Research Article

A Comparative Analysis of Machine Learning Algorithms for Detection of Organic and Nonorganic Cotton Diseases

Sandeep Kumar [1], **Arpit Jain** [2], **Anand Prakash Shukla** [3], **Satyendr Singh** [4], **Rohit Raja** [5], **Shilpa Rani** [6], **G. Harshitha** [1], **Mohammed A. AlZain** [7], **and Mehedi Masud** [8]

[1] *Sreyas Institute of Engineering & Technology, Hyderabad, India*
[2] *Teerthanker Mahaveer University, Moradabad, U.P, India*
[3] *KIET Group of Institutions, Gaziabad, India*
[4] *BML Munjal University, Gurugram, India*
[5] *Central University, Bilaspur, Chhattisgarh, India*
[6] *Neil Gogte Institute of Technology, Hyderabad, India*
[7] *Department of Information Technology, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*
[8] *Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*

Correspondence should be addressed to Mehedi Masud; mmasud@tu.edu.sa

Cotton is the natural fiber produced, and the commercial crop grown in monoculture on 2.5% of total agricultural land. Cotton is a drought-resistant crop that provides a reliable income to the farmers that grow under the area with a threat from climatic change. These cotton crops are being affected by bacterial, fungal, viral, and other parasitic diseases that may vary due to the climatic conditions resulting in the crop's low productivity. The most prone to diseases is the leaf that results in the damage of the plant and sometimes the whole crop. Most of the diseases occur only on leaf parts of the cotton plant. The primary purpose of disease detection has always been to identify the diseases affecting the plant in the early stages using traditional techniques for better production. To detect these cotton leaf diseases appropriately, the prior knowledge and utilization of several image processing methods and machine learning techniques are helpful.

1. Introduction

Day by day, all over the world, agriculture land is going to be reduced because the population is increasing rapidly and lack of water resources. Disease in the plant is one of those hazards that have to be examined at this stage. In contrast, the isolation of plants from their natural environment is being happened, and they are grown in unusual conditions. Many valuable crops and plants are very vulnerable to disease. They would have a great struggle to survive in nature without human involvement. Yield loss in harvests is regularly connected with plant illness or factors, such as climate, water availability, and supplement accessibility [1]. To improve the productivity of the crop, environmental factors or product resources such as temperature and humidity are important [2]. An important role is played by the root exudates of the plant, which helps in improving the nutrients of the soil [3]. Compared to their wild relations, cultivated plants are always more flexible to diseases. This is the large numbers of the same species or different kinds, having a similar gene grown together, sometimes over many kilometers of distance.

The microorganisms which cause diseases may spread rapidly under these conditions. Plant illnesses influence the plants' development and yield, making a socio-biological and monetary effect on agriculture [4]. The past writings reveal that the various kinds of leaf diseases, such as spots, mildews, rusts, and blight, cause scarcity and many other extreme changes in nations' economies. Plant diseases are one of the ecological factors that facilitate keeping living plants and animals in balance with each other. Plant cells mainly enhance their defenses in opposition to the animals, insects, and also pathogens through signaling pathways that are contained in them. For thousands of years, humans have selected and cultivated plants for food, medicine, clothing, shelter, fiber, and beauty with proper care. Therefore, the monitoring of regional crop diseases is important to enhance food security [5].

1.1. Motivation and Contribution. The growth of most developing countries depends on agriculture. Agriculture plays an essential role in India's economic growth. Most of the north Indian states grow cotton crops in 95% of the agricultural land. Various diseases affect the productivity of cotton crops. It is mandatory to detect and classify these diseases in the early stage for the better productivity of cotton crops. Considering the same, this review paper summarized the various cotton diseases and symptoms of the diseases. This paper analyzes various machine learning algorithms of segmentation, detection, and classification techniques to identify cotton diseases. This study would be helpful for future research in this field.

The main contributions of this paper are as follows:

(i) The various cotton diseases are classified, and their symptoms are also presented

(ii) The existing cotton disease prediction techniques are studied, and these are categorized according to the technologies used

(iii) The comparative analyses of the existing techniques are performed to assess their advantages and disadvantages

(iv) The various databases that have been used to conduct these studies are also discussed.

(v) The various performance metrics that can be used to evaluate the efficiency of disease prediction techniques are also presented.

The rest of the paper is organized as follows: Section 2 describes the background of the cotton crop. Section 3 discusses the existing literature related to cotton diseases. Section 4 presents the discussion of the existing techniques in terms of advantages, disadvantages, databases used, and performance evaluation. Section 5 concludes the paper and provides the future directions.

2. Background of Cotton Crop

Cotton is a fiber crop that fills in as a wellspring of feed, staple, biofuel production, and oilseed harvest, which provides 35% globally of the total fiber and raw material for producing textiles [6–8]. It is the most important and principal cash crop that affects India's economy in many

ways [9, 10]. It is always considered the most economically main crop, but various disease pathogens affect its production [11–13]. The largest cotton yielding country is India, as compared to other countries in the entire world, as shown in Figure 1. All these together provide most of the yield all over the world. Area-wise, India is the first among them. Production-wise, it has a third place.

Cotton plays a major role in the agrarian economy and industrial activities of India [14]. We have nine states all over India that grow cotton which is partitioned into zones, namely, north, central, and south. Haryana, Punjab, Western Uttar Pradesh, and Rajasthan are northern zone states. Maharashtra, Gujarat, and Madhya Pradesh are the central zone, whereas Andhra Pradesh, Tamil Nadu, and Karnataka fall under the south zone, as shown in Table 1 and Figure 2. All of the states in each of the zones extend to 90 percent of the land to cultivate cotton crops and produce 95% of the total cotton production in India. The other 5% of cotton production is contributed by small states such as Bihar, Meghalaya, Tripura, Orissa, and Assam. The zones in which cotton is produced differ depending upon the soil type, sources available for better crop protection, level of production, and geography.

Plant illnesses have perceptually been a challenge to plant development and harvest creation in numerous global segments. Plant infections will influence plants by numerous cycles such as the absorbance and movement of water and supplements, chemical change, flowering, development of the fruit, plant growth, cell division, and extension. Collecting, acquiring, and analyzing cotton diseases that are being processed manually leads to a lack of efficiency, making the identification difficult [15]. Cotton plant diseases may be caused due to the different types of microorganisms, nematodes, and other agents. Therefore, efforts should be directed towards avoiding disease epidemics just like blight diseases and developing effective disease management methods.

2.1. Types of Cotton. Cotton is the cumulative name given according to perpetual shrubs in the family *Malvaceae* grown as the soft fiber. The four different kinds of plants in the genus are shown in Figure 3, which preserve the plant seeds. *Gossypium hirsutum* is estimated at roughly 90% in the world. These plants have a central main stem which gives various branches upwards. The plant leaves are set out around those branches, which possess 3 to 5 triangular lobes and long petioles. The plant gives a single bud that can be red or purple, yellowish-white in color on each auxiliary branch, and forms a leathery, oval seed capsule known as "boll," generally 2 to 6 cm long. The cotton plant can reach heights of 1 to 1.5 meters (3.3–4.9 ft).

2.2. Diseases of Cotton and Their Symptoms. The natural and most frequently occurring cotton disorders are due to the loss of nutrition, environmental stress, and chemical factors that cause imbalances. These factors affect the production of cotton during the growth of the crop. There is a contrasting difference between the modification of the plant and the harmful environments, including the disorders. Verticillium
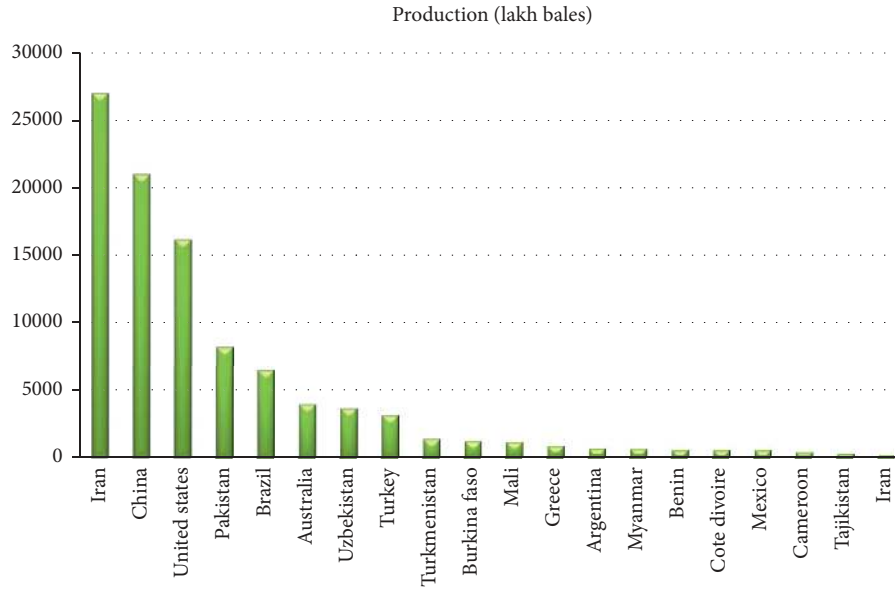
Figure 1: Cotton production by the various countries.

Table 1: In India, state-wise production (lakh bales).

State	2015-16	2016-17	2017-18	2018-19
Punjab	12.99	6.26	8.99	11.49
Haryana	22.99	15.61	20.49	23.10
Rajasthan	16.99	16.02	17.01	23.08
Gujarat	111.99	91.01	94.99	103.75
Maharashtra	80.01	77.20	87.90	86.85
Madhya Pradesh	18.99	19.30	21.01	21.49
Telangana	51.10	57.99	49.01	56.77
Andhra Pradesh	27.49	24.86	20.07	21.90
Karnataka	35.01	19.01	19.70	19.76
Tamil Nadu	6.01	6.09	5.07	6.59
Orissa	3.02	3.01	3.01	3.61
Others	2.01	2.11	2.14	2.37

wilt and cotton leaf curl diseases are major limiting factors for cotton production [16, 17]. Verticillium wilt is a highly hazardous disease that is strongly epidemic [18]. The abundance of wilt pathogens is one of the many factors which results in disease occurrence [19]. The cotton boll is also a chief component of fiber production [20]. Boll injury and symptoms of cotton boll with leaf-footed bugs were similar to those of plant bugs [21, 22]. Early detection of bollworm and pink bollworm with the help of field evolved resistance was also discussed in [23]. Fungicides were frequently applied during the planting stage to control cotton seedling disease by spraying [24]. Tobacco Streak Virus (TSV) belongs to the genus Ilavirus of the family of *Bromoviridae*. Identification of the symptoms of TSV infection by visual observation of the plants often results in misdiagnosis as this virus matches those reflecting nutritional disorders affecting cotton [25]. A sample example is shown in Figure 4.

