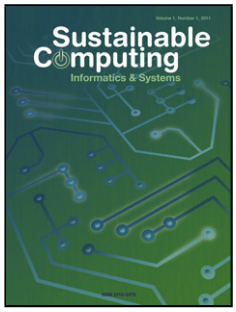

Title: Drone Agriculture Imagery System for Radish Wilt Disease Identification via Efficient Convolutional Neural Network

Authors: L. Minh Dang, Ibrahim Hassan Syed, Im Suhyeon, Arunkumar Sangaiah, Irfan Mehmood, Seungmin Rho, Sanghyun Seo, Hyeonjoon Moon

PII: S2210-5379(17)30401-8
DOI: https://doi.org/10.1016/j.suscom.2018.05.010
Reference: SUSCOM 250

To appear in:

Received date: 18-10-2017
Revised date: 12-1-2018
Accepted date: 18-5-2018

Please cite this article as: Dang LM, Syed IH, Suhyeon I, Sangaiah A, Mehmood I, Rho S, Seo S, Moon H, Drone Agriculture Imagery System for Radish Wilt Disease Identification via Efficient Convolutional Neural Network, *Sustainable Computing: Informatics and Systems* (2018), https://doi.org/10.1016/j.suscom.2018.05.010

Drone Agriculture Imagery System for Radish Wilt Disease Identification via Efficient Convolutional Neural Network

L. Minh Dang[1], Ibrahim Hassan Syed[1], Im Suhyeon[1], Arunkumar Sangaiah[2], Irfan Mehmood[1], Seungmin Rho[3], Sanghyun Seo[3], Hyeonjoon Moon[1*]

[1] Department of Computer Science and Engineering, Sejong University, Seoul, Republic of Korea;

[2] School of Computing Science and Engineering, VIT University, Vellore, India

[3]Department of Media Software, Sungkyul University, Anyang, South Korea

*Correspondence: hmoon@sejong.ac.kr

Highlights

- Detecting and categorizing the criticalness of Fusarium wilt of radish.
- Segmenting the radish regions from other regions on the field such as bare ground and mulching film.
- Classifying the severity of the Fusarium wilt of radish based on thresholding a range of color features.
- Creating two different datasets. One dataset contains images for radish, bare ground and mulching film regions which were manually verified and collected. The other dataset contains images for healthy, disease light and disease heavy radish. This dataset will be made public for the research community for further experimentation and simulations

Abstract. The significant role of plants can be observed through the dependency of animals and humans on them. Oxygen, materials, food and the beauty of the world are contributed by plants. Climate change, the decrease in pollinators, and plant diseases are causing a significant decline in both quality and coverage ratio of the plants and crops on a global scale. In developed countries, above 80 percent of rural production is produced by sharecropping. However, due to widespread diseases in plants, yields are reported to have declined by more than a half. These diseases are identified and diagnosed by the agricultural and forestry department. Manual inspection on a large area of fields requires a huge amount of time and effort, thereby reduces the effectiveness significantly. To counter this problem, we propose an automatic disease detection and classification method in radish fields by using a camera attached to an unmanned aerial vehicle (UAV) to capture high quality images from the fields and analyze them by extracting both color and texture features, then we used K-means clustering to filter radish regions and feeds them into a fine-tuned GoogleNet to detect Fusarium wilt of radish efficiently at early stage and allow the authorities to take timely action which ensures the food safety for current and future generations.

Keywords. Unmanned aerial vehicles, feature extraction, radish field clustering, Fusarium wilt of radish classification

1. Introduction

In Korea, radish is considered the national vegetable, occupying about 10% of the entire vegetable farming area. It is an indispensable ingredient in soups, stews, and other dishes. However, the radish yield has decreased sharply due to the Fusarium wilt of radish, a disease has emerged at an unprecedented rate. The symptoms include wilting, chlorosis, necrosis, premature leaf death and a deterioration of the vascular elements in roots, stems, and petioles, which lead to the imminent death of the infected plant [1]. It is a challenge to prevent and treat the disease for several reasons. When the disease appears, it spreads rapidly from infected plants to healthy plants, resulting in severe harvest losses. Early detection of the disease could prevent the widespread of the disease and mitigate the damage. Manual field inspection has been applied for a long time, but it is inefficient and time-consuming. Therefore, a surveillance system, which can precisely and automatically detect the Fusarium wilt of radish, will likely become a reality more than ever as advanced technologies such as UAVs, IoT application, and remote sensing are emerging.

Remote sensing is part of a GIS system that supports the measurements of the agricultural area. Satellite, SAR and airplanes have been three major remote sensing technologies for a quite some time. The results acquired from satellite remote sensing are affected by poor resolution images or inaccurate information due to poor weather conditions and high costs [2], whereas in airplane remote sensing, the plane is equipped with multiple sensors and cameras that provide high quality and detailed information, yet the equipment is expensive and hard to operate [3]. On the other hand, SAR remote sensing, the ground resolution is affected by the length of the microwave beam generated by the antenna [35,38]. As an alternate method, unmanned aerial vehicles (UAV) are a remote-controlled aircraft which records the field surface at relatively low-altitudes [4]. Owing to the advances in sensing technology, control, and positioning techniques, UAV is now capable of acquiring high spatial resolution surface images at low operational cost. Because it has multiple functions as well as its cost-effectiveness, UAV applications are rising rapidly. It can be applied in many areas from traffic monitoring [5] to forest fire monitoring [6] and search & rescue operations [7]. In addition, UAV has promising potential in converting precision farming into autonomous farming [8]. It assists farmers in observing crop and field information promptly, which leads to improvement in crop management and farm planning.

Beside remote sensing, Internet of things (IoT) concepts are emerging and has drawn much recent attention to the research community [21,27,28,29]. IoT lets various devices around us communicate and collaborate with each other. In the case of radish farming, multiple sensors have already been placed on the field to collect periodic scalar data such as light, temperature, humidity, etc. then the data are sent directly to the main controller for analyzing and automatically adjusting the optimal environmental parameters for radish. As UAV applications have rocketed in recent years, it can be integrated into existing IoT system on the field. By letting the drone fly over pre-programmed locations at a specific time on a daily basis to record multimedia contents such as video and images, these contents are then sent to the main controller for detecting the wilt of radish in real-time and notifying the users detailed information by sending a message if the disease is detected.

