Received 29 October 2025, accepted 02 April 2026, date of publication 30 September 2026
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.25.3.4857

Pepper Leaf Diseases Detection Using Hybrid Techniques with Deep Learning

DESALEGN TASEW DEGAGA[1], ARULMURUGAN RAMU[2], K MARIYAPPAN[3], DIRIBA GICHILE RUNDASA[4]

[1]Department of Computer Science, Jimma University, Agaro, Ethiopia
[2]Heriot-Watt International faculty, K.Zhubanov University, Aktobe, Kazakhstan
[3]Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai, India
[4]Department of Computer Science and Engineering, Adama Science and Technology University, Ethiopia

Corresponding author: Desalegn Tasew Degaga (e-mail: dtddegaga@gmail.com).

ABSTRACT Plant disease detection in agriculture is a significant challenge, with potential benefits in monitoring crop fields and identifying symptoms on plant leaves. However, most systems underperform due to varying leaf disease patterns, limiting detection precision. Automating the recognition of leaf diseases can boost farming production. To solve this problem, we proposed a novel state-of-the-art hybrid deep learning model. Our objectives are to construct a better pepper leaf disease detection method and model using hybrid deep learning techniques based on CNN-CAE and CNN-HOG models that have been trained, validated, and tested. The performance of each model for detecting pepper leaf disease was assessed. The hybrid CNN-HOG and CNN-CAE model was giving us good and promising results such as accuracy, precision, recall, Fl-Score performance metrics. The findings suggest that hybrid deep learning can detect pepper leaf illness better. Farmers and agriculture specialists can identify and manage pepper leaf illnesses faster with such approaches, improving pepper output and sustainability.

KEYWORDS Leaf-diseases, Image-processing, Convenets, Histogram of Oriented Gradients (HOG), hybrid techniques.

I. INTRODUCTION

With its vast potential for crop development, Ethiopia stands out as a country in Africa that offers a wide range of traditional medicinal and everyday use plants [1]. Ethiopia offers a diverse range of traditional medicinal and every day-use plants, making it a promising crop development destination in Africa [2]. Ethiopia's agricultural sector contributes significantly to GDP, employment, and foreign exchange earnings [3]. Agricultural products are vital to developing countries' economies. Early detection of plant illness can prevent disease spread and reduce loss [1], [4] Ethiopian farmers grow pepper as a major crop for both domestic and commercial purposes [5]. However, variables such as diseases that injure the plant during its growth have a significant impact on the crop's quality and productivity [5]. This crop produces little in areas where it is grown nationally. Despite its limited production, Capsicum annuum holds significant potential for increasing agricultural productivity and generating income for local farmers. Efforts to improve cultivation techniques and market access could enhance its yield and profitability across the country. In Ethiopia, pepper accounts for 67.98% of total vegetable production. According to financial records from 2012, Ethiopia's economy is primarily driven by agriculture [3], [5].

The hot pepper, Capsicum annuum L., is a member of the Solanaceae family, which also includes tomato, potato, tobacco, and petunia. It is known by various names in different Asian countries, including "ema" in Bhutan, "la-jiao" in China, "cabe" in Indonesia, "prik" in Thailand, and "chilli" in India. The early Aztecs of Mexico referred to it as "chilli," and this term is now widely used globally, with minor variations. Hot peppers, which originated in Central and South America, were introduced to Asia and Africa in 1493 and are now grown all over the world except in Antarctica [2]. They are the world's second-largest vegetable crop, trailing only tomatoes, and the most widely used spice for flavoring and coloring food. In Ethiopia, they are used as both a vegetable and a spice, with "karia [7]. Pepper grows best in climates ranging from 15 to 400 degrees Celsius, but is sensitive to various illnesses and produces low yearly yields

due to factors like low-yielding crop types, drought, insect pests, illnesses, and poor cultural practices. Diseases are the most concerning of these variables, as they are caused by a wide variety of fungi, bacteria, and viruses. Researchers detected over 34 pepper infections, which resulted in a 33% decrease in crop productivity. Plant diseases are costly and diminish agricultural productivity. Infections with bacteria, viruses, or fungi can produce pigmentation in leaves and stems. The virus produced marketable fruit losses ranging from 60 to 100%, whereas pepper anthracnose resulted in losses of up to 100%. Bacterial infections induced by the seed-borne disease Xanthomonas campestris pyvesicatoria can cause significant plant defoliation. This can lead to reduced crop output and lower-quality harvested fruit, especially if developing fruits are badly injured [1], [2], [3].

Figure 1. Sample Hot pepper plants in study area (own data, 2024)

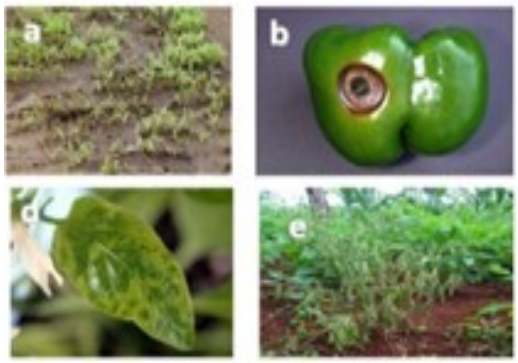

Figure 2. Some Major Pepper diseases in Ethiopia (Agenghu et al. 2023)

Early crop growth issues typically show up most directly as pepper leaf infections. For the purpose of quickly identifying development problems and facilitating precise prevention and control measures, reliable and quick identification of pepper leaf diseases is crucial. The traditional method for identifying pepper leaves is subjective, time-consuming, and expensive since it relies on visual inspection and human experience. Therefore, the development of an accurate, quick, and practical method for identifying pepper leaf infections is urgently needed [1] [2]. However advanced technology can overcome this issue through image processing in mass cultivating pepper crops. Early and accurate identification of such diseases is required to prevent the diseases accordingly before loss. Automatic detection will not only help in reducing time and effort but will also help in detecting disease at an early stage, as soon as it will start appearing on plant leaves. The advanced solution is automating the disease identification and classification process using different algorithms or approaches. Many approaches for detecting and classifying plant diseases have been proposed by researchers in the fields of computer vision and machine learning. ML and DL techniques have played an important role in agricultural sectors for the prediction, classification, and detection of plant disease and provide non-destructive, low cost, fast, and reliable means of plant disease detection [1]. DL and image processing have been investigated in the past for the purpose of classifying pepper plant leaf diseases. Image processing techniques have made it possible to identify many diseases automatically.

According to artificial intelligence theory, deep learning can solve complex visual tasks. Computer vision, Image processing, Machine learning (ML) algorithms are designed to diagnose plant infections or diseases by analyzing the distinctive color, shape, size, and look of the leaves [4]. Deep learning (DL) is a subset of machine learning (ML) algorithms that excel at image processing and data analysis. Deep learning methods surpass machine learning as image classifiers with CNN being the best. The CNN is widely recognized as the fine stand most ineffective image classier on both small and large data sets, and it serves as the foundation for all deep learning models [5, 6]. It is a basic deep learning model used for the classification of images according to some patterns or features. One of the best properties is that CNN trained in a supervised manner with the help of existing data and have the ability to automatically extract features from input photos, enabling them to accurately identify plant diseases. In particular CNN, as one of the most promising techniques, has been successfully used in crop leaf disease recognition [1]. A variety of convolutional neural network models are closely integrated with recognizing crop leaf diseases. The [1] authors performed color transformation to covert RGB image into HSV color space and also used threshold segmentation techniques, which is a similarity-based segmentation approach that segments the image based on image space region to detect diseased area, and then GLCM extracted feature is sent to the neural network. Proposed [1]. Pepper bell leaf disease detection and classification using an optimized convolutional neural network.

The author explores, they have applied preprocessing techniques like resizing, edge detection, and color conversion. They have used 1478 datasets for the healthy class and 979 for the bacterial class. Finally, detecting pepper leaf into bacteria and healthy leaves. In the plant disease identification process identifying infected and healthy is necessary but, rather than specifying healthy and unhealthy identifying the specific disease that affects the plant is preferable to get fast treatment. The authors' limitation is that validation and testing accuracy were not included. One limitation of the study was the effort made to distinguish only binary bell pepper leaves from healthy and bacteria-spot disease In general; most authors have used pre-trained models, machine learning algorithms, and binary classification techniques for infected or not infected pepper images. However, they have used an unbalanced dataset and they identify the disease of single species of pepper plant i.e. bell pepper and used CNN as both a feature extraction and classification. The researchers sum up the success of the deep learning network model for two big

reasons, the first reason is the deep learning model makes use of high power computation-based graphical processing units (GPUs) which helps a lot in the smooth training process of learning classifiers. Secondly, make use of raw data as a whole and extracting significant features itself helps the deep learning model a lot in the training process. The model trains itself on all significant features and gives promising results.

Chemical analysis, eye inspection, and laboratory testing are all expensive and time-consuming ways to diagnose and describe infections in pepper leaves. Advanced technology, like image processing, can help to solve this problem in bulk farming. Disease classification has been improved by using deep learning and image processing, and many diseases can now be detected automatically [1]. Agriculture uses image processing to identify diseased leaves, stems, or fruits; measure damaged areas; and determine color.

Deep learning, a branch of machine learning, analyzes each plant's distinct color, shape, size, and appearance to detect plant diseases or infections [1], [2]. Using hybrid deep learning approaches, the primary goal of this work is to create an automated system model for the precise and effective diagnosis of various pepper leaf diseases. This study suggested using a deep learning model—a hybrid of CNN-HOG and CNN-CAE techniques—to identify and classify pepper leaf diseases. This study uses a hybrid deep learning approach that combines CNN algorithms with convolutional neural network models for efficient image processing and data analysis [1]. The study recommends using a deep learning model that combines CNN-HOG and CNN-CAE approaches to detect and classify pepper sickness in pepper leaves. The model uses a multi-stage methodology that includes image processing, dataset collecting, noise reduction, segmentation, feature extraction, and classification. It focuses on six types of pepper leaf diseases: bacterial wilt, Cercospora leaf spot, leaf curl, healthy leaves, and bacterial spot.

This study focuses on six specific classes Pepper Leaf diseases: The diseases affecting the Pepper are Bacterial Wilt, Cercospora leaf spot, Cercospora leaf spot (frogeye), Leaf Curl, Healthy leaves, and bacterial spot. The illnesses were intentionally chosen based on the extent of damage they do to pepper production and the presence of imaging data for them. The main contributions of this work.

Lightweight Deep Learning Model: Proposes an enhanced lightweight CNN for accurate and efficient detection of multiple pepper leaf diseases, suitable for low-resource environments.

Comprehensive Dataset: Utilizes a large combined dataset (≈4,300+ images) from field data and public sources (PlantVillage, Kaggle), categorized into six disease classes with expert validation.

Advanced Preprocessing Pipeline: Implements a complete image processing workflow including noise removal, segmentation, augmentation, and feature extraction to improve model performance.

Hybrid Modeling Innovation: Introduces hybrid approaches (CNN–HOG and CNN–CAE), showing improved performance compared to existing methods.

Robust Validation: Evaluated using real-world images and multiple metrics (accuracy, precision, recall, F1-score, confusion matrix), ensuring reliability and effectiveness.

The upcoming sections are organized as follows: Section 2 interprets the related works related to pepper leaf disease classification using DL techniques. Section 3 describes the proposed methodology. Section 4 determines the results and discussion. Section 5 represents the Conclusion of the proposed study.

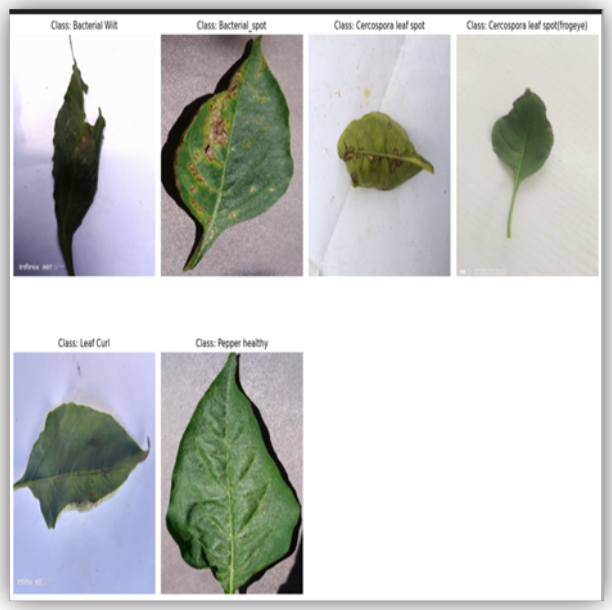

Figure 3. Sample image all diseases selected class by the study (own data, generated by system, 2024)

