SCIENTIFIC REPORTS

OPEN

Rice Blast Disease Recognition Using a Deep Convolutional Neural Network

Received: 10 April 2018
Accepted: 8 January 2019
Published online: 27 February 2019

Wan-jie Liang[1,2], Hong Zhang[2], Gu-feng Zhang[3] & Hong-xin Cao[1]

Rice disease recognition is crucial in automated rice disease diagnosis systems. At present, deep convolutional neural network (CNN) is generally considered the state-of-the-art solution in image recognition. In this paper, we propose a novel rice blast recognition method based on CNN. A dataset of 2906 positive samples and 2902 negative samples is established for training and testing the CNN model. In addition, we conduct comparative experiments for qualitative and quantitatively analysis in our evaluation of the effectiveness of the proposed method. The evaluation results show that the high-level features extracted by CNN are more discriminative and effective than traditional hand-crafted features including local binary patterns histograms (LBPH) and Haar-WT (Wavelet Transform). Moreover, quantitative evaluation results indicate that CNN with Softmax and CNN with support vector machine (SVM) have similar performances, with higher accuracy, larger area under curve (AUC), and better receiver operating characteristic (ROC) curves than both LBPH plus an SVM as the classifier and Haar-WT plus an SVM as the classifier. Therefore, our CNN model is a top performing method for rice blast disease recognition and can be potentially employed in practical applications.

Rice as a food source provides protein and energy to more than half of the world's population[1]. Moreover, rice consumption and demand are increasing with the growth of the population. To meet the increased food demand, rice production must be increased by more than 40% by 2030[2]. Unfortunately, rice diseases have caused a great deal of loss in yield, and rice blast disease is considered as one of the main culprits[3], reducing yield by between 60% and 100%[4]. Currently, the use of pesticides and deployment of blast-resistant cultivars are the main methods of combating the disease[5]. However, excessive use of pesticides not only increases the cost of rice production but also causes considerable environmental damage[6]. Moreover, in practice, diagnosis of rice blast is often manually conducted and this is subjective and time-consuming even for well-experienced experts. In modern agricultural practices, it is very important to manage pests and diseases using highly efficient methods with minimum damage to the environment[7]. In recent decades, combined with crop images, computer-aided diagnostic methods have become dominant for monitoring crop diseases and pests[8–10]. An automated rice disease diagnostic system could provide information for prevention and control of rice disease, set aside time for disease control, minimize the economic loss, reduce the pesticide residues, and improve the quality and quantity of agricultural products. In order to achieve such a system, research in effective algorithms of feature extraction and classification of rice disease is critical.

Currently, there exists no public dataset for rice blast disease classification. To fill this void, we establish in this work a rice blast disease dataset and use it for training and testing a disease classification model, based on convolutional neural network (CNN). The rice blast disease images are obtained from the Institute of Plant Protection, Jiangsu Academy of Agricultural Sciences, Nanjing, China. These images are captured in a naturally-lit environment while plant protection experts conduct field investigation. As a result, the trained CNN model on the dataset can be expected to have direct applicability. At the same time, the dataset is useful for other people who are interested in rice or even crop disease classification research.

In recent years, due to its ability to extract good features, CNN has been employed extensively in machine learning and pattern recognition research[11–17]. Hinton *et al.*[18] stated that a multi-layer neural network has excellent learning ability, and that the learned features can abstract and express raw data conveniently for classification.

[1]Institute of Agricultural Information, Jiangsu Academy of Agricultural Sciences, Nanjing, 210014, China. [2]Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada. [3]Institute of Plant Protection, Jiangsu Academy of Agricultural Sciences, Nanjing, 210014, China. Correspondence and requests for materials should be addressed to W.-j.L. (email: wanjie.liang@163.com)

Figure 1. Example images of the rice blast disease dataset: (**a**) positive samples and (**b**) negative samples.

CNN provides an end-to-end learning solution that avoids image pre-processing, and extracts relevant high-level features directly from raw images. The CNN architecture was inspired by the visual cortex of cats in Hubel's and Wiesel's early work[19]. In particular, Krizhevsky[20] performed object classification and won the first place in the ImageNet Large Scale Visual Recognition Challenge 2012 using a deep CNN. This is followed by the emergence of many improved algorithms and applications of CNN[21–23]. Since[20], similar CNN architectures have been successfully developed to solve a variety of image classification tasks.

