A Convolutional–Transformer Blend Model for Accurate Multiclass Crop Disease Identification

Dontha Sai Tejaswini[1], Sk.MD.Rafi[2]

[1]P.G Scholar, Department of M.tech(CSE), A1 Global Institute of Engineering and Technology in markapur Prakasam dist -523316, JNTU Kakinada, E-mail: saitjswn@gmail.com

[2] Associate Professor, Department of M.tech(CSE), A1 Global Institute of Engineering and Technology in markapur Prakasam dist -523316, JNTU Kakinada, E-mail: rafisk.sk133@gmail.com

Abstract: Crop diseases gift a huge challenge to international agricultural output, ensuing in great yield reductions, plant fatalities, and the ability extinction of precise vegetation, consequently negatively impacting farmers and the global food deliver. Conventional strategies for ailment identification, ordinarily depending on visual evaluation by farmers and agricultural professionals, are labour-in depth, subjective, and susceptible to errors. Recent breakthroughs in artificial intelligence (AI) present a probable opportunity through actual-time tracking, automated identity, and sensible selection-making, enabled by way of the integration of Internet of Things (IoT) and cloud computing technology. This observe investigates the utilisation of numerous algorithms for type tasks, encompassing Convolutional Neural Networks (CNN), Vision Transformers (ViT), InceptionV3, EfficientNetB0 and B7, Global Wavelet Transform (GWT), Xception, NasNetMobile, CVT (CNN ViT), ICVT (Inception ViT), and EGWT (EfficientNet GWT). We employ state-of-the-art models which include YOLOV5x6, YOLOV5s6, YOLOV8n, and YOLOV9n for the detection of anomalies in crop leaves. Additionally, datasets together with PlantVillage Data, Cassava Data, and Tomato Data are utilised to improve the precision and efficacy of crop ailment identification. This studies seeks to equip farmers with superior techniques to relieve the effects of crop sicknesses on agricultural productiveness.

"Index Terms - *Crop Disease Detection, Convolutional Neural Networks (CNN), Vision Transformers (Vit), Hybrid Models, Transfer Learning, Object Detection, YOLO, Deep Learning, Plantvillage Dataset, Cassava Dataset, Tomato Dataset, Efficientnet, Global Wavelet Transform (GWT)".*

1. INTRODUCTION

Early detection and accurate diagnosis of crop disease is essential to protect the health of crops and improve agricultural yields. Effective measures for detection and management can significantly increase the yield and quality of crops, reduce waste by source and protect agroecosystems. Between different techniques, it has been shown that deep learning models for crop image analysis are a particularly effective strategy.

Since the establishment of convolutional neural networks (CNN) in 2012, these models have seen significant development and optimization, which are promoted as essential in computer vision tasks [10]. Recently, visual transformers (VIT) have gained considerable attention to their ability to manage long -term dependence and demonstrate strong performance in visual tasks [1] [16]. Convolutional neural networks (CNN), VIT and hybrid architecture that integrate the benefits of both frames showed remarkable efficiency in the detection of crop disease [3] [19].

The efficiency of these models is supported in advance on extensive public data files, including imagENets [10], common objects in context (Coco) [11] and MIT places [12] that offer a robust basis for transmission learning. However, a significant problem in implementing these deep learning models is their increased number of parameters that directly affect computing complexity. This problem is particularly evident in defects, where performance improvements often correlate with a significant increase in model parameters [19]. Confrontation with this difficulty continues to be the primary problem in constantly improving the deep learning techniques to identify crop disease.

2. RELATED WORK

More studies have been exploring the use of artificial intelligence (AI) and deep learning methodologies in the detection and categorization of crop disease. Convolutional neural networks (CNN) were essential for this research due to their exceptional effectiveness in image -related tasks. Kurmi et al. He devised a deep CNN model to identify agricultural diseases from the pictures of the leaves and showed good accuracy of the classification [4]. Agarwal et al. They use CNN to effectively detect tomato crop disease, emphasizing the robustness and scalability of the model [5].

After the CNN breakdown, vision transformers (VIT) have been shown as significant for their ability to model dependence on a long range in the pictures. Fu et al. He introduced an increased approach based on VIT for detecting images of crop pests and demonstrated its effectiveness in controlling complex formulas [1]. Wang et al. Integrated convention networks with transformers for the development of hybrid architecture for regional classification of agricultural diseases, which provides excellent results with regard to independent models [3].

The integration of transmission learning significantly improved the efficiency of these models. Hassan et al. It used the transmission learning methodology based on CNN to diagnose plant diseases and therefore reduced the duration of the training and increased accuracy across different data sets [6]. Similarly Chen et al. The learning of deep transmission has been used to identify rice plant plants, which demonstrates the versatility of preliminary models [20] [22].

In addition to model architecture, the quality and diversity of the data set is necessary to increase the accuracy of detection. Plantvillage data file was widely used to practice and validate deep learning models [6] [20]. In addition, Cassava and tomatoes data sets offer significant standards for model performance assessment across different crops [5] [22].

Recent developments in object detection models have significantly improved this field. Architecture based on Yolo, such as Yolov5x6, Yolov5s6 and Yolov8n, have been used to accurately locate and identify the diseases in crop leaves, and provide a combination of speed and accuracy [2] [5].

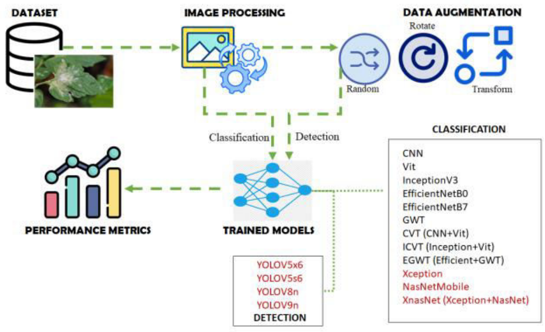
Regardless of these achievements, problems, including increased computing expenditure and the complexity of the model, especially with transformer -based proposals, will last. The initiatives to solve these challenges include the creation of light hybrid models, such as CVT (CNN + VIT) and ICVT (Inception + VIT), designed to optimize power and efficiency [3] [18].