Traditional approaches in image representation have relied heavily on extracting hand-engineered features like color, texture and shape features [9,10,36,37], scale-invariant features transform (SIFT) [11], and bag-of-visual-words [12]. Those systems relied on shallow classification methods like support vector machines [22], decision tree [24], random forest, dependencies of inter-block coefficients [30], multi-scale non-negative sparse coding [31], multi-factor feature [32,33], swarm intelligence [25], fuzzy reinforcement [34] and neural networks for identifying plant diseases [26]. However, the low discriminative ability of these features and their failure to describe high-level semantics resulted in poor classification performance. Recent successes in deep learning based methods in image classification tasks have received many interests of computer vision researchers to utilize these powerful hierarchical architectures for various tasks [13]. While modern CNN architectures do require huge computational power, efforts have been made to reduce their computational needs while keeping their performance as high as possible. As a result, several efficient CNN architectures have been proposed like GoogleNet [14], Network-in-Network [15].

Based on the above analysis and considerations, we proposed a novel approach to combine UAVs with computerized methods for detecting the critical nature of the Fusarium wilt of radish. The main contributions include four aspects:

- Detecting and categorizing the criticalness of the Fusarium wilt of radish.
- Divided the radish regions from other regions on the field such as bare ground and mulching film by using clustering algorithm.
- Classifying the severity of the Fusarium wilt of radish based on thresholding the range of color features.

Creating two different datasets. One dataset contains high-resolution radish field images obtained from the drone. The other dataset contains images of radish, bare ground and mulching film regions which were manually verified and collected. They will be made public for the research community for further experimentation and simulations.

2. Datasets

2.1. Image Acquisition

The images used in this study were captured in different areas in Korea including Jungsun, Gangwon and Hongchun between July and September 2016. Two commercial UAVs (Phantom 4, DJI co., Ltd.), equipped with RGB camera (12 megapixels), were used to obtain the images at the altitudes of approximately 10m above the ground level. A total 40 images were acquired, and each image has dimensions of 4000 x 3000 pixels at 72dpi. Fig. 1 introduces an image captured by drone on the radish field.

2.2. Dataset Preparation

From the original 40 images, two datasets were created. Dataset A contains manually cropped and labeled regions of interest (ROIs) representing three regions: radish, bare ground, and mulching film as shown in Fig. 1. In total, 1,500 ROIs were extracted; 500 ROIs for radish, 500 ROIs for bare ground, and 500 ROIs for mulching film regions. Dataset A is used for radish region classification. Dataset B contains the original 40 images of size (4000x3000) pixel.

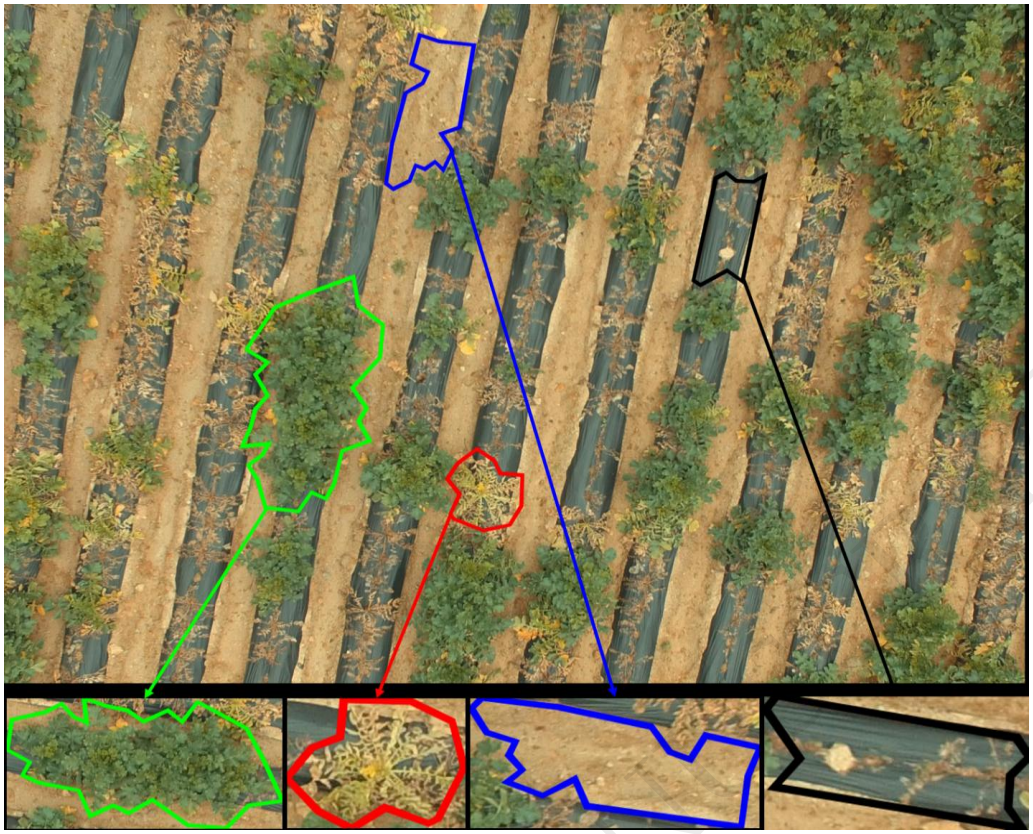

Figure 1. Sample image of radish field. Different colored outlines indicate different regions of radish fields (green indicates healthy radish, red indicates Fusarium wilt of radish, blue indicates bare ground, and black indicates mulching film)

