International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 03 | Mar 2022 www.irjet.net p-ISSN: 2395-0072

Literature Review on Classification of Cotton Species

[1]Annapoorna B R, [2]P Roopasree, [3]Shriya J, [4]Roopa Shree S P

[1]Professor, Dept.of Computer Science engineering, Dayananda Sagar College of Engineering, Karnataka, India
[2]Student, Dept.of Computer Science engineering, Dayananda Sagar College of Engineering, Karnataka, India

---

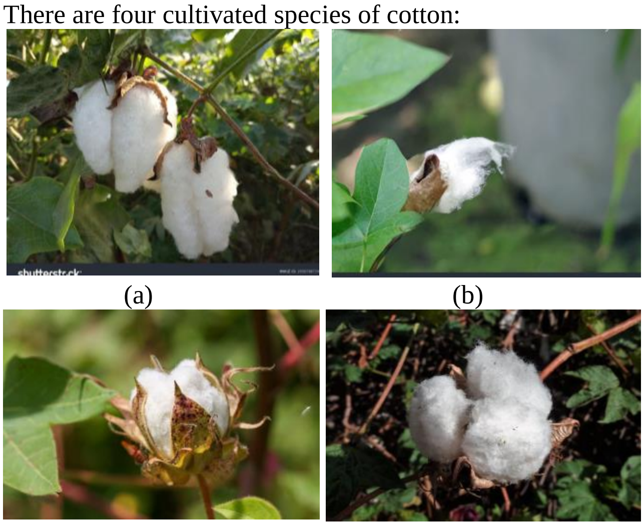
Abstract - *Species identification with traditional keys is difficult, time-consuming, and confusing for non-experts owing to the usage of specific botanical words. For beginners interested in learning about species, this poses a difficult barrier to overcome. There is a growing interest in automating species identification procedures worldwide. The availability and widespread use of key technologies such as digital cameras and mobile devices, and new image processing and pattern recognition algorithms have made the notion of automated species identification a reality. This paper provides the recent methods used to detect and classify cotton species. In the proposed study, we analyze several object recognition, image processing, and classification approaches utilized to accomplish this task by taking into account the outcomes obtained from each method. The ultimate objective is to have the finest technique available currently. This study carefully examines and compares these strategies and procedures.*

Key Words*: *Image processing, Classification, Object detection, Cotton Species.

1. INTRODUCTION

Cotton is the most widely grown plant fiber crop in the world, yet numerous components of the plant are used. Following harvest, the fluffy white fiber known as lint is collected from the seed and processed into yarn, which is woven into fabrics. The seed is used to make hulls, meat, and oil for feeding the livestock. It is a globally important agricultural product, with more than 60 producer countries and yearly sales of around US$ 10 billion. As a result, technical advancements that result in increased productivity may have a significant economic and societal impact. The development of genetically better cultivars, in particular, is a crucial factor in enhancing plantation yield. Increased herbicide and insect resistance, as well as improved germination, strength, and cotton fiber qualities, are all important features of the United States, India, Uzbekistan, China, Turkey, Brazil, Egypt, and Greece. These countries account for roughly 85% of world cotton output. India is the world's second-largest cotton producer. In India, many states produce cotton: Gujrat, Maharashtra, Karnataka, Andhra Pradesh, Telangana. Gossypium is the most significant crop for producing vegetable oil and pollen fiber, as well as high-quality cattle feed.

There are four cultivated species of cotton:



(a) (b)

(c) (d)

Figure 1: (a) Gossypium herbaceum. (b) Gossypium barbadense (c) Gossypium hirsutum (d) Gossypium arboreum

India is the only country that produces all four species of cotton. Here morphological differences have been taken into consideration for identifying the above species of the cotton plant. Botanists use various plant traits as identification keys in a manual identification process, which is evaluated sequentially and adaptively to identify plant species. Cotton species identification by field observation involves extensive botanical knowledge, putting it out of reach for the majority of nature enthusiasts. Traditional cotton species identification is nearly impossible for the general public and difficult even for experts that deal with botanical challenges daily, such as environmentalists, farmers, foresters. Even for botanists, species identification is generally challenging. Image-based approaches for species identification are considered to be promising. A user can snap a photo of a plant in the field with the built-in camera of a smartphone and analyze it with an installed identification model to identify the species or, if a single match is unavailable, to obtain a list of probable species. Non-professionals can participate in this process by using a plant identification system. As a result, a huge number of studies are devoted to automating the process of plant species identification.

