ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2026.0541

ONION IMAGE CLASSIFICATION USING CNN MODELS

J. Anusha Jajur and U. Kumar Siddamallappa
Department of Computer Science, Davangere University, India

Abstract
The quality of onions (Allium cepa), a vegetable that is consumed worldwide, is essential for food safety and agricultural economics. This study suggests a deep learning-based approach that uses convolutional neural networks (CNNs) to categorize images of onions as either healthy or unhealthy. The technique uses a proposed architecture and a Customized Dataset (CDS) and publically accessible sources like Fruit360, Onion-det, and Vegetable360. The model, which outperformed the others, reached a peak accuracy of 98.33% on the CDS dataset. The study also highlights the importance of color information in onion disease categorization, with models trained on RGB images performing better than monochrome counterparts. The model's classification skills are further confirmed using confusion matrices. The CNN architecture has a lot of potential for automated onion quality evaluation, outperforming conventional pre-trained models regarding resilience and dependability, delivering high classification accuracy and great generalization across various datasets and images circumstances.

Keywords:
CNN, MobileNetV2, VGG16, ResNet50

1. INTRODUCTION

Among the most widely consumed veggies worldwide, onions (Allium cepa) are essential to culinary traditions and make a substantial contribution to the global agricultural economy. Onions are prized for their nutritional value, which includes important vitamins, minerals, and antioxidants, in addition to their capacity to improve flavor. Onions are a notable source of sulfur-containing phytochemicals, including thiosulfinates, thiosulfonates, allicin, alliin, and ajoene, in addition to their rich historical significance. These ingredients add to its therapeutic qualities, which, when taken consistently, may help prevent and treat heart disease, rheumatism, some types of cancer, digestive issues, hyperglycemia, and frequent coughing [3]. Available in several varieties—such as yellow, red, white, and green onions differ in flavour profiles, culinary uses, and growth conditions. Onion production is still susceptible to a number of issues, especially pests and diseases, which can drastically lower crop quality and yield despite their widespread demand.

Fungal infections like Botrytis allii (causing neck rot) [1] [2], Peronospora destructor (causing downy mildew), and Sclerotium cepivorum (white rot), along with bacterial infections like Pseudomonas syringae (soft rot), are among the most common illnesses that impact. As well as to pathogens, nematodes and insect infestations can further threaten crop health. These problems don't just impact growers by reducing harvests and market value but also consumers, by lowering the availability of high-quality produce. Traditionally, visual inspection and manual evaluation have been employed to identify diseases, but these methods are labour-intensive, time-consuming and susceptible to human error. Early disease symptoms are often subtle and may go unnoticed, leading to unpredictable grading outcomes, increased waste, and harm of marketable yield.

According to the FAO's Good Agricultural Practices for Onion and Garlic (2022), poor disease management and post-harvest handling are leading contributors to onion yield losses. Manual grading processes fail to meet the rising demand worldwide for high-quality agricultural produce and export-grade onions. There's a growing need for innovative and automated solutions to improve quality evaluation, ensure timely intervention, and support sustainable farming practices.

Agricultural diagnostics could undergo a revolution because to recent developments in computer vision and artificial intelligence (AI). These technologies, particularly and deep learning, offer scalable and cost-effective tools for monitoring crop health. The Convolutional Neural Network is one of the deep learning models have proven exceptionally capable in image classification tasks, making them suitable for identifying and distinguishing healthy and unhealthy onions based on pictorial symptoms [4] [5]. Image classification challenges frequently employ Convolutional Neural Network (CNN) designs like MobileNet, VGG16, and ResNet50. Among these, many researchers prefer MobileNet due to its efficiency and effectiveness in producing accurate results [11] [25].

Fig.1. Healthy and Unhealthy Onions

This study customs a similar strategy to growth the robustness and accurateness of onion disease detection under various image conditions, taking inspiration from latest developments in tomato disease [6], citrus fruit disease classification [7], apple leaf disease[8][9] and Avocado fruit disease [10] that use diverse ensemble models combining multiple pre-processing techniques, handcrafted and deep features, and hybrid classifiers. Beyond achieving high precision in disease detection, this research also aims to advance the use of smart technologies in agriculture**,** improve early-stage disease intervention, and reduce post-harvest losses. By empowering farmers with reliable, automated tools for real-time monitoring, this strategy supports more general objectives to increase agricultural marketability, food security and promote sustainability in onion production.

2. LITERATURE REVIEW

Current developments in artificial intelligence and computer vision have made it possible to accurately classify related veggies, such as shallots and garlic. Lestari et al. (2024) developed a CNN-based sorting four convolutional layer model, achieving 98.33% training and 100% testing accuracy on a garlic-onion image dataset. This study demonstrates the effectiveness of CNNs in agricultural image classification and highlights their potential to improve sorting efficiency and reduce manual errors in post-harvest processes [18].

Techniques in computer vision can be useful, but two key issues need to be addressed which classifier provides the best classification and which image attributes are associated with product quality. Using x-ray imaging, sweet onions were line-scanned for interior flaws, and significant characteristics were chosen using the Bayesian technique. Spatial edge features and the coefficients of the discrete cosine transform were found to be reliable markers of internal flaws. Overall accuracy for a neural classifier has 90%, with limited losses and false positives [19].

Purple blotch is one of the illnesses that affect onions, the second-largest vegetable crop in the world. A pre-trained improved InceptionV3 model is suggested as a solution to identify onion illness in photos. With an 85.47% classification accuracy, the model shows the promise of deep learning for agricultural data. This study intends to establish the theoretic foundations of deep learning in agriculture by creating a unique method for the quick and precise identification of plant/crop diseases [20].

The report explores the utilizes of SWIR hyperspectral imaging to identify onions with sour peel, focusing on two key wavelengths, 1070 nm and 1400 nm, which are sensitive to water content and structural changes. The study found that the SVM (Support Vector Machine) classifier reached an accuracy of 87.14%, demonstrating its potential for rapid, non-destructive inspection. [21].

