iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: VI Month of publication: June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83467

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Soybean Leaf Disease Detection Using Convolutional Neural Network

Aarti Pandurang Chavare[1], Ajit B Patil[2], Tanvi Gopinath Patil[3], Dipti Prakash Swami[4], Manasi Dipak Shirale[5]
Department of Computer Science and EngineeringKolhapur Institute of Technology's College of Engineering, Kolhapur

Abstract: Soybean is one of the most important agricultural cropsworldwideandservesasamajorsourceoffood,edible oil, and animal feed. The productivity and quality of soybean cropsaresignificantlyaffectedbyvariousleafdiseasessuch as Caterpillar damage, Rust, Bacterial Blight, and Diabrotica Speciosa. These diseases reduce crop yield and negatively impact agricultural production and the economic condition of farmers. Therefore, early and accurate disease identification is essential foreffectivecropmanagementandimprovedproductivity.Tradi-tionaldiseasedetectionmethodsmainlyrelyonmanualinspection by agricultural experts, which is time-consuming, costly, less efficient for large-scale farming, and prone to human error.

This paper presents an automated soybean leaf disease detec-tionsystemusingimageprocessinganddeeplearningtechniques. The proposed system utilizes Convolutional Neural Networks (CNN) for automatic feature extraction and classification of healthyanddiseasedsoybeanleaves.ARandomForestclassi-fierisalsoimplementedforcomparativeanalysistoevaluate the performance of traditional machine learning approaches.The collected dataset undergoes preprocessing steps such as image resizing, normalization, filtering, and data augmentation techniques including rotation, zooming, and flipping to improve model performance and reduce overfitting.

Experimental results demonstrate that the CNN-based model achieves higher accuracy, precision, recall, and F1-score com-paredtoRandomForestmethods.Theproposedsystemprovides an efficient, reliable, and cost-effective solution for real-time soybean leaf disease detection. It can support smart agriculture systems and assist farmers in taking timely preventive measures to minimize crop damage and improve overall agricultural productivity.

Index Terms: Soybean, Leaf Disease Detection, Deep Learn-ing, CNN, Image Processing, Softmax, Random Forest

I. INTRODUCTION

Agriculture plays a vital role in the economy of many countries, especially India, where a large population depends on farming for their livelihood. Soybean is one of the most importantcommercialandoilseedcropsduetoitshighprotein content, nutritional value, and industrial applications. It is widelycultivatedinagriculturalregionssuchasMaharash-tra, Madhya Pradesh, and Karnataka. Soybean contributes significantly to food production, edible oil industries, and animal feed. However, soybean crops are highly affected by various foliar diseases such as Rust, Bacterial Blight, Cater-pillar damage, and Diabrotica Speciosa, which reduce crop qualityandproductivity.Thesediseasescanspreadrapidly under unfavorable environmental conditions and cause severe economic losses to farmers if not detected at an early stage.

Traditional disease detection methods mainly rely on man-ual visual inspection by farmers or agricultural experts. These methods are time-consuming, labor-intensive, subjective, and prone to human error, especially in large agricultural fields. Delayed or inaccurate disease identification can lead to severe crop loss and reduced agricultural productivity. In many rural areas, access to agricultural experts is limited, making timely disease diagnosis difficult.

Recent advancements in Artificial Intelligence (AI), image processing,anddeeplearninghaveprovidedefficientsolutions for automated plant disease detection. Computer vision tech-niquescombinedwithConvolutionalNeuralNetworks(CNNs) can automatically identify disease patterns from leaf images withhighaccuracy.Thesetechniquesreducehumaneffortand providefasterandmorereliableresultscomparedtotraditional approaches.

The main objective of this research is to develop an auto-mated soybean leaf disease detection system using CNN and image processing techniques. The study also aims to compare CNN performance with Random Forest methods and improve diseaseclassificationaccuracyforreal-timeagriculturalappli-cations.

II. LITERATURE REVIEW

Severalstudieshavebeenconductedonplantdiseasedetec-tionusingimageprocessing,machinelearning,anddeeplearn-ing techniques.