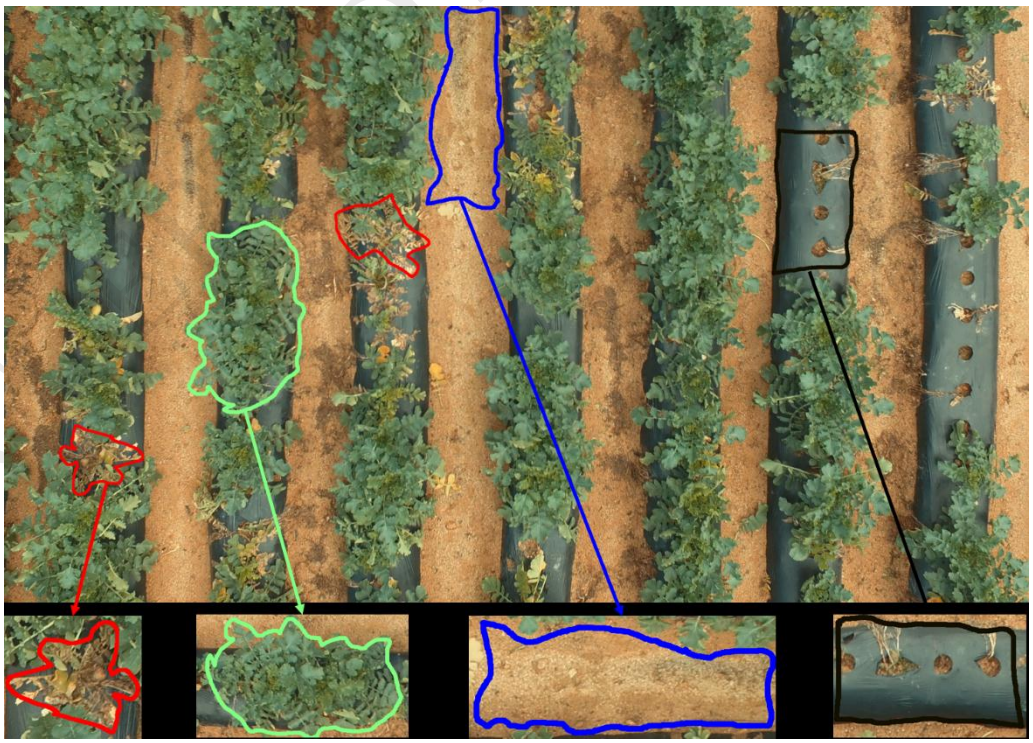

Figure 2. Another sample of radish field. Different colored outlines indicate different regions of radish fields (green indicates healthy radish, red indicates Fusarium wilt of radish, blue indicates bare ground, and black indicates mulching film)

3. Methodology

As shown in Fig. 1, the entire radish field contains three main regions (Radish, bare ground, and mulching film). The final goal of this study is to detect Fusarium wilt on radish leaves, so clustering is needed to divide the radish field into Dataset B into distinct regions. Next, a classifier is trained to recognize which region is radish using Dataset A, the output after applying this classifier to the clustered regions is radish regions. After that, by sliding a fixed size rectangular window (64x64, 128x128, 256x256) pixel through the radish regions from the clustering step, a list of small size radish images was extracted, each image was assigned a label indicates the level of disease by applying various image processing methods. Finally, a convolutional neural network (CNN) model is implemented for classifying the Fusarium wilt of radish based on the severity of the disease. The overall process of the proposed model is shown in Fig. 3.

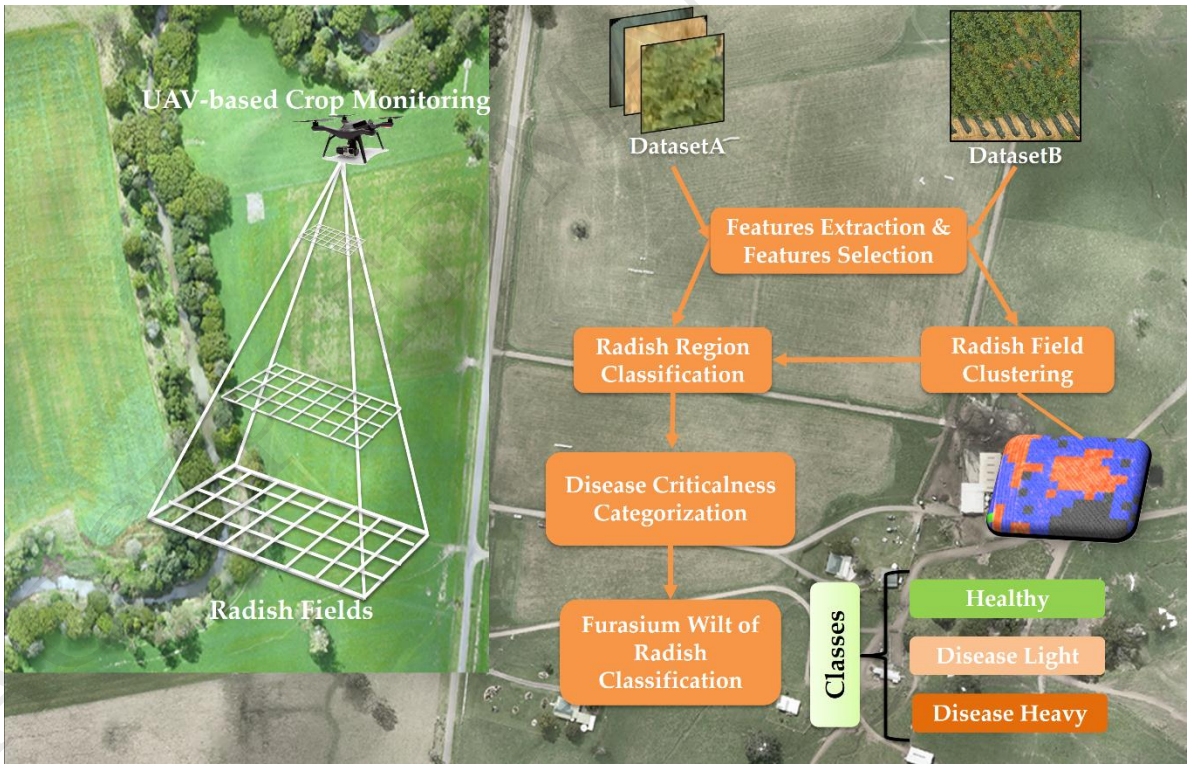

Figure 3. Overall architecture of Fusarium wilt of radish detection model

3.1. Features Extraction & Features Selection

An image contains various useful features such as color, shape, texture. Each type of feature serves a specific purpose in image processing. However, image information retrieval using only one kind of feature is not good enough for the accuracy and efficiency of the model. The high dimensional features lower the model efficiency whereas low dimensional feature decrease model accuracy, so it is better to use multi features for feature retrieval. Since color and texture are the most important visual features, they are extracted in this study. As shown in Fig. 4, Local binary pattern (LBP) in [16] was applied to extract texture features from radish image whereas color features were extracted by using color-space conversion. After that, two featured sets were combined into one feature vector then AutoEncoder (AE) [17] was utilized to reduce the feature dimension.

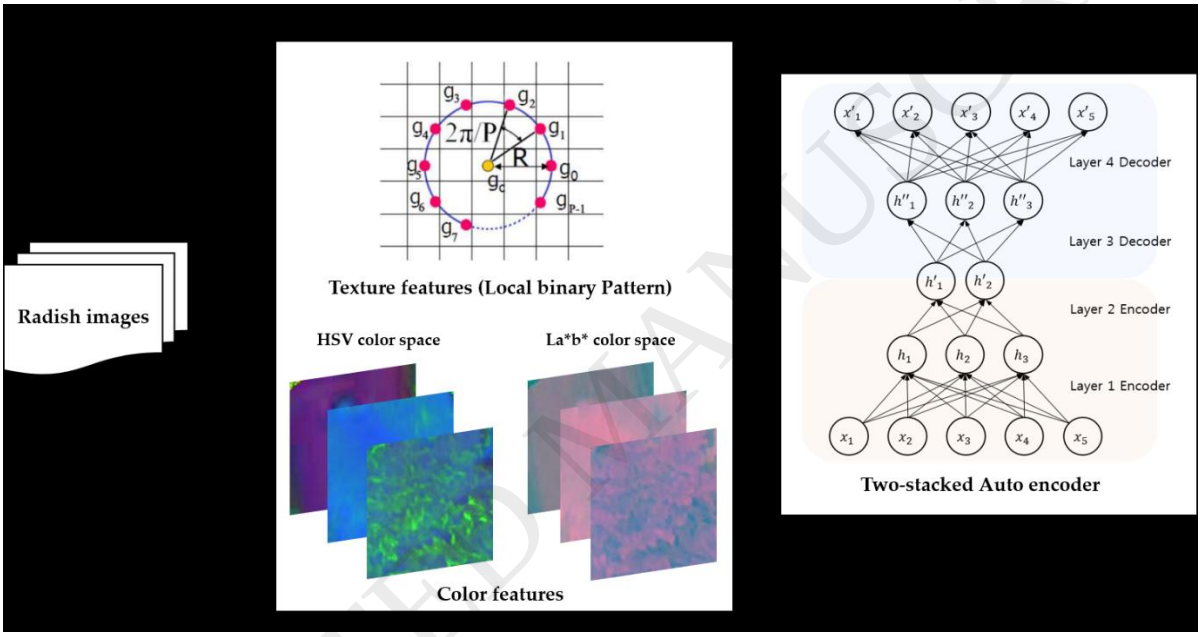

Figure 4. Features extraction and features selection process

3.1.1. Features extraction

The local binary pattern provides a robust texture descriptor which is invariant to rotation and illumination changes at low computational cost. In the image, randomly select a pixel c, LBP takes its neighboring pixels SS 3í (a set of nearby pixels which form circle shape) under a radius R and creates a binary pattern vector as shown in following formula:

$$.\$ \quad \text{ÉÆË} \; L \; \tilde{A}\tilde{a}^{\acute{E}?5}_{\tilde{a}@4} Q\, C_{\tilde{a}} \; F \; C_{\ddot{O}}; t^{\tilde{a}} \qquad (1)$$

where s[LVLI[RULI[$C_{\ddot{O}}$ and $C_{\tilde{a}}$ indicate the grayscale of the center pixel and its neighborhood pixels, respectively. However, the binary pattern vector generated above is still sensitive to rotational invariance, so Equation (2) is a solution to overcome this problem:

$$.\$ \quad \text{ÉÆË} \; L \; \bullet \quad <414 \; k.\$ \quad \text{ÉÆË} \; \text{áÆo} \quad EL \; r \; \text{á}s\text{áå} \; \text{á}2 \; F \; s \; = \qquad (2)$$

where 414 : TáE; is calculated by right shift operation on a circular bitwise which means that the same binary pattern code created by the bitwise operation is regarded a similar pattern. In our case, LBP features are calculated on gray image by computing three nearby topologies (P, R) = {(8, 1), (16, 2), (24, 3)} which generated 703,404 features.

Radish field images were captured in RGB (red, green and blue) color space then converted into HSV (hue, saturation, value) to separate color information from luminance and L*a*b color space (L stands for lightness, a and b represent the color components green±red and blue± yellow, respectively) to approximate human vision. Then the histograms were generated using hue, *a, and *b channels (one histogram contains 256 bins or features). After that, three color histograms were concatenated into one color histogram, generating approximately 768 color features.

3.1.2. Feature selection

After extracting texture features and color features, they were combined into one single features vector. However, two features cannot be directly combined because there is a huge difference between their dimensions. Therefore, the classification and clustering step will be influenced by the texture features which had many more dimensions than the color features. This imbalance will probably affect the accuracy of the model. The solution to this problem is to scale the dimensions of these two features to the same dimensions. Initially, texture features $P_Ü$ is shrunk, whereas color features $?_Ü$ is extended so each features set will occupy half of the dimensions @& of the input vector $T_Ü$

The input vector $T_Ü$ is made by concatenating texture features $6_Ü$ and color features $\%_{\phi Ü}$ which are the shrunken texture features of $P_Ü$ and extended color features of $?_Ü$ respectively. $6_Ü$ $\%_{\phi Ü}$ and $T_Ü$ are computed as follows:

$$6_Ü L\ 9_ç P_Ü E >_ç \quad (3)$$

$$\%_{\phi Ü} L\ 9_Ö ?_Ü E >_Ö \quad (4)$$

$$T_Ü L\ \succ 6_Ü \%_{\phi Ü} ? \quad (5)$$

where the matrices $9_ç Ð 4^{\times ã H \times ß}$ and $9_Ö Ð 4^{\times ã H \times î}$ represent the weights, $>_ç$ and $>_Ö$ represent their biases.

After applying scaling, the texture features $P_Ü$ was reduced from 703,404 to $6_Ü L$ 10,000, and color features $?_Ü$ was extended from 768 to $\%_{\phi Ü} L$ 10,000. As the result, the input vector $T_Ü$ contains a total of 20,000 features. The number of features was in very a high dimensional space so the features selection method is implemented to improve the system performance.

There are various available features selection techniques such as PCA, Stacked AutoEncoder, etc. PCA is a very simple technique that only performs a linear transformation on the input space to align directions of maximum variation with the directions of the axes, whereas Stacked AE [17] is a much more sophisticated technique that can model relatively complex relationships and non-linearities of the input dataset. It includes an input layer, an output layer (having the same dimension) and hidden layer(s) which learn the approximation/representation from the input. The dimensions of the input and output layers are greater than the hidden layer. The hidden layers learn the input compressed representation (encoding), i.e., getting all the input meaningful features. Finally, the input features were reduced and compressed by applying two-stacked AE [18] on 20,000 features to only 1,770 features.

3.2. Radish Region Classification

The classifier used in this study is the softmax classifier, which computes the probability to assign a specific class to a region. The class with the highest probability is appointed to the region. It is mainly deployed as a final classification layer for the neural network-based model. With a feature vector T, the softmax classifier returns the probability for each class label F: FL s áå á%; as follows:

$$2:UL\ FT;\ L\ \frac{Ø^{à:ã;Õ}}{\tilde{A}_{Õ8}^{¼}\ Ø^{à:ã;Õ}}\ á\ BKNFL\ s\ å\ ä\% \tag{6}$$

$$Q:T;_{Ý}L\ \tilde{A}\hat{A}_{Ú@5}\ S_{ú}\ ®T_{Ü}E\ \succ_{Ý} \tag{7}$$

in which $U_{Õ}$ represents the class label, $S_{ú}$ is the weight and $\succ_{Ý}$ is a bias :EL s áå á+áFL s áå á%; . + and % denotes the features and name of class (mulching film, bare ground, and radish), respectively. The weight ™ and bias „ are computed to reduce mean squared error (MSE) by iterating 200 times. Equation (6) is called softmax function that return a C-dimensional vector which contain values between 0 and 1, representing the categorical probability distribution. The class with has highest probability is assigned to T.

Support vector machine (SVM) was also implemented to compare with softmax classifier to find the optimal classifier for the dataset, SVM is proved to be one of the most efficient algorithms for classification problem [22]. It uses a decision boundary to separate sample from different classes. The SVM kernel used in this study is linear kernel because the number of features were larger than the number of training samples, so there is no need to map data to a higher dimensional space, and we only needed to search for the suitable parameter C. LibSVM [23] is a library for implementing support vector machines (SVM). It helps the users implement

and customize SVM easily to fit their case. After using cross-validation supported by the library, the best values for C was 0.5.

3.3. Radish Field Clustering

In the entire radish field image, radish is the only region of interest for further steps. Thus K-means clustering is implemented to cluster the radish field image into distinct regions, and only the radish regions are collected. Fig. 5 shows the entire process of radish field clustering. Features are extracted from original radish RGB field image. Then, K-means clustering is performed with K=3, 5, 10, 15 and 20. For each cluster, the radish field classifier from the previous section is used to assign the class name to each cluster.

After filtering out the radish regions through K-means clustering, a fixed window was created to slide through the region, and the window size was set to 64, 128, and 256 sequentially, so we can discover which window size achieves the best Fusarium wilt of radish classification results.

Figure 5. Overview of radish field clustering process

3.4. Disease Criticalness Categorization

As the criticalness of Fusarium wilt appears in the radish at many stages, farmers have a specific treatment plan for each stage. Early and accurate detection and diagnosis of plant diseases is the key factor in plant production and the reduction of both qualitative and quantitative losses in the crop yield. Fig. 6 shows sample ROIs for healthy, light and heavy disease.

Each image extracted from the clustering step was converted from RGB to HSV color space, and after that, thresholding operations were performed by applying a range of pixel

values representing the greenish region. When applying the thresholding technique, the portion of a black pixel represents the yellowish region (Fusarium wilt), while the portion of a white pixel represents the greenish region (healthy radish). Radish that suffered from the more severe disease has more black pixels in the threshold image than healthy radish or radish suffered from an early stage of the disease. Therefore, the following properties are used to set the threshold to classify each input image into one of these three categories: healthy, light disease, and heavy disease as illustrated in Fig. 6. If the percentage of white pixels in the image is over or equal to 90%, it is categorized as a healthy radish. For the light disease radish, it is less than 85% and over 65%. The image is categorized as heavy disease radish if the percentage of white pixels is less than 65%. As the result, for each window size (64, 128, 256), 6,000 images were extracted: 2,000 for healthy radish, 2,000 for light disease, and the other 2,000 images for heavy disease

Figure 6. Sample criticalness categorization on (128x128) pixel window size

3.5. Fusarium Wilt of Radish Classification

To detect and classify the criticalness of the Fusarium wilt of radish, a well-known CNN model GoogleNet [14] is adopted. This CNN model was proposed in the ImageNet large-scale visual recognition challenge [19], this CNN model achieved a remarkable performance of 5.5% top-5 classification error compared to AlexNet 15.3% top-5 classification error.

3.5.1. Convolutional neural network

This CNN model is more complicated and deeper in analyzation than all previous CNN PRGHOV 0RUHRYHU LW XVHV ¸QFHSWLRQ ZKLFK FRQFDWHQDWHV ILOWHUV RI YDULRXV VL]HV dimensions into a single new filter. It contains two convolutions, two pooling and nine ¸QFHSWLRQ ODHUV,QWKLVSDSHUWKUHHGLIIHUHQWVL]HVZHUHXVHG as an input (64x64, 128x128, 256x256) corresponding to 3 different window sizes (section 3.3). By comparing them, we can decide which size gives the best results. Fig. 7 shows the illustration of CNN architecture used in this paper.

Figure 7. Illustration of GooglH1HW¶DUFKLWHFWXUH

3.5.2. CNN training process

After the disease criticalness categorization in Section 3.4, we have three small datasets for three different window sizes: each size includes 2,000 healthy radish images, 2,000 light Fusarium wilt, and 2,000 heavy Fusarium wilt. The entire dataset was randomly divided into training and validation sets. The training and validation set occupied 75% and 25% of the entire dataset respectively, and the validation set was used for tuning the learning rate. In the training phase, the batch size was set to 90 and momentum was 0.9. The learning rate was initially set to 0.01 and gradually reduced to 0.0001 according to the error rate of the validation set, as the parameter vector bounces around chaotically with a high learning rate. Therefore it is ideal to step decay the learning rate. The training phase ran for 30 epochs, which took approximately 1 hour.

The system used for training the CNN model was NVIDIA DIGITS toolbox with Caffe framework. The experiments were implemented on a Linux machine preinstalled Ubuntu 14.04, it used Intel® Core i7-5930K processor, four 3072 Cuda cores, 4 Titan X 12GB GPUs, and 64GB of DDR4 RAM.

4. Experimental Results and Discussion

4.1. Evaluation Methodology

4.1.1. Classification evaluation protocol

K-fold cross-validation (k=3) is applied to evaluate the system performance. It seperates the entire dataset into roughly k equal-sized subsets. Two subsets are used to train the proposed method, while the remaining subset is utilized to compute the performance of the method. The training process is reiterated k times with a different collection of the remaining subset. In radish field classification, the confusion matrix is calculated to evaluate the ability of our model to distinguish different regions (mulching film, bare ground, and radish). The confusion matrix CM is calculated as:

%/ ŭ L Ã ËÐ× [NÐ 4 O QDPD= P) N K Q J @D6 N Q E= J @ N A @ E:?N PLEFK J (8)

where 4 is the ROIs and)NKQJ@6NQDPN and PNA@E?PEKJ denote the ground truth class label and predicted class label of an ROI N respectively.

4.1.2. Clustering evaluation protocol

The K-means clustering was accessed with a different value of K (K=3,5,10,15, and 20). Therefore, the results varied when the different K were applied. To find the value of K, which had the highest performance, the clustering results were evaluated by calculating the pixel-level clustering accuracy [20] (PSA). PSA is the most popular semantic clustering measure that evaluates pixel-level classification accuracy. PSA is calculated in Equation (9)

$$25\# \; L \; \frac{\tilde{A}_{O}^{\grave{U}}{}^{1\!/\!4}{}_{O}}{\tilde{A}_{O}^{U}\tilde{A}_{O}^{U}{}^{1\!/\!4}{}_{O}} \tag{9}$$

PSA was measured using a range of K values (K=3, 5, 10, 15, and 20). Thus we can decide whether the size of the clusters affects the clustering performance**.**

4.2. Radish Region Classification Results

Table 1 and Table 2 show the confusion matrix of classification results on Dataset A (radish ROIs: 500, bare ground ROIs: 500, and mulching film ROIs: 500). After that, we calculated precision, recall, and F-measure for SVM classifier and softmax classifier using the confusion matrix. The result is shown in Fig. 8; it is noticeable from the graph that softmax classifier achieves the overall performance over 90%, whereas for SVM the F-measure is 72% since correctly classified bare ground was only 342 per 500 samples. As a result, the softmax classifier is chosen to be the default classifier in the model.

