RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 59 Issue 4: 628-635 (April 2025)

Enhancing Apple Disease Detection with Pixel Pyramid Net through Pixel-Level Segmentation and Spatial Pyramid Pooling Techniques

C. Radhiya Devi[1], S.K. Jayanthi[1]

10.18805/IJARe.A-6248

ABSTRACT

Background: Apple fruit is prone to various diseases that can significantly affect fruit quality.

Methods: This paper proposes a novel approach for apple disease detection utilizing PPN-Pixel Pyramid Net, a neural network architecture designed for pixel-level segmentation tasks. The methodology incorporates spatial pyramid pooling techniques to effectively address input images of diverse sizes and aspect ratios.

Result: Experimental results showcase the model's performance metrics across different classes, providing an in-depth analysis of Accuracy, Intersection over unionand Mean BF Score.

Key words: Apple disease, Pixel Pyramid Net, Pixel-level segmentation; Spatial pyramid pooling.

INTRODUCTION

Artificial Intelligence (AI) has significantly influenced Indian agriculture by enhancing productivity, efficiencyand sustainability. AI applications such as crop monitoring and disease detection, powered by image recognition and machine learning, have enabled early detection of issues, improving crop health management. Precision farming techniques, including AI-driven irrigation and soil management, optimize resource use and boost yields. Market prediction algorithms assist farmers in making informed decisions regarding crop selection and pricing. Despite these advancements, challenges such as high costs, lack of infrastructureand the need for technical expertise remain barriers to widespread AI adoption in Indian agriculture (Chandra Vinod, 2023).

Machine learning (ML) models have revolutionized plant disease prediction and detection, significantly improving accuracy and efficiency. Supervised learning techniques, such as support vector machines (SVMs) and decision trees, are commonly used for classifying plant diseases based on image data. Convolutional neural networks (CNNs) have also proven highly effective in recognizing disease patterns in plant leaves. Additionally, ensemble methods, which combine multiple algorithms, enhance prediction reliability. Despite these advancements, challenges like data quality, computational requirements and the need for large labeled datasets remain significant hurdles are to be addressed as suggested by Shivappa and Gyanappa (2023).

Apple disease detection is crucial for ensuring crop health and yield optimization in agricultural sectors. Traditional methods for disease identification often rely on manual inspection, which is labor-intensive and time-consuming. In recent years, advancements in computer vision and machine learning have provided opportunities to automate this process using image processing techniques.

This work proposes PPN-Pixel Pyramid Net, a novel neural network architecture designed specifically for pixel-level segmentation tasks in apple disease detection. The architecture incorporates spatial pyramid pooling techniques to handle input images of varying sizes and aspect ratios, overcoming the limitations of fixed-size input requirements in traditional convolutional neural networks (CNNs). By dividing the input feature map into fixed-size grids and applying adaptive pooling within each grid, the model can effectively capture local spatial information and maintain detailed representations of the input, contributing to improved performance and robustness.

There are several techniques available to segment diseased apple fruit images, ranging from traditional image processing methods to more advanced deep learning approaches. In their research, (Han *et al.,* 2021) developed a region-aggregated attention CNN for disease detection

[1]Department of Computer Science, Vellalar College for Women, Erode-638 012, Tamil nadu, India.

Corresponding Author: C. Radhiya Devi, Department of Computer Science, Vellalar College for Women, Erode 638 012, Tamil nadu, India. Email:radhiyadevic@gmail.com

How to cite this article: Devi, C.R. and Jayanthi, S.K.(2025). Enhancing Apple Disease Detection with Pixel Pyramid Net through Pixel-Level Segmentation and Spatial Pyramid Pooling Techniques. Indian Journal of Agricultural Research. 59(4): 628-635. doi: 10.18805/IJARe.A-6248.

Submitted: 30-04-2024 **Accepted:** 01-07-2024 **Online:** 23-07-2024

in fruit images, enhancing crop yield and quality. Their method, based on Mask R-CNN, incorporates additional feature pyramids, feature map aggregationand a squeeze-and-excitation block. Emphasizing the importance of timely disease detection in agriculture, the study underscores the limitations of traditional methods and the potential of deep learning.

