Fuzzy Systems and Soft Computing
ISSN : 1819-4362

PLANT DISEASE IMAGING USING CRNN PIPELINE FOR BIG DATA PLANT CLASSIFICATION

S. Gopinath Department of Information Technology, Gnanamani College of Technology, India sgopicse@gmail.com
G. Kowsalya Department of Computer Science and Engineering, Gnanamani College of Technology, India gtkowsalya@gmail.com
K. Sakthivel Department of Computer Science and Business Systems, K.S.Rangasamy College of Technology, India sakthivelk@ksrct.ac.in
C. Kavitha Department of Computer Science, Thiruvalluvar Govt. Arts College, India kavithasakthivel@gmail.com

Abstract:
Because of the current advancements in big data technologies, data from the farming industry is now being included into big data. There are no traditional strategies for dealing with such massive volumes of data. Large volumes of data need the use of a parallel calculation and analysis paradigm. To handle massive picture collections, it is required to employ a big data analytic framework. This article describes an automated big data platform for determining the condition of plant diseases. This framework is made up of a succession of actions that lead to a last step. In the ordering, a new image classifier is applied. A CRNN classification technique was used to create the picture classifier. The classifier is intended to distinguish between normal and aberrant leaves. Classifying Images of banana, pepper, potato, and tomato plants were created using big datasets. This is compared to other existing large data asset categorization approaches as CNN, RNN, deep neural network, and forward and backward propagation artificial neural network. The results suggest that the proposed technique detects and classifies sick plants better than previous methods: CNN (96.14%), RNN (95.07%), and DNN (95%). Artificial NN through Forward (94.97%) and backward propagation technique (93.33%).

Keywords:
Big data, CRNN, Weighted Naive Bayesian Network Classifier Algorithm

1. INTRODUCTION

Plant diseases are a huge hazard to agriculture, reducing crop production and quality. Furthermore, due to the complexity of disease patterns and the absence of domain specialists available for monitoring these diseases, farmers may find it difficult to identify and diagnose plant illnesses (Pandian et al., 2022). Furthermore, manual inspection for pest identification takes time and is not always reliable. Artificial intelligence methods such as deep learning and CNN have emerged as viable solutions for plant disease diagnosis and recognition to solve these difficulties (Chandy, 2019). By automatically evaluating vast volumes of data and discovering patterns that may not be clearly discernible through manual observation, these AI systems have the potential to give optimal solutions for plant disease diagnosis.

1.1 Deep CNN for Plant Disease Identification

Deep CNN Having recently demonstrated fashion abilities in the realm of computer vision promising results in factory complaint identification and recognition. These neural networks are specifically designed to reuse and dissect visual data, making them well- suited for detecting patterns and abnormalities in factory images. The use of deep CNN for factory complaint identification is grounded on their capability to prize features from images, allowing for accurate bracket Proper and timely complaint identification, containing early forestallment, takes no way been further significant in this evolving terrain. Factory pathology can be sensed in numerous ways. Certain conditions take no egregious indications or the goods are too late to be detected and also a complex analysis is needed. Still, utmost conditions produce some incarnation in the visible diapason, so good professional examination is the primary factory discovery fashion. A factory pathologist must have good observation chops to recognize characteristic indications in order toward make an accurate opinion of factory conditions [4]. Variations in the symptoms displayed by diseased shops can lead to misdiagnosis, as identification may be more delicate for an amateur than for an expert pathologist. Gardening suckers and skilled specialists can greatly take advantage of automatic systems designed to determine factory health grounded on the factory's appearance and visual symptoms to determine diagnose the complaint. A computer vision provide openings to enlarge and decorate perfection crop protection practices as well as expand the request for perfection computer vision operations for husbandry. To descry and classify factory conditions, the use of communal digital image processing technologies similar by way of color investigation and thresholding [3] has been used. Different deep literacy styles are presently used for factory complaint discovery and the best popular among them is CNN. A literacy is a novel trend in machine literacy, with slice- edge results in numerous exploration zones, with computer vision, medicinal, and bio-informatics. Deep literacy aids from the ability to use rare data straightly without the need for homemade work [5-6]. The usage of deep literacy, for dual causes, has lately yielded worthy outcomes both educationally and technologically [5]. First, each day a huge quantum of data is created. As a result, this information may be utilized to train a deep model. Eventually, the computational rule of a plate's processor helps train and operate models that are deep in computer community. The determination of this knowledge to introduce deep literacy as a system for classifying factory conditions, fastening on factory splint images. In this work, an automatic big data frame for crop complaint bracket is accessible. This structure contains of a sequence of actions leading to the last step, where bracket is performed using a new image classifier. The photograph classifier is designed the use of the CRNN bracket set of rules. The classifier is planned in a method that allows bracket among usual and unusual leaves. Grounded on the below algorithms (Convolutional Recurrent Neural Network), it's set up that CNN or histogram parcel is used for bracket. It has difficulty directly detecting splint conditions and should only be used for conditions on one splint. But in this design, neural network is used to classify splint conditions.

2. Related Work

Dudney,J, etal.,(2021)[1], Changes in the types of factory contagious conditions are anticipated under climate change. As factory conditions multiply, arising abiotic and biotic relations are anticipated to regulate their distribution, foremost to unanticipated alterations in complaint trouble. Still, substantiation for these complex distributional changes due to climate trade remains largely academic.

Buja I, etal., 2021[2], proposes a new method to mortal conditioning that contributes significantly to the transnational spread of cerebral adversity. Pathogen- related food losses are now responsible for reduced volume and quality of yields as well as reduced value and fiscal returns. Thus, "beforehand discovery" combined with "presto, accurate and affordable" opinion has also come the fresh mantra in factory pathology and especially for arising or germline conditions. Complicated complaint spreads to asymptomatic individualities with verbose original symbols and signs but also come hard to bear.

Zhang,J., etal.,(2021)[5], proposed a area- motivated involving critical element investigation for the position of cells and their cornucopia in bitsy images.

Fuentes,A., 2022[9], suggested that factory conditions plus pest is a major concern in the agrarian sector. More perfect and briskly discovery of crop conditions and pests can help grow early treatment while significantly falling fiscal damages. Recent trends in DNN have enabled experimenters to significantly ameliorate the delicacy and common structure of object discovery. Associated to state-of-the-art technology.

Brahimi.,etal (2017)[9] use a vast data set. This, imageries of tomato leaves depicting nine conditions are registered. The work presents Convolutional Neural Network as a algorithm to train a bracket system. Automatic point birth over rare image is one and only of the biggest advantages of CNN. This study is used visual styles to dissect patterns to diagnose pathological areas of leaves. The three main sensor families considered by Fuentes,A., etal.(2017)[10], which uses briskly indigenous CNN, completely indigenous CNN, and single- shotmulti-box sensor called deep literacy meta-armature. Every of this Meta infrastructures are joined thru deep extractors similar by way of VGG networks and residual networks. This presentation of deep infrastructures and point extractors are presented and extra way is proposed to growth the delicacy as well as the number of false cons when training on original and global reflections demand as well as class data. The authors have developed comprehensive test systems for large tomato pesto and complaint datasets containing grueling complaint images that include numerous intra- andinter-class variants similar as affected and spot of the factory.

Zhang., etal (2017)[11] propose a novel system to identify cucumber conditions including 3 channels K- means parts diseased splint images, excerpts lesion information about shape and color, and classifies meager splint images. Some of the benefits of this system is that SR spatial bracket be able to effectively reduce computational cost and increase recognition performance.

Vuong,G., etal.(2017)[12] totally estimated the recital of deep networks trained from the DL method, meliorated using transfer literacy. The depth of complaint in the red-black apple image was analyzed by Plant dataset by a sequence of deep CNN.

Lu,Y., etal.(2018)[13] proposed a new system to identify rice conditions grounded on the ways used by CNN. The CNN were trained to isolate 20 common rice conditions with 510 natural images of affected and vigorous rice leaves in addition to stems taken during rice trials.

3. Proposed Method

This sector, we existent an automated frame illustrated in Figure (i) that aids categorize banana factory complaint status. This frame includes a series of processes including i) image accession ii) preprocessing iii) point birth iv) bracket by CRNN. The last stage is the bracket method, using the Convolutional Recurrent Neural Network algorithm. The classifier is framed in a mode that allows bracket among usual and unusual banana leaves.

3.1. Image acquisition

These pictures were taken with a camera with dissimilar exposures, dimensions, backgrounds, acts and lighting.

3.2 preprocessing

A big No. of images of vigorous and diseased plants can be set up in data deposited in original or worldwide depositories. All image contains three RGB canals. It needs to be tried for the connection of our method to together RGB and grayscale pictures in our trials. To do this, we perform a preprocessing step by converting each image in our dataset to a 256x256 pixel grid.

The incidence of noise is eliminated since the accrued photo sections, where a conclusion is made or not the pixel is easy or broken. When pixels are stored, a scale of 0 - 255 was used, pixels are decoded as degraded. The window range is increased to 5 cross 5 once removing the broken 3 cross 3 pixels in the original scene. The window dimension is enlarged to permit processing of total damaged pixels in a particular processing system.

The average value is formed by these damaged pixels and the window range is reduced to the average value. In the present window, the altered mean values are replaced using the damaged pixel.

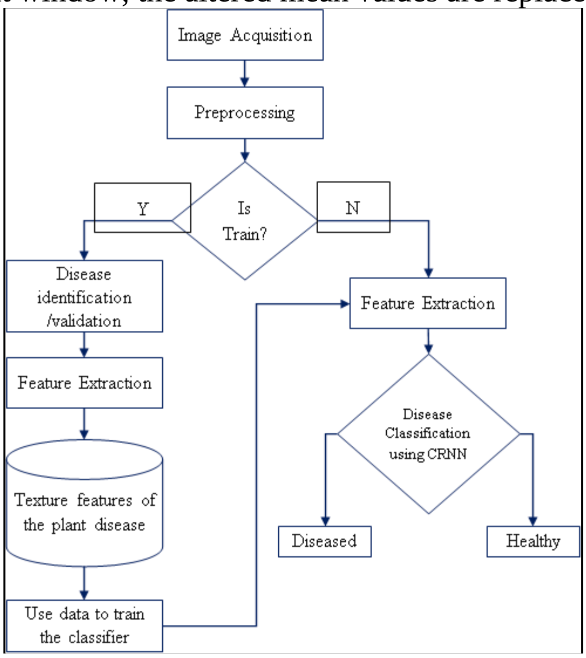

Fig 1: Architecture of Suggested Classification model [14]

Also, after all pixels in the frame are noticed to be distorted in current pixel will be exchanged by the former pixel numbers. Simply move the present window to the coming pixel handling element once handling is complete through the main window. Once the pixel is protected, it's decoded to drop from 0 to 255.

After removing the injured pixels since the frame at the main stage, the frame value is enlarged. The frame dimension is expanded to permit a processing of all broken pixels in a given system. These broken pixels produce a midpoint, through the window range dwindling to the midpoint. The changed mean values are exchanged by the spoiled pixels in the present window. Also, when complete pixels in the frame are noticed, the present pixel will be changed by the former pixel.

Simply change the current frame to the coming pixel handling element after the first frame ends.

3.3 **Feature Extraction**

Practiced (primer) features are uprooted from images to construct point vectors. For Gabor Sea transfigure (GWT) to prizemulti-scale features, color moments are used, e.g.in color moments. To prizeTexture parcels, GLCM or Gray Level-occurrence Matrix is used. The GLCM grounded SIFT includes the following texture features uprooted since the image by the equations

1. Auto correlation = $\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(x^{*}y^{*}I_{clu}\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(x^{*}y^{*}I_{clu}$(x and y)) (1)
2. Correlation= $\sum_{i,j=0}^{G-1}\frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i\sigma_j}\sum_{i,j=0}^{G-1}\frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i\sigma_j}$ (2)
3. Entropy = $\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(I_{clu}\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(I_{clu}$ (x, y) * $\log I_{clu}\log I_{clu}$(x and y) + $\in\in$)) (3)
4. Contrast = $\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(x-y)^2 * I_{clu}(x,y)\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(x-y)^2 * I_{clu}(x,y)$ (4)
5. Energy = $\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(I_{clu}(x,y))^2\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(I_{clu}(x,y))^2$ (5)
6. Cluster shade = $\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(x+y-\mu_{1-\mu_2})^3\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(x+y-\mu_{1-\mu_2})^3$ (6)
7. Cluster prominence = $\sum_{x=1}^{R}\sum_{y=1}^{R}(x+y-\mu_{1-\mu_2})^4\sum_{x=1}^{R}\sum_{y=1}^{R}(x+y-\mu_{1-\mu_2})^4 * I_{clu} I_{clu}$(x and y) (7)
8. Dissimilarity = $\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(x-y) * I_{clu}(x,y)\sum_{x=1}^{Re}\sum_{y=1}^{Cl}(x-y) * I_{clu}(x,y)$ (8)
9. Homogeneity = $\sum_{x=1}^{Re}\sum_{y=1}^{Cl}\frac{I_{clu}(x,y)}{(1+(x-y)^2)}\sum_{x=1}^{Re}\sum_{y=1}^{Cl}\frac{I_{clu}(x,y)}{(1+(x-y)^2)}$ (9)

since, x and y are directs row and column of an record in the cooccurrence matrix. Lclu (x and y) is the [x and y]th access in a gray-tone spatial dependence matrix. G is the dimension, Pij are probabilities calculated for values in GLCM is the average deviation of pi and pj, where pi & pj are partial probability density function are the mean of pi & pj. Further the shape-based features contains the following:

1. Variance $\sigma_i\sigma_i$ = $\sum_{i,j=0}^{G-1}(i-\mu_i)p(i,j)\ and\ \sigma_j\sum_{i,j=0}^{G-1}(i-\mu_i)p(i,j)\ and\ \sigma_j$ = $\sum_{i,j=0}^{G-1}(j-\mu_j)p(i,j)\sum_{i,j=0}^{G-1}(j-\mu_j)p(i,j)$ (10)
2. Median B(x,y,t) = median {I(x, t, t-i)} (11)
3. Co-variance cov (x and y) = $\frac{1}{n}\sum_{i=1}^{n}(x-\bar{x})\frac{1}{n}\sum_{i=1}^{n}(x-\bar{x})$ ($y-\bar{y}y-\bar{y}$) (12)

4. Correlation= $\left|\sum_{i,j=0}^{G-1}\frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i\sigma_j}\right|\sum_{i,j=0}^{G-1}\frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i\sigma_j}\Bigg|$ (13)

3.4 CRNN Classification

CRNN [7] consists of five convolutional layers, one retrogression subcaste, and two totally associated layers. The Convolutional Neural Network layers are used to learnmid level graphic models are analogous to the initial 5 layers of the prevalent Alex Net 7 layers. The Recurrent Neural Network sub-castes are used to acquire the spatial necessity among graphic models at the meso- position. The last 2 layers, 2 completely connected Recurrent Neural Network labors will be composed and a global demonstration of the image will be learned. Also, the soft- maximum subcaste for bracket must be applied for the N- Way (N represents the No. of layers).

CRNN processes are as follows:

1. Select the data set
2. Prepare the exercise data set:

Making our dataset for exercise will include requiring routes and for generating groups (tickets), resizing our pictures.

3. Generate preparation data

The line-up is an array with a view to include the pixel values of the image and the catalog at which the photograph is inside the categories list.

4. Merge datasets
5. Labeling and structures: In this two-list format will be used in sorting by NN
6. Normalize X then convert tickets to categorical statistics
7. Separate X, Y for usage in CRNN
8. Identify, collect and train CRNN models and
9. Model exactness and score

In these nine simple stages, you'll be prepared to train your individual CNN model and explain real world problems by these helps.

4. Results and Discussions

CRNN [7] contains 5 convolutional layers, one retrogression subcaste, and 2 totally linked layers. The CNN layers are used to learnmid-level graphic models are analogous to the starting 5 layers of the general Alex Net 7 layers. The Recurrent Neural Network sub castes are used to study the spatial dependence among graphic models at the meso- position. In the last 2 layers, 2 completely connected Recurrent Neural Network labors will be composed and a worldwide representation of the image will be learned. Also, the soft- maximum sub-caste for bracket must be smeared to the N Way (N represents the No. of layers).

Table. 4.1(a) Input

Image Type	No. of Images
Banana (Healthy)	2500
Pepper (Healthy)	2406
potato (Healthy)	2239
potato (Healthy)	2380

The entire feature set of the data is treated by the CRNN technique to control the optimal qualities for classification. The junction curve to find the optimum feature using the future method is resolute. Then, the CRNN is expert using the exercise set and verified using the test set.

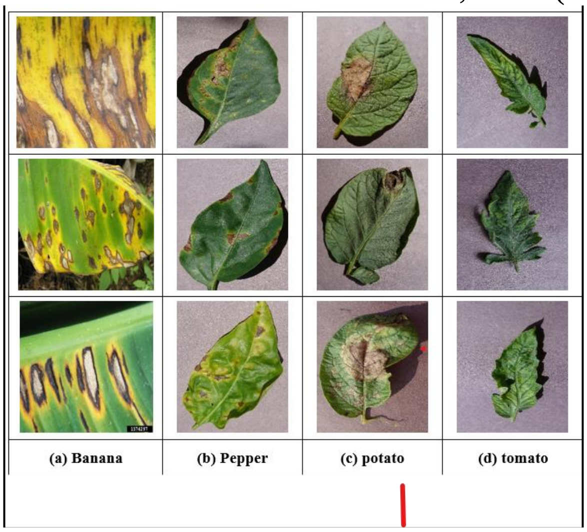

Fig 2. Disease dataset - various leaves

4.1 Results of CRNN

Table 4.1(b) shows the correctness of CRNN. We see that 20% (train) and 80% (test) has an Accuracy value of 0.979, 40% (train) and 60% (test) has an accuracy value of 0.9688, 50% (train) and 50% (test) has an accuracy value of 0.987, 60% (train) and 40% (test) takes an accuracy value of 0.979 and 80% (train) and 20% (test) has an accuracy value of 0.9123

Table. 4.1(b) Replication parameters

Classifier		**Precision**
Testing	**Training**	
80	20	.979
60	40	.9688
50	50	.987
40	60	.979
20	80	.9123

Table 4.1(c) displays the recall of Convolutional Recurrent Neural Network. We see that 20% (train) and 80% (test) has an accuracy value of 0.9760, 40% (train) and 60% (test) has an accuracy value of 0.9759, 50% (train) and 50% (test) has an accuracy value of 0.9879, 60% (train) and 40% (test) has an accuracy value of 0.9577 and 80% (train) and 20% (test) has an accuracy value of 0.9188.

Table 4.1(d). displays the accuracy value of Convolutional Recurrent Neural Network. We see that 20% (train) and 80% (test) has an accuracy value of 0.9811, 40% (train) and 60% (test) has an accuracy value of 0.9810, 50% (train) and 50% (test) has an accuracy value of 0.9952, 60% (train) and 40% (test) has an accuracy value of 0.9626 and 80% (train) and 20% (test) has an accuracy value of 0.9238.

Table. 4.3(c) Replication parameters

Classifier		**Recall**
Testing	**Training**	
80	20	.9760
60	40	.9759
50	50	.9879
60	60	.9577
20	80	.9188

Table. 4.1(d) Replication parameters

Classifier		**Accuracy**
Testing	**Training**	
80	20	.9811

60	40	.9810
50	50	.9922
40	60	.9626
20	80	.9238

5 Conclusion

In this research work, we present an automated big data frame work for factory complaint bracket. This frame contains a sequence of operations leading to the last step of the CRNN system, where bracket is performed using a new image classifier method. The CRNN bracket technique is used to create the picture classifier. The classifier is considered in a way that permits bracket among usual and unusual leaves. The study was performed on vigorous and harmful pepper, harmful banana, harmful tomato and harmful potato leaves. Leaf image bracket on large data sets is matched with other existence big data factory bracket ways and conventional bracket ways. The results show that the proposed system can ameliorate the discovery and bracket of diseased factory. In the future, a study on the optimal use of U-net and histogram- grounded point birth will be suitable for double and multiclass bracket of splint conditions.

Reference

[1] Dudney, J., et al. Nonlinear shifts in infectious rust illness due to climate change. Nat Commun, 12, 5102 (2021).
[2] Buja I, etal. Progresses in Plant Disease Finding and Monitoring: From Traditional Analyses to In-Field Diagnostics. *Sensors*. 2021; 21(6):2129.
[3]Upadhyay, etal. A approach for rice plant sicknesses classification with deep CNN. Int. j. inf. tecnol. (2021).
[4] Sladojevic, S., etal. DNN Based Recognition of Plant Infections by Image Classification. Computational Intelligence and Neuroscience, 2016, pp. 1–11
[5] Sladojevic, S., etal. DNN Based Recognition of Plant Infections by Image Classification. Computational Intelligence and Neuroscience, 2016, pp. 1–11
[6] Zhang,et al. Fast and exact automated recognition of the main cells from fecal images based on Faster R-CNN. Sci Rep 11, 10361 (2021).
[7] Mokhtar, U., etal. SVM based detection of tomato plants diseases. In *Intelligent Systems' 2015*, pp. 641-652.
[8] Ferentinos, K. P. Deep learning models for plant disease finding and diagnosis. *Computers and Electronics in Agriculture*, *145*, pp.311-318, 2018.
[9] Brahimi, M., Boukhalfa, K., & Moussaoui, A. Deep learning for tomato diseases: cataloguing and indications visualization. *Applied Artificial Intelligence*, *31*(4), pp. 299-315, 2017.
[10] Fuentes, etal. A robust deep learning based detector for real time tomato plant diseases and pests recognition. *Sensors*, *17*(9), 22.
[11] Zhang, S., et al (2017). Leaf image based cucumber sickness recognition by sparse representation classification. *Computers and electronics in agriculture*, *134*, pp. 135-141.
[12] Wang G, et al. Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. Comput Intell Neurosci. 2017,
[13] Lu, Y., et al. Identification of rice diseases using deep convolutional neural networks. *Neuro-computing*, *267*, 378-384, 17.
[14] Oppenheim, D., &Shani, G. Potato disease classification using convolution neural networks. Advances in Animal Biosciences, 8(2), 244-249, 2017. 2917536.
[15] Gopinath S, Sakthivel K, Lalitha S, A Plant Disease Image Using Convolutional Recurrent Neural Network Procedure Intended for Big Data Plant classification, Journal of Intelligent & Fuzzy Systems, 43(4), 4173-4186, 22