Computers and Electronics in Agriculture 176 (2020) 105612

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Occurrence prediction of cotton pests and diseases by bidirectional long short-term memory networks with climate and atmosphere circulation

Peng Chen[a,b,c,1], Qingxin Xiao[b,1], Jun Zhang[d,*], Chengjun Xie[e,*], Bing Wang[f,*]

[a] National Engineering Research Center For Agro-Ecological Big Data Analysis & Application, Anhui University, 230601 Hefei, Anhui, China
[b] Institutes of Physical Science and Information Technology, Anhui University, 230601 Hefei, Anhui, China
[c] School of Internet, Anhui University, 230039 Hefei, Anhui, China
[d] School of Electrical Engineering and Automation, Anhui University, 230601 Hefei, Anhui, China
[e] Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031, China
[f] School of Electrical and Information Engineering, Anhui University of Technology, 243032 Ma'anshan, Anhui, China

ARTICLE INFO

Keywords:
Occurrence of pests and diseases
Climate
Atmosphere circulation
Pest counting
Bi-directional long-short term memory

ABSTRACT

The occurrence of crop pests and diseases always affects the development of agriculture seriously, while pest meteorology showed that climate is important in affecting the occurrence. Recently, recurrent neural network (RNN) has been broadly applied in various fields, which was designed for modeling sequential data and has been testified to be quite efficient in time series problem. This paper proposes to use bi-directional RNN with long short-term memory (LSTM) units for predicting the occurrence of cotton pests and diseases with climate factors. First, the problem of occurrence prediction of pests and diseases is formulated as time series prediction. Then the bi-directional LSTM network (Bi-LSTM) is adopted to solve the problem, which can capture long-term dependencies on the past and future contexts of sequential data. Experimental results showed that Bi-LSTM shows good performance on the occurrence prediction of pests and diseases in cotton fields, and yields an Area Under the Curve (AUC) of 0.95. This work further verified that climate indeed have strong impact on the occurrence of pests and diseases, and circulation parameters also have certain influence.

1. Introduction

With global warming, the occurrence frequency of regional crop pests and diseases has increased rapidly, causing great loss in agricultural production. The crop pest and disease is one of the major natural disaster in China, whose occurrence, development, and epidemic are closely related to climate conditions, or occur with meteorological disasters. Investigating the relationship between pandemic diseases and climate is significant for establishing climate-pest forecasting model and improving the long-term prediction of pests and diseases.

Cotton is an important economic crop, which occupies a vital position in the national economy in China. However, cotton is always endangered by various pests and diseases during its growth. Perennial pests and diseases caused about 15–20% economic loss, even up to 50%, in recent years. Therefore, the control of pests and diseases is crucial to the growth of cotton, which can recover more than 900,000 tons of cotton annually by pests and diseases control (Cui et al., 2007). During the growth of cotton, there are many factors affecting the production, where the most significant one is abnormal climate change. Abnormal climate change can result in the continuous evolution of pests and further make them adaptive to the environment. In addition, studies have shown that circulation parameters have a certain correlation with the occurrence of crop pests and diseases (Zhou and Gao, 2014). All of these factors seriously influence the yield and quality of crop production, and make it very difficult to control pests and diseases (Wu et al., 2009).

Nowadays, the methods of controling pests and diseases in cotton mainly include pesticide screening, ecological control, biological control and artificial trapping (Luo et al., 2017), where pesticides are always used and they are insecticidally effective and direct when used in cotton fields. But most pesticides are highly toxic and often cause serious residual pollution. Subsequently, high efficiency, low degree and environment-friendly new types of pesticides were tried to be developed for prevention and control on pests and diseases. With the rapid development of life sciences, biological control has become a popular

* Corresponding authors.
E-mail addresses: pchen@ahu.edu.cn (P. Chen), cjxie@iim.acn.cn (C. Xie), bingwang@ustc.edu (B. Wang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.compag.2020.105612
Received 12 December 2018; Received in revised form 15 May 2020; Accepted 28 June 2020

direction more and more. Cotton bollworm is commonly known as cotton leaf hopper, which is a type of severe pests in cotton and okra. Singh et al. (2018) evaluated 15 common housekeeping genes during different developmental stages of bollworm, and tried to feed/inject sequence-specific double-stranded RNA (dsRNA). Targeting such essential genes can help in developing new generation insect resistant transgenic plants, targeted towards downregulation or knockdown of essential genes for causing mortality. Moreover, many researches have been developed in releasing natural enemies of cotton fields, exploring habits and resources related to habitat control to attract natural enemies, which have played an important role in practice (Feng, 2008; Gao et al., 2016).

Although the control of cotton pests and diseases has gotten good results, pests and diseases are often occurred in sudden and complex ways. With computer development, machine learning-based methods hold promising on the control and prevention of cotton pests and diseases. Extensive studies have focused on the pest and disease prediction of crops. Hang et al. proposed an convolutional neural network-based model to predict apple leaf diseases (Hang et al., 2019). Li et al. presented a deep learning-based pipeline for the visual localization and counting of agricultural pests by self-learning of a saliency feature map from convolutional neural network (Li et al., 2019). Xie et al. built an model with multi-level learning features for automatic classification of field crop insects (Xie et al., 2018). Ding et al. proposed an automatic detection pipeline on the basis of deep learning technique for identifying and counting pests in images taken from field traps (Ding and Taylor, 2016), which can real-time monitor and warn on the occurrence of pests in the field. Zhang et al. developed three models, multiplier feed-forward neutral networks (MLFN), general regression neutral networks (GRNN) and support vector machine (SVM), to predict the occurrence area of dendrolimus superans, where SVM performs the best (Zhang et al., 2017).

As a special kind of RNN, LSTM neural networks is verified to be efficient in modeling sequential data (Hochreiter and Schmidhuber, 1997), which introduces gate mechanism into vanilla RNN to prevent the vanished or exploding gradient problem. Moreover, Bi-LSTM neural network (Bin et al., 2016), derived from LSTM network, has advantages in memorizing information for long periods in both directions, and shows rapid improvement in comparison with LSTM for video description. Bi-LSTM has achieved good results in different fields. Jiang et al. applied a character-level bi-directional LSTM network to represent tokens and classify tags for each token, and the LSTM-based system achieved a micro-F1 score of 0.8986 in i2b2 strict evaluation (Jiang et al., 2017). Zhao et al. designed a deep neural network structure, named Convolutional Bi-directional Long Short-Term Memory networks (CBLSTM), to address raw sensory data, which was used for monitoring machine health, and experimental results have shown that the model outperforms several state-of-the-art baseline methods (Zhao et al., 2017). Xie et al. designed a deep neural network approach with the state-of-the-art Bi-LSTM Network to extract e-cigarette safety information in social media, which eventually achieved the best performance compared with three baseline models, with a precision of 94.10%, a recall of 91.80% and an F-measure of 92.94% (Xie et al., 2018).