The research (Mohana Priya *et al.,* 2023) addresses the complexities involved in segmenting various components of maize leaves for disease detection using semantic segmentation. The methodology involves class labeling, constructing image data repositoriesand training neural networks. Evaluation metrics such as Global Accuracy and Mean IOU showcase the superior performance of the proposed model compared to FCN-8s. This approach holds significant potential in agricultural applications by accurately identifying components and improving disease detection outcomes. However, further refinement is needed, particularly in lesion detection.

Kanagaraju *et al*. (2022), presents an image processing method using Convolutional Neural Network (CNN) for apple disease detection. Steps include preprocessing, area identification, highlighting affected regions and showing the result. Various diseases are testedand K-means clustering is compared with CNN for accuracy and time efficiency. Feature extraction techniques like Global Color Histogram (GCH), Color Coherence Vector (CCV), Local Binary Pattern (LBP) and Complete Local Binary Pattern (CLBP) are used to improve accuracy. CNN is used for segmentation and SVM aids in disease classification.

Dubey *et al.* (2012) propose an image processing solution for detecting and classifying apple fruit diseases, focusing on apple blotch, rot and scab. The approach involves K-Means clustering for segmentation, feature extraction and Multi-class Support Vector Machine (SVM) for training and classification. Features like Global Color Histogram (GCH), Color Coherence Vector (CCV), Local Binary Pattern (LBP) and Complete Local Binary Pattern (CLBP) are utilized, with CLBP showing the highest accuracy, notably in the HSV color space. Further enhancements are suggested for method refinement.

Dubey *et al.* (2016) propose an approach for apple disease classification using color, texture and shape-based features. The primary steps are detecting infected fruit parts with the help of K-Means clustering method, computing color, texture and shape-based features over the segmented image and combining them to form a single descriptorand using multi-class support vector machine to classify the apples into one of the infected or healthy categories. For this study, apple fruit is taken as a test case with categories of diseases, namely blotch, rotand scab as well as healthy apples.

Alharbi *et al.* (2020) address the classification of apple diseases using computer vision and deep learning techniques, focusing on fungal diseases like apple scab, blotch and rot. It compares various Convolutional Neural Network (CNN) models for accuracy and efficiency. Results indicate that model-5 achieved the highest accuracy also showed superior performance in terms of training and testing time. The study highlights the effectiveness of CNN in accurately classifying healthy apples and identifying various apple diseases, offering a promising approach for agricultural disease management.

Organic cultivation significantly enhances the quality of fruits and vegetables by improving their nutritional value and taste. Research indicates that organically grown produce contains higher levels of vitamins, minerals and antioxidants compared to conventionally grown counterparts. Organic farming practices, which eschew synthetic pesticides and fertilizers, contribute to this nutritional superiority. Additionally, organic methods promote environmental sustainability by preserving soil health and reducing chemical runoff. Despite these benefits, organic farming faces challenges such as lower yields and higher production costs as given by Mohanapriya *et al.* (2024).

This section has provided an overview of the different algorithms associated with the detection of plant leaf diseases. The overall literature survey says that various methods and classification techniques are applied and in the proposed, PPN-Pixel Pyramid Net, a neural network architecture designed for pixel-level segmentation tasks, is utilized.

Materials and Methods

The research was conducted at the research laboratory in Vellalar College for Women, Erode, Tamil Nadu, India. The experiments utilized an Intel Core i3 64-bit processor operating at 2.30 GHz with 4GB of RAM, under the Windows 10 operating system. MATLAB 2021a was employed as the primary software for the research, which was carried out from February 2023 to April 2024.

The architecture describes a neural network model for image processing, particularly suited for pixel classification. Fig 1 elucidates the methodology employed for Apple disease detection through Pixel level Segmentation and Spatial Pyramid Pooling Techniques.

1. Input image

The PPN-PixelPyramidNet processes RGB images of apples, considering each image with 3 channels as input.

2. Convolutional layer with RELU

The apple image is convolved with two 3×3 filters to extract features, followed by applying 64 convolution filters of size 3×3 with a stride of [1, 1] and padding to maintain spatial dimensions. A rectified linear unit (RELU) activation function is then applied.

