A COMPARATIVE STUDY AND ANALYSIS OF MULTIMODAL DEEP LEARNING FOR EARLY DETECTION AND SEVERITY ASSESSMENT OF SUGARCANE LEAF DISEASES

Elanchezhian D[1], Vasanthakumar S[2]

Abstract

The early detection of sugarcane leaf disease is important, as it helps to increase the yield and quality of the sugar crop and farm productivity. Current image-based disease detection approaches primarily diagnose apparent symptoms and typically exclude environmental aspects that are linked to disease development. In this paper, a multimodal deep learning framework for detecting and assessing the severity of sugarcane leaves is proposed, dubbed CaneSense-MS. EfficientNetV2-S is used for extracting the visual features from the sugarcane leaf images, and a bidirectional long short-term memory network is used for the processing of the temperature, humidity, rainfall, and soil-moisture readings. The cross-modal attention mechanism is used to fuse visual and environmental representations by highlighting the domain-specific elements from both modalities. The proposed multitask learning module classifies disease into its category and its severity level (mild, moderate, or severe) and attempts to estimate the percentage of the leaf area that is affected. Grad-CAM identifies the regions of the image that are important for the prediction, while SHAP identifies the environmental variables that are important for the prediction. The knowledge distillation technique is utilized to learn a lightweight student model for deployment on smartphones and edge devices. The framework is assessed based on classification accuracy, macro F1-score, AUC-ROC, severity MAE, and inference latency and offers an interpretable and practical approach for the early management of sugarcane diseases in agricultural fields.

Keywords : Sugarcane Disease Detection, Multimodal Learning, Environmental Sensing, EfficientNetV2-S, BiLSTM, Severity Classification, Explainable AI, Edge Computing.

Research Scholar
Department of Computer Science[1]
Karpagam Academy of Higher Education[1]
Coimbatore, Tamil Nadu, India[1]
elan15592@gmail.com[1]

Assistant Professor[2]
Department of Computer Science[2]
Karpagam Academy of Higher Education[2]
Coimbatore, Tamil Nadu, India[2]
vasanth.chml@gmail.com[2]

* Corresponding Author

I. INTRODUCTION

Sugarcane is an important commercial crop that is utilized for sugar, ethanol, biofuel, and various industrial products. It is highly susceptible to leaf diseases including brown rust, smut, grassy shoot, yellow leaf disease, pokkah boeng and brown spot. When left untreated, these diseases lower the quality of the crop, sugar yield, and recovery. The traditional diagnosis is primarily through manual field examination and is a time-consuming and expert-based process. The same type of symptoms can also cause misdiagnoses when a patient suffers from more than one disease. Thus, it is necessary to have an automatic and precise disease detection system to facilitate timely crop monitoring and treatment.

This research aims to enhance the diagnosis of sugarcane diseases using the image of sugarcane leaves and environmental data. Current image-based models may not work if diseases have similar symptoms or if images captured in the field do not have good lighting or complex backgrounds. Other environmental factors, like temperature, humidity, rainfall, and soil moisture, can give further clues to disease development. Most existing techniques are also only able to classify the disease without quantifying the disease severity. Therefore, a practical system should be able to predict accurately, explain, and be severity-aware. Furthermore, lightweight deployment is needed to enable real-time usage on smartphones and edge devices.

Previous studies on sugarcane disease detection are primarily based on convolutional neural networks for disease classification using images of sugarcane leaves. Thite et al. released a dataset consisting of 6,748 leaf images of sugarcane from various diseased and healthy classes [1]. Daphal and Koli built a deep learning model and created an application for practical disease recognition on mobile devices [2]. To increase the rate of classification accuracy while decreasing the computational complexity, EfficientNet-based models are tested [3]. To get more disease images, sugarcane GAN was proposed for enhancing the diversity of the training set [4]. Effective classification of sugarcane diseases was also achieved by other CNN and deep neural network techniques [5, 6]. In addition, lightweight MobileNetV3 models were used to support uncertainty-aware and resource-efficient deployment in the field [7].

In this paper, the authors present CaneSense-MS, a multimodal deep learning framework for detection of

sugarcane diseases and severity classification. The features of the sugarcane leaf images related to disease are extracted using EfficientNetV2-S. Environmental data, including temperature, humidity, rainfall, and soil moisture, are processed in a BiLSTM network. Cross-modal attention integrates the information from the visual and environmental modalities and prioritizes the most relevant information from both sources. The model can predict the category of the disease, the severity of the disease, and the percentage of leaf area affected. The explainable predictions and the lightweight deployment to smartphones or edge devices are achieved by Grad-CAM, SHAP, and knowledge distillation.