This paper proposes a Bi-LSTM network-based method to predict the occurrence of cotton pests and diseases. An improved Bi-LSTM based neutral network was properly designed with fully connected layers to form a classification model, with the use of climate factors and some atmosphere circulation parameters. Results showed that our model outperformed other traditional prediction models for the occurrence prediction of cotton diseases and insect pests.

2. Material and method

2.1. Dataset of pest images

The occurrence of crop pests and diseases is affected by a number of climate factors, such as temperature, humidity, rainfall, light and so on. In addition, circulation parameters also affect the occurrence of pests in specific ways. However, the interaction and the influence of various factors on pests and diseases are very complex. The study of the impact of climate factors and circulation parameters on agriculture is of great significant to the effective production of crops.

For climate factors, the datasets of cotton pests and diseases from Crop Pest Decision Support System (http://www.crida.in:8080/naip/AccessData.jsp) were used, where 15,343 cotton documents recorded weekly for 10 insect pests and diseases along with corresponding climate factors across 6 important locations in India were investigated in this work. Several climate factors are applied in the occurrence of pests, including Maximum Temperature (MaxT (°C)), Minimum Temperature (MinT (°C)), Relative Humidity in the morning (RH1 (%)), Relative Humidity in the evening (RH2 (%)), Rainfall (RF (mm)), Wind Speed (WS (kmph)), Sunshine Hour (SSH (hrs)) and Evaporation (EVP (mm)). The historical records can be used to predict future occurrence of pests and diseases. From the website, a total of 63 datasets with time-series records of cotton pests and diseases are obtained. Tables 1 and 2 provide simple statistics on different types and areas of cotton pests and diseases, respectively.

Moreover, 74 atmosphere circulation indexes from China Meteorological Administration National Climate Center (http://cmdp.ncc–cma.net/cn/index.htm) were used for additional factors, which recorded the corresponding parameters of atmospheric circulation intensity in different regions, i.e., Indian Subtropical High Area Index, Northern Hemisphere Subtropical High Intensity Index, Indian Subtropical High Intensity Index, etc. The website provided 74 circulation documents monthly from 1951 to nowadays. Table 3 lists the top 25 circulation indexes from random forests-based embedded method.

2.2. Data process and feature selection

Our research data mainly include two parts, climate-pest records weekly and circulation parameter documents monthly. In order to be unified in the time dimension and obtain enough data to train the network, circulation parameters monthly should be expanded into weekly statistics by interpolation technique. As a result, 8 climate parameters, 74 circulation parameters and corresponding cotton pest or disease value records weekly were obtained.

2.3. Climate factor combination

According to plant disease and pest meteorology, there is a certain relationship between climate factors and the occurrence of pests and diseases, which has been verified by many studies (Kelly et al., 2015; Prasetyo et al., 2017). In order to improve model performance on occurrence prediction of cotton pests and diseases, some combinations of

Table 1
Cotton pest and disease datasets in different areas.

Location	Number of samples
Akola	2028
Coimbatore	208
Lam	5265
Nagpur	3328
Pharbhani	2644
Sirsa	1870
Total	15343

Table 2
The types of pest and disease datasets.

Type of pest and disease	Number of sample
Bollworm	7183
Aphid	1032
Jassid	1974
Thrips	832
Whitefly	1508
Spodoptera	630
Mealybug/Miridbug	260
LeafBlight/LeafSpot	1924
Total	15343

Table 3
The top 25 circulation indexes from random forests-based embedded method.

1. Meridional index over Asia
2. The ridge line of the North America-Atlantic subtropical high
3. Station of the East Asian trough
4. Index of the strength of the polar vortex in the Atlantic and Europe sector
5. The ridge line of the subtropical high over the eastern Pacific
6. The ridge line of the Indian subtropical high
7. Index of the strength of the subtropical high over the Pacific
8. The ridge line of the Atlantic subtropical high
9. Index of the area of the northern hemisphere polar vortex
10. Index of the northern extend of the North America-Atlantic subtropical high
11. Index of the strength of the polar vortex in the Asia
12. The ridge line of the subtropical high over the western Pacific
13. Index of the strength of the polar vortex in the Pacific sector
14. Index of the northern extend of the subtropical high over the western Pacific
15. Index of the strength of the polar vortex in the north American
16. The ridge line of the subtropical high over the Pacific
17. Atlantic and Europe pattern E
18. Index of the northern extend of the Indian subtropical high
19. Index of the northern extend of the subtropical high over the Atlantic
20. Index of the area of the North American subtropical high
21. Index of the western extend of subtropical high over the Western Pacific
22. Meridional index over Eurasian continent
23. Index of the area of the subtropical high over Indian
24. Station of the East Asian trough
25. The ridge line of the North African subtropical high

climate features were constructed, including Light temperature product (LTP), Temperature difference (TF), Precipitation/temperature ratio (PTR) and Temperature and humidity coefficient (THC). These definitions are:

$$
\begin{aligned}
&LTP = SSH \times \frac{MaxT + MinT}{2}\\
&TF = MaxT - MinT\\
&PTR = \frac{RF}{MaxT + MinT}\\
&THC = \frac{RH1 + RH2}{2} \Big/ \frac{MaxT + MinT}{2}
\end{aligned} \tag{1}
$$

2.4. Circulation parameter selection

Due to redundancy in features affecting model training, top circulation features were selected instead of all 74 original features. The redundant features not only can not contribute to the prediction of insect pests, but also may affect the training performance of our model. All experiments in this work were implemented using Python, where the sklearn module gives us a large number of methods for feature selection. Here, random forests-based embedded method was adopted from the feature selection library for variable selection. This method mainly has two functions: investigating insights on the behavior of variable importance index based on random forests and proposing a ranking strategy of explanatory variables and a stepwise ascending variable introduction strategy (Genuer et al., 2010). Eventually 25 circulation features were obtained, as shown in Table 3. The top 25 circulation features and climate factors are fused to predict the occurrence of cotton pests and diseases.

Table 4
The distribution of records at different levels of pests and diseases in cotton.

Cotton pest	No pest	A little (< 5%)	General (5–20%)	Serious (> 20%)	Total
Records	10974	1855	1209	1305	15343

2.5. The definition of occurrence level of pests and diseases

Compared with the occurrence prediction of cotton pests and diseases, the prediction of hazard level may be more indicative significant. However, there is no standard definition to determine the hazard level of crop pests and diseases till now. Therefore, according to the occurrence of pests and diseases, we classify them into four levels: no pests, a little (< 5%), general (5–20%) and serious (> 20%). The detailed records are show in Table 4.

2.6. Data process

After climate factor combination and circulation parameter selection, a total of 38 features are obtained, consisting of 12 climate features, 25 circulation features and 1 for pest/disease value. In order to make the model generalized a certain degree, different types of cotton pests and diseases data were put together for network training. The ratio of the size of training dataset to that of validation dataset to that of test dataset was set as 7:1:2. Before model training, all features were normalization and standardization.

2.7. Problem formulation

To investigate the impact of climate factors and circulation parameters on pest/disease occurrences, the weekly records of cotton climate factors and circulation parameters are used. Our aim is to predict the occurrences of pests and diseases under different climate factors and circulation parameters.

Suppose X be the vector set of climate feature records along with the occurrences of pests and diseases, $Y = \{0, 1\}$, in one single area along all the time, which is a time series set. The prediction problem can be then converted into the problem of whether feature vector X_i, $i = 1 \dots N$ is of the non-occurrence ($Y_i = 0$), a little ($Y_i = 1$), general ($Y_i = 2$) or serious occurrence ($Y_i = 3$) of pests and diseases, where N is the number of feature vectors. We try to build a model to capture the relationship among data (X_i, Y_i), $i = 1 \dots N$, to predict future occurrence of pests and diseases under climate factors and circulation parameters. So the prediction problem can be formulated as a multiclass classification problem, according to the past climate factors (X) and pest values (Y).

2.8. Architecture of pest/disease prediction with Bi-LSTM Network

Recurrent Neural Networks (RNNs), and specifically a variant with Long Short-Term Memory (LSTM), are enjoying renewed interests and have been successfully applied in a wide range of machine learning problems that involve sequential data (Karpathy et al., 2015). LSTM is a recurrent neural network architecture (an artificial neural network) published in 1997 by Hochreiter and Schmidhuber (Hochreiter and Schmidhuber, 1997), and has been refined and promoted by Alex Graves (Graves, 2013) recently. Like most RNNs, LSTM has a memory function that can be used to handle time series problems (Sutskever et al., 2014); Unlike traditional RNNs, LSTM is well-suited for long-term dependency problems because it can solve the problems of gradient vanishing (Hochreiter and Schmidhuber, 1997; Bengio et al., 1994; Hochreiter, 1991) and gradient exploding (Mikolov, 2012; Pascanu et al., 2013) caused by Back Propagation Through Time (BPTT). Bi-LSTM neural network (Bin et al., 2016) is similar to LSTM network in structure, and both of them are constructed with LSTM units (Schuster and Paliwal, 1997). The special unit of this network is capable of learning long-term dependencies without keeping redundant information. Bi-LSTM has an end-to-end working mode like neural network, which automatically processes input data and yields desired results (Miao et al., 2015). It does not require complex feature selection and model testing compared with traditional machine learning. Once Bi-LSTM network training is completed, it only needs to update network parameters with new data, without having to rebuild the model. In recent years, researchers proposed improved structure of LSTM unit, i.e., Gated Recurrent Unit (GRU) (Chung et al., 2014), making it more applicable and more efficient in prediction performance and training time.

2.8.1. LSTM unit

To capture potential relationship between the time series data of climate factors and pest/disease values, LSTM was used in this work, where each LSTM unit contains three doors. For input x_i, the input gate decides the input entering into current cell, $i_t = \sigma(W^i \times [h_{t-1}, x_t] + b^i)$; the forget gate decides if and how much information can be forgotten for the previous memory, $f_t = \sigma(W^f \times [h_{t-1}, x_t] + b^f)$; and the output one controls the information outputting from current cell, $o_t = \sigma(W^o \times [h_{t-1}, x_t] + b^o)$. The gating operation ultimately determines which information is forgot and which information is entered into the neural network as useful information. The sigmoid function can be expressed as: $\sigma(x) = \frac{1}{1+e^{-x}}$. For the climate-pest/disease forecasting issue, it processes a series of temporal dependency inputs x_t at time t and the hidden vector h_{t-1} from the last time, therefore the memory vector of the LSTM cell can be iterated as $C_t = f_t \times C_{t-1} + i_t \times c_t$, where $c_t = tanh(W^c \times [h_{t-1}, x_t] + b^c)$ is a new short-term state vector at time t. The tanh function can be expressed as: $tanh(x) = \frac{sinh(x)}{cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. As a result, the output vector of the cell can be further predicted (Hochreiter and Schmidhuber, 1997) as:

$$h_t = o_t \times tanh(C_t), \tag{2}$$

where W is the recurrent weight matrix; b is the corresponding bias vector; the superscripts of i, f and o are the outputs of the input, forget, and output gates, respectively; and C and h are the memory vector and out vector of the cell, respectively.

2.8.2. Bi-LSTM network

A Bi-LSTM consists of two LSTM units that run in parallel, one on the input sequence and the other on the reverse of the input sequence. At each time step, the hidden state of Bi-LSTM is the concatenation of the forward hidden state (h_{t-1}, C_{t-1}) and the backward hidden state (h_{t+1}, C_{t+1}), which allows the hidden state to capture both past and future information. Bi-LSTM network was designed to capture information of sequential dataset and maintain climate-pest/disease features from past and future.

As described in Schuster and Paliwal (1997) and Kalchbrenner et al. (2015), the output of Bi-LSTM, (h_t, C_t), can be represented as a whole function LSTM (*):

$$(h_t, C_t) = LSTM([h_{t-1}, h_{t+1}, x_i], C_{t-1}, C_{t+1}, W), \tag{3}$$

where W concatenates the four weight matrices W^i, W^f, W^o and W^c.

2.8.3. Bi-LSTM based classifier

Fig. 1 illustrates the structure of the proposed networks. Supposed (X, Y) be data input, where X denotes the records of 38 feature vectors (12 climate features, 25 circulation parameters and historical value for cotton pests and diseases), and Y: {0, 1, 2, 3} denotes pest/disease hazard level. Feature selection and Feature combination for feature vectors were developed to select proper features to predict the pest/disease level. Normalization and Standardization were used to unify different statistical data, which can ensure the comparability between different types of pests and diseases. Before inputting data into the network, the forecasting problem of time series for pests and diseases must be reframed as supervised learning problem. The "Reshape" technique helped change data into format we needed. For RNN, the time series records should be converted to 3D tenser (N_{samp}, *timesteps*, N_{feat}). In this work, N_{samp} is the number of samples that is set as 15,343, *timesteps* as 4 and N_{feat} as 38 including 37 features and one for pest/disease value. The Bi-RNN has advantages in dealing with small data samples. So, Bi-RNN combining with 6 Bi-LSTM units, and a full-connected layer with 4 nodes, was constructed to build a basic LSTM network block (Zhang et al., 2017). The former can capture the temporal relationship between different features and the occurrences of pests and diseases. Since the output of the LSTM layer is a vector, then a full-connected layer was adopted to make a better abstraction and a combination of output vectors. The latter also reduces the dimensionality of output and then maps the reduced vector to a final prediction. In addition, "Dropout" technique was introduced into Bi-LSTM block to avoid over-fitting. The final prediction can be defined as below:

$$(h_i, C_i) = LSTM([h_{i-1}, h_{i+1}, x_i], C_{i-1}, C_{i+1}, W), \tag{4}$$

$$prediction = softmax(W^{fc} \times y_l + b^{fc}), \tag{5}$$

where (h_i, C_i) stands for the output of the i-th cell of Bi-LSTM; softmax (*) is softmax function; h_l is the hidden vector in the last time step of Bi-LSTM layer; W^{fc} and b^{fc} are the weight matrix and bias term in full-connection layer, respectively; $prediction = \{0, 1, 2, 3\}$, after one-hot encoding, represents the classification result of Bi-LSTM network.

