotation and are excluded from hard negative example. By this way, 7000 hard negative samples were taken and added to the training and the validation data set. The retaining steps resulted in improvements to the classifier performance, and the final retrained classifier was adopted in subsequent experiments.

IV. EXPERIMENTAL RESULTS

A Python PyTorch library [34] running on a ‘GeForce GTX 2080 Ti’ Graphic Processing Unit was used for the experiment. The data prepared in Section 3. A was used for learning or verification. To overcome the insufficient training data, the pre-trained model was fine-tuned as the backbone with the ImageNet dataset. The learning rate was 0.0001, the momentum was 0.9, and the stochastic gradient descent (SGD) method was used as the optimization. The batch size of the training data was 200 while that of the validation data was 100. The retraining started from the initial learned parameters with the hard negative examples.

A. CLASSIFICATION PERFORMANCE

To measure the classifier’s performance, 499 positive examples of damaged trees were extracted from the portion of each multispectral ortho-image that had not been previously selected for training and validation (Table 1). 1,500 positive images of 32x32 patches were made, using the same augmentation method as used for training and validation. Negative images of the same size were then randomly extracted. The confusion matrix of the classification result is shown in Table 2. For comparison the same experiment was performed with RGB images (Table 2).

Overall accuracy, recall, precision and F1 scores [35] were the metrics by which classification performance was assessed. The overall accuracy is the ratio of the total number of correctly classified samples to the total number of samples of all classes. The value in parentheses () in the following equations represent the results of the RGB images. Each of these performance metrics was determined as follows:

\[
Recall = \frac{TP}{TP + FN} = 98.91% (95.35\%) 
\]

\[
Precision = \frac{TP}{TP + FP} = 98.42% (96.61\%) 
\]

\[
F1 \text{ score} = \frac{2TP}{2TP + FN + FP} = 98.66% (95.97\%) 
\]

\[
Accuracy = \frac{TP + TN}{TP + FP + FN + TN} = 98.66% (96.0\%) 
\]

Results based on multispectral images are improved over those from RGB. Multispectral images are expected to be useful for detecting dead trees, because NIR and Red_Edge are particularly sensitive to moisture, and RGB color is less useful for detecting pine tree damage once their leaves turn brown. Unfortunately, however, the images used in our study were not gathered between late autumn and winter, so the effect of this phenomenon was not significant.

The dense network blocks as UFAB was responsible for the improved performance (Table 2). Without UFAB, recall was 98.74%, precision was 98.23%, F1-score was 98.48%, and accuracy was 98.48%, all of which were slightly under the results achieved with UFAB. Note that we didn’t compare the results with those from NDVI images, because some part of them were included in the ground truth.

In order to justify the model selection, several well-known backbone networks were compared to each other in terms of training and validation accuracy, and the number of
Figure 7. The process of retraining after hard negative example mining.

Table 2. Confusion matrix of classifier for multi-spectral and RGB images.

|                      | Damaged tree images | Other images | Total |
|----------------------|---------------------|--------------|-------|
| Multispectral images (w.o.UFAB) | TP 14811 | FN 189 | 15000 |
|                      | FP 267            | TN 14733    | 15000 |
| Multispectral images (w.UFAB)  | TP 14836 | FP 164 | 15000 |
|          | FP 238            | TN 14762    | 15000 |
| RGB images          | TP 14302 | FN 698 | 15000 |
|                      | FP 502            | TN 14498    | 15000 |

Table 3. Comparison of backbone models with respect to training/validation accuracy and number of parameters.

| Model      | Number of Training parameters | Training accuracy (%) | Val accuracy (%) |
|------------|-------------------------------|-----------------------|------------------|
| VGG-16     | 14.9M                         | 98.78                 | 97.22            |
| ResNet18   | 11.5M                         | 99.45                 | 98.48            |
| ResNet34   | 21.2M                         | 99.75                 | 98.45            |
| InceptionV3| 23.9M                         | 99.01                 | 97.97            |
| DenseNet121| 8.1M                          | 99.26                 | 98.13            |

Although it is not the test result, Table 3 shows ResNet 18 gives the best accuracy with the smallest number of parameters [36].

