iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VIII **Month of publication:** August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73497

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

AI Powered Drone Surveillance System for Real Time Object Detection in Unmanned Aerial Vehicles Using Deep Learning Approaches

Pulagam Naveen Kumar[1], Rampranay Ponnamaneni[2], Thalla Tirumaleswarlu[3], B. Rajesh[4]
[1,3]Assistant Professor, Department of Computer Science and Engineering,
[2,4]Assistant Professor, Department of Electrical and Electronics Engineering
Scient Institute of Technology, Ibrahimpatnam, R.R Dist [India]

Abstract: Deep learning algorithms, particularly CNNs, have significantly improved object detection accuracy in remote sensing applications. Unlike most UAV-based approaches, this project implements a static camera system utilizing CNN and YOLOv8 (a state-of-the-art one-stage detector) for real-time aerial image processing. The system is optimized for surveillance, environmental monitoring, and disaster response applications. While our current implementation uses fixed cameras, the architecture supports seamless UAV integration. This research examines the speed accuracy trade-offs between one-stage and two-stage detectors, demonstrating YOLOv8's ability to maintain both rapid inference and reliable detection performance. Experimental results validate the system's effectiveness in static deployments while its adaptability for dynamic drone-based scenarios.
Keywords: Object detection, multi-target detection and tracking, drone detection, deep learning, visual detection

I. INTRODUCTION

Drones, or Unmanned Aerial Vehicles (UAVs), have revolutionized various industries by providing access to remote, hazardous, or otherwise inaccessible areas. Equipped with advanced cameras, UAVs can capture high-resolution images and videos from multiple altitudes and perspectives, enabling critical applications such as aerial photography, search and rescue operations, environmental monitoring, and defense surveillance. However, real-time manual tracking and analysis of drone-captured imagery are impractical, necessitating automated solutions. Machine learning and deep learning algorithms have emerged as powerful tools for processing and interpreting visual data from UAVs, enabling tasks such as object detection, anomaly recognition, and real-time decision-making. While these AI-driven systems enhance post-capture analysis, they do not directly control the image acquisition process, which remains crucial for surveillance, inspection, mapping, and surveying applications. In a typical drone surveillance system, an onboard camera—such as an RGB, thermal, or multispectral sensor—captures aerial footage during flight. This data is either transmitted wirelessly to a ground station in real time or stored locally for later retrieval. Camera parameters, including resolution, focus, and exposure, can be dynamically adjusted to optimize image quality based on mission requirements.

However, the rapid development and widespread application of drones also bring various hidden dangers, such as public safety, personal safety, personal privacy, and so on. The occurrences of accidents caused by the illegal of drones become more frequent. Therefore it is necessary to regulate the drones. A few companies like DJI, set up zones to ensure the safety in some sensitive areas, such as airports, prisons. But the effect of no- y zones is very limited. It is therefore significant to implement real-time drone detection to give warning accurately in time.

Many existing techniques, e.g., radar, radio frequency, acoustic [2] [3] [4] and optical sensing techniques can be used for drone detection [5] [6]. Because of high detection accuracy and long effective range, video-based detection has attracted many research interests [7] [8] and has great potential for drone detection. In our work, we employ video-based detection based on deep learning approach which is a powerful tool in the computer vision area. Drone detection is essentially an object detection problem. In early years, video-based object detection is by extracting discriminant features such as Local Binary Pattern (LBP) [9], Scale Invariant Feature Transform (SIFT) [1], Histogram of Oriented Gradient (HOG) [1] and Speeded Up Robust Features (SURF) [2] then using these features to train the detector. In 2012, Krizhevsky et al. [3] showed the amazing power of the convolutional neural network (CNN) in the ImageNet grand challenge.

Since then, the developments and applications of deep learning methods increase rapidly. There are several variants in CNNs such as the R-CNN [9], SPPNet [8] and Faster-RCNN [6]. Since these networks can generate highly discriminant features, their performances are far beyond the traditional object detection techniques

This work focuses on developing an AI-powered drone surveillance system that integrates deep learning techniques to automate image analysis and enhance situational awareness. By employing convolutional neural networks (CNNs) and real-time object detection models like YOLO (You Only Look Once), the system will process live stream feeds as well as pre-recorded videos to identify and track objects of interest. Additionally, the project explores efficient data transmission methods and adaptive camera control to improve performance in static as well as dynamic environments. The proposed work aims to bridge the gaps between drone-based data acquisition and intelligent analytics, offering scalable solutions for security, disaster management, and infrastructure inspection. Through this project, we seek to demonstrate the potential of deep learning in transforming UAV surveillance into a more efficient, accurate, and autonomous process.