2.8.4. Architecture of the Bi-LSTM Network

The occurrence prediction of cotton pests and diseases is regarded as a time series problem, which uses the historical climate data and pest counting values or percentages of disease area to identify whether pests and diseases will occur in the future. We should determine the length of historical observations used for the occurrence prediction. Of course the longer the historical data is, the better the prediction will be, however the more computation the prediction will need. Here the "timesteps" is set as 4, i.e., four samples of records are inputting together into Bi-LSTM. In addition, three parameters for the whole structure of the network should be determined: the layer number of Bi-LSTM l_r, the number of full-connected layers l_{fc} and the corresponding number of hidden units denoted by *units_r*.

In order to train the network, some critical parameters have to be determined, such as optimization method, learning rate, batch size and so on. The traditional optimization method for deep neutral network is stochastic gradient descent (SGD) (Ruder, 2016), which is the batch version of gradient descent. The details of gradient descent and the parameters of network can be seen in below:

$$\begin{aligned} g_t &= \frac{df_t(\theta)}{d\theta} \\ \theta_t &= \theta_{t-1} - \eta \times g_t, \end{aligned} \tag{6}$$

where $f_t(\theta)$ is the objective function used in the Bi-LSTM network; η is the learning rate; θ is the parameter vector of network.

Here, categorical-crossentropy was adopted as the loss function of the binary classification, whose definition is shown in below,

$$f_t(\theta) = -\sum_{i=1}^{n}\sum_{t=1}^{c}(y_{i,t}^{true} \times log(y_{i,t}^{prediction})), \tag{7}$$

where n is the number of samples; c is the number of categories; $y_{i,t}^{true}$ is the actual value; $y_{i,t}^{prediction}$ is the prediction value of the network, which is calculated by Eq. (5).

However, SGD has many disadvantages in real training process. For example, it uses the same learning rate for all parameter updates and it is difficult to find the best learning rate for non-stationary objectives and different features. Sometimes, it falls into a local optimal solution.

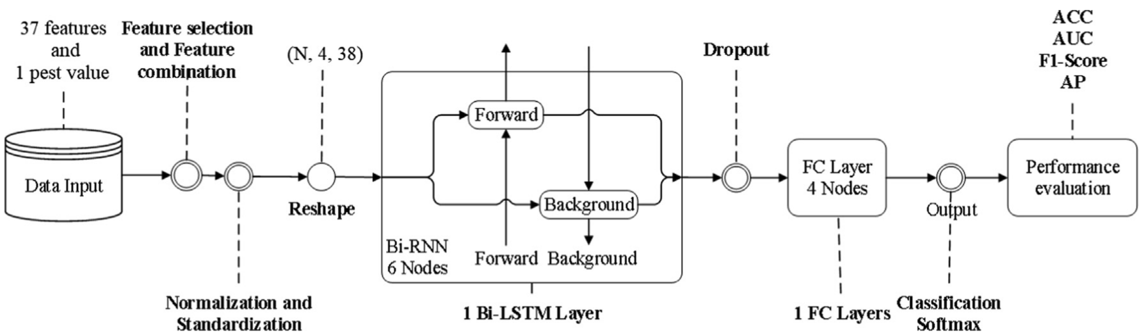

Fig. 1. Overview of the proposed network configurations.

To address these issues, Adam optimization method (Kingma and Ba, 2014) was adopted that combines the advantages of two recently popular methods: AdaGrad (Duchi et al., 2011), which works well with sparse gradients, and RMSProp (Tieleman and Hinton, 2012), which works well in on-line and non-stationary settings. Adam calculates an independent adaptive learning rate for different parameters by the first and second moment estimates of the gradients m_t and v_t. For the Bi-LSTM-based network, at timestep t, it returns the best parameters θ_t as below:

$$m_t = \beta_1 \times m_{t-1} + (1 - \beta_1) \times g_t \quad mv_t = v_{t-1} \times \beta_2 + (1 - \beta_2) \times g_t$$
$$\widehat{m_t} = \frac{m_t}{1-\beta_1^t}, \quad \widehat{v_t} = \frac{v_t}{1-\beta_2^t}, \quad \theta_t = \theta_{t-1} - \eta \times \frac{\widehat{m_t}}{\sqrt{\widehat{v_t}} + \xi}, \tag{8}$$

where m and v are the first moment estimate and second raw moment estimate, respectively; $\beta 1$ and $\beta 2$ are corresponding exponential decay rates.

2.9. Implementation and performance measurement

Other traditional classification models, i.e., k-Nearest Neighbor (KNN) and random forest, were also implemented for cotton pest occurrence prediction in comparison with our Bi-LSTM model. The experiments ran under the environment of Intel (R) Core (TM) i7-4790 CPU @3.60 GHz (8CPUs), 8G RAM, Windows 10 64 bits operating system, programming with Python 3.6. The proposed network was implemented by TensorFlow 0.11 (Abadi et al., 2016), while KNN and random forest were implemented by Scikit-learn (Pedregosa et al., 2011).

Here, the prediction of pests and diseases is a basic classification problem. In this work, Accuracy (ACC) (Accuracy (trueness and precision) of measurement methods and results, YYYY), Area Under the Curve (AUC) (Hanley and McNeil, 1983), Average Precision (AP) and F1-score are used to measure the effectiveness of prediction methods. For binary classification model outputs, there are only two types of results, positive and negative ones (denoted as P and N). Therefore bivariate model has four outcomes for the case predictions: true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

The definitions of *ACC*, *AP* and *F1-score* are shown in below:

$$Acc = \frac{TP + TN}{P + N}$$
$$Precision = \frac{TP}{TP + FP}$$
$$Recall = \frac{TP}{TP + FN}$$
$$AP = \sum_n (Recall_n - Recall_{n-1}) \times Precision_n$$
$$F1 - score = \frac{2(Precision \times Recall)}{Precision + Recall} = \frac{2TP}{P + P'}, \tag{9}$$

In multiclass classification task, the notions of precision, recall, and F-measures can be applied to each label independently, which could combine results across labels, specified by the "average" argument. Here, we set "average" as "weighted".

In addition, Receiver Operating Characteristic (ROC) curve was introduced and the area under the ROC curve (AUC) can be used to evaluate a classifier. The ROC curve is drawn by True Positive Rate (TPR) with respect to False Positive Rate (FPR). The closer the ROC curve is to the upper left corner, the better the classifier performs.

$$TPR = \frac{TP}{TP + FN}$$
$$FRP = \frac{FP}{FP + TN} \tag{10}$$

3. Results

3.1. Determination of parameters

This work used 15,343 records from the Crop Pest Decision Support System. In order to test the accuracy and generalization ability of the proposed network, 70% datasets were randomly selected to train the Bi-LSTM-based network and determine the parameters of network. First, let's set $l_r = 1$ for the layer number of Bi-LSTM, $l_{fc} = 1$ for the layer number of fully connected layers, $units_{fc} = 1$ for the unit number of fully connected layers, and choose a proper unit number of Bi-LSTM, $units_r$, from (4, 5, 6, 7, 8). Table 5 shows the prediction comparison with different values of $units_r$. The boldface one in the table represents the best performance among $units_r$, i.e., the largest ACC, AUC, F1-score and AP for different $units_r$. It can be seen from the results that the best performance occurs when $units_r = 6$. So in the following experiments, we set $units_r$ as 6.

Then, the time series sequences were used to choose a proper value for l_r from {1,2,3}, the other two parameters are set as $units_r = 6$ and $l_{fc} = 1$. Table 6 shows the prediction comparison with different values of l_r. The boldface one in the table represents the best performance. Results showed that the best performance occurs when $l_r = 1$. The reason may be due to the increasing number of weights with increasing recurrent Bi-LSTM layers, which resulted in that insufficient dataset can not be fully train the larger amount of weights. Actually, experiences showed that Bi-LSTM with more layers did not always performs good. Results in this work also showed that more Bi-LSTM layers yields unstable results more likely. Therefore, in the following experiments, l_r is set as 1.

Similarly, on the same datasets, a proper value was set for l_{fc} from {1, 2, 3} and its units. Table 7 shows the prediction comparison with

Table 5
Prediction comparison with respect to $units_r$.

$units_r$	AUC	ACC	F1-score	AP
4	0.9520	0.8728	0.8725	0.8982
5	0.9521	0.8750	0.8765	0.8995
6	**0.9543**	**0.8773**	**0.8773**	**0.9044**
7	0.9536	0.8753	0.8749	0.9016
8	0.9512	0.8711	0.8706	0.8970

Table 6
Prediction comparison with respect to l_r.

l_r	AUC	ACC	F1-score	AP
1	**0.9543**	**0.8773**	**0.8773**	**0.9044**
2	0.9527	0.8755	0.8753	0.9022
3	0.9529	0.8722	0.8728	0.8993

Table 7
Prediction comparison with respect to l_{fc}.

l_{fc}	AUC	ACC	F1-score	AP
1[4]*	**0.9543**	**0.8773**	**0.8773**	**0.9044**
2[6,4]	0.9528	0.8750	0.8748	0.8997
3[6,6,4]	0.9506	0.8691	0.8702	0.8962

* The numbers in the square brackets stand for the number of the hidden units.

respect to the parameter of l_{fc}. The boldface one in the table represents the best performance. The model achieves the best performance when $l_{fc} = 1$. The reason is similar to that in the choose of l_r, i.e., the model with more layers means that there are more weights to be trained and more computation it is required. So in the following experiments, we set $l_{fc} = 1$ and the number of the hidden units are 4. The final full connectivity layer is integrated into the Bi-LSTM model to yield the predictions of pests and diseases.

After building the basic framework of the proposed Bi-LSTM network, the other parameters have to be adjusted to make the model achieving higher performance, i.e., $dropout = 0.1$, $batch_{size} = 32$, $learning_{rate} = 0.001$. The structure of our Bi-LSTM network is shown in Fig. 1. Compared with classical machine learning methods, one advantage of the deep learning model is that it can directly update network parameters for new data of the same type, without having to repeat feature selection and build networks (LeCun et al., 2015). Bi-LSTM not only can update the network parameters in real time according to the current input data and can be applied to predict the occurrences of other kinds of pests, but also have advantages in dealing with small data samples compared with traditional neural networks.

Although deep learning models do not require the cumbersome and time-consuming feature selection generally, adequate feature inputs associated with prediction targets still result in relatively high performance. Fig. 2 shows the prediction comparison of the model with only 9 features (8 climate features and 1 for pest value) and all 38 features (12 climate features, 25 circulation parameters and 1 for pest value) on our network. From the Fig. 2, the model with all 38 features still outperforms slightly than that with the 9 climate features on our network.

Table 8
Prediction results with different features on our network.

Feature	AUC	ACC	F1-score	AP
9 features	0.9494	0.8704	0.8701	0.8968
38 features	0.9545	0.8784	0.8784	0.9059

Table 8 shows the detailed scores and the boldface items in the table represent the best performance.

3.2. Prediction comparison with other methods

To show the power of our proposed method, the prediction comparison has been investigated with other classical machine learning methods, such as KNN (Pedregosa et al., 2011), Random Forest (Hanley and McNeil, 1983) and LSTM network. The parameters of these models are set as, for LSTM network and our Bi-LSTM-based network, the same parameters of $units_r$, l_r and l_{fc} were set as 6, 1 and 1, respectively; for KNN, $weights$ = 'distance', $n_{neighbors} = 3$, $algorithm$ = 'ball_tree' and $p = 2$; for Random Forest, $n_{estimators}$ is set as 100.

Fig. 3 and Table 9 show that our network obtain good performance on the occurrence prediction of cotton pests and diseases, while Fig. 4 shows the ROC Curve of our network on occurrence prediction of pests and diseases. The boldface items in the table represent the best performance, i.e., the largest ACC, AUC, AP and F1-score. It can be seen from the results that the Bi-LSTM-based network achieves the best prediction performance, LSTM are the second, Random Forest are the third method, and KNN performs the worst. Moreover, the proposed Bi-LSTM method achieves an AUC of 0.95 and an AP of 0.90, whileas it is difficult to achieve such high performance with traditional machine learning methods. From the results, LSTM and Bi-LSTM perform similarly. Although the results are based on the small dataset with 15,343 records, Bi-LSTM performs better than LSTM for 10 insect pests and diseases cross 6 important locations in India.

3.3. Prediction comparison with other methods

Moreover, the prediction performance of our model on different types of data was investigated. According to Crop Pest Decision Support System, the data of 6 areas in India and 10 different types of cotton pests and diseases are shown in Table 1 and Table 2, respectively. Our model performed on these two types of datasets separately and the results are shown in Table 10 and Table 11. Among them, because the Coimbatore area (208 records) datasets and the Mealybug/Miridbug insect pests (260 records) datasets are too small to obtain stable prediction performance, their predictions were ignored here. From

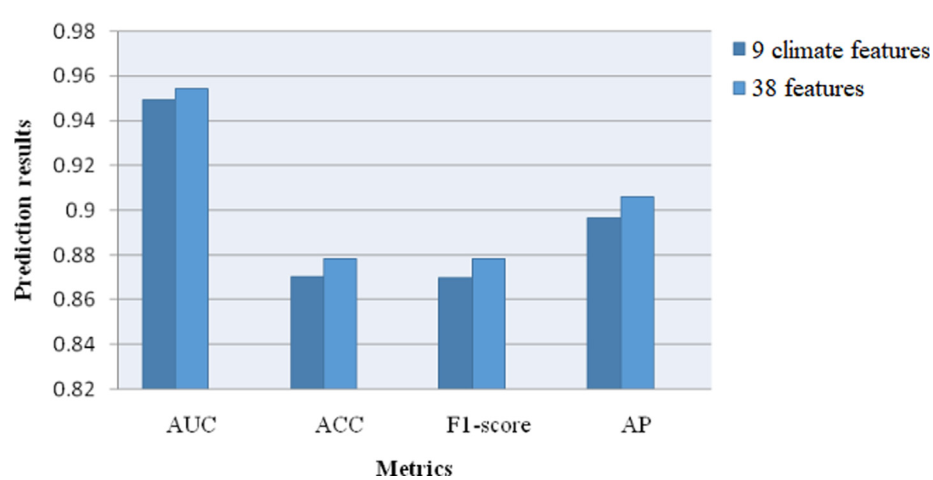

Fig. 2. Prediction comparison with different features on our network.

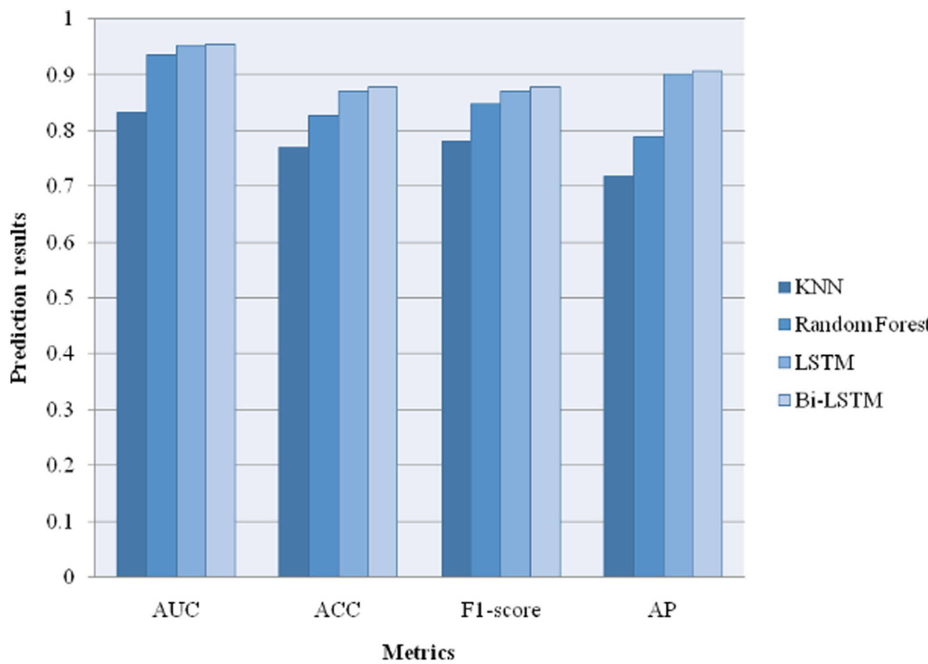

Fig. 3. Performance comparison with different methods.

Table 9
Performance comparison with other methods.

Methods	AUC	ACC	F1-score	AP
KNN	0.8332	0.7684	0.7809	0.7180
Random Forest	0.9359	0.8258	0.8478	0.7878
LSTM	0.9520	0.8701	0.8698	0.9017
Our Bi-LSTM method	**0.9545**	**0.8784**	**0.8784**	**0.9059**

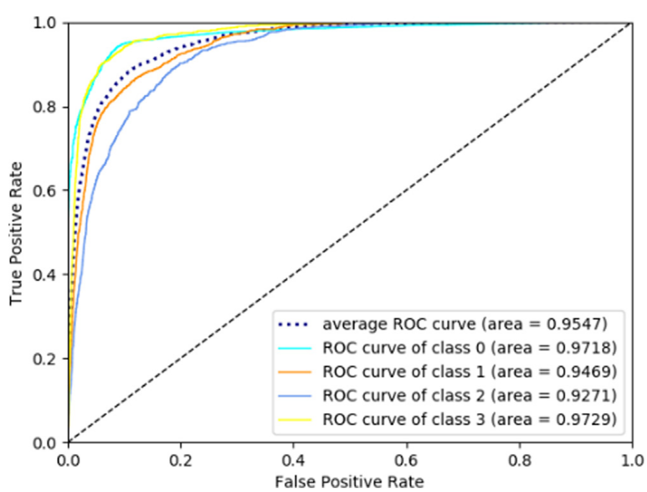

Fig. 4. ROC curves of four pest level classes with Bi-LSTM network. Here "area" means the area under each ROC curve.

Table 10
Performance comparison on pests and diseases of different areas with Bi-LSTM network.

Metrics	Akola (2028)	Lam(5265)	Nagpur (3328)	Pharbhani (2644)	Sirsa(1870)
AUC	0.9356	0.9502	0.9514	0.9578	0.9545
ACC	0.8245	0.8739	0.8674	0.8899	0.8869
F1-score	0.8234	0.8734	0.8649	0.8883	0.8837
AP	0.8578	0.8972	0.8997	0.9188	0.9148

Table 11
Performance comparison on different types of pests and diseases with Bi-LSTM network.

Datasets(size)	AUC	ACC	F1-score	AP
Aphid(1032)	0.9401	0.8298	0.8234	0.8649
Bollworm(7083)	0.9461	0.8675	0.8668	0.8941
Jassid(1974)	0.9661	0.8873	0.8871	0.9197
Spodoptera(630)	0.9225	0.8437	0.8291	0.8733
Thrips(832)	0.9548	0.8353	0.8326	0.8789
Whitefly(1508)	0.9404	0.8565	0.8473	0.8882
LeafBlight/LeafSpot(1924)	0.9669	0.9172	0.9095	0.9447

Table 10 and Table 11, results showed that our model achieved good performance both on the datasets in different regions and on those of different types of insect pests and diseases.

From Table 10, the model tested on pests and diseases of most areas achieves an AUC more than 0.95, except for Akola area. It mights because there are more complex reasons to affect the occurrence of pests and diseases on Akola. Compared with the prediction results in Table 10, the differences in Table 11 are more significant and the results are also easy to accept. There are large difference between the occurrence of different types of insect pests and that of crop diseases. From the results, our model shows the best performance in predicting the occurrence of cotton LeafBlight disease and LeafSpot disease, which may indicate that there is a large correlation between climate change and the occurrence of cotton disease.

3.4. Occurrence forcasting of pests for different future times

It is very interesting to know the effectiveness of our model in predicting the occurrence of pests and diseases in the coming weeks. Therefore, our model on the datasets were investigated the occurrence of cotton pests and diseases in the next week, two weeks and one month (four weeks). Results are shown in Table 12. Obviously, with prolonging the forecasting time, the prediction performance of our model is declining. When the predicted time is one month, our model still achieves an AUC of about 0.90, which shows that our model can make a good occurrence forcasting of cotton pests and diseases for a long future time.

Table 12
Occurrence forcasting of pests for different future times with our Bi-LSTM-based network.

Metrics	1 week	2 weeks	1 month
AUC	0.9530	0.9312	0.8979
ACC	0.8746	0.8332	0.7935
F1-score	0.8745	0.8264	0.7742
AP	0.9006	0.8699	0.8292

4. Conclusions

In this paper, the problem of cotton pest occurrence was transformed as a time series multiclass classification problem. To predict the future occurrence of pests and diseases, a Bi-LSTM network based architecture was proposed to model the temporal relationship of climate features and pests. This is the first time, to our knowledge, to use bidirectional recurrent neural network to solve the prediction problem of occurrence of pests and diseases. The proposed network utilized Bi-LSTM layer to model the time series data, and further adopted full-connected layer to map the output of Bi-LSTM layer to obtain final prediction. The model could predict the occurrence of cotton pests and diseases according to climate factors and circulation parameters in the future, so that people can take real time precaution and reduce crop economic losses.

The optimal setting of this architecture was explored by experiments and reported the prediction results of all datasets to confirm the effectiveness of the proposed method. In addition, the effects of 9 climate features and 38 features have been compared on pest and disease prediction. The results showed that the use of 38 features achieved better prediction on the occurrence of cotton pests, compared with that of 9 climate features. Moreover, some traditional machine learning methods (KNN and Random Forest) and the basic LSTM network were implemented to show the prediction comparison with Bi-LSTM model. Results showed that Bi-LSTM network has certain advantages in processing small sample dataset and time-dependent problem, and shows the importance of model selection. Although our model outperformed other methods, probably, the features the datasets contained are insufficient to achieve more accurate predictions.

In order to test the generalization ability of our model, we forecasted cotton pests and diseases in different regions of India, i.e., Akola, Nagpur, Sirsa. And the results showed good predictions in different regions. Therefore, our model is not regionally specific in predicting the occurrence of cotton pests. Moreover, we also predicted the occurrence of different types of cotton pests and diseases with Bi-LSTM network. Unlike before predictions, there have some differences among prediction results on different types of pests and diseases. Compared with cotton insect pests, our network performs better on cotton diseases (LeafBlight and LeafSpot), which achieved an AUC score about 0.97 and an AP score about 0.94. This may indicate that there have a certain relationship between climate factors and the occurrence of cotton pests and diseases, which are practical guidance in agricultural development.

Of cause the occurrence of cotton pests and diseases is not only related to climatic factors, but also closely related to the district topography, the growth of cotton, the growth cycle, evolution of pests and so on. The model only considering climatic factors in this work cannot fully investigate the occurrence of pests and diseases. Although the proposed model yielded good predictions on different datasets, it seems that it could be greatly improved and it is worth of collecting more effective features to further optimize the network. This paper only addressed the issue of the occurrence of cotton pests and diseases and predicted the occurrence with respect to climate factors. However, it is more interesting and meaningful to concern about the pest hazard level of crops in reality. It is a problem of regression prediction. Therefore, in the future, we will focus on two aspects of work. On the one side, we would like to build datasets with more factor features, including climate factors, the occurrence cycle of pests and diseases and so on. On the other side, we would try to construct more effective model to predict the hazard level of pests and diseases so that prediction results are more responsive to data, making it easier for people to develop detailed pest control strategies.

CRediT authorship contribution statement

Peng Chen: Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration. **Qingxin Xiao:** Data curation, Writing - original draft, Formal analysis. **Jun Zhang:** Visualization, Investigation. **Chengjun Xie:** Validation, Writing - review & editing. **Bing Wang:** Validation, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 61672035, 61872004 and U19A2064), Educational Commission of Anhui Province (No. KJ2019ZD05), Anhui Province Funds for Excellent Youth Scholars in Colleges (gxyqZD2016068), the fund of Co-Innovation Center for Information Supply & Assurance Technology in AHU (ADXXBZ201705), and Anhui Scientific Research Foundation for Returness.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, athttps://doi.org/10.1016/j.compag.2020.105612.

References

Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. J. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Józefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. G. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. A. Tucker, V. Vanhoucke, V. Vasudevan, F. B. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, Tensorflow: Large-scale machine learning on heterogeneous distributed systems, CoRR abs/1603.04467. arXiv:1603.04467. http://arxiv.org/abs/1603.04467.

Accuracy (trueness and precision) of measurement methods and results - part 1: General principles and definitions.

Bengio, Y., Simard, P.Y., Frasconi, P., 1994. Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Networks 5 (2), 157–166. https://doi.org/10.1109/72.279181.

Bin, Y., Yang, Y., Shen, F., Xu, X., Shen, H.T., 2016. Bidirectional long-short term memory for video description. In: In: Kompatsiaris, Y., Li, J. (Eds.), Proceedings of the 2016 ACM Conference on Multimedia Conference, MM 2016, Amsterdam, The Netherlands, October 15–19, 2016 ACM, pp. 436–440. https://doi.org/10.1145/2964284.2967258.

Chung, J., Gülçehre, Ç., Cho, K., Bengio, Y., 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling, CoRR abs/1412.3555. arXiv:1412.3555. http://arxiv.org/abs/1412.3555.

Cui, J., Chen, H., Zhao, X., 2007. Research course of the cotton ipm and its prospect. Cotton Sci. 29 (S1), 100–112.

Ding, W., Taylor, G., 2016. Automatic moth detection from trap images for pest management. Computers Electron. Agric. 123, 17–28. https://doi.org/10.1016/j.compag.2016.02.003.

Duchi, J.C., Hazan, E., Singer, Y., 2011. Adaptive subgradient methods for online learning and stochastic optimization. J. Mach. Learn. Res. 12, 2121–2159 URL: http://dl.acm.org/citation.cfm?id = 2021068.

Feng, H., 2008. Ecological regulation of cotton aphid by using non-cotton habitat. J. Anhui Agric. Sci. 36 (9), 3752–3753.

Gao, X., Hu, D., Zhang, S., Luo, J., Wang, L.L.C., Cui, J., 2016. Preliminary study of the prevention and control of lysiphlebia japonica ashmead on aphis gossypii glover. China Cotton 43 (5), 8–10.

Genuer, R., Poggi, J.-M., Tuleau-Malot, C., 2010. Variable selection using random forests. Pattern Recogn. Lett. 31, 2225–2236. https://doi.org/10.1016/j.patrec.2010.03.014.

Graves, A., 2013. Generating sequences with recurrent neutral networks, CoRR abs/1506.02078. arXiv:1506.02078. http://arxiv.org/abs/1506.02078.

Hang, J., Zhang, D., Chen, P., Zhang, J., Wang, B., 2019. Classification of plant leaf diseases based on improved convolutional neural network. Sensors 19 (19), 4161. https://doi.org/10.3390/s19194161.

Hanley, J.A., McNeil, B.J., 1983. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 148, 839–843. https://doi.org/10.1148/radiology.148.3.6878708.

Hochreiter, S., 1991. Untersuchungen zu dynamischen neuronalen netzen, Master's thesis, Institut fur Informatik. Technische Universitat, Munchen.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Jiang, Z., Zhao, C., He, B., Guan, Y., Jiang, J., 2017. De-identification of medical records using conditional random fields and long short-term memory networks. J. Biomed. Inform. 75S, S43–S53. https://doi.org/10.1016/j.jbi.2017.10.003.

Kalchbrenner, N., Danihelka, I., Graves, A., 2015. Grid long short-term memory, CoRR abs/1507.01526. arXiv:1507.01526. http://arxiv.org/abs/1507.01526.

Karpathy, A., Johnson, J., Li, F., 2015. Visualizing and understanding recurrent networks, CoRR abs/1506.02078. arXiv:1506.02078. http://arxiv.org/abs/1506.02078.

Kelly, H.Y., Dufault, N.S., Walker, D.R., Isard, S.A., Schneider, R.W., Giesler, L.J., Wright, D.L., Marois, J.J., Hartman, G.L., 2015. From select agent to an established pathogen: The response to phakopsora pachyrhizi (soybean rust) in north america 105. pp. 905–916. doi:10.1094/phyto-02-15-0054-fi.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization, CoRR abs/1412.6980. arXiv:1412.6980. http://arxiv.org/abs/1412.6980.

LeCun, Y., Bengio, Y., Hinton, G.E., 2015. Deep learning. Nature 521 (7553), 436–444. https://doi.org/10.1038/nature14539.

Li, W., Chen, P., Wang, B., Xie, C., 2019. Automatic localization and count of agricultural crop pests based on an improved deep learning pipeline 9. doi:10.1038/s41598-019-43171-0.

Luo, J., Zhang, S., Ren, X., 2017. Research progress of cotton insect pests in china in recent ten years. Cotton Sci. 19 (5), 385–390.

Miao, Y., Gowayyed, M., Metze, F., EESEN: end-to-end speech recognition using deep RNN models and wfst-based decoding. In: 2015 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2015, Scottsdale, AZ, USA, December 13-17, 2015, IEEE, 2015, pp. 167–174. doi:10.1109/ASRU.2015.7404790.

Mikolov, T.A., 2012. Statistical language models based on neural networks. Ph.D. thesis. Brno University of Technology.

Pascanu, R., Mikolov, T., Bengio, Y., 2013. On the difficulty of training recurrent neural networks. In: Proceedings of the 30th International Conference on Machine Learning, ICML 2013, Atlanta, GA, USA, 16–21 June 2013, Vol. 28 of JMLR Workshop and Conference Proceedings, JMLR.org, pp. 1310–1318. http://jmlr.org/proceedings/papers/v28/pascanu13.html.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., VanderPlas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in python. J. Mach. Learn. Res. 12, 2825–2830 http://dl.acm.org/citation.cfm?id=2078195.

Prasetyo, S.Y.J., Agus, Y.H., Dewi, C., Simanjuntak, B.H., Hartomo, K.D., 2017. Geodata: Information system based on geospatial for early warning tracking and analysis agricultural plant diseases in central java 180. 012070. doi:10.1088/1757-899x/180/1/012070.

Ruder, S., 2016. An overview of gradient descent optimization algorithms, CoRR abs/1609.04747. arXiv:1609.04747. http://arxiv.org/abs/1609.04747.

Schuster, M., Paliwal, K.K., 1997. Bidirectional recurrent neural networks. IEEE Trans. Signal Process. 45 (11), 2673–2681. https://doi.org/10.1109/78.650093.

Singh, S., Gupta, M., Pandher, S., Kaur, G., Rathore, P., Palli, S.R., 2018. Selection of housekeeping genes and demonstration of rnai in cotton leafhopper, amrasca biguttula biguttula (ishida). Amrasca biguttula biguttula 13 (1), e0191116. https://doi.org/10.1371/journal.pone.0191116.

Sutskever, I., Vinyals, O., Le, Q.V., 2014. Sequence to sequence learning with neural networks. In: In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8–13 2014, Montreal, Quebec, Canadapp. 3104–3112 http://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks.

Tieleman, T., Hinton, G., 2012. Neural networks for machine learning, Tech. rep.

Wu, K., Lu, Y., Wang, Z., 2009. Advance in integrated pest management of crops in china. Chinese Bull. Entomol. 46 (6), 831–836.

Xie, C., Wang, R., Zhang, J., Chen, P., Dong, W., Li, R., Chen, T., Chen, H., 2018. Multi-level learning features for automatic classification of field crop pests. Computers Electron. Agric. 152, 233–241. https://doi.org/10.1016/j.compag.2018.07.014.

Xie, J., Liu, X., Dajun Zeng, D., 2018. Mining e-cigarette adverse events in social media using bi-lstm recurrent neural network with word embedding representation. J. Am. Med. Inform. Assoc.: JAMIA 25, 72–80. https://doi.org/10.1093/jamia/ocx045.

Zhang, W., zhong Jing, T., Yan, S., 2017. Studies on prediction models of dendrolimus superans occurrence area based on machine learning. J. Beijing Forestry Univ. 39 (1), 85–93.

Zhang, Q., Wang, H., Dong, J., Zhong, G., Sun, X., 2017. Prediction of sea surface temperature using long short-term memory. IEEE Geosci. Remote Sensing Lett. 14 (10), 1745–1749. https://doi.org/10.1109/LGRS.2017.2733548.

Zhao, R., Yan, R., Wang, J., Mao, K., 2017. Learning to monitor machine health with convolutional bi-directional lstm networks. Sensors (Basel, Switzerland) 17. https://doi.org/10.3390/s17020273.

Zhou, Y., Gao, P., 2014. Design and implementation of ensemble forecast system for crop diseases and pests. J. Computer Appl. 34 (S1), 141–144.