The Pharma Innovation Journal 2022; SP-11(10): 627-630

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2022; SP-11(10): 627-630

www.thepharmajournal.com
Received: 09-07-2022
Accepted: 11-08-2022

Kumar Sanjeev
Mandan Bharti Agriculture College, BAU, Sabour, Bihar, India

Suneeta Paswan
Krishi Vigyan Kendra, Agwanpur, Saharsa, Bihar, India

Corresponding Author:
Suneeta Paswan
Krishi Vigyan Kendra, Agwanpur, Saharsa, Bihar, India

Enhance ANN classifier performance using feature selection technique for detection of potato tuber diseases

Kumar Sanjeev and Suneeta Paswan

Abstract
Plant growth can be hampered by disease, which has negative impacts on crop output. The disease and pests caused a 10–20% annual reduction in crop productivity. Because they are a cheap food, potatoes give the human diet a source of cheap energy. Diseases in the potato crop significantly reduce yield. For image capture, image pre-processing, image segmentation, feature extraction, feature selection, and picture recognition, the image processing technique is utilised. These photos are used to extract the 76 colour, texture, and area attributes. Both prediction and classification are done using the Feed Forward Neural Network (FFNN) Model. For the selection of features, Relieff approaches were employed. Without feature selection, the model's accuracy is 84.76%. By selecting features using the Relieff approach, the model's accuracy is 85.23%. Therefore, improving classifier performance is helpful for disease detection that is accurate.

Keywords: Potato, disease, feature extraction, feature selection, classification

Introduction
India is an agricultural country where in about 70% of the population depends on agriculture. Farmers have wide range of diversity to select suitable Fruit and Vegetable crops. However, the cultivation of these crops for optimum yield and quality produce is highly technical. It can be improved by the aid of technological support. Plant disease diagnosis is an art as well as science. Farmers experience lot of problem in detecting and identify the diseases in various plants (Oerke, 2006) [5]. Digital image processing and image analysis technology based on the advances in microelectronics and computers has many applications in biology and it circumvents the problems that are associated with traditional photography. This new tool helps to improve the images from microscopic to telescopic range and also offers a scope for their analysis. It, therefore, has many applications in biology. So, there is a necessity in detecting diseases in such kind of plants leaves as well as fruits. (Barbedo, 2018) [1] Automatic identification and classification of diseases based on their particular symptoms are very useful to farmers and also agriculture scientists. Early detection of diseases is a major challenge in horticulture/agriculture science. (Bernardes, 2013) [2] Plant diseases are important factors because its affects human being as well as animals etc. that's why as it can cause significant reduction in both quality and quantity of crops in agriculture production. Therefore, detection and classification of diseases is an important and urgent task. (Konstatinos, 2018) [1] Feature selection (FS) for classification is a well-researched problem, aimed at minimizing the dimensionality and noise in data sets. Acceptable selection of features may improve accuracy and efficiency of classifier.

Review of Literature
(Revathi, 2014) [7] Proposed improved PSO feature selection technique that accepted skew divergence process and used different features like variances of texture, edge and color. (Zhang, 2007) [9] Introduced to classify and diagnose cotton disease using computer vision. They projected the fuzzy feature selection technique, fuzzy curves (FC) and surfaces (FS) to choose features of cotton leaf diseased. The fuzzy feature selection method was performed better for detecting and identifying of diseases. (Kanjalkar, 2014) [3] Presented in their paper some significant features of diseased leaves which will help to find exact disease of plant. (Revathi, 2013) [6] Have used three number of features for identifying leaf spot diseases of cotton crop. Color variance related to skew divergence feature was calculated by histogram of color and color descriptor. Sobel method was used to calculate the shape Skew divergence feature. Canny method used for estimating of the edge variance.

The skew divergence texture feature was calculated by Gober filter and texture descriptor. (Zhang, 2008) [8] Suggested a machine vision method-based system. This system was automatic for inspection of flue-cured tobacco leaves. Computer vision approaches were used to solve difficulties of features extraction. They included the features related to color, size, shape and surface texture features of tobacco leaves. The investigational results exhibited that this system was feasible for the features extraction. It can be used for the automatic classification of tobacco leaves.

Materials and Methods

Image Processing Techniques

The image processing is very useful for agriculture field. The image processing technique is consisting of different steps like Image Acquisition, Image Preprocessing, Image Segmentation, Feature Extraction, Feature Selection and Image Recognition.

Image Acquisition

Dataset of Potato Tuber Images

Table 1 and Figure 1 provide more details about the tuber dataset.

Table 1: The detail information of tuber dataset

Sl. No.	Name of Sample Class	No. of Sample
i.	Black Scurf	54
ii.	Common Scab	96
iii.	Green Colour	60
Total		210

I. Black Scurf II. Common Scab III. Green Colour

Fig 1: Different Class of Potato Tuber Images

Image Preprocessing

Preprocessing is the core step before starting any additional procedure. The key purpose of this stage is growing classification accuracy and reducing training time. In Image Preprocessing, generally, we have done to remove the noise of the sample, resize the images as par requirement and filtering the images and to do image enhancement for better viewing the processed image than the original image.

Image Segmentation

Image segmentation is used for partitioning of image into multiple region. That divided regions refers to a problem specific. This region of interest shows diseased region or your objective region. Image segmentation can be done various method like clustering methods, histogram-based methods etc.

Feature Extraction

The features are very useful in identifying one class of objects to another. In this present work we have extracted color, shape and texture features of images. The total number of extracted features are 76. These features are utilized for training of models. We considered as following features.

Colour Feature

Color image processing is divided into two major areas: full color and pseudo color processing.

Gray level: It refers to a scalar measure of intensity that ranges from black to grays and finally white. The RGB image was renovated to a gray image, and the following conversion formula are used:

Gray = R × 0.299 + G × 0.587 + B × 0.114
Where, R = Red, G = Green, B = Blue

The RGB mean, variance, and range are computed using the following expressions.

a. **Mean:** $\bar{x} = \frac{\sum_i^n x_i}{n}$

b. **Variance:** Variance is another measure of the spread of data in a data set. In fact, it is almost identical to the standard deviation. The formula is this:

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

c. **Range:** $R = X_{max} - X_{min}$

Texture Feature

The texture is the set of the pixel which has been frequent characterize. We considered statistical methods like Mean, Standard Deviation, Smoothness, Skewness, Kurtosis, Root Mean Square, Inverse Difference Moment, Energy, Contrast, Homogeneity and Variance for texture feature extraction of images.

Shape Features

It is a set of measurements that describe a certain shape according to some of its fundamental geometric properties. We considered geometrical methods like Aspect Ratio, Rectangularity, Area ratio of convex hull, Perimeter Ratio of Convex Hull, Sphericity, Circularity, Eccentricity and Form Factor for shape feature extraction of images.

Feature Selection

It enables to achieve maximum classification performance by reducing the number of features used in classification while maintaining acceptable classification accuracy.

Relieff

It finds the weights of predictors in the case where y is a multiclass categorical variable. The algorithm penalizes the predictors that give different values to neighbours of the same

class, and rewards predictors that give different values to neighbours of different classes.

Image Recognition

Classification is a process, in which one objects distinguished from another. In this research, we employed artificial neural network (ANN) classifier. In artificial neural network (ANN), the net input can be calculated as:

$Y_{in} = X_1.W_1+X_2.W_2+X_3.W_3 \ldots X_m.X_m$ i.e., Net input, $Y_{in} = \sum X_m.W_m$

Performance Analysis

Precision Rate (%) = (TP / TP + FP) *100 where, TP is True positive, FP is False positive.

Recall Rate (%) = (TP / TP + FN) *100 where, FN is False negative.

Accuracy (%) = (TP+TN / TP+TN+FN+FP) *100 where, TN is True Negative.

Result and Discussion

ANN trained model without feature selection on tuber dataset

The 76 features have extracted regarding colour, texture and shape of images. After that feed these extracted features as input into Feed Forward Neural Network (FFNN) Model. FFNN Model has generated input layer, one hidden layer and output layer. Details of model development have demonstrated in Figure 2.

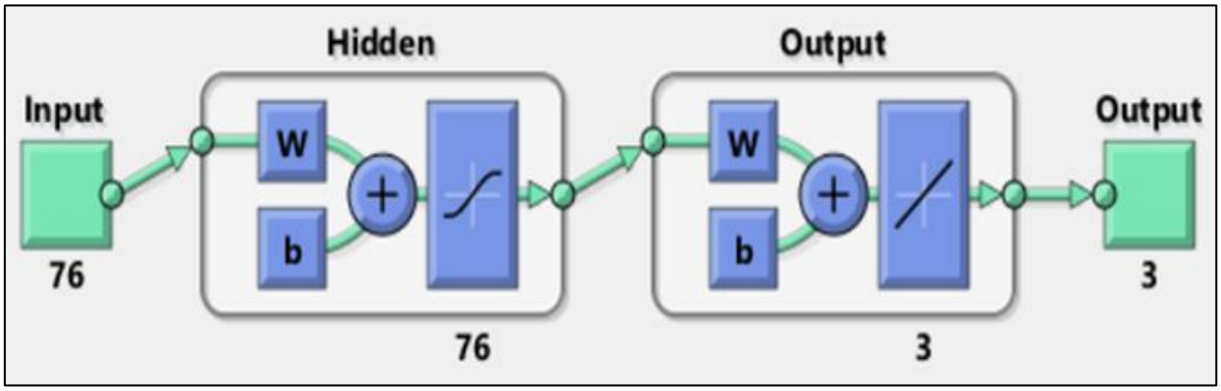

Fig 2: Development of FFNN Classifier without Feature Selection on Tuber Dataset

Table 2: Confusion Matrix Result Using ANN Classifier without Feature Selection on Tuber Dataset

Confusion Matrix		Target Class			Recall Rate (%)
		1	2	3	
Output Class	1	46	8	0	85.18
	2	11	82	3	85.41
	3	1	9	50	83.33
Precision Rate (%)		79.31	82.82	94.33	84.76

Table 2 shows 84.76% accuracy of ANN classifier. Precision rates of model are 79.31%, 82.82% and 94.33% of Class 1, Class 2 and Class 3 respectively. Class 3 provides better precision rate. Recall rates are 85.18%, 85.41% and 83.33% of Class 1, Class 2 and Class 3 respectively. Class 2 yields better recall rate.

The Performance of ANN on Tuber Dataset Training with Feature Selection Techniques

Table 3: The performance of ANN on Tuber Dataset Training with Feature Selection Techniques

Number of Features	Feature Selection Method
	Relieff Avg. Accuracy (%)
15	72.69
30	75.07
45	80.15
60	72.53
76	85.23

Table 3 shows that it estimates accuracy of ANN classifier is 85.23% by using Relieff feature selection method.

Table 4: Comparative Study of Without and With Feature Selection for ANN on Tuber Dataset Training

Classifier	Without Feature Selection	With Feature Selection Techniques
	Accuracy (%)	Relieff Method Accuracy (%)
ANN	84.76	85.23

Table 4 shows that feature selection method achieved better than without feature selection. Relieff feature selection method achieved maximum 85.23% accuracy.

Conclusion

The biggest hazard to food safety is crop disease. It's still challenging to diagnose diseases quickly and accurately. New technological developments can help to mitigate or perhaps resolve this issue. Both prediction and classification are done using the Feed Forward Neural Network (FFNN) Model. For the selection of features, Relieff approaches were employed. Without feature selection, the model's accuracy is 84.76%. The Relieff feature selection method increases model accuracy by 85.23%. Therefore, improving classifier performance is helpful for precise disease identification in potatoes.

References

1. Barbedo JGA. Factors influencing the use of deep learning for plant disease recognition. Bio systems engineering. 2018;172:84-91.
2. Bernardes AA, Rogeri JG, Marranghello N, Pereira AS, Araujo AF, Tavares RS. Identification of foliar diseases in cotton crop. Springer, Brazil; c2013.
3. Kanjalkar HP, Lokhande SS. Feature extraction of leaf diseases. International Journal of Advanced Research in

Computer Engineering & Technology. 2014;3:1.
4. Konstantinos PF. Deep learning models for plant disease detection and diagnosis. Computers & Electrical Engineering, 2018;145:311-318.
5. Oerke EC. Crop losses to pests. The Journal of Agricultural Science. 2006;144(1):31-43.
6. Revathi P, Hemalatha M. Cotton leaf spot diseases detection utilizing feature selection with skew divergence method. International Journal of Scientific Engineering and Technology. 2013;3:1.
7. Revathi P, Hemalatha M. Identification of cotton diseases based on cross information gain deep forward neural network classifier with PSO feature selection. International Journal of Engineering and Technology. 2014;5:6.
8. Zhang X, Zhang F. Images features extraction of tobacco leaves. CISP '08 Proceedings of the 2008 Congress on Image and Signal Processing. 2008;2:773-776.
9. Zhang YC, Mao HP, Xili BHM. Features selection of cotton disease leaves image based on fuzzy feature selection techniques. IEEE Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China; c2007. p. 2-4.