With full consideration of CNN's excellent performance, we propose a method that uses CNN for rice blast image feature extraction and disease classification, and we are able to obtain remarkable performance through fine tuning the structure and the parameters of a CNN model. We conduct comparative experiments for rice blast disease recognition with two traditional feature extraction methods, LBPH and Haar-WT. As well, we combine an SVM classifier with the deep features extracted from the CNN to further investigate and verify the effectiveness of deep features of CNN.

The major contributions of this paper are summarized as follows. First, we introduce a rice blast disease dataset with the assistance of plant protection experts. The dataset is used to train and verify our model. The dataset is useful for other researchers who are interested in rice or even crop disease recognition. The dataset is available from the, http://www.51agritech.com/zdataset.data.zip. Second, we propose an effective rice blast feature extraction and classification method using CNN. The evaluation results show that the high-level features extracted by the CNN are more discriminative than LBPH and Haar-WT, with classification accuracies above 95%.

The remainder of this paper is organized as follows. Section 2 describes the dataset and the feature extraction and the rice blast disease classification methods. Section 3 describes the evaluation criteria of the feature extraction and recognition methods. The experiments and results are also provided and discussed in this section. Finally, the conclusions and future work are given in Section 4.

Rice Blast Disease Dataset and Proposed Classification Method

Dataset. Rice images with rice blast disease are obtained from the Institute of Plant Protection, Jiangsu Academy of Agricultural Sciences, Nanjing, China. The Institute mainly conducts research in the mechanism and the technologies in controlling the disease and insect pests of such crops as rice, wheat, cotton, rape, fruit and vegetables in Jiangsu Province and across China. To avoid duplicates and ensure label quality, each image in our dataset is examined and confirmed by plant protection experts. There is no special requirement for rice blast disease images and their pixels, and no special preprocessing is done. All the rice blast images are patches of 128×128 pixels in size, extracted from original larger images with a moving window of a stride of 96 pixels. Then, the patches containing rice blast lesions are identified by domain experts and used as positive samples, and patches without lesions are used as negative samples. The final dataset includes 5808 image patches of which 2906 are positive and 2902 negative. Some positive and negative samples are shown in Fig. 1. In addition to scale, rotation, illumination and partial viewpoint changes, the dataset also has the following characteristics. First, the background of rice canopy texture, water body, and soil can cause great difficulty to recognition, as do dead leaves and other plant lesion. Second, rice blast lesion shape and location are not predictable. Overall, the combination of above factors poses significant challenges for rice blast disease recognition.

Feature extraction from rice blast images. Feature extraction is a key step in object recognition. It requires the features to be sufficiently discriminating to be able to separate the different object classes while retaining invariant characteristics within the same class. Feature extraction is also a dimension reduction process for efficient pattern recognition and machine learning in image analysis. In this work, CNN, Harr-wavelet and LBPH feature extraction methods are employed and compared to process rice blast images.

The CNN model. CNN[24] is a multi-layer neural network with a supervised learning architecture that is often made up of two parts: a feature extractor and a trainable classifier. The feature extractor contains feature map layers and retrieves discriminating features from the raw images via two operations: convolutional filtering and down sampling[25]. Convolutional filtering as the key operation of CNN has two vital properties: local receptive field and shared weights. Convolutional filtering can be seen as a local feature extractor used to identify the relationships between pixels of a raw image so that the effective and appropriate high-level features can be extracted to enhance the generalization ability of a CNN model[26]. Furthermore, down sampling and weight sharing can greatly reduce the number of trainable parameters and improve the efficiency of training. The classifier and the weights learned in the feature extractor are trained by a back-propagation algorithm.