Table 1. Confusion matrix for softmax classification

Predicted \ True	Radish	Bare Ground	Mulching Film
Radish	**498**	0	13
Bare ground	1	**491**	104
Mulching film	1	9	**383**

Table 2. Confusion matrix for SVM classification

Predicted \ True	Radish	Bare Ground	Mulching Film
Radish	**475**	3	2
Bare ground	18	**342**	6
Mulching film	7	155	**492**

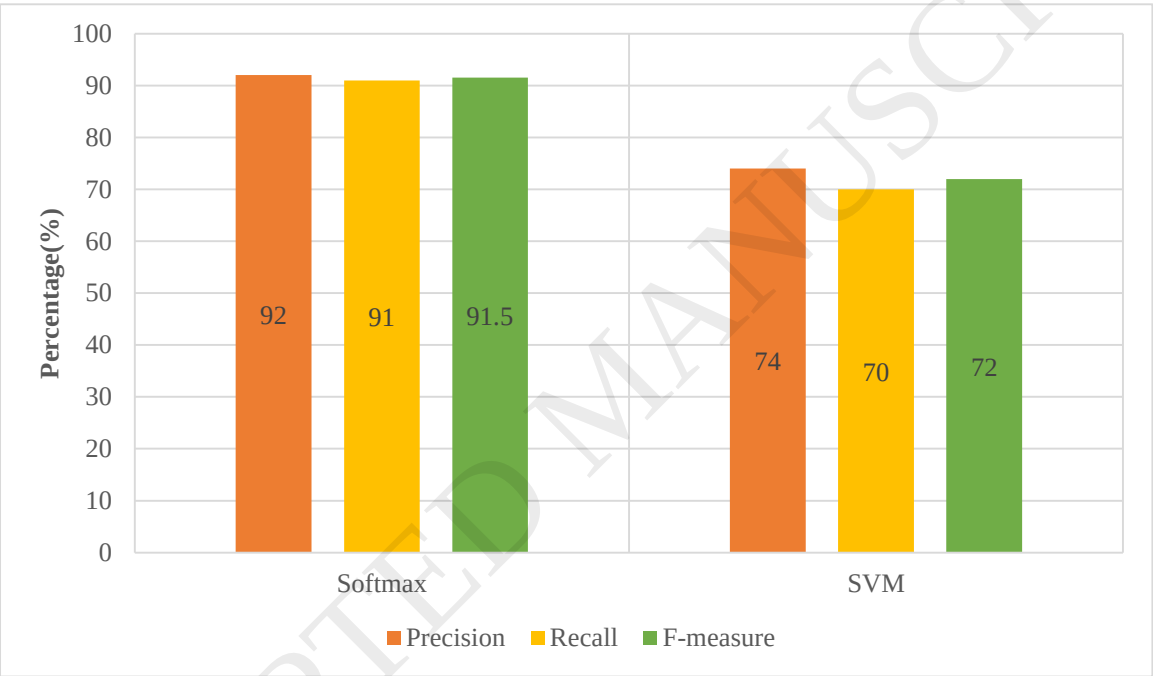

Figure 8. Comparison between SVM and softmax classifier on Dataset A

4.3. Radish Field Clustering Results

Radish field clustering accuracy was computed by pixel-level clustering accuracy (PSA), it is also validated by 3-fold cross validation with same dataset and classifier used for radish field classification evaluation.

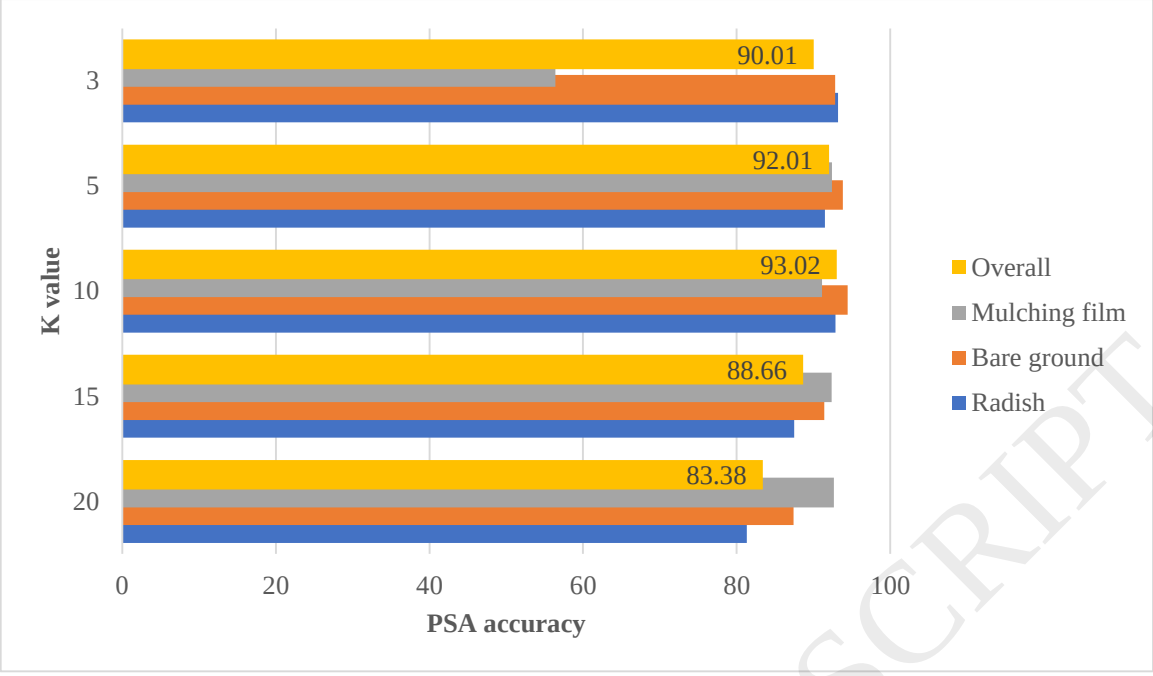

Figure 9. PSA evaluation on K-means with different K

Radish field clustering results (PSA) with different cluster (K) values are presented in Fig. 9. As K increases, the overall PSA accuracy increases gradually to 93%. However, when K is more than 10, it gradually decreases. As the optimal value of K depends solely on the dataset, in case of this dataset, PSA was used to find the most suitable K value. With K=10, the PSA accuracy was over 91%. The results confirm 10 is the optimal amount of cluster for radish field clustering.

Figure 10. Radish field clustering results

Fig. 10 shows the clustering results with K=10. The regions corresponding to radish, bare ground, and mulching film are correctly classified as labeled. Although the clustering result works great on separating radish, bare ground and mulching film, misclassified regions are also represented in Fig. 11. These include withered radishes that are mainly brown as their resemblance to ground color; these regions were clustered to bare ground by K-means clustering. The low performance is caused by the clustering method, as K-means clustering depends on 3 color channels as described in section 3.2.

Figure 11. Wilt of radish clustering results (Brown leaves was misclassified as bare ground)

4.4. Fusarium Wilt of Radish Classification Results

Fig. 12 shows the results of Fusarium wilt of radish classification on different image sizes. Overall, the 128x128 image size dataset achieves the highest classification performance at over 90% for three classes. Taking a closer look into 128x128 window size, the model misclassifies two pairs of classes, the first pair is (normal, disease light), the second pair is (disease light, disease heavy), one possible reason is because the color value range threshold to distinguish between three classes was similar. Therefore it was unable to assign the right class for some images. Besides the 128x128 size dataset, the results for 64x64 and 256x256 window sizes are not as good as 128x128 size dataset, thus in this study, the 128x128 size dataset is deemed the optimal size.

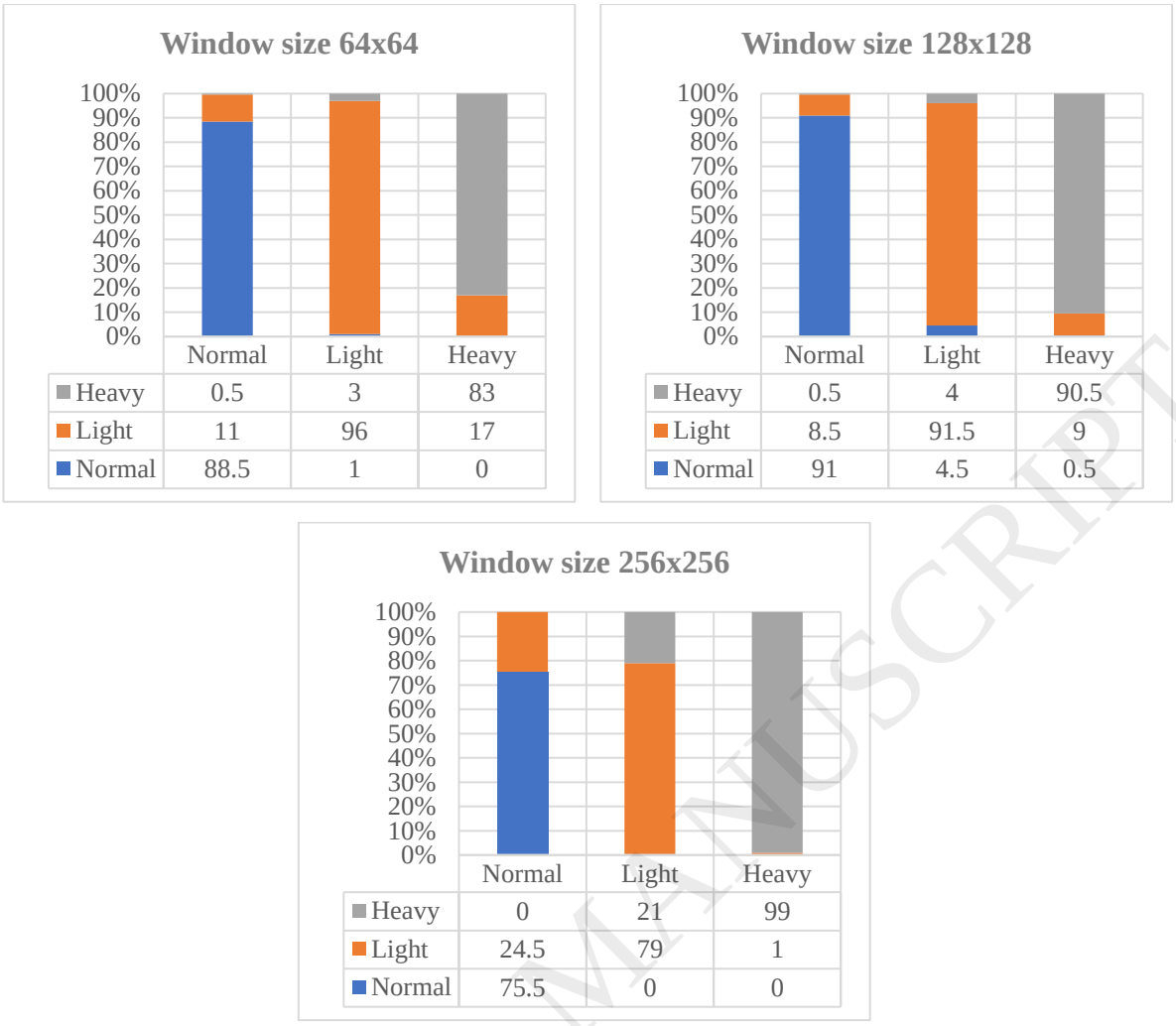

Figure 12. Fusarium wilt of radish classification results on different window sizes

Fig. 13 shows the Fusarium wilt of radish detection results on radish region images. Fig. 14 shows the detailed detection results of the Fusarium wilt of radish. The Fusarium wilt of radish grows from yellow to brown as the disease gets worse. As can be seen in Fig. 13, light-level Fusarium wilt of radishes (partly yellow) and heavy-level Fusarium wilt of radishes (mainly yellow) are successfully detected. However, some parts of the leaves are missing because browned radishes such as heavy-level Fusarium wilt of radish or dried radish are often clustered as bare ground in the clustering step. As the result, they were filtered out because only radish region images are extracted. This problem causes a reduction of Fusarium wilt of radish detection accuracy.

Figure 13. Extracted radish image (Left) and wilt of radish detection result in radish field (Right)

Figure 14. Close-range Fusarium wilt of radish detection

Conclusions

This study introduces an efficient framework to identify and detect different levels of Fusarium wilt on radish from healthy to disease heavy. Many techniques were implemented to improve the performance of the system such as extracted both color features and textures features from the image, clustered the field and selected only radish region. Finally, GoogLeNet can detect Fusarium wilt of radish with the accuracy of over 90%.

