The Pharma Innovation Journal 2022; SP-11(10): 327-330

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2022; SP-11(10): 327-330

www.thepharmajournal.com
Received: 27-08-2022
Accepted: 30-09-2022

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

A comparative study of ANN and KNN classifiers performance for detection of potato tuber diseases

Kumar Sanjeev and Suneeta Paswan

Abstract

In India, potatoes are the vegetable crop that is most grown. According to estimates, diseases played a major role in the yield loss of potatoes. It is crucial to find diseases in their early stages. These problems have been resolved using image processing techniques. These images are used to extract the 76 attributes related to colour, texture, and area. Finally, to categorise the disease, these features are fed into the Feed Forward Neural Network (FFNN) and K- Nearest Neighbors (KNN) models. The accuracy rates for the ANN and KNN classifiers were 84.76% and 63.33%, respectively. A comparison of ANN and KNN classifiers has been started, and the results demonstrate that ANN classifiers outperformed KNN for the identification of potato tuber diseases.

Keywords: Potato, tuber disease, feature extraction, classifier

Introduction

A fundamental requirement for human survival is agriculture. Therefore, it becomes crucial to increase the production of fruits, vegetables, and cereals for developing nations like India. In addition, great manufacturing quality is necessary for both increased productivity and better health. However, a number of issues, such as disease feast, which may be avoided with early detection, worry about the quality and production of food. The majority of these illnesses are infectious, which may result in decreased agricultural yield (Singh, 2018) [7]. A vital factor in the effective cultivation of crops is the identification of plants, leaves, symptoms, and determining the presence of pests and diseases, their incidence rate, and the signs of their attack (Dickson, 1997) [3].

A crucial and important challenge is the identification and classification of illnesses. Growers typically use a naked eye observation method to detect illnesses. For the quick and precise diagnosis of plant illnesses and for the early detection and treatment of those diseases, some researchers have used image processing techniques (Pujari, 2013) [5]. The researchers used picture acquisition, image pre-processing, segmentation, feature extraction, and disease classification as the phases in image processing approaches for crop disease detection. The approach utilized for illness spot identification determines how accurate the results will be occurred (Bashir, 2012) [1].

Review of Literature

According to the approach suggested by (Phadikar, 2008) [4] an entropy-based thresholding is utilized for segmentation after converting the picture to the HSI colour space to detect and distinguish between two diseases impacting rice fields. After the image had been segmented, spots were found by measuring the intensity of the green components. (Wang, 2012) [8] designed a strategy to distinguish between related diseases in grapevines and wheat. A K-means algorithm is used to segment the photos, and 50 colour, shape, and texture attributes are then retrieved. (Revathi, 2014) [6] proposed an enhanced PSO feature selection technique that utilised many features, such as variations in texture, edge, and colour, and accepted the skew divergence process.

The SVM, Back propagation neural network, and the SVM were fed the mined features as input (BPN). Six different types of diseases, including Bacterial and Leaf Blight, Wilt by Fusarium, Root Rot, Micro Nutrient, and Wilt by Verticilium, were evaluated in tests to find the best categorization model. It was introduced to detect and identify cotton illness using computer vision (Zhang, 2007) [9]. Fuzzy curves (FC) and fuzzy surfaces (FS) were projected to be used in the fuzzy feature selection technique to pick out the sick features of cotton leaves.

For identifying and diagnosing diseases, the fuzzy feature selection strategy worked better.

KNN is one of the widely used classification tools, according to (Buana, 2012) [2], who discussed this in their study. It employed all training samples in the classification, which increased the complexity of the calculations. They introduced this method in their paper and solved the issue by merging the K-Means cluster algorithm and the conventional KNN algorithm. The dataset was prepared. The word (term) was then assigned weight by using Term Frequency-Inverse Document Frequency (TF-IDF). The simulation results showed that the suggested approach had improved its accuracy to 87%, had an average f-measure of 0.8029 with a best-fit k-value of 5, and required 55 seconds to compute one document.

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 Image: In present research, sample of potato tuber images have collected from crop fields. The detail information of dataset is given in Table1 and Figure1.

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**

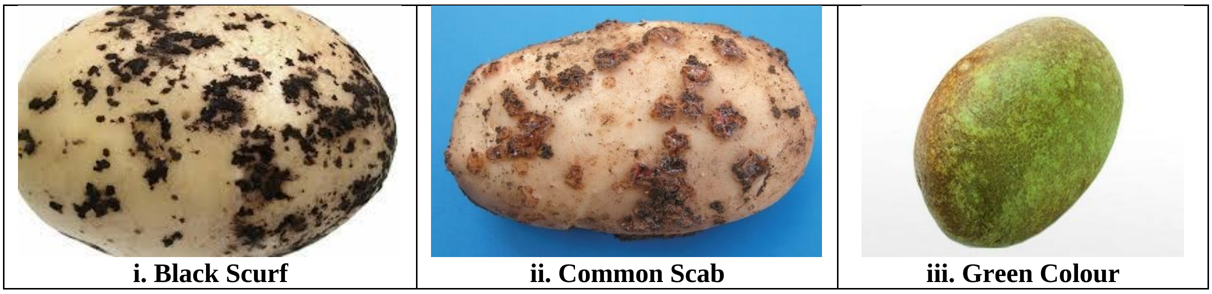

Fig 1: Different Class of Potato Tuber Images

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.

Image Recognition

- **Artificial Neural Network (ANN):** 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, $Y_{in} = \sum X_m.W_m$

- **K-Nearest Neighbours (KNN):** KNN technique has been useful in pattern recognition. KNN is built on non-parametric exercise which applied for classification and regression. It is called non-parametric supervise learning model (Sutton, 2012). In each case, the input is encompassed of the K nearest of the dataset samples within the features zone. K-NN has worked for classification or regression on the basis of output dependent.

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 on Leaf Dataset

The 76 features are extracted from 210 images of potato tuber. These 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 been evaluating and manipulating neurons and its weights. It has generated input layer, one hidden layer and output layer. Details of model development have demonstrated in Figure 2.

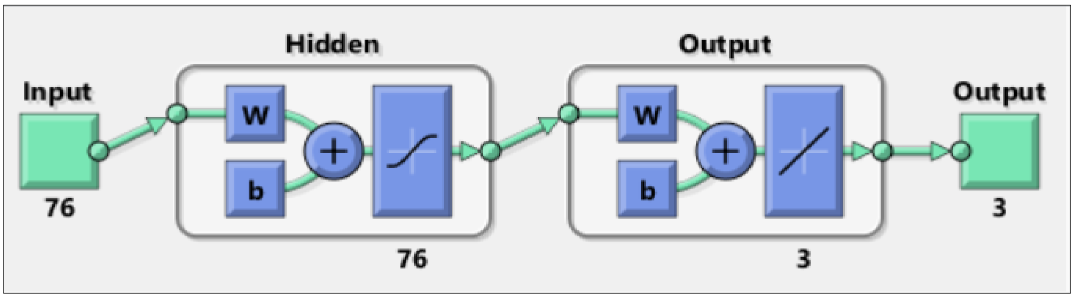

Fig 2: Development of FFNN Classifier on Tuber Dataset

Table 2: Confusion Matrix Result Using ANN on Tuber Dataset

Confusion Matrix		**Target Class 1**	**2**	**3**	**Recall Rate (%)**
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

Here, confusion matrix has been evaluated the performance between target and output class. Table 2 shows 84.76 % accuracy of ANN classifier. Precision rates are 79.31%, 82.82% and 94.33% of Class1, Class2 and Class3 respectively. Class3 provides better precision rate. Recall rates are 85.18%, 85.41% and 83.33% of Class1, Class2 and Class3 respectively. Class2 shows better recall rate.

KNN Trained Model on Tuber Dataset

The 76 features were extracted regarding colour, texture and shape of images. These 76 extracted features have given as input into K-Nearest Neighbour (KNN) Model. K value has been taken as 5. KNN Model has been evaluating and developing a model. Details of the model development have been shown in Figure 3.

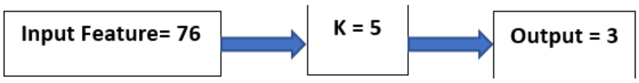

Fig 3: Development of KNN Classifier on Tuber Dataset

Table 3: Confusion Matrix Result Using KNN on Tuber Dataset

Confusion Matrix		**Target Class 1**	**2**	**3**	**Recall Rate (%)**
Output Class	1	29	22	23	39.18
	2	19	66	11	68.75
	3	9	13	38	63.33
Precision Rate (%)		50.87	65.34	57.77	63.33

Here, confusion matrix has been evaluated the performance between target and output class. Table 3 shows the accuracy 63.33 % of KNN classifier. Precision rate are 50.87%, 65.34% and 57.77% of Class1, Class2 and Class3 respectively. Class2 produces better precision rate. Recall rate are 39.18%, 68.75% and 63.33% of Class1, Class2 and Class3 respectively. Class2 performed better recall rate.

Table 4: A Comparative Study of the Performance of ANN and KNN on Tuber Dataset Training

Classifier	**Accuracy (%)**
ANN	84.76
KNN	63.33

Table 4 shows comparative study of the performance of ANN and KNN have initiated on tuber dataset training. The result of table 4 showed that the maximum accuracy was found with the ANN classifier which shows that ANN classifier performed better than KNN for detection of potato leaf diseases. It is also shown below in Figure 5.

Comparison of Accuracy

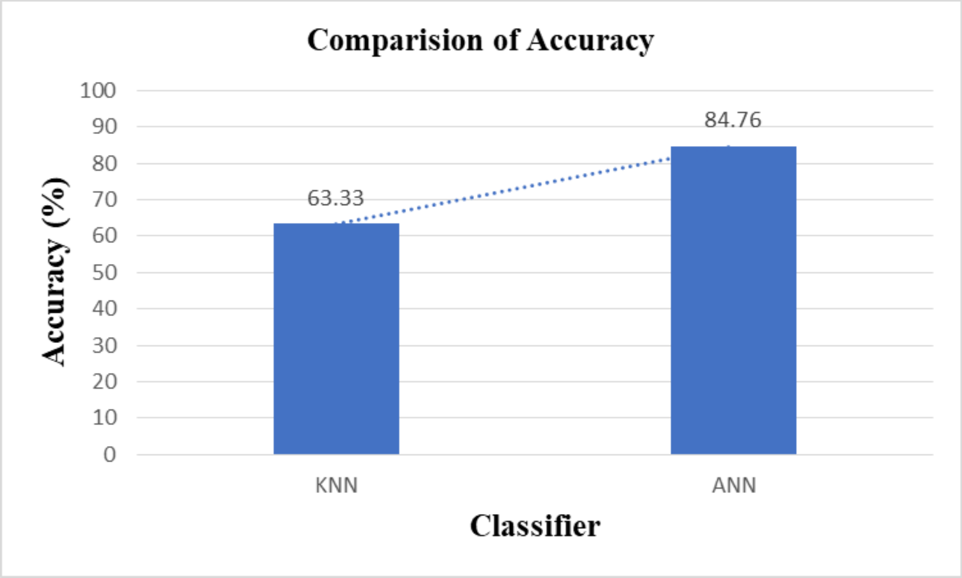

Fig 5: Comparison of Accuracy

Conclusion

A crucial and important challenge is the identification and classification of illnesses. These problems have been resolved using image processing techniques. These images are used to extract the 76 colour, texture, and area attributes. Finally, to categorise the disease, these features are fed into the Feed Forward Neural Network (FFNN) and K- Nearest Neighbors (KNN) models. The accuracy rates for the ANN and KNN classifiers were 84.76% and 63.33%, respectively. A comparison between ANN and KNN classifiers has been started, and the results suggest that ANN classifiers outperformed KNN for detecting potato tuber diseases.

References

1. Bashir S, Sharma N. Remote area plant disease detection using image processing. Amity School of Engineering and Technology, Uttar Pradesh, India; c2012.
2. Buana PW, Sesaltina JDRM, Putra I. Combination of k-nearest neighbor and k-means based on term re-weighting for classify Indonesian news. International Journal of Computer Applications. 2012;50(11):0975-8887.
3. Dickson MA, Bausch WC. Plant recognition using a neural network classifier with size and shape descriptors. Transactions of the ASAE. 1997;1:97-102.
4. Phadikar S, Sil J. Rice disease identification using pattern recognition techniques. IEEE, Khulna; c2008. p. 420-423.
5. Pujari JD, Yakkundimath R, Byadgi AS. Classification of fungal disease symptoms affected on cereals using color texture features. International Journal of Signal Processing, Image Processing and Pattern Recognition. 2013;6(6):321-330.
6. 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):4637-4642.
7. Singh KK. An artificial intelligence and cloud based collaborative platform for plant disease identification, tracking and forecasting for farmers. In: Proceedings of international conference on cloud computing in emerging markets (CCEM); c2018. p. 49-56.
8. Wang H, Li G, Ma Z, Li X. Application of neural networks to image recognition of plant diseases. Proceedings of the 2012 International Conference on Systems and Informatics (ICSAI). IEEE, Yantai; c2012. p. 2159-2164.
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. 2007. p. 2-4.