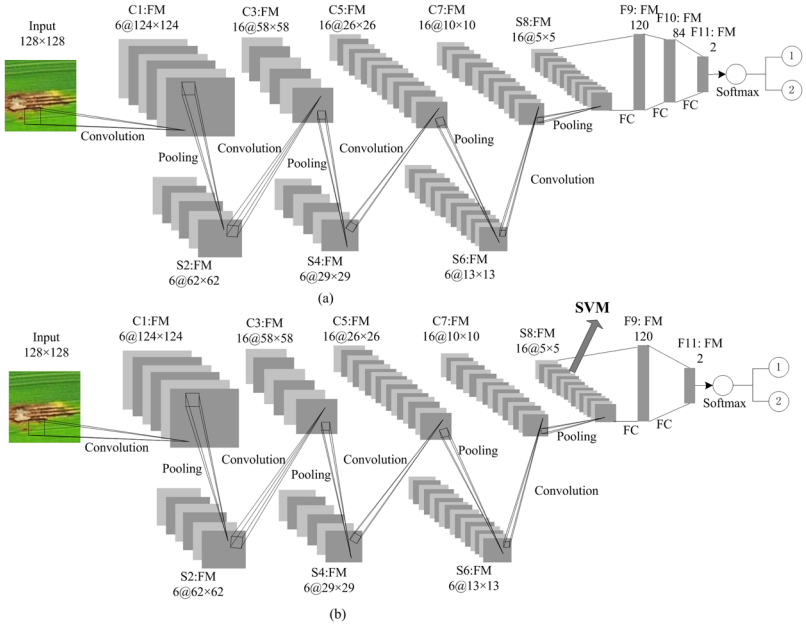

Figure 2. Structure of the two CNN models.

A convolutional layer computes feature maps by applying convolution kernels to input data followed by an activation function as follows[27]:

$$y_j^l = f\left(z_j^l\right) \tag{1}$$

$$z_j^l = \sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \tag{2}$$

where, y_j^l is the output feature maps at layer l; $f(\cdot)$ is the activation function (commonly used functions include sigmoid, tanh, and ReLU, etc., of which ReLU was chosen); z_j^i is the activation of the j channel at layer l; x_i^{l-1} is the feature maps of the $l-1$ layer; M_j is the subset of input feature maps; k_{ij}^l is convolution kernel matrix at layer l; * is the convolution operation; and b_j^l is the offset. For a more detailed explanation of convolutional neural networks, we refer the reader to LeCun *et al.*[24] and Krizhevsky *et al.*[20].

In this study, two network structures similar to Lenet5 (LeCun *et al.*)[24] are established. As shown in Fig. 2, the first network contains four convolutional layers, four max-pooling layers, and three fully connected layers, and ReLU is added after each layer (Fig. 2(a)). The second network has the same convolutional layers and max-pooling layer structure as the first network, but has two fully connected layers (Fig. 2(b)). To avoid over-fitting, one spatial dropout layer is added after the C5 layer for the both models, and another dropout layer is added after the F10 layer for the first model and after the F9 layer for the second model, respectively. The related parameters of CNN are shown in Fig. 2.

The models are implemented using Torch7 which is a scientific computing framework. The main steps of the second model are shown in Fig. 3. Stochastic gradient descent (SGD) is employed for training, and the number of training epochs is 150. Other training parameters are as shown in Fig. 3.

Comparative experiments are conducted for the two CNN models, and classification accuracy are computed. To reduce possible biases in the selection of the validation set, 5-fold cross-validation is employed. In 5-fold cross-validation, the original sample is randomly partitioned into five equal size subsamples. Of the five subsamples, a single subsample is retained as the validation set, and the other four subsamples are used for training. The cross-validation process is then repeated five times, and the results are averaged. As shown in Table 1, there is no obvious performance improvement of the first CNN model with more connected layers. In order to ensure that there is no over-fitting, the learning curves are generated. Here, 10% of the original samples are reserved as a test set, and 500 samples are randomly selected from the remaining dataset as training samples at starting point. By increasing 500 samples for training incrementally, we repeat the training process ten times in each step. The classification accuracy of training set and validation set are averaged, and the learning curves are obtained (Fig. 4). It can be seen that the two models have low bias and variance, good convergence, and high accuracy, and that there is no over-fitting. However, the stability of the first model is poor with small samples. Therefore, the second CNN model is chosen in the remainder of this study.

Haar-WT. Haar-WT is chosen as a competing hand-crafted feature in our evaluation. Haar-WT is an extension of the wavelet transform to simplify computation, and it is commonly used in image feature extraction. Haar-WT is a multi-resolution approach for image texture analysis[28] that employs two important functions of WT: the high

	1th model	2th model
Accuracy	95.37	95.83

Table 1. The 5-fold cross-validation results of the two CNN models.

```
.....
net = nn.Sequential()
net:add(nn.SpatialConvolution(3, 6, 5, 5))
net:add(nn.ReLU())
net:add(nn.SpatialMaxPooling(2,2,2,2))
net:add(nn.SpatialConvolution(6, 6, 5, 5))
net:add(nn.ReLU())
net:add(nn.SpatialMaxPooling(2,2,2,2))
net:add(nn.SpatialConvolution(6, 16, 4, 4))
net:add(nn.ReLU())
net:add(nn.SpatialDropout(0.3))
net:add(nn.SpatialMaxPooling(2,2,2,2))
net:add(nn.SpatialConvolution(16, 16, 4, 4))
net:add(nn.ReLU())
net:add(nn.SpatialMaxPooling(2,2,2,2))
net:add(nn.View(16*5*5))
net:add(nn.Linear(16*5*5, 120))
net:add(nn.Dropout(0.5))
net:add(nn.Linear(120, 2))
net:add(nn.LogSoftMax())
net:cuda()
criterion = nn.ClassNLLCriterion():cuda()
net:training()
trainer = nn.StochasticGradient(net, criterion)
trainer.learningRate = 0.001
trainer.learningRateDecay=0.9
trainer.maxIteration = 150
.....
```

Figure 3. The main code of the second model.