2. METHODOLOGIES

The following are some of the most crucial processes in object detection and image processing techniques:

1. **Obtaining images:** The images/videos are captured using a simple RGB camera and then transformed into the proper formats using specific software. The photos you take should be of high quality. As a result, using an HD camera is highly advised.
2. **Image segmentation and preprocessing:** Once the photos or videos have been collected, they must be preprocessed, which involves things like creating bounding boxes, image augmentation, turning a video into a collection of images, filtering the images as needed, adding noise to them, and so on. For practically all image processing and object detection approaches, these are the most popular. This stage also involves a process called Segmentation, which highlights the required sections of the image for accuracy.
3. **Extraction of features:** In this step, an algorithm is used to extract the relevant features of the photos or videos that are required for future analysis. This step increases the final result's precision. As a result, depending on the type of object detection, it's critical to carefully select features and algorithms. Because cotton is an agricultural product, the characteristics may vary in size, color, texture, and other factors. Each crop or species we evaluate will have its own set of requirements.
4. **Classification**: The items or parts of the object detected are sorted into distinct classes or categories depending on requirements in the last step of the procedure. The YOLO approach may be used to detect objects and machine learning algorithms such as Support Vector Machines (SVM). Artificial Neural Networks (ANN), and others can be used to classify them.

3. MOTIVATION

Agriculture is critical to the development of most developing countries. It is also critical to India's economic development. Cotton is grown in 95 percent of the agricultural land in most north and south Indian states. When each species is handled correctly with the appropriate insecticides and manure, productivity improves. For higher cotton crop productivity, it is necessary to identify and classify these species at an early stage. In light of this, this review paper compiled a list of numerous classification approaches and object detection techniques. This research will be beneficial to future research in this sector.

4. LITERATURE SURVEY

[1] Salman Qadri, Najia SaherTanveer Aslam, "A Spatial Model of K-Nearest Neighbours for Classification of Cotton (Gossypium) Varieties based on Image Segmentation".Vol.5 issue1, Jan-March 2021

This study described the classification of four cotton varieties using multiple features leaves dataset and a methodology that uses a machine learning (ML) approach to classify four (4) different cotton leaf varieties namely; Z-32, BS-15, Z-31, and S-32. Here four different machine learning techniques were employed successfully namely: Naive Bayes Tree, K, Random Forest Tree, and K-Nearest Neighbor (K-NN). Vision and Image Processing (CVIP) tools were deployed for feature extraction. The feature optimization was carried out using the Correlation-Based Feature Selection (CFS) genetic algorithm technique. The Transductive Parameter Transfer (TPT) approach was used for image segmentation and classification. This approach was evaluated in three phases:

Phase 1, the ROI selection procedure is employed for image segmentation.

Phase 2, employed the correlation feature selection (CFS) genetic search algorithm for feature reduction or feature optimization procedure.

Phase 3, these optimized feature datasets employed different machine learning classifiers for output or result. Different methodologies were applied in the description of the image dataset for the classification of the cotton leaves but the K-NN's resultant accuracy was better as compared to the other machine learning classifiers' results by achieving 98.2583% accuracy on the four cotton varieties. The achieved overall accuracy results evaluate that the proposed technique is efficient and can be employed in real-life applications.

[2] Joseph Christian Nouaze, Guoxu Liu, Jae Ho Kim, Philippe Lyonel Touko Mbouembe," YOLO-Tomato: A Robust Algorithm for TomatoDetection Based on YOLOv3".