II. RELATED WORKS

Deep learning applications are widely used around the world in multiple domains such as visual data processing, health care, audio and speech processing and social network analysis [1] [2]. In the comparison of the previously employed machine learning algorithms that can improve performance with the several feature extraction techniques whereas the deep learning model can extract features automatically and perform training according to these features expressed in the hierarchical architectures The study [3] Proposes a concatenated CNN model using VGG16 and AlexNet features for pepper disease classification. The model, which includes dataset collection, image per-processing, noise removal, segmentation, and feature extraction, achieved 100% training, 97.29% validation, and 95.82% testing accuracy [9]. The study's findings suggest that the proposed concatenated model is effective in identifying pepper leaf and fruit diseases from digital images of pepper. The authors (Prakanshu, Kritika, Vibhav, 2021} presents convolutional neural networks with transfer learning to detect cassava illness. Although the system reached 93% accuracy, specialists retrieved photos manually to detect serious illnesses. A controlled experiment on a public datasets of 54,306 utilizing CNN and automated feature extraction yielded 99.35% accuracy.

Authors in [12] researched plant disease identification using deep learning classification model specifically using CNN model. They explored early disease detection by using Laplacian filter and Unsharp masking technique as image processing technique and Canny edge detection as image segmentation technique. For image processing we applied unsharp mask technique. For segmenting images further for extracting features, we used canny edge detection technique for segmentation and GLCM technique for extracting features. They are used 3000 images pepper leaf and implemented

convolution neural network with 97.82% accuracy and average 10% validation loss. Their limitations are limited to small data size, their works show plots training and validation accuracy high in, plots shows gaps between them that expressed as over fitting.

Research paper in [8]). This paper proposes the novel hybrid Coyote Grey Wolf optimization (CO-GWO) algorithm for the classification of mango leaves as normal or diseased. The classification process is done through the extraction of significant features from the segmented image. The presented model was trained and tested with the mango leaf images collected from the Mendeley database and the performances are evaluated. Proposed algorithm that plays a major role in the enhancement of the classification accuracy. The effectiveness of the proposed model is evaluated concerning the evaluation metrics, namely accuracy, precision, recall, and F1 measure, and is attained to be 96.7111%, 97.5712%, 97.1504%, and 96.4792%, respectively. They limit data size and datasets balance not known, old image prepossessing is used, they used binary class and parameters not known validation, testing accuracy unclear.

The [4] study suggests convolutional network (CNN) and histogram-oriented gradients (HOG)-based automatic detection and classification system for mango disease. The researcher collects in the proposed study; a dataset of 2500 data sets was collected from the Ethiopian agricultural research institute, the Assosa branch, and an online repository (Kaggle). The dataset has five (5) classes. Researchers conducted three experiments with CNN, HOG, and a combination of CNN and HOG (hybrid). The best results came from combining CNN & HOG (hybrid), which outperformed other techniques with accuracy rates of 98.60% during training and 99.5% during testing. The proposed hybrid (CNN and HOG) model proposes an extended model that has high validation accuracy and minimal validation loss, outperforming training accuracy. Their work gaps are data set insufficient for DL model,data size in-balanced, over-fitting issues.

The author [1] was proposed for identifying plant diseases using images of leaves. Compared to state-of-the-art methods like VGG19, ResNet50, and InceptionResNetV2, it achieved an average classification accuracy of 99.09%. This research introduces a unique hybrid model for autonomous plant disease diagnosis that combines the Convolutional Auto encoder (CAE) network and the Convolutional Neural Network (CNN). This system is unique in literature and has not been proposed in any cutting-edge systems. The model is designed to detect Bacterial Spot disease in peach trees using leaf photos, but it can also be used to detect other plant diseases. The trials employed the Plant Village datasets, which is publicly available, and obtained 99.35% training and 98.38% testing accuracy with only 9,914 training parameters.

The study [4] presents a CNN-YOLOv7 concatenated neural network for a linked-layer classification model for pepper leaf diseases. The model uses a multi-stage procedure, optimizing hyperparameters for accuracy and efficiency. The CNN-YOLOv7 hybrid model in agriculture offers a reliable method for real-time disease diagnosis, improving food security and promoting sustainable farming practices, with an average precision of 98.92% and 99.02% respectively. The paper [2] proposed model aims to reduce complexity in classifying apple leaf disease using deep learning, with a validation accuracy of 93.3% on the pepper leaf disease dataset. The study presents a computer vision system that involves picture capture, feature extraction, and classification. Deep Belief Network (DBN) is used to classify real-time pictures. The suggested approach for detecting illness on pepper plant leaves outperforms previous methods in terms of classification rate. The categorization results indicate whether the leaf is sick or not. The authors in [4] suggested using CNN for feature extraction and SVM for classification. The hybrid model was called CNN-SVM. The CNN model's final fully connected layer uses the SVM classifier rather than the activation function. The SVM model was combined with the "rbf" kernel. The CNN-SV M was tested on the MNIST datasets to recognize handwritten digits.

In study [5] has designed a convolution neural network model (GoogLeNet-EL) with high accuracy and easy transplantation for pepper leaf disease identification. The initial step involves compressing the Inception structure to reduce model parameters, leading to a remarkable enhancement in recognition speed. the proposed improved model has been trained on the real dataset of 9183 images, containing 6 types of pepper diseases. The cross-validation results show that the model accuracy is 97.87%, which is 6% higher than that of Google-net based on Inception-V1 and Inception-V3. The memory requirement of the model is only 10.3 MB, which is reduced by 52.31%-86.69%, compared to GoogLeNet. We have also compared the model with the existing CNN-based models including Alex-net, ResNet-50 and MobileNet-V2. The result shows that the average inference time of the proposed model decreases by 61.49%, 41.78% and 23.81%, respectively. The results show that the proposed enhanced model can significantly improve performance in terms of accuracy and computing efficiency.

In [2] research conducted A hybrid deep learning model approach for automated detection and classification of cassava leaf diseases. This paper explores the application of advanced deep learning techniques to detect as well as classify cassava leaf diseases which include Efficient-net models, DenseNet169, Xception, MobileNetV2, Res-net models, Vgg19, InceptionV3, and InceptionResNetV2. A dataset consisting of around 36,000 labelled images of cassava leaves, afflicted by diseases such as Cassava Brown Streak Disease, Cassava Mosaic Disease, Cassava Green Mottle, Cassava Bacterial Blight, and healthy leaves, was used to train these models. The hybrid model (DenseNet169 + EfficientNetB0) had superior performance with classification accuracy of 89.94% while as EfficientNetB0 had the highest values of precision, recall, and F1score with 0.78 each. The novelty of the hybrid model lies in its ability to combine DenseNet169's feature reuse capability with EfficientNetB0's computational efficiency, resulting in improved accuracy and scalability.

Th authors in [6] explore a model designed to classify various plant leaf diseases effectively. To propose a model that can classify different (38 in this work) types of plant leaf images and finds the classification accuracy. Model utilized by their works various convolutional neural network (CNN) architectures, including InceptionV3, ResNetV2, MobileNetV2, and EfficientNetB0, have been developed. PlantVillage dataset is used for training and testing which contains 38 different classes and 14 different plant species in total, among 38 classes 12 of which are healthy, 26 of which are diseased. Model architectures, including AlexNet, ResNet50, VGG16, InceptionV3, MobileNetV2, and EfficientNet, were carefully selected to represent diverse

deep-learning frameworks. Each model was finely tuned for the plant leaf disease classification task, employing optimization techniques such as SGD and Adam. Transfer learning methodologies, combined with pre-trained CNN models, facilitated intricate feature extraction from plant leaf images, resulting in robust disease classification outcomes. In Efficient-net architecture, even though the input photograph size is re-sized to 132x132, it yields greater success attaining an accuracy of 97.5% than other models.

In [1] study presents a five-layered CNN model for auto matic detection of plant disease utilizing leaf images. Dataset consists of 2475 images related to pepper bell leaves that are divided into two different types that are used as true labels such as bacteria and healthy leaves. The Optimized-CNN is used with the image augmentation technique. To better train a CNN model, 20,000 augmented images are generated. Experimental results demonstrate that proposed optimized-CNN model can predict pepper bell plant leaf as healthy or bacterial with 99.99% accuracy. An analysis of the performance of the proposed approach is carried out with different numbers of epochs such as 5, 10, 15, 20, and 25 epochs. The highest accuracy, 99.99% has been achieved with 25 epochs. While the precision, recall, and F1 score are 99.99%, 99.99%, and 99.99%. The Optimized-CNN approach outperformed in the classification of healthy and bacterial pepper bell leaf images.

The study in [8] proposes a novel optimized DL model for classifying the presence and absence of pepper leaf disease using an effective feature learning process. The proposed study is implemented via the Python platform, and a publicly available Plant-Village dateset is utilized for the simulation process. The proposed study undergoes four major stages namely pre-processing, segmentation, feature extraction, and classification. Kernelized Gravity-based Density Clustering (KGDC) technique is conquered to segment the diseased portions from the leaf images. Finally, the Gated Self-Attentive Convoluted MobileNetV3 (GSAtt-CMNetV3) technique is proposed to extract the features and classify the pepper leaf disease accurately. A novel osprey optimization algorithm (Os-OA) is introduced to tune the parameters of the proposed DL model for enhancing the classification performance. Accuracy, precision and recall values achieved by the proposed pepper leaf disease classification for training percent 80 is 97.87%, 96.87% and 97.08% respectively. but their works are limited to binary class pepper leaf and only benchmark dateset of plant village, testing and validation their model performance does not mention unknown how well their model accurate.

A. RESEARCH GAP AND SOLUTION

Early plant disease detection is a critical task with far-reaching implications for agricultural sustainability and food security. Leveraging plant disease datasets for this purpose has shown promise, but several limitations must be addressed for such models to be more effective and useful. Traditional detection methods relying on expert observation and laboratory examination of leaves are empirical, particularly in most remote areas of Ethiopia. The study introduces a deep learning-based solution for automating pepper leaf diseases detection and classification, addressing gaps in existing systems such as limited datasets and narrow disease detection scopes, used only publically dataset. Although most of existing studies have focused on traditional methods, used single well known model like CNN, SVM, and RNN for pepper leaf disease detection and classification. There is a need to explore the potential of deep learning techniques for the detection and classification. ML-based image processing algorithm face limitations in automatically extracting these features while DL-overcomes this limitation by automatically extracting low-level, intermediate level and high-level semantic feature from a high dimensional features space to perform detection and classification tasks. Another crucial factor to evaluate is how well the model performs under different image situations. Environmental factors such as illumination, camera angles, and image quality can all substantially impact on disease detection model performance.

Furthermore, combining multimodal data sources may improve the accuracy and reliability of disease detection models. Combining image data with other sensor data, such as spectrum or hyper spectral imaging, meteorological data, or soil information, may provide additional information that increases disease detection accuracy and allows for more thorough monitoring of plant health. Furthermore, plant disease detection models require consistent evaluation methodologies and standards [15]. Establishing uniform datasets, evaluation measures, and benchmarking techniques would allow for fair comparisons of diverse models, accelerating advancement in the field. Future research directions for early plant disease detection using plant disease datasets should concentrate on improving model generalizability to previously unknown plant species or diseases, improving performance under changing image conditions, integrating multimodal data sources, and developing standardized evaluation protocols. Addressing these problems will help to produce more accurate, reliable and adaptable plant disease detection systems, which will have far-reaching agricultural effects.

B. HYBRID TECHNIQUES DEEP LEARNING MODEL IN IMAGE PROCCESSING

Artificial intelligence has transformed the imaging discipline, notably picture classification. Machine learning (ML) has made the process easier, while deep learning (DL) has improved accuracy and automation. Solo deep learning (SDL) is a single architecture, whereas hybrid deep learning (HDL) models incorporate many DL architectures [17] [8]. This is the first narrative HDL review, examining all areas of image classification with AI, DL methods, and ML models. HDLs are divided into spatial, temporal, and spatial-temporal categories and analyzed using attributes such as continent, publisher, DL or ML hybridization, architecture layout, application type, data set type, dataset's size, feature extraction methodology, connecting classifier, performance evaluation metrics, and risk-of-bias. The review finds that HDL models exhibit consistent and higher performance by integrating the best aspects of two or more solo DL or fusion of DL with ML models. HDL is being aggressively applied to various medical and agricultural applications. Overall as we reviewed the different papers in different approaches and different journals papers there were gaps at all [18].

C. EXPLAIN HOW TO APPLY CNN+CAE AND CNN+HOG

The effectiveness of different hybrid machine learning techniques, specifically CNN+CAE and CNN+HOG, when combined with ensemble methods like bagging and stacking. Bagging (Bootstrap Aggregating) is an ensemble method that

aims to improve the stability and accuracy of DL algorithms. It works by training multiple models on different subsets of the training data and then aggregating their predictions. Here's how we were applied bagging to CNN+CAE and CNN+HOG architectures: Feed the image (and its HOG features) through the first CNN+HOG model to get its prediction. Repeat for all N CNN+HOG models. A CNN+CAE (Convolutional Neural Network + Convolutional Auto encoder) hybrid model typically uses the CAE for feature extraction (unsupervised pre-training) and the CNN for classification or regression based on these features. for both CNN+CAE and CNN+HOG, the core idea of bagging remains the same: create multiple diverse training sets via bootstrap sampling and train a version of your hybrid model on each, then combine their outputs

III. MATERIALS AND METHODOLOGY

A. DATA COLLECTION AND DATASET PREPARATION

To train a model on pepper leaf diseases, researchers used original data from the Illu Aba Bor Zone and secondary data from the Plant Village Dataset. Over 2552 leaf photos were captured, including bacterial wilt, Cercospora leaf spot, frogeye, leaf curl, pepper bacterial spot, and healthy pepper leaf. The training dataset was increased using image augmentation. The datasets were divided into three categories: training, validating, and testing, with 80% for training and 20% for validation, validating, and testing. A total of 2552 pepper leaf photos were used to train and test the model. From early November to December 2023, the images were taken with an Infinix HOT 12play. RGB and other graphics are used to describe pepper leaf diseases. JPEG images with resolutions of 5312 x 2988 were stored. Mr. Tilahun Silash, a PHD student at Jima University, and Mr. Motuma Gemechu, a plant pathology lecturer at Mattu University's Bedelle Agriculture campus, verify the pictures. Table 3: shows dataset collected before Augmentation and Table 4 shows the pepper illness category-specific training, validation, and testing datasets.

Table 1. Dataset collected before Augmentation

S/no	Classes	No of images per classes
1	Bacterial wilt	173
2	Cercospora leaf spot,	114
3	Cercospora leaf spot (frogeye),	168
4	Leaf Curl	235
5	Pepper Bacterial spot	997
6	Healthy pepper leaf	865
Total		2552

Table 2. Numbers of datasets after Augmentation

S.No	Classes	No images per classes
1	Bacterial wilt	619
2	Cercospora leaf spot,	656
3	Cercospora leaf spot (frogeye),	604
4	Leaf Curl	700
5	Pepper Bacterial spot	897
6	Healthy pepper leaf	865
TOTAL		4341

B. RESEARCH DESIGN AND MATERIALS USED

Design science is the study of inventive problem solutions and the creation of new worlds. This study employs experimental research techniques, which entails data preparation, model construction, and prototype creation. Implementation tools (CPU) included as a software Anaconda, Python, Keras, Tensor Flow, and open-source libraries. An experimental prototype for pepper leaf disease classification models was created. The researcher used as a hardware smart mobile phone camera, the Infinix HOT 12play, to capture field photographs and a sluggish machine with an Intel Core i5-8400 processor, 8GB memory, and 1TB hard disk space for deep learning

C. ARCHITECTURAL OF THE PROPOSED MODEL

In recent years, pepper plant diseases' automatic detection has grown in popularity by using deep neural networks. We present a short description of the proposed framework that will be used for detection and classification of diseases in pepper leaves plants using deep learning and image processing. The general architecture of our suggested deep learning models is depicted in Figure 4; This study proposed using a combination of deep learning models to detect different pepper crop diseases. The suggested architecture for this study consists of three stages: testing, validation, and training. During the training phase, image data is collected, prepossessed, and separated into training, validating, and testing sets. The Image-Data Generator instance generates images through random transformations, which are then used in training and testing. The model is trained using prepossessed data and evaluated for performance. During testing, unseen data is prepossessed, and the model classifies images based on bacterial wilt, leaf curl, pepper bacterial spot, Cercospora leaf spot, Cercospora leaf spot (frogeye), or healthy leaf.

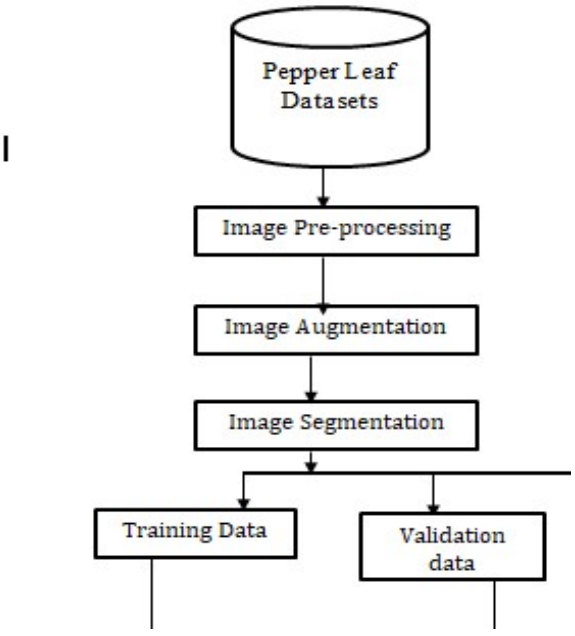

Figure 4. Proposed Model Architecture (own Design)

The system receives training and validation data and performs prepossessing operations such as filtering, noise reduction, and data re-scaling. Adam's optimization approach is used for segmentation and augmentation. Following training, the model is saved and prepossessed on test data to

ensure that the parameters match. The saved model is then used to predict the proper image class from unlabeled data.

i. **Feature Extraction:** Extraction of features is responsible for extracting the descriptive features of pepper leaf images through the extraction of the most important details extracted from the picture. It is the process of representing the image feature into distinguishable and identifiable attributes in numerical value [17]. To extract the feature of the pepper leaf image, we have used a deep extraction algorithm, CNN, and a handcrafted feature extraction algorithm, GLCM. We have used CNN to extract the deep features of pepper leaf and used GLCM to extract the texture and intensity [15].

ii. **Convolutional Auto Encoders:** Convolution Auto Encoders are feed-forward neural networks designed for unsupervised learning or data unlabeled situations. They compress input into a lower-dimensional code and reconstruct output using the compressed code. The encoder, code, and decoder are three parts of an auto encoder, with the decoder producing output and the encoder accepting input. Applications include image processing, popularity forecasting, and drug development [1].

iii. **Histogram of Oriented Gradients (HOG):** The Histogram of Oriented Gradients (HOG) is a traditional feature descriptor technique used in computer vision and image handling for object detection and recognition. It focuses on the shape and structure of objects within photographs, counting the occurrences of gradient orientation in each local region. HOG is essential for improving computer vision framework accuracy and reliability. It captures important texture and edge information by computing histograms of gradient orientations within discrete sections of a picture. HOG is employed in image processing and computer vision to detect objects by counting the instances of gradient orientation in the localized area of a picture. It is superior to edge descriptors as it uses magnitude and angle of the gradient to compute features [19]

V. RESULTS AND DISCUSSIONS

The comparative evaluation of the Hybrid CNN–HOG and Hybrid CNN–CAE models reveals that both approaches achieve high classification performance across all disease classes; however, the Hybrid CNN–CAE model demonstrates superior overall effectiveness. While the CNN–HOG model attains strong precision, recall, and F1-scores—often ranging between 96% and 100%—it exhibits minor inconsistencies in certain classes, such as Cercospora Leaf Spot and Frogeye. In contrast, the CNN–CAE model delivers more balanced and consistent results, achieving perfect or near-perfect scores in multiple classes, including Cercospora Leaf Spot and Leaf Curl. Furthermore, the CNN–CAE model outperforms CNN–HOG in aggregate metrics, with higher macro-average (82% vs. 80%) and weighted-average scores (84% vs. 82%), indicating better generalization across class distributions. This improved performance can be attributed to the Convolutional Autoencoder's ability to learn rich, hierarchical feature representations, enabling more effective capture of complex disease patterns compared to the hand-crafted feature extraction approach of HOG. Therefore, the Hybrid CNN–CAE model is identified as the more robust and reliable approach for plant disease classification in this study.

PERFORMANCE COMPARISON OF EACH MODEL

Table 3. Comparison of the experimental results of each model

Model	Evaluation Matrix				
	Total Parameter	Batch size	No Epochs	Accuracy	Test Accuracy
CNN	3455686 (13.18 MB)	32	10	96.8%	96.81%
CAE	1889 (7.38 KB)	32	10	57.84%	57.83%
HOG	169734 (35.61 MB)		10	97.68%	95.58%

The comparison aims to evaluate the performance of each model in detecting pepper sickness, using chosen metrics and class metrics. The table provides a summary and conclusions about the effectiveness of each model in detecting pepper leaves diseases [15].

A. PERFORMANCE COMPARISON OF THE EACH OF OUR HYBRID MODEL

As shown in Figure 5, the Hybrid CNN–CAE model demonstrates superior performance compared to the Hybrid CNN–HOG model. Although both models achieved high classification scores across all classes, CNN–CAE produced more consistent and balanced precision, recall, and F1-scores. Additionally, it achieved higher macro and weighted averages, indicating better generalization and robustness. This improvement can be attributed to the CAE's ability to learn deep feature representations, making it more effective in capturing complex disease patterns than the hand-crafted HOG features.

Table 4. Comparison of each hybrid model

Class	Hybrid of CNN-HOG model			Hybrid of CNN-CAE model		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Bacterial Wilt	97	97	97	96%	98	98
Bacterial Leaf Spot	98	99	98	100%	97	97
Cercospora Leaf Spot	99	98	97	100%	100	100
Cercospora Leaf Spot(Frogeye)	97	100	96	98%	98	98
Leaf Curl	99	100	100	100%	98	99
Health Leaf	100	100	100	97%	100	99
Accuracy						87%
Macro avg				80%		82
Weighted avg				82%		84

Figure 5. Graph shows results of hybrid model CNN+CAE Accuracy in training and validation

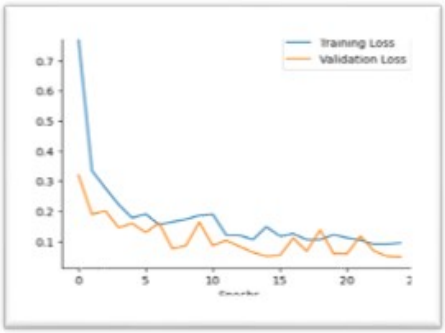

Figure 6. Graph shows results of hybrid model CNN+CAE Accuracy of training and validation loss

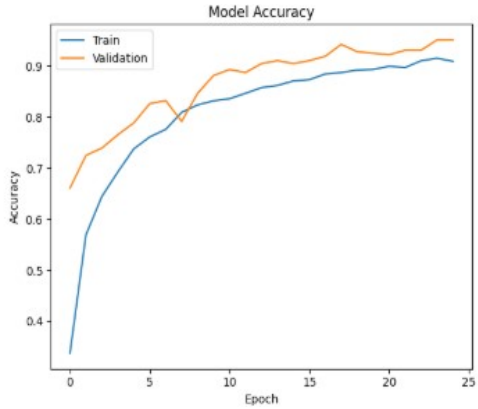

Figure 7. Graph shows result of hybrid model CNN+HOG Accuracy in training and validation

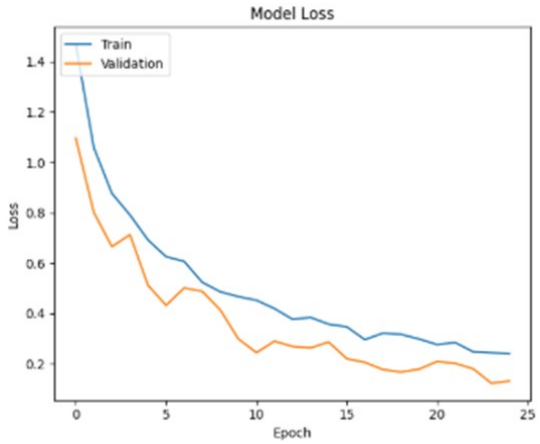

Figure 8. Graph shows results of hybrid model CNN+CAE Accuracy of training and validation loss

VI. CONCLUSIONS

Many studies have applied machine learning and deep learning techniques to diagnose pepper diseases. In this work, a deep learning system was developed for early detection using image processing steps such as preprocessing, noise removal, resizing, feature extraction, and classification. Different models, including hybrid approaches like CNN–CAE and CNN–HOG, were evaluated to accurately identify diseases on pepper leaves. In conclusion, although both Hybrid CNN–HOG and Hybrid CNN–CAE models achieve high classification performance, the Hybrid CNN–CAE model proves to be more effective and reliable. It delivers more consistent and balanced results across all disease classes, along with higher macro and weighted average scores, indicating better generalization. This superior performance is largely due to its ability to learn deep, hierarchical features, allowing it to capture complex disease patterns more effectively than the HOG-based approach. Therefore, the Hybrid CNN–CAE model is the preferred choice for plant disease classification in this study. This innovative method could greatly enhance the ability to identify diseases on pepper leaves and promote the general health of pepper plants.

References

[1] H. Mustafa, M. Umer, U. Hafeez, A. Hameed, A. Sohaib, S. Ullah, H. A. Madni, "Pepperbell leaf disease detection and classification using optimized convolutional neural network," *Multimedia Tools and Applications*, pp. 12065–12080, 2022. https://doi.org/10.1007/s11042-022-13737-8.

[2] G. Sambasivam, G. Prabu Kanna, M. Singh Chauhan, P. Raja, Y. Kumar, "A hybrid deep learning model approach for automated detection and classification of cassava leaf diseases," *Scientific Reports*, vol. 15, 7009, 2025. https://doi.org/10.1038/s41598-025-90646-4.

[3] Y. A. Bezabh, A. O. Salau, B. M. Abuhayi, et al., "CPD-CCNN: classification of pepper disease using a concatenation of convolutional neural network models. *Scientific Reports*, vol. 13, 15581, 2023. https://doi.org/10.1038/s41598-023-42843-2.

[4] W. Sema, Y. Yayeh, G. Andualem, "Automatic detection and classication of mango disease using convolutional neural network and histogram oriented gradients," *Research Square*, pp. 1-15, 2023. https://doi.org/10.21203/rs.3.rs-3073109/v1.

[5] M. Dai, W. Sun, L. Wang, Md. M. H. Dorjoy, S. Zhang, H. Miao, L. Han, X. Zhang, M. Wang, "Pepper leaf disease recognition based on enhanced lightweight convolutional neural networks," *Frontiers in Plant Science*, pp. 1-18, 2023. https://doi.org/10.3389/fpls.2023.1230886.

[6] G. S. Hukkeri, B. C. Soundarya, H. L. Gururaj, V. Ravi, "Classification of various plant leaf disease using pretrained convolutional neural network on Imagenet," *The Open Agriculture Journal*, vol. 18, pp. 1-15, 2024. https://doi.org/10.2174/0118743315305194240408034912.

[7] A. Sheneamer, "Early detection of plant leaf diseases using stacking hybrid learning," PLOS ONE, pp. 1-23, 2022. https://doi.org/10.1371/journal.pone.0313607.

[8] S. S. A. Begum, H. Syed, "GSAtt-CMNetV3: Pepper leaf disease classification using osprey optimization," *IEEE Access*, vol. 12, pp. 32493-32506, 2024, https://doi.org/10.1109/ACCESS.2024.3358833.

[9] Y. K. Afework, *Developing Bacterial Wilt Detection Model on Enset Crop Using a Deep Learning Approach*, Master Thesis, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia, 2019.

[10] S. Goumiri, D. Benboudjema, W. Pieczynski, "A new hybrid model of convolutional neural networks and hidden Markov chains for image classification," *Neural Comput. Appl.*, vol. 35, no. 24, pp. 17987-18002, 2023. https://doi.org/10.1007/s00521-023-08644-4.

[11] T. Tao, X. Wei, "A hybrid CNN–SVM classifier for weed recognition in winter rape field," *Plant Methods*, vol. 18, issue 29, pp. 1-12, 2022. https://doi.org/10.1186/s13007-022-00869-z.

[12] K. Omar, Z. A. Ebrahim, "Hybrid system for image classification using CNN and low image processing," *International Journal of Engineering Research & Technology* (IJERT), vol. 11, issue 11, pp. 333-336, 2022, DOI : 10.5281/zenodo.18397495.

[13] P. Bedi, P. Gole, "Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network," *Artificial Intelligence in Agriculture*, vol. 5, pp. 90-101, 2021, https://doi.org/10.1016/j.aiia.2021.05.002.

[14] M. P. Mathew, T. Y. Mahesh, "Leaf-based disease detection in bell pepper plant using YOLOv5," *Signal, Image and Video Processing*, vol. 16, issue 3, pp. 841–847, 2022. https://doi.org/10.1007/s11760-021-02024-y.

[15] A. Özcan and E. Dönmez, "Bacterial disease detection for pepper plant by utilizing deep features acquired from DarkNet-19 CNN model," *DUJE*, vol. 12, no. 4, pp. 573–579, 2021. https://doi.org/10.24012/dumf.1001901.

[16] P. Srivastava, K. Mishra, V. Awasthi, V. K. Sahu, P. K. Pal, "Plant disease detection using convolutional neural network," *International Journal Advanced Research (IJAR)*, vol. 9, issue 1, pp. 691-698, 2021. https://doi.org/10.21474/IJAR01/12346.

[17] S. Berta, "Isolation and characterization of wilt-causing pathogens of local growing pepper (capsicum annuum L.) in gurage zone, Ethiopia," *International Journal of Agronomy*, pp. 1-8, 2020. https://doi.org/10.1155/2020/6638683.

[18] A. K. Yimeru, *Automatic Coffee Disease and Pest Damage Identification*, Thesis, Addis Ababa University, Addis Ababa, Ethiopia, 2020. http://etd.aau.edu.et/handle/123456789/22352.

