iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: XI Month of publication: November 2025

DOI: https://doi.org/10.22214/ijraset.2025.75080

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 XI Nov 2025- Available at www.ijraset.com

An Object Detection Model Using MobileNetV3-SSD

Rajnandini Chaudhari[1], Nilesh Vani[2]
[1]PG Student, GF's Godavari College of Engineering, Jalgaon, India, 425002
[2]Associate Professor, GF's Godavari College of Engineering, Jalgaon, India, 425002

Abstract: The rapid evolution of deep learning techniques has greatly enhanced the capabilities of object detection systems. This paper investigates the integration of MobileNet with the Single Shot MultiBox Detector (SSD) for efficient and accurate object detection. MobileNet, designed with depthwise separable convolutions, significantly reduces computational complexity while maintaining high performance. By combining MobileNet's lightweight architecture with the SSD framework, which performs object localization and classification in a single pass, we achieve an effective balance between speed and accuracy. Our study benchmarks MobileNetV3 SSD against established object detection models on various datasets, highlighting its strengths in real-time applications. We demonstrate that MobileNetV3 SSD offers notable improvements in processing time and resource utilization without compromising detection quality. The findings underline MobileNetV3 SSD's suitability for deployment in environments with limited computational power, such as mobile devices and edge computing platforms.
Keywords: Computer Vision, Object Detection, MobileNetV3, Single Shot Multi-Box Detector, OpenCV.

I. INTRODUCTION

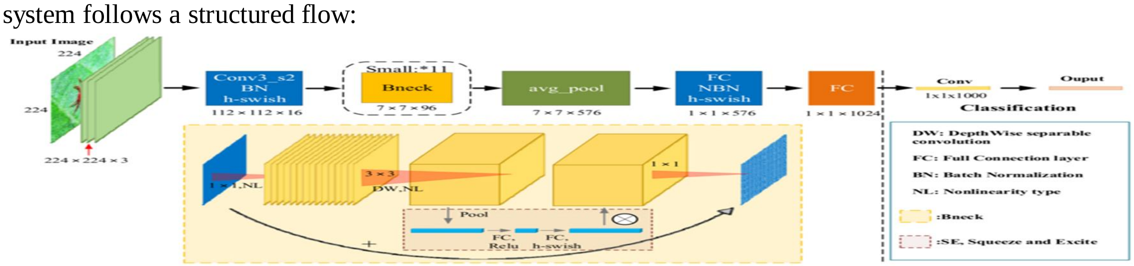
Object detection remains a pivotal challenge in computer vision, encompassing the dual tasks of identifying the presence of objects within an image and accurately localizing them with bounding boxes. This task is critical across various domains, including autonomous vehicles, surveillance systems, and interactive applications. Recent advancements in deep learning have significantly enhanced the performance of object detection systems, but deploying these models in real-time, resource-constrained environments still presents a challenge. The Single Shot MultiBox Detector (SSD) is a state-of-the-art framework that addresses these challenges by providing a unified approach to object detection. To further enhance the efficiency of SSD, particularly for deployment on mobile and embedded devices, we integrate it with MobileNetV3 — a lightweight deep learning architecture optimized through neural architecture search and squeeze-and-excitation modules.

II. RELATED WORK

Object detection has evolved from classical computer vision approaches such as Haar cascades and HOG features to deep learning-based frameworks. R-CNN and its successors like Fast R-CNN and Faster R-CNN laid the foundation for accurate region-based detection. SSD (Single Shot MultiBox Detector) revolutionized real-time detection by predicting object classes and bounding boxes in a single forward pass. MobileNetV3 further enhanced computational efficiency through architecture search and lightweight convolutional blocks. Their integration enables robust, real-time object detection suitable for mobile and embedded devices.

III. PROPOSED METHODOLOGY

The proposed object detection system leverages MobileNetV3 as the backbone network integrated with the Single Shot MultiBox Detector (SSD) for efficient and real-time object detection. As illustrated in the proposed architecture (Fig–Proposed Methodology Design), the system follows a structured flow:



Fig–Proposed Methodology Design

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 XI Nov 2025- Available at www.ijraset.com

1) Input Image/Video Frame: The input is captured from a camera or loaded from an image dataset. Each frame is resized and normalized for optimal model compatibility.
2) Feature Extraction using MobileNetV3 Backbone: MobileNetV3, optimized using Neural Architecture Search (NAS) and Squeeze-and-Excitation (SE) modules, extracts deep spatial and semantic features. Its lightweight design ensures low computational overhead while maintaining high accuracy.
3) Feature Map Generation and Multi-scale Detection (SSD): The extracted features are passed to the SSD detection head, which performs multi-scale feature mapping. This enables detection of objects of various sizes within the same frame. SSD predicts both bounding boxes and class probabilities in a single forward pass, improving inference speed.
4) Bounding Box Regression and Non-Maximum Suppression (NMS): The SSD outputs multiple bounding boxes per object. NMS is applied to remove redundant detections, retaining only the boxes with the highest confidence scores.
5) Output Visualization: The final detected objects are visualized with bounding boxes and labels. The results demonstrate effective detection under different conditions such as low light and varied object scales.
6) Optimization and Deployment: The model is optimized using quantization and pruning techniques to minimize memory usage. Deployment is done using TensorFlow Lite, enabling real-time inference on mobile and edge devices without significant accuracy degradation.