3. Spatial pyramid pooling (SPP) layer

Spatial Pyramid Pooling (SPP) is a technique employed in convolutional neural networks (CNNs) to accommodate

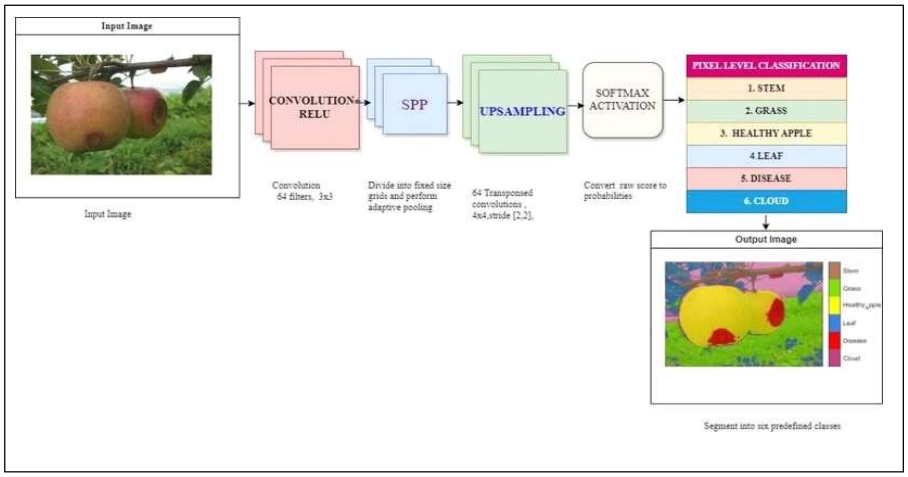

Fig 1: Architecture of PPN-Pixel pyramid Net.

input images of diverse sizes and aspect ratios without the necessity for resizing. Unlike conventional pooling layers with fixed-size requirements, SPP divides the input feature map into distinct grids and conducts pooling operations within each grid. By utilizing 2 pyramid levels in this architecture, the input feature maps are split into two sets of grids, resulting in a reduction to a 2×2 output size. Consequently, the pooled features are consolidated into a 2×2 grid, enabling effective handling of inputs with varying dimensions and aspect ratios.

The spatial pyramid pooling (SPP) layer involves a combination of operations such as grid division, adaptive poolingand concatenation.

Input

The input feature map, with dimensions H×W×C, represents an image where H denotes height, W represents width and C signifies the number of channels. Each channel captures specific information, while the height and width define the spatial dimensions, forming the basis for subsequent processing.

Grid division

In grid division, the feature map representing an apple image is divided into fixed-size grids or regions, with each grid capturing a specific level of spatial information. Initially, the dimensions of the grid, such as its width and height, are determined. Then, the feature map is subdivided into these predefined grid cells, where each cell corresponds to a distinct region of the image. This segmentation enables the extraction of features at various spatial scales, facilitating tasks like identifying different parts of the apple or distinguishing it from the background. For example, smaller grids might capture fine details like texture, while larger grids encompass broader features such as overall shape or color of the apple. The granularity of the grids influences the level of detail captured during feature extraction.

Adaptive pooling

Adaptive pooling is applied to each grid independently after the feature map of an apple image has been divided into fixed-size regions. Unlike conventional pooling methods with fixed kernel sizes, adaptive pooling adjusts dynamically to the size of each grid. This flexibility allows the pooling operation to effectively capture the most significant features within each region, regardless of its dimensions. By performing pooling independently within each grid it ensures relevant information is retained while discarding less relevant details.

Concatenation

Concatenation is a step following adaptive pooling within each grid of an apple image's feature map. The pooled results from all grids are combined by stacking or joining the pooled feature maps into a unified structure. This process creates a fixed-size output regardless of the original input size. By concatenating these feature maps, the model can consistently process apple images of varying sizes, making it adaptable to different datasets and situations.

4. Upsampling

This layer which is also known as deconvolution or upsampling layer. It applies 64 transposed convolutions of size 4×4 with a stride of [2, 2] and cropping of [1, 1, 1, 1]. It helps in upsampling the feature maps to the size of the input image.

5. Softmax activation

Softmax activation is applied to the output of the preceding layer, transforming raw scores into probabilities. This activation function is widely used in classification tasks as it generates a probability distribution across different classes.

6. Pixel classification layer

This layer uses a class-weighted cross-entropy loss function for training the network. It's specifically designed for pixel-level classification tasks, with classes like 'Cloud', 'Disease', 'Leaf', 'Healthy_Apple', 'Grass' and 'Stem'.

RESULTS AND DISCUSSION

Table 1 shows the pixel count classes (PCC) distribution across different dataset classes. 'Grass' has the most

pixels (51556), while 'Disease' has the fewest (4140). 'Healthy_Apple' notably dominates the dataset with a high PCC. 'Cloud', 'Leaf' and 'Stem' have moderate PCC values, indicating they contribute moderately to the overall pixel distribution.

