Object Detection Using AI (Mobile Application)

Mrs. Ahirekar S.Y
Computer Engineering
BSCOER,POLY.Narhe

Shubham Patil
Computer Engineering
BSCOER,POLY.Narhe

Harshad Salunke
Computer Engineering
BSCOER,POLY.Narhe

Anup Pawar
Computer Engineering
BSCOER,POLY.Narhe

Nitesh Kachare
Computer Engineering
BSCOER,POLY.Narhe

ABSTRACT

Object detection is a fundamental task in computer vision that involves identifying and locating objects within images or video streams. With the advancement of Artificial Intelligence (AI) and Deep Learning, object detection systems have achieved high accuracy and real-time performance. This paper presents an AI-based Object Detection System that uses deep learning algorithms to detect and classify objects in real time.

The proposed system utilizes Convolutional Neural Networks (CNN) and pre-trained models such as YOLO (You Only Look Once) and SSD (Single Shot Detector) to recognize objects efficiently. The system processes input images or live camera feeds, detects multiple objects, and provides bounding boxes along with labels and confidence scores.

The system can be implemented in various applications such as surveillance, autonomous vehicles, healthcare, and mobile applications. Experimental results demonstrate that the proposed system provides fast and accurate object detection, making it suitable for real-time applications.

Keywords: Object Detection, Artificial Intelligence, Deep Learning, CNN, YOLO, Computer Vision, Image Processing

I.INTRODUCTION:

Artificial Intelligence (AI) has emerged as one of the most transformative technologies in recent years, significantly impacting various domains such as healthcare, transportation, security, and automation. One of the most important applications of AI is in the field of computer vision, where machines are trained to interpret and understand visual data from the real world. Object detection is a key task in computer vision that involves not only identifying objects present in an image or video but also determining their exact location using bounding boxes.

Object detection plays a crucial role in many real-world applications. In autonomous vehicles, it helps in detecting pedestrians, traffic signals, and other vehicles to ensure safe navigation. In surveillance systems, it is used to monitor activities and detect suspicious behavior. In healthcare, object detection assists in analyzing medical images such as X-rays and MRIs for disease diagnosis. Similarly, in retail and manufacturing, it is used for inventory management, quality control, and automation. Traditional object detection methods relied on manual feature extraction techniques such as edge detection, color segmentation, and shape analysis. These methods were often time-consuming and lacked robustness when dealing with complex images, varying lighting conditions, and occlusions.

With the advancement of machine learning and deep learning, especially Convolutional Neural Networks (CNNs), object detection has become more accurate and efficient. CNNs automatically learn important features from data, eliminating the need for manual feature engineering.

Modern object detection techniques are mainly divided into two categories: two-stage detectors and single-stage detectors. Two-stage detectors, such as R-CNN, Fast R-CNN, and Faster R-CNN, first generate region proposals and then classify them. These methods provide high accuracy but are relatively slower. On the other hand, single-stage detectors like YOLO (You Only Look Once) and SSD (Single Shot Detector) perform detection and classification in a single step, making them faster and suitable for real-time applications.

In this paper, we propose an AI-based object detection system that utilizes deep learning models for real-time

object identification and localization. The system can process images or live video streams and detect multiple objects simultaneously with high accuracy. By leveraging pre-trained models and transfer learning techniques, the system reduces training time while maintaining performance. The proposed solution is efficient, scalable, and can be integrated into mobile and embedded systems for practical deployment.

II.   LITERATURE REVIEW:

Object detection has evolved significantly over the years with the development of machine learning and deep learning techniques. Earlier methods such as Haar Cascades and Histogram of Oriented Gradients (HOG) were widely used but had limitations in accuracy and scalability.

With the emergence of deep learning, models like R-CNN, Fast R-CNN, and Faster R-CNN improved detection accuracy by using region-based approaches. However, these models were computationally expensive and slower for real-time applications.

To overcome these limitations, real-time object detection models such as YOLO (You Only Look Once) and SSD (Single Shot Detector) were introduced. YOLO divides the image into grids and predicts bounding boxes and class probabilities in a single pass, making it extremely fast. SSD also performs object detection in a single shot with good accuracy.

