International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(2s) 2026 pp. 1269-1276
https://doi.org/10.70917/ijcisim-2026-9998
Article

AN ANALYSIS OF YOLO VARIANTS FOR EFFECTIVE CLASSIFICATION AND DETECTION OF INDIVIDUAL PALM TREES IN UAV IMAGERY

Harjinder Singh[1]*, Ranjana Sharma[2], Anil Kumar[3]

[1] College of Computing Sciences & IT, Teerthanker Mahaveer University, Moradabad, India
Email: harjinder.mca07@gmail.com
[2] College of Computing Sciences & IT, Teerthanker Mahaveer University, Moradabad, India
Email: sharmaranjana04@gmail.com
[3] Photogrammetry & Remote Sensing Department (PRSD), Indian Institute of Remote Sensing (IIRS), ISRO, Dehradun, India
Email: aniliirsisro@gmail.com
Corresponding Author: Harjinder Singh (Email: harjinder.mca07@gmail.com)

Abstract: Advancements in modern information technology, particularly artificial intelligence, have reduced the need for human interaction while increasing productivity, accuracy, and dependability. Intelligent irrigation systems, along with customized fertilizer and pesticide applications, can greatly enhance the health and productivity of trees. For precise tree classification, sophisticated Deep Neural Network based techniques as You Only Look Once (YOLO) model and Convolutional Neural Networks (CNNs) can be utilized. These methods enable farmers to accurately identify and monitor a wide range of tree species, improving the management of growing conditions and health monitoring. Ultimately, this results in higher output and lower expenses. In this study, we have focused on Individual Palm Tree Identification. Various iterations of the YOLO method, including versions 5, 6, 7, and 8, have been thoroughly trained over a data-set of multiple variety of images of Palm trees and other trees taken from urban as well as plant estates through UAV to classify and detect individual tree canopies present in an scene captured by drone and further contrasted each YOLO algorithm version's performance. YOLO v8 version performed at its best with highest 85.7% mAP, 83.0%precision and 79.2% recall. For researchers and professionals in this field, this research enhances decision-making by providing valuable insights into which YOLO models are best suited for identifying palm trees in UAV datasets.

Keywords: YOLO, Object detection, UAV, Deep Learning, Palm trees..

1. INTRODUCTION

Globally, palm trees have a substantial economic impact, especially through the production of palm oil. Millions of people are employed in the palm oil industry, especially in Southeast Asia, Latin America, and Africa [1]. They represent both practical utility and cultural significance, making them essential parts of ecosystems and civilisations around the world [2].Tree monitoring via satellite and unmanned aerial vehicle surveillance is essential for comprehensive forest management and conservation efforts. These instruments make it possible to track reforestation initiatives, measure rates of deforestation, and continuously evaluate the condition of the woods. Deep learning algorithms have the potential to successfully execute the classification and detection of individual trees to help the aforementioned initiatives [3].

Deep learning has ushered in this modern age of precision and efficiency in object detection, fundamentally reshaping how machines perceive and interact with their visual environments. [4],[5],[6]. This capability is further enhanced by architectures like YOLO and Faster R-CNN, which streamline the detection process while maintaining durability in various data sets and real-world scenarios[7].

In This research paper, we have compared different versions of YOLO Algorithm on the basis of mAP, Precision and Recall. By providing superior speed, accuracy, simplicity, and training efficiency, YOLO distinguishes itself from more established deep learning models like R-CNN and SSD.YOLO enables real-time image processing at high frame rates (e.g., 45 FPS for YOLOv3 on a strong GPU), whereas the multi-stage process of R-CNN makes it significantly slower than YOLO.YOLO's end-to-end architecture is more straightforward and easier to implement than R-CNN's intricate, multi-stage methodology. Moreover, YOLO facilitates end-to-end training, optimizing the entire model at once, in contrast to R-CNN, which requires distinct training phases.