The bar chart depicted in Fig 2 visually represents the distribution of pixel count across different classes, offering insights into the relative presence of each class within the dataset.

In evaluating the performance of a segmentation model for diseased apple fruit across six classes Cloud, Disease, Leaf, Healthy_Apple, Grass, Stem, the key metrics Accuracy, Intersection over Union (IOU) and Mean Boundary F1 Score (Mean BF Score) are employed. The terms that are fundamental in key metrics are,

Table 1: Pixel count classes (PCC) across classes.

Classes	PCC
Cloud	10154
Disease	4140
Leaf	16218
Healthy_Apple	39680
Grass	51556
Stem	7812

Table 2: Comparative analysis of accuracy, IOU and mean BF score across six classes.

Classes	Accuracy	IOU	MeanBFScore
Cloud	63.75%	44.99%	84.15%
Disease	98.24%	64.12%	45.21%
Leaf	76.94%	54.01%	56.26%
Healthy_Apple	92.70%	90.55%	69.82%
Grass	87.22%	84.44%	52.92%
Stem	86.82%	68.76%	73.72%
Average	84.28%	67.81%	63.68%

TP (True Positives): Number of pixels correctly classified as diseased apple fruit.

TN (True Negatives): Number of pixels correctly classified as not diseased apple fruit (belonging to other classes).

FP (False Positives): Number of pixels incorrectly classified as diseased apple fruit (but belong to other classes).

FN (False Negatives): Number of pixels belonging to diseased apple fruit class but incorrectly classified as other classes.

Accuracy

Accuracy measures the percentage of correctly classified pixels among all the pixels in the segmentation masks. It's calculated as the ratio of the number of correctly classified pixels to the total number of pixels. In the context of apple diseased fruit segmentation, accuracy indicates how effectively the model identifies different classes such as diseased areas, healthy apple regionsand background elements like clouds or grass.

$$\text{Accuracy} = \frac{TP}{(TP + FN)}$$

For ex. the model achieved an accuracy of 0.98237 for the 'Disease' class, indicating that it correctly predicted approximately 98.24% of the pixels belonging to this class.

Intersection over union (IOU)

IOU measures the degree of overlap between the predicted segmentation mask generated by the model and the ground truth mask, which is either manually annotated or represents the true segmentation. It quantifies how well the predicted regions align with the actual regions of interest on apple trees, including disease spots, healthy areas, leaves, stemsand potential obstructions like clouds.

$$\text{IOU} = \frac{TP}{TP + FP+FN}$$

For ex. an IOU value of 0.90554 for the 'Healthy_Apple' class indicates a high degree of overlap between predicted

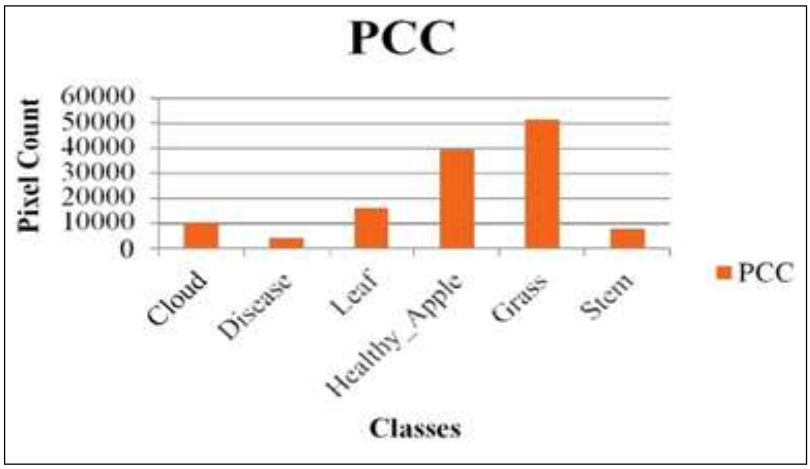

Fig 2: Barchart of distribution of Classes.

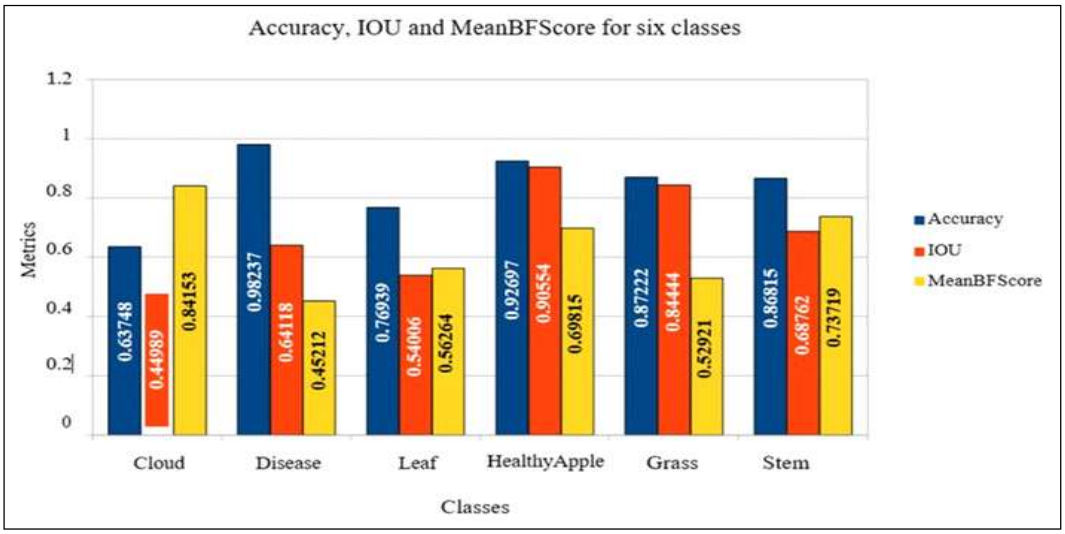

Fig 3: Comparative Barchart of Accuracy, IOU and Mean BF Score across six different classes.

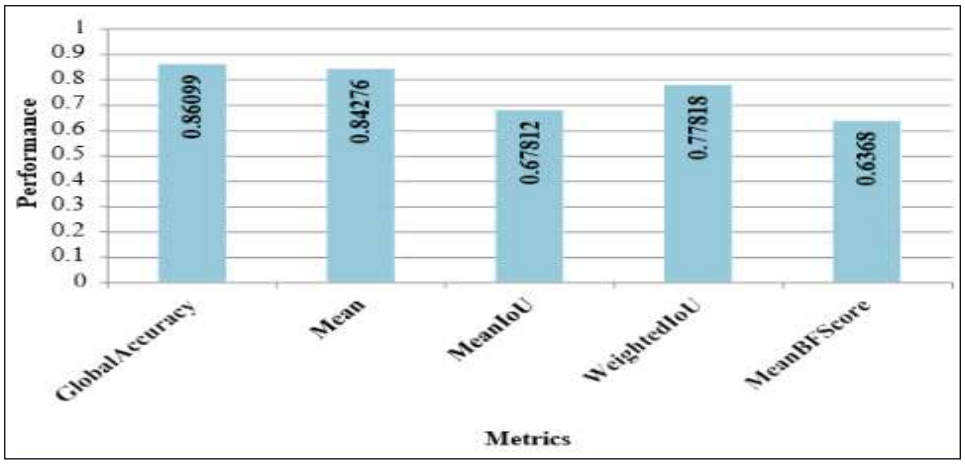

Fig 4: Barchart of various performance metrics.

Table 3: Performance metrics.

Global accuracy	Mean accuracy	Mean IoU	Weighted IoU	Mean BF score
86.10%	84.28%	67.81%	77.82%	63.68%

Table 4: Training evolution: Accuracy and loss over epochs.

Epoch	Accuracy	Loss
1	12.30%	5.09%
50	79.86%	0.67%
100	82.88%	0.55%
150	84.24%	0.51%
200	85.33%	0.47%
250	86.08%	0.45%

and ground truth masks. This means that around 90.55% of the pixels classified as healthy apple by the model correspond to actual healthy apple regions. This strong IOU value reflects the model's ability to accurately identify and delineate healthy apple regions, demonstrating its effectiveness in distinguishing between healthy and diseased areas. This high IOU value signifies strong performance in accurately segmenting healthy apple regions, crucial for apple disease detection tasks where distinguishing healthy areas is as important as identifying diseased regions.

Mean BF score (Boundary F1 score)

Mean BF score assesses the model's ability to accurately predict object boundaries, such as boundaries between diseased and healthy areas on apple. It computes the F1 score for boundary prediction, which is the harmonic mean of precision and recall. A higher mean BF score indicates better boundary delineation accuracy, which is crucial for accurately segmenting different features of interest in apple images.

