International Journal of New Computer Architectures and their Applications (IJNCAA) 9(1): 25-30
The Society of Digital Information and Wireless Communications, 2019 ISSN 2220-9085 (Online); ISSN 2412-3587 (Print)

# Automatic Identification of Plant Physiological Disorders in Plant Factories with Artificial Light Using Convolutional Neural Networks

Shigeharu Shimamura Kenta Uehara Seiichi Koakutsu
Chiba University
1-33, Inage-ku, Yayoi-cho, Chiba-shi 263-8522, Japan
s.shimamura@hanmo.jp

## ABSTRACT

Plant factories with artificial light (PFAL) are attracting worldwide attention as a technology for stably producing crops. One of the problems of PFAL is tipburn which is a physiological disorder of crops. Tipburn is a phenomenon in which plant growth point cells are necrotized. Lettuce cultivated in PFAL in particular has a high frequency of tipburn. When tipburn occurs, its identification is done by human eye observation, and tipburn leaves are trimmed by hand or tipburn lettuce is removed from products. These operations require much labor and cost. If tipburn identification can automatically be done using machine learning, the economic effect will be great and it will be a driving force for spreading PFAL. In this study, we aim to perform binary discrimination of tipburn occurrence and its non-occurrence about lettuce cultivated in PFAL using machine learning with convolutional neural networks. In particular, we aim to recognize the symptom of tipburn which means the early stages of tipburn immediately before leaf tips discolor blackly and the commercial value as the vegetables is damaged. The results of the experiments indicate that the recognition of the symptom of tipburn can be performed with high accuracy.

## KEYWORDS

machine learning, convolutional neural network, diagnostic imaging, plant factory, plant physiological disorder, tipburn

## 1 INTRODUCTION

In recent years, smart agriculture, which is a new agriculture that promotes labor-saving and high-quality production by utilizing robot technology, information and communication technology (ICT), and artificial intelligence (AI) technogogy, is regarded as important, because farmers are rapidly aging and labor shortages are becoming serious in Japan's agricultural field. By utilizing smart agriculture, labor saving and lightening can be promoted in farm work, and it is expected to secure new farmers and inherit cultivation technology. In particular, plant factories with artificial light (PFAL) which is one of key technologies for smart agriculture are paid much attention not only in horticuture but also in engineering. PFAL enable to stably produce crops in a closed space where environmental conditions are highly controlled[1].

One of the problems of PFAL is tipburn which is a physiological disorder of crops. In PFAL, the growth of crops is promoted by creating the condition suitable for crops and the period from seeding to harvesting can be reduced, whereas PFAL tends to cause tipburn. Tipburn is a phenomenon in which plant growth point cells are necrotized[2][3][4]. Especially, lettuce cultivated in PFAL has a high frequency of tipburn. When tipburn occurs, leaf tips discolor blackly and the commercial value as vegetables is damaged. Identification of tipburn is done by human eye observation, and tipburn leaves are trimmed by hand or that lettuce is removed from products. These operations require much labor and cost. If tipburn identification can automatically be done using machine learning, economic effect is great and it will be a driving force for spreading PFAL.

Tipburn can be classified into two types, A and B, in case of frill lettuce. Type A tipburn occurs at central leaves, whereas type B tipburn occurs at peripheral ones. Image diagnosis by machine learning using convolutional neural networks (CNNs), which have excellent performance in image recognition, is effective for automatic identification of the occurrence of type A or type B tipburn. We have performed binary discrimination of type B tipburn occurrence and its non-occurrence with accuracy 0.929[5], and binary discrimination of type A tipburn occurrence and its non-occurrence with accuracy 0.986[6] using machine learing with CNNs. The accuracy exceeds target yield (90%) of farm products of PFAL.

In this study, we aim to perform binary discrimination of the symptom of type A tipburn occurrence and its non-occurrence about lettuce cultivated in PFAL using machine learing with CNNs[**?**]. The symptom of tipburn means the early stages of tipburn immediately before leaf tips discolor blackly and the commercial value as vegetables is damaged. Lettuce irreversibly transits from normal to tipburn through a stage of symptom. It is important to detect the symptom of tipburn because lettuce has commercial value and can be sold at the stage of symptom. In addition, we adopt support vector machine (SVM)[7], which is a representative binary discrimination method, as a comparison method, and conduct comparison experiments among various CNN models and SVM. The results of the experiments show that CNNs are effective for binary discrimination of the symptom of type A tipburn occurrence and its non-occurrence.

## 2 TRAINING DATA

Machine learning requires a large number of training data which include normal lettuce images, symptom ones, and tipburn ones. To obtain the training data, we cultivate normal lettuce, symptom one and tipburn one on purpose. Data labels obtained in this manner are considered highly accurate.

### 2.1 Test crop and cultivation method

Frill lettuce is one of the crops with the biggest production in PFAL and its economic effect is big. Besides frill lettuce is known as one of crops which have a high frequency of tipburn. Frill ice (Snow Brand Seed Co. LTD.) which is one of frill lettuce is adopted as a test crop.

We adopt two type of PFAL as cultivation systems. One is showcase-type PFAL for normal lettuce, and another is house-type PFAL for symptom one and tipburn one. Both type of PFAL are located at center for environment, health and field sciences in Chiba University. Showcase-type PFAL (HM-PF-DC-AL01, Hanmo Co.) is compact size and easy to control cultivation condition because it installed in an interior and doesnot undergo influence of an outdoor weather fluctuation. House-type PFAL is also easy to control cultivation condition such as temperater and $CO_2$ level. Temperature and $CO_2$ level are very imporadant factors to create the clutivation conditions which promote tipburn.

**Table 1.** System specification (showcase-type FPAL).

| | |
|---|---|
| ♯cultural shelves | 3 (seed bed:1, cultivation:2) |
| controller | lighting cycle, fan control |
| lighting | white LED (18W) 6 |
| ventilator | 8 fans |
| nutrient solution dispenser | circulator, tank, water level sensor |

**Table 2.** System specification (house-type FPAL).

| | |
|---|---|
| ♯cultural shelves | 4 |
| controller | lighting cycle, air conditioner, $CO_2$ control |
| lighting | RGB LED (14.5W) 10 |
| nutrient solution dispenser | deep float technique, aeration |

Hydroponic culture is adopted as a cultivation method. White LED light in showcase-type FPAL and RGB LED light in house-type FPAL are adopted as artificial lights, because both LED lights contaion wave length which is suitable for photosynthesis, and are low heat generation. In showcase-type PFAL, lighting heat is removed with room air using ventilation fans. In house-type PFAL, air conditoner is used to remove lighting heat and control room air temperature. In addition, $CO_2$ level is raised in house-type PFAL in order to promote the growth of lettuce and induce tipburn.

Table 1 and 2 show the system specification of showcase-type FPAL and that of house-type FPAL respectively.

### 2.2 Cultivation condition

In general, tipburn is caused by (1) high temperature, (2) low humidity, (3) high light intensity, (4) long day length, and (5) high nutrient concentration. Parameters for cultivating tipburn lettuce are estimated and tuned in consideration of these factors.

**Table 3.** Cultivation conditions.

| | normal lettuce | tipburn lettuce |
|---|---|---|
| cultivation system | showcase-type | house-type |
| temperature [°C] | 18 | 24 |
| humidity [%] | 65∼75 | 65∼75 |
| day length [h] | L15-D9 (15h) | L20-D4 (20h) |
| light intensity [$\mu$mol/m$^2$/s] | PPFD200 | PPFD200∼250 |
| nutrient concentration [dS/m] | EC1.5 | EC2.4 |

In both type of FPAL, enshi formula nutrient solution which is the most commonly used for hydroponic culture is adopted as nutrient solution.
In showcase-type FPAL, nutrient solution is circulated by a pump from a tank to culture racks 24 hours a day. The flow of nutrient solution is 10 L/min for each culture rack. Ventilation fans operate 18 hous a day. Other conditions are shown in Table 3.
In house-type FPAL, deep float technique with 24-hours aeration is adopted as nutrient solution dispenser. $CO_2$ level is raised in order to increase photosynthesis and promote growth of lettuce. Liquefied $CO_2$ gas cylinder is installed in the house-type PFAL, and $CO_2$ level is controlled to be 2,000 ppm using $CO_2$ sensor. Other conditions are shown in Table 3.
In Table 3, light intensity is meared with photosynthetic photon flux density (PPFD [$\mu$mol/m$^2$/s]), and nutrient concentration is measured with electrical conductivity (EC [dS/m]).