(1) *Angular Leaf Spot or Black Arm Disease.*

   Causal Organism. Xanthomonas axonopodis pv. *malvacearum* (Smith) Vauterin.

Symptoms. Small spots appear under the cotyledons, which may dry and fall off. Such similar spots may also appear on the leaves. They become angular bound by leaflets and turn brown. Several small spots may combine. The infected petiole may get damaged. Elongated dark brown lesions appear on the stem as well as petioles and branches. The young stems may be enclosed and killed in the dark arm phase. Sunken black lesions are seen on the bolls. Young bolls may wither. The disease attacked stem may become weak. Discoloration of lint may appear.

Best Methods Used for Detection. KNN and principal component analysis.

(2) *Vascular Wilt Disease.*

   Causal Organism. Fusarium oxysporum f. sp. *vasinfectum* (Atk.) Snyder and Hansen.

   Symptoms. The soils with pH values ranging from 6 to 8.00 will affect this disease. Almost all the parts of the plant get affected. The yellowing of cotyledons, sautéing of petioles, and filling of the dried leaves occur at the seedling stage. In contrast, in the young and grown-up plants, deficiency of bloat, leaf hanging, delicate shoots, caramelizing, and, lastly, the demise of the plants takes place.

   Best Methods Used for Detection. Edge detection.

(3) *Grey Mildew or Dahiya Disease.*

   Causal Organism. Ramularia areola Atk and *Mycosphaerella areola.*

   Symptoms. The organism typically assaults the more seasoned leaves making nonuniform rakish, pale, and lackluster spots. They are generally bound by the veinlets and are generally below or above the piece of the leaf. Usually, some spots which might be up to hundreds are recognized on the leaf surface. In
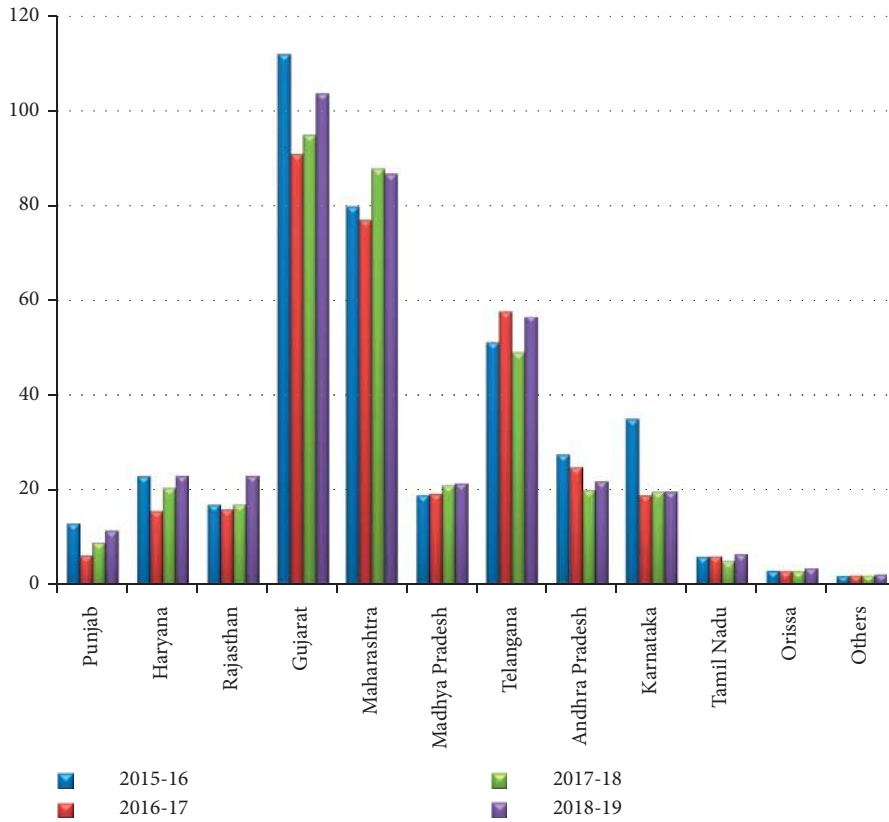
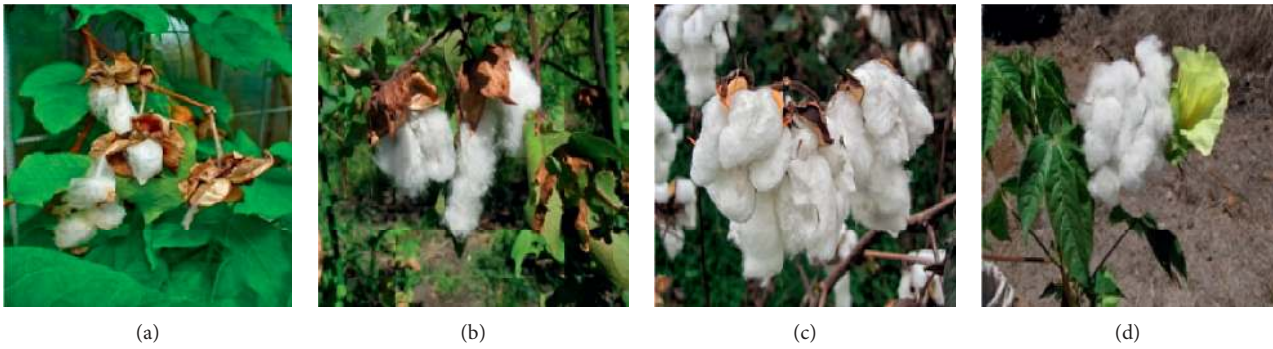Figure 2: Cotton production by the various states in India.

Figure 3: (a) *Gossypium herbaceum*. (b) *Gossypium arboretum*. (c) *Gossypium hirsutum*. (d) *Gossypium barbadense*.
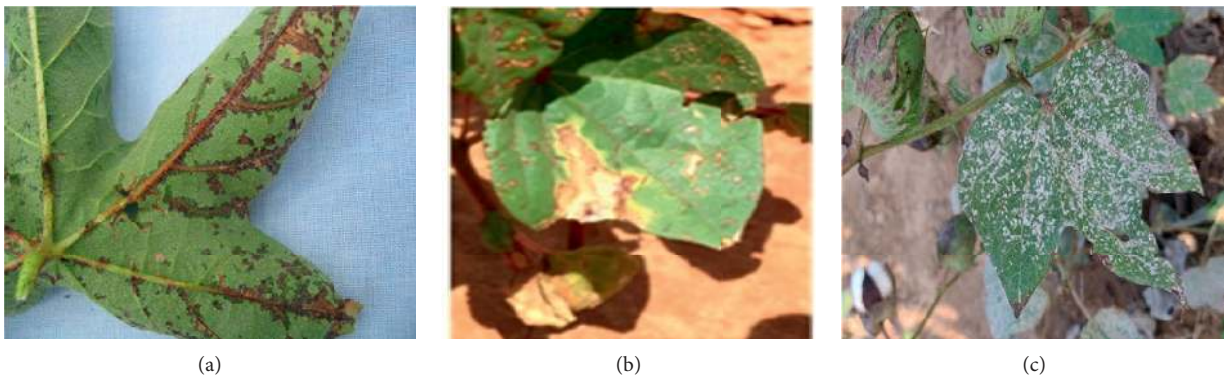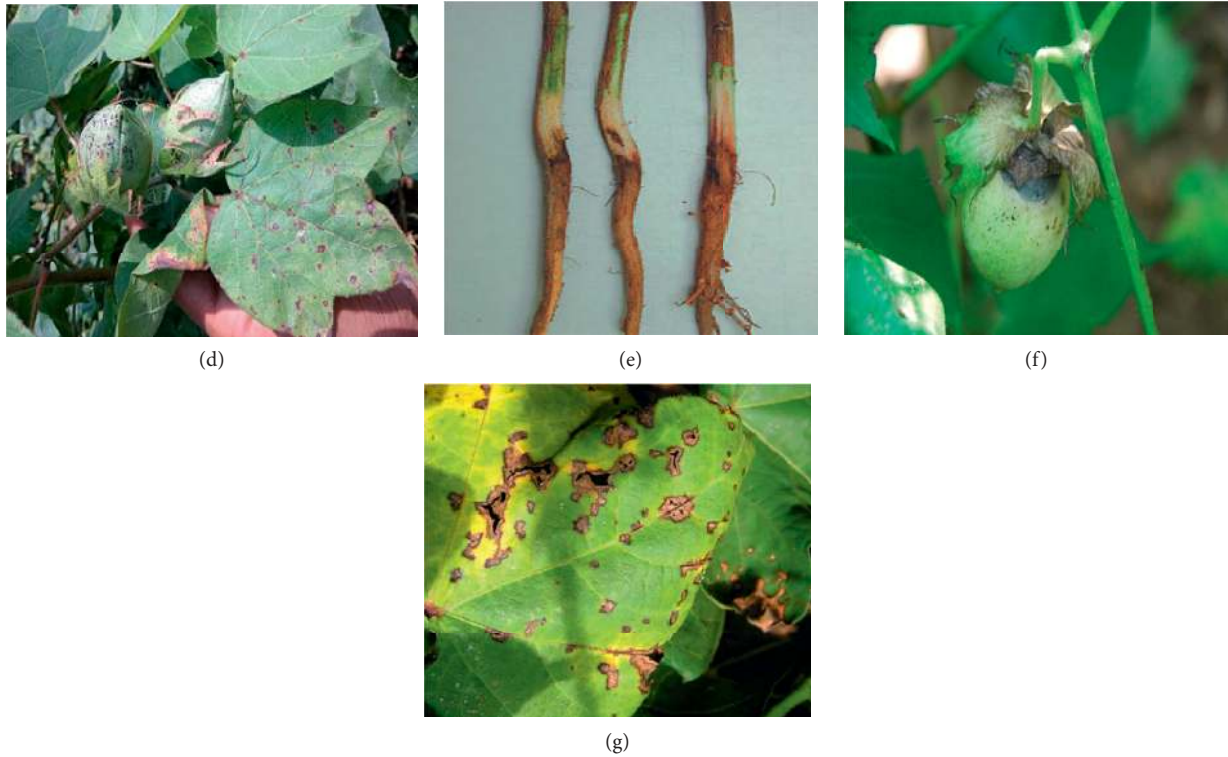

Figure 4: Continued.