Recent research focuses on improving detection speed, accuracy, and efficiency using advanced neural networks and transfer learning techniques. Many applications now use mobile-based object detection systems for real-world implementations.

III.   SYSTEM ARCHIETECTURE:
The proposed system consists of four main III. SYSTEM ARCHITECTURE

The proposed Object Detection System using Artificial Intelligence is designed as a comprehensive and integrated pipeline that efficiently processes images or real-time video streams to detect, classify, and localize multiple objects with high accuracy. The system begins with the input module, which captures visual data from various sources such as mobile cameras, webcams, CCTV feeds, or stored image and video datasets. This raw input is then passed to the preprocessing module,

where several important operations such as image resizing, normalization of pixel values, noise reduction, and format conversion are performed to ensure that the data is consistent and suitable for deep learning models.

After preprocessing, the data is forwarded to the feature extraction module, which uses advanced Convolutional Neural Networks (CNNs) such as ResNet, VGG, or MobileNet to automatically extract meaningful features like edges, textures, patterns, and object shapes. These extracted features play a crucial role in identifying objects accurately. The processed features are then given to the core component of the system, the object detection module, which uses state-of-the-art algorithms such as YOLO (You Only Look Once) or SSD (Single Shot Detector). These algorithms are capable of detecting multiple objects in a single frame, generating bounding boxes around each object, and assigning confidence scores indicating the probability of correct detection.

Following detection, the classification module assigns appropriate labels to each detected object, such as person, car, animal, or other predefined categories, while also filtering out low-confidence predictions to improve accuracy. The output module then displays the final results in an understandable format, showing bounding boxes, object names, and confidence levels on the screen in real time, which can be viewed on mobile devices, computers, or embedded systems. Additionally, the system includes a database or model storage component that stores pre-trained models, datasets, and detection results for future use and analysis.

IV. FEEDBACK LOOP

The feedback loop is an essential component of the proposed Object Detection System using Artificial Intelligence, as it enables continuous improvement and adaptability of the model. In this system, the output generated after object detection and classification is not considered final; instead, it is analyzed and fed back into the system to enhance future performance. When the system processes images or video streams, it may sometimes produce incorrect detections, missed objects, or low-confidence predictions. These errors are identified and stored along with the corresponding input data.

This collected feedback data is then used to update and retrain the deep learning model, allowing it to learn from its mistakes and improve its accuracy over time. The feedback loop involves comparing predicted outputs with expected results, identifying deviations, and adjusting model parameters accordingly through training processes such as backpropagation. Additionally, new data collected from real-world environments can be continuously added to the dataset, making the system more robust and capable of handling diverse conditions such as lighting variations, object occlusion, and complex backgrounds.

V. CONCLUSION

The Object Detection System using Artificial Intelligence demonstrates the effective use of deep learning techniques in identifying and localizing objects within images and real-time video streams. By integrating advanced algorithms such as Convolutional Neural Networks (CNNs) along with modern detection models like YOLO and SSD, the system achieves high accuracy and fast processing speed, making it suitable for real-time applications. The proposed system successfully overcomes the limitations of traditional object detection methods by automating feature extraction and improving detection performance in complex environments.

The architecture of the system, which includes modules such as input processing, feature extraction, object detection, classification, and output visualization, ensures a structured and efficient workflow. Additionally, the incorporation of a feedback loop mechanism allows the system to continuously learn from new data and improve its accuracy over time. This makes the system adaptable and scalable for various real-world scenarios.

VI. REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified Real-Time Object Detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[2] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "SSD: Single Shot MultiBox Detector," *European Conference on Computer Vision (ECCV)*, 2016.

[3] R. Girshick, "Fast R-CNN," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015.

[4] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation (R-CNN)," *CVPR*, 2014.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *CVPR*, 2016.

[6] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems (NIPS)*, 2012.

[7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *NIPS*, 2015.

[8] M. Everingham et al., "The PASCAL Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, 2010.

[9] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," *ECCV*, 2014.

[10] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *CVPR*, 2015.