INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

Agriculture Technology Articles Classification Using Deep Learning Based textCNN Model

[1]D. RAMOHANREDDY, [2]M. NARESH,

[1]Associate Professor, Dept. of CSE, Newton's Institute of Engineering, Guntur, (A.P)

[2]Associate Professor, Dept. of CSE, Newton's Institute of Engineering, Guntur, (A.P)

Abstract: *Agriculture and its requirements are, at the time, quite challenging to handle. The bulk of the country's residents are dependent on agriculture for their income. Food production should also be increased to keep up with the World's population growth. Agriculture has benefited significantly from recent technological advancements. Agricultural experts are becoming excited by current technology. this paper implements the article classification system of textCNN convolutional neural network based on Pytorch framework. Python crawler technology is used to crawl the agricultural science and technology articles of China Agriculture Network, and calibrate them according to the original classification information, and divide them into training dataset and test dataset according to the ratio of 2/8. On the training obtained m odel, the best effect of the test set classification is 93.33%, and this model can be used to assist relevant technical personnel to achieve rapid sorting and classification of agricultural scientific and technical articles.*

Keywords: *Agricultural, Deep learning, Pytorch framework, textCNN.*

I. INTRODUCTION

Agriculture is a backbone of our country. Farmers have enough crops for their farm. Regardless, growing crops to maximize known income and manufacturing is generally medical. This may have been developed with technical assistance. Supervising that routine plants always want to have excellent strength, especially to combat diseases that can have important consequences on the elements of production for financial gain. The imaging method is of the highest quality, which has driven activity in agricultural applications. Detection of infection from plant photography (Saradhambal et al. 2018). This impact can be mitigated through aid for agricultural development. Most signs and symptoms number one are microscopic, so identity disease is restricted with the help of visible human

ISSN: 2057-5688

capabilities. . This method is boring and intensive over time. There is a need for a fashion tool that recognizes, categorizes and quantifies disorder symptoms.

In the case of an industrial disease indicating a wasted stock, the disease certainly appears because the traditional method of barbecue is the performance of the general body of the country. From the current evidence, symptoms can be identified. The disease is caused by microorganisms, and this is any factor that causes the disease. Control of the disease can be a difficult task. To a large extent, disease facts appear in the leaves of the plant, or the culmination or stems of the plant. Because of the difficulty of visual patterns, the precise identification of these diseases, lesions and developments that have been diagnosed visually has not been studied. In most cases, diseases are revealed in the leaves or stems of the plant.

The Indian economic system depends on agricultural productivity. Over 70% of country houses depend on agriculture. Agriculture pays about 17% of GDP (Santhosh Kumar et al. 2019) and employs more than 60% of the population. Therefore, the detection of plant diseases plays a significant key role in agriculture. Indian agriculture is made up of many plants like rice and wheat. Indian farmers grow sugar cane, oilseeds, potatoes, and non-food items such as coffee, tea, cotton, and rubber. All of these plants developed based on leaf and root electricity. Some things lead to a particular mess of plant leaves that spoil plants and ultimately affect the economy of the country. These huge losses can be avoided by early identification of plant diseases. Careful detection of plant diseases is required to enhance the field and the agricultural economy in our country. Various types of conditions kill the leaves on the plant. Farmers have greater difficulty in detecting these diseases, and they cannot take precautions regarding plant life due to a lack of knowledge about these diseases. Biomedicine is an area that has been found with plant diseases. Currently, in this specialty, image processing methods are a suitable, green, and reliable area for disease detection with the help of foliage. Farmers need fast and useful techniques to stumble upon all kinds of plant diseases that can save time. These systems can reduce the efforts and use of pesticides. As for the dimensions of yield in agriculture, scientists propose unusual ideas with the help of laboratories and structures for the environmental identification of plant leaf

ISSN: 2057-5688

diseases. The document we provide here is a survey of various forms of plant diseases and disease detection strategies by exclusive researchers

In order to promote the transformation of China's agricultural science and information technology, China has established many agricultural science and technology databases, and China Agriculture Network is one of them, aiming to provide farmers with agriculture-related technical support. Nowadays, most of the articles shared by agricultural information websites are collected from the network and need to be sorted manually, w hich often consum es a lot of time and effort. Based on this background, this paper proposes to use the Pytorch deep learning framework to establish an agricultural science and technology neural network classification model based on textCNN convolutional neural network [1] to achieve accurate and efficient classification of agricultural science and technology articles, greatly reducing the cycle of classifying articles on agricultural science and technology websites, so that farmers can get the latest technical support in a timely manner.