[19] A. A. Agranovsky, "Virus diseases of pepper (Capsicum annuum L.) in Ethiopia, *Journal of Phytopathology*, vol. 138, pp. 89-97, 2008. https://doi.org/10.1111/j.1439-0434.1993.tb01364.x.

[20] V. Sharma, A. Verma, N. Goel, "Classification techniques for plant disease detection," *International Journal of Recent Technology and Engineering (IJRTE)*, pp. 5423-5430, 2020. https://doi.org/10.35940/ijrte.F9902.038620.

[21] M. Cethana, K. Raksha, V. Thanuja, R. Sridhar, "Classification and Identification of Leaf Diseases using Deep Learning," International Research Journal of Engineering and Technology (IRJET) , pp. 4422-4427, June 2020.

[22] S. Jana, A. Rijuvana Begum, S. Selvaganesan, "Design and analysis of pepper leaf disease detection using deep belief network," *European Journal of Molecular & Clinical Medicine*, pp. 1724-1731, 2020.

[23] A. Thamizharasi, C. Mary, N. V. Radhakrishnan, "Detection and assessment of leaf spot disease in black pepper leaf," *Proceedings of the 3rd National Conference on Recent Trends in Computer Science and Engineering and Sustainability in Civil Engineering (TECHSYNOD'16)*, Thiruvananthapuram, Kerala, India, 19-21 December 2016.

[24] G. M. Kaba, *Characterization of Fusarium species from hot pepper (Capsicum anuum L.) and In vitro antagonistic effect of Trichoderma species*, Thesis Jimma University, Jimma, Ethiopia, 2019.

[25] N. Fufa, G. Abera, E. Etissa, "Agronomic and economic performance of hot pepper (Capsicum annuum L.) in response to blended fertilizer supply at Asossa, Western Ethiopia," *International Journal of Plant & Soil Science*, vol. 26, issue 2, pp. 2018. https://doi.org/10.9734/IJPSS/2018/46132.

[26] Y. Kahsay, "Evaluation of hot pepper varieties (capsicum species) for growth, dry pod yield and quality at M/Lehke District, Tigray, Ethiopia," *International Journal of Engineering Development and Research*, pp. 144-152, 2017.

[27] A. D. A. (. R. A. R. I. (. I. o. A. R. (. M. o. A. a. R. D. International Center for Agricultural Research in the Dry Areas (ICARDA), "Participatory adaptation of hot pepper (Capsicum species) varieties for green pod production under irrigation condition," Addis Ababa, Ethiopia , JUNE 2016.

[28] M. B. Devi, Amarendra K, "Machine learning based application to detect pepper leaf diseases using HistGradientBoosting Classifier with fused HOG and LBP Features," *International Journal of Emerging Trends in Engineering Research*, pp. 7371-7376, 2020. https://doi.org/10.30534/ijeter/2020/1148102020.

[29] T. Tsehaynesh, M. Abdi, S. Hassen, W. Taye, "Aspergillus species and aflatoxin contamination in pepper (Capsicum annuum l.) in West Gojjam, Ethiopia," African Journal of Food, Agriculture, Nutrition and Development, vol. 21, issue 1, pp. 17178-17194, 2021. https://doi.org/10.18697/ajfand.96.18815.

[30] B. Habtie, T. Dejen, Y. Dessalegn, "Morphological characterization of hot pepper (Capsicum annuum.L) land races of ethiopia for qualitative characters," *International Journal of Research Studies in Science, Engineering and Technology*, pp. 4-9, 2017.

[31] S. Mekonen, Alemayehu Chala, "Assessment of hot pepper (capsicumspecies) diseases in Southern Ethiopia," *International Journal of Science and Research (IJSR)*, pp. 91-95, 2014.

[32] A. Abebe, M. Abera, "Pepper disease assessment and identification in major growing districts of West Gojam zone in Northwestern Ethiopia," *International Journal of Sustainable Agricultural Research*, pp. 8-20, 2019. https://doi.org/10.18488/journal.70.2019.61.8.20.

[33] B. Biri, P. Gomathinayagam, "Response of hot pepper (Capsicum annuum L.) to major fungal diseases under field and greenhouse conditions in Horo Guduru Wollega, Oromia, Ethiopia," *African Journal of Agricultural Research*, pp. 924-932, 2021.

***Desalegn Tasew Degaga** received his Msc and Bsc degrees in Information Technology from Mattu University 2024 and 2018 respectively. He is currently serving as Lecturer in Computer Science at Jimma University, Agaro, Ethiopia. His research interests include NLP, Machine Learning, Image processing, IoT, Network Security. He has published 3 papers in Scopus Indexing Journal and Conducting some research, preparing Lab Manuals.*

***ARULMURUGAN RAMU** is an accomplished academic and researcher specializing in computer vision, artificial intelligence, machine learning, image processing, and natural language processing. He holds a Ph.D. in Computer Science and Engineering from Anna University, India, where he focused on object recognition techniques using wavelet neural networks. With over 15 years of teaching experience, Dr. Ramu has held positions at Anna University, Bannari Amman Institute of Technology, Presidency University, CMR University, and is currently an Associate Professor in the Department of Computer Science at Heriot-Watt University's Kazakhstan campus. He has published more than 100 research papers in leading international outlets such as IEEE, Springer, Elsevier, and ACM, with significant contributions in image analysis, sustainable computing, big data analytics, and intelligent systems.*

***Dr. K. MARIYAPPAN,** with qualifications in BE, MTech, and PhD, boasts an impressive teaching career spanning over 23 years. He currently holds the position of Professor in the computer science and engineering Department at Chennai institute of Technology. His extensive experience includes teaching at both national and international universities, where his expertise and dedication have significantly contributed to the academic growth of these institutions. Dr. Mariyappan's research interests lie in the areas of Information Security, IoT (Internet of Things), and Wireless Sensor Networks. With a robust research background, Dr. Mariyappan has published numerous papers in reputable Scopus and SCI journals.*

***DIRIBA GICHILE RUNDASA** received his Msc and Bsc degree in Computer Science and Engineering from Adama Science and Technology University 2023 and BSc degree in Computer Science from Mattu University 2018 respectively. He was currently serving as Lecturer in Computer Science at Mattu University, Mattu, Ethiopia. Besides this he has been working at different community service and research centers at Mattu University. His research interests include NLP; Machine Learning and He has published papers in Scopus Indexing Journal and Conducting some research and Projects.*