



Enhanced Object Detection and Distance Prediction for Autonomous Vehicles Using Deep Neural Network

¹R Gomathi Jayam and ^{*2}R Shobana

¹Head & Assistant Professor, Department of Computer Science, Bon Secours College for Women (Autonomous), Thanjavur, Tamil Nadu, India.

^{*2}Student of II Year M.Sc., Department of Computer Science, Bon Secours College for Women (Autonomous), Thanjavur, Tamil Nadu, India.

Abstract

The proposed system is an efficient real-time video processing framework designed for accurate object detection and visual analysis using advanced deep neural network techniques. It supports both live camera feeds and prerecorded videos, making it suitable for various real-world applications. The system preprocesses video frames through resizing, normalization, and blob conversion to ensure fast and stable performance under real-time conditions. A deep learning model detects objects with high precision, while confidence-based filtering removes unreliable predictions to improve accuracy. Additional features such as gesture analysis and object tracking enhance adaptability for specialized tasks. The framework includes a real-time performance monitor to maintain low latency and displays annotated video output for easy interpretation. Overall, it provides a scalable, high-accuracy solution for applications like intelligent surveillance, robotics, industrial automation, and autonomous navigation.

Keywords: Real-Time Video Processing, Object Detection, Deep Neural Networks (DNN), Computer Vision, Frame Preprocessing, Confidence Filtering, Gesture Analysis, Object Tracking, Intelligent Surveillance, Autonomous Systems.

Introduction

Real time video processing and object detection have emerged as critical technologies across various industries, including security, robotics, autonomous vehicles and human computer interaction. The ability to analyze video streams in real time and accurately detect objects has become a cornerstone of intelligent systems, enabling applications ranging from surveillance to gesture recognition. This paper introduces a robust and adaptable video processing system designed for real time object detection and analysis, leveraging advanced deep neural network algorithms to process both live video streams and prerecorded video files. The system aims to provide high accuracy, fast processing speeds, and versatility, making it suitable for a wide array of real world applications. The workflow of the system begins by capturing video frames from either a live camera feed or a prerecorded video file. These frames are resized to ensure efficient processing, reducing computational overhead while preserving the essential information needed for detection. Once resized, the frames are transformed into image blobs, a common data format used in deep learning for easier manipulation and analysis. This conversion step optimizes the frames, preparing them for deep learning models by normalizing the data and allowing the system to handle the frames quickly and efficiently. At the heart of the system is a deep neural network algorithm capable of accurately detecting objects within the

video frames. The network is trained to recognize a variety of object types, including people, vehicles and various other entities commonly encountered in real world settings. By utilizing cutting edge deep learning techniques such as convolutional neural networks the system is able to achieve high levels of precision in object detection, even in complex and dynamic environments. To further enhance the accuracy of the system, the detection results are filtered using confidence thresholds, which help minimize the impact of false positives and ensure that only reliable detections are retained. This results in more accurate and actionable data for the user. Once objects are detected, the system generates bounding boxes around each identified object. These bounding boxes are then annotated with labels, allowing users to easily visualize the objects in the frame. This visual representation is not only intuitive but also essential for applications that require quick decision making, such as surveillance or autonomous systems. Additionally, for more specialized tasks like human object interaction detection or gesture recognition, the system incorporates advanced methods such as source to skin distance.

Review of Literature

Real-time object detection has gained significant attention with the advancement of deep learning and computer vision technologies. Earlier approaches relied on traditional image

processing techniques and handcrafted features such as Histogram of Oriented Gradients (HOG) and Haar cascades. Although these methods were computationally efficient, they lacked robustness and accuracy in complex environments.

With the introduction of deep convolutional neural networks (CNNs), object detection performance improved significantly. Models such as R-CNN, Fast R-CNN, and Faster R-CNN enhanced detection accuracy by using region-based proposals, but they required higher computational resources. Later, single-stage detectors like YOLO (You Only Look Once) and SSD (Single Shot Detector) were developed to achieve real-time detection with improved speed while maintaining competitive accuracy.

Recent studies focus on optimizing detection frameworks for real-time video processing by incorporating frame preprocessing, normalization, and efficient blob conversion techniques. Confidence-based filtering and non-maximum suppression methods are widely used to eliminate redundant or low-confidence detections. Furthermore, research has expanded into advanced applications such as gesture recognition, human tracking, and intelligent surveillance systems.

Overall, the literature indicates that combining optimized preprocessing pipelines with modern deep neural network architectures significantly enhances detection speed, accuracy, and scalability, making real-time object detection suitable for surveillance, robotics, industrial automation, and autonomous navigation systems.

Swarm Optimization with Neural Networks for Effective Classification Techniques" by K. Kalyani (2021) introduces a hybrid EHBMO-NN model, combining Extended Honey Bee Mating Optimization with Artificial Neural Networks to improve classification accuracy and reduce training time. It uses HBMO to select optimal weights for neural network hidden layers, outperforming conventional methods on benchmark datasets. The accurate cancer classification is very important task for cancer treatment. Recently the informative genes are identified from the thousands of genes for correct cancer classification. The collection of microscopic Deoxyribo Nucleic Acid (DNA) microarray is attached in the solid surface. In this study, DNA microarray data is used for cancer classification. The accurate cancer classification is very important task for cancer treatment. Recently the informative genes are identified from the thousands of genes for correct cancer classification. The collection of microscopic Deoxyribo Nucleic Acid (DNA) microarray is attached in the solid surface. In this study, DNA microarray data is used for cancer classification (6).

Existing System

Existing video processing systems for real time object detection generally rely on traditional image processing techniques, such as motion detection, edge detection and color segmentation, for identifying objects within video frames. These systems can process live video streams and prerecorded files but are often limited by their reliance on simpler, rule based algorithms that do not adapt well to complex environments or varying object types. Existing systems typically involve frame capture, resizing and basic filtering to prepare images for detection, but the accuracy and speed are often constrained by the simplicity of the models used. They can provide basic object identification, but their performance degrades significantly under conditions of high computational load, poor image quality or cluttered environments. Furthermore, many existing systems do not offer real time

performance monitoring or efficient visualization of detection results, limiting their usability for high stakes or time sensitive applications like surveillance and autonomous systems. Many existing systems cannot process high resolution or high frame rate videos in real time due to the heavy computational load imposed by conventional algorithms. This leads to delays in detection and feedback, which is unacceptable in critical applications like surveillance or robotics.

Proposed Algorithm

The proposed system is an advanced, real time video processing solution that utilizes deep neural network algorithms for object detection and analysis. It is capable of processing both live video streams and prerecorded video files, making it adaptable to various real time applications. The system begins by capturing video frames, resizing them for uniform processing and converting them into image blobs for efficient deep learning analysis. This preprocessing ensures faster and more reliable detection, even under computationally heavy tasks. The core of the system is a deep neural network designed to detect and classify objects with high precision. To enhance accuracy, the system uses confidence thresholds to filter out false positives and ensure that only the most reliable detections are presented. Bounding boxes are then drawn around the detected objects, providing clear visual feedback. For specialized applications such as human object interaction and gesture recognition, the system integrates advanced techniques, such as source to skin distance calculations, allowing for more context aware detection. The systems output is user friendly, featuring annotated frames with labels and bounding boxes for each detected object. Real time performance monitoring is built in with frames per second counter, ensuring smooth operation. This system can be applied across various domains, including security, robotics and human computer interaction, with the flexibility to scale and adapt to future needs.

Overview of the Proposed Algorithm

- **Video Input Acquisition:** Capture video from a live camera or load a prerecorded video file.
- **Frame Extraction:** Convert the video into individual frames for processing.
- **Preprocessing:** Resize frames, normalize pixel values, and convert them into image blobs for efficient model inference.
- **Deep Neural Network Detection:** Pass the processed frames to the trained deep learning model to detect objects and generate bounding boxes with confidence scores.
- **Confidence Filtering:** Remove low-confidence predictions to improve detection accuracy and reduce false detections.
- **Output Display:** Display annotated video frames with detected objects and labels.
- **Performance Monitoring:** Calculate and display real-time FPS (Frames per Second) to ensure low latency and smooth operation.