II. LITERATURE SURVEY

Trimi Neha et al. (2017) document discusses a method primarily based on digital image processing, which was used to detect and classify leaf diseases that are a gift in an agricultural plant of its kind. This will help design a unique approach to managing disturbance in an effort to be beneficial in agriculture. Automatic detection and analysis of the disease is installed in its exact signs and depth of drawing can be very beneficial to farmers. It is a key company for early detection of diseases in agricultural sciences. Living organisms such as fungi, bacteria, viruses and many others are the primary causes of plant diseases, so it is very important to improve the correct method in certain areas. All of these investigations are directed to the Early Detection and Plant Diseases category**.**

Sujeet Varshney et al. (2016) Agriculture has turned into more than just a way to feed the ever-growing population. It is very important where more than 70% of the population depends on agriculture in India. This way, it nourishes a wonderful number of people. Plant diseases affect humans immediately, or not directly, through health or economically. To find those plants the diseases we want in a fast and automated way. Diseases are analyzed

ISSN: 2057-5688

through different digital image processing techniques. In this article, we conducted a survey on exceptional digital image processing techniques to identify plant diseases**.**

Shitala Prasad et al. (2015) the plant disease detection system at a glance is a difficult and very expensive exercise, especially in developing countries like India. Designing and delivering a fast, reliable, and cellular face-to-face cellular response to such duties is a very good, realistic contribution to society. In this document a mobile client server structure is proposed for leaf disease detection and analysis, and a unique mixture of GWT and GLCM is proposed, which opens a new dimension to the sample's reputation. . The Mobile Disorder Diagnostic System is a sick patch in a distinctive vector of multiple precision and multiple paths. The mobile client takes the paper image and preprocesses it, splits the sick spots, and sends them to the pathology server, which reduces the sending cost. The server performs calculation functions: GWT-GLCM and (k) nearest neighborhood type properties. The result is sent to the customer's screen via SMS (Courier Service Provider) at a cost of 93% accuracy, in first class conditions. On the other hand, the document additionally brings awareness about human-mobile interface design (HMI), which is useful even for illiterate farmers, to automatically display their area at any level with a mobile click. Currently, Android is used to power this device, which can be easily expanded to different mobile operating structures.

Sonal et al. (2014) In order to get more goods at additional costs, there is a great need to handle the product better. Much research shows that world-class agricultural products can be reduced for many reasons. One of the most vital elements of these cute plant diseases. Hence, reducing plant diseases greatly improves the high quality of the product. The correct diagnosis of crop diseases in this topic can be very important for high manufacturing. Foliar is the primary fungal disease of cotton and occurs in all areas of cotton cultivation in India. In this article, I explain that technology strategies use markers captured by cotton-leaf spot shots cells and classify diseases using an aid vector machine. The workbook is trained in the harvest of smart agriculture, which consists of early detection of turbulence within forests, the selective benefit of fungicides, etc. These fully proposed

ISSN: 2057-5688

panels are based on hash strategies in which first-time captured images are processed to enrich them. Then, texture and shadow feature extraction techniques are used to extract features along with the boundaries, shape, color and texture of disease stains to stop disease.

Keyvan Asefpour et al. (2013) nowadays, artificial intelligence solutions such as digital image processing and artificial neural networks (ANN) have become important applicable techniques in phytomonitoring and plant health detection systems. In this research, an autonomous device was designed and developed for detecting two types of fungi (Pseudoperonospora cubensis,

Sphaerotheca fuliginea) that infect the cucumber (Cucumis sativus L.) plant leaves. This device was able to recognize the fungal diseases of plants by detecting their symptoms on plant leaves (downy mildew and powdery mildew). For leaves of cucumber inoculated with different spores of the fungi, it was possible to estimate the amount of our post inoculation (HPI) by extracting leaves' image parameters. Device included a dark chamber, a CCD digital camera, a thermal camera, a light dependent resistor lightening module and a personal computer. The proposed programmed for precise disease detection was based on an image processing algorithm and ANN. Three textural features and two thermal parameters from the obtained images were measured and normalized. Performance of ANN model was tested successfully for disease recognition and detecting HPI in images using back-propagation supervised learning method and inspection data. Such this machine vision system can be used in robotic intelligent systems to achieve a modern farmer's assistant in agricultural crop fields.

Chaudhary et al. (2012) this paper contains a set of rules for detecting disturbances Fractionation of image processing strategies on plant leaves done. This is the basic and vital stage of computing. Detecting plant diseases and their type. Sick spots Special in color but now not in depth unlike the plant Paper coloring. Therefore, RGB image coloring can be replayed more fragmentation of disease stains. In this work evaluation the effect of the CIELAB, HSI and YCbCr staining region within the immediate disorder detection procedure is complete. An average filter's output is used for Image smoothing. Finally, the threshold can be

ISSN: 2057-5688

calculated by Apply the Otsu method to the shadow component to detect turbulence Stain. A set of rules independent of background noise, the type of plant develops and shades of sick spots and experience Transferred to one family "Monocot" and "Dicot" Plant leaves with loose noise (white) and high heritage.

III. METHODOLOGY

Data Acquisition

First of all, open the home page of China's agricultural network of the amount of agricultural technology section, as shown in Figure 1. It mainly contains ten main sections, such as vegetables, garden flowers, and each section is divided into several subsections, and the more representative five subsections are selected here as the samples for this study. The target plates and their corresponding URLs are shown in Table 1

Table.1 Plate name and websites

Plate Name	url
planting technology	http://www.agronet.com.cn/Tech/list_c20663.html
process technology	http://www.agronet.com.cn/Tech/list_c20664.html
cultivation techniques	http://www.agronet.com.cn/Tech/list_c20858.html
reprouduction breeding	http://www.agronet.com.cn/Tech/List_c20332.html
cultivation technique	http://www.agronet.com.cn/Tech/List_c20334.html

The NetWork package capture analysis by the browser developer tool found that the web pages do not use dynamic web ajax and other related anti-crawling measures. Each page contains 30 article links, so when fetching article data, each panel needs to crawl 67 pages, and only 20 articles are fetched on the 67th page. The following is an example of the URL of the data page of the planting technology section. The following URL is the URL of the planting technology section, where x in _px indicates the number of pages. http://www.agronet.com.cn/Tech/list_c20663_px.html

After completing the renaming o f the plates, the sample data were split into training and test data in a two to eight ratio and then prepared for use in the training and test models

Once you understand the elements that make up the URL o f a web page, simply

ISSN: 2057-5688

cycle through the data from the web page and change the format to a json format to store it locally. **Data Cleansing**

1) Separation sample: After acquiring the source data, in order to process the data more easily later, and to distinguish the plates more easily, each plate name was renamed with the first three capital letters, and the results o f the plate name processing are shown in Table 2

Table.2 RENAME

original name	rename
planting technology	PLA
process technology	PRO
cultivation techniques	CUL
reprouduction breeding	REP
cultivation technique	CUL

The headings of the sample data are partitioned through Python's jieba library, and specific stop words are removed to eliminate interference when partitioning, and the word list is obtained through this operation. A partial example of the lexicon is shown in Table 3.

Table.3 DATASET LEXICON

technology	0	3952
Prevention	1	2401
Management	2	1967
method	3	1831
how	4	1707

PyTorch

PyTorch is an open-source machine learning platform for research and development (ML). This is a utility that's used to aid the DNN model development. It includes a diverse set of deep learning algorithms. It is based on Torch, a framework for developing and training deep neural networks. PyTorch provides a straightforward methodology for training and creating deep learning models, making it easy to do both

IV. PROPOSED NEURAL NETWORK MODEL

A. model design

In this paper, the agricultural science and technology article classification model mainly contains an embedding layer, a convolutional layer, a dropout layer, and a full connection layer to realize the data processing input, feature extraction, and result output. The convolutional network model is shown in Figure 1

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

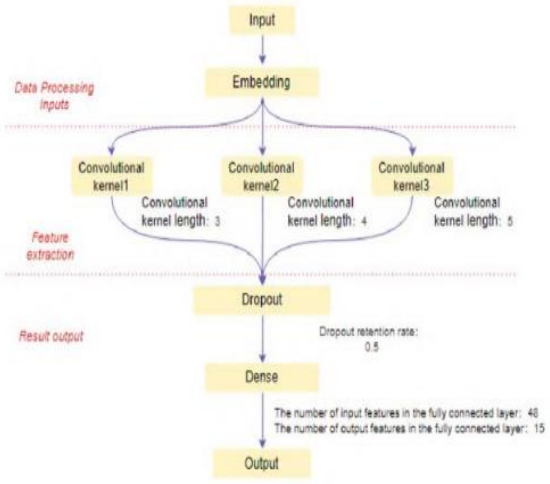

Fig.1 textCNN convolutional neural network model

In the Embedding layer set the size of the embedding dictionary to 11622 and the depth of each embedding vector to 60. A total of three convolutional kernels are set in the convolutional layer, set the channel of the depth signal of the input data body to 1, the depth of the output data body to 16, and the length of the convolutional kernels to 3, 4, and 5. in the Dropout layer randomly discard the neuron probability settings is 0.5. the number of input features in the fully connected layer are 48 and the number of output features is 15. In order to prevent overfitting when training the convolutional network model, the Dropout layer is introduced and 50% of the neurons are randomly discarded to achieve a smaller neural network with half of the neurons randomly discarded each time the layer is iteratively trained, and the overfitting and time-consuming problems are well solved.

