International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 16, No.1 (Sep-24)

http://dx.doi.org/10.12785/ijcds/160195

Optimizing Multi-Level Crop Disease Identification using Advanced Neural Architecture Search in Deep Transfer Learning

Hicham Slimani[1], Jamal El Mhamdi[1] and Abdelilah Jilbab[1]

[1] *E2SN Laboratory, ENSAM/ENSIAS of Rabat, Mohammed V University in Rabat, Rabat, Morocco*

Received 1 Feb. 2024, Revised 27 May 2024, Accepted 31 May 2024, Published 15 Sep. 2024

Abstract: Efficiently managing crop diseases holds immense potential for optimizing farming systems. A crucial aspect of this process is accurately identifying infection levels to enable targeted and effective disease treatment. Despite recent advancements, developing a reliable system for identifying and localizing crop diseases in complex, unstructured field environments remains challenging. Such a system requires extensive annotated data. This study comprehensively evaluates deep transfer learning techniques for identifying the degree of rust disease infection in Morocco's Vicia faba L. production systems. A vast dataset captured under natural lighting conditions and various crop growth stages was created to facilitate this research. Ten deep learning models were rigorously assessed through transfer learning, establishing a benchmark for this task. Deep transfer learning achieved high classification accuracy, with F1 scores consistently surpassing 90.0%. Training time for all models was reasonably short, under 2.5 hours. The NVIDIA Quadro P1000, known for its exceptional performance, was pivotal in achieving this outcome. The Neural Architecture Search-based model emerged as the top performer, achieving an impressive overall F1 score of 90.84%. Three models achieved F1 scores near or above 90.0%, highlighting the effectiveness of deep transfer learning for rust infection identification. This research illuminates the potential of deep transfer learning in detecting and diagnosing crop diseases, specifically rust infection in Vicia faba L. production systems. The findings contribute to developing robust disease management strategies, improving agricultural practices, and enhancing crop yield.

Keywords: Deep Learning, Neural Architecture Search, CNN, Crop Diseases, Real-time Object Detection, Precision Agriculture.

1. Introduction

Agriculture plays a vital role in many economies, such as Morocco, but it faces risks from temperature changes, insects, and diseases. However, given the size of agricultural lands and the scarcity of trained workers, conventional disease detection techniques are frequently arbitrary and ineffective [1]. We are creating cutting-edge methods that use deep learning, image sensors, and computer vision to overcome these obstacles. We aim to develop an intelligent system that can recognise and track crop conditions in real-time. By replacing labour-intensive human inspections with an automated system that provides improved precision and efficacy in detecting and monitoring crop diseases, this system seeks to increase agricultural productivity and resource efficiency. Our research is focused on developing deep learning techniques for accurate diagnosis and efficient detection of multi-level rust disease in Vicia faba L. crops. This innovative method of diagnosing plant diseases holds great promise for the agricultural sector [2], where quick identification and close monitoring are essential. By analyzing data by artificial intelligence (AI) [3] from sensors [4], cameras [5], and microcontrollers [6] inside a networked system [7], artificial intelligence (AI) can identify crop diseases. Machine learning models examine factors such as temperature, humidity, and leaf colour using image recognition techniques to diagnose plant diseases [8]. Convolutional neural networks (CNNs), in particular, are deep learning models that are capable of classifying leaf pictures to differentiate between healthy and unhealthy plants, as well as identifying the crop species [9][10][11]. We employ state-of-the-art deep learning models and advanced sensors to create an automated system that can accurately and quickly diagnose rust disease in Vicia faba L. [12]. This novel strategy will allow for prompt interventions and effective disease management, giving farmers the resources to tackle this enduring danger.

Our goal in investigating deep learning's potential applications in agriculture is to provide new opportunities for enhancing crop health and preserving farmers' livelihoods. With this research, we can incorporate AI-based solutions into agricultural operations. Technology can change the way we manage and prevent crop diseases completely.

E-mail address: hicham_slimani2@um5.ac.ma, j.elmhamdi@um5r.ac.ma, Abdelilah.jilbab@ensam.um5.ac.ma

2. Related Work

Crop disease identification using different CNN models has been extensively studied. Dingju Zhu et al. proposed a hybrid model called MSCVT that combines the advantages of CNN in extracting local disease information and a vision transformer in obtaining global receptive fields [13]. Orlando Iparraguirre-Villanueva et al. used CNN models such as DenseNet-201, ResNet-50, and Inception-v3 to identify and classify diseases in crop plants [14], achieving high accuracy rates of 98% and 97%. Liu et al. introduced DFF-ResNet, a feature fusion residual block that surpasses baseline models like ResNet in accurately identifying insect pests for maintaining a stable agricultural economy. Their innovative approach acknowledges the crucial role of insect pest recognition in maintaining agricultural stability and food security [15]. Sahil Verma et al. introduced a meta-learning-based framework that recommends suitable models for plant disease detection, improving resource utilization and implementation efficiency [16]. These studies demonstrate the effectiveness of CNN models in crop disease identification and highlight the importance of model selection and optimization for accurate and efficient disease detection. Agbaje and Tian used pre-trained models such as MobileNetV2, EfficientNet-B5, and Incep-tionV3, achieving high accuracy in classifying diseased and healthy plants [17]. Kumar, Tyagi, and Poonia compared three CNN architectures, VGG16, ResNet50, and EfficientNetV2, for plant disease identification [18], with EfficientNetV2 achieving an accuracy of 96.06%. Bondre and Patil discussed using deep learning and transfer learning in agricultural disease recognition, highlighting the importance of transfer learning given the existing agricultural disease data tools [19]. A classification scheme for the stages of white scale disease (WSD) infestation in date palm trees was presented by Hessane et al. They assess the performance of the SVM, KNN, RF, and LightGBM algorithms using GLCM texture features and HSV color moments. When GLCM and HSV features are combined, SVM achieves 98.29% accuracy. The framework supports preventive actions for crop productivity and tree protection in oasis agriculture by assisting in the early detection of date palm white scale disease (DPWSD) [20]. This study focuses on the effectiveness of transfer learning algorithms for identifying crop diseases, specifically rust disease in Vicia faba L. crops. The aim was to classify the disease into three levels: healthy, moderate, and severe. The study used the RetinaNet, Fully Convolutional One-Stage Object Detection (FCOS), Faster R-CNN, and YOLO-family architectures, which are one-step object detection approaches using RGB images. The evaluation compared different versions and sizes of these architectures based on various metrics. The proposed architecture showed significant advancements, particularly in precision, enabling accurate diagnosis of crop rust disease and reducing false positives. The real-time effectiveness of the system allows for quick analysis and supports agricultural experts in combating crop diseases. This research highlights the groundbreaking benefits of controlling Vicia faba L. rust disease.

3. Materials and Experimental Procedures

A. Algorithm structure

This scientific study proposes a new method for diagnosing multi-level rust disease in Vicia faba L. crops. It uses a sophisticated image recognition system based on transfer learning, specifically designed for farmers. The system comprises essential components such as parameter configuration, data collection, preprocessing, annotation, and deep learning model training. By utilizing a well-trained model, the system accurately detects the level of infection of rust disease and assesses the effectiveness of the models based on acquired data. This approach shows promise in providing precise information to farmers for effective crop disease management. It is expected to improve decision-making and lead to better disease management methods. Algorithm 1 provides a detailed description of the algorithm. The system flow chart is shown in Figure 1 to visually depict the suggested method's process.

B. Dataset Construction

The Vicia faba L. rust disease dataset utilized in this study was sourced from a farm in Morocco. To capture the progression of the multi-level rust disease, a Sony DSLR-A230 camera was employed, positioned meticulously between 30 and 50 cm from the crop pods. A systematic approach was adopted, capturing photographs at regular intervals of three to four days throughout March and April 2023. This method ensured a comprehensive photographic timeline depicting the various stages of the rust illness. The dataset was compiled using the Joint Photographic Group (JPG) format to preserve an extensive collection of photographs, deliberately capturing various angles and orientations to introduce complexity and variation. A meticulous selection process yielded 3296 high-resolution photos with a pixel resolution of 3872x2592. These expertly captured images showcase the lesions and external signs of rust disease on Vicia faba L. pods. The dataset encompasses three infection severity categories: healthy, moderate, and critical. The healthy pod category includes 1,124 images, while the moderately infected pod category contains 1,279 images and 893 images of the pods with severe infections, providing valuable insights into the advanced stages of rust disease. Figure 2 illustrates the number of instances of each class.