Furthermore, methods such as EGWT (efficiency + GWT) have been introduced to increase the efficiency of resources while maintaining high accuracy [19].

The amalgamatization of sophisticated methodologies of deep learning, hybrid architectures, transmission learning and excellent data sets permanently drive progress in the diagnosis of crop disease, which brings great prospects of sustainable agriculture.

3. MATERIALS AND METHODS

The system to be proposed seeks to transform crop disease detection and classification by incorporating state-of-the-art artificial intelligence algorithms in a strong framework. Utilizing an end-to-quit technique, the gadget makes use of loads of category algorithms, including Convolutional Neural Networks (CNN)[6], Vision Transformers (ViT)[1][16][19], InceptionV3[10], EfficientNetB0 and EfficientNetB7[4], Group-Wise Transformer (GWT)[19], Xception[4], and NasNetMobile[22], to correctly analyze plant health. Apart from those standalone algorithms, the system additionally consists of hybrid models like CVT (CNN + ViT)[3][18], ICVT (Inception + ViT)[1][7], and EGWT (Efficient + GWT)[19] to in addition enhance class accuracy. For detection purposes, the system makes use of the YOLO circle of relatives of fashions, namely YOLOV5x6, YOLOV5s6, YOLOV8n, and YOLOV9n[2][5], to quick and accurately locate abnormalities in crop leaves. The model works on famous datasets like Plant Village Data[6][20], Cassava Data[22], and Tomato Data[5][20], supplying balanced coverage of vegetation illnesses. The integration of the numerous techniques offers actual-time information to farmers, making it easier to manipulate vegetation health and maximize agricultural productiveness.



“Fig.1 Proposed Architecture”

This image (Fig. 1) defines the workflow of a deep image processing system and emphasizes classification and detection tasks. The method begins with a data file followed by image processing techniques, including random transformations and rotations for data augmentation. Models such as CNN, Vision Transformers (VITS) and hybrid architecture such as EffecientB0, EffecientB7 and Xception are used to classify images. Yolo models, including Yolov5, are used to detect tasks. The system evaluates performance through the relevant indicators, ensures effective analysis and accurate results in categorization and detection activities.

i) Dataset Collection:

The proposed approach uses many established data sets for training and verifying models of crop disease. These data sets offer different and extensive data, which guarantees the power of the model in different types of crops and disease situations.

Plant Village Dataset

Plantvillage data file is widely used to classify plants disease, including essential collection of photos that show more than 50 crops and 14 categories of diseases[6][20]. This data set is essential for training deep learning models in identifying plants, offering high quality, marked photos that support classification and detection tasks. A variety of crops and diseases guarantee that models developed by this data set can generalize across many types of crops, thereby improving their importance in practical agricultural contexts.

“Fig.2 Dataset Collection 38 classes”

Cassava Dataset

The Cassava data file is dedicated to identifying diseases in Kasava plants, in an important crop in a number of global regions. It includes annotated photographs to identify the disease, including cassava brown streak disease (CBSD) and cassava mosaic disease (CMD). This data file was widely used in the development of machine learning models, especially for problems associated with the detection of Crops diseases [22]. Incorporation guarantees that the proposed system is applicable to the cultivation of the cassava, where the supervision of diseases is essential for revenue.

“Fig.3 Dataset Collection - 5 classes – cassava bacterial blight (CBB), cassava brown streak disease (CBSD), cassava green mottle (CGM), cassava mosaic disease (CMD), healthy”

Tomato Dataset

The data set for tomatoes includes an extensive collection of photographs on many diseases of tomatoes, causing a significant source to detect and classify tomato disease. This data set has photographs annotated for prevailing tomato diseases, such as early mold, late fungal and tomato virus (TylCV). The diversity and caliber of images in this data file make it easier to identify the model and classify disease diseases in tomato plants, making it a necessary source for tomato growers.

“Fig.4 Dataset Collection - 11 classes –bacterial spot, early blight, healthy, late blight, leaf mold, powdery mildew, septoria leaf spot, spider mites two spotted spider mite, target spot, tomato mosaic virus, tomato yellow leaf curl virus”.

These data sets are essential for training, verifying and testing algorithms of the proposed system and guaranteeing effective performance in different types of crops and disease conditions. Using these diverse data sets, the system can provide real -time farmers that can be done in real time, allowing them to successfully manage crops.

ii) Image Pre-Processing:

Pictures are necessary to increase the performance and generalization of deep learning models in the detection of crop disease. The proposed system uses the ImagedDatagener method for enlargement and pre -work input images. This procedure ensures that the model is obtained by various and carefully processed data and therefore increases its robustness and ability to control a number of image deviations.

The subsequent procedures of preliminary processing are implemented:

(a) Re-scaling the Image: images are normalized by dividing pixels by 255, so ensuring that the input data falls within the range [0, 1]. The normalization procedure facilitates faster convergence of the model during training by ensuring consistent input values, reducing the duration of the training and increasing accuracy [4].

(b) Shear Transformation: shear transformations are used to accidentally push the image in a specified direction, generate new views and adjust the angles of image. This strategy increases the model's invariance to less geometric distortion, thus increasing its robustness to different observation angles and crop orientation [6]. The editing increases data enlargement by increasing the diversity of the data set of training.

(c) Zooming the Image: In the images, a random zoom is used to replicate fluctuations at a distance of the object inside the image. This step of pre - processing increases the generalization of the model by exposing the diverse levels of zoom and therefore increases its ability to detect disease on multiple scales. In addition, Zoom ensures that the model shows less sensitivity to minor fluctuations in the size and trim of plants [1] [5].

(d) Horizontal Flip: Horizontal overturning is used to reflect images through vertical axis. This segment of initial processing introduces modifications within the records record, in particular for fashions that may stumble upon photographs of plants captured from extraordinary angles. Horizontal overturning lets in the model to attain extra generalized capabilities and will increase its robustness to exclusive orientation of flora at some point of deployment. [4] [19].

(e) Reshaping the Image: The pictures are changed into a uniform measurement ordered through the neural network structure. This phase guarantees a uniform processing of all snap shots, allowing their organized get right of entry to to the version. The uniform image dimensions are important for green version schooling and guarantee that the community can examine photos with out mistakes or inconsistencies. [3] [6].

The practise techniques achieved via the Imagedatagenerator appear to growth training statistics and ensure that the version will encounter a variety of photograph versions. This expands the capability of the version to generalize and increase its accuracy whilst it is used to hit upon crop sickness in realistic agricultural contexts. [6] [5] [19].

iii) Training & Testing:

Processed images for training the proposed model of crop disease detection are divided into data sets of training and validation. The training set of training is used to instruct the model by entering the network, allowing the machine to recognize the formulas and properties associated with various crops. The model uses deep learning methods such as CNN, VIT and Hybrid models (eg CVT, ICVT) to extract knowledge from data. Optimization methods such as stochastic gradient descent (SGD) or Adam are used during training, to minimize loss function, usually categorical cross entropy, to increase the accuracy of the model [6] [4].

The test data file is different from training data and is used to assess the ability to generalize the model. After training, the model undergoes an evaluation of this new data set to determine its efficiency in accuracy, precision, recall and F1- score. The testing phase ensures that the model is durable and proficient in accurately identifying crop disease in new real images [1] [5].

iv) Algorithms:

(a) Classification Models:

CNN (Convolutional Neural Network) deep learning algorithms are designed to analyze the structured data of the grid, especially photographs. In the identification of crop disease they acquire CNN spatial hierarchy functions through convolutional layers, which demonstrates expertise in extraction and classification of elements. These models are particularly capable of identifying disease patterns in crops and therefore improve the accuracy of disease detection and facilitate rapid intervention suggestions. [4] [6].

ViT (Vision Transformer) uses transformers on image data by interpreting image patches as sequences, facilitating rapid extraction of elements and classification of agricultural diseases. Through the application of the self -knowledge mechanisms, Vitus increases the ability of the model to identify complex formulas and relationships within images, increasing the effectiveness of the disease detection [3] [7].

InceptionV3 uses several filters inside one layer to effectively capture a variety of functions, which increases its efficiency in diagnosing crop disease. This model is weighed for its ability to recognize complex patterns, which guarantees great accuracy and at the same time maintains the computational economy necessary for diagnosing disease in real time in agriculture [5] [6].

EfficientNetB0 is a light-weight convolution neural community that optimizes the electricity and length of the model via the scaling get right of entry to. It is most suitable for identifying crop disorder, offering high performance and accuracy in spite of low computational resources, that's appropriate for implementation in agricultural surroundings restricted to assets [5] [6].

***EfficientNetB7*,** multiplied iteration of EfficientNet, presents outstanding classification performance by means of scaling of depth, width and backbone. The optimized layout lets in the model to correctly identify numerous plants, which is suitable for actual -time packages in unique agriculture [5] [6].

GWT (Group-Wise Transformer) integrates the convolutional networks into organization attention, permitting effective extraction and illustration of functions. This hybrid method successfully captures complicated relationships in crop snap shots and improves the accuracy of category by way of specializing in relevant picture regions for ailment prognosis [7].

Xception makes use of an in -intensity removable concretion to growth the performance of the factors. It suggests know-how of the learning of complex patterns and is adept into distinguishing among one of a kind symptoms of the disorder, causing it to be an impressive opportunity of type of crop disorder [5] [6].

NasNetMobile is optimized for mobile and marginal gadgets and gives excessive accuracy while minimizing computational fees. Its structure helps green operation of equipment with a limited source, permitting monitoring plants in the agricultural surroundings in actual time. [6] [7].

CVT (CNN + ViT): Amalgamates Convolutional neural network with vision transformers to take advantage of each frames. This hybrid methodology notably improves the performance of the extraction and category of features and therefore will increase the accuracy of crop detection by means of figuring out complex formulas in crops [4] [7].

ICVT (Inception + ViT): Integrates the Inception structure with ViT to earn the benefits of both models. The adepts seize complex functions and hyperlinks, which could be very beneficial for the exact detection and classification of agricultural illnesses in various environments. [4] [7].

EGWT (Efficient + GWT): Integrates an efficient GWT module to improve the effectiveness of type through emphasizing the relevant snap shots while keeping a computing economic system. This method increases the detection of crop disease and ensures rapid interventions in agricultural management. [5] [7].

XNasNet (Xception + NasNet): amalgamates Xception with NasNet to reap stability among accuracy and performance. This architecture is top-rated for detecting and distinguishing numerous agricultural sicknesses, with a mild layout that makes it simpler to screen real -time illnesses on gadgets with useful resource limits [5] [6].

(b) Detection Models:

YOLOV5x6 is an advanced version of the Yolo identification method (you handiest appearance once) this is diagnosed for its incredible accuracy and speedy performance. When detecting crop sickness, Jolov5x6 anomalies in crop pix, facilitate actual -time symptoms and promote speedy agricultural interventions [4] [6].

YOLOV5s6 is a smaller, optimized variant of Yolo object detection techniques. It achieves balance among velocity and accuracy, that is ideal for fast identity of symptoms of crop sickness. This paradigm is in particular superb for real -time applications that require speedy evaluation and intervention. [4] [7].

YOLOV8n is a complicated model of Yolo, which has been designed to enhance the effectiveness of object detection tasks. It offers high accuracy in figuring out traits associated with plants, allowing powerful monitoring and manipulate of crop health. [6] [7].

YOLOV9n, the latest Yolo version, gives terrific overall performance detection in real time objects. Its sophisticated architecture enables accurate identification and categorization of crop ailment and consequently promotes powerful and speedy responses to agricultural risks [6][7].

4. RESULTS & DISCUSSION

Accuracy: The accuracy of the check worries its potential to effectively distinguish among affected person and healthy instances. In order to assess the accuracy of the test, one ought to calculate the ratio of true positives and true negatives in all evaluated cases. This can be mathematically articulated as:

$$\text{"}Accuracy\text{"} = \frac{\text{"TP + TN"}}{\text{"TP + FP + TN + FN"}} \quad (1)$$

Precision: The precision evaluates the share of exactly labeled instances among instances identified as fine. As a result, the formulation for calculating precision is expressed:

$$\text{"}Precision\text{"} = \frac{\text{"True Positive"}}{\text{"True Positive + False Positive"}} \quad (2)$$

Recall: The recall is a meter in ML that evaluates the capability of the model to understand all applicable instances of a specific elegance. It is the share of exactly anticipated high quality observations of overall real positives and offers perception into the efficiency of the version in identifying the occurrence of a particular class.

$$"Recall" = \frac{"\mathrm{TP}"}{"\mathrm{TP} + \mathrm{FN}"} (3)$$

F1-Score: The F1 score is a metric to evaluate the precision and recall. The metric of accuracy quantifies the frequency of real predictions generated by the model throughout the data file.

$$"F1\,Score" = "2" * \frac{"Recall \text{ X } Precision"}{"Recall + Precision"} * "100"(4)$$

mAP: Mean Average Precision (MAP) is a statistic for comparing ranking first-rate. It evaluates the quantity of pertinent tips and their placement within the listing. MAP at K is determined because the mathematics mean of the Average Precision (AP) at K for all customers or queries.

$$"mAP" = "\frac{1}{n}\sum_{k=1}^{k=n} AP_k" (5)$$

"In Table 1 to 4, the performance metrics—accuracy, precision, recall, F1-score —are evaluated for each algorithm. The Extension and YoloV5s6 for Classification and Detection achieves the highest scores. Other algorithms' metrics are also presented for comparison".

"Table.1 Performance Evaluation Metrics of Classification - Cassava"

Model	Accuracy	Precision	Recall	F1 Score
CNN	0.688	0.710	0.673	0.685
ViT	0.615	0.615	0.615	0.615
InceptionV3	0.515	0.523	0.385	0.431
EfficientNetB0	0.615	0.615	0.615	0.615
EfficientNetB7	0.615	0.615	0.615	0.615
GWT	0.615	0.615	0.615	0.615
Xception	0.978	0.978	0.978	0.978
NASNetMobile	0.979	0.980	0.979	0.979
CVT	0.649	0.734	0.600	0.645
ICVT	0.513	0.512	0.423	0.453
EGWT	0.607	0.655	0.572	0.600
XNASNet	0.999	1.000	0.997	0.999

"Table.2 Performance Evaluation Metrics of Classification – Plant Village"

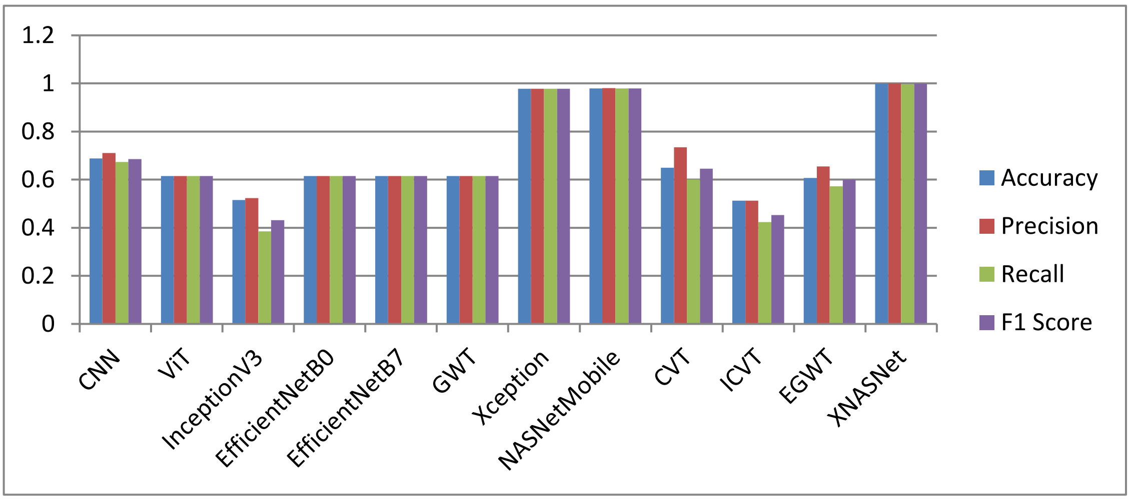
Model	Accuracy	Precision	Recall	F1 Score
CNN	0.895	0.914	0.885	0.898
ViT	0.992	0.993	0.992	0.992
InceptionV3	0.987	0.987	0.987	0.987
EfficientNetB0	0.198	0.198	0.198	0.198
EfficientNetB7	0.197	0.197	0.197	0.197
GWT	0.928	0.936	0.924	0.929

Xception	1.000	1.000	1.000	1.000
NASNetMobile	1.000	1.000	1.000	1.000
CVT	0.989	0.989	0.989	0.989
ICVT	0.995	0.995	0.995	0.995
EGWT	0.852	0.947	0.765	0.839
XNASNet	1.000	1.000	1.000	1.000

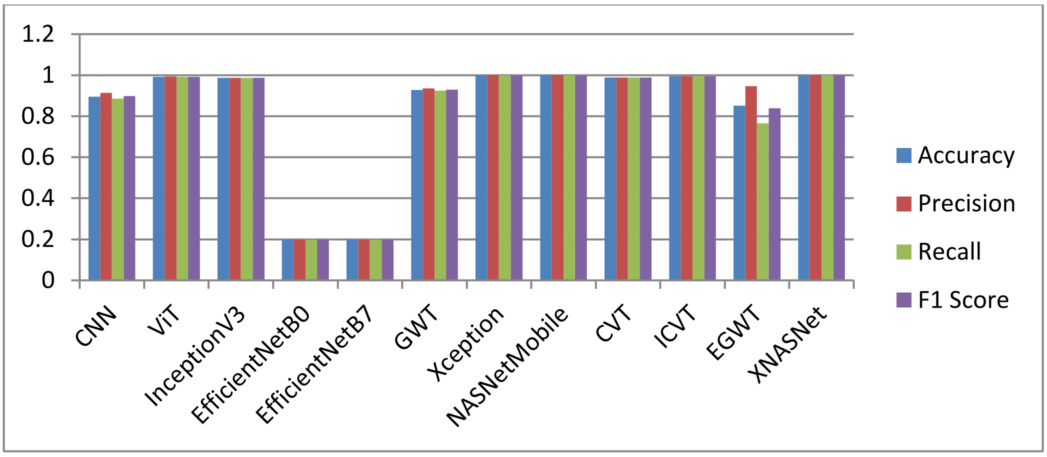
"Table.3 Performance Evaluation Metrics of Classification – Tomato"

Model	Accuracy	Precision	Recall	F1 Score
CNN	0.898	0.923	0.880	0.894
ViT	0.119	0.000	0.000	0.000
InceptionV3	0.827	0.833	0.826	0.828
EfficientNetB0	0.134	0.134	0.134	0.134
EfficientNetB7	0.119	0.119	0.119	0.119
GWT	0.118	0.118	0.118	0.118
Xception	0.991	0.995	0.984	0.988
NASNetMobile	0.867	0.868	0.866	0.988
CVT	0.906	0.942	0.861	0.888
ICVT	0.861	0.887	0.829	0.848
EGWT	0.145	0.000	0.000	0.000
XNASNet	0.918	0.940	0.893	0.908

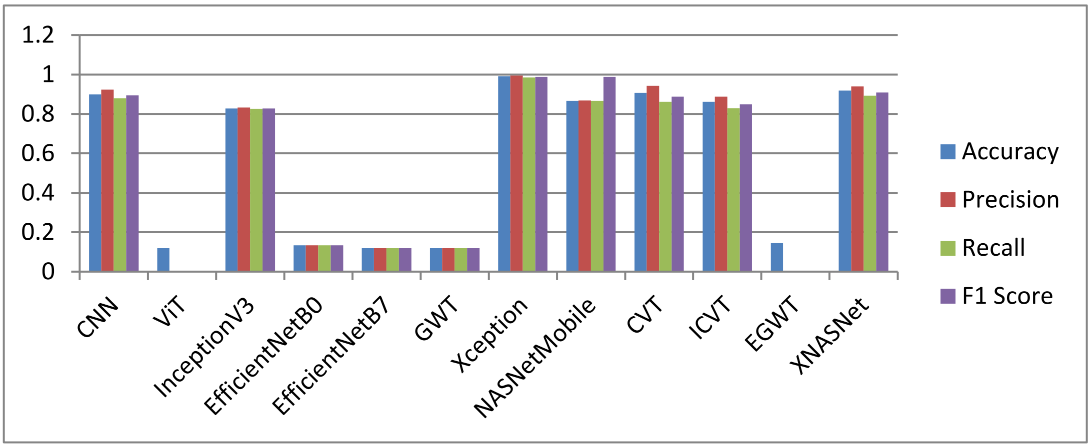
Graph.1 Comparison Graphs of Classification – Casavaa



"Graph.2 Comparison Graphs of Classification – Plant Village"



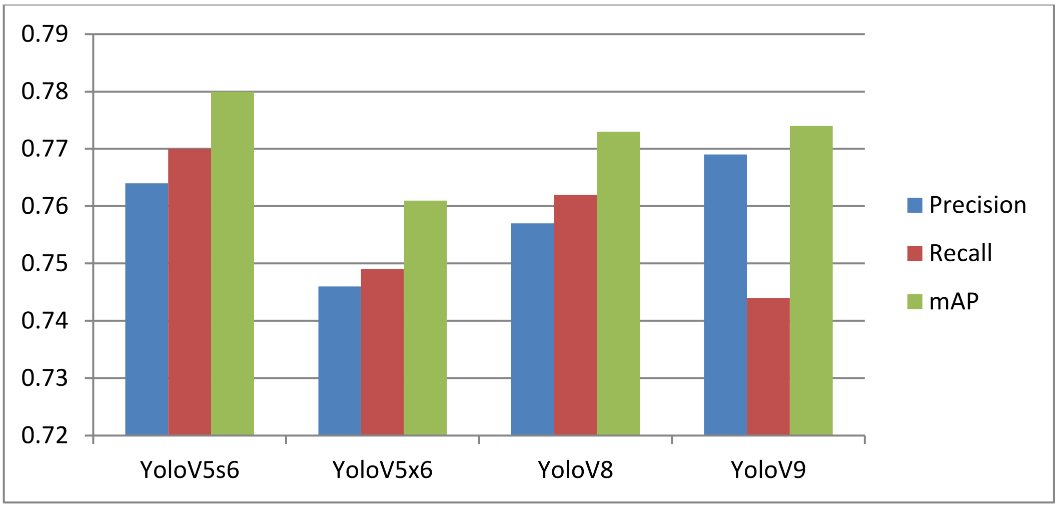
"Graph.3 Comparison Graphs of Classification – Tomato"



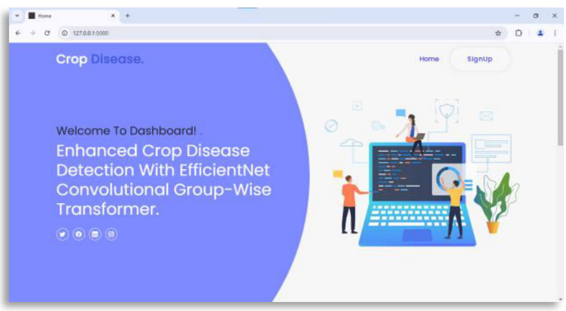
"Table.4 Performance Evaluation Metrics of Detection- All 3 Datasets"

Model	Precision	Recall	mAP
YoloV5s6	0.764	0.770	0.780
YoloV5x6	0.746	0.749	0.761
YoloV8	0.757	0.762	0.773
YoloV9	0.769	0.744	0.774

"Graph.4 Comparison Graphs of Detection – All 3 Datasets"

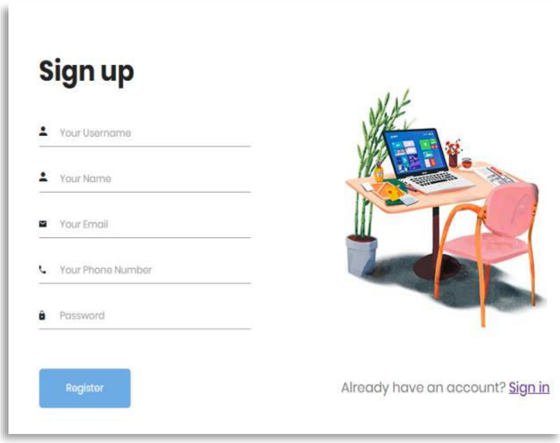


"In classification graphs, accuracy is represented in light blue, precision in maroon; recall in green, F1-score in violet, In detection graphs, precision is represented in light blue, recall in maroon, mAP in green ***Graphs (1 to 4)***. In comparison to the other models, the Xception and YOLO shows superior performance across all achieving the highest values. The graphs above visually illustrate these findings".



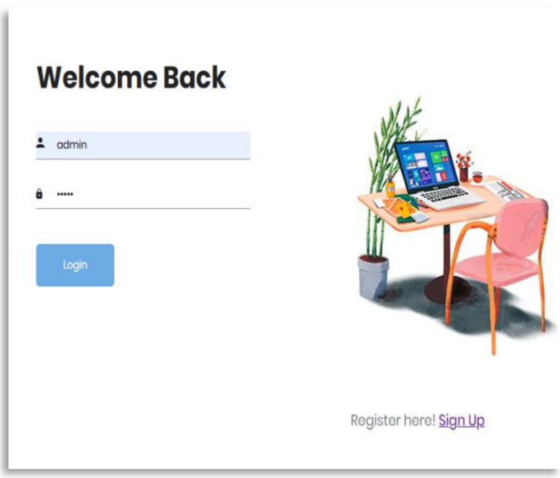
"Fig.4 Home Page"

A web dashboard showcasing "Enhanced Crop Disease Detection With EfficientNet," featuring depictions of individuals engaging with technology.



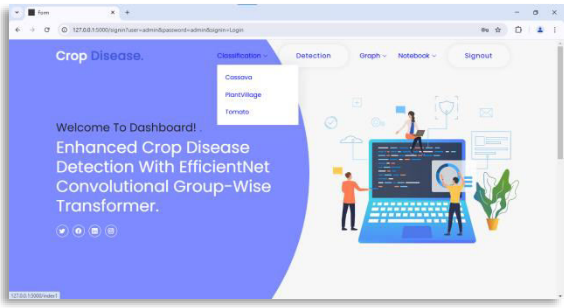
"Fig.5 Signup Page"

A registration page containing fields for user information



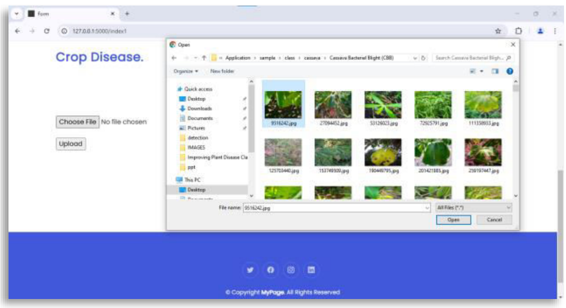
"Fig.6 Signin Page"

A login interface featuring input boxes for username and password.



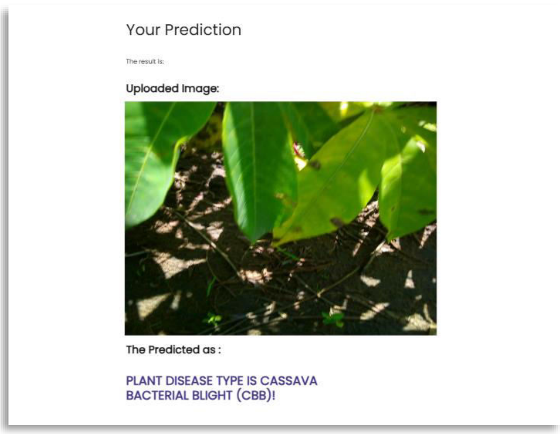
“Fig.7 Classification Dashboard”

A dashboard for agricultural disease diagnosis featuring categorisation, detection, and graphical capabilities, utilising EfficientNet and transformers.



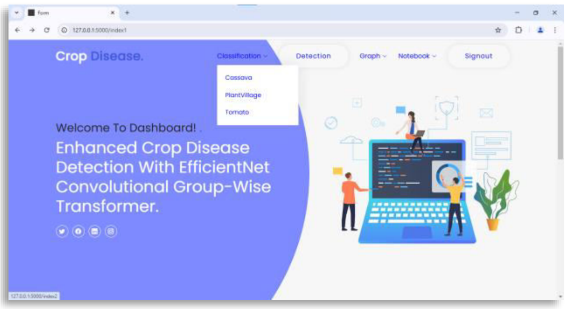
“Fig.8 Test case 1 for Classification Dataset-1”

This graphic seems to depict a file upload interface for your crop disease detection system.



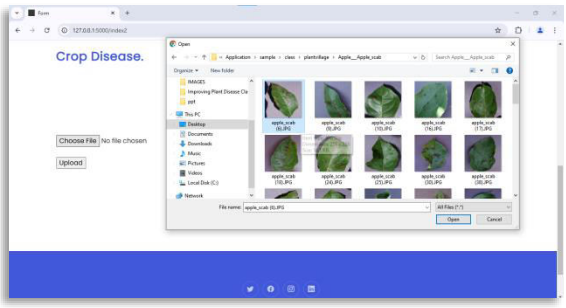
“Fig.9 Results of Test case 1 for Classification Dataset 1”

This picture displays the predictive outcome of your crop disease detection system. The uploaded image was accurately identified as "Cassava Bacterial Blight (CBB)."



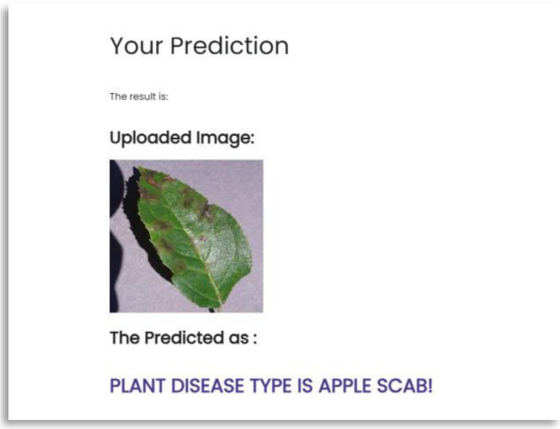
“Fig.10 Dataset-2 Classification”

A dashboard for agricultural disease diagnosis that includes categorisation, detection, and graphical features, utilising EfficientNet and transformers.



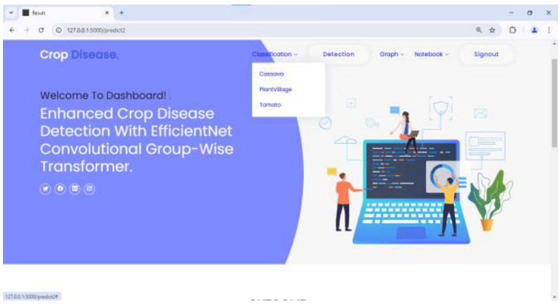
“Fig.11 Test case 1 for Classification Dataset-2”

This image depicts a file selection interface for uploading an image of an apple leaf afflicted with "Apple Scab" disease within a crop disease detection system. The interface enables users to select and upload an image for analysis.



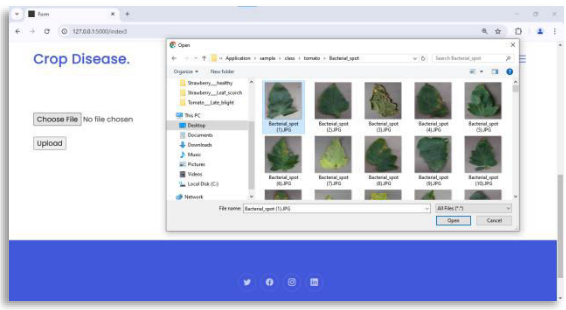
"Fig.12 Results of Test case 1 for Classification Dataset 2"

This figure illustrates the predictive outcome of a plant disease classification system. The uploaded image depicts an apple leaf, and the model accurately diagnoses the illness as Apple Scab.



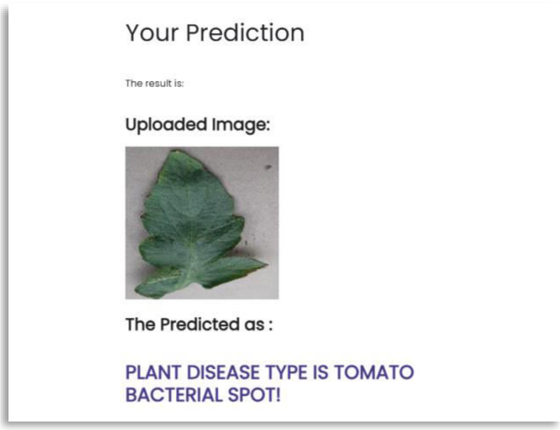
"Fig.13 Dataset-3 Classification"

A dashboard for agricultural disease diagnosis featuring categorisation, detection, and graphical capabilities, utilising EfficientNet and transformers.



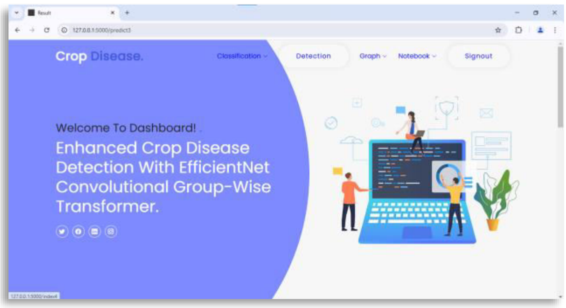
"Fig.14 Test case 1 for Classification Dataset-3"

The file choice display for uploading a tomato leaf photo exhibiting Bacterial Spot in your Crop Disease Detection on line application.



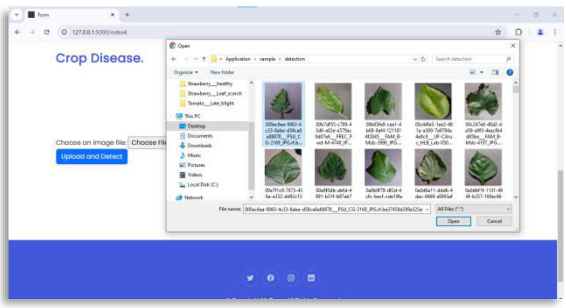
"Fig.15 Results of Test case 1 for Classification Dataset-3"

The version as it should be diagnosed the plant infection as Tomato Bacterial Spot.



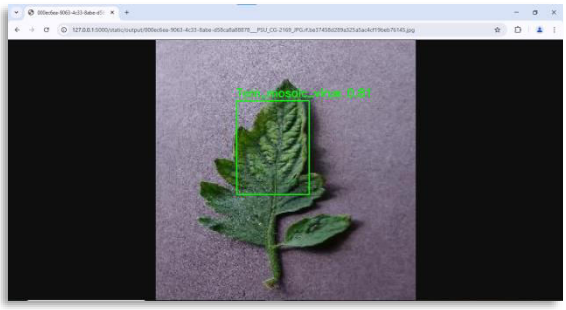
"Fig.16 Dataset-3 Detection"

A dashboard for agricultural disease prognosis incorporating categorisation, detection, and graphical factors, utilising EfficientNet and transformers.



"Fig.17 Test case 1 for Detection Dataset-1"

A web program for the analysis of plant illnesses, with a report add interface for the choice and analysis of various leaf photographs.



"Fig.18 Results of Test case 1 for Detection Dataset-1"

A plant sickness detection set of rules recognised "Tomato Mosaic Virus" with 81% self belief, delineating the troubled leaf place with a bounding box.

5. CONCLUSION

In conclusion, the combination of Xception, NasNetMobile, and XNasNet models has set up amazing efficacy within the analysis of crop sickness datasets. These models performed awesome category consequences due to their state-of-the-art feature extraction competencies and deep analyzing frameworks. Xception's implementation of depthwise separable convolutions facilitated efficient and accurate reputation of complicated sickness patterns, at the same time as NasNetMobile's light-weight however strong structure assured advanced universal performance on mobile and resource-constrained gadgets, rendering it premiere for actual-time monitoring in agricultural environments. XNasNet, through the use of amalgamating the advantages of each Xception and NasNet, provided a great equilibrium of accuracy and performance, markedly improving the overall category accuracy all through the dataset.

The YOLO family of algorithms—YOLOV5x6, YOLOV5s6, YOLOV8n, and YOLOV9n—was effectively utilized for disease detection by using figuring out irregularities in leaf pix. The YOLO fashions, recognized for his or her real-time object detection proficiency, facilitated fast and unique identification of illness symptoms. Among them, YOLOV9n exhibited extremely good overall performance, achieving excessive detection accuracy whilst preserving processing speed. Collectively, those advanced algorithms presented a comprehensive solution for the identification and categorization of crop illnesses, guaranteeing set off and dependable consequences for green agricultural control.

The future scope may additionally expand this version to tackle similarly agricultural difficulties, along with insect identification and yield prediction, whilst additionally analyzing its applicability across many real-international scenarios, which includes distinct crop kinds and environmental occasions. Additional optimization of the version's hyperparameters will be completed to reap an ideal equilibrium among performance and complexity, as a result making sure scalability and performance. Furthermore, integrating a broader variety of datasets and augmenting the version's generalization abilities could be important for boosting its applicability in realistic agricultural contexts.

REFERENCES

[1] X. Fu, Q. Ma, F. Yang, C. Zhang, X. Zhao, F. Chang, and L. Han, ''Crop pest image recognition based on the improved ViT method,'' Inf. Process. Agricult., Feb. 2023.

[2] R. Reedha, E. Dericquebourg, R. Canals, and A. Hafiane, ''Transformer neural network for weed and crop classification of high resolution UAV images,'' Remote Sens., vol. 14, no. 3, p. 592, Jan. 2022.

[3] Y. Wang, Y. Chen, and D. Wang, ''Convolution network enlightened transformer for regional crop disease classification,'' Electronics, vol. 11, no. 19, p. 3174, Oct. 2022.

[4] Y. Kurmi, P. Saxena, B. S. Kirar, S. Gangwar, V. Chaurasia, and A. Goel, ''Deep CNN model for crops' diseases detection using leaf images,'' Multidimensional Syst. Signal Process., vol. 33, no. 3, pp. 981–1000, Sep. 2022.

[5] M. Agarwal, S. K. Gupta, and K. K. Biswas, ''Development of efficient CNN model for tomato crop disease identification,'' Sustain. Comput.,Informat. Syst., vol. 28, Dec. 2020, Art. no. 100407.

[6] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska, ''Identification of plant-leaf diseases using CNN and transfer-learning approach,'' Electronics, vol. 10, no. 12, p. 1388, Jun. 2021.

[7] H.-P. Wei, Y.-Y. Deng, F. Tang, X.-J. Pan, and W.-M. Dong, ''A comparative study of CNN- and transformer-based visual style transfer,'' J. Comput. Sci. Technol., vol. 37, no. 3, pp. 601–614, Jun. 2022.

[8] R. Ahuja and S. C. Sharma, ''Transformer-based word embedding with CNN model to detect sarcasm and irony,'' Arabian J. Sci. Eng., vol. 47, no. 8, pp. 9379–9392, Aug. 2022.

[9] J. Fang, H. Lin, X. Chen, and K. Zeng, ''A hybrid network of CNN and transformer for lightweight image super-resolution,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2022, pp. 1103–1112.

[10] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, ''ImageNet large scale visual recognition challenge,'' Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015.

[11] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick, ''Microsoft COCO: Common objects in context,'' in Computer Vision—ECCV, Zurich, Switzerland. Springer, 2014, pp. 740–755.

[12] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, ''Places: A 10 million image database for scene recognition,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 6, pp. 1452–1464, Jun. 2018.

[13] S. U. Khan, N. Khan, T. Hussain, and S. W. Baik, ''An intelligent correlation learning system for person re-identification,'' Eng. Appl. Artif. Intell., vol. 128, Feb. 2024, Art. no. 107213.

[14] A. Parashar, A. Parashar, A. F. Abate, R. S.Shekhawat, and I. Rida, ''Real-time gait biometrics for surveillance applications: A review,'' Image Vis. Comput., vol. 138, Oct. 2023, Art. no. 104784.

[15] H. Wang, Y. Zhu, B. Green, H. Adam, A. Yuille, and L.-C. Chen, ''Axial-DeepLab: Stand-alone axial-attention for panoptic segmentation,'' in Proc. Eur. Conf. Comput. Vis. Springer, 2020, pp. 108–126.

[16] Y. Xu, Z. Zhang, M. Zhang, K. Sheng, K. Li, W. Dong, L. Zhang, C. Xu, and X. Sun, ''Evo-ViT: Slow-fast token evolution for dynamic vision

transformer,'' in Proc. AAAI Conf. Artif. Intell., 2022, vol. 36, no. 3, pp. 2964–2972.

[17] M.-H. Guo, C.-Z. Lu, Z.-N. Liu, M.-M. Cheng, and S.-M. Hu, ''Visual attention network,'' Comput. Vis. Media, vol. 9, no. 4, pp. 733–752, Dec. 2023.

[18] X. Pan, T. Ye, Z. Xia, S. Song, and G. Huang, ''Slide-transformer: Hierarchical vision transformer with local self-attention,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2023, pp. 2082–2091.

[19] Y. Fang, X. Wang, R. Wu, and W. Liu, ''What makes for hierarchical vision transformer?'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 45, no. 10, pp. 12714–12720, Oct. 2023.

[20] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, ''Using deep transfer learning for image-based plant disease identification,'' Comput. Electron. Agricult., vol. 173, Jun. 2020, Art. no. 105393.

[21] A. Sagar and J. Dheeba, ''On using transfer learning for plant disease detection,'' BioRxiv, 2020.

[22] J. Chen, D. Zhang, Y. A. Nanehkaran, and D. Li, ''Detection of Rice plant diseases based on deep transfer learning,'' J. Sci. Food Agricult., vol. 100, no. 7, pp. 3246–3256, May 2020.