Traditional image processing methods used color, texture, and shape features for disease identification,but their performance was affected by lighting conditions and image quality [10], [19].
MachinelearningtechniquessuchasRandomForestand LogisticRegressionimprovedclassificationaccuracybutre-quired manual feature extraction [2], [3]. These methods were less effective for complex disease patterns and large datasets. Recent deep learning approaches, especially Convolutional NeuralNetworks(CNNs),haveshownbetterperformance inplantdiseasedetection. Modelssuchas VGG16,ResNet, MobileNet, and YOLOv5 achieved high accuracy in soybean leafdiseaseclassification [1],[4],[8]. Imageaugmentation andtransferlearningtechniquesfurtherimprovedmodelper-formance[5],[12].
Although previous studies achieved high accuracy, some challengesstillremain,suchasdifficultyindetectingsimi-lar disease symptoms, high computational requirements, and lack of real-time implementation. The proposed system aimsto overcome these limitations using an efficient CNN-based approach for accurate soybean leaf disease detection.

III. METHODOLOGY

Theproposedsystemconsistsofthefollowingsteps:

A. Data Collection

A dataset of soybean leaf images is collected containing both healthy and diseased samples. The dataset includes im-ages capturedunder varyingenvironmental conditionssuch as lighting, background, and leaf orientation to improve model generalization.

B. Dataset Description

The dataset used in this research contains soybean leaf images collected from online agricultural repositories and publicly available datasets. The dataset includes both healthy and diseased soybean leaf samples.
Theimagesbelongtothefollowingclasses:

- Healthy
- Rust
- BacterialBlight
- CaterpillarDamageSoftmax
- DiabroticaSpeciosa

Σ

The dataset contains images captured under different light-ing conditions, backgrounds, and leaf orientations to improve model robustness.

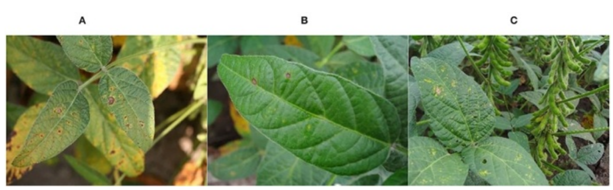


Fig.1.SampleImagesofHealthyandDiseasedSoybeanLeaves

C. Preprocessing

The collected images are resized to a fixed dimension (e.g., 224 ×224) to maintain uniformity. Pixel values are normal-ized to improve training efficiency. Noise is reduced using filtering techniques to enhance image quality. Additionally, data augmentation methods such as rotation, flipping, and zooming are applied to increase dataset diversity and prevent overfitting.

D. FeatureExtraction

The CNN model automatically extracts important features from soybean leaf images without manual intervention. These features include color variations, texture patterns, and shape details that help distinguish between healthy and diseased leaves.
During this process, convolutional layers scan the image using filters to detect edges, spots, and disease-specific pat-terns. As the network goes deeper, it learns more complexand meaningful features, which are then used for accurate classification.

E. Model Architecture

A Convolutional Neural Network (CNN) is used for clas-sification of soybean leaf diseases. The model consists of the following layers:

- InputLayer

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

- ConvolutionLayers
- ReLUActivation
- MaxPoolingLayer
- FullyConnectedLayer
- SoftmaxOutputLayer

The convolution layers extract important features from the images, while the ReLU activation function introduces non-linearity.Themaxpoolinglayerreducesthespatialdimensions and helps in feature selection. The fully connected layer processes the extracted features for final classification.

TheSoftmax classifierisused intheoutput layertoclassify the images into multiple classes. It converts the output values into probability scores.

Softmax Function:

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}}$$

This function converts the output into a probability distri-bution and is used for multi-class classification.

F. Working of CNN

Convolutional Neural Networks (CNNs) are deep learning models mainly used for image classification tasks. CNN au-tomatically extracts important features from images without manual feature engineering.

Theconvolutionlayerappliesfilterstotheinputimage to identify important patterns such as edges, textures, and spots. The ReLU activation function introduces non-linearity into the model. Max pooling reduces image dimensions and computational complexity. Finally, fully connected layers and the Softmax classifier perform final disease classification.

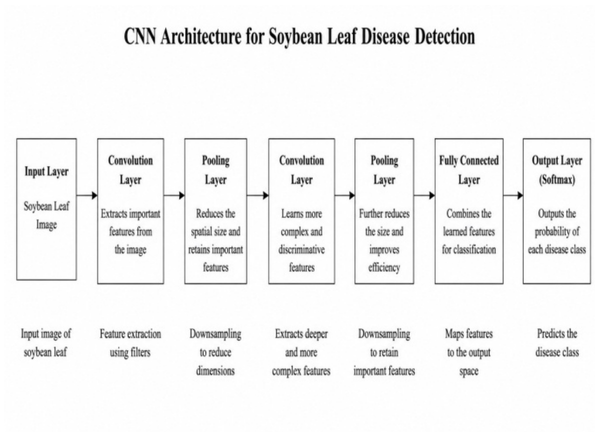


Fig.2.ProposedCNNArchitectureforSoybeanLeafDiseaseDetection

G. Training

The model is trained using labeled data with cross-entropy loss and optimized using Adam optimizer.

IV. SYSTEM ARCHITECTURE

Theproposedsoybeanleafdiseasedetectionsystemconsists of multiple stages including image acquisition, preprocessing, feature extraction, CNN-based classification, and result pre-diction.

Initially, soybean leaf images are collected from the dataset and preprocessed using resizing, normalization, and augmen-tationtechniques.Theprocessedimagesarethenpassedtothe CNN model for feature extraction and disease classification. Finally, the predicted disease result is displayed through the web-based interface.

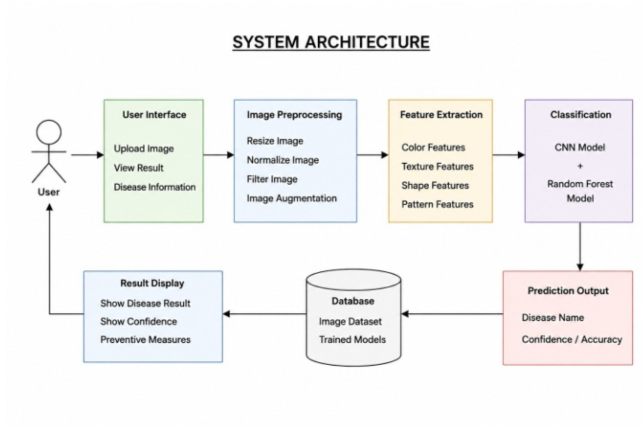


Fig.3.SystemArchitectureofProposedModel

V. ALGORITHM

- Step1:Collectsoybeanleafimagedatasetcontaining healthy and diseased leaf samples.
- Step2:Performpreprocessingtechniquessuchasimage resizing, normalization, and noise removal.
- Step3:Applydataaugmentationmethodsincludingrota-tion, flipping, and zooming.
- Step4:Splitthedatasetintotrainingandtestingdatasets.
- Step5:TraintheCNNmodelusinglabeledsoybeanleaf images.
- Step6: Extract image features automatically using convo-lution layers.
- Step7:PerformclassificationusingtheSoftmaxoutput layer.
- Step 8: Evaluate model performance using accuracy, preci-sion, recall, and F1-score.

VI. RESULTS AND DISCUSSION

Theproposedmodelachieveshighclassificationaccuracy. Experimental results show:
Accuracy:97%–99%

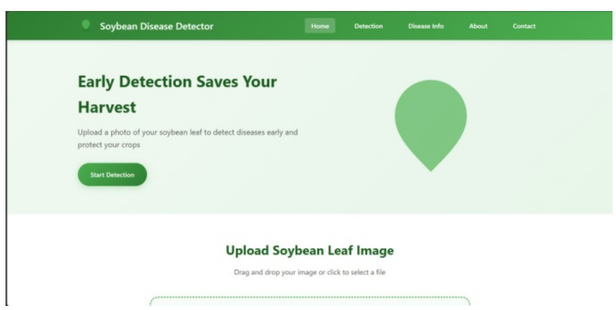


Fig.4.HomepageofSoybeanLeafDiseaseDetectionSystem

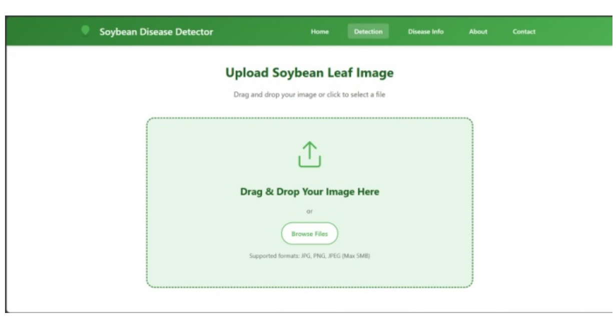


Fig.5.ImageUploadModule

Precision:High
Recall:High
Thesystemperformswellevenwithvariationsinlighting and background conditions.