The model is capable of detecting the)XVDULXPZLOWRIUDGLVKIURP8$9¶LPDJHV which have a great potential for reducing the labor cost in managing and preventing the disease, as well as ensuring sustainable radish production. The proposed system can detect Fusarium wilt on another type of crops including tomato, tobacco, banana because it is a common vegetable disease. Therefore, the model plays an important role in maintaining the sustainability of crop yields.

In the future, several related issues will be studied. First, besides RGB image, infrared is often employed to monitor diseases, and the infrared filter is easy to equip on existing UAV¶ camera lens. A methodology that combines RGB images and infrared images will probably

improve the performance of the model. Secondly, several pre-processing techniques and deep learning models to improve the performance of this problem will be studied.

Acknowledgments

This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries(IPET) through Agri-Bio Industry Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs(MAFRA) (316033-04-2-338 SB030)

References

1. Mace, M., Fungal wilt diseases of plants. 2012: Elsevier.
2. Candiago, S., et al., Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. Remote Sensing, 2015. 7(4): p. 4026-4047.
3. Tang, L. and G. Shao, Drone remote sensing for forestry research and practices. Journal of Forestry Research, 2015. 26(4): p. 791-797.
4. Matese, A., et al., Intercomparison of UAV, aircraft, and satellite remote sensing platforms for precision viticulture. Remote Sensing, 2015. 7(3): p. 2971-2990.
5. Khan, M.A., et al., UAV-Based Traffic Analysis: A Universal Guiding Framework Based on Literature Survey. Transportation Research Procedia, 2017. 22: p. 541-550.
6. Yuan, C., Y. Zhang, and Z. Liu, A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. Canadian journal of forest research, 2015. 45(7): p. 783-792.
7. Qi, J., et al., Search and Rescue RotaryæWing UAV and Its Application to the Lushan Ms 7.0 Earthquake. Journal of Field Robotics, 2016. 33(3): p. 290-321.
8. Ren, D.D., S. Tripathi, and L.K. Li, Low-cost multispectral imaging for remote sensing of lettuce health. Journal of Applied Remote Sensing, 2017. 11(1): p. 016006-016006.
9. Ahmad, J., et al., Saliency-weighted graphs for efficient visual content description and their applications in real-time image retrieval systems. Journal of Real-Time Image Processing, 2015: p. 1-17.
10. Ahmad, J., et al., Multi-scale local structure patterns histogram for describing visual contents in social image retrieval systems. Multimedia Tools and Applications, 2016. 75(20): p. 12669-12692.
11. Lowe, D.G., Distinctive image features from scale-invariant key points. International journal of computer vision, 2004. 60(2): p. 91-110.
12. Wang, C., et al. Spatial weighting for bag-of-features based image retrieval. in International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making. 2013. Springer.
13. Krizhevsky, A., I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. in Advances in neural information processing systems. 2012.
14. Szegedy, C., et al. Going deeper with convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
15. Lin, M., Q. Chen, and S. Yan, Network in network. ICLR, 2014.
16. Ojala, T., M. Pietikainen, and T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on pattern analysis and machine intelligence, 2002. 24(7): p. 971-987.
17. Vincent, P., et al., Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. Journal of Machine Learning Research, 2010. 11(Dec): p. 3371-3408.
18. Simonyan, K. and A. Zisserman, Very deep convolutional networks for large-scale image recognition. CoRR, abs/1409.1556, 2014.
19. Russakovsky, O., et al., Imagenet large scale visual recognition challenge. International Journal of Computer Vision, 2015. 115(3): p. 211-252.
20. Csurka, G., et al. What is a good evaluation measure for semantic segmentation? in BMVC. 2013.
21. Li, Li, and Ken Choi. "Activity-driven optimised bus-specific-clock-gating for ultra-low-power smart space applications." IET communications 5.17 (2011): 2501-2508.
22. Dang, Minh, and Duc Duong. "Improvement methods for stock market prediction using financial news articles." Information and Computer Science (NICS), 2016 3rd National Foundation for Science and Technology Development Conference on. IEEE, 2016.
23. Chang, Chih-Chung, and Chih-Jen Lin. "LIBSVM: a library for support vector machines." ACM transactions on intelligent systems and technology (TIST) 2.3 (2011): 27.
24. Kieu, Tung, et al. "Mining top-k co-occurrence items with sequential pattern." Expert Systems with Applications 85 (2017): 123-133.
25. Cui, Zhihua, and Xiaozhi Gao. "Theory and applications of swarm intelligence." (2012): 205-206.
26. Udoyara Sunday Tim, and Sugam Sharma. Harnessing the Power of Big Data and Big Data Analytics in Agriculture: Opportunities and Challanges. 2014.
27. Chen, Chen, et al. "Latency estimation based on traffic density for video streaming in the internet of vehicles." Computer Communications 111 (2017): 176-186.
28. Qiu, Tie, et al. "A Robust Time Synchronization Scheme for Industrial Internet of Things." IEEE Transactions on Industrial Informatics (2017).
29. Samuel, Oluwarotimi Williams, et al. "An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction." Expert Systems with Applications 68 (2017): 163-172.

30. Liao, Xin, et al. "Medical JPEG image steganography based on preserving inter-block dependencies." Computers & Electrical Engineering (2017).
31. Zhang, Ruijie, et al. "Medical image classification based on multi-scale non-negative sparse coding." Artificial Intelligence in Medicine (2017).
32. Liang, Wei, et al. "SIRSE: A secure identity recognition scheme based on electroencephalogram data with multi-factor feature." Computers & Electrical Engineering (2017).
33. Samuel, Oluwarotimi Williams, et al. "Pattern recognition of electromyography signals based on novel time domain features for amputees' limb motion classification." Computers & Electrical Engineering (2017).
34. Sangaiah, Arun Kumar, et al. "Towards an efficient risk assessment in software projects±Fuzzy reinforcement paradigm." Computers & Electrical Engineering (2017).
35. Bi, Chujian, Haoxiang Wang, and Rui Bao. "SAR image change detection using regularized dictionary learning and fuzzy clustering." Cloud Computing and Intelligence Systems (CCIS), 2014 IEEE 3rd International Conference on. IEEE, 2014.
36. Chen, Qi, et al. "Single image shadow detection and removal based on feature fusion and multiple dictionary learning." Multimedia Tools and Applications (2017): 1-24.
37. Zhang, Shanwen, Harry Wang, and Wenzhun Huang. "Two-stage plant species recognition by local mean clustering and Weighted sparse representation classification." Cluster Computing (2017): 1-9.
38. Wu, Zeling, and Haoxiang Wang. "Super-resolution Reconstruction of SAR Image based on Non-Local Means Denoising Combined with BP Neural Network." arXiv preprint arXiv:1612.04755 (2016).