II. LITERATURE SURVEY

Recent research in drone detection has increasingly focused on deep learning approaches due to their effectiveness in real-time object recognition. Traditional models such as SSD, Faster R-CNN, and Mask R-CNN have been used with moderate success, but often face limitations in detecting small objects like drones or maintaining high detection speed[1]. Among the newer methods, YOLO-based models have gained popularity for balancing accuracy and speed. The model is trained on a diverse dataset of drone images and fine-tuned it for better accuracy. The results showed that YOLOv5 performs well in real-time detection scenarios, even under low-light conditions and complex environments.

The model pruning can optimize performance[2] for resource-constrained systems while maintaining accuracy, enabling practical applications in surveillance, search-and rescue, and delivery missions where traditional tracking methods fail. Through hardware experiments with a DJI Tello drone to make it cost-effective solution for real-world UAV deployment. Real-time object detection and tracking system for UAVs can be developed by using a pruned YOLOv4 model and Siam Mask tracker within a ROS framework. This can overcome the challenge of autonomous drone operation in GPS-denied environments by combining lightweight deep learning with PID-based flight control.

There are four primary drone detection technologies[9]—radar, RF, acoustic, and vision-based systems—along with emerging sensor fusion approaches. Radar excels in long-range detection but struggles with small drones; RF methods effectively intercept communication signals but require known frequencies; acoustic systems are cost-effective yet limited by ambient noise; while vision-based techniques offer high accuracy but depend on lighting conditions. But machine learning, particularly deep learning algorithms (CNNs, RNNs, YOLO variants), significantly enhances detection accuracy across all modalities the sensor fusion combining multiple technologies (e.g., radar+vision or RF+acoustic) achieves superior performance (90-99% accuracy) by compensating for individual limitations.

Hybrid deep learning model (AVUHBO)[2] for object detection in drone imagery, combines improved RCNN with optimized Bi-GRU/LSTM networks to handle challenges like changing camera angles and lighting. The system works in three steps: prepares images, extracts key features (using LGBP, ResNet, and SIFT), then recognizes objects. AVUHBO helps the system learn better, achieving maximum accuracy - better than other methods. This helps search-and-rescue missions where drone images are tricky to analyse. Tests show it outperforms existing techniques by combining traditional and deep learning features smartly.

III. METHODOLOGY

In this section, we present a deep convolution neural network model, YOLOv8 [9], which is the state-of-the-art on standard detection tasks. Then we propose a new model for real-time drone detection by modifying the structure and tuning parameters of YOLOv8, which makes the model better adapt to the drone detection.Deep learning models have transformed the field of object detection, enabling machines to automatically identify and locate objects within images and videos with remarkable accuracy and efficiency. These models leverage advanced neural network architectures, to learn complex features from visual data.

A. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are specialized deep learning models designed to process grid-like data, such as images. In object detection, CNNs analyze images through layers of convolutional filters to extract features, followed by pooling layers to reduce dimensionality, ultimately predicting the presence and location of objects within the image.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

1) Working of CNNs in Object Detection

- Input Image: The process begins with an input image that is fed into the CNN.
- Feature Extraction: The CNN applies multiple convolutional and pooling layers to extract hierarchical features from the image. Early layers capture low-level features (edges, textures), while deeper layers capture high-level features (shapes, objects).
- Classification and Localization: In object detection tasks, the CNN not only classifies the objects present in the image but also predicts their locations using bounding boxes. This is often achieved through additional layers that output both class scores and bounding box coordinates.
- Training: CNNs are trained using labeled datasets, where the model learns to minimize the difference between predicted and actual outputs through backpropagation and optimization techniques.

B. YOLO (You Only Look Once)

YOLO (You Only Look Once) is a real-time object detection algorithm that processes images in a single pass, making it both fast and efficient. It divides the image into a grid and predicts bounding boxes and class probabilities for each grid cell, allowing for simultaneous detection and classification of multiple objects.

1) Working of YOLO in Object Detection

a) Image Input:

• The input image is resized to a fixed size (e.g., 448 x 448) for consistent processing.

b) Feature Extraction

The CNN extracts features from the image using convolutional layers, which helps in predicting bounding boxes and class probabilities.

c) Prediction

• Each grid cell predicts multiple bounding boxes and their associated confidence scores.
• The model combines these predictions to generate a final list of detected objects.

d) Training

• YOLO is trained on labeled datasets, where it learns to minimize the difference between predicted and actual bounding boxes and class probabilities through backpropagation.