.

IV. RESULTS

The MobileNetV3-SSD model demonstrated effective object detection performance in both static and dynamic scenarios. The model was able to detect various objects such as cars, bicycles, and bottles with high confidence levels even under reduced lighting and resolution conditions.

Fig -1: Objects (Car) detected

Fig -2: Objects (Aeroplane) detected

Fig -3: Objects (Bicycle) detected

Fig -4: Objects (Horse) detected

Fig -5: Objects (Bottle) detected

Fig -6: Objects (Bird) detected

V. CONCLUSION

In this study, we presented an object detection approach based on the MobileNetV3-SSD architecture. The model achieved an optimal balance between speed and accuracy, making it highly suitable for deployment in real-time applications. Experimental results confirmed that MobileNetV3-SSD delivers robust detection performance even on devices with limited computational capacity. Future work may involve integrating attention mechanisms and feature pyramid networks to further improve detection accuracy for small and overlapping objects.

VI. FUTURE SCOPE

1) Integration of Advanced Attention Mechanisms: Incorporating attention modules (e.g., CBAM or Transformer-based blocks) can help the model focus more on relevant features, improving detection accuracy, especially for small or overlapping objects.
2) Use of Feature Pyramid Networks (FPN): Combining SSD with an FPN structure can enhance multi-scale feature representation and improve object detection in complex scenes.
3) Edge and IoT Optimization: Extending optimization techniques such as mixed-precision quantization and model distillation can make the model even more suitable for IoT and embedded devices with minimal power consumption.
4) Dataset Expansion and Real-World Training: Training the model on more diverse and real-world datasets will improve its adaptability to various environments, including autonomous navigation, smart surveillance, and agriculture monitoring
5) Integration with Cloud and Edge Collaboration: A hybrid approach can be developed where edge devices perform preliminary detection while cloud servers handle fine-tuned post-processing, enabling large-scale distributed object detection
6) Real-Time Video Analytics and Tracking: Future development can incorporate object tracking algorithms (e.g., SORT or Deep SORT) to extend the model for real-time surveillance and intelligent traffic systems.

REFERENCES

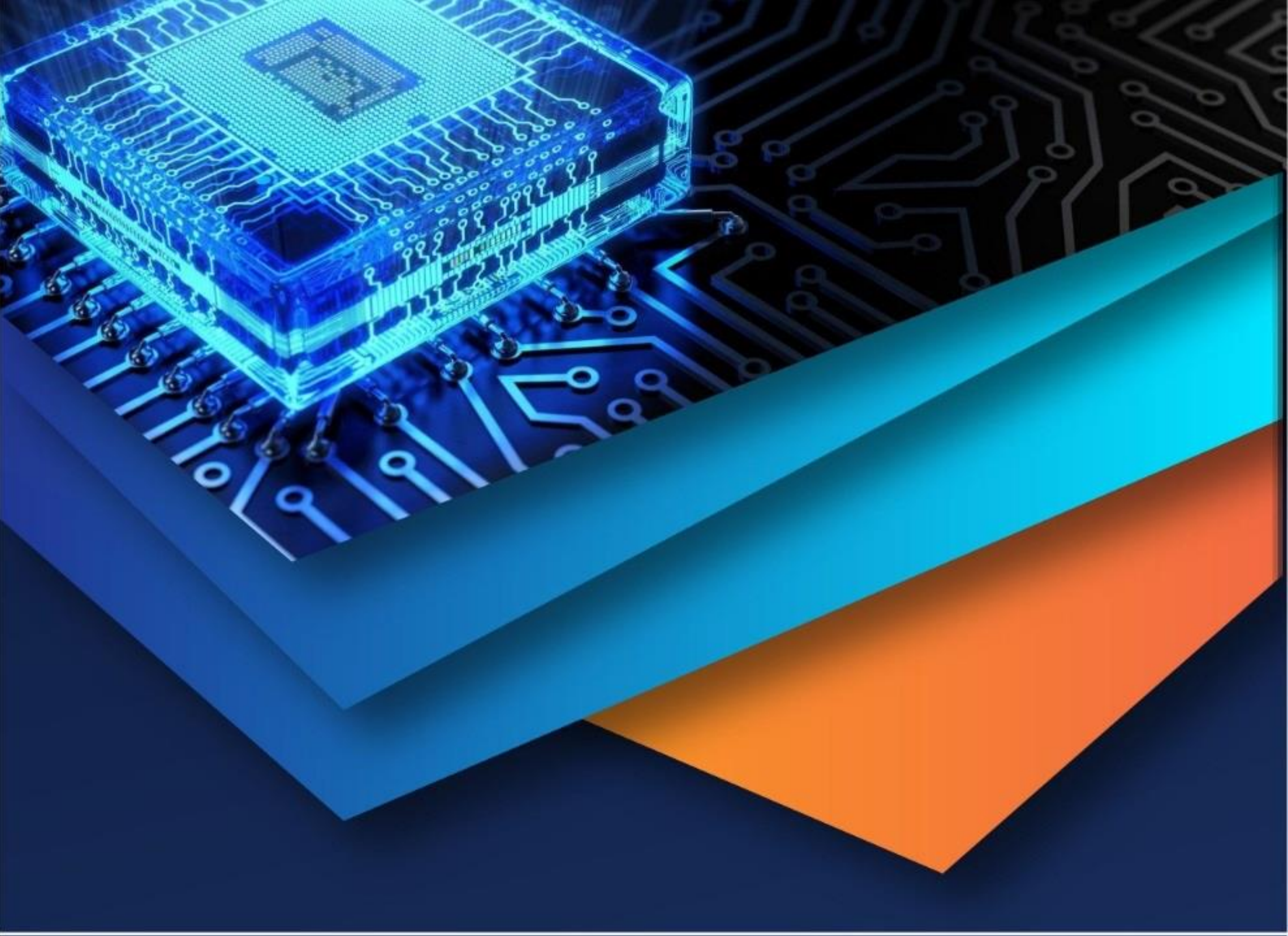
[1] R. S. Sukthanker, A. Zela, B. Stafler, S. Dooley, J. Grabocka, and F. Hutter, “Multi-objective Differentiable Neural Architecture Search,” *arXiv preprint arXiv:2402.XXXX*, Feb. 2024
[2] M. Koteswararao and P. R. Karthikeyan, “Accurate and Real-Time Object Detection System using YOLO v3-320 in Comparison with MobileNet SSD Network,” International Journal of Advanced Computer Science and Applications, 2023.
[3] S. Guo, Y. Liu, and W. Ni, “Lightweight SSD: Real-Time Lightweight Single Shot Detector for Mobile Devices,” IEEE Transactions on Neural Networks and Learning Systems, 2023.
[4] S. Liu, L. Huang, and X. Wu, “Lightweight Single Shot Multi-Box Detector: A Fabric Defect Detection Algorithm Incorporating Parallel Dilated Convolution and Dual Channel Attention,” Journal of Visual Communication and Image Representation, Oct. 2023.
[5] X. Deng and S. Li, “An Improved SSD Object Detection Algorithm Based on Attention Mechanism and Feature Fusion,” Sensors, Mar. 2023.
[6] N. Anggraini, S. H. Ramadhani, L. K. Wardhani, N. Hakiem, I. M. Shofi, and M. T. Rosyadi, “Development of Face Mask Detection using SSDLite MobileNetV3 Small on Raspberry Pi 4,” International Journal of Advanced Computer Science and Applications, 2022.
[7] T. Repák, “Fast Object Detection on Mobile Platforms using Neural Networks,” Brno University of Technology, Spring 2021.
[8] K. G. Mills, F. X. Han, M. Salameh, S. S. C. Rezaei, L. Kong, W. Lu, S. Lian, S. Jui, and D. Niu, “L2NAS: Learning to Optimize Neural Architectures via Continuous-Action Reinforcement Learning,” arXiv preprint arXiv:2109.XXXX, Sept. 2021.

[9] R. V. Iyer, P. S. Ringe, and K. P. Bhensdadiya, "Comparison of YOLOv3, YOLOv5s and MobileNet-SSD V2 for Real-Time Mask Detection," Proceedings of the International Conference on Intelligent Systems and Communication Networks, July 2021.

[10] M. Hajizadeh, M. Sabokrou, and A. Rahmani, "MobileDenseNet: A New Approach to Object Detection on Mobile Devices," Pattern Recognition Letters, 2020.

[11] X. Chu, B. Zhang, and R. Xu, "MoGA: Searching Beyond MobileNetV3," arXiv preprint arXiv:1908.XXXX, Aug. 2019.

[12] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, "Searching for MobileNetV3,"



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY