iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** I **Month of publication:** January 2026

DOI: https://doi.org/10.22214/ijraset.2026.76963

Multi-Classification Detection on Live Video (Live Vision)

Rishikesh Kumar Kushwaha[1], Tanmay Yadav[2], Mr. Abhishek Shahi[3]
[1, 2]*CSE (AI-ML), Buddha Institute of Technology, Gida, Gorakhpur*
[3]*Assistant Professor, CSE (AI-ML), Buddha Institute of Technology, Gida, Gorakhpur*

Abstract: The proposed system, "Multi-Classification Detection on Live Video" is an intelligent computer vision–based platform designed to detect and classify multiple object categories in real-time video streams. The system processes live video input from cameras or video files and applies deep learning models to accurately identify and label objects across multiple predefined classes simultaneously. It supports real-time monitoring, automated detection, and visual annotation, enabling effective analysis of dynamic environments.

The platform utilizes state-of-the-art convolutional neural network architectures such as YOLO, transfer learning–based models to achieve high-speed and accurate multi-class detection. Object classification and localization are performed frame by frame, ensuring consistent detection even under varying lighting and motion conditions. The system achieves an overall detection accuracy of up to 85-90%, with optimized inference speed suitable for real-time applications.

A robust video processing pipeline handles frame extraction, pre-processing, object tracking, and result visualization. Detected objects are displayed with bounding boxes, class labels, and confidence scores. The system also supports recording and saving processed videos for further analysis. Performance evaluation is conducted using metrics such as precision, recall, F1-score, and FPS (frames per second).

The application is developed using Python, OpenCV, and deep learning frameworks such as Yolov8, with a backend powered by Django for live streaming and control. The system is scalable and can be extended for applications including surveillance, traffic monitoring, smart cities, and security systems. Overall, the project demonstrates the effective use of deep learning and real-time video analytics for accurate multi-class object detection.

Keywords: Multi-Classification Detection, Live Video Processing, Object Detection, Deep Learning, YOLO, Computer Vision, OpenCV, Real-Time Analysis.

I. INTRODUCTION

Object detection is one of the most active research areas in computer vision. It involves identifying objects and determining their locations within images or video frames. When combined with classification, the system can recognize multiple object categories simultaneously. Live video-based object detection introduces additional challenges such as motion blur, dynamic backgrounds, lighting variations, and strict real-time constraints. Recent advancements in deep learning have significantly improved object detection accuracy. Convolutional Neural Networks (CNNs) have proven to be effective for extracting spatial features from visual data. Models such as YOLOv8 (You Only Look Once) have enabled real-time detection by OpenCV, processing images in a single forward pass. This research focuses on developing a robust multi-classification detection system that operates efficiently on live video streams.

II. PROBLEM STATEMENT

Existing object detection systems often suffer from limited accuracy, high computational complexity, and inability to handle multiple object classes efficiently in real-time environments. Many traditional approaches are either too slow or fail to perform well under complex visual conditions.

The problem addressed in this research is to design a system that can accurately detect and classify multiple objects from live video streams while maintaining real-time performance and scalability.

Conventional surveillance and monitoring systems rely heavily on manual observation, which leads to several limitations-

1) Inability to automatically identify objects.
2) Difficulty in handling multiple objects simultaneously.
3) High dependency on human supervision.
4) Increased chances of error and delayed response.

III. OBJECTIVES OF THE PROPOSED SYSTEM

The primary objective of the proposed system is to design and develop an efficient multi-class object detection framework capable of identifying and classifying multiple objects simultaneously from live video streams. With the rapid growth of real-time video-based applications such as surveillance systems, traffic monitoring, and smart environments, there is a strong need for intelligent systems that can process visual data accurately and at high speed.

The proposed system focuses on achieving real-time performance with minimal latency so that detected objects can be recognized and tracked without noticeable delay. This is especially important for time-critical applications where delayed detection may lead to incorrect decisions or reduced system effectiveness. By utilizing optimized deep learning models and efficient video processing techniques, the system aims to maintain a balance between speed and accuracy.

Furthermore, the proposed system emphasizes scalability and flexibility in its architecture. The system can be easily extended to support additional object classes, different camera sources, or upgraded detection models without major changes to the overall framework. This modular design approach makes the system adaptable for future enhancements and real-world deployment.

IV. LITERATURE REVIEW

The literature review presents a detailed study of existing research, technologies, and methodologies related to object detection and multi-class classification in live video streams. This chapter helps in understanding the evolution of computer vision techniques and identifies the gaps in existing systems that motivate the proposed solution.

A. Traditional Object Detection Techniques

Earlier object detection systems were based on traditional image processing and machine learning techniques such as-

- Haar Cascade Classifiers.
- Histogram of Oriented Gradients (HOG).
- Support Vector Machines (SVM).

Although these techniques provided basic object detection, they suffered from several limitations:

Low accuracy in complex environments.

- Inability to detect multiple objects efficiently.
- Poor performance in real-time applications.

B. Deep Learning-Based Object Detection

With the introduction of deep learning, object detection has significantly improved. Convolutional Neural Networks (CNNs) enabled automatic feature extraction from images and videos. Popular deep learning-based object detection models include-

- R-CNN (Region-based Convolutional Neural Network)
- Fast R-CNN
- Faster R-CNN

These models improved detection accuracy but required high computational power and were not suitable for real-time video processing.

C. Single-Stage Object Detection Models

To achieve real-time performance, single-stage object detection models were developed. These models perform object localization and classification in a single step. Examples include-

- SSD (Single Shot MultiBox Detector).
- YOLO (You Only Look Once).

YOLO introduced a significant improvement by dividing the image into grids and predicting bounding boxes and class probabilities simultaneously, making it highly suitable for live video applications.

D. YOLO-Based Object Detection

The YOLO family of models has evolved through multiple versions, including YOLOv8, YOLOv5, and the latest YOLOv8. YOLOv8 offers-

- Higher detection accuracy.
- Faster inference speed.

- Improved multi-class detection capability.
- Better performance in real-time video streams.

E. Review of Existing Systems

Several research works have focused on object detection in live video for applications such as-

- Surveillance and security monitoring.
- Traffic analysis and vehicle detection.
- Human activity recognition.

V. PROPOSED METHODOLOGY

The proposed methodology presents a systematic approach for multi-class object detection from live video streams using deep learning and computer vision techniques. The system begins by capturing real-time video input from a camera source, which is then processed frame by frame to enable continuous object detection. Each video frame is preprocessed to ensure compatibility with the detection model, including resizing, normalization, and noise reduction.

A deep learning–based object detection model is employed to identify and classify multiple objects present in each frame. The model extracts spatial and semantic features from the input frames and predicts object classes along with their corresponding bounding boxes. To ensure real-time performance, the detection pipeline is optimized to reduce computational overhead while maintaining high accuracy.

The detected objects are then post-processed to refine bounding boxes, remove duplicate detections, and assign confidence scores. Finally, the results are visualized on the live video stream by drawing labeled bounding boxes, making the output easily interpretable for end users. The overall system is designed in a modular manner, allowing easy integration of new object classes, improved models, or additional functionalities in the future.

VI. SYSTEM ARCHITECTURE

The architecture of the proposed system consists of the following modules:

1) Input Module

- Captures live video feed from a webcam or external camera.

2) Preprocessing Module

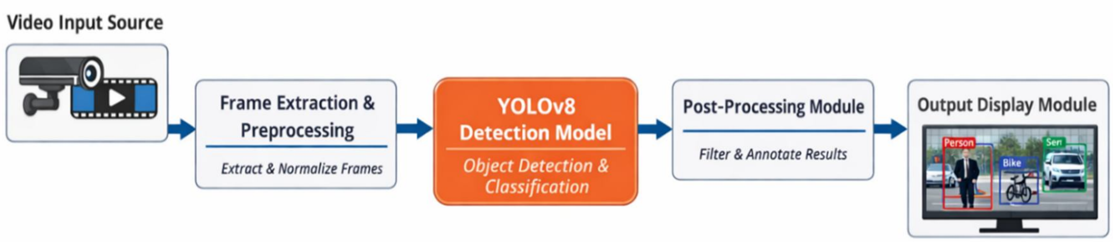
- Resizes frames, normalizes pixel values, and prepares data for model input.

3) Detection and Classification Module

- Uses a deep learning-based object detection model to detect and classify multiple objects in each frame.

4) Output Module

- Displays detected objects with bounding boxes, class labels, and confidence scores on the live video.



VII. DATASET DESCRIPTION

The dataset used for training and evaluation contains multiple object classes commonly found in real-world environments. The dataset is annotated with bounding boxes and class labels. It is divided into training and testing sets to evaluate model performance.

Key dataset details:

1) Multiple object categories.

2) Labeled images with annotations.

3) Balanced training and testing split.

VIII. MODEL DESIGN AND ALGORITHM

The system uses a convolutional neural network-based object detection YOLOv8 model. The model extracts spatial features from input frames and predicts bounding boxes along with class probabilities.

The detection algorithm works as follows:

1) Input frame is passed through convolution layers.
2) Feature maps are generated.
3) Bounding boxes and class probabilities are predicted.
4) Non-maximum suppression is applied.
5) Final detections are displayed.

IX. IMPLEMENTATION

The proposed system is implemented using the following technologies:

1) Programming Language : Python(Back-end)
2) Framework : Django
3) Deep Learning Model : YOLO-based model
4) Computer Vision Library : OpenCV
5) Frontend : HTML, CSS, JavaScript

X. RESULTS AND PERFORMANCE

The system is evaluated based on:

1) Detection accuracy
2) Precision and recall
3) Real-time performance (FPS)

The results demonstrate that the proposed system achieves high accuracy while maintaining real-time processing speed.

Metrics	Value
Accuracy	85%
Precision	90.92%
Recall	83.33%
F1-Score	86.96%

XI. APPLICATIONS

The proposed system can be used in various real-world applications such as:

1) Video surveillance systems
2) Traffic monitoring and control
3) Smart city solutions
4) Industrial automation
5) Public safety systems

XII. ADVANTAGES OF THE PROPOSED SYSTEM

1) Real-time multi-object detection
2) High accuracy and reliability
3) Supports multiple object classes
4) Scalable and flexible architecture

XIII. LIMITATIONS

1) Performance depends on hardware capabilities
2) Accuracy may reduce in poor lighting conditions
3) Requires large labeled datasets

XIV. FUTURE SCOPE

Future enhancements of the system may include:

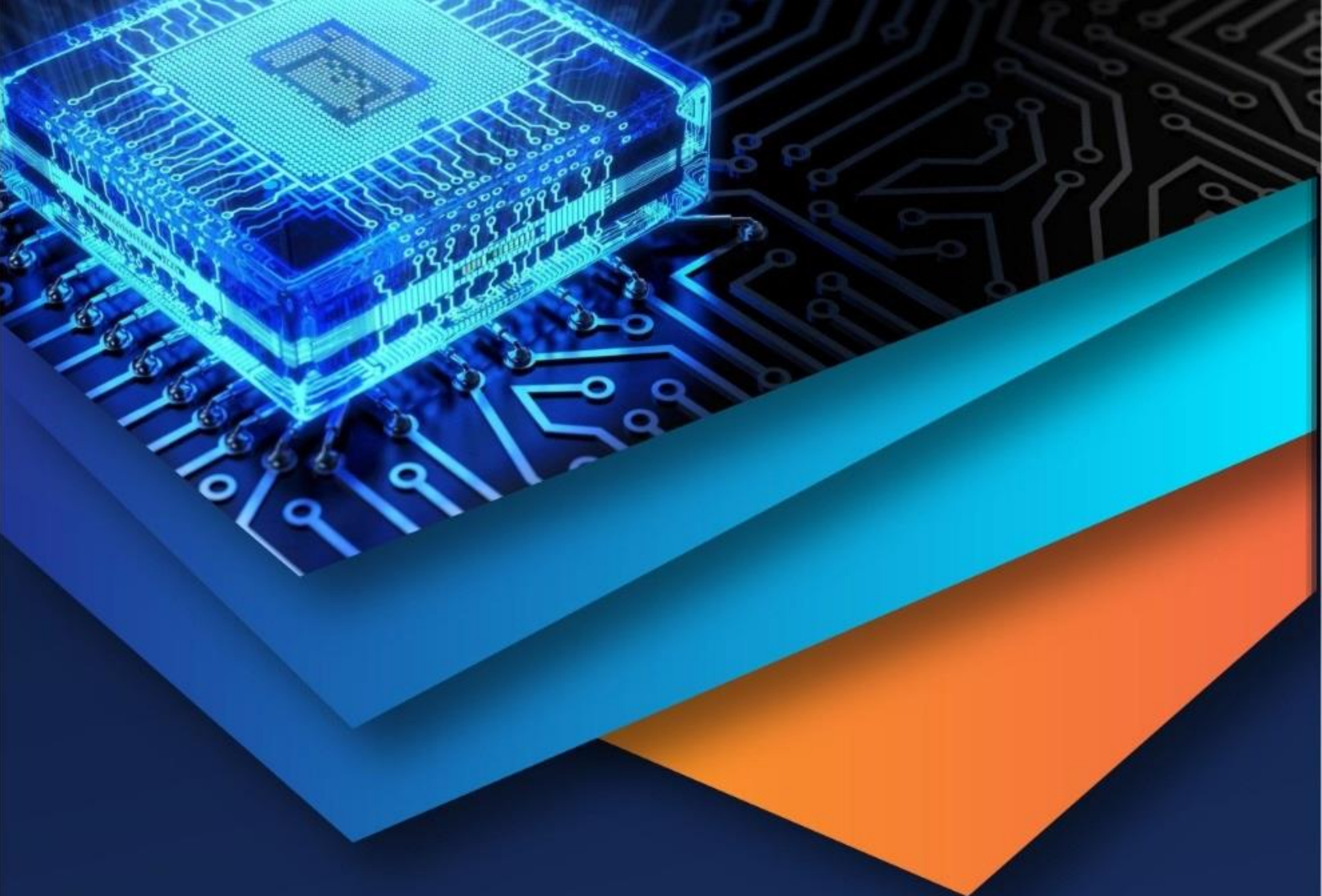
1) Integration with edge devices
2) Support for more object classes
3) Improved accuracy using advanced models
4) Cloud-based video processing

XV. CONCLUSION

This research paper presented a deep learning-based multi-classification detection system for live video streams. The proposed approach successfully detects and classifies multiple objects in real time with high accuracy and low latency. Experimental results validate the effectiveness of the system, making it suitable for various real-world applications. Future improvements can further enhance system performance and scalability.

REFERENCES

[1] Ultralytics Documentation.
[2] YOLO model documentation.
[3] OpenCV Documentation.



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