B. HEAT MAP RESULTS

Using the classifier (Fig. 5), the heat map is generated for test images (actually parts of multispectral image in Fig. 1), which had not been previously chosen for training and validation. Magnified heat map images are shown in Fig. 8 including the extracted 900x900 suspicious regions from the ortho-images. To generate the heat maps in Figs. 8, 9 and 10, the stride was 4x4 to obtain smooth heat maps.

Figs. 8, 9 and 10 reflect locations of the PWD damaged trees. In Fig. 8, the PWD manifests in a row along the highway, whereas in Fig. 9, positions of the diseased trees predicted on the heat map correspond well with the ground truth (yellow dots). Although most of the damaged trees that appeared in a row along the highway in Fig. 8 (a) were removed in 2019, some remain in the upper part of the image of Fig. 8 (b). Fig. 10 (a) and (c) show the separated damaged trees. On the other hand, in Fig. 10 (b) and (d), the areas of concentrated damage are observable on the heat map. (The true values are shown on the left of each pair of pictures. The yellow numbers stand for the number of ground truths.) The yellow boxes in Fig. 8, as well as the blue ellipses in Fig. 10, reflect the roads, rocks, or shaded areas where false detection (FP) occurred. These instances of false detection show the limits of this system. However, when actual images are overlapped into heat map images, false detections can be semi-automatically corrected, and the heat map can be used to identify and predict the spreading patterns of the pine wilt disease.

In addition, multispectral images taken over time may allow us to monitor the spreading pattern of PWD on heat maps by recent change detection techniques [37], [38]. Because the truth value was based on NDVI or expert opinion, our results cannot be said to be improved over them. However, particularly in coniferous areas where damaged trees tend to be located, what may appear to us initially to be a false detection may in fact not be. In other words, there is a possibility that ResNet classifier’s ability to generalize may actually make it more accurate at determining true than NDVI or expert opinion.

Fig. 11 shows the heat map images obtained using (1) the classifier using RGB images, (2) the initially trained classifier with multispectral images (as in Fig. 10 (a)), and (3) the retrained classifier after adding hard negative examples. The heat map of the RGB images is less clear, suggesting its false detection area is larger. These results confirm that multispectral images are advantageous, and that a classifier retrained with hard negative examples can significantly improve results.
FIGURE 8. Heat map and partial magnification.
C. COMPARISON OF DISEASE DETECTION RESULTS

Using the methods discussed in Section 3. D, we performed disease detection on the test images from the multispectral dataset and determined relative accuracy by counting True positives (TP), False positives (FP), False negatives (FN). Table 4 reflects how classification results improved after hard negative examples were added, with a particularly notable reduction in FP. According the data in Table 5, it is obvious the detection based on the retrained classifier with hard negative sample data added outperformed. Note that the deep learning-based detection models such as Fast R-CNN cannot be applied because of the limited number of positive examples.

### TABLE 4. The results of detection by finding the local maxima.

| Ground Truth | TP   | FP   | FN   |
|--------------|------|------|------|
| 50           | 42/46| 5/3  | 8/4  |
| 2            | 2/2  | 6/3  | 0/0  |
| 85           | 66/69| 36/24| 19/16|
| 362          | 339/349| 45/32| 23/13|
| 499          | 449/466| 92/62| 50/33|

### TABLE 5. Comparison of detection results with or without hard negative examples.

|                  | Recall    | Precision | F1-score |
|------------------|-----------|-----------|----------|
| With Hard Negative Samples | 93.39%    | 88.26%    | 90.75%   |
| Without Hard Negative Samples | 89.98%    | 82.99%    | 86.34%   |

D. BAND ANALYSIS

To determine which band is most critical to the classification result, an exhaustive search of 31 combinations of 5 bands has performed and compared in terms of accuracy. Table 6 presents the experimental results of the top 10 performers of these 31 combinations. As the table states shows clear, the presence or absence of the red band produces the maximum impact on classification accuracy. For classification and recognition of PWD, the best classification results were obtained by band input in the following order: R > B > NIR > Red_Edge > G. These results may depend on the dataset, however, as our analysis shows that R, B, and NIR bands are adequate if only three channels are allowed to simplify the sensor complexity.