In order to assess onion quality factors holistically and non-destructively, a multimodal machine vision system was created. The system combined a LabVIEW program with X-ray, 3D, and hyperspectral imaging sensors. By precisely measuring the weight, volume, diameter, and density of onions, the technology could tell the difference between onions that were in good health and those that weren't. Onion packing houses and other agricultural products can adopt this promising strategy. An overall accuracy of 88.9% was achieved by the collaborative SVM with cascaded SVMspectral, SVMX-ray, and SVMphysical. [22].

The research introduces an automated onion grading scheme that makes use of image processing algorithms to classify onions based on size, color, and texture, ensuring real-time sorting. The system uses Gaussian blur, edge detection, dilation, and Blob detection in RGB space. Validation tests show the system reduces manual errors and increases efficiency in onion sorting processes. [23].

The review of recent studies highlights the widespread application of DL (Deep Learning) models for plant infection detection, particularly using the Plant Village dataset [24]. Across various crops such as beans, tomatoes, grapes, and general leaf diseases, used various architectures including CNN, MobileNetV2, Generative Adversarial Networks (GANs), and hybrid models like CNNC and IKNN. Among them, MobileNetV2 [14] achieved the accuracy of 98% using 8000 tomato leaf images, demonstrating the effectiveness of lightweight yet powerful architectures. Hybrid models like IKNN also showed notable performance with 98.07% accuracy on grape leaves [15]. Most studies utilized moderate to large datasets, and consistently reported accuracies above 90%, indicating that deep learning approaches, when combined with well-curated image datasets, are highly effective in automated plant disease classification. This provides a solid basis for additional research into similar architectures for crops like onions, especially when extended to include real-world field conditions beyond the Plant Village dataset.

Table.1. Comparison of different diseases detected by various models

Ref.	Disease on	Models	Data sets	Datasets Size	Accuracy
[12]	Beans leaf	CNN with different optimizer	Plant Village	1295 images	91.74%
[13]	Leaf Disease	CNN	Plant Village	38 different classes of leaves	96%
[14]	Tomato plants	MobileNetV2	Plant Village	8000 leaf images	98%
[15]	Grape leaf	CNNC and IKKN	Plant-Village	4062	CNNC: 96.60 IKNN : 98.07
[16]	Tomato Leaf	Deep Convolutional GAN	Plant Village	1500	94.33%
[17]	Leaf Image	Multi-layer perceptron & SVM	Plant Village	16012	94.35%

3. PROPOSED METHOD

In this study, the proposed method aims to develop an effective and automated onion a classification scheme that can precisely distinguish between healthy and unhealthy onions using deep learning techniques. These images will be used to train CNNs to detect important visual indicators such as surface texture irregularities, discoloration, and structural deformation. Considering the complexity of visual features in agricultural produce, the technique uses cutting-edge Convolutional Neural Network (CNN) models with transfer learning, specifically proposed method, MobileNetV2, ResNet50, and VGG16. These models are evaluated comparatively to determine the most efficient architecture for real-time, non-destructive quality inspection in post-harvest processing.

3.1 DATA AUGMENTATION AND PRE-PROCESSING

Before training, considerable data augmentation and pre-processing methods were employed to enhance the model's robustness and generalization capacity. In the given onion image dataset's small size, data augmentation was very crucial. It included real-time image manipulations like rotation, zooming, shearing, width/height changing, and flipping images both vertically and horizontally. By simulating several angles and views that onions might display in real-world settings, these processes helped lower the likelihood of overfitting. Additionally, augmentation added diversity to the training set without requiring the time-and resource-intensive process of gathering new data.

All input images were downsized to a set resolution (224×224 pixels) for pre-processing in order to guarantee compatibility with our proposed model's input layer. In order to preserve crucial color qualities that are necessary for unhealthy identification, images were then transformed to RGB format, if they weren't already in RGB. The model was able to converge more effectively during training by normalizing pixel values to the range of [0,1] by splitting each pixel value by 255.0. Additionally, Tensor Flow's ImageDataGenerator tool was used to manage image rescaling and batching, enabling smooth on-the-fly augmentation during model training. The binary cross-entropy loss function is used to assemble these RGB models, and the RMSprop optimizer with a learning rate of 0.001 is used for optimization. Training leverages extensive data augmentation—such as rotation, zoom, shifts, shear, and horizontal flips—to artificially increase dataset variability and robustness, helping the models generalize better on unseen images.

In situations where the symptoms of visual diseases relied on minute variations in color, shape, or texture, this combination of pre-processing and augmentation not only increased training efficiency but also enhanced model performance.

3.2 FEATURE EXTRACTION

A vital stage in image classification process is a feature extraction, which gives the model ability to recognize and represent the most instructive patterns in the input images. In this work, pre-trained CNNs (convolutional neural networks) including proposed method, MobileNetV2, VGG16, and ResNet50—which were optimized for the onion health classification task—were used in order to extract characteristics. These models are renowned for extracting deep spatial information from images by using a large number of convolutional and pooling layers. The convolutional layers of the RGB image-based method automatically learned properties like color distributions, edges, forms, textures, and spatial hierarchies. Deep learning-based models extract both low-level like edges and corners and high-level like disease patterns and lesions features directly from the image data, which makes the system more flexible and scalable than traditional methods that rely on handcrafted features like color histograms or Gray-Level Co-occurrence Matrix (GLCM) features.

In these architectures, the Global Average Pooling (GAP) layer assisted in reducing the feature maps to a smaller dimension while keeping the most pertinent data. Following extraction, these features were sent to fully linked layers for categorization. The pre-trained weights from ImageNet enabled quicker convergence and enhanced accuracy through transfer learning, particularly in situations with little training data. All things considered, the feature extraction phase was crucial in helping the model differentiate between samples of onions that were healthy and those that weren't with high recall and precision.