### 2.3 Cultivation of lettuce

Lettuce is cultivated as follws. (1) Urethane sheet soaked in tap water is installed in a container designed for raising of seedling, and lettuce seeds are sown there. (2) Lettuce checked on root is transplanted to growing racks three days after sprouting. (3) Lettuce is settled planting in culture racks two weeks after transplant, and cultivated.

### 2.4 Acquisition of image data

Images of all 24 lettuces are obtaind daily from 35th day after seeding. Indivisual lettuce is checked occurrence or non-occurrence of tipburn by human eye observation, and captured (F/4.0) from above under white lighting and white backgraound with digital camera. Fig. 1 shows an example of normal lettuce, Fig. 2 shows that of symptom of type A tipburn one, and Fig. 3 shows that of type A tipburn one. In Fig. 3, the central leaf tips discolored blackly are tipburn.

**Figure 1.** An example of normal lettuce.

**Figure 2.** An example of symptom of type A tipburn lettuce.

**Figure 3.** An example of type A tipburn lettuce.

## 3 EXPERIMENTS

We examine binary discrimination of the symptom of type A tipburn occurrence and its non-occurrence about lettuce cultivated in PFAL using machine learning with CNNs. CNNs are suitable for these experiments because CNNs are effective in processing of two-dimensional images.

VGGNet[8], GoogLeNet[9], ResNet[10], and WideResNet[11] are adopted as CNN models. VGGNet is a simple CNN model composed of convolution layers and pooling layers. In VGGNet, a small-scale convolution layer is stacked between the input layer or pooling layers to reduce the number of parameters. Table 4 shows the network structure of VGGNet. GoogLeNet is a CNN model that

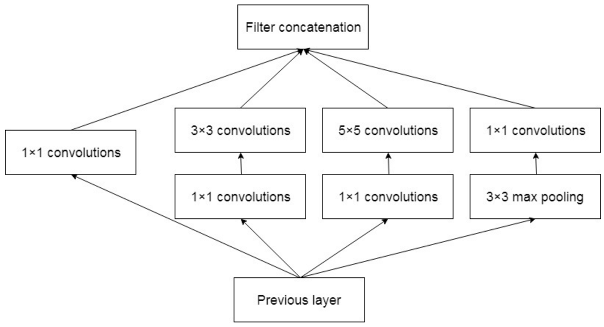


**Figure 4.** Inception module.

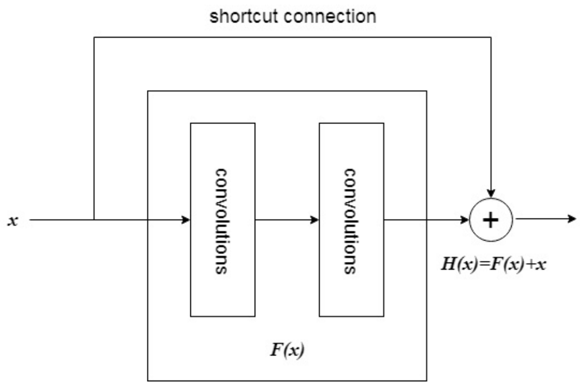


**Figure 5.** Residual module.

constitutes one large CNN by introducing inception modules in which multiple convolution layers and pooling layers are arranged in parallel, and stacking them like a normal convolution layer. Fig. 4 shows the inception module, and Table 5 shows the network structure of GoogLeNet. ResNet is a CNN model that introduces residual modules. In residual modules, instead of simply passing the transformation $F(x)$ by a processing block to the next layer, the input $x$ to that processing block is shortcut and $H(x) = F(x) + x$ are passed to the next layer. Fig. 5 shows the residual module, and Table 6 shows the network structure of ResNet. WideResNet is a CNN model that increases the input channels of the convolution layer in each residual module instead of deepening the layer in ResNet. Table 7 shows the network structure of WideResNet.

We also examine the same binary discrimination with classical networks in order to exhibit the difficulty of discriminating tipburn lettuce from normal one. Support vector machine (SVM)[7] with histogram of oriented gradients (HOG)[12] is adopted as one of classical networks. SVM is a classifier used in machine learning for the purpose of classification and regression, and HOG is a feature descriptor used in computer vision and image processing for the purpose of object detection. SVM is suitable for our experiments because SVM is effective in binary discrimination.

**Table 4.** The network structure of VGGNet17

| layer | filter | stride | output |
|---|---|---|---|
| input | | | (224, 224, 3) |
| convolution×2 | (3, 3) | (1, 1) | (224, 224, 64) |
| max pooling | (2, 2) | (2, 2) | (112, 112, 64) |
| convolution×2 | (3, 3) | (1, 1) | (112, 112, 128) |
| max pooling | (2, 2) | (2, 2) | (56, 56, 256) |
| convolution×3 | (3, 3) | (1, 1) | (56, 56, 512) |
| max pooling | (2, 2) | (2, 2) | (28, 28, 512) |
| convolution×3 | (3, 3) | (1, 1) | (28, 28, 512) |
| max pooling | (2, 2) | (2, 2) | (14, 14, 512) |
| convolution×3 | (3, 3) | (1, 1) | (14, 14, 512) |
| max pooling | (2, 2) | (2, 2) | (7, 7, 512) |
| full connected | | | 4096 |
| full connected | | | 4096 |
| full connected | | | 1000 |
| full connected | | | 2 |

**Table 5.** The network structure of GoogLeNet22

| layer | filter | stride | output |
|---|---|---|---|
| input | | | (224, 224, 3) |
| convolution | (7, 7) | (2, 2) | (112, 112, 64) |
| max pooling | (3, 3) | (2, 2) | (56, 56, 64) |
| convolution | (3, 3) | (1, 1) | (56, 56, 192) |
| max pooling | (3, 3) | (2, 2) | (28, 28, 192) |
| inception×2 | | | (28, 28, 480) |
| max pooling | (3, 3) | (2, 2) | (14, 14, 480) |
| inception×5 | | | (14, 14, 832) |
| max pooling | (3, 3) | (2, 2) | (7, 7, 832) |
| inception×2 | | | (7, 7, 1024) |
| avg. pooling | (7, 7) | (1, 1) | (1, 1, 1024) |
| full connected | | | 256 |
| full connected | | | 2 |

### 3.1 Condition

The size of image is 100KB. The number of normal lettuce images, that of symptom ones, and that of tipburn ones are 1200, 1500 and 1100 respectively. GoogLeNet is traind using Adam[13], and the number of training epoch is 50. 90% of image data are used for training, and rest for validation.

### 3.2 Results

Experimental results are indicated in Table 8 and Fig. 6∼8. Table 8 shows the comparison of various CNN models with SVM in terms of accuacy of test data in case of binary discrimination of normal and tipburn, that of normal and symptom, and that

**Table 6.** The network structure of ResNet50

| layer | filter | stride | output |
|---|---|---|---|
| input | | | (224, 224, 3) |
| convolution | (3, 3) | (1, 1) | (224, 224, 17) |
| residual | $\begin{bmatrix} 1\times 1, 17 \\ 3\times 3, 4 \\ 1\times 1, 17 \end{bmatrix}\times 3$ | (1, 1) | (224, 224, 17) |
| residual | $\begin{bmatrix} 1\times 1, 17 \\ 3\times 3, 8 \\ 1\times 1, 32 \end{bmatrix}\times 4$ | (2, 2) | (112, 112, 32) |
| residual | $\begin{bmatrix} 1\times 1, 32 \\ 3\times 3, 17 \\ 1\times 1, 64 \end{bmatrix}\times 6$ | (2, 2) | (56, 56, 64) |
| residual | $\begin{bmatrix} 1\times 1, 64 \\ 3\times 3, 32 \\ 1\times 1, 128 \end{bmatrix}\times 3$ | (2, 2) | (28, 28, 128) |
| avg. pooling | (28, 28) | (1, 1) | (1, 1, 128) |
| full connected | | | 2 |

**Table 7.** The network structure of WideResNet10

| layer | filter | stride | output |
|---|---|---|---|
| input | | | (224, 224, 3) |
| convolution | (3, 3) | (1, 1) | (224, 224, 17) |
| residual | $\begin{bmatrix} 3\times 3, & 17 \\ 3\times 3, & 48 \end{bmatrix}$ | (1, 1) | (224, 224, 48) |
| residual | $\begin{bmatrix} 3\times 3, & 48 \\ 3\times 3, & 96 \end{bmatrix}$ | (2, 2) | (112, 112, 96) |
| residual | $\begin{bmatrix} 3\times 3, & 96 \\ 3\times 3, & 192 \end{bmatrix}$ | (2, 2) | (56, 56, 192) |
| residual | $\begin{bmatrix} 3\times 3, & 192 \\ 3\times 3, & 384 \end{bmatrix}$ | (2, 2) | (28, 28, 384) |
| avg. pooling | (28, 28) | (1, 1) | (1, 1, 384) |
| full connected | | | 2 |

**Table 8.** Comparison of test accuracy.

| | WRN | RN | GN | VGG | SVM |
|---|---|---|---|---|---|
| nor.-tipb. | 1.000 | 0.993 | 0.986 | 0.977 | 0.734 |
| nor.-symp. | 0.990 | 0.981 | 0.977 | 0.978 | 0.536 |
| symp.-tipb. | 1.000 | 0.995 | 0.970 | 0.870 | 0.647 |

of symptom and tipburn respectivly. Fig.6∼8 are examples of learning curves of various CNN models. $x$-axis indicates learning epochs and $y$-axis indicates test accuracy.

CNNs are better than SVM in all cases, and test accuracy of CNNs exceeds target yield (90%) of farm products of PFAL in almost all of the cases. Type A tipbarn and its symptom can be recognized with a high degree of accuracy sufficiently using various CNNs.

### 3.3 Discussion

The results in Table 8 indicate the difficulty of discriminating tipburn lettuce from normal one. The classical network such as SVM cannot distinguish tipburn lettuce from normal one with high degree of accuracy because SVM cannot accurately capture the features of tipburn in the image data.

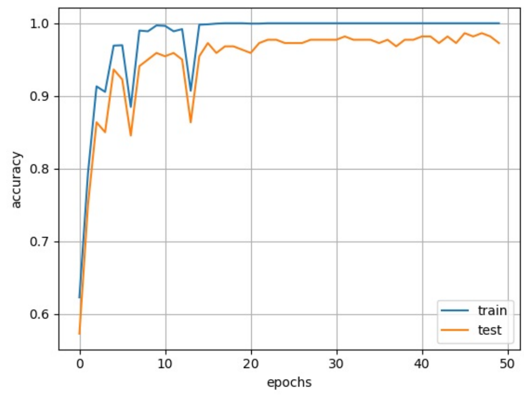


**Figure 6.** Accuracy (normal-tipburn).

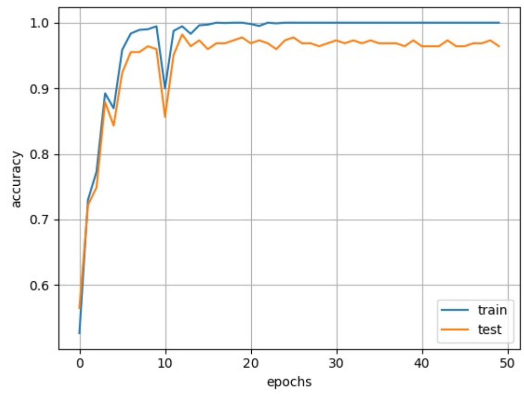


**Figure 7.** Accuracy (normal-symptom).

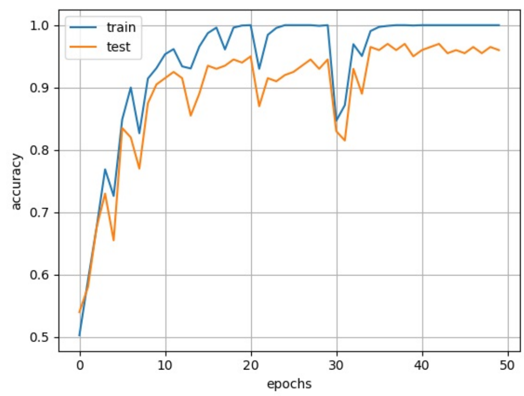


**Figure 8.** Accuracy (symptom-tipburn).

The results of CNNs indicate binary discrimination of symptom and tipburn is the most difficult, that of normal and symptom is the second most difficult, and that of normal and tipburn is the easyest. The performance of WideResNet is the best, that of

ResNet is the second best, that of GoogLeNet is the third best and VGGNet is the worst.

The reasons of the high degree of accuracy of CNNs are the quality of training image data and the correctness of data labels which indicate normal, symptom, and tipburn. The features of tipburn sufficiently appear in the image data to be captured by CNNs.

The reasons of false recognition of CNNs are described as follows. In case of normal lettuce is falsely recognized with tipburn one, the shadowed area of leaves is falsely recognized with tipburn. On the other hand, in case of tipburn lettuce is falsely recognized with normal one, the features of tipburn in the image data is so slight that CNNs cannot recognise tipburn correctly.

## 4 CONCLUSION

Binary discrimination of the symptom of type A tipburn occurrence and its non-occurrence could be done by machine learning using CNNs by the high precision as accuracy more than 0.9 about lettuce cultivated in PFAL. It would be commercially meaningful if we could harvest lettuce at the symptom stage of tipburn because lettuce has the commercial value as vegetables at this stage. To test other training algorithms can be future works. Besides in the real PFAL environment, noise on the image, the camera angle and the value of illuminance when the taking a sample photo to recognize should be considered.

## REFERENCES

[1] T. Kozai: “Plant Factory with Artificial Light,” pp. 193-212, Ohmsha (2012) (in Japanese).

[2] T. Maruo: “Resarch about Lettuce Tip Burn. (1)Ca Concentration of Individual Lettuce Leaves by Leaf Position,” Hort. Res.(Japan), Vol. 58 (Sppl.2), pp. 193-212 (1983) (in Japanese).

[3] T. Maruo et.al.: “Which Stage the Tip Burn Occurrence Is Determined for Individual Lettuce Leaf?,” Hort. Res.(Japan), Vol. 18 (Sppl.1), pp. 152 (2019) (in Japanese).

[4] E. Goto and T. Takakura: “Reduction of Lettuce Tipburn by Shorting Day/Night Cycle,” J. Agric. Meteorol., Vol. 59, No. 3, pp. 219-225 (2003).

[5] S. Shimamura, K. Uehara, and S. Koakutsu: “Automatic Identification of Plant Physiological Disorders in Plant Factory Crops,” Proc. IEEJ EIS Conf., pp. 1207-1210 (2018) (in Japanese).

[6] S. Shimamura, K. Uehara, and S. Koakutsu: “Automatic Identification of Lettuce Tipburn in Plant Factory Crops,” Proc. Fuzzy, Artificial Intelligence, Neural Networks and Computational Intelligence Symposium (FAN2018), pp. 101-104 (2018) (in Japanese).

[7] I. Tsochantaridis, T. Joachims, T. Hofmann and Y. Altun: “Large Margin Methods for Structured and Interdependent Output Variables,” J. of Machine Learning Research, No.6, pp.1453-1484 (2005).

[8] K. Simonyan and A. Zisserman: “Very Deep Convolutional Networks for Large-scale Image Recognition,” Proc. Int. Conf. on Learning Representations, pp.1-14 (2015).

[9] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich: “Going Deeper with Convolutions,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1-9 (2015).

[10] K. He, X. Zhang, S. Ren and J. Sun: “Deep Residual Learning for Image Recognition,” IEEE Conf. on Computer Vision and Pattern Recognition, pp.770-778 (2016).

[11] S. Zagoruyko and N. Komodakis: “Wide Residual Networks,” arXiv preprint arXiv:1605.07146 (2016).

[12] N. Dalal and B. Triggs: “Histograms of Oriented Gradients for Human Detection,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp.886-893 (2005).

[13] D. P. Kingma and J. L. Ba: “Adam: A Method for Stochastic Optimization,” Proc. Inter. Conf. on Learning Representations, pp. 1-13 (2015).