### TABLE 6. Top10 population results obtained by genetic algorithm.

| R | G | B | NIR | Red Edge | Result(%) |
|---|---|---|-----|----------|-----------|
| 1 | 1 | 1 | 1   | 1        | 98.66     |
| 1 | 0 | 1 | 1   | 1        | 98.45     |
| 1 | 1 | 1 | 0   | 1        | 98.37     |
| 1 | 1 | 0 | 1   | 0        | 98.19     |
| 1 | 0 | 1 | 0   | 1        | 98.08     |
| 1 | 1 | 0 | 0   | 1        | 98.01     |
| 1 | 0 | 0 | 1   | 0        | 97.23     |
| 1 | 1 | 0 | 0   | 0        | 96.25     |
| 1 | 0 | 1 | 0   | 0        | 95.22     |
V. CONCLUSION

This paper proposed a deep learning-based method to identify PWD damaged pine trees using multispectral ortho-images originated from UAVs. A patch-based classifier with ResNet18 backbone was used to overcome the challenges associated with an insufficiency of data samples. A modified UFAB based on DenseNet was attached in front of the ResNet18 to provide richer feature (or band) information. The results of this patch-based classification were accumulated and visualized as a heat map. The heat map produced on the basis of the multispectral ortho-images more precisely reflected the distribution and the spreading patterns of the disease than those produced on the basis of RGB images acquired by the same method.

To train the patch-based classifier, initially trained network was retrained with added hard negative examples, which was effective to improve performance by reducing the FPs. The patch-based classifier configured in this study classified diseased trees with 98.66% accuracy. The dense network to increase the number of features (or bands) also helpful to improve classification performance. In addition, the paper proposed a detection method by finding the local maxima on the heat map generated from the patch-based classifier. After retraining with hard negative sample, recall was 93.39%, precision was 88.26%, and F1-score 90.75%. Also, 31 possible combinations of 5 bands were evaluated to determine which bands, and in which order, impacted classification performance, and experimentally determined that \( R > B > NIR > Red\_Edge > G \). With this information, future models and designers are better positioned to improve efficiency, reduce costs, and miniaturize testing equipment by appropriately reducing unnecessary sensors. In summary, this study has shown that multispectral ortho-images are more effective in detecting PWD damaged areas than RGB images.

VI. ACKNOWLEDGMENT

This research was supported by a grant from the Basic Science Research Program administered through the National Research Foundation of Korea and funded by the Ministry of Education.

REFERENCES

[1] Y. Mamiya, “History of pine wilt disease in Japan,” Journal of nematology, vol. 20, no. 2, pp. 219–219, 1988.
[2] I. Kim, W. K. Lee, T. H. Kwon, D. A. Kwak, Y. S. Kim, and S. H. Lee, “Early detecting damaged trees by pine wilt disease using DI (Detection Index) from portable near infrared camera,” Journal of Korean Society of Forest Science, vol. 100, no. 3, pp. 374–381, 2011.
[3] T. S. Kwon, J. H. Shin, J. H. Lim, Y. K. Kim, and E. J. Lee, “Management of pine wilt disease in Korea through preventative silvicultural control,” Forest ecology and management, vol. 261, no. 3, pp. 562–569, 2011.
[4] N. Ponpandian, S. O. Rim, G. Shanmugam, J. Jeon, Y. H. Park, S. K. Lee, and H. Bae, “Phylogenetic characterization of bacterial endophytes from four Pinus species and their nematocidal activity against the pine wood nematode,” Scientific reports, vol. 9, no. 1, pp. 1–11, 2019.
[5] M. Syifa, S. J. Park, and C. W. Lee, “Detection of the Pine Wilt Disease Tree Candidates for Drone Remote Sensing Using Artificial Intelligence Techniques,” Engineering, vol. 6, pp. 919–926, 2020.
[6] G. Yang, J. Liu, C. Zhao, Z. Li, Y. Huang, H. Yu, B. Xu, X. Yang, D. Zhu, and X. Zhang, “Unmanned aerial vehicle remote sensing for field-based crop phenotyping: content status and perspectives,” Frontiers in plant science, vol. 8, pp. 1111–1117, 2017.
[7] H. Son, W. K. Lee, S. H. Lee, H. K. Cho, and J. H. Lee, “Natural spread pattern of damaged area by pine wilt disease using geostatistical analysis,” Journal of Korean Society of Forest Science, vol. 95, no. 3, pp. 240–249, 2006.
[8] A. Aggarwal and M. Kumar, “Image surface texture analysis and classification using deep learning,” Multimedia Tools and Applications, vol. 80, pp. 1329–1309, 2021.
[9] S. Pang, “VGG16-T: A Novel Deep Convolutional Neural Network with Boosting to Identify Pathological Type of Lung Cancer in Early Stage by CT Images,” International Journal of Computational Intelligence Systems, vol. 13, pp. 771–780, 2020.
[10] T. Kwak, A. Song, and Y. Kim, “The Impact of the PCA Dimensionality Reduction for CNN based Hyperspectral Image Classification,” Korean Journal of Remote Sensing, vol. 35, no. 6, pp. 959–971, 2019.
[11] A. Mililito, P. Lottes, and C. Stachniss, “Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs,” 2018 IEEE international conference on robotics and automation (ICRA), pp. 2229–2235, 2018.
[12] J. A. López-Leyva, E. Guerra-Rosses, and J. Álvarez Borrego, “Multi-Class Diagnosis of Skin Lesions Using the Fourier Spectral Information of Images with Additive Color Model by Artificial Neural Network,” IEEE Access, vol. 9, pp. 35207–35216, 2021.
[13] S. K. Punia, M. Kumar, and A. Sharma, “Intelligent Data Analysis with Classical Machine Learning,” pp. 793–799, 2021.
[14] J. You, W. Liu, and J. Lee, “A DNN-based semantic segmentation for detecting weed and crop,” Computers and Electronics in Agriculture, vol. 178, p. 105750, 2020.
[15] H. Chen, L. Zhang, J. Ma, and J. Zhang, “Target heat-map network: An end-to-end deep network for target detection in remote sensing images,” Neurocomputing, vol. 331, pp. 375–387, 2019.
[16] M. D. Jordache, V. Mantas, E. Baltazar, K. Pauly, and N. Lewycky, “A machine learning approach to detecting Pine Wilt Disease using airborne spectral imagery,” Remote Sensing, vol. 12, no. 14, p. 2280, 2020.
[17] “Annual Report of Monitoring for Forest Insect Pests and Diseases in Korea,” Seong Mun Sa, pp. 1968–2018.
[18] K. Park, Y. ki Hong, G. iwan Kim, and J. Lee, “Classification of apple leaf conditions in hyper-spectral images for diagnosis of Marssonina blotch using mRMR and deep neural network,” Computers and Electronics in Agriculture, vol. 148, pp. 179–187, 2018.
[19] R. Devadas, D. W. Lamb, S. Simplendorfer, and D. Backhouse, “Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves,” Precision Agriculture, vol. 10, no. 6, pp. 459–470, 2009.
[20] D. Ashourloo, M. R. Mobasheri, and A. Huete, “Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements,” Remote Sensing, vol. 6, no. 6, pp. 5107–5123, 2014.
[21] M.-H. Jo, J.-B. Kim, J.-S. Oh, and K.-J. Lee, “Extraction method of damaged area by pine tree pest (Bursaphelenchus xylophilus) using high resolution IKONOS image,” Journal of the Korean Association of Geographic Information Studies, vol. 4, no. 4, pp. 72–78, 2001.
[22] G. Shakhnarovich, T. Darrell, and P. Indyk, “Nearest-neighbor methods in learning and vision,” IEEE Trans. Neural Networks, vol. 19, no. 2, pp. 377, 2008.
[23] J. Platt, “Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods,” Advances in large margin classifiers, vol. 10, no. 3, pp. 61–74, 1999.
[24] S. S. Keerthi and C.-J. Lin, “Asymptotic behaviors of support vector classifiers,” vol. 10, no. 3, pp. 61–74, 1999.
[25] M. Pal, “Random forest classifier for remote sensing machines with Gaussian kernel,” Neural computation, vol. 15, no. 7, pp. 1667–1689, 2003.
[26] A. Aggarwal and M. Kumar, “Image surface texture analysis and classification using deep learning,” Multimedia Tools and Applications, vol. 80, pp. 1329–1309, 2021.
[27] X. Deng, Z. Tong, Y. Lan, and Z. Huang, “Detection and location of dead trees with pine wilt disease based on deep learning and UAV remote sensing,” AgriEngineering, vol. 2, no. 2, pp. 294–307, 2020.
for automated yellow rust disease detection from high-resolution hyperspectral UAV images,” Remote Sensing, vol. 11, no. 13, p. 1554, 2019.