The main contributions of this paper are,

- A multimodal approach is created using the sugarcane leaf images along with the temperature, humidity, rainfall, and soil-moisture data.
- Visual and environmental features are extracted using EfficientNetV2-S and BiLSTM, respectively.
- In order to enhance the fusion of both data types, cross-modal attention is introduced.
- The model predicts disease class, disease severity level, and percentage leaf area affected simultaneously.
- The Grad-CAM and SHAP are used to get interpretable disease predictions.
- Knowledge distillation is employed to obtain a light model on smartphones and edge devices.

This paper continues as follows: In Section II, the authors summarize the existing research on sugarcane disease detection, multimodal learning, severity assessment, and lightweight deployment. The proposed CaneSense-MS framework and major components are explained in Section III. The data used, data preprocessing techniques, experimental setup, and assessment criteria are detailed in Section IV. The experimental results and performance comparisons are covered in Section V. Lastly, Section VI summarizes the paper and suggests future research directions.

II. LITERATURE SURVEY

A. Explainable Sugarcane Disease Detection

El-Behery et al. proposed explainable few-shot learning systems for disease detection in rice and sugarcane [8]. They have developed a method that overcomes this issue by using convolutional feature extraction and interpretation techniques to boost the transparency of the predictions, even when the dataset is small and has limited labeled samples. The study demonstrated that explainability plays an important role in establishing whether the model is able to concentrate on the real disease symptoms or irrelevant background areas.

B. Sugarcane Disease Severity Assessment

Wang et al. [9] suggested the intelligent grading method of the severity of sugarcane leaf disease by physiological indicators and machine learning. The method was an extension of traditional classifications of diseases and uses an estimation of infection level. Their results indicated that severity grading can offer greater assistance in treatment selection, crop monitoring, and field-level disease management.

C. Multimodal Crop Disease Diagnosis

Lee et al. combined the crop images with the environment variables like temperature and humidity to diagnose the diseases [10]. Their findings showed that with weak or similar visual symptoms, environmental information can help predict. Lu et al. used a multimodal Transformer for agricultural disease detection and question answering [11]. In the attention mechanism, the interaction between visual and contextual information was improved in the sense that the former was better integrated with the latter.

D. Cross-Modal Feature Fusion

To facilitate the diagnosis of plant diseases, Wu et al. proposed a graph-based multimodal semantic interaction framework (PlantIF) [12]. The approach captured associations between different types of features and enhanced information sharing between modalities. Liu et al. suggested a vision-language cross-modal fusion approach to crop disease recognition [13]. Their work verified that they can learn complementary information from cross-modal learning that goes beyond diagnosis with only pictures.

E. Disease Severity Estimation and Interpretation

Nasir et al. built an interpretable multimodal model that can be used to diagnose and estimate the severity of tomato diseases [14]. The framework jointly executed disease classification, disease severity prediction, and explanation. Zhao et al. introduced segmentation networks to automatically estimate the severity of plant diseases in natural fields [15]. The approach identified the infected areas and quantified the infected leaf area to determine the severity of the disease with greater accuracy.

F. Research Gap

The studies reviewed have verified the effectiveness of explainable learning, cross-modal fusion, multimodal diagnosis, and severity estimation. The current approaches are mainly directed toward one or two of these aspects. There is a limited amount of research that integrates images of sugarcane leaves, environmental parameters, cross-modal attention, disease classification, disease severity assessment, explainability, and lightweight edge deployment in an integrated framework. CaneSense-MS is a proposed model by the researcher that is intended to fill this void.

III. PROPOSED WORK

The CaneSense-MS framework aims to identify cane leaf diseases and quantify the disease severity with multimodal data. It is a combination of sugarcane leaf images and environmental parameters like temperature, humidity, rainfall, and soil moisture. The image branch is used to obtain features of the image related to the disease, and the environmental branch is used to learn the conditions in which the disease develops. Cross-modal attention fuses the most salient information from both avenues. The fused features are utilized to recognize the disease category, to determine severity of the disease as mild, moderate, or severe, and to estimate percentage of leaf area affected. In addition, explainable predictions and lightweight deployment are enabled by adding Grad-CAM, SHAP, and

- Visual and environmental features are extracted using EfficientNetV2-S and BiLSTM, respectively.
- In order to enhance the fusion of both data types, cross-modal attention is introduced.
- The model predicts disease class, disease severity level, and percentage leaf area affected simultaneously.
- The Grad-CAM and SHAP are used to get interpretable disease predictions.
- Knowledge distillation is employed to obtain a light model on smartphones and edge devices.

This paper continues as follows: In Section II, the authors summarize the existing research on sugarcane disease detection, multimodal learning, severity assessment, and lightweight deployment. The proposed CaneSense-MS framework and major components are explained in Section III. The data used, data preprocessing techniques, experimental setup, and assessment criteria are detailed in Section IV. The experimental results and performance comparisons are covered in Section V. Lastly, Section VI summarizes the paper and suggests future research directions.

II. LITERATURE SURVEY

A. Explainable Sugarcane Disease Detection

El-Behery et al. proposed explainable few-shot learning systems for disease detection in rice and sugarcane [8]. They have developed a method that overcomes this issue by using convolutional feature extraction and interpretation techniques to boost the transparency of the predictions, even when the dataset is small and has limited labeled samples. The study demonstrated that explainability plays an important role in establishing whether the model is able to concentrate on the real disease symptoms or irrelevant background areas.

B. Sugarcane Disease Severity Assessment

Wang et al. [9] suggested the intelligent grading method of the severity of sugarcane leaf disease by physiological indicators and machine learning. The method was an extension of traditional classifications of diseases and uses an estimation of infection level. Their results indicated that severity grading can offer greater assistance in treatment selection, crop monitoring, and field-level disease management.

C. Multimodal Crop Disease Diagnosis

Lee et al. combined the crop images with the environment variables like temperature and humidity to diagnose the diseases [10]. Their findings showed that with weak or similar visual symptoms, environmental information can help predict. Lu et al. used a multimodal Transformer for agricultural disease detection and question answering [11]. In the attention mechanism, the interaction between visual and contextual information was improved in the sense that the former was better integrated with the latter.

D. Cross-Modal Feature Fusion

To facilitate the diagnosis of plant diseases, Wu et al. proposed a graph-based multimodal semantic interaction framework (PlantIF) [12]. The approach captured associations between different types of features and enhanced information sharing between modalities. Liu et al. suggested a vision-language cross-modal fusion approach to crop disease recognition [13]. Their work verified that they can learn complementary information from cross-modal learning that goes beyond diagnosis with only pictures.

E. Disease Severity Estimation and Interpretation

Nasir et al. built an interpretable multimodal model that can be used to diagnose and estimate the severity of tomato diseases [14]. The framework jointly executed disease classification, disease severity prediction, and explanation. Zhao et al. introduced segmentation networks to automatically estimate the severity of plant diseases in natural fields [15]. The approach identified the infected areas and quantified the infected leaf area to determine the severity of the disease with greater accuracy.

F. Research Gap

The studies reviewed have verified the effectiveness of explainable learning, cross-modal fusion, multimodal diagnosis, and severity estimation. The current approaches are mainly directed toward one or two of these aspects. There is a limited amount of research that integrates images of sugarcane leaves, environmental parameters, cross-modal attention, disease classification, disease severity assessment, explainability, and lightweight edge deployment in an integrated framework. CaneSense-MS is a proposed model by the researcher that is intended to fill this void.

III. PROPOSED WORK

The CaneSense-MS framework aims to identify cane

leaf diseases and quantify the disease severity with multimodal data. It is a combination of sugarcane leaf images and environmental parameters like temperature, humidity, rainfall, and soil moisture. The image branch is used to obtain features of the image related to the disease, and the environmental branch is used to learn the conditions in which the disease develops. Cross-modal attention fuses the most salient information from both avenues. The fused features are utilized to recognize the disease category, to determine severity of the disease as mild, moderate, or severe, and to estimate percentage of leaf area affected. In addition, explainable predictions and lightweight deployment are enabled by adding Grad-CAM, SHAP, and knowledge distillation [16] to the list.

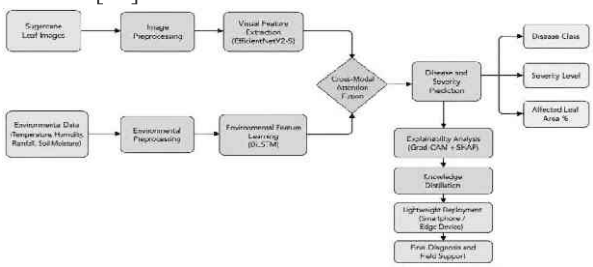

Fig.1. Proposed CaneSense-MS Framework for Multimodal Sugarcane Leaf Disease Detection and Severity Assessment

A. Dataset Description

The Sugarcane Leaf dataset consists of 6,748 high-resolution images used for the classification of sugarcane diseases. It contains 9 disease categories, healthy leaves, and the dried leaves. The diseases included are banded chlorosis, brown rust, brown spot, grassy shoot, mosaic, pokkah boeng, sett rot, smut, and yellow leaf disease. This is the original resolution of the images, which is approximately 768×1024 pixels. To train the model, all images are resized to 224×224 and split into a training set, validation set, and test set.