Figure 4: (a) Angular leaf spot, (b) vascular wilt, (c) grey mildew, (d) anthracnose, (e) root rot, (f) boll rot, and (g) leaf spot or blight.

severe infections, the leaves fall off earlier, changing their color to yellowish-brown.

Best Methods Used for Detection. KNN and decision tree algorithm.

(4) *Anthracnose Disease.*

Causal Organism. Colletotrichum gossypii Southw and *Colletotrichum capsici* (Syd.) Butl and Bisby and *Glomerella gossypii* (Southw) Edgerton.

Symptoms. In the seedling stage, minor, reddish round spots appear on the cotyledons and primary leaves. When the lesions are on the upper region, the stem may be enclosed, causing seedlings to wilt and die. In adult plants, the fungus infects the stem, causing it to separate and shred-off the bark. The symptoms are prominent on bolls as water-soaked, circular, slightly sunken reddish-brown spots which turn black. As a result of boll infection, they open prematurely. As a result, the lint becomes stained, hard, and compact.

Best Methods Used for Detection. Decision tree and fuzzy feature.

(5) *Root Rot Disease.*

Causal Organism. Rhizoctonia bataticola (Taub.) Butl.

Symptoms. The symptoms might be visible in the seedling phase, where we can find brown spots in color on the cotyledons. At the collar area, there is dark shading, which might be extended towards the lower parts. The sinewy roots go through rotting. Rotting and shredding of the roots appear by the bark of the roots. Influenced plants can be pulled out effortlessly. The patches of infection are shown up in the field.

Best Methods Used for Detection. Neural network.

(6) *Boll Rot Disease.*

Causal Organism. Macrophomina phaseolina (Tassi) Goid.

This is a complicated disease that occurs through various fungal organisms: Fusarium moniliforme, Colletotrichum capsici, Aspergillus flavus, A. niger Rhizopus nigricans, Nematosporanagpuri, and Botryodiplodia sp.

Symptoms. In the beginning, the disease's small spots, which are brown or black in color, appear that expand to enclose the complete bolls. The contamination goes to tissues present inside, resulting in the infected stem or lint. The bolls never crack open, and they do not immediately fall off in the early stages. The rotting of the stem may be external in some cases, which causes infection to the pericarp that leaves the internal tissues freely. A large number of fungi may be observed on the infected bolls.

Best Methods Used for Detection. Neural network.

(7) *Leaf Spot or Blight Disease.*

Causal Organism. Alternaria macrospora Zimm.

Symptoms. The disease can occur in many stages. However, the most severe one occurs in the leaves after 45 to 60 days of planting. Tiny, pale brown, irregular round spots, measuring 0.5 to 6 mm diameter, may be present on the leaves. Each spot contains a central lesion surrounded by concentric rings. Several spots merge to form blighted areas. The affected leaves become weak and wither. Stem lesions are observed rarely. But in extreme cases, these spots may appear on bracts and bolls.

Best Methods Used for Detection. Image processing and support vector machine.

2.3. Organic Cotton. Organic cotton is created by the utilization of techniques and materials that lowly affect the climate. Organic production structures top off and hold soil ripeness, lessen the utilization of hurtful pesticides and composts, and develop organically unique agribusiness. Outsider confirmation partnerships affirm that regular makers use exclusively techniques and materials permitted in organic production. Organic cotton is developed without the utilization of harmful and consistent pesticides or manufactured composts. Moreover, government rules confine the utilization of hereditarily designed seeds for natural cultivating. Organic cotton still solely occupies a tiny position of far much less than 1% of international production of cotton. However, the quantity of farms changing to organic cotton and the range of tasks for the growth are continuously increasing. There are several motives to grow cotton organically. The poor effects of farming traditional cotton on the environment and fitness are apparent. It is necessary to adopt a range of integrated measures in a device approach, ensuring that the interplay amongst the soil, plants, surroundings, and people is nicely balanced to get satisfactory yields and income in natural cotton farming. The "elements for success" all need to be used together:

(1) Suitable measures to improve and hold soil fruitfulness

(2) Establishment of yield revolution and harvest variety; cultivating characteristic equilibrium

(3) Selection of assortments appropriate to the conditions (soil, accessibility of water system, and market necessities)

(4) Appropriate sorts and amounts of excrements at the correct time

(5) Timely harvest the board, for example, intercultural activities, weeding, and water system

(6) Careful checking of the yield and enough assurance towards bothers as per the idea of the money-related limit level

(7) Timely and alluring picking of the cotton

(8) Sufficient documentation for examination and affirmation

(9) Capacity building and testing for constant improvement

2.4. Diseases in Organic Cotton. In most of the semiarid tropical regions, illnesses are not a big problem in organic cotton. However, diseases that from time to time happen and methods for stopping or treating them are given below:

(i) Bacterial blight: leaves show slick dark spots; stems flip dark; defoliation and boll shedding if the invasion is high. A strategic distance can be followed by methods for the use of safe assortments.

(ii) Root rot and boll rot: generally brought about through fungi and bacteria.

(iii) Fusarium wilt: crop rotation practice; getting rid of cotton stalks after harvesting.

3. Literature Survey

Computer technology has been analyzed and had been made into use broadly in agricultural applications to disclose and recognize the infection of the plant and classify them. A computer vision-based algorithm (SLIC-Simple Linear Iterative Clustering) was developed to detect the infected part of the disease and classify the disease using the support vector machine to the type which it belongs to (fungal or bacterial). Using this algorithm, three types, namely, alternate-fungal, bacterial blight, and whitefly diseases, were detected [26]. The emphasis on recent studies that reflect image processing techniques and their benefaction in detecting the plants and their diseases followed by the classification techniques under several circumstances is presented in this study. The structure of methodology for leaf disease identification and the classification techniques are shown in Figures 5 and 6, respectively.

The main focus of the pathologists is to identify the diseases that may occur in other various parts such as roots, branches, stem, and leaf of the plant. As mentioned before, this paper focuses distinctly only on the leaves of a cotton plant. The information related to the identification of disease and its classification is available in previous research. Most of them usually involve a preprocessing stage and methods of segmentation accompanied by identifying the feature and level of classification using appropriate measures. All the techniques have applied some sort of preprocessing to eliminate the noise and segmentation procedure for selecting the part of the disease that occurred. The extension of this paper is based on the classification of diseases which is also an important aspect of proper detection. Classification is a procedure that is related to the division where assets are categorized according to the field. It helps in recognition of ideas and objects that are further differentiated and understood.

3.1. Disease Detection Analysis with Image Segmentation and Feature Extraction. Rothe and Kshirsagar [27] discussed the identification of diseases utilizing picture preparing strategies that extracted the highlights, for example, region, major and minor axes, direction, and so on, from pictures of influenced ailment leaves. The paper has two strategies, in particular, image enhancement and image segmentation.
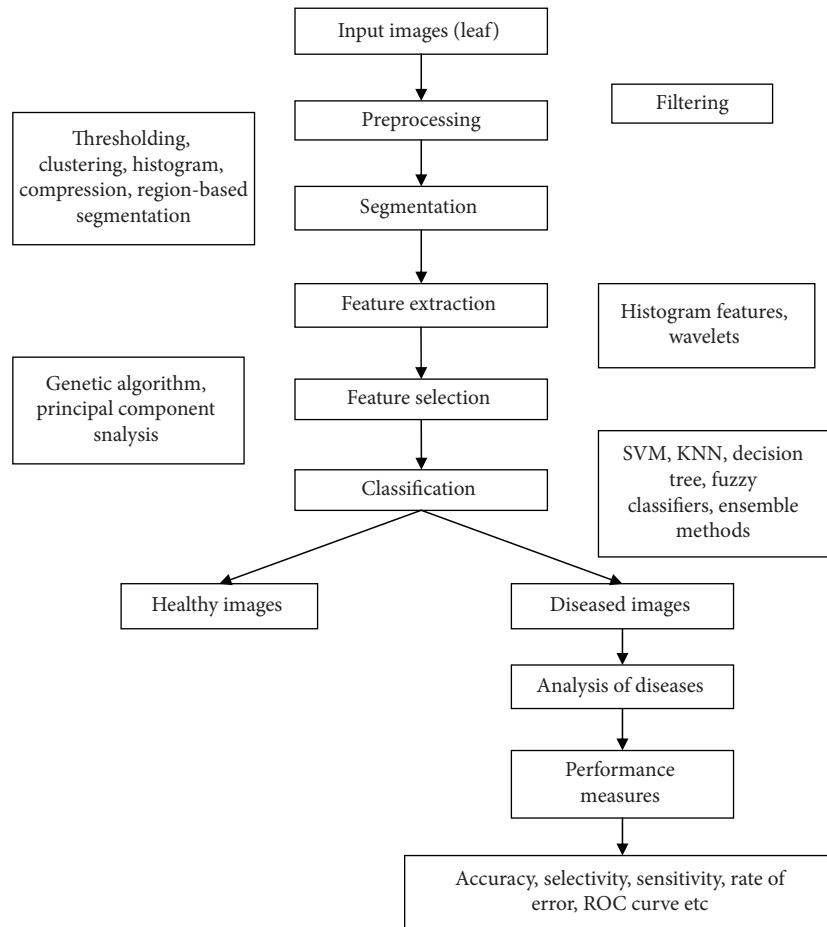
```
                    ┌─────────────────────┐
                    │  Input images (leaf) │
                    └─────────────────────┘
                               │
                               ▼
                    ┌─────────────────────┐        ┌─────────────┐
                    │     Preprocessing    │        │  Filtering  │
                    └─────────────────────┘        └─────────────┘
                               │
                               ▼
┌──────────────────┐  ┌─────────────────────┐
│  Thresholding,    │  │     Segmentation     │
│  clustering,      │  └─────────────────────┘
│  histogram,       │           │
│  compression,     │           ▼
│  region-based     │  ┌─────────────────────┐        ┌──────────────────┐
│  segmentation     │  │  Feature extraction  │        │ Histogram features,│
└──────────────────┘  └─────────────────────┘        │     wavelets      │
                               │                      └──────────────────┘
┌──────────────────┐           ▼
│ Genetic algorithm,│  ┌─────────────────────┐
│ principal component│ │   Feature selection  │
│    snalysis       │  └─────────────────────┘
└──────────────────┘           │
                               ▼
                    ┌─────────────────────┐        ┌──────────────────┐
                    │    Classification    │        │ SVM, KNN, decision│
                    └─────────────────────┘        │ tree, fuzzy       │
                       /          \                 │ classifiers,      │
                      /            \                │ ensemble methods  │
         ┌────────────────┐   ┌────────────────┐    └──────────────────┘
         │ Healthy images │   │ Diseased images│
         └────────────────┘   └────────────────┘
                                      │
                                      ▼
                            ┌─────────────────────┐
                            │  Analysis of diseases│
                            └─────────────────────┘
                                      │
                                      ▼
                            ┌─────────────────────┐
                            │    Performance       │
                            │     measures         │
                            └─────────────────────┘
                                      │
                                      ▼
                   ┌──────────────────────────────────────┐
                   │ Accuracy, selectivity, sensitivity,   │
                   │ rate of error, ROC curve etc          │
                   └──────────────────────────────────────┘
```