TABLEI
PERFORMANCECOMPARISONOFMODELS

Model	Accuracy	Precision	Recall	F1Score
CNN(Softmax)	96–98%	High	High	High
RandomForest	85–88%	Moderate	Moderate	Moderate

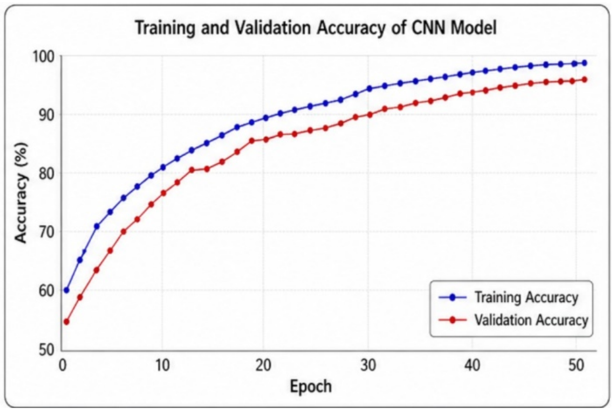


Fig.6.TrainingandValidationAccuracyofCNNModel


```
for i, category in enumerate(categories):
    print(f"{category}: {prediction[0][i]*100:.2f}%")
```

```
Caterpillar: 0.43%
Diabrotica speciosa: 99.57%
Healthy: 0.00%
```

Fig.7.PredictedOutputResult

Observations:

- CNNperformsbetterduetoautomaticfeatureextraction.
- RandomForeststruggleswithcomplexpatterns.

- CNNhandlesimagevariationsmoreeffectively.

CNNmodelshaveshownaccuracyupto99%insomeplant disease datasets.

VII. ADVANTAGES

The proposed system offers several benefits in soybean leaf disease detection:

- Automateddetectionofsoybeanleafdiseases
- Highclassificationaccuracy
- Real-timepredictioncapability
- Time-efficientandreducesmanualeffort
- Usefulforfarmersandagriculturalexperts

WhyCNNisBetter:

- Automaticallylearnsspatialfeaturesfromimages
- Eliminatestheneedformanualfeatureextraction
- Effectivelyhandlesnoiseandvariationsininputdata

LimitationsofRandomForest:

- Requiresmanualfeatureextraction
- Limitedperformancewhenapplieddirectlytoimagedata
- Requireslargedatasetfortraining
- Performancedependsonimagequality
- Highcomputationalrequirementsduringtraining

ApplicabilityofRandomForest:

- Suitableforsmalldatasets
- Requireslowercomputationalresources

VIII. CONCLUSION

This paper presents an automated soybean leaf disease detectionsystemusingConvolutionalNeuralNetworks(CNN) and image processing techniques. The proposed model effec-tivelyidentifieshealthyanddiseasedsoybeanleaveswithhigh accuracy,precision, andreliability.Experimental resultsshow that the CNN-based approach performs better than traditional machine learning methods such as Random Forest due to its automatic feature extraction capability.

Thesystemreducesmanualeffortandsupportsearlydisease detection,whichcanhelpfarmerstaketimelypreventivemea-sures and reduce crop loss. The proposed method is efficient, cost-effective, and suitable for real-time smart agriculture applications. In the future, the system can be integrated with mobile and IoT-based platforms for practical field usage.

IX. FUTURE WORK

Future improvements of the proposed soybean leaf disease detection system include:

- Integration with mobile applications for easy access and real-time disease diagnosis by farmers
- Real-timediseasemonitoringusingIoTandsmartagri-culture devices
- Incorporationofvideo-baseddiseasedetectionforcon-tinuous crop monitoring
- Extensionofthemodeltodetectdiseasesinmultiplecrop species
- Useofcloud-basedstorageandmonitoringsystemsfor large-scale agricultural analysis
- Improvementofmodelaccuracyusinglargerandmore diverse datasets
- Developmentofmultilingualsupportsystemsforfarmers from different regions

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

REFERENCES

[1] M.A.Jahin,S.Shahriar,M.F.Mridha,J.Hossen,andN.Dey,"SoybeanDisease Detection via Interpretable Hybrid CNN-GNN: IntegratingMobileNetV2 and GraphSAGE with Cross-Modal Attention," arXivpreprint, Mar. 2025.