The dataset was partitioned into three subsets, ensuring a comprehensive evaluation process. The testing, validation, and training subsets were allocated in a ratio of 4:1:1, respectively. To meet the model requirements while preserving the appropriate aspect ratio, all photographs were uniformly scaled to 640x640 pixels. This standardized resizing enhances the dataset by incorporating specific information about image capture techniques. This results in a more extensive and relatable representation of rust disease progression in Vicia faba L. crops. The enriched dataset is a valuable resource for farmers and researchers, empowering them with a precise and effective tool for managing rust disease.

Algorithm 1: Multi-Level Disease Recognition with Deep Learning Approach.

1. Capture images of Vicia faba L. crops from the production systems.
2. Preliminary image filtering (poor quality, blurred, small size, etc.).
3. Annotate the Vicia faba L. pods images with the degree of rust disease infection.
4. Split the training, validation, and testing sets from the Vicia Faba L. images dataset.
5. Pre-process the annotated images (e.g., resize, normalize) for training and testing.
6. Apply data augmentation techniques to the training set to improve data variety.
7. Select a suitable DL model architecture for multi-level rust disease infection detection.
8. Initialize the selected deep learning model with appropriate parameters.
9. Train the deep learning model using the training set.
10. Optimize the model's hyperparameters using techniques.
11. Validate the trained model using the validation set and evaluate its performance metrics.
12. Fine-tune the model based on the validation results.
13. Save the trained Rust Disease Infection Detection Model.
14. Test the trained model using the testing set.
15. Apply the trained model to new Vicia faba L. pods images for rust disease infection detection.
16. Analyze the model's classification and assess the accuracy of rust disease infection detection.
17. Compare the performance of different models (e.g., RetinaNet, Faster R-CNN, FCOS ResNet) regarding accuracy and efficiency.
18. Select the best-performing model based on the evaluation metrics and requirements.
19. Fine-tune the selected model further to improve its performance.
20. Deploy the final Rust Disease Infection Detection Model for practical use in Vicia Faba L. production systems.

C. Data annotation

Before commencing the training process for our model, an essential step awaits—meticulously labeling the images obtained from the resized dataset. This process requires meticulous attention to detail and a particular focus on accuracy [21]. To accomplish this task, we employed a remarkable Python-written tool called "LabelImg" and the open-source "MakeSense." This user-friendly graphical image annotation program proved valuable, facilitating image normalization and streamlining the annotation process. With "LabelImg" and "MakeSense," we meticulously annotated the images in the training, validation, and test sets, adhering to the widely recognized VOC format [22]. Through this meticulous annotation process, we obtained XML and Txt files that captured essential information about each image. These files encompassed crucial details such as image names, sizes, class names, target image positions, and other pertinent data. It demonstrates the depth of information in the dataset, emphasizing the exact locations of the target photographs and offering a thorough rundown of the annotated data. We ensure that our model has the data it needs to learn and generate predictions by taking the time and trying to annotate the dataset precisely.

D. Data Augmentation

In deep learning, data augmentation is crucial in mitigating overfitting [23]. It involves introducing artificial perturbations and controlled variations to enhance the training dataset, expanding its diversity and generalizability. Various techniques, such as rotation, translation, saturation adjustments, and geometric transformations, were employed to create a diverse and robust training process. To address the class disparity, proactive measures were taken to balance the dataset by adjusting images of healthy and critical pods. Adaptive scaling and filling procedures were also conducted to ensure uniformity in image preparation. With standardized input image size, the model was given a fair and consistent analysis platform [24].

E. Proposed and Trained Models

The dataset utilized in this study was partitioned into distinct subsets, namely training, validation, and test, maintaining a balanced distribution with a ratio of 4:1:1. Each image, portraying the condition of Vicia faba L. pods, underwent meticulous manual annotation, ensuring comprehensive documentation of relevant attributes. The annotation process was executed precisely, capturing intricate details vital for subsequent analysis.A diverse range of deep learning architectures and pre-trained models were employed to address rust disease identification in Vicia faba L. production systems. Each architectural variant underwent meticulous selection and fine-tuning to optimize performance in accurately and efficiently detecting rust diseases. Here's a detailed overview of each architectural variant utilized in our research, YOLONeural Architecture Search (NAS)L and YOLO-NASM; these architectures amalgamate the YOLO framework with neural architecture search techniques. They autonomously uncover optimal network architectures for object detection tasks, are particularly adept at enhancing performance in complex detection scenarios and are vital for precisely identifying rust diseases amidst diverse environmental conditions. Faster R-CNN ResNet-50 and Faster R-CNN ResNet-101, Rooted in the Faster R-CNN framework, these architectures comprise a region proposal network (RPN) and a region-based convolutional neural network (R-CNN) for object detection. Utilizing ResNet-50 and ResNet-101 backbones, we aimed to scrutinize the impact of varying depths on model performance,

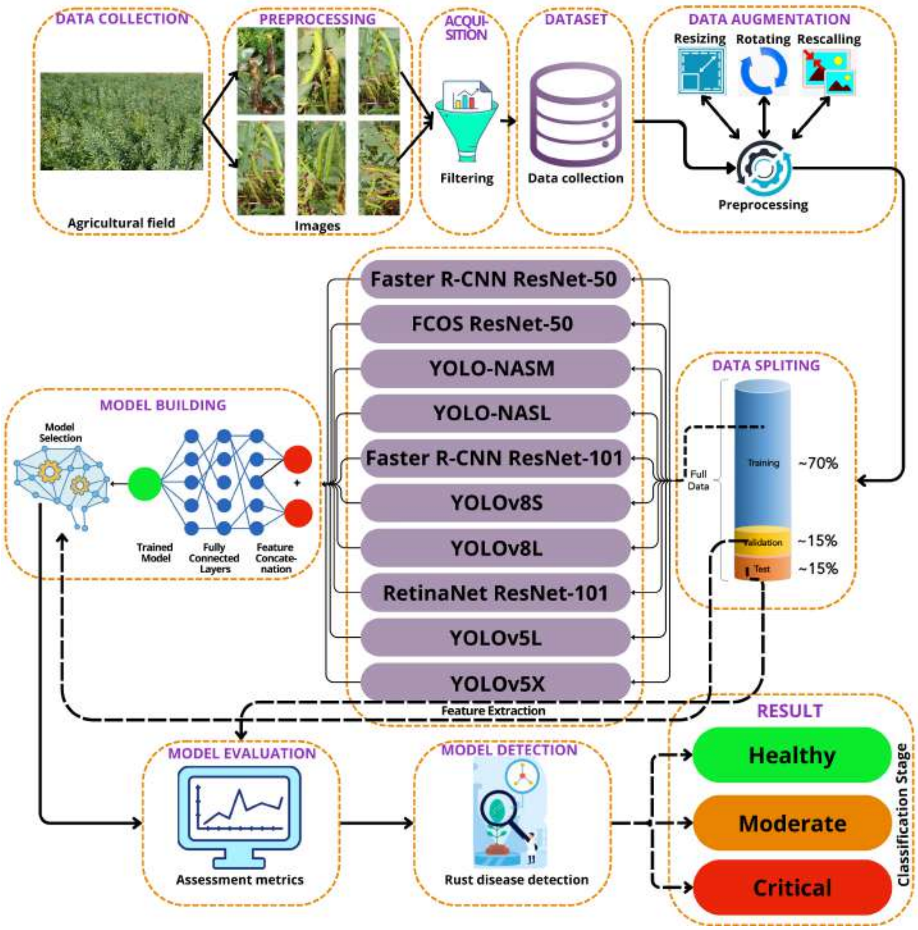

Figure 1. **The proposed model is visually represented through a flowchart.**

discerning the most effective configuration for rust disease detection. Fully Convolutional One-Stage Object Detection (FCOS) ResNet-50 is a one-stage object detection framework that directly predicts bounding boxes and class labels from feature maps. Leveraging the ResNet-50 backbone facilitated efficient feature extraction, which is crucial for quickly and precisely identifying rust diseases. RetinaNet ResNet-101, another one-stage object detection framework, leverages a feature pyramid network (FPN) and focal loss to address class imbalance. Utilizing the ResNet-101 backbone enhanced feature representation, bolstering the model's ability to detect rust diseases with heightened precision. YOLOv8S, YOLOv8L, YOLOv5L, and YOLOv5X represent variants of the YOLO framework, renowned for their simplicity and efficiency in object detection tasks. Diverse versions of YOLO were explored to evaluate their efficacy in rust disease identification and ascertain the most fitting variant for our specific application. Each model underwent a rigorous training process involving multiple iterations and encompassed several batches comprising many images. To ensure the robustness and effectiveness of our models, we employed transfer learning techniques, leveraging pre-trained models on large-scale datasets. Specifically, we fine-tuned each pre-trained model on our multi-level rust disease images dataset in Vicia faba L. crops. To customize the final layers of the network to our particular goal while preserving the characteristics acquired from the original dataset, we had to alter them. To improve the model's capacity for generalization and lessen overfitting, we employed data augmentation techniques to add random rotations, flips, and shifts to the dataset during training. We also used batch normalization and dropout methods to enhance the model's functionality and keep it from learning the noise in the training set. Each model underwent extensive hyperparameter tuning to optimize its performance. We experimented with various learning rates, batch sizes, and optimization algorithms to find the optimal configuration for each architectural variant. Our methodology encompasses a comprehensive approach to exploring and evaluating various architectural variants for rust disease detection. By leveraging transfer learning techniques, extensive data augmentation, and rigorous model evaluation, we aimed to develop robust and effective models capable of accurately identifying rust disease in agricultural settings.

We employed a meticulous assessment and comparison procedure to compare and select the most proficient model systematically. Performance metrics and evaluation criteria were carefully considered to accurately gauge the models' capabilities. Technical proficiency, precision, recall, and accuracy were among the key performance indicators utilized in this evaluation. Furthermore, the ability of each model to effectively capture and distinguish rust disease symptoms within the Vicia Faba L. pod images was of paramount importance. A standout architectural variant emerged through this rigorous evaluation process, displaying exceptional performance in rust disease detection. This model exhibited superior accuracy, showcasing its ability to effectively identify and classify rust disease symptoms within the Vicia faba L. crops. Its remarkable performance was a testament to its robust design, encompassing advanced algorithms and innovative methodologies. The optimal model selection culminated in our meticulous evaluation and analysis, driven by a quest for scientific excellence. It represents a significant milestone in rust disease detection in Vicia faba L. crops, offering valuable insights and potential solutions for agricultural practitioners and researchers alike. The findings of this study contribute to the body of knowledge surrounding rust disease detection, shedding light on the effectiveness of various architectural variants. Furthermore, the chosen model is a reliable tool for accurately identifying and classifying rust disease symptoms, facilitating improved management strategies and interventions for Vicia faba L. crops. Figure 3 depicts the architecture of the top-performing model based on a neural architecture search.

F. Experimental setting and configuration environment

Through an extensive training process, ten distinct architectural alternatives were developed, each offering unique capabilities. Among these variants were FCOS ResNet 50, YOLO NAS L, Faster R CNN ResNet 101, YOLO NAS M, RetinaNet ResNet 101, YOLOv8S, YOLOv8L, Faster R CNN ResNet 50, YOLOv5L, and YOLOv5X, each bringing its own set of advancements and features to the table: the training procedure utilized eight and ten batches, each containing multiple photographs. The models were trained on the experimental platform's Windows 10 Professional desktop PC. The system configuration was carefully selected for optimal compatibility and performance, incorporating Python version 3.10.12, CUDA version 11.8, and Torch version 2.0.1. The hardware setup included an Intel® Xeon® W-2223 CPU running at 3.6 GHz, 16 GB of RAM, and an NVIDIA GeForce Quadro P1000 graphics card, leveraging its processing power for efficient model training. Detailed information regarding the specific training strategies employed for each model iteration is provided in Table 1.

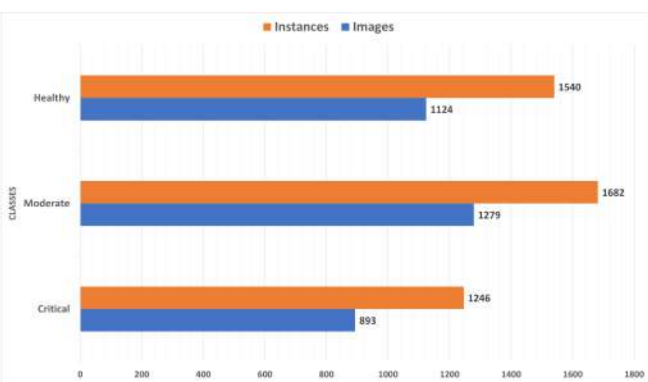

Figure 2. **Number of images and instances per class.**

This training phase's primary objective was to thoroughly evaluate the performance exhibited by each architectural variant. The ultimate aim was to identify the optimal model to detect rust disease in Vicia faba L. crops, showcasing superior precision and recall. A meticulous assessment and comparison procedure selected the architectural variant displaying the highest performance as the best model for rust disease detection. This selection process involved rigorous evaluation metrics, considering accuracy, precision, recall, and F1 score. The chosen model demonstrated an exceptional ability to identify and classify rust disease symptoms within Vicia faba L. crops. Its selection was based on robust scientific reasoning and precise evaluation methodologies, ensuring the model's reliability and effectiveness.

G. Model evaluation index

We used several evaluation measures to evaluate the performance of the suggested designs in this experimental investigation, including mAP, F1 Score, precision, and recall. These metrics are quantitative tools for assessing the performance and precision of the models used in our study. The percentage of accurately anticipated positive cases among all expected positive instances is known as precision. Precision and recall are combined to provide the F1 Score, a fair metric considering coverage and precision. The percentage of accurately anticipated positive cases among the actual positive instances in the dataset is measured by recall. Last, the average precision scores across several groups or categories are shown by mAP. By utilizing these evaluation metrics, we can provide a comprehensive assessment of the performance and capabilities of the proposed architectures.

1) $Precision = \frac{TP}{TP+FP} \times 100\%$
2) $Recall = \frac{TP}{TP+FN} \times 100\%$
3) $F1\_Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\%$
4) $mAP = \frac{1}{3} \sum_{i=1}^{3} AP \times 100\%$
5) with : $AP = \int_0^1 P(R)\, d(R) \times 100\%$

The assessment metrics in this situation might be clarified further: The number of items the model successfully recognized is represented by the True Positives (TP) count,

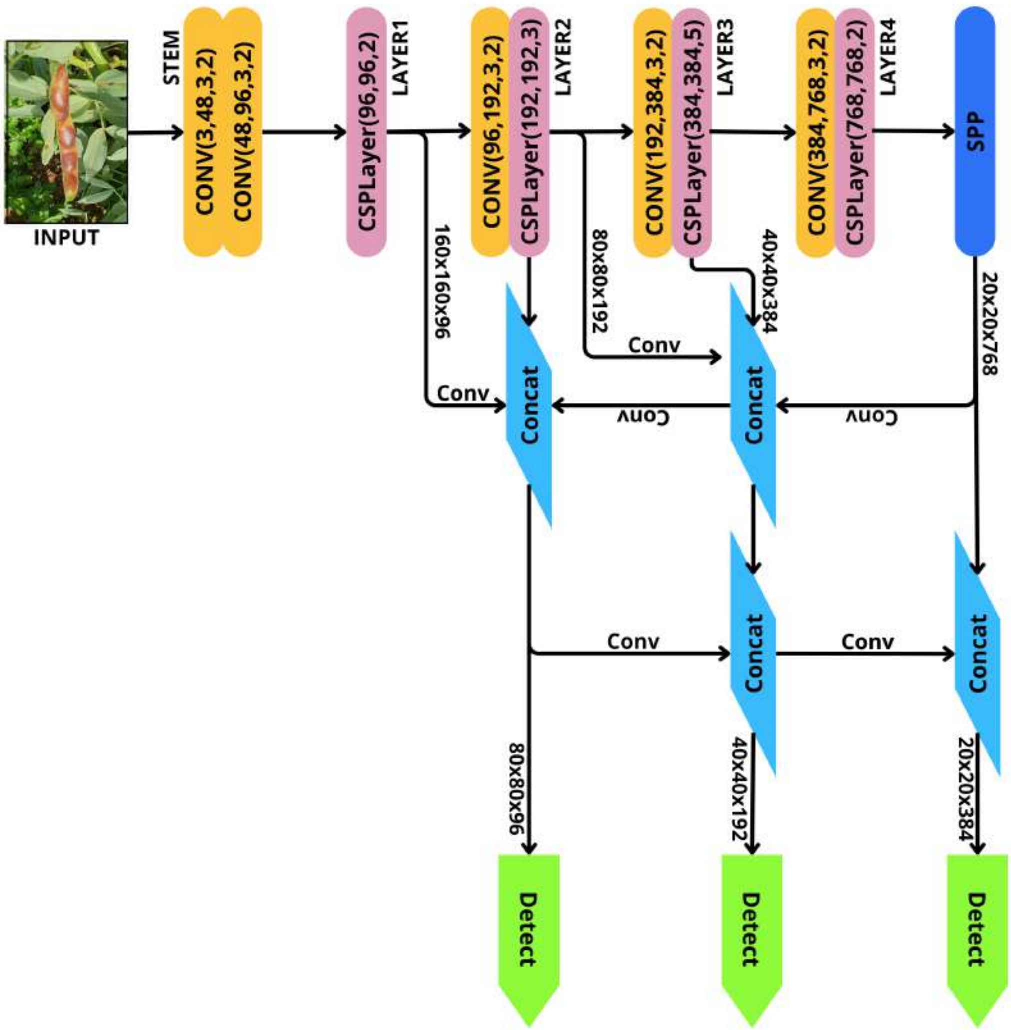

Figure 3. **Architecture of the optimal neural network model identified through neural architecture search.**

which stands for the count of correctly detected objects. The number of erroneously identified items, or False Positives (FP), is the number of times the model misclassified an object as falling into a specific category. False Negatives (FN) are the number of items the model did not detect and indicate cases in which an object that was supposed to be categorized was missed. The number of categories or classes that need to be classified is indicated by the variable ”3”.

4. Results and Discussion

A. *Metrics for Models Performance*

In this section, we rigorously evaluated our trained models using key metrics. These included the number of network parameters, Precision (P), F1 Score, Recall (R), mean Average Precision (mAP), and detection speed. Network parameters gauged model complexity and computational requirements; mAP assessed overall performance in precision and recall, measured accurate identification, minimized false positives, recall ensured comprehensive coverage of relevant objects, and detection speed focused on timely and efficient identification. We maintained consistency with an intersection over union (IOU) threshold of 0.7, ensuring fair model evaluation. This comprehensive approach allowed us to pinpoint model strengths and weaknesses, aiding in selecting the most effective solution for crop disease detection.

B. *Learning and Training Models*

Multiple architectures were developed for training, validation, and testing purposes. The training process utilized SGD and ADAM optimizers with a learning rate of 0.01 to 0.022. Several vital observations emerge when evaluating various stage detection models when considering performance metrics, model size, and speed. Percentages are integrated to underscore relative differences and trends among the models. YOLO-NASM, with a 15.6% reduction

TABLE I. **Training procedures implemented during the training process.**

Models	Param.(M)	Train	Val	Test	Opt	Lr	Size	Batch	Epochs
YOLOv5X	86.7	2197	824	275	SGD	0.010	640x640	8	40
YOLO-NASL	66.9	2197	824	275	Adam	0.022	640x640	8	50
RetinaNet ResNet-101	63.7	2197	824	275	SGD	0.012	800x800	10	50
Faster R-CNN ResNet-101	63.4	2197	824	275	SGD	0.015	800x800	10	50
YOLO-NASM	51.1	2197	824	275	Adam	0.022	640x640	8	50
YOLOv8L	43.7	2197	824	275	SGD	0.010	640x640	8	40
Faster R-CNN ResNet-50	41.5	2197	824	275	SGD	0.015	800x800	10	50
FCOS ResNet-50	32.2	2197	824	275	SGD	0.012	800x800	10	50
YOLOv8S	11.2	2197	824	275	SGD	0.010	640x640	8	40
YOLOv5S	07.2	2197	824	275	SGD	0.010	640x640	8	40

in model size (256 Mb to 196 Mb), exhibits substantial performance gains compared to YOLO-NASL, boasting an 8.6% increase in precision, a 2.0% boost in recall, and a noteworthy 4.2% gain in mAP. This suggests that YOLO-NASM not only optimizes resources but also enhances overall accuracy. Faster R-CNN ResNet-101, despite a 47.9% increase in model size compared to ResNet-50, does not consistently outperform its counterpart. The marginal improvement in precision (3.4%) is offset by lower recall, mAP, and speed, questioning the added complexity's justification for this specific task. In the YOLOv8 series, YOLOv8L, with a 274.5% larger model size than YOLOv5S, delivers superior performance across metrics. However, YOLOv5S is the fastest model, operating at 75.5% more frames per second (FPS). This highlights a balance between efficiency and real-time processing capabilities, with YOLOv5S offering a compelling trade-off. FCOS ResNet-50 outperforms RetinaNet ResNet-101 in speed (28.7% faster) despite the latter's 85.5% larger model size. This indicates that, for this task, the simpler model may be more practical, emphasizing the importance of evaluating efficiency against complexity. In the YOLOv5 series, YOLOv5X, with a 1071.4% larger model size compared to YOLOv5S, demonstrates superior performance but at the cost of speed, running at 72.9% fewer FPS. The decision between these models hinges on the specific application requirements, highlighting the necessity of balancing accuracy with real-time processing capabilities [25].

YOLO-NASM is a compelling choice in comparing object detection models, especially when considering the percentage differences compared to other evaluated models. YOLO-NASM is approximately 63.3% smaller than YOLO-NASL, indicating a significant reduction in size while maintaining or improving performance. Compared to Faster R-CNN ResNet-101, YOLO-NASM represents a reduction of approximately 19.1%, emphasizing its potential for efficiency gains, and about 15.7% larger than YOLOv5X but may provide enhanced performance. It demonstrates a favorable balance between compactness and performance, including an 8.6% increase in precision, a 2.0% boost in the recall, and a noteworthy 4.2% gain in mAP. This suggests that YOLO-NASM optimizes resources effectively, offering a more efficient and accurate solution for the given task. The reduction in model size is particularly advantageous for scenarios with resource constraints. YOLO-NASM appears to strike a favorable balance between model efficiency and performance, making it a strong contender among the evaluated models. The YOLO-NASM architecture was explicitly developed to illustrate the dynamic monitoring of training progress and model functionality, enabling an evaluation of the impact of training intervals on model performance. The results of this analysis are presented in Table 2. Notably, as the model underwent iterations from 0 to 14 epochs, its parameters exhibited significant changes. However, beyond the 30–50 epoch range, the model's performance reached a stable state. The evaluation index demonstrated stability after 30 to 50 model epochs, with the mean Average Precision (mAP) reaching approximately 95.1% before reaching a plateau (refer to Figure 7).