Experimented Results

The proposed real-time object detection system was evaluated using both live camera feeds and prerecorded video datasets. The model successfully detected multiple objects with high accuracy under different lighting and environmental conditions. Confidence-based filtering reduced false

detections and improved overall precision. The system maintained stable real-time performance with consistent frame rates and low latency. Experimental results demonstrate

that the framework provides reliable, accurate, and efficient object detection suitable for real-world applications such as surveillance and automation.



Datasets Used

This image shows the output folder of the object detection project. After running the detection model, all the processed images and videos are automatically saved inside the folder named exp4. The images contain detected objects with bounding boxes and labels, and the video file shows object detection results in real time. In simple words, this folder stores the final results generated by the system after detection.

Conclusion

We presented successfully demonstrates the development of a robust and adaptable video processing system for real time object detection and analysis, utilizing advanced deep neural network algorithms. By combining efficient preprocessing techniques, such as frame resizing and conversion into image blobs, with high performance deep learning models, the system ensures fast and reliable object detection even under heavy computational loads. The integration of confidence thresholds for filtering false positives and generating bounding boxes around detected objects significantly enhances the accuracy and clarity of the detection results, making them more interpretable and actionable for users. Additionally, the inclusion of specialized methods, such as source to skin distance calculations, allows the system to cater to more complex use cases, like human object interaction and gesture recognition. These capabilities make the system highly adaptable to various real time applications in fields such as surveillance, robotics and automation. The user friendly output, with annotated frames and a real time frames per second counter, provides clear visual feedback and ensures smooth system performance. By addressing the challenges of real-time video analysis, the project offers a significant contribution to advancing technologies in security, human computer interaction and autonomous systems..

References

1. Ma J, Lin M, Zhou G, Jia Z. *Joint Image Restoration For Domain Adaptive Object Detection In Foggy Weather Condition*. 2025 Oct.
2. Wang Z, Zhou G, Ma J, Xue T, Jia Z. *Beyond the Snowfall: Enhancing Snowy Day Object Detection Through Progressive Restoration and Multi-Feature*
3. Ros D, Dai R. *A Flexible Fall Detection Framework Based on Object Detection and Motion Analysis*. 2023 Feb.
4. Tang Z, Liu Y, Shang Y. *A New GNN-Based Object Detection Method for Multiple Small Objects in Aerial Images*. 2023 Jun.
5. Zhang Y, Zeng W, Yang F. *Towards Effective Microalgae Object Detection Solutions to IEEE UV 2022 "Vision Meets Algae" Object Detection Challenge*. 2022 Oct.
6. Wu Z, Liu C, Huang C, Wen J, Xu Y. *Deep Object Detection with Example Attribute Based Prediction Modulation*. 2022 May.
7. Shreyas E, Sheth MH, Mohana. *3D Object Detection and Tracking Methods using Deep Learning for Computer Vision Applications*. 2021 Aug.
8. Huang F, Taol D, Wang L. *DTB-Net: A Detection and Tracking Balanced Network for Fast Video Object Detection in Embedded Mobile Devices*. 2021 May.
9. Dong E, Zhang Y, Du S. *An Automatic Object Detection and Tracking Method Based on Video Surveillance*. 2020 Oct.
10. Li W, Li F, Luo Y, Wang P, sun J. *Deep Domain Adaptive Object Detection: a Survey*. 2020 Dec.
11. Kalyani K. *Swarm Optimization with Neural Networks for Effective Classification Techniques*. Annals of the Romanian Society for Cell Biology. 2021;25(4):7413-7419.
12. Kalyani K. *Classification of Microarray Gene Expression using Artificial Neural Network (ANN)*. Turkish Journal of Computer and Mathematics Education. 2021;12(7):1372-1378.
13. Kalyani K, Chakravarthy T. *An Algorithmic Approach with Improved Replacement in Bee Optimization Algorithm*. ICTACT Journal on Soft Computing. 2015 Jan 1;5(2):905-910.
14. Kalyani K. *Microarray Data Classification using Artificial Neural Network*. International Journal of Engineering and Advanced Technology (IJEAT). 2019 Dec;9(1S2):54-56.