FIGURE 5: General structure of cotton leaf diseases detection.

```
┌──────────────────────────────────────┐
│        ┌──────────────────────┐       │
│        │ Step 1: segmentation │       │
│        └──────────────────────┘       │
│  ┌──────────────┐   ┌──────────────┐  │
│  │  Nonground   │   │ Ground class │  │
│  │    class     │   │              │  │
│  └──────────────┘   └──────────────┘  │
└──────────────────────────────────────┘
         ⤵                    ⤵
┌──────────────────────────────────────┐
│        ┌──────────────────────┐       │
│        │ Step 2: segmentation │       │
│        └──────────────────────┘       │
│  ┌──────────────┐   ┌──────────────┐  │
│  │ Nondiseased  │   │Diseased class│  │
│  │    class     │   │              │  │
│  └──────────────┘   └──────────────┘  │
└──────────────────────────────────────┘
```
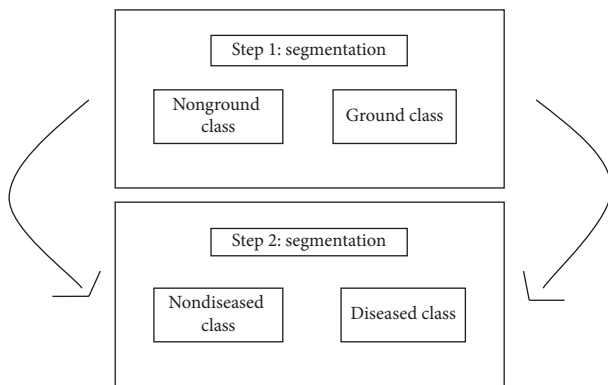
FIGURE 6: Stages for segmentation of diseased images.

For image enhancement, a low pass channel is utilized to evacuate the commotion present in the picture. The Gaussian channel is utilized as the edge that compares to high recurrence is to be held. It does not have a sharp cut-off; however, it comprises rich and regular stop band reaction that permits adequate higher recurrence parts and subsequently shows little league transfer speed item. The element extraction process includes shading and shape-based extraction. The essential nature of the visual substance is shading; along these lines, it is utilized to relate and speak to a picture. The shape-based highlights that are chosen to distinguish the sicknesses involve region, sharpness, edge, Eigen esteems, and viewpoint proportion. The quantity of pixels speaks to the territory of the article; along these lines, it tends to be outlined by checking the number of item pixels. The proportion of the length of the significant tomahawks to minor tomahawks is called angle proportion which produces data regarding the flat stretch of the infected spot. By computing normal separation between two limit focuses and the significant hub, the sharpness of the state of the spot is resolved, as shown in Table 2.

Revathi and Hemalatha [28] introduced an image processing procedure alongside edge detection. At first, they took the computerized pictures utilizing an advanced portable camera. At that point, picture handling strategies are applied to the obtained pictures to remove RGB highlights, which are important for additional examination. Some expository strategies were utilized to group the pictures as per the particular issue from that point onward. Users take pictures of the leaves utilizing sensors and then forward them to the PC framework where the info pictures are dissected utilizing the homogenous edge recognition calculations, and the pixels call work rationale utilized for the location of the influenced pieces of the leaf disease-wise. The homogeneous-based edge detector gains the outcome and

TABLE 2: Shape features of the cotton leaf.

Name of disease	Area	Perimeter	Aspect ratio	Sharpness
Leaf blight	98.00000	0.7817	1.55051	490.11292
Myrothecium	71.00000	1.6598	1.00619	194.20520
Alternaria	85.00000	2.00000	2.24798	502.62412

partitions through the normal scope of the part. This partition takes out the effect of nonlevel lighting in the picture. RGB picture changes over into the dark scale picture, utilizing edge identification procedures to recognize the leaf spot picture malady, influenced part. The use of shading channel checking and RGB highlights aids in recognizing different parts of images that yield in distinguishing the infected part. Kirthi Pilli et al. [29] discussed the eAGRO technique for early crop disease detection. The agro advisory system is used between farmers and agricultural experts by bridging a gap between them [30].

Comparative analysis of various image processing techniques has been carried out, and it is shown in Table 3.

3.2. Disease Detection Analysis Using K-Means and KNN.
Revathi and Hemalatha [31] described the recognition of cotton diseases. The following related examination on quick and exact recognition and arrangement of plant infections utilized by K-means and KNN result is 94%, demonstrating that the mix of K-means may give preferred outcomes over other related research considerations.

Parikh et al. [32] examined the recognition of the severity of the grey mildew with white spots blend that creates white spots that are huge in size on the cotton leaf. After the extreme contamination, the shrinking of leaves occurs. Stage one and two offer comparable infection attributes (white spots) yet can be isolated depending on the white spots' size. The second and third stages have distinctive infection side effects (white spots versus dried dark colored leaf) and are isolated dependent on the side effects. Stage one utilizes the K-nearest neighbor (KNN) classification method with a series of 15 highlights that are measurable. So the stage two classification utilizes the KNN classifier with two essential highlights: tone and luminance. The classification and the general component extraction are outlined. Narmadha and Arulvadivu [33] presented the image processing technique along with the color and shape features extraction and detected leaf blasts, brown spots, and narrow brown spots with the help of the methods mentioned. Schuster et al. [34] utilized a two-advance k-implies bunching calculation to distinguish the zones of the executives. One of the basic components while utilizing K-means is recognizing the best characteristics or factors for the executive's zone portrayal. Previously, scientists have investigated utilizing yield or dirt properties such as incline, rise, and electrical conductivity. The k-implies calculation is a, for the most part, utilized bunching technique pertinent broadly for unaided learning and distinguishing structure in the dataset. The comparative analysis of K-means and KNN techniques is shown in Table 4.

The algorithm implements clustering by limiting a target work in a registering way, generally the total over the square of separation of each point from the comparing group hub. While the AI approach laid out right now presents the administration zones for cotton, a significant issue keeps on existing. It includes the presentation of the progress rate into the machine learning calculation directly. Such progress might help in exhibiting machine learning to coordinate the current operational conditions. Recognizing the cool air territories through temperature detecting innovation and machine learning calculations speaks to another new application for the examination right now.

3.3. Disease Detection Analysis Using Principal Component Analysis.
Gulhane and Kolekar [35] predicted distinguishing the diseases on cotton leaves utilizing machine learning techniques; it is conceivable to expand the degree for the location of different diseases inside obvious just, and wavelength regions are invisible. In actualizing multivariable KNN and PCA systems, dissecting the factual information got practical, identified with the Green (G) channel of RGB picture. In a large portion, sicknesses that appear on the plant leaves are specific, leaf necrosis, grey mold, Alternaria, and magnesium deficiency. PCA is the direct technique that licenses for perceiving the uncorrelated parts. In the PCA technique, each element in the group highlights is considered an irregular component space plotted in the space, structuring a haze of highlights. Here, Figure 7 signifies the bend between the recurrence of sicknesses (%) over the 110 test samples. Table 5 shows the result of detection using the PCA technique.

3.4. Disease Detection Analysis Using Ensemble Methods (Random Forest) and Decision Trees.
Mehta et al. [36] presented the utilization of random forests with decision trees faring nearly the same all through the machine learning models. MLP classifier is utilized for the characterization, yet it did not get the sickness arrangement properly. The multiyield regressor performed all right with a random forest system as utilized. Random forest is an outfit technique where an example is taken from the preparation set with the trade for building a tree in the gathering. During the tree development, when a hub is split, the split that is best among the highlights will be picked, which is taken randomly. Chopda et al. [37] gave the execution of this project depending on the small scale. When all the functionalities are executed, the project would be extended for enormous scope.

Decision tree classifier is a straightforward and generally utilized characterization technique. The advantage of utilizing this calculation is simple as it bids a direct plan for the classification problem to be solved. It infers many inquiries concerning the characteristics of the test record. The Boa constrictor programming alongside the Jupyter Notepad is utilized to foresee the reasonable yield. The primary strides of the undertaking include the following: get cultivating information utilizing sensors and human contribution via portable applications, driving the gathered information into

TABLE 3: Comparison of image processing techniques.

Author and year	Methodology	Detected diseases	Remarks	Gaps identified
Rothe and Kshirsagar [27]	Image enhancement Image segmentation Feature extraction	Bacterial blightAlternaria Myrothecium	Sharpness of bacterial blight 490.11292Alternaria 194.20520Myrothecium 502.62412	The detection of diseases is limited
Revathi and Hemalatha [28]	Edge detection	Fusarium wiltVerticillium wilt Leaf blight	Accuracy of 98.1%	If the size of the image is enlarged, quality would be reduced
Kirthi Pilli et al. [29]	Image acquisitionPreprocessingSegmentationFeature extraction	Bacterium blight Magnesium deficiency	Accuracy 90%	Wider images result in the incorrect classification

TABLE 4: Comparison of K-means and KNN techniques.