2. RELATED WORKS:

Zuxiang Situa et.al.(2023)[8] This study evaluated how effective a Transfer Learning (TL)-enhanced YOLO network is at identifying five different types of sewage faults compared to four object detection methods (ODMs). The study used 11 pretrained backbone CNNs. The results indicated that, in terms of detecting precision, computing efficiency, and Intersection over Union (IoU) scores, the TL-based YOLO models performed better than the other ODMs overall. Among the models examined, Resnet18 was found to be the most successful CNN architecture in this scenario, while Inceptionresnetv2 showed the least success. Youliang Chen et.al.(2023)[9] By improving the YOLO-v4 model, this study offers a better technique for identifying bayberry trees using UAV data. The model's performance is improved by combining DIoU NMS for accurate bounding box selection with Leaky_ReLU activation for quicker extraction. The recall rates for object detection are increased when K-Means clustering and DIoU NMS are combined. The ideal YOLO-v4 threshold for improved extraction is determined to be 0.25 after training with UAV photos. 98.16% recall and 97.78% detection accuracy are attained by the improved model. In a variety of settings, comparisons with YOLO-v4, YOLO-v4 tiny, YOLO-v3, and Faster R-CNN demonstrate superior detection and recall rates, up to 97.45%. ANASTASIIA SAFONOVA et.al.(2022)[10] In order to identify infested trees using UAV photos, Three You Only Look Once deep neural network architectures—YOLOv2, YOLOv3, and YOLOv4—were assessed and their execution outcomes contrasted. They created a brand-new dataset just for these models' testing and training in order to make our analysis easier. Furthermore, we applied a pre-processing method called Balance Contrast Enhancement Technique (BCET), which worked well to improve the models' capacity for generalization. Our tests showed that YOLOv4 produced noticeably better results, especially when combined with BCET pre-processing. With a mAP of 95%, YOLOv4 achieved the highest test performance out of all the YOLO models tested. Manuel Pérez-Carrasco et.al.(2022)[7] In this study, researchers tackled the challenging tasks of counting, identifying, and segmenting high-resolution RGB images captured by unmanned aerial vehicles (UAVs) using two neural network techniques: YOLO and Mask R-CNN. They compared different age and density forest stands as part of our evaluation. When it came to tree counting, Mask R-CNN displayed a lesser overestimation of 4.7%, whereas YOLO showed an average overestimation of 8.5%. In tree detection, Mask R-CNN outperformed YOLO by an average precision of 0.82% and recall of 0.80%, whereas YOLO averaged 0.72% and 0.68% in recall.

3. MATERIALS AND METHODS:

3.1: Study Area & Dataset:

The study area includes two separate but related environments: the bustling Prince Sultan University campus and the large palm tree plantation in the Saudi Arabian Kharj region[11]. In this study we have used a dataset which contains a total of 349 photos. 91 photos were taken from the Prince Sultan University campus, and 258 were captured from an aerial view of a palm tree plantation in the Kharj region of Saudi Arabia. These two sets of images show notable variations in aspect ratio and palm tree density. The DJI Phantom 4 Pro, which has a DJI FC6310 camera that can take pictures at 4864 × 3648 and 4096 × 2160 resolutions, and the DJI Mavic Pro, which has a DJI FC220 camera that can take pictures at 4000 × 3000 resolution, were the two drones used to shoot the images.

Figure 1.1 to 1.4 showcases the images picked from the dataset. This dataset has been extracted from Kaggle. A vast number of datasets, a collaborative environment, and a vibrant learning and networking community are all provided by Kaggle.

Figure 1.1 to 1.4: Images from Prince Sultan University campus, Saudi Arabia and tree plantation in the Kharj region of Saudi Arabia (Left to Right)[11]

3.3 Methodology Used:

Rather than imparting traditional baselines like Faster R-CNN and SSD, this study concentrates on the YOLO variants. Faster R-CNN mainly designed for higher accuracy in detection but involves lots of computational overhead, so making it less suitable for the real-time scenarios. SSD offer a balance between speed and accuracy. however, recent YOLO architectures keep on achieving superior performance as well as maintaining faster inference, so making them a more appropriate choice for this work. The YOLO method and its following versions (YOLOv2, YOLOv3, YOLOv4, YOLOv5, etc.) are extensively accepted for object detection jobs because of their excellent accuracy and speedy processing time[12]. Figure 2 is a comprehensive procedure for detecting individual palm trees using the YOLO algorithm.

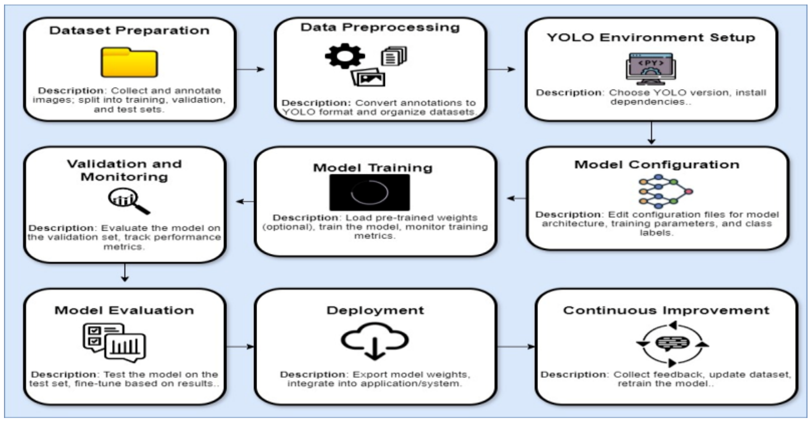


Figure 2: Operational Flow Diagram.

Figure 2 illustrates the procedure of creating and implementing an object detection model using the YOLO framework [13], [14], [15]. The process involves several crucial steps, beginning with preprocessing and dataset preparation. This includes gathering, labeling, and organizing relevant images. Next, the process moves on to setting up the environment, configuring the model, and training it. Throughout this process, careful validation and monitoring are emphasized. After the training phase, the model is rigorously evaluated to measure its performance. The following are the primary points that were crucial to these analytics.

Data Collection

A dataset consists of 349 images overall was used by us in which 258 images were shot from an overhead perspective of palm tree plantation in Saudi Arabian Kharj region, while 91 images were taken from the Prince Sultan University campus as palm trees are also available in the campus. Two drones DJI Phantom 4 Pro and DJI Mavic Pro were used by the dataset publisher. To solve the issue of imbalance between the palm and non-palm tree classes, data augmentation followed by balanced sampling strategies were applied at the time of training in order to reduce class bias and to improve learning across two of the categories.

Data Preprocessing

Data Annotation: The dataset consists of around 13,071 labeled objects; contain 11,150 palm trees and 1,921 other classes of trees. These labels were manually assigned using Roboflow's annotator tool. Two classes of objects have been created. One is Palm Tree and second is Tree. The quality of annotation was maintained using following clear as well as consistent labelling guidelines for all the object classes.

Data Augmentation: Enhancing the training dataset's robustness and variety by applying data augmentation is recommended to raising the performance of the model. For that we have used rotation Augmentation technique to generate images where radius lies between -15 degree to +15 degree. After then, the data was divided into training, testing, and validation sets at random. Training accounted for 80% of the data, testing for 10%, and validation for 10%.

Different YOLO models, performance parameters and Training: From version 5 to version 8, YOLO underwent significant evolution. Each release brought architectural enhancements to improve performance. YOLOv8, the latest version, enhances accuracy and performance with a redesigned backbone and improved algorithms[16], [17], [18],[23],[24],[25].

Experimental Setup

We utilized Google Colab along with Python 3 and Keras & Tensorflow, PyTorch library for YOLO v5, v6, v7, and YOLO v8 model's customized training on the mentioned dataset and the analysis of results. The evaluations were conducted utilizing a Google Colab Python environment equipped with a Tesla T4 GPU (15102 MiB), 2 CPUs, and 12.7 GB of RAM[19], [20], [21].

Hyper parameters used:

S.No.	Parameters used	Value
1	Batch Size	32
2	Input Image Size	640x640
3	Learning Rate	0.01
4	Momentum	0.937
5	IoU training threshold	0.20
6	Image rotation (+/- degrees)	-15 to 15 degree

Table 1: Hyper Parameters used

Evaluation metric:

EQUATION NO.	FORMULA
1	$\text{Precision} = \frac{TP}{TP + FP}$
2	$\text{Recall} = \frac{TP}{TP + FN}$
3	$\text{IoU} = \frac{\text{IntersectionAreaofPredictedBoundingBoxandGroundTruth}}{\text{UnionAreaofPredictedBoundingBoxandGroundTruth}}$
4	$\text{mAP} = 1/n \sum_{k=1}^{n} APk$
TN=TRUE NEGATIVE;TP= TRUE POSITIVE; FN=FALSE NEGATIVE; FP= FALSE POSITIVE;	

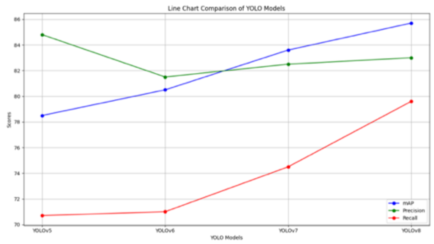
Table 2: Methods used for evaluation of performances of different models

Table 1 describes about the hyper parameters and its values used to train models. Table 2 describes the different evaluation methods used to assess the performance of each model used for the study. We assess model performance using precision, recall, IoU, and mean Average Precision (mAP@50-95). The Intersection over Union (IoU) measures the overlap between the ground truth and expected bounding boxes. In order to determine uncertainty and evaluate the consistency of the observed performance gains, bootstrapped evaluation of precision, recall, and mAP was carried out, considering that numerous independent runs were constrained by computational resources.

4. RESULTS AND DISCUSSION

Performance Evaluation for Different YOLO models

Each model is evaluated based on three primary measures: precision, mean average precision (mAP), and recall. YOLOv5 stands out with the highest precision of 84.8%, making it the most effective at reducing false positives. YOLOv8 demonstrates strong performance with 85.7% mAP and 79.6% recall, showcasing its capability for precise object detection and classification. YOLOv7, although slightly less precise than YOLOv5 and YOLOv8, still performs well with high mAP (83.6%) and recall (74.5%).With the lowest precision of all the models, YOLOv6 provides a balanced performance with competitive mAP and recall, indicating a trade-off between accuracy and precision. The merits and trade-offs of each model are depicted in this chart, making it easier to choose the optimal YOLO variant depending on your needs. Figures 3.1 and 3.2 show the performances of different YOLO Versions assessments in the form of bar diagrams and line graphs.



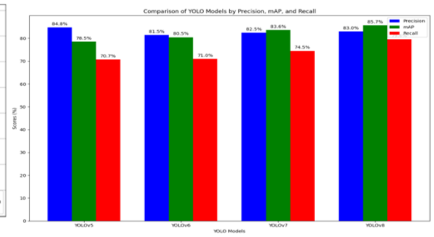


Figure 3.1, 3.2: Performance comparison using Line chart & Bar Chart

The performance of different YOLO variations on our dataset is displayed in Table 3. While inference time, FPS, and computational cost are approximate figures reported from common benchmarks for reference, precision, mAP, and recall are measured experimentally.

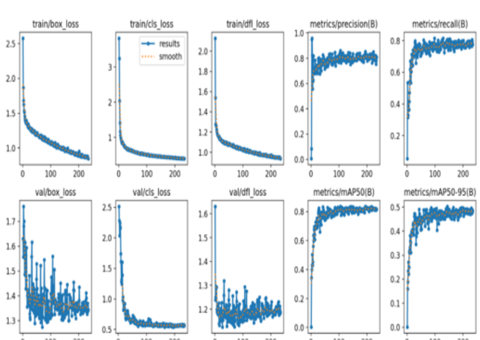
YOLO Variant	Precision (%)	mAP (%)	Recall (%)	Inference Time(s/image)	FPS	Parameters (M)	FLOPs (G)
YOLOv5	80.8	79.5	69.5	0.015	67	31.0	90
YOLOv6	81.5	80.5	70.7	0.018	56	26.0	75
YOLOv7	82.5	82.5	74.5	0.020	50	37.0	92
YOLOv8	85.7	83.0	79.2	0.012	83	11.2	17.3

Table 3: Performance comparison Table

Training Graphs for YOLO v8:

Performance assessment of different models have show the solid performance of YOLO v8.The plots have been drafted to show different loss functions and evaluation metrics at the time of training iterations. Training measures such as detection loss (dfl_loss), class loss (cls_loss), and bounding box loss (box_loss) are showcased in the top row. These metrics all show the continuous improvements in the model's operendy over time. The precision and recall metrics for the class B also improved, reaching high values (about 0.8) as training progressed. The bottom row shows the corresponding validation metrics, which are more volatile but generally follow the training trend. The validation losses (box_loss, cls_loss, dfl_loss) stabilize after initial drops, indicating about the model that it is learning without over fitting. By the end of training, the mAP50 and mAP50-95 metrics—which serve as benchmarks for assessing object detection performance—have steadily improved, reaching approximately 0.8 and 0.5, respectively. In order to reduce overfitting, early stopping was employed to cease training when validation performance stopped improving. Stabilising convergence and enhancing generalisation were aided by a learning rate schedule with weight decay.

These plots collectively suggest that the YOLO v8 model learns effectively on both the training and validation datasets, with good generalization and increased accuracy on object detection tasks. Figure 4.1collectively shows all the possible training graphs explained above. Figure 4.2 shows The confusion matrix derived from the reported precision and recall values of YOLOv8 to reflect overall classification behaviour in the absence of raw prediction logs.



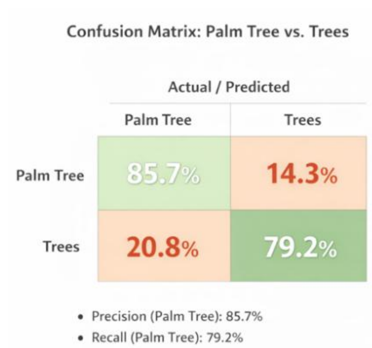


Figure 4.1, 4.2: Training Graphs and confusion matrix for YOLO v8

Model Output:

In Figure 5, The model can classify and detect individual Palm trees and other trees very effectively with an impressive confidence score.

Figure 5.1: Input Image

Figure 5.2: Output Image

Figure 5.3: Input Image

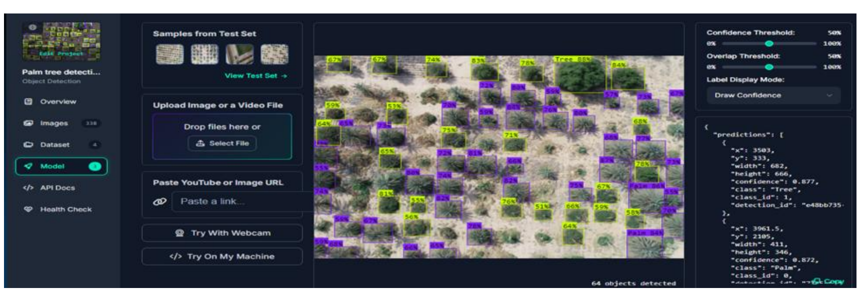
Figure 5.4: Output Image

Figure 5.1,5.2,5.3, 5.4 Input images v/s Output Images

Model Deployment:

The most effective model for identifying individual palm trees was YOLOv8's model, which was implemented on the Roboflow's cloud platform, trained, and validated over 100 epochs as displayed as figure 6. Experiments show that the model can analyse data in almost real-time on UAV-compatible platforms because to its high computational efficiency and inference speed. Future work to better evaluate real-world performance will involve actual UAV field validation.

Figure 6: Screenshot of deployment of YOLO v8 model



5. CONCLUSION

The research concentrated on evaluating YOLOv5, YOLOv7, and YOLOv8 models for Individual palm tree detection and classification in UAV images. We trained YOLO v5,v6,v7 and v8 on a dataset which consists of 349 drone shots to check the performance of the mentioned models. The YOLO v8 model learns effectively on both the training and validation datasets, with good generalization and increased accuracy on object detection tasks. YOLOv5 stands out with the highest precision of 84.8%, making it the most effective at reducing false positives. YOLOv8 demonstrates strong performance with 85.7% mAP and 79.6% recall, showcasing its capability for precise object detection and classification. YOLOv7, although slightly less precise than YOLOv5 and YOLOv8, still performs well with high mAP (83.6%) and recall (74.5%).With the lowest precision of all the models, YOLOv6 provides a balanced performance with competitive mAP and recall, indicating a trade-off between accuracy and precision. When it comes to speed, accuracy, simplicity, and training efficiency, We recognise that generalisation may be limited by the tiny dataset. Larger datasets will be used in future research to increase robustness.

6. FUTURE SCOPE

In future research, we will explore advanced strategies such as unsupervised synthetic data creation methods like Generative Adversarial Networks (GANs) to enhance the dataset and advanced object detection algorithms could potentially improve visualization systems and spatial analysis in any tree detection challenge[22]. We will use datasets with bigger size to improve robustness.

Author Contributions: The conceptual framework, data collection, in-depth analysis, inquiry, technique design, validation, visualisation, and preliminary article draughting were all contributed to by Harjinder Singh. Dr. Ranjana Sharma and Dr. Anil Kumar oversaw the project, managed resources, gave advice on how to conceptualise the study, and edited and critically assessed the article. Additionally, Both were in charge of the manuscript's general supervision and final review.

Data Availability Statement

The dataset utilized in this research can be found in Kaggle [11]. The corresponding author can also provide the data upon request.

References:

1. Y. Prabowo et al., “Palm Trees Counting Using MobileNet Convolutional Neural Network in Very High-Resolution Satellite Images,” in 2022 IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology (AGERS), 2022, pp. 79–83. doi: 10.1109/AGERS56232.2022.10093287.
2. “Agriculture | Free Full-Text | Oil Palm Tree Detection and Health Classification on High-Resolution Imagery Using Deep Learning.” Accessed: Feb. 19, 2024. [Online]. Available: https://www.mdpi.com/2077-0472/11/2/183
3. I. Nurhabib, K. Seminar, and others, “Recognition and counting of oil palm tree with deep learning using satellite image,” in IOP Conference Series: Earth and Environmental Science, IOP Publishing, 2022, p. 012058.
4. M. Beloiu, L. Heinzmann, N. Rehush, A. Gessler, and V. C. Griess, “Individual tree-crown detection and species identification in heterogeneous forests using aerial RGB imagery and Deep Learning,” Remote Sens., vol. 15, no. 5, p. 1463, 2023.
5. Y. Huang et al., “Tree Species Classification from UAV Canopy Images with Deep Learning Models,” Remote Sens., vol. 16, no. 20, p. 3836, Oct. 2024, doi: 10.3390/rs16203836.
6. U. Sirisha, S. P. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, “Statistical analysis of design aspects of various YOLO-based deep learning models for object detection,” Int. J. Comput. Intell. Syst., vol. 16, no. 1, p. 126, 2023.
7. M. Pérez-Carrasco, B. Karelovic, R. Molina, R. Saavedra, P. Cerulo, and G. Cabrera-Vives, “Precision silviculture: Use of UAVs and comparison of deep learning models for the identification and segmentation of tree crowns in pine crops,” Int. J. Digit. Earth, vol. 15, no. 1, pp. 2223–2238, 2022.
8. Z. Situ, S. Teng, X. Liao, G. Chen, and Q. Zhou, “Real-time sewer defect detection based on YOLO network, transfer learning, and channel pruning algorithm,” J. Civ. Struct. Health Monit., vol. 14, no. 1, pp. 41–57, 2024.
9. Y. Chen, H. Xu, X. Zhang, P. Gao, Z. Xu, and X. Huang, “An object detection method for bayberry trees based on an improved YOLO algorithm,” Int. J. Digit. Earth, vol. 16, no. 1, pp. 781–805, 2023.
10. A. Safonova, Y. Hamad, A. Alekhina, and D. Kaplun, “Detection of Norway spruce trees (Picea abies) infested by bark beetle in UAV images using YOLOs architectures,” IEEE Access, vol. 10, pp. 10384–10392, 2022.
11. A. Ammar, A. Koubaa, and B. Benjdira, “Deep-learning-based automated palm tree counting and geolocation in large farms from aerial geotagged images,” Agronomy, vol. 11, no. 8, p. 1458, 2021.
12. P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, “A Review of Yolo algorithm developments,” Procedia Comput. Sci., vol. 199, pp. 1066–1073, 2022.
13. C. Zhao, X. Shu, X. Yan, X. Zuo, and F. Zhu, “RDD-YOLO: A modified YOLO for detection of steel surface defects,” Measurement, vol. 214, p. 112776, 2023.
14. Z. Zhang, Y. Tanimoto, M. Iwata, and S. Yoshida, “Data generation using Pix2Pix to improve YOLO v8 performance in UAV-based Yuzu detection,” Smart Agric. Technol., vol. 10, p. 100777, Mar. 2025, doi: 10.1016/j.atech.2025.100777.
15. W. Xinming, T. S. Hong, and others, “Comparative study on Leaf disease identification using Yolo v4 and Yolo v7 algorithm,” AgBioForum, vol. 25, no. 1, 2023.
16. S. N. Appe, G. Arulselvi, and G. Balaji, “CAM-YOLO: tomato detection and classification based on improved YOLOv5 using combining attention mechanism,” PeerJ Comput. Sci., vol. 9, p. e1463, 2023.

17. W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, “YOLO-face: a real-time face detector,” Vis. Comput., vol. 37, pp. 805–813, 2021.
18. Q. Ding et al., “Cf-yolo: Cross fusion yolo for object detection in adverse weather with a high-quality real snow dataset,” IEEE Trans. Intell. Transp. Syst., vol. 24, no. 10, pp. 10749–10759, 2023.
19. C. Wei and W. Wang, “RFAG-YOLO: A Receptive Field Attention-Guided YOLO Network for Small-Object Detection in UAV Images,” Sensors, vol. 25, no. 7, p. 2193, Mar. 2025, doi: 10.3390/s25072193.
20. Y. Zhang, M. Ye, G. Zhu, Y. Liu, P. Guo, and J. Yan, “FFCA-YOLO for small object detection in remote sensing images,” IEEE Trans. Geosci. Remote Sens., 2024.
21. Z. Liu, Y. Gao, Q. Du, M. Chen, and W. Lv, “YOLO-extract: Improved YOLOv5 for aircraft object detection in remote sensing images,” IEEE Access, vol. 11, pp. 1742–1751, 2023.
22. N. Vishnoi and R. K. Dwivedi, "Comprehensive Survey on the Reliability of Wireless Sensor Networks," 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2022, pp. 496-502, doi: 10.1109/SMART55829.2022.10047428.
23. N. Vishnoi and R. K. Dwivedi, "Designing dynamic adjustment approach of event reliability in wireless sensor network", Journal of Discrete Mathematical Sciences and Cryptography,2023, pp. 2045–2060, DOI: 10.47974/JDMSC-1850
24. Gupta, Mohan Vishal, Dwivedi, Rakesh Kumar & Kumar, Anil(2023) Enhancing remote sensing satellite image classification through parameter optimization in hybrid CNN-LSTM model, Journal of Discrete Mathematical Sciences and Cryptography, 26:7, 2061–2079, DOI: 10.47974/JDMSC-1855
25. Gupta, Mohan Vishal, Dwivedi, Rakesh Kumar & Kumar, Anil(2023) Impact analysis of convolutional neural network in classification of satellite imagery, Journal of Information and Optimization Sciences, 44:6, 1151–1166, DOI: 10.47974/JIOS-1453