3.3 MOBILENETV2

The lightweight CNN architecture known as MobileNetV2 was first created for embedded and mobile vision applications. The ImageNet dataset, which comprises more than a million photos in 1000 classes, is used to pre-train the standard MobileNetV2 model. It includes a series of depth wise separable convolutions and inverted residual blocks, making it extremely computationally and memory-efficient. In its default form, MobileNetV2 is designed for multi-class classification tasks and includes a top classification layer that outputs probabilities across all ImageNet classes. However, in our customized implementation, Transfer learning has been used to modify this underlying model to suit a binary classification problem specifically, to differentiate between healthy and unhealthy onions using RGB image data. The first layers of the original MobileNetV2 model are removed (include top=False), and instead, a new classification head is added. This includes a GlobalAveragePooling2D layer to decrease the feature dimensions, then a Dropout layer to prevent overfitting, and fully connected (Dense) layers with a final sigmoid activation function to output binary class predictions. Additionally, the weights of the pre-trained layers in the base MobileNetV2 are frozen, meaning only the newly added top layers are instructed on the onion dataset. This method makes learning more effective from a smaller, task-specific dataset while continuing to gain from the general feature extraction capability of MobileNetV2 trained on large-scale RGB data. The utilize of RGB images is crucial here, As long as the model can still identify minute color variations in the input data, which often carry important visual cues for identifying disease symptoms in onions.

The deep CNN architecture known as Google's MobileNetV2 was developed for situations where mobility and resources are scarce. It incorporates two essential innovations: inverted residual blocks and linear bottlenecks. The inverted residual block allows for more expressive features and lower dimensionality as an alternative to traditional residual connections. Information that may otherwise be lost as a result of activation function saturation is preserved by the linear bottlenecks at the conclusion of each inverted residual block. Beginning with a typical 3x3 convolution, the entire MobileNetV2 architecture progresses via a series of inverted residual blocks, followed by a fully linked layer, global average pooling layer, and 1x1 convolution. Because the ReLU6 activation function restricts output values to the interval [0, 6], it is especially useful for low-precision arithmetic in edge and mobile computing settings.

3.4 CLASSIFICATION BASED ON PROPOSED MODEL

In our proposed method, we conduct a comprehensive comparison of DL models trained on RGB onion images for binary classification of healthy versus unhealthy samples. For

RGB images, we employ state-of-the-art pre-trained CNNs proposed model with a focus on transfer learning. To incorporate a unique, task-specific classification head created for our onion health classification assignment, the previously trained model is loaded without its initial top classification layers. The input to these models consists of RGB images resized to 224×224×3, which is the standard input size for these architectures.

The base pre-trained networks extract rich hierarchical feature representations from the input images. Our proposed model employs efficient depth-wise separable convolutions and inverted residual blocks to capture complex spatial features with fewer parameters, while VGG16 and ResNet50 utilize deeper and more complex convolutional blocks. After the base feature extraction, a 2D layer for Global Average Pooling decreases spatial dimensions of feature maps to fixed-length vector, thereby reducing model parameters and mitigating overfitting. Following this, further avoid overfitting and enhance generalization, a Dropout layer with a rate of 0.5 randomly disables half of the neurons during training. This is succeeded by 512 neurons in a dense layer with ReLU activation that learns task-specific complex feature interactions. Finally, just one neuron the probability that the image is from the positive class is output by a dense layer that is activated by sigmoid, enabling binary classification.

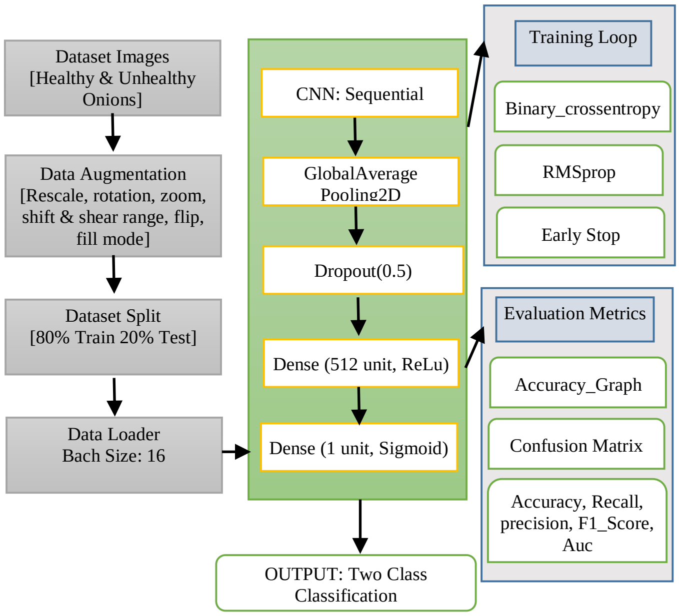

Fig.2. Architecture of proposed model

Table.2. Comparison of the performance of four DL models (Proposed model, MobileNetV2, VGG16, ResNet50) across four different datasets (Customized DS, F360, V360, OD)

	Proposed model				**MobileNetV2**				**VGG16**				**Resnet50**			
	CDS	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**
Accuracy	98.33	93.62	89.62	72.35	97.50	89.30	88.68	71.03	92.50	63.37	83.02	69.31	71.25	51.03	50.94	53.84
Precision	100	91.73	87.50	85.65	99.14	83.75	90.20	79.12	98.11	61.65	85.71	87.24	98.11	100	50.48	52.92
Recall	96.67	95.88	92.45	53.70	95.83	97.53	86.79	57.14	86.67	70.78	79.25	45.24	43.33	20.6	100	69.58
F1-Score	98.31	93.76	89.91	66.02	97.46	90.11	88.46	66.36	92.04	65.90	82.35	59.58	60.12	40.3	67.09	60.11
AUC	99.94	99.03	96.97	76.60	99.90	96.91	95.05	77.35	97.22	68.39	88.71	74.70	73.76	35.31	64.93	56.27

To avoid overfitting and maximize learning, training makes use of early halting and learning rate reduction call-backs. A test set is used to evaluate the models metrics with accuracy, precision, recall, AUC, and F1-score complemented by visualizations like accuracy curves, confusion matrices, and ROC curves. This approach allows us to systematically evaluate the effect of input color information and model architecture difficulty on classification performance. The pre-trained RGB models leverage rich color features and transfer learning to typically achieve higher accuracy and robustness.