[29] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[30] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708.

[31] M. V. Herk, “A fast algorithm for local minimum and maximum filters on rectangular and octagonal kernels,” Pattern Recognition Letters, vol. 13, no. 7, pp. 517–521, 1992.

[32] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” IEEE transactions on pattern analysis and machine intelligence, vol. 32, no. 9, pp. 1627–1645, 2009.

[33] K. K. Sung and T. Poggio, “Example-based learning for view-based human face detection,” IEEE Transactions on pattern analysis and machine intelligence, vol. 20, no. 1, pp. 39–51, 1998.

[34] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, and L. Antiga, “Pytorch: An imperative style, high-performance deep learning library,” Advances in neural information processing systems, vol. 32, pp. 8026–8037, 2019.

[35] J. Kong, H. Wang, X. Wang, X. Jin, X. Fang, and S. Lin, “Multistream hybrid architecture based on cross-level fusion strategy for fine-grained crop species recognition in precision agriculture,” Computers and Electronics in Agriculture, vol. 185, p. 106134, 2021.

[36] W. Rawat and Z. Wang, “Deep convolutional neural networks for image classification: A comprehensive review,” Neural computation, vol. 29, no. 9, pp. 2352–2449, 2017.

[37] Q. Guo, J. Zhang, and Y. Zhang, “Multitemporal images change detection based on AMMF and spectral constraint strategy,” IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 4, pp. 3444–3457, 2020.

[38] F. Jiang, M. Gong, T. Zhan, and X. Fan, “A semisupervised GAN-based multiple change detection framework in multi-spectral images,” IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 7, pp. 1223–1227, 2019.

RUIRUI ZHANG received the B.S. degree in communication engineering from Beijing University of Posts and Telecommunications and the M.S. degree in computer application from Kunming University Of Science And Technology, in 2010 and 2015, respectively. He is currently working toward the Ph.D. degree with the Artificial Intelligence Lab, Department of Computer Engineering, Chonbuk National University, Jeonju, South Korea. From 2012 to 2015, he studied in the Yunnan Provincial Key Laboratory of Computer Application, Kunming University Of Science And Technology. In 2016, he joined the School of Computer Science and Engineering, Cangzhou Normal University, China. His research interests include deep learning applications of remote sensing image, image processing, and computer vision.

JOONWHOAN LEE was born in Kyunggi-do, South Korea, in 1957. He received the B.S. degree in electronic engineering from Hanyang University, Seoul, South Korea, in 1980, the M.S. degree in electric and electronic engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 1982, and the Ph.D. degree in electrical and computer engineering from the University of Missouri, Columbia, MO, USA, in 1990. In 1985, he joined the Department of Electronic Engineering, Chonbuk National University, Jeonju, South Korea, where he has been a Professor with the Department of Computer Science and Engineering since 1998. Also, he was a Visiting Scholar with the School of Computing Science, Simon Fraser University, Canada, during his sabbatical leave from August 2013 to August 2014. He is currently the Chairman of AI Research Institute, Chonbuk National University. His research interests include machine learning applications of signal processing, and fuzzy and rough sets.

JIE YOU received the B.S. degree in computer science from the Jiangxi University of Science and Technology, GanZhou, China, and the M.S. degree in computer science and engineering from Jeonbuk National University, South Korea, in 2019. Currently, he is pursuing the Ph.D. degree on the Artificial Intelligence Laboratory. His research interests include GIS image processing and analysis, object detection and pattern recognition.

***