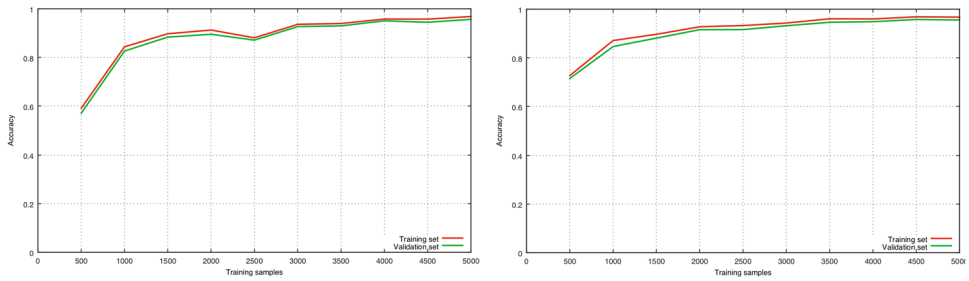

Figure 4. Learning curves of the two CNN models: (**a**) for the first model and (**b**) for the second model.

	3th level	4th level	5th level
Train accuracy	83.74	83.8	70.29
Test accuracy (c:2.0,g: 0.03125)	75.78	83.85	69.06

Table 2. The 5-fold cross-validation results of the 3th, 4th and 5th layer Haar-WT.

	1×1	2×2	4×4
Train accuracy	83.73	82.7	80.83
Test accuracy (c:8.0,g: 0.125)	82.59	80.08	51.82

Table 3. The 5-fold cross-validation result of three LBPH feature extraction method.

pass filter and the high pass filters[29]. At each level, a 2D-image is processed through low pass and high pass filters, separately. The result includes four sub level images which are one sub level of approximation of the original image (LL) and three sub levels of detail in horizontal, vertical and diagonal directions, respectively (LH, HL and HH). This process is called one level decomposition. With repeated decomposition on the approximation sub level, more sub level decomposition of an image can be obtained. The low pass filtering and high pass filtering of Haar-WT are computed as follows[30]:

$$A_i = \frac{x_i + x_{i+1}}{2},\ i \in [1, N] \tag{3}$$

$$wc_i = \frac{x_i - x_{i+1}}{2},\ i \in [1, N] \tag{4}$$

where x_i and x_{i+1} are two adjacent elements, A_i is the low-pass filtering, wc_i is the high-pass filtering, and N is the number of elements along row and column of input 2D data.

In this study, we perform Haar-WT decomposition of the rice blast image in the RGB color space. The decomposition is done up to level 5, and approximation sub levels are integrated as a single feature vector on each level. The feature vectors of 3rd, 4th and 5th level of decomposition are obtained. Using SVM as the classifier, comparative experiments are conducted, and classification accuracy is computed via 5-fold cross validation (Table 2). It could be seen that the 4th level obtains higher classification accuracy than any of the other levels. Therefore, the fourth level is chosen in our study as the Haar-WT feature, and compared with CNN.

LBPH. Local Binary Pattern Histograms (LBPH) is chosen as the second competing hand-crafted feature in our study. The LBP is a simple and efficient operator, which has been used for texture discrimination and image feature extraction and has shown to be robust with respect to the variations in rotation and illumination[31,32]. The operator labels the pixels by thresholding the 3×3 neighbourhood of each pixel with the center value to produce a binary patch. LBPH uses the histogram of the labels as a texture descriptor of the patch. Later the operator is extended to a circular neighborhood of different sizes, named as circular LBP[33]. Another extension of the original operator is called uniform pattern[34,35].

In our study, we first obtain the circular LBP of all images from the dataset, and then compute the uniform LBP patterns. The LBP feature image is then divided into $m \times m$ local blocks[36], and the histogram of each local block is extracted and integrated as a single feature vector. Using SVM as the classifier, comparative experiments are conducted, and classification accuracy is computed via 5-fold cross validation (Table 3). It can be seen that the 1×1 division obtained a higher classification accuracy than any of the others. Therefore, the undivided uniform LBPH patterns are chosen as the image feature, and compared with CNN.

SVM. The SVM is a powerful classifier that works well on a wide range of complex classification problems[25]. SVM with different kernel functions can transform a nonlinear separable problem into a linear separable problem by projecting data into a higher dimensional space to maximize the classification distance and achieve the desired classification. In this study, the radial basis function (RBF)[37], a popular kernel function of SVM, is chosen as the kernel function. The LIBSVM[38], as an efficient open source tool, is chosen to build SVMs in our experiments. Szarvas *et al.*[39] have evaluated the automatically optimized features learned by CNN on pedestrian detection, and showed that the CNN + SVM combination can achieve a very high accuracy. Therefore, we employ SVM as classifier for two purposes: comparison of feature extraction methods and improvement of the performance of rice blast disease classification.

Results and Discussion

Evaluation metrics. To evaluate the performance of the competing methods, several statistical parameters are used to be as the performance metrics. The selected quantitative measures are accuracy, ROC, and AUC, all of which are popular evaluation metrics for classification methods. The classification accuracy is the principal indicator; the higher the accuracy, the better the performance by a classifier. The accuracy can be computed by Eq (5). ROC is another important objective evaluation metric in the task of image classification, which is defined by true positive rate and false positive rate; the larger the area under the ROC curve, the better the classification

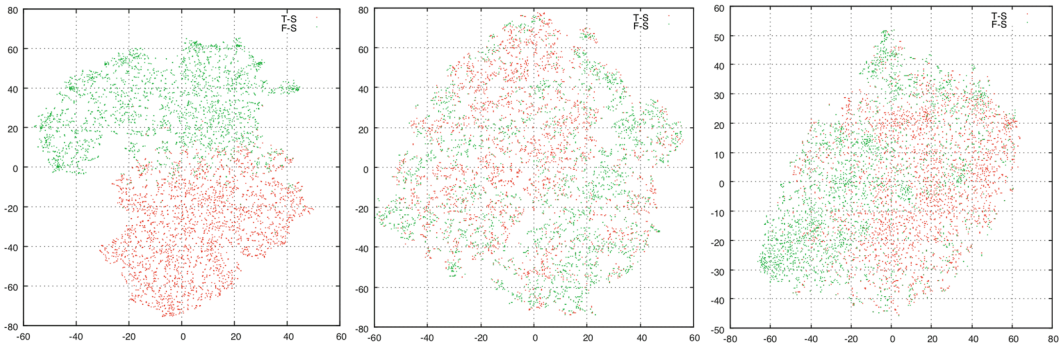

Figure 5. The t-SNE maps of (**a**) CNN, (**b**) LBPH, and (**c**) Haar-WT.

	Accuracy	AUC
CNN	95.83	0.99
CNN + SVM	95.82	0.99
LBPH + SVM	82.59	0.9
Haar-WT + SVM	83.85	0.92

Table 4. Comparison results of the four recognition methods.

performance. To analyze the reliability and the generalization ability of the feature extraction and classification methodology, the 5-fold cross-validation (CV) technique[40] is applied.

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{5}$$

where TP, FP, TN and FN are the numbers of true positives, false positives, true negatives, and false negatives in the detection results, respectively.

To assess the performance of feature extraction, the t-distributed stochastic neighbor embedding (t-SNE)[41] method is employed. The t-SNE was proposed by Hinton and has proven to be an effective qualitative indicator. In this study, we select the two-dimensional space as the mapping space for visualization, and a more linearly separable two-dimensional map implies better feature extraction performance.

Results and observations. To investigate the performance of three feature extraction methods, the t-SNE method is used to visualize the feature maps of CNN, LBPH, and Haar-WT using the same dataset. Figure 5(a–c) present the maps of the S8 layer features of CNN, LBPH features, and Harr-WT, respectively. The map of CNN in Fig. 5(a) clearly indicates that samples are almost separated in the two-dimensional space. In contrast, it is difficult to separate the two classes using LBPH and Haar-WT features, shown in Fig. 5(b,c). This result suggests that the features extracted using CNN are more discriminative than those extracted using LBPH and Haar-WT.

To further explore the effect of the features extracted by CNN, we conduct comparative experiment and quantitatively analyze in terms of accuracy, ROC, and AUC. For consistency, SVM is employed as the classifier, RBF is used for the kernel function, and the grid method is used to select the optimal c (cost) and g (gamma) parameters. To reduce possible biases in the selection of the validation set, all evaluation metrics were computed in 5-fold cross validation experiments.

First, CNN is primarily used to obtain high-level features from raw images and the Softmax is often used to classify and evaluate the accuracy. In addition, we employ SVM for classification combined with the CNN features (generated from its S8 layer). After parameter optimization, the accuracy of SVM-based classifier reached 95.82% ($c = 8.0$, $g = 0.0078125$; Table 4). Using the SVM combined with LBPH and Haar-WT features, on the other hand, we obtain two sets of comparison results shown in Fig. 6 and Table 4. To obtain accurate comparison results, we ensure that the same set of training and testing datasets were used for every method.

As shown in Table 4, it can be observed that the feature extraction method of CNN model combined with the SVM classifier achieves a remarkable performance in terms of the recognition rates, far superior to the LBPH and Haar-WT. The results of quantitative analysis in terms of accuracy, ROC and AUC are in agreement with the qualitative analysis using t-SNE. Therefore, the results verify that the features extracted using CNN can be effective in solving the rice blast classification problem.

For the same CNN features, the SVM and Softmax obtain higher accuracy (SVM:95.82%, Softmax:95.83%) and AUC value (SVM:0.99, Softmax: 0.99) than LBPH and Haar-WT (Fig. 6). Hence, the CNN and CNN + SVM showed a remarkable performance and are better for rice blast identification than LBPH + SVM and Haar-WT + SVM. In comparison, the SVM classifier has similar accuracy and AUC value to the Softmax. However, CNN is a black box model with random parameter initialization and, as a result, the output features of each trained model are different. SVM is a data-driven classifier that needs to optimize its parameters for different

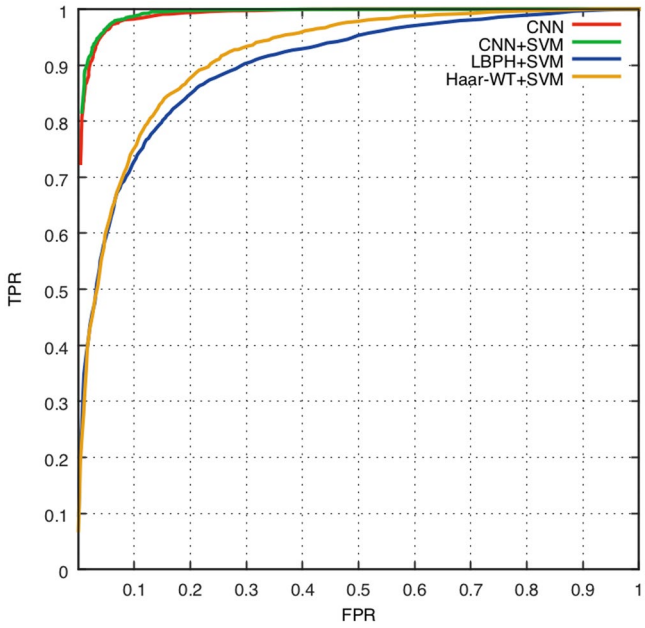

Figure 6. ROC curves of CNN, CNN + SVM, LBPH + SVM, and Haar-WT + SVM.

Figure 7. Example images of misclassification: (**a**) false negatives and (**b**) false positives.

feature data. Therefore, CNN + SVM is less convenient than CNN + softmax in terms of efficiency and system implementation, although it can be considered as a strong competitive method for rice blast recognition.

In order to understand the reasons of the misclassification, we analyze the misclassification samples some of which are shown in Fig. 7. We can observe from Fig. 7 that most notable mistakes in images of Fig. 7(a) are caused by blur, water droplets and small or incomplete lesion, and the main reason of misclassification in Fig. 7(b) are shadow, light spot, water droplets and complex background.

Finally shown in Fig. 8 is an example classification result of the CNN model presented by this study on a complete original image. This example demonstrates that the CNN model can correctly and effectively recognize almost all of the rice blast lesions.

Conclusions and Future Work

In this study, we present a rice blast feature extraction and disease classification method based on deep convolutional neural networks (CNN). Because of the absence of image dataset for this particular recognition research, as our first contribution, we established a rice blast disease dataset with the assistance of plant protection experts. The dataset can be combined with other rice disease images to build a content-rich dataset. Our hope is that this dataset will be useful for other people who are interested in rice or even crop disease recognition research.

Figure 8. The recognition result of rice blast using CNN.

In addition, we conduct comparative experiment based on the dataset and analyze the experimental results. Qualitative assessment by t-SNE indicates that the high-level features extracted by CNN are more discriminative and representative than LBPH and Haar-WT. Quantitative analysis results indicate that CNN with Softmax and CNN + SVM have almost the same performance, which is better than that of LBP + SVM and Haar-WT + SVM by a wide margin.

The occurrence of rice disease is regular, and the type and the probability of the rice disease vary with the stages of rice growth. Therefore, different rice disease identification systems should and can be established using the method presented by this study, and then the automated rice disease diagnosis can be realized by combining identification models and domain knowledge of rice disease.

Although our method of automatic identification of rice blast has achieved satisfactory results, substantial further work is needed to improve its accuracy and reliability in rice disease diagnosis systems. In particular, we plan to address the following two issues in future studies:

(1) Expand the dataset of rice disease, and establish a comprehensive tool for rice disease diagnosis system. The data augmentation method will be employed for building a good classifier when the number of samples is insufficient.
(2) Study other deep neural network architectures and take full advantage of the deep learning algorithms to improve the classification accuracy, and enhance the reliability and robustness of the rice disease diagnosis systems.

Data Availability

The rice blast disease dataset used for training and testing CNN model is available from the, http://www.51agritech.com/zdataset.data.zip, and all the data generated during and/or analyzed during the current study are included in the manuscript.

References

1. Khush, G. S. What it will take to Feed 5.0 Billion Rice consumers in 2030. *Plant Molecular Biology.* **59**, 1–6 (2005).
2. Roy-Barman, S. & Chattoo, B. B. Rice blast fungus sequenced. *Current Science* **89**, 930–931 (2005).
3. Abed-Ashtiani, F., Kadir, J. B., Selamat, A. B., Hanif, A. H. B. & Nasehi, A. Effect of Foliar and Root Application of Silicon Against Rice Blast Fungus in MR219 Rice Variety. *Plant Pathology Journal.* **28**, 164–171 (2012).
4. Kihoro, J., Bosco, N. J., Murage, H., Ateka, E. & Makihara, D. Investigating the impact of rice blast disease on the livelihood of the local farmers in greater Mwea region of Kenya. Springerplus. **2** (2013).
5. Dadley-Moore, D. Fungal pathogenesis - Understanding rice blast disease. *Nature Reviews Microbiology.* **4**, 323–323 (2006).
6. Wu, Y. *et al.* Characterization and evaluation of rice blast resistance of Chinese indica hybrid rice parental lines. *The Crop Journal.* **5**, 509–517 (2017).
7. Abed-Ashtiani, F. *et al.* Plant tonic, a plant-derived bioactive natural product, exhibits antifungal activity against rice blast disease. *Industrial Crops & Products.* **112**, 105–112 (2018).
8. Lu, Y., Yi, S., Zeng, N., Liu, Y. & Zhang, Y. Identification of rice diseases using deep convolutional neural networks. *Neurocomputing.* **267**, 378–384 (2017).
9. Sengupta, S. & Das, A. K. Particle Swarm Optimization based incremental classifier design for rice disease prediction. *Computers and Electronics in Agriculture.* **140**, 443–451 (2017).
10. Phadikar, S., Sil, J. & Das, A. K. Rice diseases classification using feature selection and rule generation techniques. *Computers and Electronics in Agriculture.* **90**, 76–85 (2013).
11. Jiao, Z. C., Gao, X. B., Wang, Y. & Li, J. A deep feature based framework for breast masses classification. *Neurocomputing* **197**, 221–231 (2016).
12. Ypsilantis, P. P. *et al.* Predicting Response to Neoadjuvant Chemotherapy with PET Imaging Using Convolutional Neural Networks. *Plos One.* **10** (2015).

13. Liu, Z. Y., Gao, J. F., Yang, G. G., Zhang, H. & He, Y. Localization and Classification of Paddy Field Pests using a Saliency Map and Deep Convolutional Neural Network. Scientific Reports. **6** (2016).
14. Johnson, J., Karpathy, A. & Fei-Fei, L. DenseCap: Fully Convolutional Localization Networks for Dense Captioning. 2016 IEEE Conference on Computer Vision And Pattern Recognition (Cvpr). 4565–4574 (2016).
15. Kang, M. J. & Kang, J. W. Intrusion Detection System Using Deep Neural Network for In-Vehicle Network Security. *Plos One*. **11** (2016).
16. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature.* **521**, 436–444 (2015).
17. Duan, M., Li, K., Yang, C. & Li, K. A hybrid deep learning CNN–ELM for age and gender classification. *Neurocomputing.* **275**, 448–461 (2018).
18. Hinton, G. & Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science.* **313**, 504–507 (2006).
19. Hubel, D. H. & Wiesel, T. N. Receptive fields,binocular interaction and functional architecture in the cat's visual cortex. *J.Physiol.* **160**, 106–154 (1962).
20. Krizhevsky, A., Sutskever, I. & Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the Acm.* **60**, 84–90 (2017).
21. Brahimi, S., Ben Aoun, N. & Ben Amar, C. Very Deep Recurrent Convolutional Neural Network for Object Recognition. Ninth International Conference on Machine Vision (Icmv 2016). **10341** (2017).
22. Zeiler, M. D. & Fergus, R. Visualizing and Understanding Convolutional Networks. *Computer Vision - Eccv 2014, Pt I.* **8689**, 818–833 (2014).
23. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Networks.* **61**, 85–117 (2015).
24. Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE.* **86**, 2278–2324 (1998).
25. Niu, X. X. & Suen, C. Y. A novel hybrid CNN–SVM classifier for recognizing handwritten digits. *Pattern Recognition.* **45**, 1318–1325 (2012).
26. Liu, X. *et al*. Localization and diagnosis framework for pediatric cataracts based on slit-lamp images using deep features of a convolutional neural network. *PLOS ONE.* **12**, e0168606 (2017).
27. Ding, W. & Taylor, G. Automatic moth detection from trap images for pest management. *Computers and Electronics in Agriculture.* **123**, 17–28 (2016).
28. Jadid, M. A.& Rezaei, M. Facial Age Estimation Using Hybrid Haar Wavelet and Color Features with Support Vector Regression. 2017 Artificial Intelligence And Robotics. (Iranopen): 6–12 (2017).
29. Lionnie, R. & Alaydrus, M. An Analysis of Haar Wavelet Transformation for Androgenic Hair Pattern Recognition. 2016 International Conference on Informatics And Computing (ICIC). 22–26 (2016).
30. Sarker, M. Content-based Image Retrieval Using Haar Wavelet Transform and Color Moment. The Smart Computing Review. **3** (2013).
31. Kamencay, P. *et al*. Accurate Wild Animal Recognition Using PCA, LDA and LBPH. 2016 Elektro 11th International Conference. 62–67 (2016).
32. Ojala, T., Pietikäinen, M. & Harwood, D. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition.* **29**(1), 51–59 (1996).
33. Benzaoui, A., Kheider, A. & Boukrouche, A. Ear Description and Recognition Using ELBP and Wavelets. 2015 International Conference on Applied Research In Computer Science And Engineering (Icar) (2015).
34. Ojala, T., Pietikainen, M. & Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Ieee Transactions on Pattern Analysis And Machine Intelligence* **24**, 971–987 (2002).
35. Ahonen, T., Hadid, A. & Pietikainen, M. Face recognition with local binary patterns. Computer Vision - Eccv 2004, Pt 1. 3021: 469–481 (2004).
36. Liao, S. C., Zhu, X. X., Lei, Z., Zhang, L. & Li, S. Z. Learning multi-scale block local binary patterns for face recognition. *Advances In Biometrics, Proceedings.* **4642**, 828–837 (2007).
37. Fan, R. E., Chen, P. H. & Lin, C. J. Working set selection using second order information for training support vector machines. *Journal Of Machine Learning Research.* **6**, 1889–1918 (2005).
38. Chang, C. C. & Lin, C. J. LIBSVM: A Library for Support Vector Machines. Acm Transactions on Intelligent Systems And Technology. **2** (2011).
39. Szarvas, M., Yoshizawa, A., Yamamoto, M. & Ogata, J. Pedestrian detection with convolutional neural networks. 2005 IEEE Intelligent Vehicles Symposium Proceedings. 224–229 (2005).
40. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. International Joint Conference on Artificial Intelligence1995. p. 1137–45 (1995).
41. van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *Journal Of Machine Learning Research.* **9**, 2579–2605 (2008).

Acknowledgements

This research was supported and funded by the Fund of Jiangsu Academy of Agricultural Sciences (No. 6111646), by the Program of Foshan Innovation Team (Grant No. 2015IT100072), and by NSFC (Grant No. 61673125), and by the Natural Science and Engineering Research Council of Canada.

Author Contributions

Wan-jie Liang performed the experiments, analyzed the results and wrote the manuscript. Hong Zhang supervised the work and revised the manuscript. Gu-feng Zhang and Hong-xin Cao performed data collection and labeling. All authors reviewed the manuscript and and agree with its contents.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.