Table 5: Precision-recall analysis of disease detection classes with AUC calculation.

Class	Precision	Recall	Interpolated Precision at IOU	AUC
Cloud	95.00%	90.00%	95.00%	85.50%
Disease	96.00%	92.00%	96.00%	88.32%
Leaf	97.00%	95.00%	97.00%	92.15%
Healthy_Apple	98.00%	94.00%	98.00%	92.32%
Grass	96.00%	93.00%	96.00%	89.28%
Stem	97.00%	91.00%	97.00%	88.27%
Average	96.50%	92.50%	96.50%	89.59%

Table 6: Comparative performance analysis of disease detection models using average precision.

Model	AP	AP 50
Mask R-CNN	66.91%	87.33%
SSD	52.88%	74.08%
Retinanet	54.87%	82.35%
YOLOv3	63.86%	88.62%
Region-aggregated attention CNN	72.26%	88.62%
Proposed PPN-PixelPyramidNet	89.31%	89.81%

Table 2 shows the comparative analysis of accuracy, IOU and Mean BF Score across six classes and Fig 3 shows the corresponding bar chart.

From the above table and the barchart, it is clear that the performance metrics provide a comprehensive evaluation of PPN-Pixel Pyramid Net's performance in detecting apple diseases through pixel-level segmentation with SPP. They evaluate pixel classification accuracy, mask prediction alignment and boundary prediction accuracy, aiding system evaluation and enhancement. For instance, the 'HealthyApple' class achieved a mean BF Score of 0.69815, indicating effective boundary prediction. 'Disease' and 'HealthyApple' show high accuracy and IOU, while 'Cloud' encounters challenges with lower accuracy and IOU, particularly in cloud detection.

Accuracy, IOU and mean BF score focus on individual object or pixel-level accuracy, while global accuracy, mean accuracy, mean IOU, weighted IOU and mean BF score provide aggregate evaluations over multiple instances or classes.

For the task of apple disease detection using PPN-Pixel Pyramid Net with the evaluation metrics such as,

Global accuracy

measures overall precision in identifying pixels across all classes including 'Cloud', 'Disease', 'Leaf', 'Healthy Apple', 'Grass' and 'Stem'.

Mean accuracy

evaluates average accuracy per class, considering the importance of identifying different areas like diseased, healthy, leaf, grass, stem and cloudy regions.

Mean IOU

Quantifies the average intersection over union for all classes, evaluating segmentation accuracy.

Weighted IOU

Gives more weight to classes with larger pixel counts like 'HealthyApple', 'Grass' and 'Leaf'

Mean BF Score

Evaluates boundary prediction accuracy for each class, crucial for delineating boundaries between healthy and diseased areas, leaves, stemsand potential obstructions like clouds.

Table 3 shows the performance metrics to assess effectiveness of the PPN-Pixel pyramid net model for apple disease detection and Fig 4 shows the corresponding bar chart.

Han *et al.* (2021) presented the disease detection results for apple images through visualizing bounding boxes. In their description, they noted that ground truth bounding boxes were denoted by dotted lines, while predicted bounding boxes were illustrated with solid-colored lines. They emphasized that red arrows highlighted false positive detections, signifying instances where the model erroneously identified disease when it was not present. Additionally, they mentioned that blue arrows indicated false negative detections, representing cases where the model failed to detect actual disease.

The proposed PPN-Pixel pyramid net method performs well in pixel-level semantic segmentation by reducing false positives and false negatives compared to the Region Aggregated CNN Han *et al.* (2021) method which is shown in Fig 5.

This Table 4 represents the training progress of a semantic segmentation neural network over multiple epochs. The table shows how both accuracy and loss change as the training progresses through different epochs. Typically, as the number of epochs increases, the accuracy tends to improve while the loss decreases. This indicates that the model is learning and becoming more accurate in its predictions over time.

Fig 6 represents the evolution of accuracy and loss over epochs during the training of a neural network. The line chart shows the trend of accuracy and loss as training progresses. At the beginning (epoch 1), both accuracy and loss values are relatively poor, indicating that the model is performing poorly and has high uncertainty. As training progresses, the accuracy steadily increases, while the loss gradually decreases.

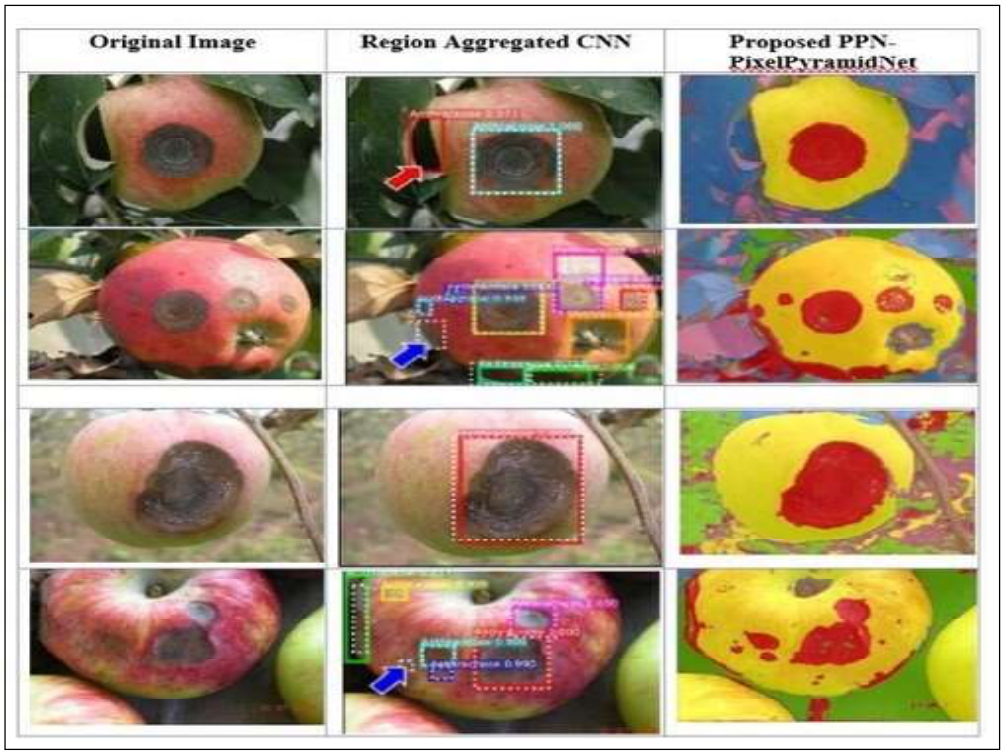

Fig 5: Comparison of proposed PPN-Pixel Pyramid Net with the Region Aggregated RCNN.

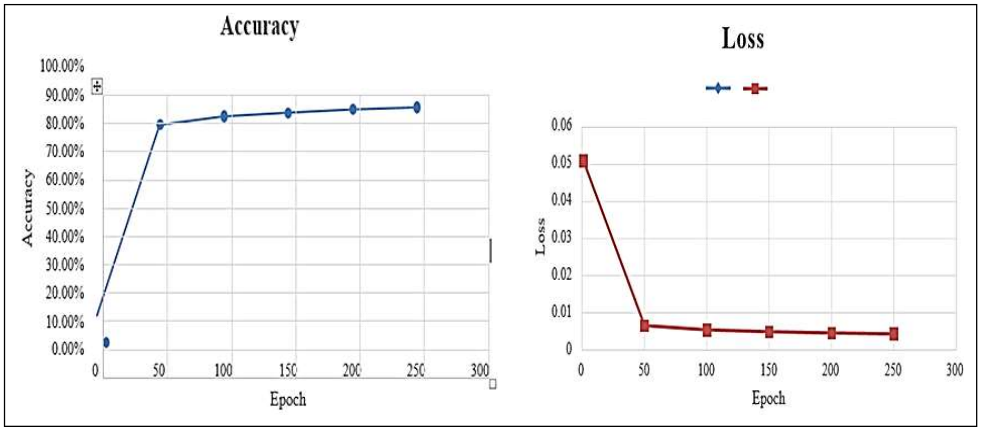

Fig 6: Chart representation of accuracy and loss over epochs.

Table 5 summarizes the precision, recall, interpolated precision at IOU and the area under the precision-recall curve (AUC) for each class, along with the computed average precision over intersection-over-union thresholds (AP). The formula for Average Precision (AP) is:

Where: $$AP = \frac{1}{N} \sum_{i=1}^{N} \text{AUC-PR}_i$$

N= Total number of classes.

$AUC\text{-}PR_i$= Area under the precision-recall curve for class *i*.

This formula gives the average of the AUC-PR values across all classes, providing a single scalar metric to evaluate the overall performance of the model across multiple classes.

In this comparison of disease detection models, the performance metrics of average precision (AP) shed light on their effectiveness in accurately identifying diseases from images. The proposed PPN-Pixel pyramid net emerges as the top performer with an impressive AP of 89.31% and an outstanding AP50 score of 89.81%. These scores indicate its exceptional ability to achieve high precision and recall rates, even at the stringent IOU threshold of 50%. Moreover, the Region-aggregated attention CNN also demonstrates notable performance with

an AP of 72.26% and a strong AP50 of 88.62%. These results highlight the efficacy of these models in disease detection tasks, offering promising solutions for accurate and efficient diagnosis and treatment planning.

The proposed PPN-Pixel Pyramid Net and the Region-aggregated attention CNN exhibit superior precision and recall compared to other models like Mask R-CNN, SSD, Retinanet and YOLOv3. These findings highlight the potential of advanced deep learning architectures in transforming disease detection, which could lead to substantial advancements in early diagnosis and intervention strategies. Table 6's comparative performance analysis shows that the PPN-Pixel Pyramid Net model outperforms all others in terms of Average Precision (AP) and AP50.

CONCLUSION

The novel approach for automating apple disease detection using proposed PPN-Pixel Pyramid Net, integrates spatial pyramid pooling techniques to effectively handle input images of varying sizes and aspect ratios, overcoming limitations of traditional convolutional neural networks (CNNs). Through experimentation and analysis, this work demonstrates the model's effectiveness in accurately classifying pixels and delineating boundaries between different regions of interest, such as diseased areas, healthy apple regions, leaves, stems and potential obstructions like clouds. Key performance metrics including accuracy, intersection over union and boundary F1 score provide insights into the model's capabilities across various classes, facilitating the evaluation and refinement of the apple disease detection system.

Conflict of interest

All the authors here by declare that there is no conflict of interest.

REFERENCES

Alharbi, A.G. and Arif, M. (2020). Detection and classification of apple diseases using convolutional neural networks. In 2020. 2nd International Conference on Computer and Information Sciences (ICCIS). Sakaka, Saudi Arabia. (pp. 1-6).

Chandra Vinod, S.S. (2023). Role of artificial intelligence in Indian agriculture: A review. Agricultural Reviews. 44(4): 558-562. doi: 10.18805/ag.R-2296.

Dubey, S.R. and Jalal, A.S. (2012). Detection and classification of apple fruit diseases using complete local binary patterns. In 2012 Third International Conference on Computer and Communication Technology. Allahabad, India. (pp. 346-351). https://doi.org/10.1109/ICCCT.2012.76.

Dubey, S.R. and Jalal, A.S. (2016). Apple disease classification using color, texture and shape features from images. SIViP. 10: 819-826. https://doi.org/10.1007/s11760-015-0821-1.

Han, C.H., Kim, E., Doan, T.N.N., Han, D., Yoo, S.J. and Kwak, J.T. (2021). Region-aggregated attention CNN for disease detection in fruit images. PLoS ONE. 16(10): e0258880. https://doi.org/10.1371/journal.pone.0258880

Kanagaraju, P., Aushiq, N.M. and Vanan, R.T. (2022). Disease detection and analysis in fruits using image processing. International Journal of Health Sciences. 6(S8): 1198-1211. https://doi.org/10.53730/ijhs.v6nS8.9879.

Mohana, Priya, C. *et al.* (2023). Customized semantic segmentation for enhanced disease detection of maize leaf images. International Journal on Recent and Innovation Trends in Computing and Communication. 11(11): 31-37. https://doi.org/10.17762/ijritcc.v11i11.9074.

Mohanapriya, R., Kalpana R., Aravinth Vijay, K. (2024). Quality enhancement of fruits and vegetables through organic Cultivation: A review. Agricultural Reviews. 45(2): 207-217. doi: 10.18805/ag.R-2351.

Saleem, M.H., Potgieter, J. and Arif, K.M. (2019). Plant disease detection and classification by deep learning. Plants. 8: 468. https://doi.org/10.3390/plants8110468.

Shivappa, M.M. and Gyanappa A.W. (2023). Machine learning models for plant disease prediction and detection: A review. Agricultural Science Digest. doi:10.18805/ag.D-5893.