B. Dataset Configuration:

Table I Datasets used in the Proposed Framework

Dataset Attribute	Description
Dataset name	Sugarcane Leaf Dataset
Total images	6,748
Original image resolution	Approximately (768 \times 1024) pixels
Input image size	(224 \times 224) pixels
Total classes	11
Disease classes	Banded chlorosis, brown rust, rown spot, grassy shoot, mosaic, okkah boeng, sett rot, smut, and yellow leaf disease
Additional classes	Healthy leaf and dried leaf
Image format	JPEG
Training split	70%
Validation split	15%
Testing split	15%
Augmentation methods	Rotation, flipping, zooming, shifting, and brightness adjustment

C. Environmental Dataset

The environmental data comprises temperature, relative humidity, rainfall, and soil moisture data relevant to the cultivation of sugarcane. These parameters also give additional information on the environmental conditions that could affect the development of disease. Use data from NASA POWER or a local meteorological station for temperature, humidity, and precipitation, and use field sensors for soil moisture. Individual environmental records need to be associated with the leaf image by using the observation date, location, and sampling period [17].

The Sugarcane Leaf Dataset is not synchronized with environmental data; therefore, field observations or location-based weather data needs to be taken separately. Interpolation or median imputation is used to fill in missing data, and all features are normalized prior to training the model.

Table II Environmental Dataset

Attribute	Description
Environmental features	Temperature, humidity, rainfall, and soil moisture
Weather-data source	NASA POWER or local weather station
Soil-data source	Field soil-moisture sensor
Synchronization	Matched using date, location, and image-capture period
Preprocessing	Missing-value handling and normalization
Observation window	Seven days before image capture

D. Visual Feature Extraction

The preprocessed sugarcane leaf images are given as input to the network EfficientNetV2-S for feature extraction. The model detects significant disease characteristics like lesions, rust spots, yellowing, discoloration, dry areas, and leaf shape. The model is initialized using transfer learning with previously learned image features. The last classification layer is discarded, and the visual representation extracted is kept for multimodal learning. Dropout is to mitigate overfitting during training. The resulting feature vector is passed on to the cross-modal fusion stage.

Equation(1) - Visual Feature Extraction

$$F_e = BiLSTM\ (X_e) \quad (1)$$

Where I represent the preprocessed sugarcane leaf image and denotes the extracted visual feature vector containing disease-related characteristics such as lesions, discoloration, rust spots, and damaged regions.

C. Environmental Feature Processing

The environmental measurements are recorded as sequential records of temperature, humidity, rainfall, and soil moisture. Mathematical imputation methods (either interpolation or median imputation) are used prior to normalization to handle missing measures. The BiLSTM network reads the environmental sequence both forwards and

backwards. This facilitates the model to learn the short-term and long-term conditions, which are linked to disease development. The output from the BiLSTM generates a small environmental feature vector. This vector is then added to the pictorial representation.

Equation(2) - Environmental Feature Extraction

$$F_e = BiLSTM\ (X_e) \quad (2)$$

Where $X_e = \{x_1, x_2, \ldots x_T\}$ represents the sequence of environmental observations, including temperature, humidity, rainfall, and soil moisture, and F_e is the resulting environmental feature vector.

D. Cross-Modal Feature Fusion

The visual and environmental feature vectors are supplied to the cross-modal attention module. This module analyzes the associations between the visible symptoms of the leaves and the other environmental factors that are likely to be relevant. More emphasis is placed on features that are significant for disease recognition. Both branches' selected information is merged into a single multimodal representation. This fusion aids in the differentiation of diseases that may have similar visual symptoms under varying environmental conditions. The fused representation is conveyed to the prediction module [18].

E. Disease and Severity Prediction

The fused multimodal features are used to produce three prediction outputs. The first output is the label of the category of the sugarcane disease among the possible classes. The second output is the severity of the infection, which can be mild, moderate, or severe. The third output is an estimate of the percentage disease incidence of the leaf area. These outputs can be learned simultaneously with multitask learning, sharing a common feature representation. Class weighting is used when needed to lessen the impact of an imbalanced disease sample. The final prediction offers disease identification and infection intensity [19].

Equation (3) - Disease and Severity Prediction:

$$\widehat{y_d} = Softmax\left(W_d F_f + b_d\right)$$

$$\widehat{y_s} = Softmax\left(W_s F_f + b_s\right)$$

$$\widehat{y_a} = W_a F_f + b_a \quad (3)$$

Where $\widehat{y_d}$ denotes the predicted disease class, $\widehat{y_s}$ represents the predicted severity level, and $\widehat{y_a}$ represents the estimated percentage of affected leaf area.

A. Explainability Analysis

Finally, Grad-CAM is used on the image branch to show the parts of the leaf that affect the prediction of the disease. The created heat maps can be used to check if the model is concentrating on lesions, spots, the yellow areas, or damaged tissue. The role of temperature, humidity, rainfall, and soil moisture is calculated using the model called the SHAP. Each prediction has positive and negative feature effects shown. These explanations enhance clarity and are used to provide a basis for validation by agricultural experts. They also play a role in identifying false predictions due to background areas or unrelated environmental patterns [20].

Equation (4) - Total multitask loss

$$L_{total} = \lambda_1 L_{disease} + \lambda_2 L_{severity} + \lambda_3 L_{area} \quad (4)$$

Where $L_{disease}$ and $L_{severity}$ are categorical cross-entropy losses, L_{area} is the regression loss for the affected leaf area, and λ_1 , λ_2 and λ_3 control the importance of each task.

A. Knowledge Distillation and Deployment

The entire CaneSense-MS framework serves as the teacher model in knowledge distillation. The smaller model of students learns from correct labels and prediction probabilities made by the teacher. MobileNetV3 can be used as the lightweight image encoder in the student network. This process decreases the number of parameters, model size, memory requirement, and inference time. The compressed model is then converted to a TensorFlow Lite model or ONNX. It can then be integrated on mobile phones, Raspberry Pi devices, or other agricultural edge devices [21].

B. Performance Evaluation

The proposed framework is benchmarked with accuracy, macro F1 score, AUC-ROC, mean absolute error of severity, and inference latency. The overall measure of accuracy is the percentage of all the samples of sugarcane correctly classified. The macro F1-score measures the balanced performance of the classification for all disease categories. AUC-ROC provides the accuracy of the model in distinguishing between classes of the disease. Severity MAE determines the difference between leaf area percentages (LAP) for the actual values and predicted values. Inference latency is the time for processing one multimodal input on the target device [22] [23].

C. Performance Evaluation

The framework is validated with the use of the service latency, carbon emission, energy consumption, makespan, and meeting deadlines rate given as normal. The experiments are conducted at low, medium, high, and burst loads. The comparison is done with the round robin, min–min, particle swarm optimization, and standard grey wolf optimization. All methods are performed using the same tasks, fog node configurations, and carbon conditions. Tables and graphs are used to make comparisons, and percentage enhancements are used. Workload balance is considered an optimization goal

and not a main metric that is reported. The computational performance and environmental sustainability trade-off are discussed in the final analysis [24].

IV. RESULT ANALYSIS

As illustrated in Table III, the proposed CaneSense-MS model exhibits the optimum overall performance among all the considered models. It records the highest accuracy of 96.48%, a macro F1-score of 96.02%, and an AUC-ROC of 0.982. In addition, the lowest severity MAE produced by the model is 3.84%, which means that it is more accurate at estimating the affected areas. It has an inference latency of 38 ms, which is lower than the current models and thus is suitable for real-time field usage. The improvements are primarily achieved through the usage of multimodal feature fusion, cross-modal attention, and knowledge distillation.

Table III Performance Comparison of Existing Models and Proposed CaneSense-MS

Model	Accuracy (%)	Macro F1-Score (%)	AUC-ROC	Severity MAE (%)	Inference Latency (ms)
ResNet50	90.42	89.86	0.934	8.72	84
MobileNetV3	91.28	90.73	0.941	7.95	46
InceptionV4	93.10	92.56	0.958	6.48	92
EfficientNet-B6	93.39	92.91	0.961	6.15	78
Proposed CaneSense-MS	**96.48**	**96.02**	**0.982**	**3.84**	**38**

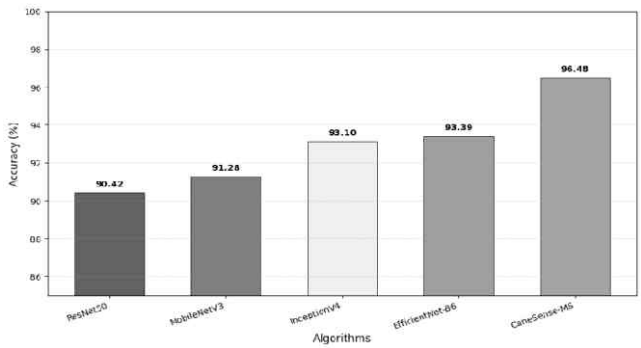

Fig.2. Comparative Model Accuracy Analysis

The accuracy of the classification of the selected models is compared in Figure 2. MobileNetV3 is next with the lowest accuracy of 91.28%, followed by ResNet50 with 90.42%. The accuracies of InceptionV4 and EfficientNet-B6 are 93.10% and 93.39%, respectively. The proposed CaneSense-MS model gets the accuracy of 96.48%, which is the best. Multimodal feature learning and cross-modal attention are credited for this improvement.

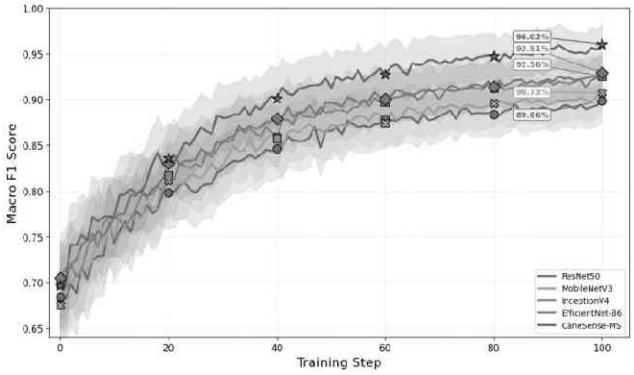

Fig.3. Macro F1-Score Comparison

Figure 3 presents the Macro F1 score improvement of the models in comparison over the course of 100 training steps. ResNet50 and MobileNetV3 get final scores of 89.86% and 90.73%, respectively. The performance is enhanced by 92.56% and 92.91% for InceptionV4 and EfficientNet-B6, respectively. The proposed CaneSense-MS model achieves the best macro F1 score of 96.02%. This result shows good overall classification over all categories of sugarcane diseases. The smoother convergence also indicates that the CaneSense-MS has more stable learning when training.

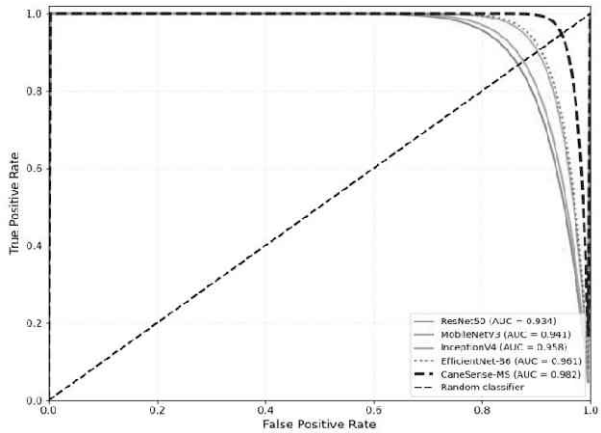

Fig.4. Comparative AUC-ROC Performance Analysis

The ROC performance of the selected models is compared in terms of their AUC values in Figure 4. The AUCs for ResNet50 and MobileNetV3 are 0.934 and 0.941, respectively. The improved values of 0.958 for InceptionV4 and 0.961 for EfficientNet-B6 are achieved. The proposed CaneSense-MS model achieves the best AUC-ROC value of 0.982, thereby having better class separation ability. Its curve is the closest to the upper left part, where it's more sensitive but has a low false positive rate. Multimodal feature fusion leads to improved ROC performance, which demonstrates its capability of differentiating between similar sugarcane diseases from a visual perspective. It also shows that the

classification thresholds CaneSense-MS is able to detect are reliable.

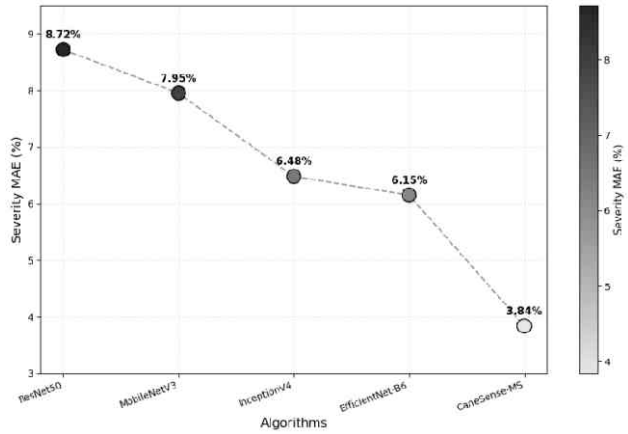

Fig.5. Severity MAE Comparison Analysis

The MAE values for the selected models on severity estimation were compared in Figure 5, with the lowest value as the best model. The top two with the least error are ResNet50 with 8.72% and MobileNetV3 with 7.95%. The error is lowered to 6.48% with InceptionV4 and 6.15% with EfficientNet-B6. The lowest severity MAE is obtained when using the proposed CaneSense-MS model (3.84%). This demonstrates the better estimate of affected leaf area made by the model as compared to existing methods. Multimodal feature fusion and joint disease-severity learning are the primary factors for improving performance. Thus, CaneSense-MS can assist in making better decisions for timely field-level treatment.

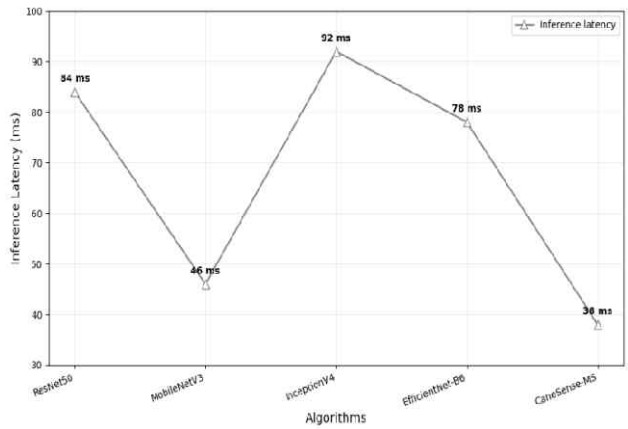

Fig.6. Comparative Deadline Satisfaction Analysis

To compare, the inference latency of the selected models is compared (Figure 6) with a lower value representing a faster prediction. The latency of EfficientNet-B6 is 78ms, while that of MobileNetV3 is 46ms. ResNet50 follows with a latency of 84ms, and InceptionV4 has the worst latency of 92ms. The proposed CaneSense-MS model provides a minimum inference latency of 38 ms. this is because knowledge distillation and a lightweight design lead to a reduction in computational efficiency. Hence, CaneSense-MS is better suited to be deployed in real time at the edge of the network, such as on smartphones.

V. CONCLUSION

The proposed CaneSense-MS framework is the integration of sugarcane leaf images and the environmental measurements for better disease detection and severity assessment. The EfficientNetV2-S, BiLSTM, and cross-modal attention networks add up to good visual symptom and environmental learning, and predictions of EfficientNetV2-S, BiLSTM, and cross-modal attention networks are interpretable with Grad-CAM and SHAP. The model has demonstrated better classification accuracy, macro F1 score, AUC-ROC, severity estimation, and inference latency compared to the current methods. Knowledge Distillation also facilitates knowledge transfer to lightweight edge devices and smartphones for real-world diagnosis. The framework can help farmers and agricultural experts to detect disease conditions early and choose appropriate management measures. The future direction of the project will concentrate on obtaining larger synchronized multimodal datasets, testing the model across other varieties, seasons, and regions of the sugarcane, incorporating thermal or hyperspectral images, and creating real-time disease progression monitoring. Further research will also examine uncertainty estimation, missing-sensor handling, and the cloud-based decision support for the broader agricultural application.

REFERENCES

[1] S. Thite, Y. Suryawanshi, K. Patil, and P. Chumchu, "Sugarcane leaf dataset: A dataset for disease detection and classification for machine learning applications," Data in Brief, vol. 53, Art. no. 110268, 2024.

[2] S. D. Daphal and S. M. Koli, "Enhanced deep learning technique for sugarcane leaf disease classification and mobile application integration," Heliyon, vol. 10, no. 8, Art.no. e29438, 2024.

[3] İ. Kunduracıoğlu and I. R. S. Yılmaz, "Deep learning-based disease detection in sugarcane leaves: Evaluating Efficient Net models," Journal of Open Innovation: Technology, Market, and Complexity, 2024.

[4] X. Li, X. Li, M. Zhang, Q. Dong, G. Zhang, Z. Wang, and P. Wei, "SugarcaneGAN: A novel dataset generating approach for sugarcane leaf diseases based on lightweight hybrid CNN–Transformer network," Computers and Electronics in Agriculture, vol. 219, Art. no. 108762, 2024.

[5] D. Suryaprabha, S. Saraswathi, V. S. Lekha, and J. Loghesh, "Sugarcane leaf disease classification using convolutional neural network," in Proc. IEEE International Conference on Trends in Quantum Computing and Emerging Business Technologies, 2024.

[6] S. Srinivasan, S. M. Prabin, S. K. Mathivanan, H. Rajadurai, S. Kulandaivelu, and M. A. Shah, "Sugarcane leaf disease classification using deep neural network approach," BMC Plant Biology, vol. 25, Art. no. 282, 2025.

[7] P. Pugazhendi, "Uncertainty-aware deep learning for sugarcane leaf disease detection using Monte Carlo dropout and MobileNetV3," AgriEngineering, vol. 8, no. 1, Art. no. 31, 2026.

[8] H. El-Behery et al., "An explainable deep learning framework for few-shot crop disease detection in rice and sugarcane using CNN-based feature extraction," Scientific Reports, 2026.

[9] X. Wang et al., "Intelligent grading of sugarcane leaf disease severity by integrating physiological indicators and machine learning," Frontiers in Plant Science, vol. 16, 2025.

[10] H. Lee et al., "A deep learning-based crop disease diagnosis method using multimodal mixup augmentation," Applied Sciences, vol. 14, no. 10, Art. no. 4322, 2024.

[11] Y. Lu, X. Lu, L. Zheng, M. Sun, S. Chen, B. Chen, T. Wang, J. Yang, and C. Lv, "Application of multimodal transformer model in intelligent agricultural disease detection and question-answering systems," Plants, vol. 13, no. 7, Art. no. 972, 2024.

[12] X. Wu, J. Zhang, Z. Zou, C. Chen, Y. Yu, P. Yu, Y. Xiao, Q. Wang, W. M. W. W. Kandegama, and G. Hao, "PlantIF: Multimodal semantic interactive fusion via graph learning for plant disease diagnosis," Plant Phenomics, Art. no. 100132, 2025.

[13] K.Kathirvel,Javahar S. Deep Learning Techniques for Pneumonia Detection: A Review of CNN Applications, In Proceedings of the 9th International Conference on Electronics Communication and Aerospace Technology ICECA 2025. IEEE.

[14] Sivaraman, M., & Shriman, B. (2026, May). Air Pollution Level Prediction using Machine Learning. In 2025 3rd International Conference on Sustainable Computing and Smart Systems (ICSCSS). IEEE.

[15] C. Zhao, C. Li, X. Wang, X. Wu, Y. Du, H. Chai, T. Cai, H. Xiang, and Y. Jiao, "Plant disease segmentation networks for fast automatic severity estimation under natural field scenarios," Agriculture, vol. 15, no. 6, Art. no. 583, 2025.

[16] P. Bedi, P. Gole, and V. Jindal, "PDSE-Lite: Lightweight framework for plant disease severity estimation," Frontiers in Plant Science, vol. 14, Art. no. 1319894, 2024.

[17] R. Polly and E. A. Devi, "Semantic segmentation for plant leaf disease classification and damage detection: A deep learning approach," Smart Agricultural Technology, vol. 9, Art. no. 100526, 2024.

[18] N. Nigar, H. M. Faisal, M. Umer, O. Oki, and J. Lukose, "Improving plant disease classification with deep learning-based prediction model using explainable artificial intelligence," IEEE Access, vol. 12, pp. 100005–100014, 2020.

[19] Krishnaveni A, Krishnaraj S, "Hybrid GAN-CNN Approach for Predicting Tomato Leaf Diseases",9th International Conference on Inventive Systems and Control (ICISC 2025) 12-13, August 2025, IEEE, JCT College of Engineering and Technology, 12-13 August 2025, IEEE, Coimbatore, Tamilnadu, India.

[20] Waiker, V., Thangarasu, N., Bhalla, A., Das, S., Patel, D. J., & Acharjya, K. (2025). Predictive crop yield analysis using deep learning and optimized sensor data in precision agriculture. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 16(3), 155-177.

[21] S. M. Alhammad et al., "Deep learning and explainable AI for classification of potato leaf diseases," Frontiers in Artificial Intelligence, vol. 7, Art. no. 1449329, 2025.