[2] S.Kaur,A.Pandey,andS.Goel,"Semi-automaticleafdiseasedetectionand classification system for soybean," IET Image Processing, vol. 12,no. 5, pp. 1038–1048, 2021.

[3] R.Bhavani,"DetectionofLeafDiseasesinSoybeanPlantusingAutoen-coderandMultinomialLogisticRegression,"LegumeResearch,vol.48,no. 5, pp. 876–884, 2025.

[4] X.Peng,L.Wang,andY.Zhang,"EnhancedDetectionofSoybeanLeafDiseases Using an Improved YOLOv5 Model," International Journal ofKnowledge and Innovation Studies, vol. 2, no. 1, pp. 45–56, 2024.

[5] "Deep Learning Model for Classifying and Evaluating Soybean LeafDisease Damage," International Journal of Molecular Sciences, vol. 25,no. 1, pp. 106–118, 2024.

[6] A. Mohanty, D. Hughes, and M. Salathe´, "Using Deep Learning forImage-Based Plant Disease Detection," Frontiers in Plant Science, vol.7, pp. 1419–1425, 2016.

[7] S.Sladojevic,M.Arsenovic,A.Anderla,D.Culibrk,andD.Stefanovic,"Deep Neural Networks Based Recognition of Plant Diseases by LeafImage Classification," Computational Intelligence and Neuroscience,vol. 2016, pp. 1–11, 2016.

[8] K. P. Ferentinos, "Deep Learning Models for Plant Disease DetectionandDiagnosis,"ComputersandElectronicsinAgriculture,vol.145,pp.311–318, 2018.

[9] J.G.A.Barbedo,"FactorsInfluencingtheUseofDeepLearningfor Plant Disease Recognition," Biosystems Engineering, vol. 172, pp.84–91, 2018.

[10] P.Sharma,Y.Hans,andS.Gupta,"ClassificationofPlantLeafDiseasesUsing Machine Learning and Image Processing," International Journalof Engineering Research and Technology, vol. 9, no. 6, pp. 234–240,2020.

[11] R. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep Learning forTomato Diseases: Classification and Symptoms Visualization," AppliedArtificial Intelligence, vol. 31, no. 4, pp. 299–315, 2017.

[12] S. Too, L. Yujian, S. Njuki, and Y. Yingchun, "A Comparative StudyofFine-TuningDeepLearningModelsforPlantDiseaseIdentification,"Computers and Electronics in Agriculture, vol. 161, pp. 272–279, 2019.

[13] M. Picon, A. Alvarez-Gila, A. Seitz, et al., "Deep Convolutional NeuralNetworks for Mobile Capture Device-Based Crop Disease Classifica-tion," Agronomy, vol. 9, no. 1, pp. 3–15, 2019.

[14] H. Durmus, E. Gunes, and M. Kirci, "Disease Detection on the Leavesof the Tomato Plants by Using Deep Learning," IEEE InternationalConference on Agro-Geoinformatics, pp. 1–5, 2017.

[15] S. Ramesh, D. Vydeki, and K. Venkataraman, "Plant Disease DetectionUsing Machine Learning," International Conference on Design Innova-tions for 3Cs Compute Communicate Control, pp. 41–45, 2018.

[16] A. Ramcharan, K. Baranowski, P. McCloskey, et al., "Deep LearningforImage-BasedCassavaDiseaseDetection,"FrontiersinPlantScience,vol. 8, pp. 1852–1860, 2017.

[17] T. Rahman, M. K. Hasan, and M. S. Islam, "CNN-Based Plant LeafDisease Detection Using Image Processing," Journal of Artificial Intel-ligence and Capsule Networks, vol. 2, no. 1, pp. 5–14, 2020.