IV. ARCHITECTURE OF THE MODEL

Let us assume if we are working with images or videos, and your task is to identify objects within them accurately. This is where YOLOv8, the latest iteration of the YOLO (You Only Look Once) series, comes into play. YOLOv8 represents a significant milestone in computer vision, designed to enhance the efficiency and precision of object detection. Its underlying principle simplifies the process by processing an entire image through a neural network known as a convolutional neural network (CNN) in a single pass, eliminating the need for laboriously inspecting each part of the image individually. This streamlined approach, often referred to as “You Only Look Once,” leads to impressive gains in both speed and efficiency.

The architecture of the model is shown in Fig 1:

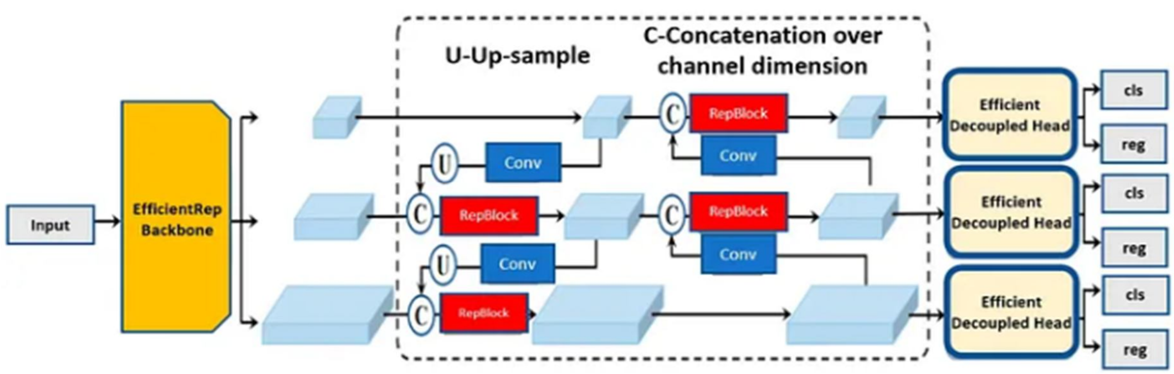


Figure 1 : Architecture of the YOLOV8

At its core, YOLOv8 follows a straightforward set of principles. It all starts with an input image, which can be of varying sizes, making it a versatile tool for a wide range of applications. The image then undergoes a process of feature extraction within a neural network. This network, carefully engineered and trained, excels at identifying and highlighting the most relevant features within the image, akin to a detective piecing together crucial clues at a crime scene. YOLOv8 stands out by not just detecting objects at a single scale but by simultaneously spotting them at different scales, ensuring it can identify objects both large and small effectively.

V. EXPERIMENTAL RESULTS

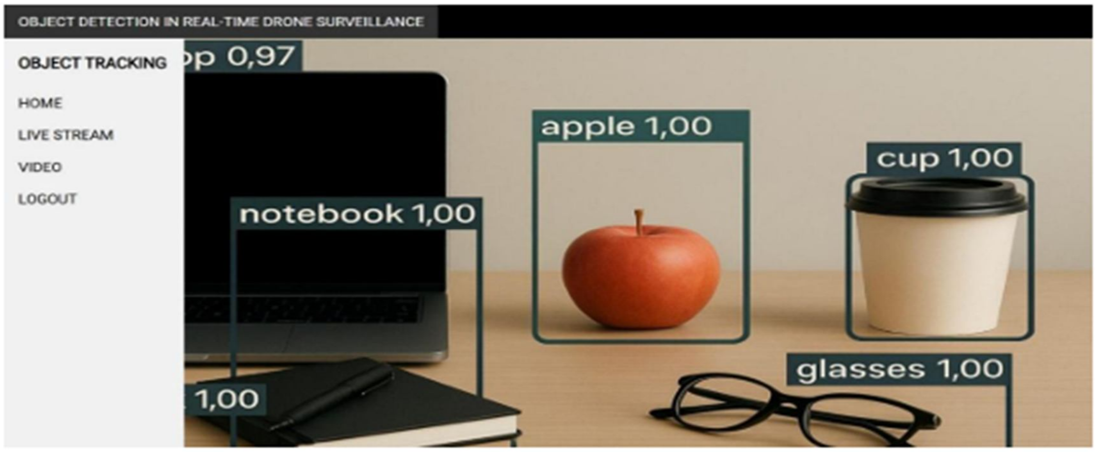


Fig 2 : Object detection from UAV

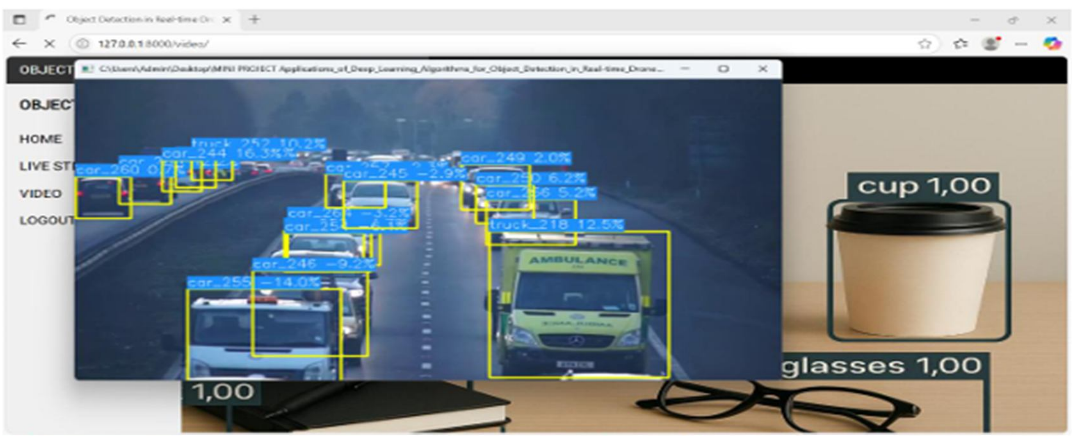


Fig 3 : Object detection from videos

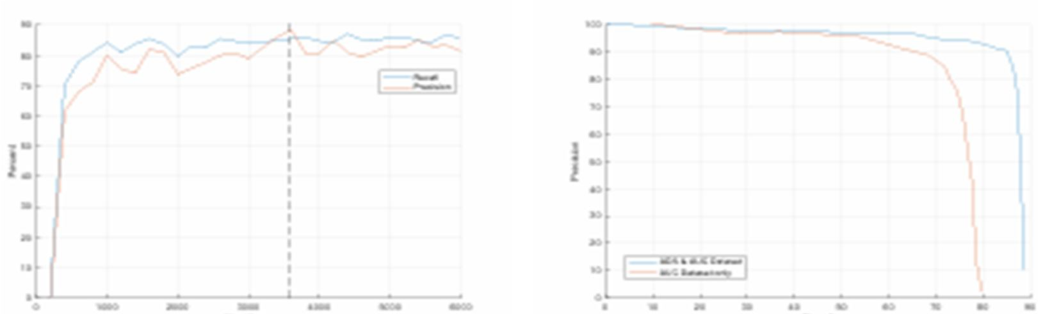


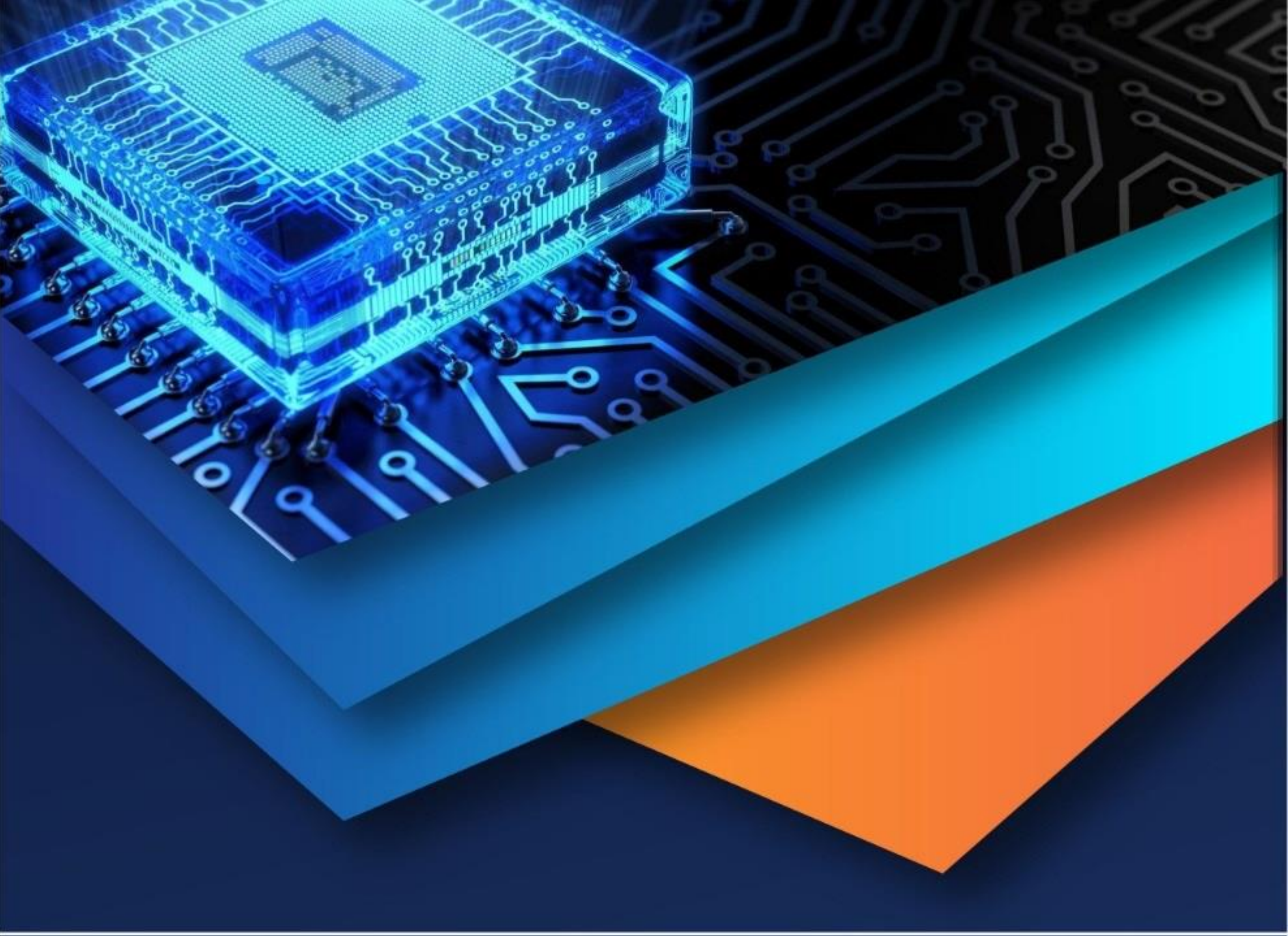
Fig 4 : The left figure shows the change of precision and recall as iterations increase. The right figure shows comparison of the drone detector trained byonly USC dataset and combined with KCF labelled dataset.

VI. CONCLUSION

In this work we developed a real time AI powered drone surveillance system that integrates deep learning techniques for object detection in Unmanned Aerial Vehicles(UAV). The proposed YOLOV8 model processed process live stream feeds as well as pre-recorded videos to identify and track objects of interest. The proposed model uses drone-based data acquisition and intelligent analytics, offering scalable solutions for security, disaster management, and infrastructure inspection Deep learning (DL) is often considered a "black-box" solution for many problems, although ongoing research aims to address this perception. In the field of remote sensing, DL has already made significant contributions across various applications. Future directions focuses on usage of Transfer learning techniques involves leveraging pre-trained models on large-scale datasets to transfer learned knowledge from a source task to a target task. This approach conserves computational resources, accelerates training, and enhances model performance.

REFERENCES

[1] Aytar, Y. (2014). Transfer learning for object category detection. PhD Thesis, 146. https://www.robots.ox.ac.uk/~vgg/publications/2014/Aytar14a/aytar14a.pdf
[2] Chamarty, A. (2020). Fine-Tuning of Learning Rate for Improvement of Object Detection Accuracy. Proceedings - 2020 IEEE India Council International Subsections Conference, INDISCON 2020, 135–141. https://doi.org/10.1109/INDISCON50162.2020.000 38
[3] Cui, X., Zhang, W., Tüske, Z., & Picheny, M. (2018). Evolutionary stochastic gradient descent for optimization of deep neural networks. Advances in Neural Information Processing Systems, 2018- Decem (NeurIPS), 6048–6058.
[4] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 580–587. https://doi.org/10.1109/CVPR.2014.81
[5] Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. 32nd International Conference on Machine Learning, ICML 2015, 1, 448–456.
[6] Koo, K. M., & Cha, E. Y. (2017). Image recognition performance enhancements using image normalization. Human-Centric Computing and Information Sciences, 7(1). https://doi.org/10.1186/s13673-017-0114-5
[7] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2020). Focal Loss for Dense Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(2), 318–327. https://doi.org/10.1109/TPAMI.2018.2858826
[8] Liu, W., Anguelov, D., Erhan, D., & Fu, C.-Y. (2016). SSD:Single Shot MultiBox Detector. European Conference on Computer Vision, 1(January 2017), 398–413. https://doi.org/10.1007/978-3-319-46448-0_2
[9] MuhammadHussain, YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. https://doi.org/10.3390/machines11070677.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY