jet INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2024, VOL. 70, NO. 2, PP. 455-464
Manuscript received November 7, 2023; revised June, 2024. doi: 10.24425/ijet.2024.149566

CNN ensemble approach for early detection of sugarcane diseases – a comparison

Kavitha K J, and Krishna Prasad K

***Abstract*—This paper mainly concentrates and discusses on sugarcane crop, the variety of cane seeds available for sowing; various cane diseases and its early detection using different approaches. Machine Learning (ML) and Deep Learning (DL) techniques are used to analyze agricultural data like temperature, soil quality, yield prediction, selling price forecasts, etc. and avoid crop damage from a variety of sources, including diseases. In the proposed work, with particular reference to eight specific sugarcane crop diseases and including healthy crop database, the neural network algorithms are tested and verified in terms quality metrics like accuracy, F1 score, recall and precision.**

***Keywords*—Sugarcane diseases; CNN; KNN; Soil quality; quality metrics**

I. Introduction

THE population across the world is escalating day by day demanding more groceries in addition to the natural resources like territory and water. Under this constraint and bearing in view that, there is a need of more land and water for agriculture to produce enough food grains for feeding 7-8 billion worlds' population. Also, the natural resources' availability is depleting at faster rate, and there of it is a very challenging task for the farmer's community to fulfill the needs of upcoming population. As, India is a country of villages, agriculture plays an important role and main occupation of farmers in villages.

India is known to be the world's largest consumer and the second largest producer of sugar. Sugarcane is one of the important cash led industrial crop in India. The sugarcane crop has considerably contributed to the growth of Indian cultivation and National Gross contributed Domestic Products - NGDP. The rural economy in conventional sugarcane growing region is mainly associated with sugarcane crop and sugar or its allied industries. Since from early 20th century, a lot of emphasis has been laid in research and Development (R&D) field associated with sugarcane cash crop, which in fact enabled the sugar industries to progress through years.

Sugarcane farming has been practiced in India since the Vedic era i.e. Sugarcane farming is first mentioned in texts from India between 1400 and 1000 B.C and now it is generally acknowledged that the Saccharum species is originated in India and it has been identified that cultivated canes in India may be categorized as two main groups: Thin and hardy north-Indian types as Saccharum-barberi and Saccharum-Sinense and Thick and juicy noble high price cane Saccharum-officinarum Sugarcane-Saccharum is the major source of sugar in India.

Kavitha K J is with GM Institute of Technology, VTU (email: kavithakj219@gmail.com).
Krishna Prasad K is with Srinivas University, Mangalore (e-mail: karanikrishna@gmail.com).

The crop also occupies an important position in the Indian-Agricultural scenario as this crop has a wider adoption in different agro-climatic conditions across various regions in the country. As it is mentioned earlier, the crop plays a significant role towards the national economy and provides raw/unrefined material to extract sugar and more than twenty-five other major industries like alcohol producing industry, paper industry, chemical industry, cattle feed, pharmaceutical industry, and to extract ethanol for bio-fuel production is exclusively based on sugarcane production as a raw material. In addition to the sugar factories and many other industries which are rely on its by-products, the crop also supports rural and cottage industry of jaggery-Gur and brown sugar-khandsari which together produce about 7-10 million tones of sweeteners. Due to multi-purpose uses of sugarcane crop in different industries, the demand for the crop is also increasing and the crop is also finding its sustainability in the country.

The study presented in this paper was motivated by the rapid growth of a variety of sugarcane disease classes and the incompetence of farmers in identifying and diagnosing diseases. Computer vision and deep learning are used to address this problem.

II. Literature Review

The sugarcane cultivation can be done under two divergent agro-climatic regions – tropical and sub-tropical regions. In our country, the tropical region is of about 45% area which in turn contributes 55% of the total sugarcane production, whereas, the sub-tropical region is of 55% area which contributes 45% of total production of sugarcane in the country and thereby the average sugarcane yields in the country is about 69.4 t/ha.

The sugarcane crop can grow well under warm humid conditions, i.e. it requires an average temperature of 28 to 32°C and a higher temperature above 45°C results in the reduction of tillering and thereby stops its growth, whereas the low temperature such as below 20°C results in slow growth of the crop. And, also the areas with a very low temperature less than 5°C are not appropriate for sugarcane cultivation. The sugarcane crop grows well in the region which usually receives an annual rainfall of 75 to 120 cm.

The cane crop is also well known as a lazy crop as it requires a long growing season of nearly 10 to 18 months based on the cane seeds and a maximum of 12 month duration is common in most of the regions. Also a relative humidity which is usually less than 50% during growing season is not suitable for sugarcane cultivation but a relative humidity of 70 to 85% during growth and 55 tob75% during ripening phase is ideal for getting high yield.

The tropical regions in India includes the states Maharashtra, Karnataka, Tamil Nadu, Andhra Pradesh, Telangana, Gujarat, Madhya Pradesh and Kerala whereas the sub-tropical region in India includes the states Bihar, Uttarakhand, Haryana, Punjab, West Bengal, Assam, Chhattisgarh, Jharkhand, Odisha and Rajasthan. The planting seasons of sugarcane crops varies in India i.e. in subtropical region of India, planting seasons are autumn-October, spring-February to March),and summer-April to May whereas in Peninsular India, cane cultivation is done in the month between January to February.

A. Cane Seeds

- Based on the tropical and sub-tropical regions, there are varieties of cane seeds provided by the various sugarcane industries. Some of the present popular cane seeds are CoLk 12207 (Ikshu-6), CoLk 12209 (Ikshu-7), CoLk 11203 (Ikshu-5), CoLk 09204 (Ikshu-3), CoLk 9709, CoLk 94184 (Birendra), CoLk8001, CoLk8102, Co06030, Bo99, CoP9301, CoSe98231,Co86002, Co86032, CoSi 95071,CoJ 64, CoS 8436,Co 94012, CoC 671, Co92020, Co8014, Co86032, Co62175, CoC671, Co86032, Co8011, Co94012, Co62175, CoA89085, CoJ85, CoJ88, Co94012, Co94010, CoS8436, Coj64, CoS88230, CoS98231,CoS 8436, CoS88230, Cos767 and etc.

The cane seed varieties may be categorized as Early varieties, Mid-late varieties, and Late varieties based on the amount of sucrose contribution and the period during which the sucrose content or the plant become matured. The cane plants which contribute almost 16.5 % of sugar and maturity of the crop in 10 months are categorized as early sugarcane varieties; the cane plants which contribute nearly 16 % of sugar and maturity of the crop in 12 months are categorized under mid-late sugarcane varieties; whereas the cane plants similar to that of mid-late varieties contributing a sugar level of 16 % but the plant maturity at more than 12 months and maintaining the same sugar level up to 14 months are categorized under late sugarcane varieties. The next section discusses the type of soil suitable for growing sugarcane crop and the necessary land preparation to be made before planting the cane seeds.

B. Field Preparation

From sandy loam to clay loam, sugarcane may be grown in every type of soil. But it does best on well-drained soils. It can also be successfully grown on lighter soils with suitable irrigation systems in place, as well as on hard clays with proper drainage and organic matter addition. Acidic, alkaline, and saline soils are completely inappropriate for sugarcane. The physical and chemical properties of the soils should be considered for sugarcane cultivation. Sugarcane is considered as a good managerial crop that would result in maximum biomass with maximum expose to the sunlight under a certain set of management strategies. The productivity and fertility of soils vary widely. It is possible to increase the production levels of a particular soil by using optimal soil, fertiliser, and crop management strategies. This is often accomplished by enhancing the soil's physical and biological properties through correct drainage, aeration, soil particle aggregation, addition of sufficient quantities of bulky organic matter, and application of fertilizers in a balanced manner based on the results of soil tests. The soil and plant relationship is centered mostly on variations in soil moisture, which in turn depends on the irrigation techniques used and the type of soil where the crop is allowed to grow.

The amount of organic matter in the soil, the level of microbial activity, changes in soil pH, and the condition of soil moisture all have an impact on where the soil supplies nutrients. It is emphasized that efforts should not only focus on increasing soil fertility but also on increasing soil productivity by halting the emergence of problems including salinity, alkalinity, compaction, weed infestation, and a lack of micronutrients.

In the cultivation of cane, the pH of the soil plays a crucial role. Electrical conductivity (EC), which is used to measure salt concentration in soil, should not be greater than 4 m.mhos/cm. Despite there being enough water in the soil, the plant cannot obtain or absorb enough soil moisture when there is such a high concentration of soil salts. Additionally, if the soil has a higher sodium concentration, the plant's capacity to access potassium and magnesium is severely hampered, which results in lower yields. Alkaline soils and water used to irrigate sugarcane produce juices with low sucrose levels that are difficult to process for jaggery production or sugar production. In acid soil having pH of 5.6, inter-nodal elongation and tillering were reduced.

In salty soils, young shoots may not grow normally, and the leaves typically have a yellowish hue. In severe situations, the canes' growth is impeded and they become thin. The leaves also turn white with black areas of dead tissue. The clumps' abundance of short, dead roots makes them simple to pick out. High soil salinity and compaction cause sick roots, chlorotic and stunted shoot growth, and subpar cane quality. In general, a well-structured, appropriately aerated soil profile with no compaction, hard pans, lime bands, or salt zones in the subsoil is needed for sugarcane crops to grow. This depth should be between 40 and 60 cm.

A field preparation in the sugarcane cultivation involves the following necessary steps such as; Supervision of earlier crop residues, Tillage (ploughing/ harrowing/sub-soiling etc.), Leveling, Inclusion of organic/natural manures and Field layout. To enhance and maintain soil fertility and productivity, organic or natural manure input during soil preparation is crucial. Organic matter aids in enhancing microbial activity, water holding capacity, and soil structure. Additionally, it aids in the release of various plant nutrients, including micronutrients and phosphorus. Farmyard manure (FYM), the conventionally applied manure, is quickly becoming scarce as the number of animals decreases.

Press-mud and sugarcane waste from sugar factories can be utilized in sugarcane fields as a supplement to FYM, either immediately or after composting. Waste from sugarcane is a great source of organic material. It could be utilized and composted. Growing a green manure crop before sugarcane and in situ inclusion, such as daincha (Sesbania aculeata) or sun-hemp (Crotolaria juncea), is a highly helpful method to increase soil fertility and production.

This is crucial in areas where organic manure application is insufficient or non-existent. An addition of 7.5 to 25 tonnes of green matter and 10 to 30 kg of nitrogen per ha are provided by a leguminous green manure. Sun-hemp has 0.75% N, 0.12% P2O5, and 0.51% K2O, while daincha has about 0.62% N. 50 kg of seeds can be distributed in the prepared land to establish a green manure crop.

The ideal time for integration is during the flowering stage because the crop is at its most delicate and nutrient-rich. Before planting sugarcane, there should be enough time after the green manure has been introduced to allow for proper decomposition. The next section discusses the various sugarcane diseases and pest attack that would happen during the growth of the crop.

C. Various diseases and Pest attack in Sugarcane Plant

An annual crop, sugarcane is susceptible to numerous diseases and pests. They may harm the crop and crop products in terms of quality and economics. It is crucial to control insect pests and illnesses that affect the sugarcane crop, and choosing an efficient management approach is just as crucial. These methods could include:

➢ Applying chemicals to soils, seeds, plants, or other materials; examples include fungicides, nematicides, insecticides, herbicides, and bio-regulators.

➢ The employment of beneficial insects, pathogens (viruses, bacteria, and fungi), antagonists, resistant cultivars, induced resistance, organic fertilizers, etc. are examples of biological control techniques.

➢ Biotechnological methods, such as the employment of pheromones, insect hormones for growth and reproduction, sterile male techniques, and physical and/or chemical attractants.

➢ Agronomic methods, such as proper seeding and planting, location and improvement, culture, crop rotation, removal of inoculum sources, or alternative and intermediate hosts Physical techniques, including mechanical, thermal, exclusions (such as nets), and radiation.

During the survey, several sugarcane diseases caused by fungi, bacteria, and viruses were discovered in diverse locations. When compared to bacterial diseases, viruses and fungi-related illnesses were shown to be more common. Some of the most common sugarcane diseases are listed below: Red rot, Wilt, Grassy shoot, Smut, Leaf scald, Red striped disease, Mosaic disease, Pokkahboeng, Rust, Sugarcane yellow leaf disease and Crop Stage-wise IPM practices for Sugarcane.

Numerous diseases and insect pests have the potential to damage sugarcane. An estimate states that illnesses and insect pests cause productivity of sugarcane to fall by 20, and 19, respectively. The control of insects, illnesses, and other pests is crucial for increasing agricultural output. Insect pests and disease relevance vary according to variation in agro-ecological settings, thus a management plan should be established in accordance with this. About 288 insects, almost two dozen of which significantly reduce crop quality and yield, are present in sugarcane.

The sub-tropical and tropical belts of sugarcane have different levels of disease and insect infestation. While internodes borer and early shoot borer, as well as the diseases rust and eye spot, are more common in tropical regions, top borer and stalk borer are more common in sub-tropical places.

D. Time of harvest:

Harvesting should be done as far away from weather extremes as feasible. It has been demonstrated that in subtropical India, spring harvesting of plant crops will provide greater yields than fall harvesting. After reaching maturity, the sugarcane crop is harvested; typically, in subtropical states, it begins in the month of October and lasts until the month of May, while in tropical states, it begins in the month of December and lasts until the month of May.

E. Lacunas in Sugarcane farming

Based on the survey carried out in the local regions, we have identified some of the real time problems faced by the local farmers in growing and harvesting sugarcane crop. As we know that Karnataka is the third-largest sugarcane-producing state in India; however, the farmers in the state continue to follow traditional methods of farming such as:

❖ Lack of testing of seed canes and soil test.

❖ Lack of knowledge disease spread and pest attack time.

❖ Sugarcane harvesting time is based on the length of time of planting of cane seeds.

This type of procedure followed by the farmers usually leads in low production of the crop. Also, the farmers take suitable action on the crop by estimating the period of each stage of crop growth like:

➢ Spray pesticides to the crop based on the time of cane seeds are planted.

➢ Decides the time of harvesting based on the type of planted cane seeds.

In general, diseases and soil fertility typically have an impact on agricultural productivity. Therefore, it is essential to employ cutting-edge approaches to boost agricultural product output and, in turn, farmers' financial revenue. The majority of diseases associated to plants are caused by bacteria, viruses, fungi, and pathogens.

To overcome the lacunas mentioned above, this paper proposes the use of digital image processing techniques in the farming methods to predict the early detection of diseases and to decide the harvesting time of sugarcane crop based on sugar content in the crop so as to boost crop production. The proposed block diagram for disease prediction and crop harvesting is as shown below in figure 1:

The overall procedure used in the process of early detection of sugarcane plants are image capturing, image pre-processing, image segmentation, feature extraction [19] and detection of diseases based on the extracted features. In the proposed work eight types of sugarcane diseases i.e. red rot, red rust, red striped disease, wilt, smut, mosaic, grassy shoot and sugarcane yellow leaf diseases along with healthy leaf are identified. The detailed explanation of the proposed work is explained in the next section.

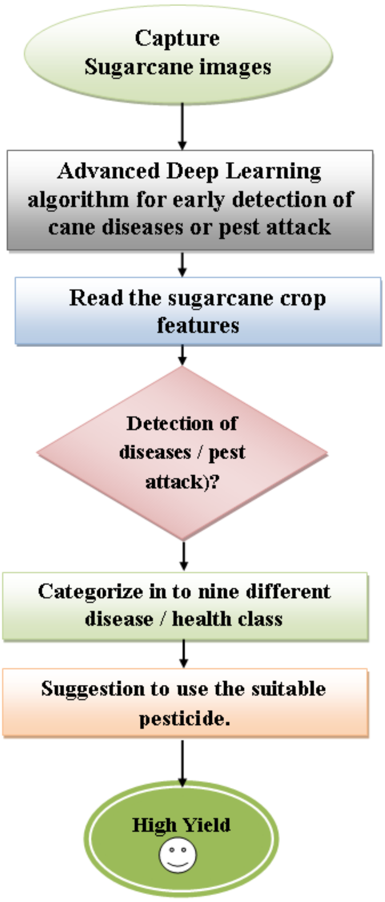


Fig. 1. Proposed Block Diagram for Early prediction of crops and harvesting period

III. Methodology

To detect the sugarcane diseases in the proposed work, mainly colour and texture part of the plant is used in the feature extraction procedure using neural network technologies [17 &18]. The overall procedure used in the proposed work is explained using pseudo code as below:

- ➢ RGB color image acquisition and store in a database.
- ➢ Colour transformation structure creation.
- ➢ Convert the colour values in RGB to the space specified in the colour transformation structure.
- ➢ Apply pre trained clustering technique to categorize.
- ➢ Masking green-pixels.
- ➢ Remove the masked cells inside the boundaries of the infected clusters.
- ➢ Convert the infected (cluster / clusters) from RGB to HSI Translation.
- ➢ SGDM Matrix Generation for H and S. (Another name for gray-level co occurrence.
- ➢ Matrix is gray-level spatial dependence matrix.
- ➢ Calling the GLCM function to calculate the features.
- ➢ Texture Statistics Computation.
- ➢ Configuring Neural Networks for Recognition

The pseudo code for the above steps is given below:

$$I = read(Sugarcane_{plant} - image)$$
$$I = resize(I,(128,128,3))$$
$$I = \text{savitzky}_{\text{golay}}(\text{I}); \text{Noise removal}$$
$$I = Predefined\_cnn(I)$$
$$I = hsi(I)$$
$$I = eucli\_dist(I)$$
$$I_{seg} = cnn(I)$$
$$I_{feature} = disease_{detection(I)}$$

As explained in the previous section, the first step in the procedure is sugarcane image capturing. In the present work, the open image dataset from the Kaggle open source, nearby local regions of Karnataka and other sources of internet has been used for the analysis purpose and the details of sugarcane dataset is described below: In the proposed work, a total of 3256 sugarcane images categorized into 8 different classes as red rot, smut, wilt, yellow leaf, grassy shoot, red striped, mosaic and rust sugarcane diseases plus 1 more class to represent healthy sugarcane leaf. Infected patches of various patterns may be seen on most of the image dataset. Using various patches in accordance, each of these places was individually marked. Each of these spots was carefully annotated using the appropriate patches. The neural simple model is as shown below in figure 2:

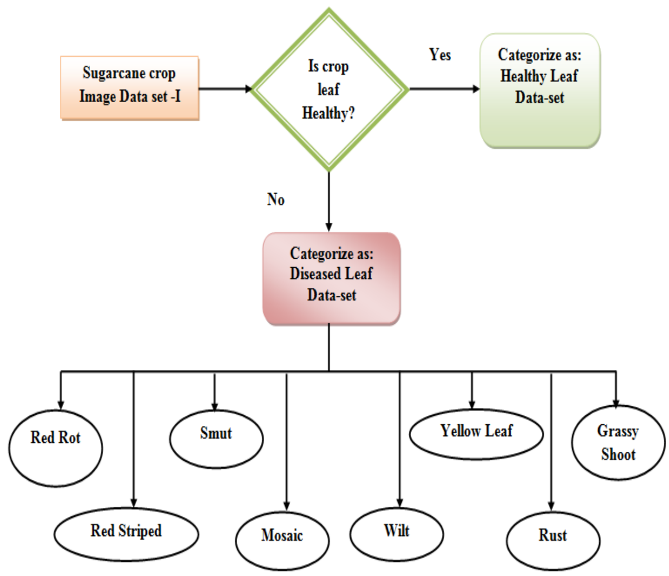


Fig. 2. Neural Model for Disease Categorization

The above model may also be expressed in terms of set theory as defined by the equation 1:

$$S = \{H, D, X\} \quad (1)$$

Where,

$S = A\ total\ data\ set\ of\ sugarcane\ crop\ leaves\ comprising\ subsets\ healthy\ leaf\ H, diseased\ leaf\ D\ and\ incorrect\ data.$

$H = A\ subset\ of\ Healthy\ cane\ leaves.$

$D = A\ set\ of\ subset\ of\ class\ of\ diseases = \{D1, D2, \ldots \ldots \ldots D8\}$

$X = Incorrect\ data$

A pictorial representation of data set designed and used in the proposed work may be shown as in figure 3:

The overall working principle may be explained using the block diagram shown in the below figure 4:

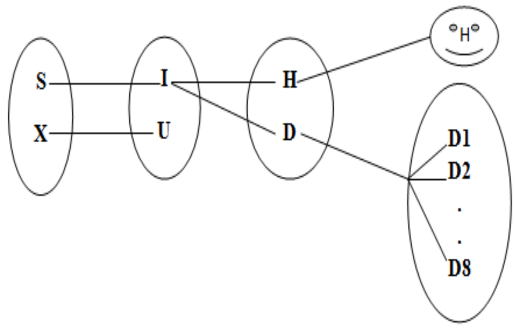


Fig. 3. Pictorial representation of data set

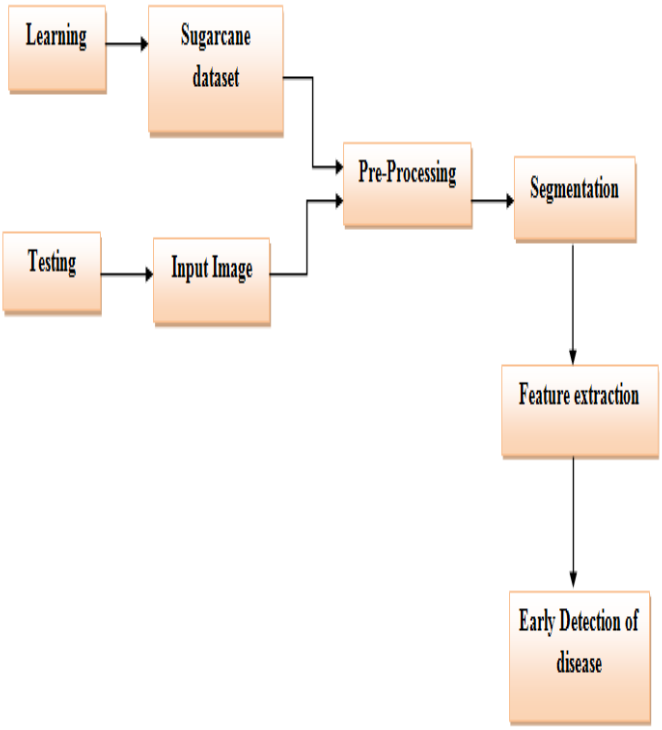


Fig. 4. Proposed Block diagram

The images which are in the form of RGB colour structure are transformed to device independent colour-Space transformation and this is achieved by the following colour transformation equations defined in [1& 2].

The captured sugarcane- dataset then subjected to advanced image processing techniques for the early detection of various sugarcane diseases which includes the various steps as explained below:

Pre-Processng:

In the pre-processing stage as shown in figure 5, the sugarcane image dataset undergoes various processing techniques to enhance the information in the images which would help in analyzing various factors.

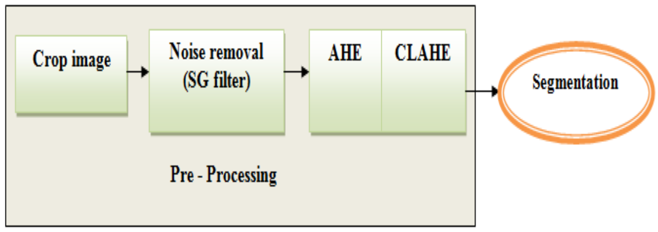


Fig. 5. Pre-Processing technique in the proposed system

In the work presented, digital smoothing polynomial filters or least-squares smoothing filter which is well known as Savitzky-Golay (SG) filter [3] is used for noise removal purpose as it helps in edge preservation by retaining high frequency components of the crop and is implemented using the equation (2):

$$p_i = \sum_{k=pl}^{pr} w_k\, f_i + kw \qquad (2)$$

Where,

$w_k \dashrightarrow$ *pixel weight contribution*

$f_i \dashrightarrow$ *ith pixel to be smoothened*

The smoothening step is followed by the image enhancement technique where the contrast of the image is enhanced using adaptive histogram equalization (AHE) followed by contrast limited AHE (CLAHE) where it spreads the image intensities using cumulative distribution function and applied with gamma correction rule [4 & 5] and also the contrast crop image helps in identifying crop suffering from particular disease or not.

The histogram of a leaf may be calculated using the equations 3 & 4:

$$h(i) = \frac{cdf(I) - cdf_{min}(i)}{leaf\_size(i)} \qquad (3)$$

$$h_{cl}(i) = bilinear\_interp\{h(tile(i))\} \qquad (4)$$

Where,

$h(\,) \rightarrow$ *histogram equalized value of image*

$cdf \rightarrow$ *cumulative distribution function of image 'i'*

$tile \rightarrow$ *sections of image 'i'*

Segmentation:

The enhanced image is the subjected to segmentation technique [2] to select the region of interest to predict the diseased area of the plant and is implemented using CNN algorithm [6], and each image of a sugarcane sample is segmented into individual segments. The image is then changed from RGB to HSI format after the infected items have been located.

First, the RGB images of leaves are converted into Hue Saturation Intensity (HSI) colour space representation as shown in figure 4. The purpose of the colour space is to facilitate the specification of colors in some standard, generally accepted way. HSI (hue, saturation, intensity) colour model is a popular colour model because it is based on human perception [10, 11 & 12].

Hue is a colour attribute that refers to the dominant colour as perceived by an observer. Saturation refers to the relative purity or the amount of white light added to hue and intensity refers to the amplitude of the light. Colour spaces can be converted from one space to another easily. After the transformation process, the H component is taken into account for further analysis.

S and I are dropped since it does not give extra information. Equation 5 shows the H, S and I components. The R, G, B values are divided by 255 to change the range from 0.

$$\left.\begin{array}{l} R1 = R/255 \\ G1 = G/255 \\ B1 = B/255 \\ Cmax = max(R1, G1, B1) \\ Cmin = min(R1, G1, B1) \\ \Delta = Cmax - Cmin \\ Hue: \\ H = \begin{cases} 60^{\circ} \times \left[\left(\frac{G1-B1}{\Delta}\right) mod\ 6\right] & with\ Cmax = R1 \\ 60^{\circ} \times \left[\left(\frac{B1-R1}{\Delta}\right) + 2\right] & with\ Cmax = G1 \\ 60^{\circ} \times \left[\left(\frac{R1-G1}{\Delta}\right) + 4\right] & with\ Cmax = B1 \end{cases} \\ Saturation: \\ S = \begin{cases} 0 & ; Cmax = 0 \\ \frac{\Delta}{Cmax} & ; Cmax \neq 0 \end{cases} \\ \text{Intensity:} \\ I = \dfrac{R1 + G1 + B1}{3} \end{array}\right\} \quad (5)$$

Only H and S images are then used to build the Spatial Gray-level Dependence Matrices - SGDM matrices for each pixel map of the image. The SGDM calculates the likelihood that a given pixel with a specific grey level will be located a specific distance and orientation angle [1] away from another pixel with a different grey level which is calculated using the equations 6 & 7 as:

$$G_s(p,q) = \sum_{x=1}^{m} \sum_{y=1}^{n} \begin{cases} 1 & ; I(x,y) = p\ and\ I(x+\Delta x, y+\Delta y) = j \\ 0 & ; otherwise \end{cases} \quad (6)$$

$$I_{Cont} = \sum_p \sum_q (p-q)^2 G_s(p,q) \quad (7)$$

Where,

$G_s(p,q) \rightarrow SGDM\ matrix$

$I_{Cont} \rightarrow Enhanced\ contrast\ leaf\ image$

Each image's texture statistics were created from the SGDM matrices. We suggest developing software based on neural networks for automatically identifying and categorizing leaf diseases.

Feature Extraction:

After the segmentation process, various features of the sugarcane plant such as Color, texture, morphology, edges etc. [7] is extracted which can be used in the early detection of sugarcane crop diseases. The three most popular techniques used for feature extraction are texture-based, color-based, and shape-based. Feature extraction is a crucial step to extract the characteristics of an image's interesting region [8 & 13]. Hardness, roughness, and colour distribution [9] over the whole image are all indicators of the image's texture. The texture features such as skewness, correlation, entropy, smoothness, homogeneity etc required for the detection of plant diseases are extracted using Gray Level Co-occurrence matrix (GLCM). The architecture used in the feature extraction is as shown in figure 6.

The training data is used to build a number of folds, usually five or 10. After training each base model with training data from each fold, predictions are produced using the validation set. Every base model's prediction on the validation set serves as an input feature for the meta-model. The ensemble model is trained on how to integrate the predictions from the validation set as efficiently as possible to get a final prediction. The test data may be used to make predictions once the ensemble model has been trained. Stacking has shown to be an effective method for improving prediction performance in several machine learning applications. By integrating the predictions of several models, stacking can assist to reduce overfitting, improve generalisation, and raise prediction accuracy. It is essential to extract stage-specific spatial characteristics to prevent the loss of critical spatial information in order to anticipate leaf diseases that have a great deal of morphological similarities [23].