Figure 4 showcases our trained model's exceptional performance in identifying rust disease on Vicia faba L. pods. It accurately locates disease positions, classifies them based on pod state, and avoids missed detections and false positives. The model's proficiency extends to small and numerous targets, making it a reliable tool for effective disease management. Figure 5 presents a comprehensive analysis of the "Cls_Loss" function utilized by the best-trained model for identifying and classifying different levels of rust disease infection. The network is optimized through the stochastic gradient descent approach by adjusting its parameters during the learning process, decreasing the loss function value. This reduction in the loss function value demonstrates a strong correlation with other performance indicators, such as recall rate (refer to Figure 7a) and mean average precision (refer to Figure 7b). Additionally, the graphs demonstrate how the loss values quickly drop

TABLE II. **Comparative evaluation of trained models.**

Model	Weight. (Mb)	Prec.	Rec.	mAP50	F1 Score	mAP50@95	Speed (ms)	FPS
YOLO-NASL	256,0	82,10	94,82	90,90	87,87	79,99	5,67	177
Faster R-CNN ResNet-101	243,0	73,74	79,20	85,14	76,37	65,02	7,82	128
RetinaNet ResNet-101	228,0	73,74	86,40	85,15	79,24	67,75	6,44	156
YOLO-NASM	196,0	84,80	96,96	94,10	90,84	85,13	4,84	207
Faster R-CNN ResNet-50	167,0	75,51	82,80	81,40	78,07	64,29	5,06	198
YOLOv5X	166,0	88,61	87,64	88,72	88,09	69,67	13,81	73
FCOS ResNet-50	123.6	74,88	88,51	82,34	81,12	66,79	5,73	175
YOLOv8L	83,7	95,13	89,52	93,73	92,21	76,53	10,13	100
YOLOv8S	22,3	89,92	90,30	91,81	90,09	72,72	4,80	209
YOLOv5S	14,0	75,25	68,91	74,87	71,91	45,61	2,82	358

Figure 4. **Detection results from the proposed model: illustrative examples.**

during the first few epochs before stabilizing after the model achieves steady performance. This suggests that the model is rapidly gaining knowledge from the data and reaching a sound conclusion. The classification loss plays a crucial role in assessing the algorithm's ability to predict specific item categories accurately. As the value of the loss function decreases, the classification accuracy tends to increase. Hence, minimizing the loss function value is of utmost importance in achieving higher accuracy levels.

C. Memory

The proportion of system memory used during the various models' training times is represented by the curves in Figure 6. Memory usage curves are plots to show the amount of memory a deep learning model consumes during training or inference. To understand the efficiency and scalability of the model, as well as to identify potential bottlenecks or errors. We chose four of the ten models- Faster R-CNN ResNet-50, RetinaNet ResNet-101, YOLO-NASM, and FCOS ResNet-50, to compare from a system memory point of view and found that the YOLO-NASM model curve is the best and the most stable, showing no sudden spikes or dips, does not change radically over time, indicates that the model uses memory consistently and that there are no memory leaks or fragmentation[26]. We can say that the model uses neither too much nor too little memory and that there are no problems of memory overflow or under-utilization.

Initially, there is a sharp increase in memory usage, from around 40% to just over 55% in a short time (from 0 to about 40 on the x-axis). This rapid increase suggests that a process or set of methods has been launched and has consumed a significant amount of memory. After this initial peak, memory usage stabilizes, fluctuating between 55% and 60%. This pattern could be due to regular processing tasks occurring at regular intervals, causing slight fluctuations in memory usage. We want to choose a deep learning model with low peak, mean, and variance memory usage while maintaining high accuracy and speed. However, these goals may have trade-offs or constraints, so we need to balance them according to our specific needs and preferences.

D. Performance Evaluation: Precision, Recall, inference speed, mAP, and F1 Score

Figures 6a and 6b comprehensively analyze recall and precision performance for various models under different conditions in object detection tasks. Recall, which measures a model's ability to detect all relevant objects in an image, is directly linked to effectiveness, with higher recall indicating fewer missed objects. Conversely, precision measures a model's ability to avoid false positives, ensuring that only relevant objects are identified. Higher precision values indicate better model performance. In

Figure 7a, the radar chart focuses on recall performance. Ten models are compared across three Vicia faba L. pod conditions: Healthy, Moderate, and Critical. The proposed YOLO-NASM model consistently exhibits the highest recall performance, making it the most effective and reliable model for detecting infection levels. RetinaNet ResNet-101 and YOLOv8L demonstrate similar recall performance under Healthy and Moderate conditions. However, Faster R-CNN ResNet-101 experiences a notable decrease in recall under the Critical condition. Faster R-CNN ResNet-101 and YOLOv5S exhibit the lowest recall performance across all conditions, indicating limited effectiveness in object detection, particularly under critical conditions. This graph enables users to select the most suitable model based on recall performance and robustness. Figure 7b shifts the focus to precision performance. The radar chart depicts precision values for each model under the same three conditions. YOLOv8L consistently delivers the highest precision across all conditions, accurately identifying relevant objects and minimizing false positives. This high precision is crucial for precision agriculture, where accurately identifying healthy and diseased crops can prevent unnecessary treatments and optimize resource use. YOLOv5X consistently outperforms other models, demonstrating superior precision regardless of image quality. This robustness makes it suitable for field conditions where image quality can vary, ensuring reliable disease detection and management. YOLO-NASL and RetinaNet ResNet-101 exhibit similar precision performance under Healthy and Moderate conditions. However, YOLO-NASL experiences a significant drop in precision under critical conditions, suggesting its vulnerability to image degradation and limited performance in challenging environments [27]. This indicates that YOLO-NASL may be less reliable for detecting severe diseases, potentially leading to missed detections and delayed interventions. In contrast, RetinaNet ResNet-101 showcases more resilience, maintaining higher precision even under Critical conditions.

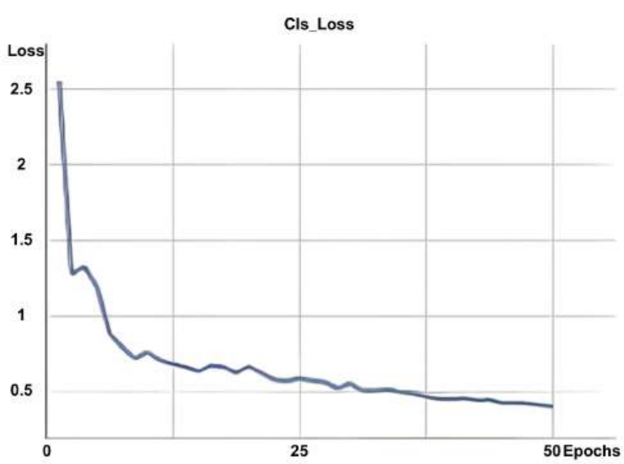

Figure 5. **The 'cls_loss' function for identifying rust disease infection levels.**

This resilience makes RetinaNet ResNet-101 a better choice for environments with higher disease severity, ensuring timely and accurate identification of critical cases. Faster R-CNN ResNet-101 exhibits lower precision performance across all conditions. This may limit its effectiveness in precision agriculture applications where high precision is needed to minimize false positives and ensure accurate disease identification. At the same time, FCOS ResNet-50 displays the poorest precision, generating numerous false positives, especially under Moderate conditions. This high rate of false positives can lead to unnecessary treatments and increased costs, making FCOS ResNet-50 less suitable for practical agricultural applications. These precision insights aid users select models that align with their application requirements. For instance, models like YOLOv8L and YOLOv5X are recommended for their high precision and robustness across various conditions, ensuring reliable performance in real-world agricultural settings. On the other hand, models like YOLO-NASL may require further optimization for critical situations. In contrast, models with lower precision, such as FCOS ResNet-50, might need to be more practical for highly accurate applications. Understanding the precision performance of different models allows for more informed decisions in deploying these technologies for crop disease management [28]. High-precision models can significantly enhance the effectiveness of precision agriculture by ensuring accurate disease detection, optimizing resource allocation, and ultimately improving crop health and yield. Figures 6a and 6b provide valuable information on different models' recall and precision performance under varying conditions, facilitating informed decision-making regarding model selection.