1) Data Processing Inputs: Data input is the first step of the training model, but also a crucial step, after the training data for data cleaning, building character tables, building word vectors and other operations, as long as the converted two-dimensional matrix into the textCNN convolutional neural network can start the training of the network

2) Feature Extraction: With the help of textCNN convolutional neural network can quickly select the same class of agricultural science and technology articles with the same features, this paper constructs three convolutional kernels with the kernel lengths of 3, 4, and 5 respectively, and performs convolutional processing on the input word vectors, through different convolutional kernels can detect the different sizes of multiple adjacent characters, so the location of the characters can be disregarded during training.

3) Result Output: There may be overfitting when training the network, so this paper uses the Dropout layer to implement overfitting of the convolutional

ISSN: 2057-5688

and fully connected layers, i.e., half of the neurons are randomly discarded each time the layer is trained to reduce the chance of overfitting during training.

B. Experimental Testing

In this paper, Chinese character-level text classification using TextCNN convolutional neural network is used to achieve classification of agricultural science and technology articles mainly through article titles, and a relatively high classification accuracy is obtained to som e extent. When the textCNN model was trained, the classification accuracy was found to be close to 0.9333, w hich confirmed that the accuracy of the model was ideal for classification and could replace manual classification to some extent.

Model validation

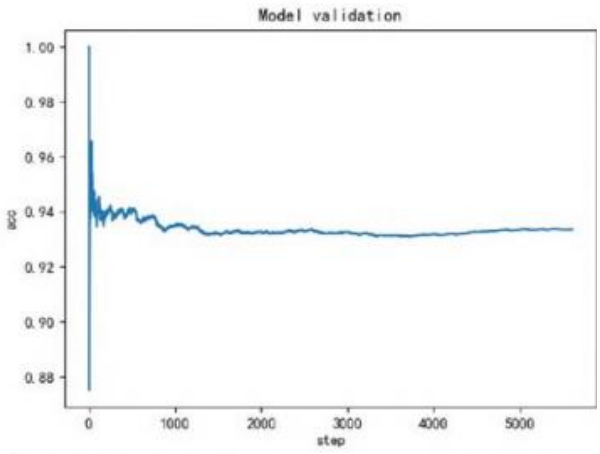

Fig.2 Model validation

V. CONCLUSION

In this paper, a TextCNN-based deep convolutional neural network classification system for agricultural science and technology articles is proposed to address the problem of time-consuming and laborious classification o f a large number of agricultural science and technology articles. Through the TextCNN convolutional neural network model, the system can analyse the relationship between the attributes of each keyword and its corresponding keywords, determine the nature of article titles, and then classify the article categories to replace manual classification. The system can be initially applied to provide assistance to relevant personnel and improve the classification efficiency, but there is still potential room for improvement of the model in this paper, which needs further research and improvement.

REFERENCES

[1]Bao Guo,Chunxia Zhang,Junmin Liu,Xiaoyi Ma. Improving text classification with weighted word embeddings via a multi-channel TextCNN model[J]. Neurocomputing,2019,363.

ISSN: 2057-5688

[2] Hong Sheng Xu,Yong Liang Li,Chao Li. Application of AJAX Technology and SSH Framework in Development of Electronic Commerce System[J]. Advanced Materials Research,2014,3245.

[3] Dongwei Xu,Chenchen Wei,Peng Peng,Qi Xuan,Haifeng Guo. GE-GAN: A novel deep learning framework for road traffic state estimation[J]. Transportation Research Part C,2020,117.

[4] Shoko Chisaki,Ryoh Fuji-Hara,Nobuko Miyamoto. Combinatorial designs for deep learning[J]. Journal of Combinatorial Designs,2020,28(9).

[5] Qingchen Zhang,Changchuan Bai,Zhuo Liu,Laurence T. Yang,Hang Yu,Jingyuan Zhao,Hong Yuan. A GPU-based residual network for medical image classification in smart medicine[J]. Information Sciences,2020,536.

[6] Henrietta Baker,Matthew R. Hallowell,Antoine J.-P. Tixier. AI-based prediction of independent construction safety outcomes from universal attributes[J]. Automation in Construction,2020,118.

[7] Amit Dvir,Angelos K. Marnerides,Ran Dubin,Nehor Golan,Chen Hajaj. Encrypted video traffic clustering demystified[J]. Computers & Security,2020,96

[8] Kussul N, Lavreniuk M, Skakun S, Shelestov A. Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters 2017;14(5):778–782.

[9] Li H, Ota K, Dong M. Learning IoT in edge: Deep learning for the Internet of things with edge computing. IEEE Network 2018;32(1):96–101.

[10]Liakos K, Busato P, Moshou D, Pearson S, Bochtis D. Machine learning in agriculture: A review. Sensors 2018;18(8):267