4. RESULT AND DISCUSSION

In this our main aim is to classifying the healthy and unhealthy onions with CNN's assistance architecture with different models. There are several different techniques for evaluating the results by considering testing, examining and experimentation. The integration and interpretation of results, including the tests carried out, the results obtained, and the associated analytical insights, constitute the last phase of the system workflow. To assess the performance of the suggested algorithm, a number of methodical tests are conducted. The size of the dataset, its structural characteristics, and the type of input parameters given to the algorithm are some of the variables that affect this evaluation. In our approach, we are using supervised learning CNN model with proposed model as shown in the Fig.2.

4.1 DATASET CREATION

The dataset used in our work includes a customized dataset (CDS) and data collected from three publicly available sources on Kaggle: Fruit360, Onion-det, and Vegetable360. The CDS was manually curated and includes 1,200 images of onions in all, equally separated into 600 healthy and 600 unhealthy samples. The Fruit360 dataset contributed 809 healthy and 809 unhealthy onion images, while the Onion-det dataset provided 1,260 healthy and 1,260 unhealthy samples. Additionally, the Vegetable360 dataset included 176 healthy and 176 unhealthy onion images. These datasets were selected to ensure a diverse and balanced representation of onion conditions, covering various sources and imaging environments.

In the course of training, the dataset size was effectively doubled through augmentation techniques in both RGB and grayscale formats. After that, these datasets were utilized to carry out to perform comparative analysis across multiple deep learning models, including the proposed model, MobileNetV2, VGG16, and ResNet50. Each model's performance was evaluated on different datasets and image formats to regulate the most real approach for accurate onion health classification.

The table presents a qualified evaluation of four DL models are Proposed Model, MobileNetV2, ResNet50 and VGG16, according to five performance indicators: Accuracy, F1-Score, Recall, Precision, and AUC, across four onion datasets: Customized Dataset's (CDS), Fruit 360 (F360), Vegetable 360 (V360), and Onion-det (OD).

To assess the different DL models' classification performance in detecting onion diseases, a comprehensive comparison was completed using four datasets: CDS, F360, V360, and OD, across three popular pre-trained CNN architectures: MobileNetV2, VGG16, and ResNet50. In addition, a proposed custom model was included for benchmarking. The evaluation was based on five performance metrics: Accuracy, F1-Score, Recall, Precision, and AUC (Area Under the Curve). Across all datasets in Table 2, the suggested model showed better performance, particularly on the CDS dataset, where it reached an outstanding accuracy of 98.33%, perfect precision of 100%, high recall of 96.67%, an F1-score of 98.31%, and an almost perfect AUC of 99.94%. This suggests that the model did more than just forecast the classes correctly in most cases but also minimized false positives and false negatives significantly. MobileNetV2 followed closely, with accuracies of 93.62% and 89.62% on F360 and V360 respectively, and maintained strong precision and recall values across these datasets. Its lightweight architecture and efficient feature extraction likely contributed to its robust and consistent performance as shown in Fig.3.

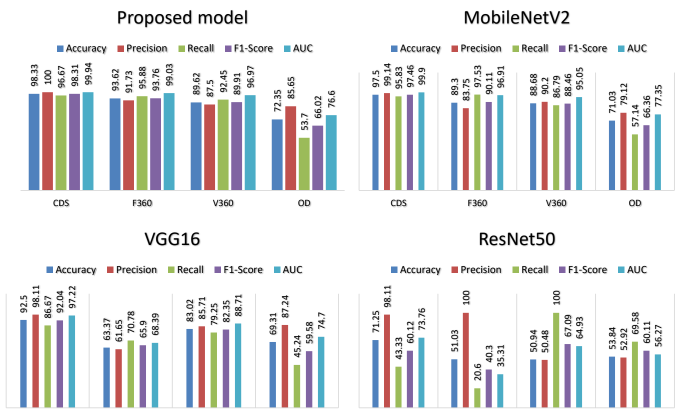

Fig.3. Performance Comparison of Deep Learning Models Across Four Onion Disease Datasets

VGG16 and ResNet50, although successful on the CDS dataset with accuracies of 97.50% and 92.50% respectively, exhibited more variability on challenging datasets like OD and V360. For example, VGG16's accuracy dropped to 71.03% on the OD dataset, with a sharp decline in recall (20.6%), suggesting that it failed to correctly identify many true positive cases, possibly due to overfitting or limited feature adaptability. Similarly, ResNet50's performance on the OD dataset was suboptimal, achieving only 69.31% accuracy and an AUC of 56.27%, reflecting a reduced ability to distinguish between classes in more complex or noisy environments. Interestingly, ResNet50 showed a recall of 100% on V360, but this came at the cost of low precision (50.48%), implying that while it identified all positives, it also misclassified many negatives, and taking the lead to a high false positive rate.

The proposed model's better efficacy across all datasets is further supported by the F1-score, which strikes a compromise between recall and precision. For example, its F1-score was 98.31% on CDS, whereas the highest F1-score for ResNet50 on OD was only 60.11%, indicating inconsistent performance. The model's ability is gauged by the AUC values to distinguish among classes across thresholds, reinforce these findings: the proposed model achieved the highest AUCs consistently, while pre-trained models saw significant drops in more complex datasets. These outcomes suggest that while pre-trained networks like MobileNetV2 can provide a reliable baseline for transfer learning, the suggested model is more adept at capturing the domain-specific features of onion disease images, particularly when the datasets are relatively clean and balanced. However, the sharp decline in performance on the OD dataset across all models highlights the need for additional pre-processing, data balancing, or advanced augmentation techniques to address the variability and noise in real-world images. Overall, the suggested model demonstrates excellent generalization and outperforms standard pre-trained architectures in both controlled and moderately complex conditions.

Fig.4 illustrates the accuracy of validation and training curves of four models: the proposed RGB-based model, MobileNetV2, VGG16 and ResNet50. The projected model (a) demonstrates higher performance, achieving a validation accuracy of approximately 98.5% with minimal gap between training and validation curves, indicating strong generalization and stability across epochs. Similarly, MobileNetV2 (b) shows a smooth and consistent rise in both training and validation accuracy, converging close to 97.5%, highlighting its efficiency and suitability for image classification tasks on RGB datasets.

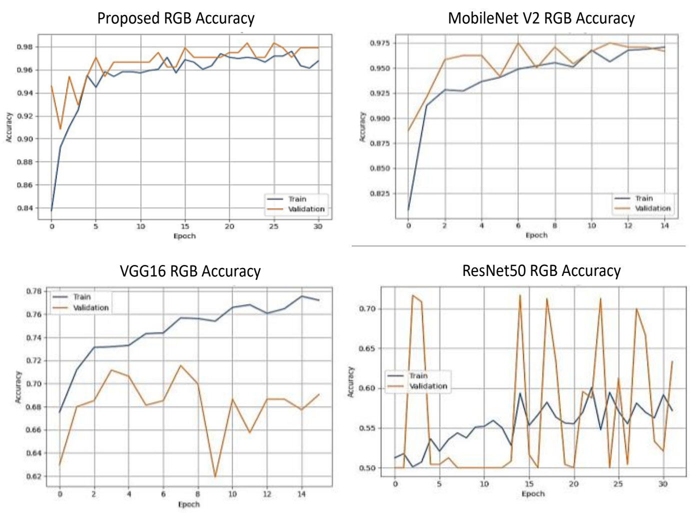

Fig.4. Training and Validation Accuracy Curves of Deep Learning Models Using RGB Images for our Customized datasets

However, in another case (c), VGG16 trained on a different dataset exhibits a notable gap between training (up to 79%) and validation accuracy (nearly 65–72%), suggesting overfitting and reduced generalization, possibly due to dataset imbalance or noise. In contrast, ResNet50 (d) shows highly unstable and fluctuating performance throughout training, with both training and validation accuracy failing to converge beyond 60%. This inconsistency indicates challenges in learning and generalizing, possibly due to high model complexity or inappropriate hyper parameters for the given dataset. Overall, the suggested model and MobileNetV2 (on CDS/F360) outperform the others regarding of accuracy, consistency, and robustness.

Table.3. Comparative analysis for different models with different datasets excluding RGB
(CDS: Custom Datasets, F360: Fruit360, V360: Vegetable360, OD: Onion-Det)

	Proposed model				**MobileNetV2**				**VGG16**				**Resnet50**			
	CDS	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**
Accuracy	97.92	90.74	87.74	71.30	96.67	86.01	89.62	69.58	92.50	61.32	81.13	69.71	59.17	50	50	54.63
Precision	100	87.22	84.48	82.86	100	78.88	86.21	77.01	99.04	60.15	80	85.65	57.05	50	50	54.28
Recall	95.83	95.47	92.45	53.70	93.33	98.35	94.34	55.82	85.83	67.08	83.02	47.35	65.83	100	100	58.73
F1-Score	97.87	91.16	88.29	65.17	96.55	87.55	90.09	64.72	91.96	63.42	81.48	60.99	64.49	66.67	66.67	56.42
AUC	99.97	98.55	96.55	75.73	99.74	96.33	96.05	75.52	97.83	68.39	87.82	74.2	73.84	49.51	64.54	55.62

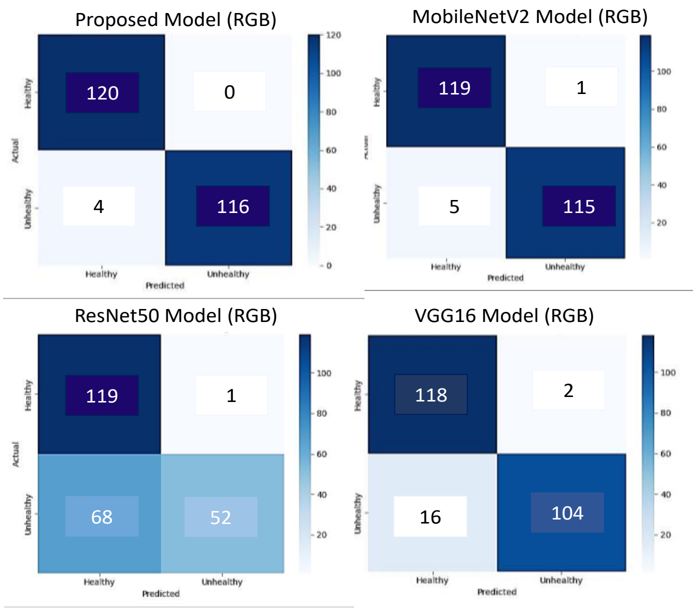

Fig.5. Confusion Matrix for DL Models Using RGB Images for our Customized datasets (a) proposed model for customized datasets, (b) MobileNetV2 for customized datasets, (c) ResNet50 for customized datasets, (d) VGG16 for customized datasets

The Fig.5 illustrates the confusion matrices of four different deep learning models namely the Proposed Model, MobileNetV2, VGG16, and ResNet50 trained on RGB onion images for classifying them as healthy or unhealthy. The Proposed Model demonstrates the highest classification performance, correctly identifying all 120 healthy onions and 116 out of 120 unhealthy onions, with only 4 misclassifications, resulting in no false negatives. This indicates strong precision and recall, making the model highly reliable for real-world classification tasks. MobileNetV2 also performs well, with 119 true positives and 115 true negatives, accompanied by only 1 false negative and 5 false positives, indicating slightly reduced performance compared to the Proposed Model. In contrast, VGG16 shows a decline in performance, with 16 unhealthy onions incorrectly classified as healthy and 2 healthy onions misclassified, indicating lower specificity and a higher risk of false positive identification. ResNet50, although achieving a comparable number of true positives (119), significantly underperforms in detecting unhealthy onions, with 28 false positives and only 92 true negatives. Because of its lowest specificity, this model might not be appropriate for applications that demand high disease detection reliability. Overall, the Suggested Model overtakes the other models with the peak classification accuracy, followed closely by MobileNetV2, while VGG16 and ResNet50 show comparatively lower performance due to increased misclassification rates.

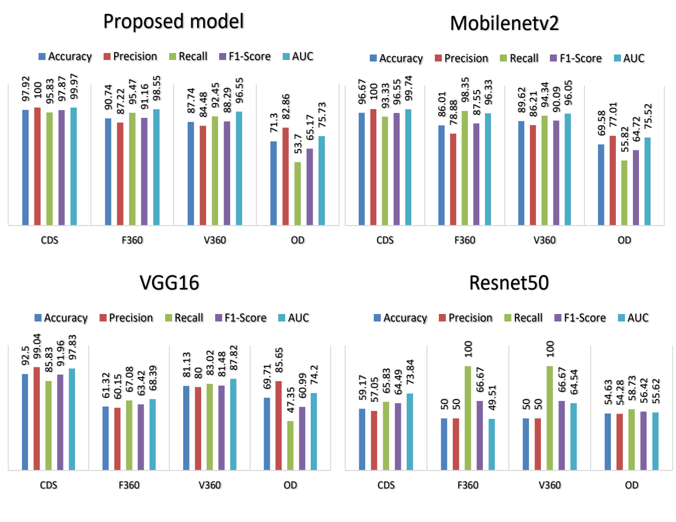

Fig.6. Graphical representation of evaluation metrics

4.2 ABLATION STUDY

4.2.1 Without RGB Images:

This was carried out without utilizing RGB images means the original images and achieved lower accuracy in contrast to the RGB format as shown in Table.3. The accuracy across all categories, notably 97.92% in CDS and above 87% in F360 and V360.

Table.4. Comparative analysis for different models with different datasets including grey scale

	Proposed model				**MobileNetV2**				**VGG16**				**Resnet50**			
	CDS	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**	**CDS**	**F360**	**V360**	**OD**
Accuracy	97.92	93.00	91.51	70.63	97.5	90.95	89.62	69.71	92.92	57.00	79.25	67.20	58.33	48.15	63.21	56.75
Precision	99.15	91.30	87.93	87.50	100	86.72	90.38	85.31	98.13	56.44	82.98	75.39	55.49	43.48	71.88	55.92
Recall	96.67	95.06	96.23	48.15	95	96.71	88.68	47.62	87.50	61.32	73.58	51.06	84.17	12.35	43.40	63.76
F1-Score	97.89	93.15	91.89	62.12	97.44	91.44	89.52	61.12	92.51	58.78	78	60.88	66.89	19.23	54.12	59.58
AUC	99.80	98.54	97.12	75.68	99.94	98.69	97.19	74.13	98.34	60.62	86.01	72.57	74.35	48.93	70.45	56.87

Even in the most challenging category, OD, it maintains 71.30% accuracy, along with strong precision (82.86%) and AUC (75.73%), indicating a balanced and reliable classification capability.

MobileNetV2 also shows strong performance, with high accuracy (96.67% in CDS and 89.62% in V360) and an exceptional recall of 98.35% in F360, making it particularly effective at identifying true positives. However, its performance in the OD category slightly lags behind the Modified Model. VGG16 delivers moderate results, performing fairly well in CDS and V360 but showing a noticeable drop in accuracy and recall for F360 and OD, indicating some inconsistency across categories. In contrast, ResNet50 performs the weakest, with only 59.17% accuracy in CDS and merely 50% in both F360 and V360. Although it achieves perfect recall in F360 and V360, this comes at the cost of extremely low precision (50%), resulting in a high false positive rate and poor F1-Scores and AUC values as shown in Fig.6.

The Fig.7 presents the confusion matrices of Customized datasets for four deep learning models like Proposed model, MobileNetV2, VGG16, and ResNet50—used to However, it also includes 1 false positive and 17 false negatives, indicating slightly lower recall for the unhealthy class compared to the MobileNetV2 models. In contrast, the ResNet50 model exhibits noticeably weaker performance, with only 89 healthy and 79 categorize images of onion as either healthy or unhealthy. . In our proposed model, the classifier accurately identified 120 healthy onions and 115 unhealthy onions, with 6 false positives and 5 false negatives. The MobileNetV2 shows slightly better performance, perfectly classifying all 120 healthy samples and misclassifying only 8 unhealthy ones, achieving strong overall precision and recall. For the VGG16 model, the matrix shows high accuracy with 119 correctly predicted healthy onions and 103 unhealthy ones. However, it also includes 1 false positive and 17 false negatives, indicating slightly lower recall for the unhealthy class compared to the MobileNetV2 models. In contrast, the ResNet50 model exhibits noticeably weaker performance, with only 89 healthy and 79 unhealthy onions correctly identified. With 31 false positives and 41 false negatives, it has the highest amount of misclassifications, indicating that it is difficult to reliably distinguish between the two classes.

4.2.2 For Grey Scale Images:

The models trained on original RGB images and those trained on grayscale images reveals a clear advantage in retaining full color information during training. When using RGB images directly without conversion or removal of color the models demonstrated steadily improved performance in every evaluation metrics and datasets has changed to grey scale image as shown in Table 4. For instance, the maximum accuracy was attained using the suggested model of 98.33% on CDS and 93.00% on F360**,** compared to 97.92% and 90.74% respectively with grayscale input. Similarly,

MobileNetV2 and VGG16 exhibited stronger results with RGB images, particularly on datasets where color is essential in distinguishing healthy from unhealthy samples. Notably, MobileNetV2 showed an accuracy of 91.51% on V360 with RGB, while this dropped to 89.62% in grayscale mode.

The confusion matrices for the grayscale models illustrate in Fig.8, the performance of four different architectures Proposed Model, MobileNetV2, VGG16, and ResNet50 when trained and evaluated using grayscale onion images. The proposed model demonstrated the highest accuracy, correctly classifying 120 healthy and 115 unhealthy onions, with only three misclassifications in total, indicating excellent precision and recall even without color information. MobileNetV2 also showed strong performance, achieving 119 correct predictions for healthy and 114 for unhealthy onions, with a balanced number of six false positives and six false negatives. VGG16 followed closely, accurately identifying 118 healthy and 109 unhealthy onions, although it recorded a somewhat higher number of misclassifications, particularly 11 false negatives. However, ResNet50 showed the weakest performance among the four, correctly classifying only 93 healthy and 103 unhealthy onions, while misclassifying a notable number of samples 27 healthy and 13 unhealthy. These findings imply that the proposed model and MobileNetV2 are more robust to grayscale input, whereas VGG16 has moderate adaptability, and ResNet50 appears to be more dependent on color features for effective classification.

Further, key metrics like F1-score, recall and precision were generally higher for RGB-trained models. For example, in the OD dataset, the suggested model recorded a precision of 87.50% and F1-score of 62.12 with RGB input, outperforming the grayscale-trained model which achieved 82.86% precision and 65.17 F1-score—highlighting that while grayscale may sometimes offer higher recall, RGB provides a better balance across metrics. The AUC (Area Under the Curve) values further confirm this pattern; the proposed model reached 99.85 on CDS and 98.54 on F360 with RGB, compared to 99.97 and 98.55 in grayscale indicating slightly more stable and generalizable learning with RGB images as shown in Fig.9.

The models trained directly on RGB images captured richer spatial and chromatic features, leading to superior classification performance, especially in complex datasets. This demonstrates the significance of color cues in accurately detecting onion diseases and validates the decision to retain RGB input in model training. Overall, the Modified Model stands out as the most robust and consistent performer across all onion classes. MobileNetV2 follows closely, particularly excelling in recall. VGG16 shows moderate potential but struggles with category-specific performance, while ResNet50 exhibits considerable limitations in accuracy and balanced classification, making it the least effective among the compared models.

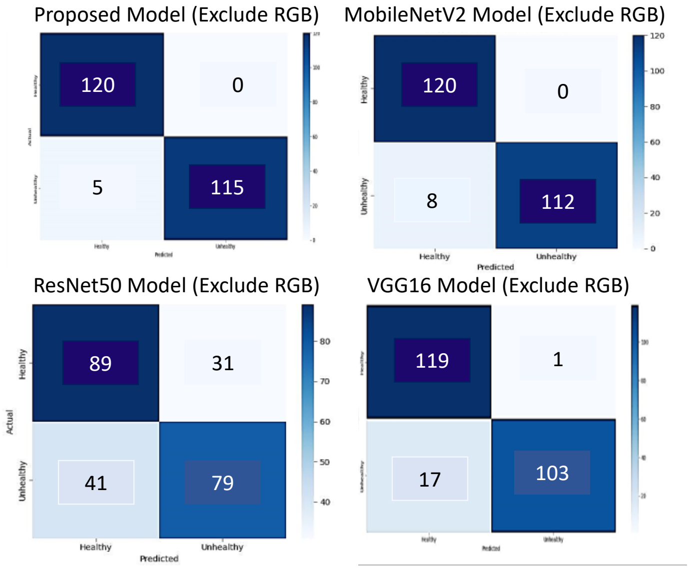

Fig.7. Confusion matrix for excluding RGB format (a) customized MobileNetV2 for customized datasets, (b) MobileNetV2 for customized datasets, (c) ResNet50 for customized datasets, (d) VGG16 for customized datasets

Fig.8. Confusion matrix for grey-scale images (a) Customized MobileNetV2 for customized datasets, (b) MobileNetV2 for customized datasets, (c) ResNet50 for customized datasets and (d) VGG16 for customized datasets

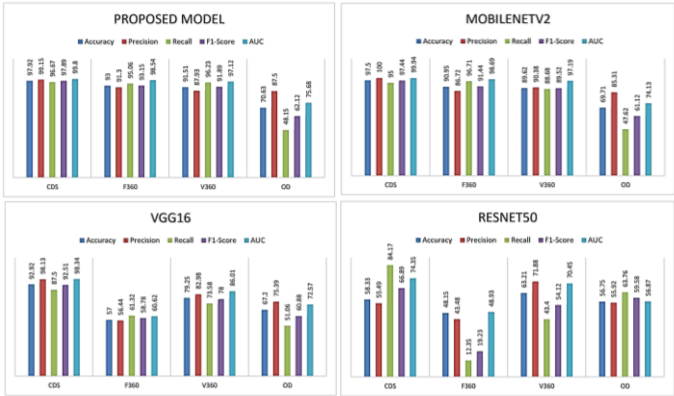

Fig.9. Graphical representation of evaluation metrics

5. CONCLUSION

In this work, the performance of various DL models were assessed using RGB, grayscale, and non-RGB (color-filtered) image formats across four onion disease datasets (CDS, F360, V360, and OD). The results consistently showed that models trained on RGB images achieved the highest accuracy, F1-score, precision, and AUC across all datasets. The suggested model, for example, greatly outperformed other formats on the CDS dataset utilizing RGB photos, with an accuracy of 98.33% and an AUC of 99.94%. Models like MobileNetV2 and VGG16 also showed robust performance under RGB, especially with relation to recall and F1-score. In contrast, the performance of models trained on grayscale images, while slightly lower than RGB, remained comparatively competitive, particularly for the proposed model and MobileNetV2. Grayscale inputs yielded higher precision and F1-scores than non-RGB formats and performed notably well on datasets like F360 and CDS, suggesting that spatial and texture features remain effective even without color. However, models trained on non-RGB formats showed the lowest performance, especially on complex datasets like OD and V360, where recall and AUC dropped significantly. For example, the ResNet50 model on OD with non-RGB input recorded only 50.94% accuracy and AUC of 56.27%. This suggests that color information is essential to the discrimination of diseases. Overall, RGB remains the most informative input format, but grayscale can serve as a lightweight alternative with reasonably good performance. Non-RGB formats, though less effective in this context, may still be useful in specific scenarios with further optimization.

REFERENCES

[1] Food and Agriculture Organization of the United Nations, “Good Agricultural Practices for Onion and Garlic Production” Rome: FAO, 2022”, Available at: https://www.manage.gov.in/publications/eBooks/Good%20Agricultural%20Practices%20in%20Onion%20and%20Garlic.pdf, Accessed on 2025.

[2] J. Sun, R. Kunnemeyer, A. McGlone, N. Tomer and K. Sharrock, “A Spatially Resolved Transmittance Spectroscopy System for Detecting Internal Rots in Onions”, *Postharvest Biology and Technology*, Vol. 163, pp. 111141-111163, 2020.