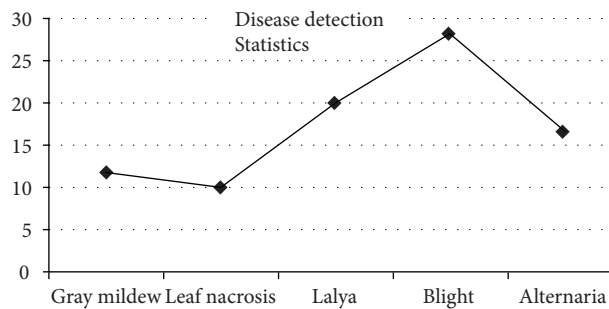| Author and year	Methodology	Detected diseases	Remarks	Dataset	Gaps identified
Revathi and Hemalatha [31]	K-means nearest neighbour	Grey mildew Bacterial blight Leaf curl virus disease-gemini virusAlternaria leaf spot	Accuracy of 92%	N/A	The accuracy of KNN is less compared to others
Parikh et al. [32]	KNN classification	Grey mildew	Accuracy of 82.5%	150 images 40 images (1024 × 1024 pixels)	Size of the block is chosen depending upon the presence of some disease pattern
Narmadha and Arulvadivu [33]	K-means technique	Blast Brown spot	Accuracy of 94.7%	MATLAB image library	Time consuming
Schuster et al. [34]	K-means clustering Artificial neural network	N/A	Accuracy of 92.5%	N/A	KNN not specified for a particular metric may be contiguous

FIGURE 7: Graph showing frequency of diseases for test samples [35].

TABLE 5: Detection using PCA Technique.

Author and year	Methodology	Detected diseases	Remarks	Dataset
Gulhane and Kolekar [35]	Principal component analysisKNN	Blight Narcosis Alternaria Grey mildew Magnesium deficiency	Accuracy of 95%	110 test samples

Thingspeak server, examining and anticipating the gathered information utilizing the decision tree classifier, and building up the versatile application for ranchers for crop-related forecasts and cautions. Steps engaged with the decision tree classifier preparing information are an option of a certain level of all-out database alongside testing

information utilized for training the model. As a result, more proficiency of the model is obtained, which depends on the training data. Testing information is utilized to allocate names for the sort of infection in the decision tree classifier. The comparative analysis of various decision tree and random forest techniques and their results are shown in Table 6.

3.5. Disease Detection Analysis Using Fuzzy Logic Analysis.

To detect a common cotton disease in the plants, various parameters can be used, and a list of these parameters is shown in Table 7. Rothe and Kshirsagar [38] implemented the detection of cotton diseases such as Alternaria, Myrothecium, and bacterial blight. Initially, the picture was taken through the digital camera, and then, the implementation of preprocessing image methodology takes place to smoothen the image. The region-based segmentation algorithm and the neural network were utilized for the benchmark datasets' classification. Testing data estimate the overall accuracy of the system. Zhao et al. [39] displayed the picture acknowledgment after-effects of an example. The locales of the creepy-crawly nuisance of sugarcane cotton aphids have been separated viably to give the condition to the programmed ID and analyze bug irritation of sugarcane cotton aphids. The overall evaluation results have proven 85% segmented correctly. Zhang et al. [40] introduced the identification of cotton diseases with the assistance of a fuzzy inference system. In recognizing and diagnosing the cotton malady utilizing PC vision in farming, determination of proper methods influences the classifier's plan and execution.

The fuzzy component choice methodology, which incorporates fuzzy surfaces (FS) and fuzzy curves (FC) techniques, is currently the preferred choice for classifying cotton infected leaf images [41]. To get the best data for recognizing, a subclass of noteworthy autonomous highlights is distinguished, misusing the methodology of the fuzzy element. Firstly, use fuzzy curves to confine a little arrangement of critical highlights rapidly from the arrangement of unique highlights as per their need, take out the misleading element, and afterward utilize fuzzy surfaces to dispose of the highlights subject to the enormous highlights. Brewer and Glover [21] discussed the study that describes that root rot of cotton occurs in similar areas in the fields in repeating years even if there is a difference in the diseased pattern above these years. The authors of [42] present the capability of evaluation of the airborne multispectral detection of cotton plants. The comparative analysis of fuzzy system techniques is shown in Table 8.

3.6. Disease Detection Analysis Using Support Vector Machine.

Bhimte and Thool [44] initiated the work to identify leaf spot disease of cotton using SVM classification and image processing. The objective of this work confers the utilization of a straightforward picture handling perspective for the programmed analysis of the leaf diseases of the cotton plant. The

method of classification depends upon choosing the fitting features, for example, shading, and the surface of pictures is finished by using the SVM classifier. Different diseased leaf features dependent on the surface are misused by the Grey Level Coefficient Matrix (GLCM) for the identification of diseases. GLCM technique for learning the second request factual highlights the extraction framework for checking on the surface that thinks about the spatial relationship of pixel dim-level dispersal [45]. Recognizable proof of disease pursues the means such as staking the picture, differentiate upgrade, changing RGB to HSV, and extracting highlights [46]. The HSV algorithm for leaf disease can be determined [47].

Prashar et al. [48] presented the paper with a solution for the problem identified with farming, which includes recognizing leaf disease by utilizing the visible features. The multilayered perceptron (MLP) comes with overlapping pooling, an adaptable layering system to characterize the plant leaves to identify contaminated leaves. MLP is combined with ANN to forecast the content of chlorophyll from leaves [49]. For the overlapping arrangement of the layer to limit the curves by two-fold surfaced representation, KNN and SVM are used. Different procedures such as morphological division, design coordinating, and tint coordinating are consolidated which restrict the infection area with over 96% accuracy. Figure 8 shows the pictures of the cotton leaf at the beginning phase of improvement.

Prajapati et al. [50] proposed the strategy which utilizes the SVM Classifier. To distinguish the cotton leaf infections precisely, the need for machine learning and image processing is beneficial. The background images are handled through the thresholding process of Otsu. Distinctive partitioned images will help remove the highlights, such as color, shape, and surface, from the images. These removed highlights are needed as the contributions for the classifier. Various authors used SVM to classify the cotton disease, and the comparative analysis is shown in Table 9.

3.7. Disease Detection Analysis Using Neural Networks.

Shah and Jain [52] actualized the use of ANN to detect cotton leaf diseases. The layman approach to distinguish infection from the cotton plant was made by examining it outwardly, which brought about a high measure of mistake because of various visual observations along with lighting; however, with the assistance of the artificial neural system, it was simple to recognize the nature of the leaf of the cotton plant. For the testing, 18 leaf pictures were utilized to recognize 6 unique kinds of diseases. The outcomes were "1" and "0" for the leaf with disease and leaf without the disease, respectively. A comparison of ANN and the actual result is shown in Figure 9.

Pujari et al. [53] proposed the work where the center has been around the early recognizable proof of infectious contamination from visual signs. This supportive work, in actuality, involves errands such as image acquisition, image preprocessing, feature selection, and headway of approaches

TABLE 6: Comparison of decision tree and random forest techniques.

Author and year	Methodology	Detected diseases	Remarks	Dataset	Gaps identified
Mehta et al. [36]	Decision tree Random forest	N/A	For cotton disease prediction, RF 95.30% Decision tree 96.73% Multioutput regressor 89.61% Sensitivity 82.21%	30 images of size 1504×1000	MLP did not do well while classification
Chopda et al. [37]	Decision tree classification	Anthracnose Grey mildew Wilt	N/A	N/A	The model needs training

TABLE 7: Parameters for three common cotton diseases.

Features	Cotton diseases		
	Anthracnose	Black spot	Red leaf blight
Area	75.02	13.00	56.73
Major axis	12.69	4.31	9.42
Minor axis	8.01	4.31	6.15
Orientation	25.76	27.44	5.40
The ratio of principle axes	2.42	1.00	1.531
Equivalent diameter	10.14	4.07	7.233
Eccentricity	1.73	0.42	0.613
Solidity	1.02	1.00	0.943
Extent	0.71	0.52	0.65
Hydraulic radius	0.99	1.28	1.265
Complexity	13.00	9.85	12.00
Euler number	1.01	1.00	1.00
Moments of inertia xx	7.55	4.08	8.98
Moments of inertia yy	8.01	8.98	12.54

TABLE 8: Comparison of fuzzy system techniques.

Author and year	Methodology	Detected diseases	Remarks	Dataset
Rothe and Kshirsagar [38]	Adaptive neuro-fuzzy inference system Graph cut method	Bacterial blight Myrothecium Alternaria	Accuracy of 90%	N/A
Zhao et al. [39]	Rough set fuzzy C-means clustering	Cotton aphids	The accuracy rate of segmentation is 85%	20 images
Zhang et al. [40]	Fuzzy feature selection	Anthracnose Black spot Red leaf blight	N/A	150 images
Chengai Yang [43]	Fuzzy technique	Root rot	N/A	N/A

for the proof of transmissible disease manifestations that impact several farming or growing crops, leading towards the CVS design improvement. The correlation of image processing systems claimed to distinguish the parasitic disease arrangement and proof that has been influenced on various farming or cultivation crops is condensed. The database is made to store the yields of the element extractions. The database is utilized to monitor parasitic disease manifestations that influenced natural product crops, business harvests, and grain and vegetable crops. Kashinath Kamble [54] predicted that identifying the disease is the primary need of this system. Comparative analysis of neural network techniques is shown in Table 10.