The analysis of Figure 8, which showcases the inference speed of various models measured in milliseconds, provides valuable insights into the performance of these models. Inference speed, representing the time a model takes to process input and generate output, is critical, with lower speeds indicating faster models. The graph utilizes color coding to represent health status, with green showing healthy, orange for moderate, and red for critical, based on specific thresholds. One notable observation is the significant variability in inference speeds among the different models. Models like YOLOv5S and Faster R-CNN ResNet-50 exhibit high speeds, while YOLOv5X performs slower. This observation highlights the correlation between detection and inference speed, where faster models are generally healthier and more suitable for real-time applications. On the other hand, slower models may struggle to meet performance requirements. Model architecture also plays a role in determining inference speed. This distinction may be attributed to design differences, such as the presence of convolutional layers and attention mechanisms. Several factors influencing the data must be considered. Fluctuations in input data types and sizes have a discernible impact on the speed and accuracy of the models. Given the need for tailored resizing or cropping methodologies for each model, selecting the right hardware is crucial for efficient

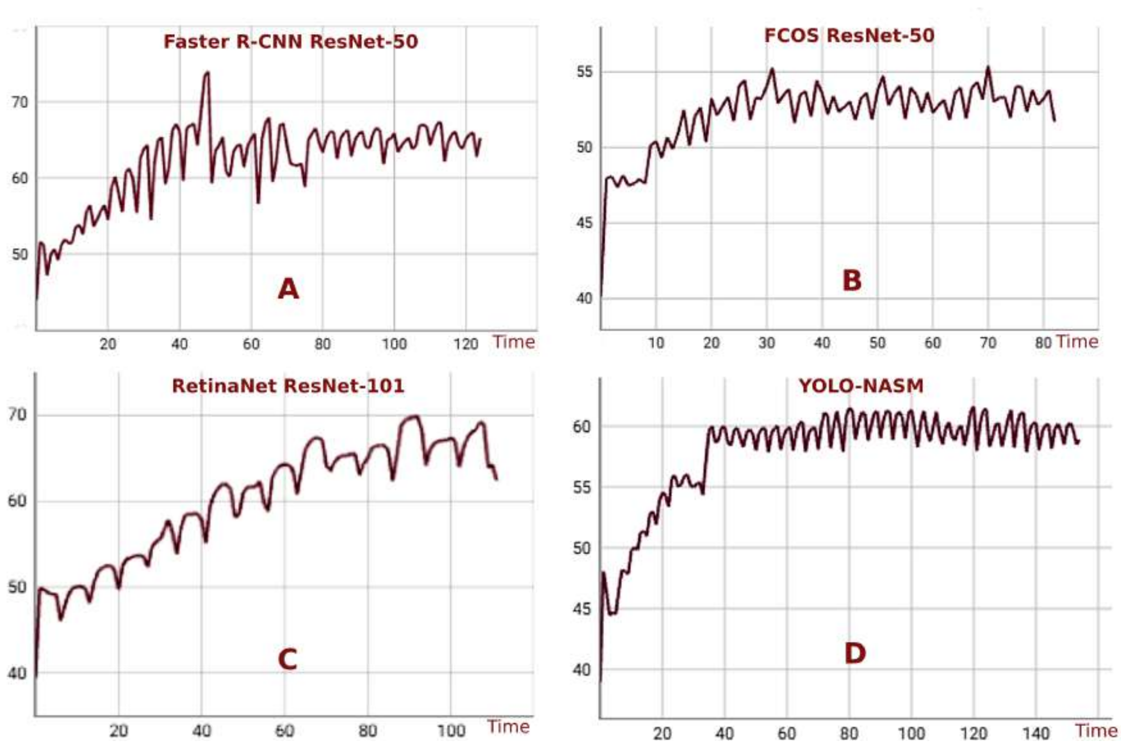

Figure 6. **System memory usage curves for various models during training.**

model training. We chose the NVIDIA Quadro P1000 for its exceptional computational capabilities, including robust CUDA cores and high memory bandwidth. These features enable accelerated training of complex neural networks, leading to faster convergence and reduced training times. The Quadro P1000's ample memory capacity is essential for handling large-scale datasets, improving model generalization and accuracy. Its seamless integration with deep learning frameworks ensures optimized performance and compatibility, facilitating smooth training processes and advanced techniques implementation. By carefully evaluating the results and considering the delicate balance between accuracy and detection speed, the proposed YOLO-NASM model emerges as a recommendation due to its exceptional speed/accuracy ratio. This nuanced analysis equips us with the necessary insights to make well-informed decisions when deploying machine learning models across diverse applications. By considering factors such as inference speed, model architecture, and the specific requirements of the task at hand, users can select the most suitable model that strikes the right balance between speed and accuracy.

The F1 scores of six models across three classes—Healthy, Moderate, and Critical—are depicted in Figure 9a. The F1 score varies significantly between classes, indicating this metric's sensitivity to each class's complexity and characteristics. For instance, the Healthy class may possess more distinct and precise features than the Moderate and Critical classes, making it easier for the models to detect and classify. On the other hand, the Critical class may exhibit more overlapping or ambiguous features, posing challenges for accurate identification and differentiation by the models [29]. Each model demonstrates varying performance across the different classes. YOLONASM and YOLOv8L consistently improve the F1 score across all classes, while YOLOv5X experiences a significant drop in the F1 score for the Moderate and Critical courses. Additionally, Faster R-CNN R-101 performs poorly in the Moderate class but emerges as one of the top performers in the Critical class. These observations suggest that Faster R-CNN R-101 may be more susceptible to errors or overfitting in moderate scenarios than other models. The evaluation of F1 scores highlights the importance of assessing the models' performance for each class independently rather than relying solely on the overall score. This approach provides more comprehensive insights. These findings have significant implications for precision agriculture and crop disease management strategies. The ability of models like YOLONASM and YOLOv8L to maintain high F1 scores across all classes suggests they are reliable tools for identifying and classifying different levels of crop health. This reliability is crucial for early detection and intervention, allowing farmers to apply targeted treatments and prevent the spread of disease, ultimately optimizing resource use and minimizing crop loss. In contrast, the variable performance of YOLOv5X and Faster R-CNN R-101 indicates that these models may require further tuning

Figure 7. **Training Progress: (a) Recall, (b) Precision.**

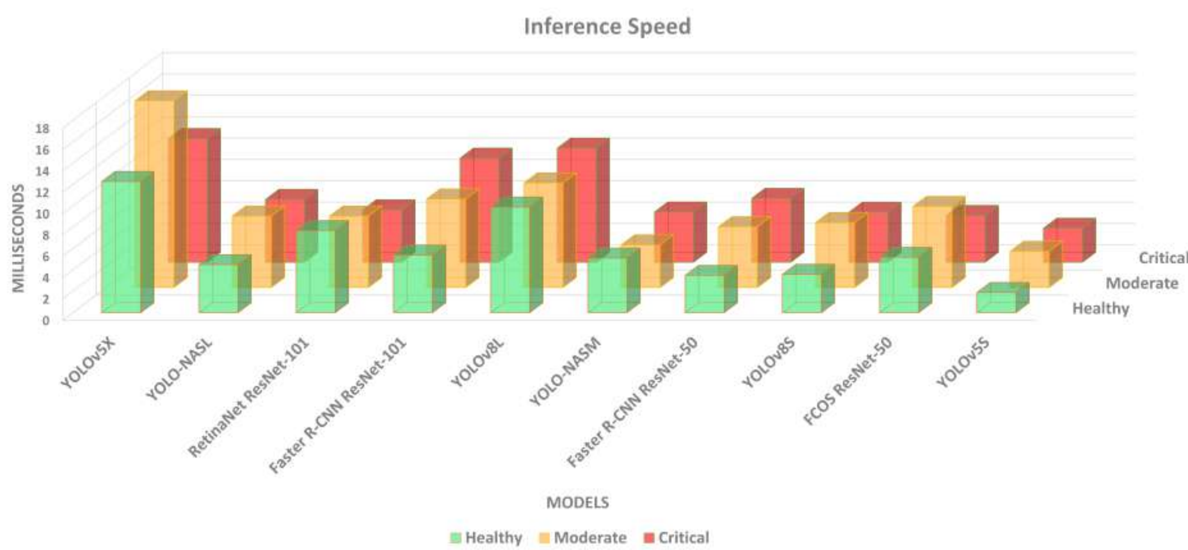

Figure 8. **Analysis of inference speed performance in multi-level infection detection.**

or hybrid approaches to effectively handle the complexity of Moderate and Critical cases. Understanding these nuances can guide the development of more robust models or integrate multiple models to cover their weaknesses. Furthermore, the insights gained from evaluating F1 scores by class can help refine disease management protocols. For example, recognizing that the Critical class poses challenges for specific models can lead to focused research on improving detection algorithms for severely affected crops, ensuring timely and accurate responses. This approach provides more comprehensive insights into the models' strengths, weaknesses, opportunities, and challenges. Based on previous findings, the model with the highest performance, as depicted in Figure 9a, belongs to the "Critical" class. This suggests that YOLO-NASM, compared to the other models, is more specialized or finely tuned for complex scenarios. Figure 9b illustrates multi-level infection detection models' Mean Average Precision (mAP) performance across the Healthy, Moderate, and Critical classes. Retina Resnet-50 and Faster RCNN Resnet-101 exhibit modest performance, achieving scores of approximately 86% and 87%, respectively, in both the 'Healthy' and 'Critical' classes. In contrast, FCOS Resnet-50 consistently performs at a lower overall mAP of 77%. However, it shows improvement in the "Critical" class, attaining scores between 80% and 84%. These findings indicate that FCOS Resnet-50 may struggle to distinguish between healthy, moderate, and

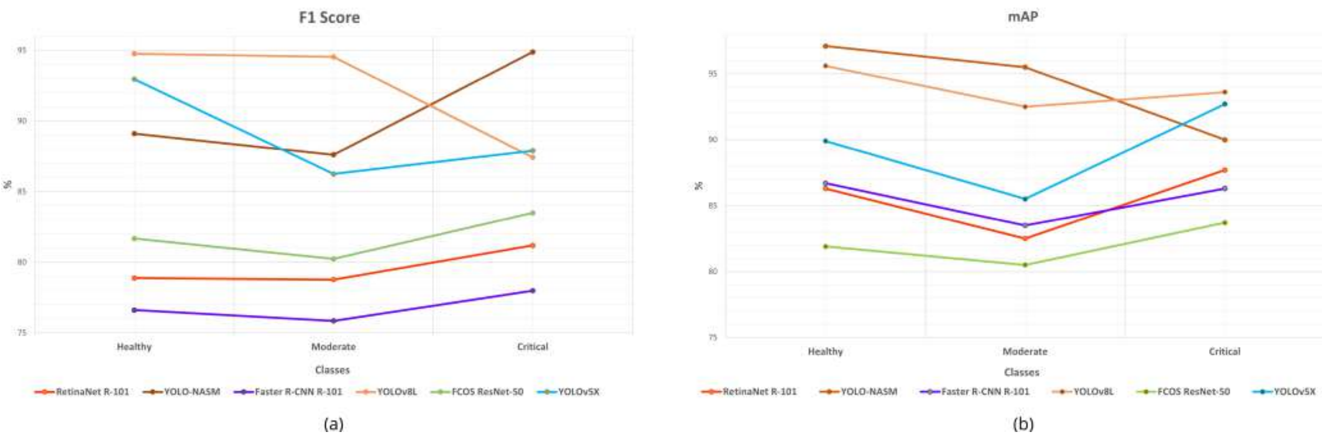

Figure 9. **Training progress evaluation: (a) F1 score, (b) mAP.**

critical conditions. Conversely, the model based on neural architecture search consistently delivers strong performance across all three classes, with an mAP above 90%. It excels in the "Healthy" class, achieving the highest score of 96%. These outcomes underline the robustness and reliability of the proposed model for multi-level disease detection, positioning it as a preferred choice for general use.

The proposed model based on neural architecture search demonstrated exceptional performance with an important mean Average Precision. This accuracy in disease diagnosis highlights the practical feasibility of implementing deep learning-based approaches in agriculture [30]. The developed models exhibit notable strengths, including high relevance, real-time efficiency, cost-effectiveness, and adaptability to diverse agricultural contexts. These strengths have the potential to revolutionize agricultural practices, increase crop output, and enhance food security. We will focus on specialized disease identification in Vicia faba L. and other crops, refining classification algorithms for greater specificity. By incorporating additional data sources such as crop type and environmental variables, we can further enhance the accuracy of disease identification. Furthermore, we plan to enrich the dataset with new classes that capture essential information on crop infections, disease evolution, spread, and severity [31].

5. Conclusions and Future Work

This study represents a significant leap forward in detecting agricultural disease, specifically identifying rust in Vicia faba L. pods. Through a comprehensive evaluation and comparison of advanced object identification models, such as Faster R-CNN ResNet-50, FCOS ResNet-50, RetinaNet ResNet-101, YOLOv8S, YOLOv8L, Faster R-CNN ResNet-101, YOLO-NASL, YOLO-NASM, YOLOv5L, and YOLOv5X, we have achieved remarkable progress. The tailored dataset generated in this study holds immense value for practical application in natural agricultural settings, offering valuable insights into the cutting-edge field of disease identification. The compiled dataset serves as an invaluable resource for future research, aiding in identifying agricultural diseases and facilitating the practical training of disease detection models. Notably, our model based on neural architecture search (NAS) exhibited exceptional performance, achieving a mean Average Precision (mAP) of 94.10%. This outstanding performance in precise disease diagnosis underscores the practical viability of implementing deep learning-based approaches in agriculture. Our developed models possess notable strengths, including high relevance, real-time efficiency, cost-effectiveness, and adaptability to diverse agricultural contexts. These strengths have the potential to revolutionize agricultural practices, boost crop output, and enhance food security. Moreover, they contribute to developing more effective disease monitoring systems for the benefit of society at large. We will focus on specialized disease identification in Vicia faba L. and other crops, refining classification algorithms for greater specificity. By incorporating additional data sources such as crop type and environmental variables, we can further enhance the accuracy of disease identification. Additionally, we plan to enrich the dataset with new classes that capture vital information on crop infections, disease evolution, spread, and severity. This study marks a significant milestone in leveraging deep learning for agricultural disease detection. It demonstrates the immense potential of these advanced models to revolutionize farming practices, improve crop management, and contribute to global food security. By continuing our research and dataset enrichment efforts, we aim to make even more significant strides in disease identification and pave the way for a more resilient and productive agricultural sector.

References

[1] H. Slimani, J. El Mhamdi, and A. Jilbab, "Assessing the advancement of artificial intelligence and drones' integration in agriculture through a bibliometric study," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 1, pp. 878–890, 2024.

[2] H. Yu, F. Yang, C. Hu, X. Yang, A. Zheng, Y. Wang, Y. Tang, Y. He, and M. Lv, "Production status and research advancement on root rot disease of faba bean (vicia faba l.) in china," *Frontiers in Plant Science*, vol. 14, p. 1165658, 2023.

[3] M. A. Goralski and T. K. Tan, "Artificial intelligence and sustainable development," *The International Journal of Management Education*, vol. 18, no. 1, p. 100330, 2020.

[4] A.-K. Mahlein, E.-C. Oerke, U. Steiner, and H.-W. Dehne, "Recent advances in sensing plant diseases for precision crop protection," *European Journal of Plant Pathology*, vol. 133, pp. 197–209, 2012.

[5] A. Bouguettaya, H. Zarzour, A. Kechida, and A. M. Taberkit, "A survey on deep learning-based identification of plant and crop diseases from uav-based aerial images," *Cluster Computing*, vol. 26, no. 2, pp. 1297–1317, 2023.

[6] D. M. Rajesh, K. Bhuvanesh, C. V. Chakravarthi, C. S. Gopi, M. A. Hussain *et al.*, "A real-time precision monitoring and detection system for rice plant diseases using machine learning approach," *Journal of Survey in Fisheries Sciences*, vol. 10, no. 3, pp. 289–298, 2023.

[7] P. Venkatamohan, D. Nagpal, M. K. Bheemireddy, and R. K. Ketharapu, "Identification of paddy plant diseases using artificial intelligence (ai)," in *2023 4th International Conference on Intelligent Engineering and Management (ICIEM)*. IEEE, 2023, pp. 1–6.

[8] P. Nayar, S. Chhibber, and A. K. Dubey, "Detection of plant disease and pests using coherent deep learning algorithms," in *2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*. IEEE, 2023, pp. 8–12.

[9] B. P. Zen, D. C. Fransisca *et al.*, "Applications for detecting plant diseases based on artificial intelligence," *Sinkron: jurnal dan penelitian teknik informatika*, vol. 7, no. 4, pp. 2537–2546, 2022.

[10] L. Kamelia, T. K. Abdul Rahman, R. R. Nurmalasari, and K. K. Hamdani, "Citrus tree nutrient deficiency classification: A comparative study of ann and svm using colour-texture features in leaf images," *International Journal of Computing and Digital Systems*, vol. 14, no. 1, pp. 1–13, 2023.

[11] A. Y. Krishna, S. T. Sri, V. Sravya, P. S. Praneetha, B. V. Vardhan *et al.*, "Disease recognition of crops using resnet and mdfc-resnet," in *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*. IEEE, 2023, pp. 738–744.

[12] H. Slimani, J. El Mhamdi, and A. Jilbab, "Artificial intelligence-based detection of fava bean rust disease in agricultural settings: an innovative approach," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 6, 2023.

[13] D. Zhu, J. Tan, C. Wu, K. Yung, and A. W. Ip, "Crop disease identification by fusing multiscale convolution and vision transformer," *Sensors*, vol. 23, no. 13, p. 6015, 2023.

[14] O. Iparraguirre-Villanueva, V. Guevara-Ponce, C. Torres-Ceclén, J. Ruiz-Alvarado, G. Castro-Leon, O. Roque-Paredes, J. Zapata-Paulini, and M. Cabanillas-Carbonell, "Disease identification in crop plants based on convolutional neural networks," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 3, 2023.

[15] W. Liu, G. Wu, F. Ren, and X. Kang, "Dff-resnet: An insect pest recognition model based on residual networks," *Big Data Mining and Analytics*, vol. 3, no. 4, pp. 300–310, 2020.

[16] S. Verma, P. Kumar, and J. P. Singh, "A meta-learning framework for recommending cnn models for plant disease identification tasks," *Computers and Electronics in Agriculture*, vol. 207, p. 107708, 2023.

[17] A. O. Agbaje and J. Tian, "Plant disease recognition using transfer learning and evolutionary algorithms," in *TENCON 2022-2022 IEEE Region 10 Conference (TENCON)*. IEEE, 2022, pp. 1–7.

[18] S. Tiwari, S. Kumar, S. Tyagi, and M. Poonia, "Crop recommendation using machine learning and plant disease identification using cnn and transfer-learning approach," in *2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*. IEEE, 2022, pp. 1–4.

[19] S. Bondre and D. Patil, "Recent advances in agricultural disease image recognition technologies: A review," *Concurrency and Computation: Practice and Experience*, vol. 35, no. 9, p. e7644, 2023.

[20] A. Hessane, A. El Youssefi, Y. Farhaoui, B. Aghoutane, and F. Amounas, "A machine learning based framework for a stage-wise classification of date palm white scale disease," *Big Data Mining and Analytics*, vol. 6, no. 3, pp. 263–272, 2023.

[21] H. Slimani, J. El Mhamdi, and A. Jilbab, "Drone-assisted plant disease identification using artificial intelligence: A critical review," *International Journal of Computing and Digital Systems*, vol. 14, no. 1, pp. 10 433–10 446, 2023.

[22] Y. J. Hwang, G. H. Kim, M. J. Kim, and K. W. Nam, "Deep learning-based monitoring technique for real-time intravenous medication bag status," *Biomedical Engineering Letters*, pp. 1–10, 2023.

[23] S.-A. Rebuffi, S. Gowal, D. A. Calian, F. Stimberg, O. Wiles, and T. Mann, "Fixing data augmentation to improve adversarial robustness," *arXiv preprint arXiv:2103.01946*, 2021.

[24] P. Bachhal, V. Kukreja, and S. Ahuja, "Real-time disease detection system for maize plants using deep convolutional neural networks," *International Journal of Computing and Digital Systems*, vol. 14, no. 1, pp. 10 263–10 275, 2023.

[25] D. Dharrao, S. Deokate, S. Rajput, and S. Ambala, "Hyperspectral image (hsi)-assisted multi-level crop classification with spinalnet-enabled fractional light spectrum optimizer," *Biomedical Engineering: Applications, Basis and Communications*, vol. 36, no. 02, p. 2450004, 2024.

[26] Z. Liu, Y. Zheng, X. Du, Z. Hu, W. Ding, Y. Miao, and Z. Zheng, "Taxonomy of aging-related bugs in deep learning libraries," in *2022 IEEE 33rd International Symposium on Software Reliability Engineering (ISSRE)*. IEEE, 2022, pp. 423–434.

[27] M. Chilakalapudi and S. Jayachandran, "Optimized deep learning network for plant leaf disease segmentation and multi-classification using leaf images," *Network: Computation in Neural Systems*, pp. 1–34, 2024.

[28] H. Qu, C. Zheng, H. Ji, R. Huang, D. Wei, S. Annis, and F. Drummond, "A deep multi-task learning approach to identifying mummy berry infection sites, the disease stage, and severity," *Frontiers in Plant Science*, vol. 15, p. 1340884, 2024.

[29] H. Slimani, J. El Mhamdi, and A. Jilbab, "Advancing disease identification in fava bean crops: A novel deep learning solution integrating yolo-nas for precise rust," *Journal of Intelligent & Fuzzy Systems*, no. Preprint, pp. 3475–3489, 2024.

[30] S. U. Islam, S. Zaib, G. Ferraioli, V. Pascazio, G. Schirinzi, and G. Husnain, "Enhanced deep learning architecture for rapid and accurate tomato plant disease diagnosis," *AgriEngineering*, vol. 6, no. 1, pp. 375–395, 2024.

[31] J. Hu, X. Jiang, J. Gao, and X. Yu, "Lfmnet: a lightweight model for identifying leaf diseases of maize with high similarity," *Frontiers in Plant Science*, vol. 15, p. 1368697, 2024.

Hicham Slimani obtained a research master's degree in electrical engineering from Mohammed V University in Rabat, Morocco's ”École Nationale Superieur des Arts et Métiers” (ENSAM Ex-ENSET). Currently enrolled in the ”École Nationale Supérieure d'Arts et Métiers (ENSAM)” Ph.D. program, he works in the research group ”Electronic Systems, Sensors and Nanobiotechnologies.” He has had over 4 research articles published in reputable national and international journals. His research interests include Wireless Sensor Networks, Artificial Intelligence and Data Communication, and the Internet of Things. He is working as a Professor of science engineering at the Ministry of National Education in Morocco.

Jamal El Mhamdi is a professor of Electrical Engineering at the esteemed ”École Nationale Supérieure des Arts et Métiers” (ENSAM, formerly known as ENSET), Mohammed V University in Rabat, Morocco; he brings a wealth of expertise and experience to his academic role. He has a Ph.D. in electrical, electronic, and telecommunications engineering from the University of Rennes, France, in 1988. He has had over 50 research articles published in reputable national and international journals and conferences, and he has established himself as a leading authority in Signal Processing, Communication Systems, and Artificial Intelligence.

Abdelilah Jilbab is a professor of Electrical Engineering at ”École Nationale Superieur des Arts et Métiers” (ENSAM, Ex-ENSET), Mohammed V University in Rabat, Morocco. He has made significant contributions with a Ph.D. in Computer and Telecommunication from Mohammed V-Agdal University. He has published more than 100 research papers in reputed international and national journals. His research primarily focuses on Image Processing, Engineering, Information Systems, and Artificial Intelligence. His profound research interests and contributions have significantly impacted these domains, paving the way for technological advancements.