[3] M. Singh, R.B. Kale and V. Mahajan, "Status of Onion and Garlic Production in India and World", Available at: rajiv.kale@icar.gov.in, Accessed on 2025.

[4] L. Goyal, C.M. Sharma, A. Singh and P.K. Singh, "Leaf and Spike Wheat Disease Detection and Classification using an Improved Deep Convolutional Architecture", *Informatics in Medicine Unlocked*, Vol. 25, pp. 100642-100653, 2021.

[5] S.E. Pawar, V. Surana, A. Sharma and R. Pujeri, "Fruit Disease Detection and Classification using Machine Learning and Deep Learning Techniques", *Proceedings of International Journal of Intelligent Systems and Applications in Engineering*, Vol. 12, No. 4, pp. 440-453, 2023.

[6] M. Astani, M. Hasheminejad and M. Vaghefi, "A Diverse Ensemble Classifier for Tomato Disease Recognition", *Computers and Electronics in Agriculture*, Vol. 198, pp. 107054-107065, 2022.

[7] M. Hassam, M.A Khan, A. Armghan, S.A. Althubiti, M. Alhaisoni and A. Alqahtani. "An Single Stream Modified MobileNet V2 and Whale Controlled Entropy based Optimization Framework for Citrus Fruit Diseases Recognition", *IEEE Access*, Vol. 10, pp. 91828-91839, 2022.

[8] H. Yu, X. Cheng, C. Cheng, A.A. Heidari, J. Liu, Z.N. Cai and H. Chen, "Apple Leaf Disease Recognition Method with Improved Residual Network", *Multimedia Tools and Applications*, Vol. 81, pp. 7759-7782, 2022.

[9] Y. Zhong and M. Zhao, "Research on Deep Learning in Apple Leaf Disease Recognition", *Computers and Electronics in Agriculture*, Vol. 168, pp. 105146-105165, 2020.

[10] S. Mishra, T. Hailu, V. Ellappan, D. Rathee and H. Kalla, "Avocado Fruit Disease Detection and Classification using Modified SCA-PSO Algorithm-based MobileNetV2 Convolutional Neural Network", *Iran Journal of Computer Science*, Vol. 5, pp. 345-358, 2022.

[11] E. Elfatimi, R. Eryigit and L. Elfatimi, "Beans Leaf Diseases Classification using MobileNet Models", *IEEE Access*, Vol. 10, pp. 9471-9482, 2022.

[12] V. Singh, A. Chug and A.P. Singh, "Classification of Beans Leaf Diseases using Fine-Tuned CNN Model", *Procedia Computer Science*, Vol. 218, pp. 348-356, 2023.

[13] B.V. Nikith, N.K.S. Keerthan, M.S. Praneeth and T. Amrita, "Leaf Disease Detection and Classification", *Procedia Computer Science*, Vol. 218, pp. 291-300, 2023.

[14] G. Mukherjee, A. Chatterjee and B. Tudu, "Identification of the types of Disease for Tomato Plants using a Modified Gray Wolf Optimization Optimized MobileNetV2 Convolutional Neural Network Architecture Driven Computer Vision Framework", *Concurrency and Computation: Practice and Experience*, Vol. 34, pp. 1-22, 2022.

[15] M. Shantkumari and S. Uma, "Grape Leaf Image Classification based on Machine Learning Technique for Accurate Leaf Disease Detection", *Multimedia Tools and Applications*, Vol. 82, pp. 1477-1487, 2023.

[16] Q. Wu, Y. Chen and J. Meng, "DCGAN based Data Augmentation for Tomato Leaf Disease Identification", *IEEE Access*, Vol. 8, pp. 98716-98728, 2020.

[17] Y. Kurmi, S. Gangwar, D. Agrawal, S. Kumar and H. Shrivastava, "Leaf Image Analysis-Based Crop Diseases Classification", *Signal, Image and Video Processing*, Vol. 15, pp. 589-597, 2021.

[18] A.D. Lestari, A.U. Khan, D.A.A. Pertiwi and M.A. Muslim, "A New CNN Model Integrated in Onion and Garlic Sorting Robot to Improve Classification Accuracy", *Journal of Soft Computing Exploration*, Vol. 5, No. 1, pp. 80-85, 2024.

[19] M. Shahin, E. Tollner, R. Gitaitis, D. Sumner and B. Maw, "Classification of Sweet Onions based on Internal Defects using Image Processing and Neural Network Techniques", *Transactions of the ASAE*, Vol. 45, No. 5, pp. 1613-1618, 2002.

[20] M. Zaki, S. Narejo, M. Ansari, S. Zai, M. Anjum and N. Memon, "Image-Based Onion Disease (Purple Blotch) Detection using Deep Convolutional Neural Network", *International Journal of Advanced Computer Science and Applications*, Vol. 12, pp. 448-458, 2021.

[21] W. Wang, C. Li, E. Tollner, R. Gitaitis and G. Rains, "Shortwave Infrared Hyperspectral Imaging for Detecting Sour Skin (Burkholderia Cepacia)-Infected Onions", *International Journal of Food Engineering*, Vol. 109, pp. 36-48, 2012.

[22] W. Wang and C. Li, "A Multimodal Machine Vision System for Quality Inspection of Onions", *Journal of Food Engineering*, Vol. 166, pp. 291-301, 2015.

[23] B. Deplomo, J. Dela Cruz and J. Balbin, "Classifying Quality Grading and Size of Allium Cepa using Digital Image Processing Algorithms", *Proceedings of International Conference on Recent Trends on Computer Science*, pp. 123-126, 2020.

[24] Kaggle Dataset, Available at: https://www.kaggle.com/datasets/emmarex/plantdisease, Accessed on 2025.

[25] Kumar Siddamallappa and Anusha Jajur, "A Comprehensive Study on Onion Classification based on Various Algorithms", *GRENZE International Journal of Engineering and Technology*, Vol. 11, No. 2, pp. 13239-13245, 2025.