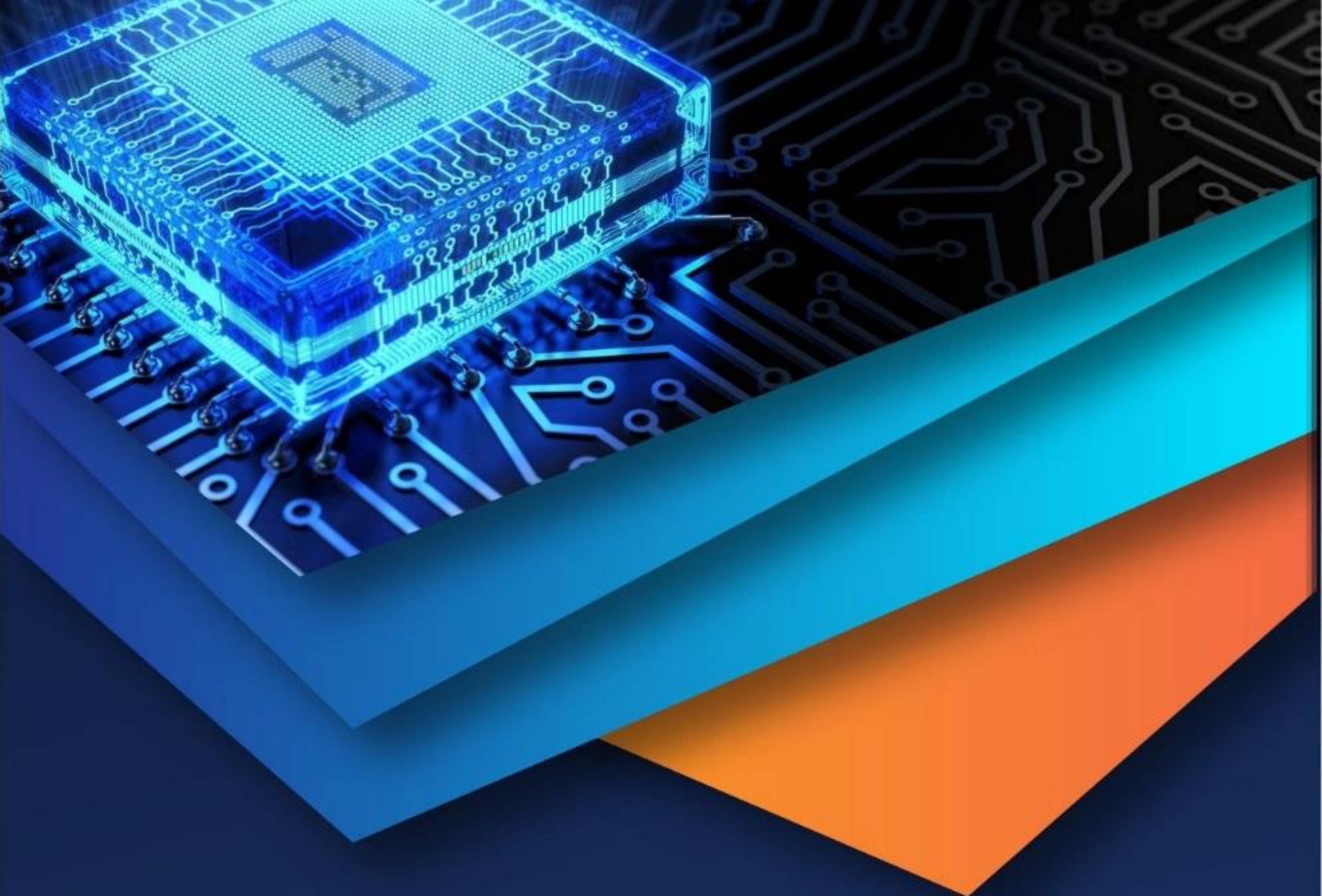
[18] N. P. Patel and D. Shah, "Soybean Leaf Disease Detection UsingConvolutional Neural Networks," International Journal of ComputerApplications, vol. 183, no. 12, pp. 20–25, 2021.

[19] V. Singh and A. Misra, "Detection of Plant Leaf Diseases Using ImageSegmentationandSoftComputingTechniques,"InformationProcessingin Agriculture, vol. 4, no. 1, pp. 41–49, 2017.

[20] S. Pawar and M. Kokare, "Leaf Disease Detection Using FeatureExtraction and Neural Networks," International Journal of AdvancedComputer Research, vol. 5, no. 2, pp. 115–120, 2019.

[21] R. Sujatha and Y. Chatterjee, "A Deep Learning-Based Approach forPlant Disease Detection," Journal of Physics: Conference Series, vol.1717, no. 1, pp. 1–8, 2021.

[22] P. Revathi and M. Hemalatha, "Classification of Cotton Leaf SpotDiseases Using Image Processing Edge Detection Techniques," IEEEInternational Conference on Emerging Trends in Science, Engineeringand Technology, pp. 169–173, 2012.



10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY