A

Mini-Project Report on

# Real-Time Object Detection using Computer Vision

Submitted in partial fulfillment of the requirements for the degree of

BACHELOR OF ENGINEERING

IN

Computer Science & Engineering

Artificial Intelligence & Machine Learning

by

Tanveer Angane (22106057)

Anish Gawade (22106109)

Siddharth Chaurasiya (22106060)

Under the guidance of

# Prof. Bharti Khemani



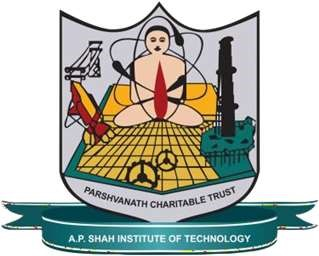
Department of Computer Science & Engineering

(Artificial Intelligence & Machine Learning)

A. P. Shah Institute of Technology

G. B. Road, Kasarvadavali, Thane (W)-400615

University Of Mumbai 2024-2025



# A. P. SHAH INSTITUTE OF TECHNOLOGY

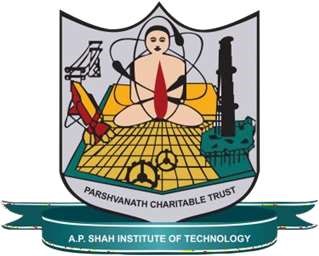
CERTIFICATE

This is to certify that the project entitled “Real-Time Object Detection using Computer Vision” is a bonafide work of Tanveer Angane (22106057), Anish Gawade (22106109), Siddharth Chaurasiya (22106060) submitted to the University of Mumbai in partial fulfillment of the requirement for the award of Bachelor of Engineering in Computer

Science & Engineering (Artificial Intelligence & Machine Learning).

|  |  |
| --- | --- |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| Prof. Bharti Khemani | Dr. Jaya Gupta |
| Mini Project Guide | Head of Department |

# A. P. SHAH INSTITUTE OF TECHNOLOGY



Project Report Approval

This Mini project report entitled “Real-Time Object Detection using Computer Vision” by Tanveer Angane, Anish Gawade and Siddharth Chaurasiya is approved for the degree of Bachelor of Engineering in Computer Science &Engineering, (AIML) 2024-25.

External Examiner:

Internal Examiner:

Place: APSIT, Thane

Date:

Declaration

We declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that We have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Tanveer Angane Anish Gawade Siddharth Chaurasiya

(22106057) (22106109) (22106060)

ABSTRACT

Object detection is a fundamental task in computer vision, enabling machines to identify and classify objects within images and videos. Real-time object detection is crucial for applications in autonomous vehicles, surveillance, robotics, healthcare, and augmented reality. The proposed system integrates computer vision techniques with deep learning frameworks such as TensorFlow and OpenCV to enhance detection accuracy and speed. The study highlights the trade-offs between accuracy and speed in different models and demonstrates how real-time processing can be achieved by optimizing neural network architectures. Experimental results show that modern object detection frameworks can achieve high accuracy with low computational costs, making them suitable for real-world applications. The model is optimized for efficient inference on edge devices and GPUs, ensuring minimal latency. Various datasets, including COCO and Pascal VOC, are utilized to train and evaluate performance. This research concludes that real-time object detection using deep learning provides significant advancements in automation and artificial intelligence, paving the way for future innovations in various industries.

Keywords: YOLO (You Only Look Once), OpenCV, Computer Vision.

## Index

|  |  |  |  |
| --- | --- | --- | --- |
| Index | | | Page no. |
| Chapter-1 | | | 1-3 |
|  | Introduction | | 2 |
|  |  |  |  |
| Chapter-2 | | | 4-7 |
|  | Literature Survey | |  |
|  | 2.1 | History | 5 |
|  | 2.1 | Review | 6 |
|  |  |  |  |
| Chapter-3 | | | 8-9 |
|  | Problem Statement | | 9 |
|  |  |  |  |
| Chapter-4 | | | 10-12 |
|  | Experimental Setup | |  |
|  | 4.1 | Hardware Setup | 11 |
|  | 4.2 | Software Setup | 11 |
|  |  |  |  |
| Chapter-5 | | | 13-15 |
|  | Proposed system and Implementation | |  |
|  | 5.1 | Block Diagram of proposed system | 14 |
|  | 5.2 | Description of Block diagram | 15 |
|  | 5.3 | Implementation | 16 |
|  |  |  |  |
| Chapter-6 | | | 20 |
|  | Conclusion | | 20 |
|  |  |  |  |
| References | | | 21 |
|  |  |  |  |

# CHAPTER 1

INTRODUCTION

1

## 1. INTRODUCTION

The integration of real-time object detection with IoT (Internet of Things) and computer vision has revolutionized numerous industries by enabling intelligent decision-making and automation. This technology allows IoT-connected devices, such as cameras, sensors, and embedded systems, to identify, track, and classify objects in real time without human intervention. By leveraging AI-driven computer vision algorithms, edge computing, and cloud processing, real-time object detection enhances efficiency, security, and automation across multiple domains, including healthcare, manufacturing, agriculture, security, and smart cities. Computer vision, powered by deep learning models such as YOLO enables machines to interpret and process visual data from IoT devices. These models allow real-time analysis of images and videos, making it possible to detect objects with high accuracy and speed. The IoT framework connects these smart devices over networks, enabling seamless data sharing and decision-making based on detected objects. In transportation and smart cities, real-time object detection optimizes traffic monitoring, vehicle tracking, and accident prevention. AI-powered smart cameras analyze traffic flow, detect violations, and improve road safety. Autonomous vehicles also rely on object detection to navigate, identify pedestrians, and avoid collisions.

Key Benefits:

* Enhanced Automation & Smart Systems - IoT-connected devices equipped with computer vision can automate tasks such as inventory management, facial recognition, smart surveillance, and predictive maintenance.
* Improved Security & Surveillance - Real-time object recognition in surveillance systems enhances security by detecting unauthorized personnel, weapons, or abandoned objects.
* AI-powered security cameras integrated with IoT can send instant alerts to authorities when suspicious activities are detected.
* Smart Traffic Management & Transportation - Traffic monitoring systems use AI-powered cameras to detect vehicles, pedestrians, and traffic violations, improving road safety.
* Smart parking systems use IoT sensors and computer vision to detect vacant parking spots in real time. 2

The integration of real-time object detection with IoT and computer vision has transformed industries by enhancing automation, security, and efficiency. As AI models continue to evolve, and edge computing technology advances, the adoption of smart vision-based IoT systems will increase across various sectors.

3

CHAPTER 2

# LITERATURE SURVEY

4

## 2. LITERATURE SURVEY

2.1-HISTORY

The evolution of Finance Tracking and Prediction systems has been shaped by advancements in financial management practices and technology over the past several decades. Initially, financial management relied on manual bookkeeping and basic accounting, with forecasting based on simple extrapolation of historical trends. The 1960s brought the introduction of computerized accounting systems, which improved data processing and accuracy, while early statistical methods like regression analysis began to enhance financial forecasting.

In the 1990s, the advent of Enterprise Resource Planning (ERP) systems integrated various business processes, including finance, providing real-time data access and more comprehensive analysis. This era also saw the incorporation of more complex statistical models into financial forecasting, supported by better data availability and computing power.

The early 2000s introduced Business Intelligence (BI) tools, which improved data visualization and reporting, making financial data more actionable. Concurrently, machine learning began to be applied to financial data, offering predictive capabilities beyond traditional methods. The 2010s further advanced these systems with the integration of big data and advanced analytics, enabling real-time processing and more accurate predictions. Machine learning and AI became central to financial forecasting, capturing complex relationships in financial data.

Today, Finance Tracking and Prediction systems are highly sophisticated, integrating real-time data, machine learning, and AI to provide comprehensive insights. Current trends focus on enhancing predictive accuracy, integrating diverse data sources, and improving user interfaces. The literature reflects a clear progression from manual methods to advanced, data-driven systems that are becoming increasingly crucial for effective financial management and decision-making.

5

Machine learning techniques applied to financial market prediction

The search for models to predict the prices of financial markets is still a highly researched topic, despite major related challenges. The prices of financial assets are non-linear, dynamic, and chaotic; thus, they are financial time series that are difficult to predict. Among the latest techniques, machine learning models are some of the most researched, given their capabilities for recognizing complex patterns in various applications. With the high productivity in the machine learning area applied to the prediction of financial market prices, objective methods are required for a consistent analysis of the most relevant bibliography on the subject. This article proposes the use of bibliographic survey techniques that highlight the most important texts for an area of research. Specifically, these techniques are applied to the literature about machine learning for predicting financial market values, resulting in a bibliographical review of the most important studies about this topic. Fifty-seven texts were reviewed, and a classification was proposed for markets, assets, methods, and variables. Among the main results, of particular note is the greater number of studies that use data from the North American market. The most commonly used models for prediction involve support vector machines (SVMs) and neutral network. It was concluded that the research theme is still relevant and that the use of data from developing markets is a research opportunity. [1]

Financial Tracker using NLP

NLP (Natural Language Processing) is a mechanism that helps computers to know natural languages like English. In general, computers can understand data, tables etc. which are well formed. But when it involves natural languages, it's unacceptable for computers to spot them. NLP helps to translate the tongue in such a fashion which will be easily processed by modern computers. Financial Tracker is an approach which will use NLP as a tool and can differentiate the user messages in various categories. the appliance of the approach will be seen at multiple levels. At a personal level, this permits users to filtrate useful financial messages from an large junk of text messages. On the opposite hand, from an industrial point of view, this can be useful in services like online loan disbursal, which are hitting the market nowadays. These services attempt to provide online loans to individuals in an exceedingly faster and quicker manner. But when it involves business view, loan recovery from customers becomes a really important & crucial aspect. As most such services can’t take strict legal actions against the fraud customers, it becomes a requirement that loan should be provided only to those customers who deserve it. At that time, this model can come under the image. As a business we will find the user’s messages from their inbox (after taking permission from the users). These messages are often filtered using NLP which might help to differentiate various types of messages within the user's inbox which might further be used as a content for further prediction and analysis on user’s behaviour in terms of cash related transactions. [2]

6

Prediction of Financial Crisis with Artificial Neural Network: An Empirical Analysis on Turkey

Prediction of economic crisis, financial distress or bankruptcy has attracted great deal of attention in financial literature and in many other fields among the researchers over the past few decades. Although there are a variety of different methods that can be used to predict the future financial crisis, due to the complexity of the existing factors, prediction of financial crisis is a very difficult case. With the advent of Artificial Neural Networks (ANNs), researchers had the chance to solve various problems in finance. ANN approach is the application of artificial intelligence, which has been improved by the simulation of cognitive learning process of human brain. ANNs are commonly used in recent years, due to major advantages that they offer such as their ability to perform nonlinear statistical modeling that provides new alternative to other statistical methods and to learn directly from examples without needing or providing an analytical solution to the problem. In this study, a monthly dataset covering the period of 1990 and 2014 that belong to the Turkish economy will be used. The purpose of this study is to develop an early-warning system to predict financial crisis. To realize this aim, multi-layered feedforward neural networks (MLFNs) will be used. By using monthly data of 7 key macroeconomic and financial indicators of Turkish economy during 1990 and 2014, we find that predictive power of ANN is quite striking. Our outof-sample forecasts indicate that the Turkish economy remains at high risk due to major negative developments and potential political instability between 2014 and 2016. [3]

Financial indicators tracker application (FIT)

The Internet has become a source of vast range of financial information from financial institutions. The existing financial data provider such as Google finance, Yahoo finance, Bloomberg and Reuters are those leaders in providing financial data for all company in the world. Currently, these providers only provide financial data without any financial analysis. In order to do a financial analysis, users would have to browse the web sites of these providers and calculate manually the financial ratios. The application developed in this project aims to address this predicament. The main objective of this project is to develop an application that integrates financial data from various providers and calculate the important financial ratios of public listed companies (PLC). The ratios provided in this application will identify the financial performance of the PLCs. The requirements of this application are given by Bursa Malaysia. For the scope of the project, ten financial ratios were identified for this application. These ten financial ratios are considered as among the important ratios in determining the performance of a company. The significance of this application is that it allows users to compare financial ratios of PLCs. The application ease the users to compare these ratios and assist them in deciding which companies stocks that they want to pursue. [4]

7

CHAPTER 3

# Problem Statement

8

## 3. Problem Statement

Real-time object detection via IoT and computer vision is critical for applications like security, healthcare, smart cities, and industrial automation. Yet, traditional systems face challenges such as outdated sensors, inefficient processing, and high computational demands, causing delays and inaccuracies. They struggle with occlusions, poor lighting, fast-moving objects, and scalability within IoT networks. To overcome these, modern solutions must integrate deep learning, edge computing, and energy-efficient algorithms to boost accuracy in tough environments, reduce latency, and ensure cybersecurity. By addressing these gaps, real-time detection can enhance safety, automation, and efficiency across industries.

9

CHAPTER 4

# EXPERIMENTAL SETUP

10

#### 4. EXPERIMENTAL SETUP

#### 4.1 HARDWARE SETUP

#### 1. Processor (CPU & GPU):

#### CPU: Intel Core i5/i7/i9 (10th Gen or higher) or AMD Ryzen 5/7/9 (3000 series or higher)

#### GPU (Recommended for Faster Training): NVIDIA GeForce RTX 3060/RTX 3080/Tesla V100

#### 2. Memory (RAM):

#### Minimum: 8GB

#### Recommended: 16GB or higher

#### 3. Storage:

#### Minimum: 256GB SSD

#### 4.2 SOFTWARE SETUP

#### 1. Programming Languages and tools:

#### Python

#### HTML (Flask)

#### Visual Studio Code : VS Code is a versatile and highly efficient tool for developing an image caption generator.

#### 11

#### 2. Operating System:

#### Windows 10/11 or macOS

#### 3. Libraries and dependencies :

#### Ultralytics

#### CVzone

#### CV2

#### Math

#### Machine Learning & Deep Learning:

#### YOLO

12

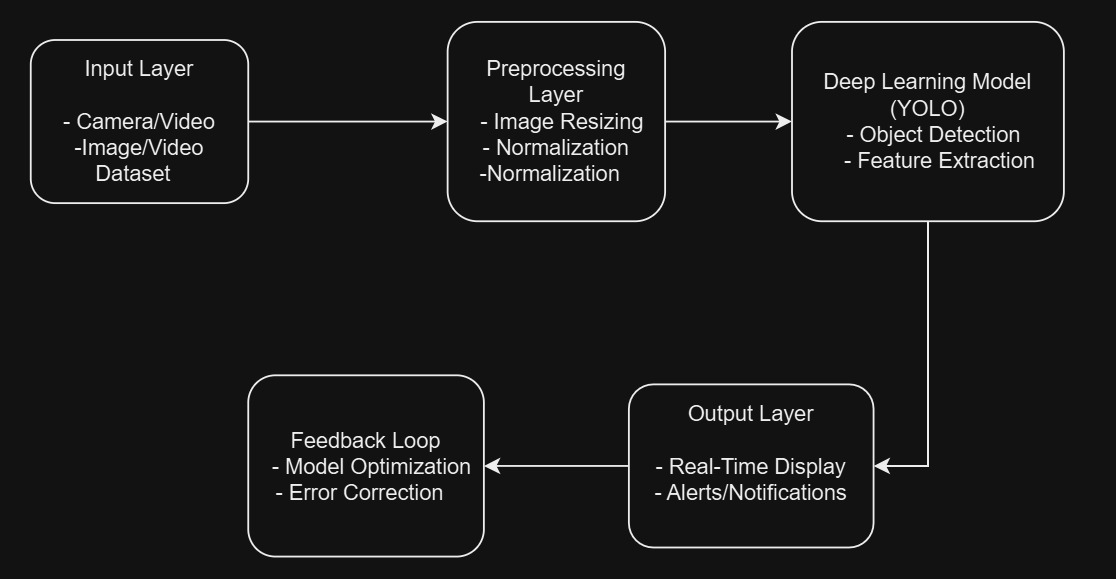
# CHAPTER 5

# PROPOSED SYSTEM & IMPLEMENTATION

#### 13

#### 5.PROPOSED SYSTEM & IMPLEMENTATION

5.1 BLOCK DIAGRAM OF PROPOSED SYSTEM



5.2 Description of Block diagram

This diagram outlines a real-time object detection system using IoT and computer vision, structured into sequential layers:

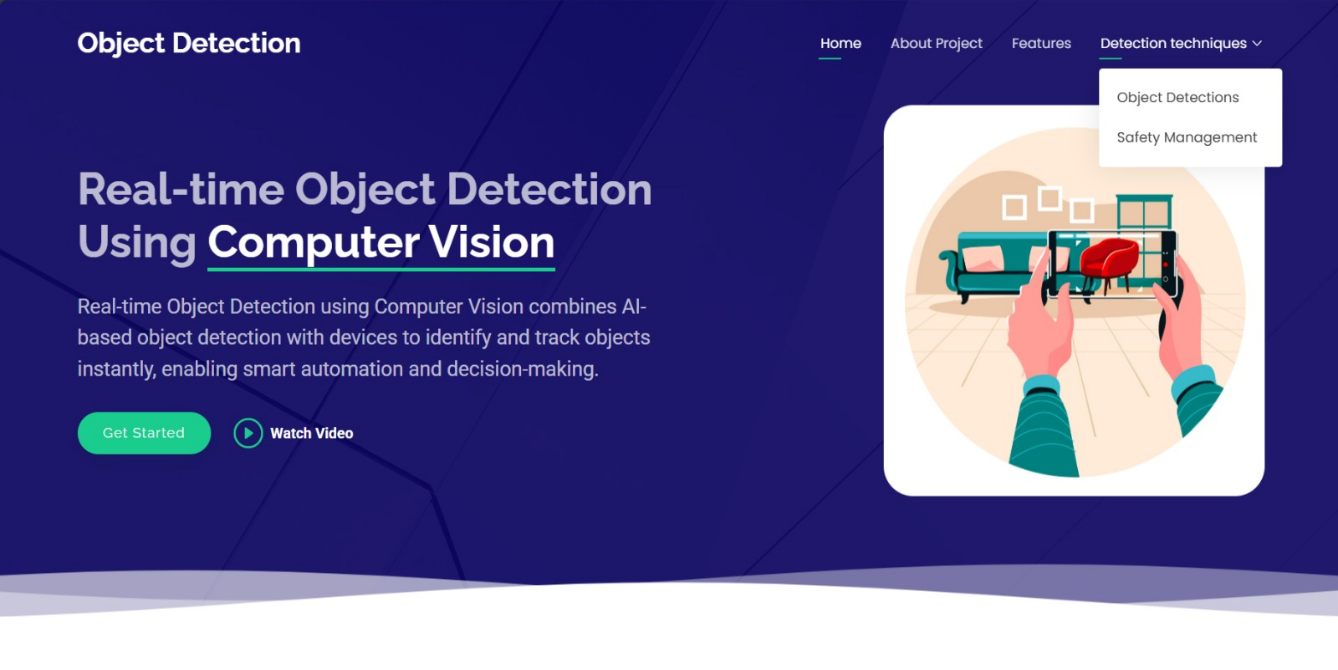
1. **Input Layer**
   * Camera/Video: Captures live visual data (e.g., surveillance feeds, traffic cameras).
   * Image/Video Dataset: Provides pre-recorded or labelled data for training or testing the system.
2. **Preprocessing Layer**
   * Image Resizing: Adjusts input dimensions to match the model’s requirements (e.g., YOLO’s 416x416 resolution).
   * Normalization: Scales pixel values (e.g., 0-255 to 0-1) to improve model training efficiency and accuracy.

14

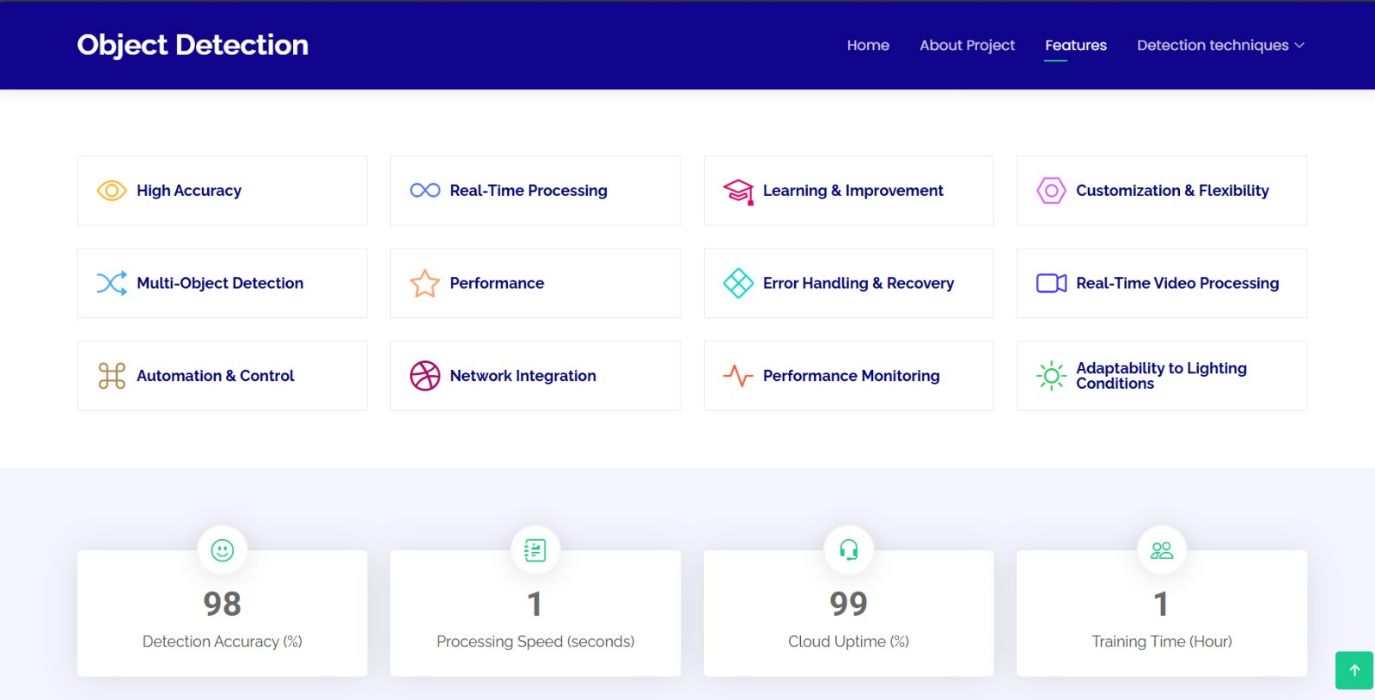
1. **Deep Learning Model (YOLO)**
   * Object Detection: Identifies and classifies objects in real time using YOLO’s single-pass architecture.
   * Feature Extraction: Analyses visual patterns (edges, textures, shapes) to distinguish objects.
2. **Feedback Loop**
   * Model Optimization: Continuously refines the model using new data or performance feedback.
   * Error Correction: Reduces false positives/negatives by adjusting parameters or retraining.
3. **Output Layer**
   * Real-Time Display: Shows detected objects with bounding boxes and labels on a live interface.
   * Alerts/Notifications: Triggers automated responses (e.g., security alerts, traffic warnings) based on detections.

15

5.3. IMPLEMENTATION



5.3.1. Dashboard

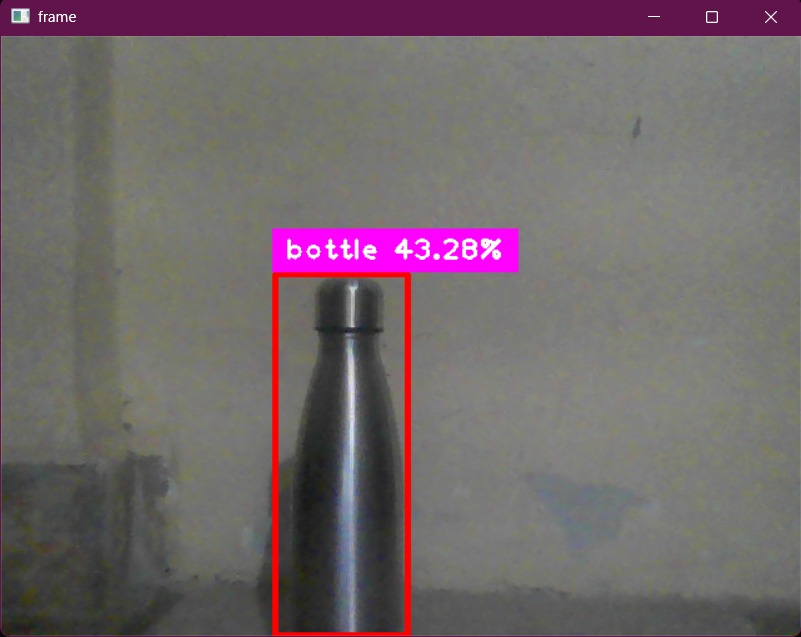


5.3.1. Applications and accuracy

16



5.3.1. Cell phone detected



5.3.1. Bottle detected

17



5.3.1. Scissor detected

18

# CHAPTER 5

# CONCLUSION

19

Conclusion

This project successfully implements real-time object detection using deep learning and computer vision techniques, optimizing accuracy and speed for practical applications. By leveraging frameworks like TensorFlow and OpenCV, the system achieves efficient detection with minimal computational overhead, making it suitable for surveillance, automation, and smart systems. Despite challenges such as occlusions and varying lighting conditions, the integration of AI and edge computing ensures low-latency processing and scalability. Future enhancements can focus on improving model efficiency and expanding its application scope to further refine real-time object detection capabilities.

20

References

* Henrique, Bruno Miranda, Vinicius Amorim Sobreiro, and Herbert Kimura. "Literature review: Machine learning techniques applied to financial market prediction." Expert Systems with Applications 124 (2019): 226-251. [1]
* Baliyan, Kartik, Shubham Vishnoi, and Swati Sharma. "Financial tracker using NLP." International Journal of Engineering and Management Research 11.3 (2021): 124-125. [2]
* Aydin, Alev Dilek, and Seyma Caliskan Cavdar. "Prediction of financial crisis with artificial neural network: an empirical analysis on Turkey." International journal of financial research 6.4 (2015): 36. [3]
* Kalid, Khairul Shafee, Muhammad Zulhilmi Bachok, and Md Akhir Mohd Sharif. "Financial indicators tracker application (FIT)." 2012 International Conference on Computer & Information Science (ICCIS). Vol. 1. IEEE, 2012. [4]

21