RESEARCH ARTICLE

An efficient neural network approach for plant disease detection

Dr.Nita Thakare [1] , Anchal Ramtekkar [2], Mrunalini Dhoble [3] , Ayushi Saxena [4], Anchal Shahu [5], Achal Giradkar [6]

Priyadarshini College Of Engineering - Nagpur

ABSTRACT

Convolutional neural networks have incontestable progressive performance in image classification associate degreed varied alternative laptop vision tasks. disease detection is a crucial space of deep learning that has been addressed by several recent techniques. However, there's a dire have to be compelled to optimize these solutions for resource-constrained moveable devices corresponding to smartphones. this is often a difficult downside as a result of deep learning models are resource in depth in nature. This paper proposes an economical method to consistently classify plant disease symptoms victimization convolutional neural networks. These networks are memory efficient and once plus the projected coaching configuration it permits speedy development of business applications by reducing the coaching times. Another vital downside arises with the improper distribution of samples among categories referred to as the category imbalance problem, that could be addressed by using an easy applied math methodology. Transfer learning is a legendary technique for training tiny datasets which transfers pre-trained weights learned on an oversized dataset. However, throughout transfer learning, negative transfer learning is a common problem. Therefore, a stepwise transfer learning approach is projected which might facilitate in quick convergence, whereas reducing overfitting and preventing negative transfer learning during data transfer across domains. The system is trained and evaluated on 2 disease datasets i.e., PlantVillage (a in public out there dataset) and pepper disease dataset provided by the National Institute of agriculture and flavorer Science, Republic of Korea. The pepper dataset is especially difficult because it contains pictures from totally different elements of the plant corresponding to the leaf, pulp, and stem. The projected system has dominated the previous works on the PlantVillage dataset and achieved 90.08% accuracy on the Pepper dataset and PlantVillage datasets.

Keywords:- Disease detection, Convulational neural networks, image classification, transfer learning..

I. INTRODUCTION

India is an agricultural country wherein most of the population depends on agriculture. Research in agriculture is aimed towards increase of productivity and food quality at reduced expenditure, with increased profit. Agricultural production system is an outcome of a complex interaction of soil, seed, and agro chemicals. Vegetables and fruits are the most important agricultural products. In order to obtain more valuable products, a product quality control is basically mandatory. Many studies show that quality of agricultural products may be reduced due to plant diseases. Diseases are impairment to the normal state of the plant that modifies or interrupts its vital functions such as photosynthesis, transpiration, pollination, fertilization, germination etc. These diseases are caused by pathogens viz., fungi, bacteria and viruses, and due to adverse environmental conditions. Therefore, the early stage diagnosis of plant disease is an important task [1]. Farmers require continuous monitoring of experts which might be prohibitively expensive and time consuming. Therefore looking for fast [2], less expensive and accurate method to automatically detect the diseases from the symptoms that appear on the plant leaf is of great realistic significance. This enables machine vision that is to provide image based automatic inspection, process control and robot guidance.

II. LITERATURE REVIEW

A Proliferation of literature is available in plant leaf disease detection. We will highlight some of the key contributions. A methodology for detecting plant diseases early and accurately using diverse image processing techniques has been proposed by Anand H.Kulkarni et al. [1], where Gabor filter has been used for feature extraction and ANN based classifier has been used for classification with recognition rate up to 91%. F. Argenti, et al. [2] proposed a fast algorithm for calculating parameters of co-occurrence matrix by supervised learning and maximum likelihood method for fast classification. Homogenize techniques like sobel and canny filter has been used to identify the edges by P.Revathi et al. [3]. These extracted edge features have been used in classification to identify the disease spots. The proposed homogeneous pixel counting technique for cotton diseases detection (HPCCDD) algorithm has been used for categorizing the diseases. They claim the accuracy of 98.1% over existing algorithm. Tushar H Jaware et al. [4] proposed a novel and improved k-means clustering technique to solve low-level image segmentation. Spatial gray-level dependence matrices (SGDM) method has been used for extracting statistical texture features by Sanjay B. Dhaygude et al. [5]. RGB images have been converted into Hue Saturation Value (HSV) color space representation and showed the H, S and V components. Mokhled S. Al-Tarawneh [6] presented an empirical investigation of olive leaf spot disease using auto-cropping segmentation and fuzzy c-means classification. Rgb to Lab colorspace and median filter used for image enhancement. At end present comparative assessment of fuzzy c-means and k-mean clustering. Detection of unhealthy region and classification using texture features has been proposed by S. Arivazhagan, et al. [10]. Their algorithm has been tested on ten species of plants namely banana, beans, jackfruit, lemon, mango, potato, tomato and sapota. 94.74% accuracy has been achieved by Support vector machine (SVM) classifier. Dheeb Al Bashish, et al. [11] developed neural network classifier based on statistical classification and could successfully detect and classify the diseases with a precision of around 93%. A Research of maize disease image recognition of corn based on BP networks effectively identified by Song Kai et al. [12] where YCbCr color space technology is used to segment disease spot, Co-occurrence matrix (CCM) spatial gray level layer is used to extract disease spot texture feature, and BP neural network has been used to classify the maize disease. The applications of K-means clustering, BP neural networks had been formulated for clustering and classification of diseases that affect on plant leaves by H. Al-Hiary, et al. [13]. They provide adequate support for accurate detection of leaf diseases. The proposed algorithm has been tested on five diseases viz. Early and late scorch, cottony and ashen mold , tiny whiteness. Menukaewjinda et al. [14] tried another ANN, i.e. back propagation neural network (BPNN) for efficient grape leaf color extraction with complex background. They also explore modified self organizing feature map (MSOFM) and genetic algorithm (GA) and found that these techniques provide automatic adjustment in parameters for grape leaf disease color extraction. Support vector machine (SVM) has been also found to be very promising to achieve efficient classification of leaf diseases. 21 color, 4 shape and 25 texture features has been extracted by Haiguang Wang et al. [15] and principal component analysis (PCA) has been performed for reducing dimensions in feature data processing, then back-propagation (BP) networks, radial basis function (RBF) neural networks, generalized regression networks (GRNNs) and probabilistic neural networks (PNNs) has been used as the classifiers to identify diseases.

III. METHODOLOGY

There are five main steps used for the detection of plant leaf diseases as shown in fig.3. The processing scheme consists of image acquisition through digital camera or web, image pre-processing includes image enhancement and image segmentation where the affected and useful area are segmented, feature extraction and classification. Finally the presence of diseases on the plant leaf will be identified. In the initial step, RGB images of leaf samples were picked up. The step-by-step procedure as shown below: 1) RGB image acquisition; 2) convert the input image into color space; 3) Segment the components; 4) obtain the useful segments; 5) Computing the texture features; 6) Configuring the neural networks for recognition. 3.1. Image acquisition Firstly, the images of various leaves acquired using a digital camera with required resolution for better quality. The construction of an image database is clearly dependent on the application. The image database itself is responsible for the better efficiency of the classifier which decides the robustness of the algorithm. Image pre-processing In the second step, this image is pre-processed to improve the image data that suppress undesired distortions, enhances some image features important for further processing and analysis task. It includes color space conversion, image enhancement, and image segmentation. The RGB images of leaves are converted into color space representation. The purpose of the color space is to facilitate the specification of colors in some standard accepted way. RGB images converted into Hue Saturation Value (HSV) color space representation. Because RGB is for color generation and his for color descriptor. HSV model is an ideal tool for color perception. Hue is a color attribute that describes pure color as perceived by an observer. Saturation termed as relative purity or the amount of white light added to hue and value means amplitude of light. After the color space transformation process, hue component used for further analysis. Saturation and value are dropped since it does not give extra information [5]

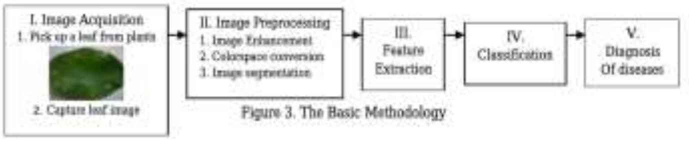


Figure 3. The Basic Methodology

Ycbcr color system is a common color space, which is applied by the most widely used jpeg image. Y, cb and cr, indicates a luminance component and two color component signals respectively. Different from other color space, Ycbcr color space is orthogonal, which fully takes important factors of composition of RGB from other colors into account. Ycbcr color space model is often used in image compression [12]. A , U, and Cr components from LAB [6], UVL , and Ycbcr color space used to extract affected leaf color with the purpose of less illumination effects [14].

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms.

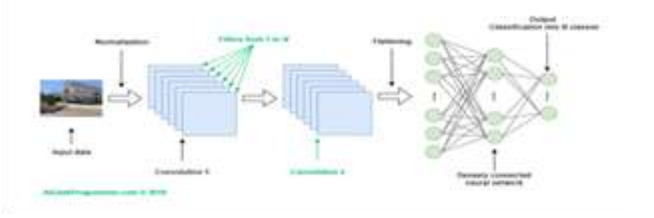
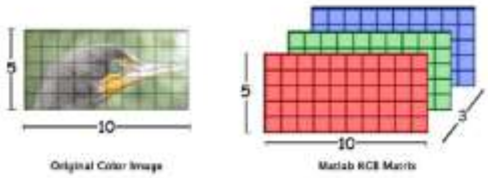
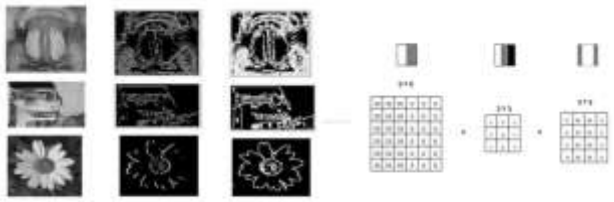


Image representation is process used to simplify the representation of an image into something that is more meaningful and easier to analyze. It refers to the way that brings information, such as color is coded digitally, and how the image is stored, i.e., how an image file is structured.

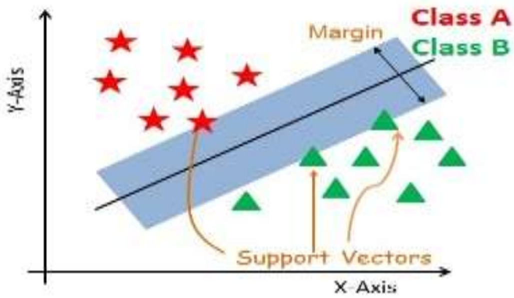


We have also used; Edge based Segmentation can also be done by using edge detection techniques. There are various techniques viz. gradient, log, canny, sobel, laplacian, robert. In this technique the boundary is identified to segment. Edges are detected to identify the discontinuities in the image. For classification they use both fixed and adaptive feature of support vector machine

Feature extraction: After segmentation the area of interest i.e. diseased part extracted. In the next step, significant features are extracted and those features can be used to determine the meaning of a given sample. Actually, image features usually includes color, shape and texture features. Currently most of the researchers targeting plant leaf texture as the most important feature in classifying plants. With the help of texture features, plant diseases are classified into different types.



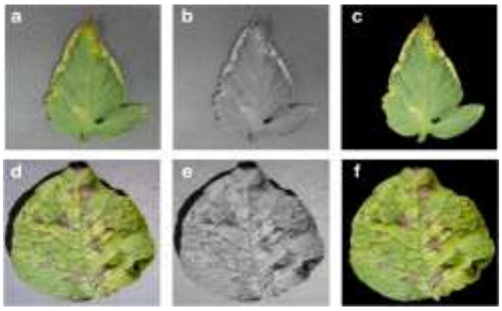
Classifier: Training and testing performed via several neural network classifier.
K-nearest neighbour :- K-nearest neighbor classifier is used to calculate the minimum distance between the given point and other points to determine the given point belongs to which class.
Artificial neural networks :- ANNs are popular machine learning algorithms that are in a wide use in recent years. Multilayer Perception (MLP) is the basic form of ANN that updates the weights through back propagation during the training. There are other variations in neural networks, which are recently, became popular in texture classification. Probabilistic Neural Network (PNN): It is derived from Radial Basis Function (RBF) network and it has parallel distributed processor that has a natural tendency for storing experiential knowledge. PNN is an implementation of a statistical algorithm called kernel discriminate analysis in which the operations are organized into a multilayered feed forward network having four layers viz. input layer, pattern layer, summation layer, and output layer [15]. Convolutional neural network: It is a neural network that has convolution input layers acts as a self learning feature extractor directly from input images. Hence, it can perform both feature extraction and classification under the same architecture. Back propagation network: A typical BP network consists of three parts: input layer, hidden layer and output layer. Three parts in turn connect through the collection weight value between nodes [11, 15]. The largest characteristic of BP network is that network weight value reach expectations through the sum of error squares between the network output and the sample output, and then it continuously adjusted network structure's weight value [8,12,13]. It is popular and extensively used for training feed forward networks. Also it has no inherent novelty detection, so it must be trained on known outcomes for training feed forward networks.
Support vector machine: Support vector machine (SVM) is a non-linear classifier, and is a newer trend in machine learning algorithm. SVM is popularly used in many pattern recognition problems including texture classification [14]. SVM is designed to work with only two classes. This is done by maximizing the margin from the hyper plane. The samples closest to the margin that were selected to determine the hyper plane is known as support vectors. Multi-class classification is applicable and basically built up by various two class SVMs to solve the problem, either by using one-versus-all or one.



DATASET DESCRIPTION

For experiments, we have used one datasets. The one that is available publicly as PlantVillage dataset. We analyze 54,306 images of plant leaves, which have a spread of 38 class labels assigned to them. Each class label is a crop-disease pair, and we make an attempt to predict the crop-disease pair given just the image of the plant leaf. Figure 1 shows one example each from every crop-disease pair from the PlantVillage dataset. In all the approaches described in this paper, we resize the images to 256 × 256 pixels, and we perform both the model optimization and predictions on these downscaled images.

Across all our experiments, we use three different versions of the whole PlantVillage dataset. We start with the PlantVillage dataset as it is, in color; then we experiment with a gray-scaled version of the PlantVillage dataset, and finally we run all the experiments on a version of the PlantVillage dataset where the leaves were segmented, hence removing all the extra background information which might have the potential to introduce some inherent bias in the dataset due to the regularized process of data collection in case of PlantVillage dataset. Segmentation was automated by the means of a script tuned to perform well on our particular dataset. We chose a technique based on a set of masks generated by analysis of the color, lightness and saturation components of different parts of the images in several color spaces (Lab and HSB). One of the steps of that processing also allowed us to easily fix color casts, which happened to be very strong in some of the subsets of the dataset, thus removing another potential bias.
This set of experiments was designed to understand if the neural network actually learns the “notion” of plant diseases, or if it is just learning the inherent biases in the dataset. Figure 2 shows the different versions of the same leaf for a randomly selected set of leaves.



IV. CONCLUSIONS

Across all our experiments, we use three different versions of the whole PlantVillage dataset. We start with the PlantVillage dataset as it is, in color; then we experiment with a gray-scaled version of the PlantVillage dataset, and finally we run all the experiments on a version of the PlantVillage dataset where the leaves were segmented, hence removing all the extra background information which might have the potential to introduce some inherent bias in the dataset due to the regularized process of data collection in case of PlantVillage dataset. Segmentation was automated by the means of a script tuned to perform well on our particular dataset. We chose a technique based on a set of masks generated by analysis of the color, lightness and saturation components of different parts of the images in several color spaces (Lab and HSB). One of the steps of that processing also allowed us to easily fix color casts, which happened to be very strong in some of the subsets of the dataset, thus removing another potential bias

ACKNOWLEGMENT

This study and research is carried out in Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India, under the guidance of C R Pote. I also thank our Head of the Department Dr. Nita Thakre and J P. Gabhane . The Principal of the institution, Dr. S A Dhale for providing us all the required facilities and suitable environment to successfully complete this research.

REFERENCES

[1] Peng Jiang , Yuehan Chen ,Bin Liu , Dongjian He , Chunquan Liang ,' Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks', (Volume: 7), pp. 06 May 2019

[2] Zhou, R., Kaneko, S., Tanaka, F., Kayamori, M., Shimizu, M., 'Disease detection of Cercospora Leaf Spot in sugar beet by robust template matching', Computers and Electronics in Agriculture, Volume 108, pp.58-70, 2014

[3] Barbedo, J.G.A., Godoy, C.V., 'Automatic Classification of Soybean Diseases Based on Digital Images of Leaf Symptoms', SBI AGRO,2015

[4] Barbedo, J.G.A., 'A review on the main challenges in automatic plant disease identification based on visible range images', 2016 ,Biosystems Engineering, Volume 144, pp. 52-60. [5] Bashish, D.A., Braik, M., Ahmad, S.B., 'A Framework for Detection and Classification of Plant Leaf and Stem Diseases', International Conference on Signal and Image Processing, pp. 113-118, 2010

[6] Punajari, J.D., Yakkundimath, R., Byadgi, A.S., 'Image Processing Based Detection of Fungal Diseases In Plants', International Conference on Information and Communication Technologies, Volume 46, pp. 1802-1808, 2015

[7] Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff Rohani Binti S Mohamed Farook, “Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques”, 2012 Third International Conference on Intelligent Systems Modelling and Simulation.

[8] Chunxia Zhang, Xiuqing Wang, Xudong Li, “Design of Monitoring and Control Plant Disease System Based on DSP&FPGA”, 2010 Second International Conference on Networks Security, Wireless Communications and TrustedComputing.

[9] Omrani, E., Khoshnevisan, B., Shamshirband, S., Saboohi, H., Anuar, N.B., Nasir, M.H.N., 'Potential of

radial basis functionbased support vector regression for apple disease detection', Journal of Measurement, pp. 233-252, 2014

[10] Gharge, S., Singh, P., 'Image Processing for Soybean Disease Classification and Severity Estimation', Emerging Research in Computing, Information, Communication and Applications, pp.