3.8. Various Cotton Disease Segmentation and Classification Techniques. Many authors have used traditional and machine learning techniques for segmentation and classification purpose. In this paper, a comparative analysis has been performed for the better analysis of segmentation and classification techniques. This study is shown in Table 11.
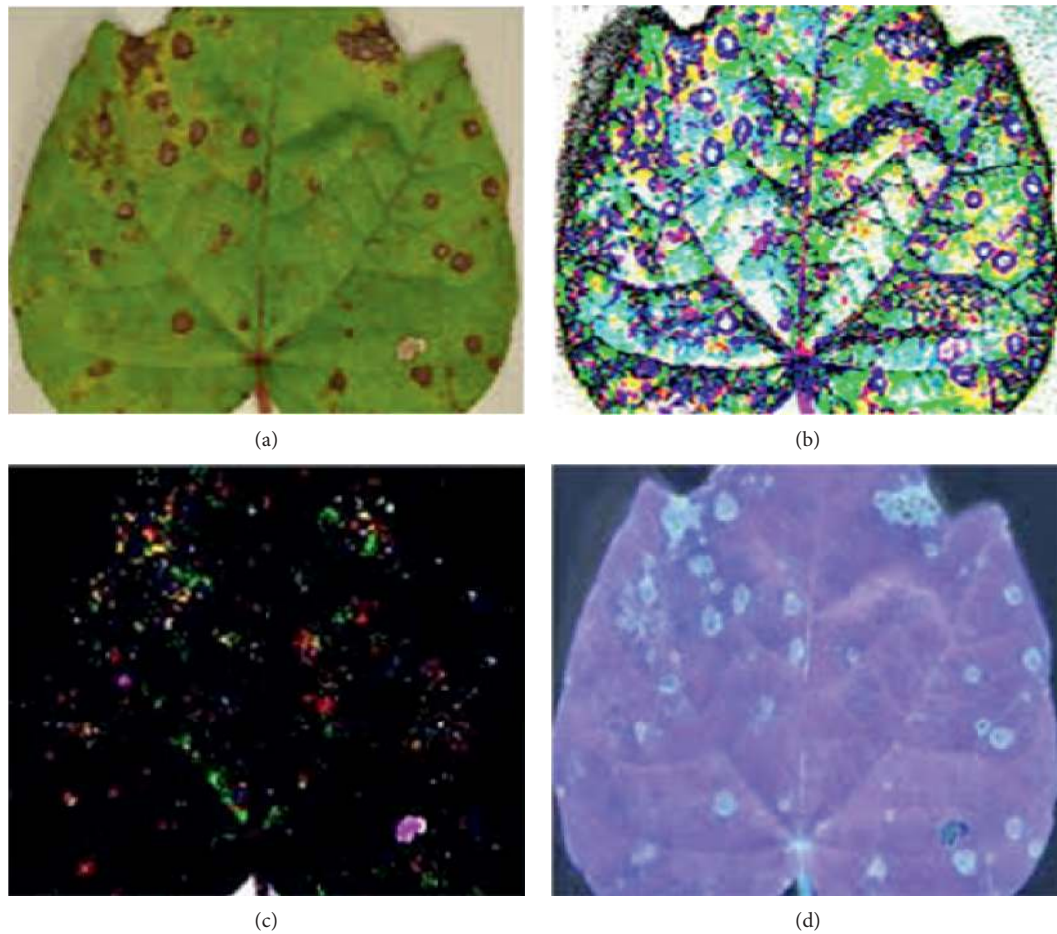
(a)

(b)

(c)

(d)

FIGURE 8: (a) Actual image. (b) Segmented image. (c) Colour variation. (d) Disease detected image.

TABLE 9: Comparison of support vector machine techniques.

Author and year	Methodology	Detected diseases	Remarks	Dataset
Bhimte and Thool [44]	Support vector machine classification	Bacterial blight Magnesium deficiency	Accuracy of 98.46%	130 images 100 infected images 30 noninfected images
Prashar et al. [48]	Support vector machine KNN	N/A	Accuracy of 96%	N/A
Vijaya Kishor et al. [51]	SVM tool classification	N/A	Accuracy of 96%	N/A
Prajapati et al. [50]	Support vector machine KNN Neural network	Alternaria Bacterial blight Cercospora	N/A	N/A
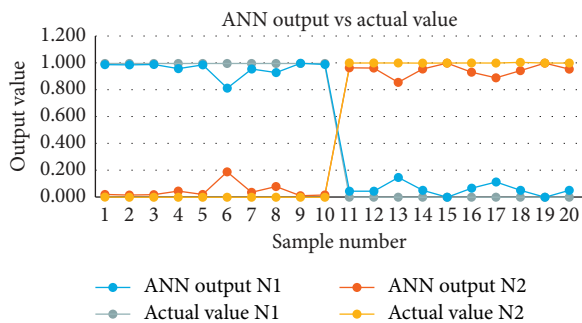
FIGURE 9: Comparison between actual output and ANN [52].

4. Discussion

4.1. Advantages and Disadvantages of Existing Techniques.
Tables 12 and 13 represent the advantages and disadvantages of existing techniques on segmentation and classification techniques.

The review on foundation end and segment methodology was moreover examined. Through this examination, they contemplated that the foundation expulsion shading space change from RGB to HSV is important and discovered that the thresholding framework gives extraordinary results that appeared differently concerning other foundation end

TABLE 10: Comparison of neural network techniques.

Author and year	Methodology	Detected diseases	Remarks	Dataset
Shah and Jain [52]	Artificial neural network	N/A	Relative error 0.05%	18 leaf images 6 images are detected
Pujari et al. [53]	PCA KNN ANN PNN SVM	Fungal diseases	Accuracy of PNN 86.48 ANN 84.11% KNN 91.54% SVM 85%	N/A
Kashinath Kamble [54]	Artificial neural network	N/A	N/A	750 images

TABLE 11: Comparative study of various segmentation and classification techniques.

S. no	Name of the author and year	Segmentation	Classification	Remarks	Dataset
2	Prashar et al. [48]	KNN	(i) Support vector machine	Accuracy 96% Precision = 91%	Self-database
3	Usha Kumari et al. [45]	K-means clustering	Artificial neural networks	Precision = 90% Recall = 80% Accuracy 92.5%	Self-database
5	Bhimte and Thool [44]	PCA	Support vector machine	Accuracy 98.46%	130 images
6	Masud et al. [55]	Image segmentation using Gaussian kernel function	—	Segmentation Accuracy = 63.99%	Self-database
7	Mehta et al. [36]	Decision tree	Random forest	RR = 95.30% Sensitivity = 82.12%	Self-database
12	Sarangdhar and Pawar [9]	Machine learning using regression IoT	—	Accuracy 83.26% Precision = 81% Recall = 79.1%	900 images of cotton leaves 629 are trained 271 are for testing
13	Vijaya Kishor et al. [51]	SVM tool classification	—	Accuracy 96%	Postgre SQL
14	Parikh et al. [32]	K-means clustering	SVM	Accuracy 82.5%	150 images 40 images of 1024 × 1024 pixels
17	Pujari et al. [53]	K-means clustering	ANN PNN SVM	Accuracy of ANN 84.11% PNN 86.48% SVM 85%	Self-database
21	Schuster et al. [34]	K-means clustering	Artificial neural network	Accuracy 88% F-score = 87.91%	Self-database

TABLE 12: Advantages and disadvantages of segmentation techniques.

Segmentation techniques	Advantage	Disadvantage
K-means clustering	If a colossal number of pictures are present in the dataset, then k-implies are significant for the division; it places near pixels in one group and assorted ones in the other group	Tedious; we need to choose which cluster gives a better outcome physically
Otsu thresholding	If one or more classes (closer view and establishment) are in the picture, then the Otsu method is fitting; moreover, it is found that Otsu produces better results that appeared differently concerning k-implies gathering for picture division	As a matter of course, the grey thrush capacity of MATLAB takes a limit estimation of 0.5; in any case, this worth may not be ideal for various situations, trouble in choice of edge esteem
Canny and Sobel	This method furnishes finer edge recognition, whereas the Sobel method gives precise corners and edges	For our dataset, canny edge discovery does not discover edges and corners; moreover, Sobel edge recognition does not function admirably when there are dainty and smooth lines in the images in the case of nitrogen deficiency

TABLE 13: Advantages and disadvantages of classification techniques.

Classification techniques	Advantages	Disadvantages
ANN	It can manage noisy data as the technique is self-adaptive	Neural networks are time-consuming due to the training and selection of such NN is difficult
SVM	SVM helps for linear as well as nonlinear classification; it is easy to grasp, and when compared to other classification methods, it produces precise results	Hard to choose the parameters for the kernel function
PCA	Sensitivity is less to noisy data; the need for memory is decreased	Linear separation of data cannot be performed; it is complex to obtain the data of the covariance matrix precisely

techniques. Thus, at long last, SVM gives extraordinary results, similar to precision, for the course of action of diseases.

4.2. Various Benchmark Datasets Used.
A dataset is an assortment of information. We as a whole perceive that to make up a machine learning approach; we need a dataset. In general, these machine learning datasets are used for analyzing the information. A dataset is the collection of homogeneous information. Dataset is utilized to mentor and worth the machine learning model. It has a major role in the system that has efficiency and reliability. If the database is commonplace and noise-free, it will give better outcomes. However, we are advanced with various datasets. It may frequently be business-related data or regular clinical data and much additional information. However, the actual problem is to find out the relevant ones according to the system requirements.

    (i) *MNIST Database.* The MNIST (Modified National Institute of Standards and Technology) information could be a huge dataset that is conventionally utilized for training various image processing systems

    (ii) *Postgre SQL.* Postgre SQL is an open-source object-social information-based framework with more than 30 years of dynamic improvement that has acquired a solid standing for dependability, power, and execution

4.3. Evaluation Parameters Used.
Different evaluation parameters are utilized to assess distinctive machine learning calculations. For instance, a classifier is used to recognize pictures of various items; we can utilize characterization parameters, for example, Logarithmic Loss, Precision, and AUC. Furthermore, if the machine learning model is attempting to predict a stock value, RMSE (root mean square error) can be utilized to compute the proficiency of the model. Another metric for assessing machine learning calculations is accuracy or NDCG, which can be utilized for arranging calculations utilized via web indexes. Along these lines, we see that various measurements are required to gauge the effectiveness of various calculations, additionally relying on the current dataset. Figure 10 shows the confusion matrix.

    (i) Confusion matrix
    (ii) ROC curve

The ROC curve (Receiver Operating Characteristics Curve) is another measurement to Figure 11. The introduction of a classifier model, the ROC curve, depicts the locus of certified positives' locus concerning the pace of genuine positives. This highlights the affectability of the classifier model. The ideal classifier will have a ROC where the outline would hit a certifiable positive pace of 100% with zero enhancements of x-center point (genuine positives). Since that is impossible when in doubt, we measure what numbers of good positive characterizations are being picked up with an increase in the rate of true positives. An example case of a ROC curve is referred to as follows.

Various parameters can be used to predict the outcome of different models, and a list of parameters is shown in Table 14. Most of the authors have used these parameters for the evaluating the performance of the algorithms. Table 15 gives a summarized view of theses parameters used by various authors.

4.4. Gaps and Findings.
After surveying the available research paper on a cotton disease, we have identified some of the gaps in the currently available work. There are some gaps and findings which are discussed below:

    (i) Some of the techniques are time-consuming by late identification of the disease, and the process for the detection may be lengthy in few cases [57–59].