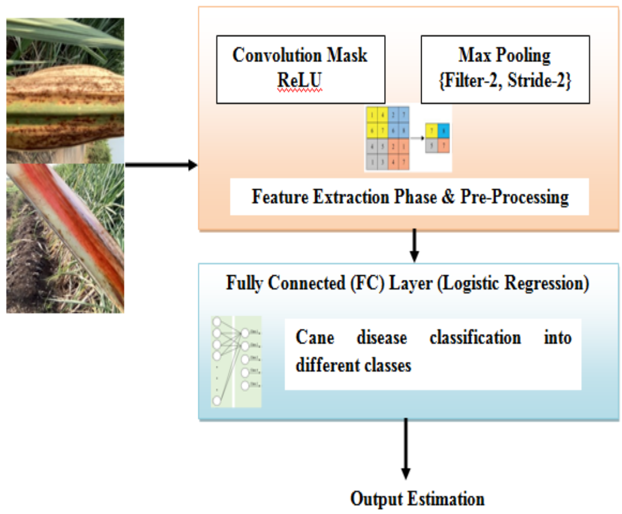


Fig. 6. Feature extraction architecture using CNN

IV. Results and Discussion

In our proposed work, a library of 3256 images of diseased sugarcane plant was gathered to serve as the system's input. The dataset for the proposed study was divided into learning and test phases, with learning phase using 80% of the images. Random Forest was employed as a classifier to categorize the sugarcane crop diseases. Figures 3 to 5 exhibit the tabulated outcomes from this technique, together with the statistical findings. When leaf images are acquired, noise is introduced, so as to reduce noise, we employ Savitzky-Golay filtering methods. We design a colour space transformation framework that is device agnostic, as a result, we develop the framework for colour transformation, which specifies colour space conversion. The image's colour values are changed in the next step to the colour space provided in the colour transformation structure using device-independent colour space transformation.

A device independent colour space is one in which the equipment used to create the colour has no bearing on the final hue. For instance, the brightness and contrast of the colour generated by a pixel with the specified RGB values will vary depending on the display device being utilized.

The colour space's main objective is to make it easier to specify colours in a uniform, widely recognized manner. Due to its foundation in human vision, the HSI (hue, saturation, intensity) colour model is well-liked [10]. Hue is a colour property that describes the prevailing colour that an observer would notice. Saturation describes the quantity of white light added to a colour, whereas intensity describes the size of the light beam. It is simple to change colour spaces from one to another. The H component is considered for more examination after the transformation procedure. S and I are eliminated since it provides no more details. The majority of the pixels that are green are identified in this stage. Following that, the predominantly green pixels are masked as follows based on the threshold value that is computed for these pixels: if the green component of the pixel intensity is less than the pre-computed threshold value, the red, green, and blue components of this pixel are assigned a value of zero. This is done because the green pixels primarily depict the healthy leaf portions and offer nothing to the disease diagnosis process. It also speeds up processing time by a large margin.

Plant diseases are divided into distinct categories using textural properties. The borders of the infected cluster and all pixels with red, green, and blue component values of zero are totally erased. This is advantageous since it provides a more precise disease categorization and drastically shortens the processing time. The colour representation of the infected cluster is changed from RGB to HSI.

When a leaf is affected by more than one disease, CNN [15] is better choice compared to other classifiers such as K-means clustering employed to divide it into 9 groups, one of which will contain healthy dataset and in remaining 8 classes, at least one of which will contain the diseased leaf. According to a set of characteristics, the K means clustering algorithms seeks to categorize objects (in this example, pixels) into K number of classes.

By reducing the sum of squares of distances between the objects and the associated cluster or class centroid, classification is accomplished. Images are segmented in this stage in order to distinguish the leaves from the backdrop. K-means clustering is used for segmentation, with one cluster centre for the background and one for the foreground. K-means clustering is an unsupervised learning approach [16 & 17] that divides data points into a predetermined number (k) of clusters [21 & 22] or groups according to how similar the data points are to one another.

The patch's size was selected such that the important data would not be lost. The patch size used in this method is 32X32 pixels. Extraction of the relevant portions is the following step. Not every part includes a substantial quantity of information.

In this study, neural networks are utilized to identify sugarcane crop diseases automatically. Due to the well-known method of the neural network as a good classifier [14] for many practical applications, it is chosen as a classification tool.

One of the crucial elements in creating a precise process model with NNs is the training and validation procedures. The training feature set is used to train the NN model, while the testing feature set is used to check the learned NN model's correctness.

Together, these two feature sets make up the dataset used for training and validation procedures. The correct network architecture, including the kind of the network and training procedure, must be set up before the data can be supplied to the ANN model [20]. The results of the proposed system for cane crop disease detection are shown in figure 7.

The parameters based on which the evaluation of CNN ensemble model and KNN model has been done is tabulated in the table I as shown for nine different classes of the crop.

The results of proposed algorithm for the detection of correct and wrong disease identification using CNN model and its comparison with KNN approach is as shown in figures 8 & 9.

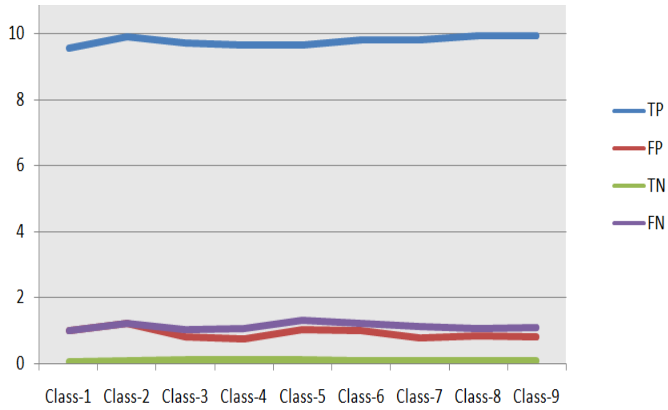


Fig. 8. ROC of Correct and Wrong disease Identification using CNN model

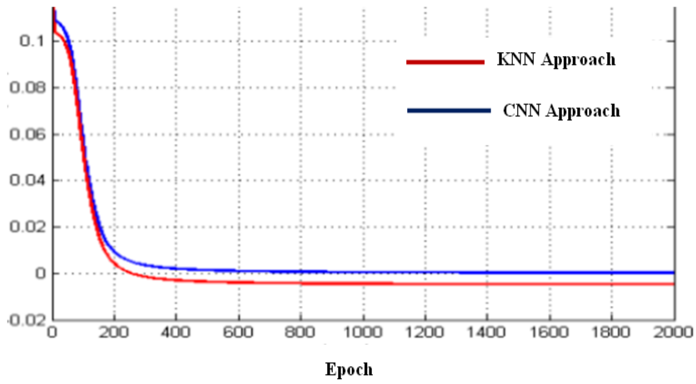
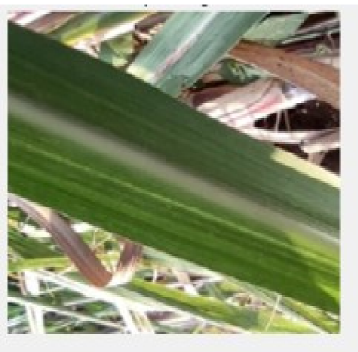
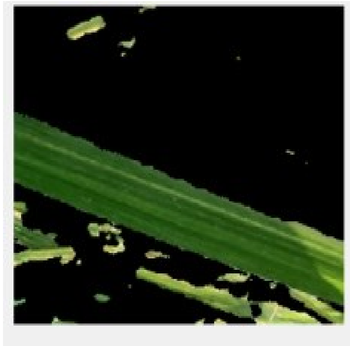
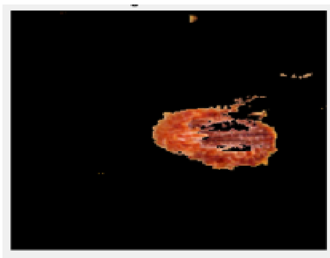
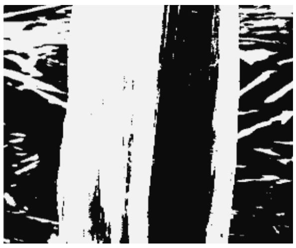
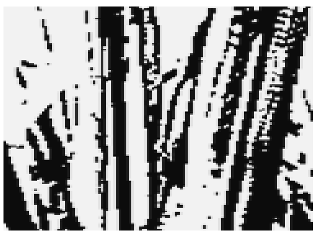


Fig. 9. Comparison between KNN and CNN approach

V. Conclusion And Future Enhancement

In this study, K-means clustering and Convolution Neural Networks (NNs) applications for categorizing and clustering diseases that affect sugarcane plantations have been developed, analyzed and compared respectively. The major goal of the suggested technique, which may detect sugarcane crops' disease with minimal computing effort, is to identify the disease is obtained with CNN technique. The CNN approach incorporating ensemble model gives better results than KNN approach at least by an improvement of 5.21% accuracy. With the right classifier, this method may be utilized for agricultural applications including the detection and categorization of diseases in different plant sections. In this research, use of potential hybrid methods like SG filter, adaptive histogram equalization, GLCM and ensemble approach for extracting low-level visual features and other characteristics improves the results. An extension of this work will focus on the detection of percentage of sucrose level and harvesting period for the sugarcane crop using advanced techniques as a future work.

Input Image	**Enhaced Contrast Image**	**Segmented Image**	**Type of disease**	**% of area affected**
			Healthy leaf	None
			Red rot	15.003
			Red striped	15.0015
			Red rust	15.0108
			Yellow leaf	15.0023
			Grassy shoot	15.023

Mosaic 15.002

Fig. 7. Results of Proposed System for disease detection

Table I
Metrics evaluation of the proposed system

CNN Approach incorporating ensemble model					
Disease Type	**Class Type**	**Recall**	**F1 score**	**Accuracy**	**Precision**
Healthy	Class 0	96	95	95.613	93.141
Red rot	Class 1	97	93	96.772	97.428
Rust	Class 2	98	96	97.243	98.210
Red Striped	Class 3	96	95	96.721	97.826
Smut	Class 4	97	96	96.869	98.221
Wilt	Class 5	97	94	97.105	97.327
Yellow leaf	Class 6	98	97	97.351	98.024
Mosaic	Class 7	96	98	96.025	96.692
Grassy shoot	Class 8	97	96	96.715	97.342
KNN Approach					
Disease Type	**Class Type**	**Recall**	**F1 score**	**Accuracy**	**Precision**
Healthy	Class 0	92.1	92	89.13	87.023
Red rot	Class 1	92.02	91	88.32	88.324
Rust	Class 2	90.42	92	90.63	89.213
Red Striped	Class 3	93.20	91	91.22	90.705
Smut	Class 4	92.12	90	90.39	91.313
Wilt	Class 5	91.92	89.8	92.32	91.438
Yellow leaf	Class 6	93.13	90	91.17	90.211
Mosaic	Class 7	92.70	89	88.57	90.327
Grassy shoot	Class 8	91.207	90	89.62	90.256

Declaration:

Data Availability

The datasets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that, we have no competing interests relevant to the content of this article.

Funding

The work does not have any funding details.

Authors' contributions

The authors' collected details of soil, and plants data nearby local regions to test and validate the algorithm developed.

Acknowledgements

The authors would like to thank all the farmers who helped in collecting data from their fields and agricultural department for providing the information.

REFERENCES

[1] Wenjiang Huang, Qingsong Guan, Juhua Luo, Jingcheng Zhang, Jinling Zhao, Dong Liang, Linsheng Huang, and Dongyan Zhang, (2014) "New Optimized Spectral Indices for Identifying and Monitoring Winter Wheat Diseases", IEEE journal of selected topics in applied earth observation and remote sensing,7(6), 2516-2524. https://doi.org/10.1109/JSTARS.2013.2294961

[2] K.Thangadurai and K.Padmavathi, (2014) "Computer Vision image Enhancement For Plant Leaves Disease Detection," Proc. of 2014 World Congress on Computing and Communication Technologies, Trichirappalli, India, 173-175. https://doi.org/10.1109/WCCCT.2014.39

[3] Monica Jhuria, Ashwani Kumar, and Rushikesh Borse, (2013) "Image Processing For Smart Farming: Detection Of Disease And Fruit Grading," Proc. of 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013), Shimla, India, 521-526.

[4] Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff Rohani Binti and S Mohamed Farook, (2012) "Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques," Proc. of 2012 Third International Conference on Intelligent Systems Modelling and Simulation, Kota Kinabalu, Malaysia, 291-296. https://doi.org/10.1109/ISMS.2012.33

[5] Mrunalini R. Badnakhe, Prashant R. Deshmukh, (2012) "Infected Leaf Analysis and Comparison by Otsu Threshold and k-Means Clustering", International Journal of Advanced Research in Computer Science and Software Engineering, 2(3), 449-452.

[6] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. ALRahamneh, (2011) "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications, 17(1), 31-38. https://doi.org/10.5120/2183-2754
[7] Chunxia Zhang, Xiuqing Wang and Xudong Li, 2010 "Design of Monitoring and Control Plant Disease System Based on DSP&FPGA," Proc. of 2010 Second International Conference on etworks Security, Wireless Communications and Trusted Computing, Wuhan, China, 479-482. https://doi.org/10.1109/NSWCTC.2010.246
[8] Meunkaewjinda, P. Kumsawat, K. Attakitmongcol and A. Srikaew, 2008 "Grape leaf disease detection from color imagery using hybrid intelligent system," Proc. of 2008 ECTI-CON, Krabi, Thailand, 513-516. https://doi.org/10.1109/ECTICON.2008.4600483
[9] Santanu Phadikar and Jaya Sil, (2008) "Rice Disease Identification using Pattern Recognition," Proc. of 2008 11th International Conference on Computer and Information Technology (ICCIT 2008), Khulna, Bangladesh, 420-423. https://doi.org/10.1109/ICCITECHN.2008.4803079
[10] Ricardo Jardim and F. Morgado-Dias, (2020) "Savitzky–Golay filtering as image noise reduction with sharp color reset", Microprocessors and Microsystems, 74 (2020), 1-9. https://doi.org/10.1016/j.micpro.2020.103006
[11] Iqbal, M. A., & Talukder, K. H. (2020, August). "Detection of potato disease using image segmentation and machine learning." In 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET) (pp. 43-47). IEEE.
[12] Deepak, A. H., Gupta, A., Choudhary, M., &Meghana, S. (2019, December). "Disease Detection in Tomato plants and Remote Monitoring of agricultural parameters". In 2019 11th International Conference on Advanced Computing (ICoAC) (pp. 28-33).IEEE. https://doi.org/10.1109/ICoAC48765.2019.246812
[13] Hidayatuloh, A., Nursalman, M., & Nugraha, E. (2018, October). "Identification of tomato plant diseases by Leaf image using squeezenet model". In 2018 International Conference on Information Technology Systems and Innovation (ICITSI) (pp. 199-204).IEEE. https://doi.org/10.1109/ICITSI.2018.8696087
[14] Sarangdhar, A. A., &Pawar, V. R. (2017, April). "Machine learning regression technique for cotton leaf disease detection and controlling using IoT." In 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA) (Vol. 2, pp. 449-454).IEEE https://doi.org/10.1109/ICECA.2017.8212855
[15] Hatuwal, B. K., Shakya, A., & Joshi, B. (2020). "Plant leaf disease recognition using random forest, KNN, SVM and CNN." Polibits, 62, 13-19. https://doi.org/10.17562/PB-62-2
[16] Dhingra, G., Kumar, V., & Joshi, H. D. (2018). „Study of digital image processing techniques for leaf disease detection and classification." Multimedia Tools and Applications, 77(15), 19951-20000.
[17] Albattah, W., Nawaz, M., Javed, A., Masood, M., & Albahli, S. (2022). "A novel deep learning method for detection and classification of plant diseases." Complex & Intelligent Systems, 8(1), 507-524.
[18] Hungilo, G. G., Emmanuel, G., & Emanuel, A. W. (2019, April). "Image processing techniques for detecting and classification of plant disease: a review." In Proceedings of the 2019 international conference on intelligent medicine and image processing (pp. 48-52). https://doi.org/10.1145/3332340.3332341
[19] Khirade, S. D., & Patil, A. B. (2015, February). "Plant disease detection using image processing." In 2015 International conference on computing communication control and automation (pp. 768-771). IEEE. https://doi.org/10.1109/ICCUBEA.2015.153
[20] Ramesh, S., & Vydeki, D. (2020). "Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm." Information processing in agriculture, 7(2), 249- 260. https://doi.org/10.1016/j.inpa.2019.09.002
[21] Das, D., Singh, M., Mohanty, S. S., & Chakravarty, S. (2020, July). "Leaf Disease Detection using Support Vector Machine." In 2020 International Conference on Communication and Signal Processing (ICCSP) (pp. 1036-1040).IEEE. https://doi.org/10.1109/ICCSP48568.2020.9182128
[22] Samajpati, B. J., & Degadwala, S. D. (2016, April). „Hybrid approach for apple fruit diseases detection and classification using random forest classifier." In 2016 International conference on communication and signal processing (ICCSP) (pp. 1015-1019),IEEE. https://doi.org/10.1109/ICCSP.2016.7754302
[23] Swapnil & Sanjay (2023). "Enhancing sugarcane disease classification with ensemble deep learning: A comparative study with transfer learning techniques, 9(2023), 1-19, Heliyon. https://doi.org/10.1016/j.heliyon.2023.e18261