YOLOTomato is a powerful tomato recognition system based on YOLOv3. YOLOv3's rich architecture facilitates feature reuse for a more compact and accurate model. In addition, this approach replaces the traditional rectangular bounding box with a circular bounding box for tomato localization. The new bounding box matches well with tomatoes and improves the calculation of non-maximum suppression cross-linking (IOU). This article presents a DCNN-based detection model for detecting tomatoes in complex environments. Two major strategies have been proposed to improve cognitive performance. First, the YOLOv3 model added a high-density architecture to simplify feature reuse and support model capture of more features for tomato representation. Second, CBbox has been proposed as an alternative to box for tomato localization.

[3] Ojus Thomas Lee, Gargi Chandrababu," Identification of Plant Species using Deep Learning". IJERT, NCREIS - 2021 Conference Proceedings

This research focuses on plant species identification, which is a critical task due to the huge variety of plant species, necessitating the development of an automatic plant identification system. This research focuses on several deep learning-based methodologies that are used for plant species identification. Methods based on deep learning technologies such as convolutional neural networks are used to extract features and classify leaf photos. Classification techniques such as Artificial Neural Networks, Support Vector Machine, Decision Tree, and k-Nearest Neighbor are used in state-of-the-art procedures. The following are the main steps in the plant recognition process:

Features Extraction: Important stage in recognizing plant species. Global and local features are different types of feature extraction. The term "global features" refers to the depiction of features such asThe general content of the image is described by the color, texture, and shape. Local features are the parts of an image that are located within a certain location.

Classification: After all of the collected features have been concatenated into a feature vector, it is then categorized.

The suggested technique achieved 90.54 and 99.6 percent accuracy, on the LeafSnap dataset and the Flavia dataset.

[4] Prabhpreet Kaur, Surleen Kaur, "Plant Species Identification based on Plant Leaf Using Computer Vision and Machine Learning Techniques ", Journal of Multimedia Information System VOL. 6, NO. 2, June 2019 (pp. 49-60): ISSN 2383-7632(Online)

This study presents a method for automatically identifying plant species by classifying plant leaf pictures using machine learning and computer vision techniques. The research was divided into four stages: image segmentation, picture pre-processing, feature extraction, and image classification. Photographs from a Swedish leaf dataset containing 1,125 images of 15 *different species were used. After that, a Gaussian filtering mechanism was used to pre-process the image, and then a combination of texture and color features was retrieved. Finally, a Multiclass-support vector machine was used to classify the data. The system had an average accuracy of 93.26 percent on a Swedish dataset. The program was able to classify 15 different plant species automatically. In comparison to techniques that focus solely on morphological shape features, texture and color feature space fared admirably. In addition, as compared to PNN and k-NN, SVM performed significantly better as a classifier. The method described is both easy to use and powerful. Although the model attained an accuracy of more than 90%, it still falls short of solutions that use deep learning or neural networks techniques. Here, the aim is to make automatic plant species identification more real by working with a live dataset.*

[5] Vishwanath Burkpalli, Bhagya M. Patil, "A Perspective View of Cotton Leaf Image Classification Using Machine Learning Algorithms Using WEKA" Hindawi Advances in Human-Computer Interaction Volume 2021, Article ID 9367778, 15 pages

This research proposes a method for the classification of leaf diseases as,In the sphere of agriculture, this is a crucial responsibility. The disease recognition helps the farmer in determining what additional safeguards can be taken. The categorization can be done using a variety of machine learning techniques, and it's done with the cotton leaf database. The segmentation is done as the cotton photos are taken from the field, with a complicated background. Color and texture features are extracted from the segmented output images afterward. Color signals are sufficient for discriminating between good and dangerous images, according to this study. The WEKA tool, which aids in the analysis of several classifiers, uses the same attributes.

Four color features, eight texture features, and twelve (texture and color) features are compared. Color characteristics, it can be shown, are sufficient for enhancing classification accuracy. According to the survey, artificial neural networks outperform traditional classifiers such as AdaBoost, Random Forest, KNN, SVM, and Naive Bayes in terms of accuracy. Because color features can help discriminate between good and diseased leaves, it may be decided that extracting multiple texture descriptors is unnecessary. In the future, the research might be expanded to incorporate disease classification using textural cues or a deep learning system.

[6] Fabián R. Jiménez López, Wolfgang D. Niño Pacheco, "Tomato classification according to organoleptic maturity (coloration) using machine learning algorithms K-NN, MLP,
and K-Means Clustering", IEEE International Conference conducted on 24-26 April 2019 at Bucaramanga, Colombia.

This post describes the design, development, implementation, development, and design evaluation of various machine learning methods for classifying Milan and Chongto tomatoes based on physical parameters such as color (ripeness). First, capture the image using the Embedded Board Raspberry Pi 2 Model B Fiber Optic Camera Module. The distance between the camera and the fruit was kept constant for image capture, and the background of the scene was white to aid in image processing. The tomato image database is built through fieldwork in the greenhouse. I took a total of 600 shots of Milan Ichiban and Chongto tomatoes of various sizes. Tomato photos were digitally processed with OpenCV. 3.1.0 software platform and Python
V.2.7.13 artificial vision library. After the photo is captured in RGB format, it is converted to grayscale to simplify image preprocessing. The photo was then cleaned and scaled-down on the filter stage. Unsupervised learning algorithms such as MLP, KKNN (Multilayer Perceptron) type neural networks, and KMeans are some of the classification techniques used and evaluated here. The algorithm is evaluated using a global confusion matrix and performance metrics such as accuracy, accuracy, sensitivity, and specificity.

[7] Rohit Raja, Shilpa Rani, G. Harshitha, Mohammed A. AlZain, d Mehedi Masud, Sandeep Kumar, Arpit Jain, Anand Prakash Shukla, Satyendra Singh, "A Comparative Analysis of Machine Learning Algorithms for Detection of Organic and Nonorganic Cotton Diseases".

To better apprehend the segmentation and class approaches, this paper describes comparative analysis. Learn approximately the diverse illnesses that natural and inorganic cotton is at risk of and the quality device gaining knowledge of algorithms for each. Leaf spot or black arm ailment is an indexed ailment. [ANN and principal component analysis are the best detection methods.]. An avascular wilt ailment is an ailment that impacts arteries and veins. [Best detection edge detection method]. Botrytis cinerea is another call for Botrytis cinerea. [Fuzzy function] and [Best recognition decision tree method]. Root rot is an ailment that impacts the roots of plants. [Neural network is the best detection method]. Leaf spots, or rot, are illnesses that affect plants. [Optimal recognition method Image processing and support vector machine] Describes diverse class algorithms generally utilized in device gaining knowledge of and deep gaining knowledge of additionally mentions the 4 foremost sorts of cotton generally utilized in our work.

[8] Liping Di, Ziheng Sun, Annie Burgess, Senior Member, IEEE, Hui Fang, "Deep Learning Classification for Crop Types in North Dakota"[2020].

This study presented a classification method for producing high-quality in-season crop maps from Satellite pictures for North Dakota using a deep neural network. As training labels, ground gathered datasets, NASS reports, CDL and regional crop maps are employed, and Landsat 8 images are used as model inputs. Crop field boundaries are obtained by digitizing high-resolution photos. The majority of the ground data is gathered here through field surveys and roadside picture samples. Landsat pictures are processed into input batches and agricultural map images are used as output masks, in addition to the standard preparation operations such as spatial augmentation, calibration, and image enhancement. Then converting the machine learning outputs back to agricultural maps, which involves several post-processing steps. A soft-max function is used to the neural network's output, which converts values into a probability matrix reflecting the likelihood of the input object's predicted class. Finally, it was discovered that a deep neural network is better at detecting large farmlands than it is at recognizing wetlands and residential areas. The model that has been trained on various situations over many years and months has better accuracy. These findings show that deep neural networks can provide valid in-season maps for main crops on large farms, as well as a generally accurate reference for minor crops in scattered wetland areas.

5. CONCLUSIONS

Image processing and object identification have a wide range of real-world applications and can significantly increase productivity when used appropriately. The articles mentioned in the preceding section provide us with vital knowledge about the many outcomes that can be obtained as well as the most practical strategy that can be used in our suggested project. We employ the Decision Tree (DT), k-Nearest Neighbor (k-NN), and Support Vector Machine (SVM), however, the speed may suffer because it is slower than other methods such as Faster R-CNN, YOLO, and so on. When all of the characteristics in this scenario are considered, we can infer that using YOLO-v5 (the most recent version of YOLO) for object detection will yield better results. It has a higher chance of being useful in the actual world and producing accurate findings. As a result, YOLO-v5 with an R-CNN-based classification method will be more efficient than any other technique now available for object detection and species categorization.

REFERENCES

[1]Tanveer Aslam', Salman Qadri *, Najia Saher' "A Spatial Model of K-Nearest Neighbours for Classification of Cotton (Gossypium) Varieties based on Image Segmentation".Vol.5 issue1, Jan-March 2021.

[2]Guoxu Liu 1,2 , Joseph Christian Nouaze 2, Philippe Lyonel Touko Mbouembe 2andJae Ho Kim 2,*

[3]Gargi Chandrababu, Ojus Thomas Lee, Ojus Thomas Lee," Identification of Plant Species using Deep Learning". IJERT, NCREIS - 2021 Conference Proceedings.

[4]Surleen Kaur1*, Prabhpreet Kaur2, ``Fiber Species Identification based on Plant Leaf Using Computer Vision and Machine Learning Techniques ", Journal of Multimedia Information System VOL. 6, NO. 2, June 2019 (pp. 49-60): ISSN 2383-7632(Online)

[5]Bhagya M. Patil, Vishwanath Burkpalli, "A Perspective View of Cotton Leaf Image Classification Using Machine Learning Algorithms Using WEKA" Hindawi Advances in Human-Computer Interaction Volume 2021, Article ID 9367778, 15 pages

[6]Wolfgang D. Niño Pacheco, Fabián R. Jiménez López, "Tomato classification according to organoleptic maturity (coloration) using machine learning algorithms K- NN, MLP, and K-Means Clustering", IEEE International Conference conducted on 24-26 April 2019 at Bucaramanga, Colombia.

[7]Sandeep Kumar, Arpit Jain, Anand Prakash Shukla, Satyendra Singh, Rohit Raja, Shilpa Rani, G. Harshitha, Mohammed A. AlZain, d Mehedi Masud, "A Comparative Analysis of Machine Learning Algorithms for Detection of Organic and Nonorganic Cotton Diseases

[8]Ziheng Sun, Liping Di, Senior Member, IEEE, Hui Fang, and Annie Burgess, "Deep Learning Classification for Crop Types in North Dakota"[2020]

[9]A shill, MD. Asif Rahman, "Plant Disease Detection Based on YOLOv3 and YOLOv4"[2021]

[10]Kenneth LI-Minn Ang, Jasmine Kah Phooi Seng, "Big Data and Machine Learning With Hyperspectral Information in Agriculture"[2021].

[11]Showkat Ahmad Bhat, Nen-Fu Huang, "Big Data and AI Revolution in Precision Agriculture: Survey and Challenges.", [2021].

[12]Adhao Asmita Sarandhar, Prof . Dr. V.R. Pawaer, "Machine Learning Regression Technique for Cotton Leaf Disease Detection and Controlling using IoT", [2017].

[13]Sandeep Kumar, Manish Kashyap, Akashdeep Saluja, Mahua Bhattacharya, "Segmentation of Cotton Bolls by Efficient Feature Selection Using Conventional Fuzzy C Means Algorithm With Perception of Colour".

[14]Ravali Koppaka, Teng-Sheng Moh(2020)," Machine Learning in Indian Crop Classification of Temporal Multi-Spectral Satellite Image".

[15]Kshyanaprava Panda Panigrahi, Himansu Das, Abhaya Kumar Sahoo, and Suresh Chandra Moharana," Maize Leaf Disease Detection and Classification Using Machine Learning Algorithms".