    (ii) In most countries, no proper availability of detected disease management after the crop being affected affects the productivity of the cotton crops. It affects the economic growth of the country [55, 60].

    (iii) Most of the cotton disease detection methods need basic knowledge of operating the mobile application. In India, most of the farmers have lack of knowledge of such application and their usage.

    (iv) The available traditional and machine learning algorithms are giving accurate results, but it would be better if these algorithms could detect the diseases in the early stage [56, 61].

    (v) Researchers need more standard online datasets; therefore, research can be done at a faster pace.

    (vi) Most of the researchers applied the algorithms on standalone datasets; therefore, there is a need to build a model which can be used for multiple datasets.

		Real	
		Y	N
Expected	Y	TP	FP
	N	FN	TN

TN: true negative   TP: true positive

FP: false positive   Y = yes

FN: false negative   N = no

FIGURE 10: Structure of the confusion matrix.
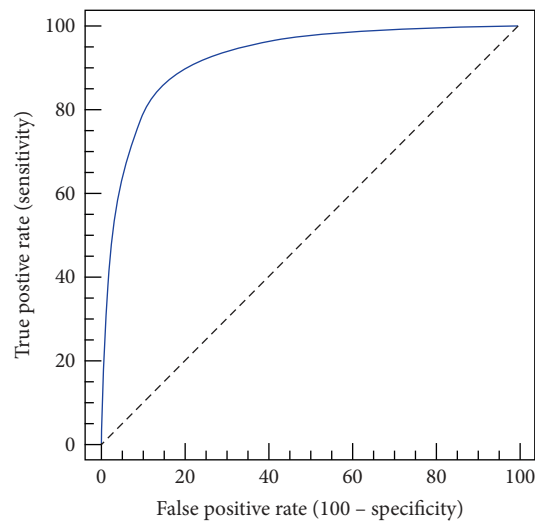
FIGURE 11: ROC curve.

TABLE 14: List of evaluation parameters.

S. no.	Parameter name	Formula
1	Accuracy	ACC = (TP + TN)/(TP + TN + FP + FN)
2	Precision	Precision = true positive/(true positives + false positives)
3	Recall	Recall = true positive/(true positives + false negatives)
4	Sensitivity	Sensitivity = true positive/(true positives + false negative)
5	Specificity	Specificity = true negative/(true negatives + false positives)
6	True positive rate	TPR = TP/(TP + TN)
7	True negative rate	TNR = TN/(TN + FP)
8	F-1	F1 = 2TP/(2TP + FP + FN)
9	Error rate	1 − accuracy

TABLE 15: Comparison of various parameters used by authors.

Reference no.	Accuracy	Precision	Recall	F-score	Sensitivity	Specificity	Self-data	Std. data
[48]	Y	Y	N	N	N	N	Y	N
[45]	Y	Y	Y	N	N	N	Y	N
[44]	Y	N	N	N	N	N	Y	N
[55]	Y	N	N	N	N	N	Y	N
[36]	Y	N	N	N	Y	N	Y	N
[9]	Y	Y	Y	N	N	N	N	Y
[51]	Y	N	N	N	N	N	N	Y
[32]	Y	N	N	N	N	N	Y	N
[53]	Y	N	N	N	N	N	Y	N
[56]	Y	N	N	Y	N	N	N	N

Here, N means no and Y means yes.

5. Conclusion

Accurate recognition and disease classification help us improve growth, leading to the development of several advanced techniques to be used in agriculture. Due to rapid growth in various diseases and adequate knowledge, identification and detection have become a major challenge. By monitoring the stages continuously, one can notice the disease with the naked eye, but the only drawback would be the time-consuming and increasing cost. It shows that the image processing had substantiated itself as a viable instrument for recognizable proof and arrangement of plant diseases, while the computerized camera capacities are viewed as a superior alternative for the human eye subbed by a learning calculation. To conquer the troubles of the procedure physically, a few strategies dependent on computer technology vision are created lately to perceive the disease and recognizable proof of agribusiness and agriculture crops. The quality analysis becomes difficult in captured images as they contain noise. Another restriction is that the images are to be captured under the environment in the conditions of lightning controlled. The methods of image processing execution in genuine use might be difficult because of a portion of the issues. To reduce the impact of light, point of the camera, catching gadget, and separation between objects, powerful adjustment is required. Different issues may emerge because of the changeability of shading under characteristic conditions. The absence of specialized information or image processing methods is critical for successfully executing any computer framework.

The greater parts of the strategies proposed by authors are equipped for managing just a single species group among a few diseases in the plants. Thus, to discover plant infections of various types, a global system has to be built. A significant number of the creators do not give their data of testing and processing information, which is the principal perspective during the assessment of the approval of results. Therefore, it is essential to give specific solid measures. For better sicknesses' discovery and their acknowledgment, the dataset of great pictures is suggested. The lighting arrangement should likewise be in the correct situation as it can influence the image captured. As provided by some of the authors of what tools they had used clearly, others are suggested to dedicate more time to the proper implementation of the techniques and thorough understanding of the tools to be used.

In the future, we will collect several images from enriching the cotton plant database and develop a deep neural network-based architecture to improve the accuracy. We will be developing a complete module to detect the diseases of cotton plants with a trained model, and information access will be through smart mobiles. It will also involve the recognition of cotton diseases based on different lands.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

References

[1] S. Satraj Sohrab, E. I. Azhar, M. A. Kamal, P. S. Bhattacharya, and D. Rana, "Genetic variability of cotton leaf curl betasatellite in Northern India," *Saudi Journal of Biological Sciences*, vol. 21, no. 6, pp. 626–631, 2014.

[2] O. Kulkarni, "Crop disease detection using deep learning," in *Proceedings of the 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, August 2018.

[3] S. Kumar, K. M. V. V. Prasad, A. Srilekha, T. Suman, B. Pranav Rao, and J. Naga Vamshi Krishna, "Leaf disease detection and classification based on machine learning," in *Proceedings of the 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*, October 2020.

[4] Y. Dong, F. Xu, L. Liu et al., "Monitoring and forecasting for disease and pest in crop based on web-GIS system," in *Proceedings of the 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, Istanbul, Turkey, September 2019.

[5] Z. Diao, H. Wang, Y. Song, and Y. Wang, "Image segmentation method for cotton mite disease based on colour features and area thresholding," *Journal of Theoretical and Applied Information Technology*, vol. 48, no. 1, pp. 527–533, 2013.

[6] H. Zhang, Y. Lan, C. P.-C. Suh et al., "Fusion of remotely sensed data from airborne and ground-based sensors to enhance detection of cotton plants," *Computers and Electronics in Agriculture*, vol. 93, pp. 55–59, 2013.

[7] L. A. Suarez, A. Apan, and J. Werth, "Hyperspectral sensing to detect the impact of herbicide drift on cotton growth and yield," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 120, pp. 65–76, 2016.

[8] V. Kiranmai, D. Ghai, and S. Kumar, "A review on classification of land use/land cover change assessment based on normalized difference vegetation index," *Journal of Critical Reviwes*, vol. 7, no. 14, pp. 2416–2431, 2020.

[9] A. A. Sarangdhar and V. R. Pawar, "Machine learning regression technique for cotton leaf disease detection and controlling using IOT," in *Proceedings of the International Conference on Electronics, Communication and Aerospace Technology (ICECA 2017)*, Coimbatore, India, April 2017.

[10] P. R. Rothe and R. V. Kshirsagar, "Cotton leaf disease identification using pattern recognition techniques," in *Proceedings of the International Conference on Pervasive Computing (ICPC)*, Pune, India, January 2015.

[11] P. Yang, Z.-X. Sun, S.-Y. Liu, H.-X. Lu, Y. Zhou, and M. Sun, "Combining antagonistic endophytic bacteria in different growth stages of cotton for control of verticillium wilt," *Crop Protection*, vol. 47, pp. 17–23, 2012.

[12] X. M. Suo, Y.-T. Jiang, M. Yang, S. K. Li, K.-R. Wang, and C.-T. Wang, "ANN to predict leaf population cholorophyll

content from cotton plant images," *Agricultural Sciences in China*, vol. 9, no. 1, pp. 38–45, 2009.

[13] V. C. Delfosse, M. F. Casse, Y. C. Agrofoglio et al., "Agro-inoculation of a full-length cDNA clone of cotton leafroll dwarf virus (CLRDV) results in systematic infection in cotton and the model plant Nicotiana benthamiana," *Virus Research*, vol. 175, no. 1, pp. 64–70, 2013.

[14] K. D. Bodhe, H. V. Taiwade, V. P. Yadav, and N. V. Aote, "Implementation of prototype for detection and diagnosis of cotton leaf diseases using rule-based system for farmers," in *Proceedings of the International Conference on Communication and Electronics Systems (ICCES 2018)*, Coimbatore, India, October 2018.

[15] Q. Zhao, C. Wang, G. Jin, and X. Yin, "Design and implementation of the pests and diseases information collection systems for cotton fields based on mobile terminals," in *Proceedings of the International Conference on Information Technology and Applications 2013*, Chengdu, China, November 2013.

[16] P. Chen, L. I. Shao-Kun, K.-R. Wang et al., "Spectrum character of cotton canopy infected with verticillium wilt and applications," *Agricultural Sciences in China*, vol. 7, no. 5, pp. 561–569, 2007.

[17] B. B. Kalbande and A. S. Patil, "Plant tissue culture independent agro bacterium tumefaciens medicated in-plant transformation strategy for upland cotton (*Gossypium hirsutum*)," *Journal of Genetic Engineering and Technology*, vol. 14, no. 1, pp. 9–18, 2016.

[18] R. W. Briddon and P. G. Markham, "Cotton leaf curl virus disease," *Virus Research*, vol. 71, no. 1-2, pp. 151–159, 2000.

[19] W. Zhang and Y. Du, "Rhizosphere fungal community structure succession of Xinjiang continuously cropped cotton," *Fungal Biology*, vol. 123, pp. 42–50, 2018.

[20] S. Sun, C. Li, A. H. Paterson, P. W. Chee, and J. S. Robertson, "Image processing algorithms for infield single cotton boll counting and yield prediction," *Computers and Electronics in Agriculture*, vol. 166, Article ID 104976, 2019.

[21] M. J. Brewer and J. P. Glover, "Boll injury caused by leaf footed bug in late-season cotton," *Crop Protection*, vol. 119, pp. 214–218, 2019.

[22] H. E. Dao-Fang, X. Zhao, L. Cheng-Zhen et al., "Genetic variation in LBL1 contributes to depth of leafblades lobes between cotton subspecies, *Gossypium hirsutum* and *Gossypium barbadense*," *Journal of Integrative Agriculture*, vol. 17, no. 11, pp. 2394–2404, 2018.

[23] J. H. Hancock, J. B. Wilkerson, F. H. Moody, and M. A. Newman, "Seed specified placement of in-furrow fungicides for control of seedling disease in cotton," *Crop Protection*, vol. 23, no. 9, pp. 789–794, 2003.

[24] L. Amrao, I. Amin, M. S. Shahid, R. W. Briddon, and S. Mansoor, "Cotton leaf curl disease in resistant cotton is associated with a single begomovirus that lacks an intact transcriptional activator protein," *Virus Research*, vol. 152, no. 1-2, pp. 153–163, 2010.

[25] S. Pandurang Gawande, K. P. Raghavendra, D. Monga, D. T. Nagrale, and S. Kranthi, "Rapid detection of tobacco steak virus (TSV) in cotton (gissopyium hirustum) based on reverse transcription loop medicated isothermal amplification (RTLAMP)," *Journal of Virological Methods*, vol. 270, 2019.

[26] Y. K. Dubey, M. M. Mushrif, and S. Tiple, "Super pixel based roughness measure for cotton leaf diseases detection and classification," in *Proceedings of the Fourth International Conference on Recent Advances in Information and Technology (RAIT-2018)*, Dhanbad, India, March 2018.

[27] P. R. Rothe and R. V. Kshirsagar, "Automated extraction of digital images features of three kinds of cotton leaf diseases," in *Proceedings of the International Conference on Electronics, Communication and Computational Engineering (ICECCE)*, Hosur, India, November 2014.

[28] P. Revathi and M. Hemalatha, "Classification of cotton leaf spot diseases using image processing edge detection techniques," in *Proceedings of the 2012-International Conference on Emerging Trends in Science, Engineering, and Technology*, Tiruchirappalli, India, December 2012.

[29] S. Kirthi Pilli, B. Nallathambi, S. J. George, and V. Diwanji, "eAGROBOT—a robot for early crop disease detection using image processing," in *Proceedings of the Second International Conference on Electronics and Communication Systems (ICECS-2015)*, Coimbatore, India, February 2014.

[30] S. Chaudhary, M. Bhise, A. Banerjee, A. Goyal, and C. Moradiya, "Agro advisory system for cotton crop," in *Proceedings of the AGRINETS Workshop, COMSNETS-2015*, Bangalore, India, January 2015.

[31] P. Revathi and M. Hemalatha, "Advance computing enrichment evaluation of cotton leaf spot disease detection using image edge detection," in *Proceedings of the 2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*, Coimbatore, India, July 2012.

[32] A. Parikh, M. S. Raval, C. Parmar, and S. Chaudhary, "Disease detection and severity estimation in cotton plant from unconstrained images," in *Proceedings of the 2016 IEEE International Conference on Data Science and Advanced Analytics*, Montreal, Canada, October 2016.

[33] R. P. Narmadha and G. Arulvadivu, "Detection and measurement of paddy leaf disease symptoms using image processing," in *Proceedings of the International Conference on Computer Communication and Informatics (ICCCI-2017)*, Coimbatore, India, January 2017.

[34] E. W. Schuster, S. Kumar, S. E. Sarma, J. L. Willers, and G. A. Milliken, "Infrastructure for data driven agriculture: identifying management zones for cotton using statistical modeling and machine learning techniques," in *Proceedings of the 2011 8th International Conference & Expo on Emerging Technologies for a Smarter World*, Hauppauge, NY, USA, November 2011.

[35] V. A. Gulhane and M. H. Kolekar, "Diagnosis of diseases on cotton leaves using PCA classifier," in *Proceedings of the 2014 Annual IEEE India Conference (INDICON)*, Pune, India, December 2014.

[36] V. Mehta, C. Jain, K. Kanchan, and V. Sawant, "A machine learning approach to foretell the probability of a contracting a disease," in *Proceedings of the 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, August 2018.

[37] J. Chopda, H. Raveshiya, S. Nakum, and V. Nakrani, "Cotton crop disease detection using decision tree classifier," in *Proceedings of the 2018 International Conference on Smart City and Emerging Technology (ICSCET)*, Mumbai, India, January 2018.

[38] P. R. Rothe and R. V. Kshirsagar, "Adaptive neuro-fuzzy inference system for recognition of cotton leaf diseases," in *Proceedings of the International Conference on Innovative Applications of Computational Intelligence on Power, Energy, and Controls with their Impact on Humanity (CIPECH-14)*, Ghaziabad, India, November 2014.

[39] J. Zhao, M. Liu, and M. Yao, "Study on image recognition of insect pest of sugarcane cotton aphids based on roughset fuzzy

C-means clustering," in *Proceedings of the Third International Symposium on Intelligent Information Technology Applications*, Nanchang, China, November 2009.

[40] Y.-C. Zhang, H.-P. Mao, Bo hu, and M.-X. Li, "Features selection of cotton disease leaf images based on fuzzy feature selection techniques," in *Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition*, Beijing, China, November 2007.

[41] H. Tang, J. Ye, L. Zhou, and S. Zhou, "Agriculture disease diagnosis expert system based on knowledge and fuzzy reasoning: a case study of flower," in *Proceedings of the 2009 A Sixth International Conference on Fuzzy Systems and Knowledge Discovery*, Tianjin, China, August 2009.

[42] A. Raheem, N. Esmaeili, M. O'Connell, and J. Zhang, "Progress and perspective on drought and salt stress tolerance in cotton," *Industrial Crops and Production*, vol. 130, pp. 118–129, 2018.

[43] C. Yang, G. N. Odvody, J. A. Thomasson, T. Isakeit, and R. L. Nichols, "Change detection of cotton root rot infection over 10-year intervals using airborne multispectral imagery," *Computers and Electronics in Agriculture*, vol. 123, pp. 154–162, 2016.

[44] N. R. Bhimte and V. R. Thool, "Disease detection of cotton leaf spot using image processing and SVM classifier," in *Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018)*, Madurai, India, June 2018.

[45] C. H. Usha Kumari, S. Jeevan Prasad, and G. Mounika, "Leaf disease detection: feature extraction with K-means clustering and classification with ANN," in *Proceedings of the Third International Conference on Computing Methodologies and Communications (ICCMC 2019)*, Erode, India, March 2019.

[46] H. Krishnan, K. Priyadarshini, M. Gowsic et al., "Plant disease analysis using image processing in MATLAB," in *Proceedings of the International Conference on Systems Computation Automation and Networking 2019*, Pondicherry, India, March 2019.

[47] G. Karthik and N. Prabhuram, "Detection and prevention of banana leaf diseases from banana plant using embedded linux board," in *Proceedings of the 2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, Coimbatore, India, November 2016.

[48] K. Prashar, R. Talwar, and C. Kant, "CNN based on overlapping pooling method and multi-layered learning with SVM and KNN for American cotton leaf diseases recognition," in *Proceedings of the International Conference on Automation, Computational and Technology Management (ICACTM)*, London, UK, April 2019.

[49] W. D. Yan, W. M. Shi, B. H. Li, and M. Zhang, "Overexpression of a foreign Bt gene in cotton affects the low molecular weight components in root exudates," *Pedosphere*, vol. 17, no. 3, pp. 324–330, 2007.

[50] B. S. Prajapati, V. K. Dabhi, and H. B. Prajapati, "A survey on detection and classification of cotton leaf diseases," in *Proceedings of the International Conference on Electrical, Electronics and Optimization Techniques (ICEEOT 2016)*, Chennai, India, March 2016.

[51] R. Vijaya Kishor, K. Pooja Shatrughan, M. Komal Balasaheb et al., "Agromet expert system for cotton and soybean crops in regional area," in *Proceedings of the International Conference on Advances in Communication and Computing Technology (ICACCT 2018)*, Sangamner, India, February 2018.

[52] N. Shah and S. Jain, "Detection of disease in cotton leaf using artificial neural network," in *Proceedings of the 2019 Amity International Conference on Artificial Intelligence (AICAI)*, Dubai, UAE, February 2019.

[53] J. D. Pujari, R. Yakkundimath, and A. S. Byadgi, "Identification and classification of fungal diseases affected on agriculture/horticulture crops using image processing techniques," in *Proceedings of the 2014 IEEE International Conference on Computational Intelligence and Computing Research*, Coimbatore, India, December 2014.

[54] J. Kashinath Kamble, "Plant disease detector," in *Proceedings of the 2018 International Conference on Advances in Communication and Computing Technology (ICACCT)*, Sangamner, India, February 2018.

[55] M. Masud, N. Sikder, A.-A. Nahid, A. K. Bairagi, and M. A. Alzain, "A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework," *Sensors*, vol. 21, no. 3, p. 748, 2021.

[56] B. Gupta, M. Tiwari, and S. Singh Lamba, "Visibility improvement and mass segmentation of mammogram images using quantile separated histogram equalisation with local contrast enhancement," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 2, pp. 73–79, 2019.

[57] N. Gianchandani, A. Jaiswal, D. Singh, V. Kumar, and M. Kaur, "Rapid COVID-19 diagnosis using ensemble deep transfer learning models from chest radiographic images," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–13, 2020.

[58] S. Ghosh, P. Shivakumara, P. Roy, U. Pal, and T. Lu, "Graphology based handwritten character analysis for human behaviour identification," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 55–65, 2020.

[59] M. D. Kamrul Hasan, S. Ahmed, Z. M. Ekram Abdullah et al., "Deep learning approaches for detecting pneumonia in COVID-19 patients by analyzing chest X-ray images," *Mathematical Problems in Engineering*, vol. 2021, Article ID 9929274, 8 pages, 2021.

[60] H. S. Basavegowda and G. Dagnew, "Deep learning approach for microarray cancer data classification," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 22–33, 2020.

[61] D. Singh, V. Kumar, V. Yadav, and M. Kaur, "Deep neural network-based screening model for COVID-19-infected patients using chest X-ray images," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 35, no. 3, Article ID 2151004, 2021.