##### A Project report on

**Object Detection for Visually Impaired using YOLO v7**

###### A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the academic requirements for the award of the degree.

**Bachelor of Technology**

**in**

**Computer Science and Engineering**

Submitted by

N. SREEKAR

(20H51A0540)

A. SAI PRASHANTH

(20H51A05A9)

V. VANDANA

(20H51A05M4)

Under the esteemed guidance of

**Dr. SIVA SKANDHA SANAGALA**

(Associate Professor & HOD Dept. of CSE)



**Department of Computer Science and Engineering**

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

(UGC Autonomous)

\*Approved by AICTE \*Affiliated to JNTUH \*NAAC Accredited with A+ Grade

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

#### 2020- 2024

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**



#### CERTIFICATE

This is to certify that the Major Project Phase I report entitled **"Object Detection for Visually Impaired using YOLO v7"** being submitted by N. Sreekar (20H51A0540), A. Sai Prashanth (20H51A05A9), V. Vandana (20H51A05M4) in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering** is a record of bonafide work carried out his/her under my guidance and supervision.

###### The results embodied in this project report have not been submitted to any other University or Institute for the award of any Degree.

**Dr. Siva Skandha Sanagala**

**Associate Professor and**

**HOD Dept. of CSE**

#### ACKNOWLEDGEMENT

With great pleasure, we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project a grand success.

We are grateful to **Dr. Siva Skandha Sanagala, Associate Professor** and Head of the Department of Computer Science and Engineering, CMR College of Engineering and Technology for his valuable technical suggestions and guidance during the execution of this project work and for being a major driving force to complete my project work successfully.

We are very grateful to **Dr. Vijaya Kumar Koppula**, Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in successfully carrying out the project work.

We are highly indebted to **Major Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for permitting us to carry out this project in a successful and fruitful way.

We would like to thank the **Teaching & Non- teaching** staff of the Department of Computer Science and Engineering for their cooperation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, we extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project work.

N. Sreekar 20H51A0540

A. Sai Prashanth 20H51A05A9

V. Vandana 20H51A05M4

Object Detection for Visually Impaired using YOLO v7

**TABLE OF CONTENTS**

**CHAPTER**

**NO. TITLE PAGE NO.**

LIST OF FIGURES ii

LIST OF TABLES iii

ABSTRACT iv

**1** **INTRODUCTION** 1

1.1 Problem Statement 2

1.2 Research Objective 2

1.3. Introduction to YOLOv7 3

**2** **BACKGROUND WORK** 4

2.1 Research Paper 1 5

2.2 Research Paper 2 5

2.3 Research Paper 3 6

2.4 Research Paper 4 7

2.5 Research Paper 5 7

2.6 Research Paper 6 8

**3 RESEARCH COMPARISON TABLE 9**

3.1 Table 10

**4 RESULTS AND DISCUSSION** 13 4.1. Performance metrics 14

5 **CONCLUSION** 15

5.1 Conclusion 16

**REFERENCES** 17

CMRCET B. Tech (CSE) Page No i

Object Detection for Visually Impaired using YOLO v7

**List of Tables**

**FIGURE**

**NO. TITLE PAGE NO.**

3.1 Research Comparision Table 11-13

CMRCET B. Tech (CSE) Page No ii

Object Detection for Visually Impaired using YOLO v7

# **ABSTRACT**

Visually impaired individuals face many challenges in their daily lives. One of the biggest challenges is navigating the world around them. This can make it difficult to perform everyday tasks such as walking down the street, shopping, and socializing. It is important to address this problem because it can have a significant impact on their quality of life. Without adequate assistance, they may be at risk of injury, isolation, and depression.

This project proposes a mobile application that uses YOLOv7 for image and face detection, and live camera feed description as audio input to provide detailed audio descriptions, allowing users to interpret visual cues through auditory feedback. The project also features object exploration, color recognition, text reading, currency recognition, etc. The application is designed to help visually impaired users navigate the world around them more easily and independently. The application has the potential to make a significant difference in the lives of visually impaired users. By providing them with real-time audio descriptions of their surroundings and other helpful features, the application can help them to live more independent and fulfilling lives.

CMRCET B. Tech (CSE) Page No iii

Object Detection for Visually Impaired using YOLO v7

# **CHAPTER 1**

**INTRODUCTION**

CMRCET B. Tech (CSE) Page No 1

Object Detection for Visually Impaired using YOLO v7

**CHAPTER 1**

**INTRODUCTION**

* 1. **Problem Statement**

Many visually impaired individuals face significant obstacles in their daily lives, particularly when it comes to navigating and understanding the visual world around them. This can make simple tasks like going shopping, getting around, and socializing very difficult, and can significantly affect their overall quality of life. Without effective assistance, they are at greater risk of accidents, social isolation, and emotional distress Our project proposes a mobile app that uses YOLOv7 for image and face detection, providing detailed audio descriptions of the user's surroundings. It also incorporates features such as object exploration, color recognition, text reading, and currency identification. This app aims to help visually impaired individuals navigate with greater ease and independence, promising to significantly enhance their daily lives.

**1.2. Research Objective**

The primary objective of this research endeavor is to develop a mobile application catering to the unique needs of visually impaired individuals, with a focus on enhancing their navigational capabilities and overall independence in daily activities. Leveraging the YOLOv7 framework for image and face detection, the research aims to implement a robust system that utilizes live camera feed narration as an audio input mechanism. The core goal is to provide detailed and real-time audio descriptions of the users' surroundings, enabling them to interpret visual information through auditory feedback. Furthermore, the research seeks to integrate advanced functionalities, including but not limited to object exploration, color recognition, text reading, and currency identification within the application. By harnessing cutting-edge technology, the objective is to create a comprehensive solution that addresses the multifaceted challenges faced by visually impaired individuals in their daily lives. The research endeavors to optimize the application's usability and effectiveness, ensuring that it becomes a valuable tool for facilitating independent navigation and fostering a heightened quality of life for the target user demographic.

CMRCET B. Tech (CSE) Page No 2

Object Detection for Visually Impaired using YOLO v7

**1.3. Introduction to YOLOv7**

Real-time object detection is an important component in many computer vision systems, including multi-object tracking, autonomous driving, robotics, and medical image analysis. In recent years, real-time object detection development has focused on designing efficient architectures and improving the inference speed of various CPUs, GPUs, and neural processing units (NPUs). YOLOv7 supports both mobile GPU and GPU devices, from the edge to the cloud.

Unlike traditional real-time object detectors that focus on architecture optimization, YOLOv7 introduces a focus on the optimization of the training process. This includes modules and optimization methods designed to improve the accuracy of object detection without increasing the inference cost, a concept known as the "trainable bag-of-freebies".

YOLOv7 is a state-of-the-art real-time object detector that surpasses all known object detectors in both speed and accuracy in the range from 5 FPS to 160 FPS. It has the highest accuracy (56.8% AP) among all known real-time object detectors with 30 FPS or higher on GPU V100. Moreover, YOLOv7 outperforms other object detectors such as YOLOR, YOLOX, Scaled-YOLOv4, YOLOv5, and many others in speed and accuracy. The model is trained on the MS COCO dataset from scratch without using any other datasets or pre-trained weights. Source code for YOLOv7 is available on GitHub.

CMRCET B. Tech (CSE) Page No 3

Object Detection for Visually Impaired using YOLO v7

**CHAPTER 2**

**BACKGROUND WORK**

CMRCET B. Tech (CSE) Page No 4

Object Detection for Visually Impaired using YOLO v7

# **2.1 Research Paper 1**

# **Title:** Object Detection for Visually Impaired People Using SSD Algorithm [1]

**Abstract:**

Visually impaired people are unaware of the danger that they are facing in their life. They may face many challenges while performing their daily activity even in their familiar environments. Vision is the necessary human senses and it plays the important role in human perception about surrounding environment. Hence, there are variety of computer vision products and services which are used in the development of new electronic aids for those blind people. In this paper we designed to provide navigation to those people. It guides the people about the object as well as provides the distance of the object. The algorithm itself calculates the distance of the object. Here it also provides the audio jack to insist them about the object. Here we are using SSD Algorithm for object detection and calculating the distance of the object by using monodepth algorithm.

**Functions**: It helps in Navigation to people, it guides the people, object detection, distance from the person to object, provide audio description.

**Technologies used**: SSD Algorithm, Monodepth algorithm, computer vision, TTS

# **2.2 Research Paper 2**

**Title:** Category-Blind Human Action Recognition: A Practical Recognition System [2]

**Abstract:**

Existing human action recognition systems for 3D sequences obtained from the depth camera are designed to cope with only one action category, either single-person action or two-person interaction, and are difficult to be extended to scenarios where both action categories co-exist. In this paper, we propose the category-blind human recognition method (CHARM) which can recognize a human action without making assumptions of the action category. In our CHARM approach, we represent a human action (either a single-person action or a two-person interaction) class using a co-occurrence of motion primitives. Subsequently, we classify an action instance based on matching its motion primitive co-occurrence patterns to each class representation. The matching task is formulated as maximum clique problems. We conduct extensive evaluations of CHARM using three datasets for single-person actions, two-person interactions, and their mixtures. Experimental results show that CHARM performs favorably when compared with several state-of-the-art single-person action and two-person interaction-based methods without making explicit assumptions of action category.

CMRCET B. Tech (CSE) Page No 5

Object Detection for Visually Impaired using YOLO v7

**Functions** – It does not make assumptions of actions; it subsequently classifies actions based on its motion primitive co-occurrence patterns.

**Technologies used** – Motion Unit models, Action using MSRC12-Guesture, MU combination model, Hybrid actions3D.

# **2.3 Research Paper 3**

# **Title:** IoT-Enabled Automated Object Recognition for the Visually Impaired [3]

# **Abstract:**

Visual impairments have become one of the most predominant problems for the last few decades. To keep doing their daily tasks, vision-impaired people usually seek help from others. An automated common object and currency recognition system can improve the safe movement and transaction activity of visually impaired people.

To develop a system that can identify indoor and outdoor objects, notify the users, and send all information to a remote server repeatedly at a fixed time interval. The proposed system assists the visually impaired to recognize several objects and provides an audio message to aware the user. Four laser sensors are used in the system to detect the objects in the direction of the front, left, right and ground. The proposed system uses Single Shot Detector (SSD) model with MobileNet and Tensorflow-lite to recognize objects along with the currency note in the real-time scenario in both indoor and outdoor environments.

**Functions**: This paper represents an IoT-enabled automated object recognition system that simplifies the mobility problems of the visually impaired in indoor and outdoor environments. The overall accuracy of the proposed system in object detection and recognition is 99.31% and 98.43% respectively. In addition, the proposed system sends all processed data to a remote server through IoT.

**Technologies used**: SSD, MobileNet, TensorFlow etc.

CMRCET B. Tech (CSE) Page No 6

Object Detection for Visually Impaired using YOLO v7

# **2.4 Research Paper 4**

# **Title:** CICERONE- A Real-Time Object Detection for Visually Impaired People [4]

**Abstract:**  Our work provides a solution to people who are suffering from partial blindness due to diseases of the eye or due to accidents. Unlike people who are born blind, they depend a lot on other people for doing their day to day activities. Our project is a boon to such people as our end product; a smart walking stick detects the trained objects that are used daily by the person and intimates the person using audio messages. This project changes the visual world into an audio world by informing the visually impaired of the objects in their environment. These people can use this particular prototype for self -navigating their way. A YOLO (You look only once) algorithm is used in our project. Real-time objects in an image are detected with their names represented on a bounding box and these names are converted to speech signals. The conversion to audio signals is done by using an e-Speak tool which forms Googles Text to Speech (gTTS) system. The prototype consists of several modules. The Raspberry Pi camera module takes the image and transfers it to the Raspberry Pi desktop. Then, real-time object detection is carried out by using YOLO network. The detected image is converted to speech by using the gTTS module and the audio result is provided to the user through a headset. For achieving the required portability, the battery backup is being used.

**StakeHolders:** Partial Blind, people who lost vision in accidents, people who are suffering from eye diseases.

**Functions**: It is a smart walking stick detects the trained objects, it provides audio description

**Technologies used:** YOLO Algorithm, e-spark tool, gTTS Module

# **2.5 Research Paper 5**

**Title:** Acoustic scene classification based on three-dimensional multi-channel feature-correlated deep learning networks [5]

**Abstract:** As an effective approach to perceiving environments, acoustic scene classification (ASC) has received considerable attention in the past few years. Generally, ASC is deemed a challenging task due to subtle differences between various classes of environmental sounds. In this paper, we propose a novel approach to perform accurate classification based on the aggregation of spatial–temporal features extracted from a multi-branch three-dimensional (3D) convolution neural network (CNN) model.

CMRCET B. Tech (CSE) Page No 7

Object Detection for Visually Impaired using YOLO v7

The novelties of this paper are as follows. First, we form multiple frequency-domain representations of signals by fully utilizing expert knowledge on acoustics and discrete wavelet transformations (DWT). Secondly, we propose a novel 3D CNN architecture featuring residual connections and squeeze-and-excitation attentions (3D-SE-ResNet) to effectively capture both long-term and short-term correlations inherent in environmental sounds. Thirdly, an auxiliary supervised branch based on the chromatogram of the original signal is incorporated in the proposed architecture to alleviate overfitting risks by providing supplementary information to the model. The performance of the proposed multi-input multi-feature 3D-CNN architecture is numerically evaluated on a typical large-scale dataset in the 2019 IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events (DCASE 2019) and is shown to obtain noticeable performance gains over the state-of-the-art methods in the literature.

**Function:** Classifies based on the aggregation of spatial-temporal features using CNN.

**Technologies used:** CNN, 3D-CNN etc.

# **2.6 Research Paper 6**

**Title:** Supervised Deep Learning Techniques for Image Description [6]

**Abstract:**

Automatic image description, also known as image captioning, aims to describe the elements included in an image and their relationships. This task involves two research fields: computer vision and natural language processing; thus, it has received much attention in computer science. In this review paper, we follow the Kitchenham review methodology to present the most relevant approaches to image description methodologies based on deep learning. We focused on works using convolutional neural networks (CNN) to extract the characteristics of images and recurrent neural networks (RNN) for automatic sentence generation. As a result, 53 research articles using the encoder-decoder approach were selected, focusing only on supervised learning. The main contributions of this systematic review are: (i) to describe the most relevant image description papers implementing an encoder-decoder approach from 2014 to 2022 and (ii) to determine the main architectures, datasets, and metrics that have been applied to image description.

**Function:** It gives automatic image description using the CNN

**Technologies used**: CNN, RNN, Deep Learning

CMRCET B. Tech (CSE) Page No 8

Object Detection for Visually Impaired using YOLO v7

**CHAPTER 3**

**RESEARCH**

**COMPARISON  
 TABLE**

CMRCET B. Tech (CSE) Page No 9

Object Detection for Visually Impaired using YOLO v7

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.No** | **Authors,**  **Journal Name, YOP** | **Problem**  **Statement** | **Proposed solution** | **Solution** | **Remarks** |
| 1 | Tausif Diwan  G. Anirudh  Jitendra V.  Tembhur ne  Multime dia  Tools and  Applicati ons  (2023)  82:9243–  9275  8 August  2022 | Object  Detection  using YOLO:  challenges,  architectur al  successors,  datasets and  application s | One-stage  Object  Detection  Using  YOLO &  Two-stage  Object  Detection  Using  RCNN,  Fast- RCNN and Faster-RCNN | In this paper, we  explored two stage  object detectors  viz. RCNN, Fast  RCNN, and  Faster-RCNN  along with their  important  applications. We  majorly reviewed  single stage object  detectors  especially YOLOs,  their architectural  advancements,  underlying  pretrained CNN  architectures, and  loss function in  details. | The paper provides a comprehensive review of two-stage and single- stage object detectors. It details the architectural advancements, underlying CNN structures, and loss functions in YOLOs.  The review establishes YOLOs' superiority in terms of detection accuracy and inference time compared to two- stage detectors. The paper concludes by emphasizing ongoing research and suggesting avenues for further exploration in the realm of single-stage object detectors. |
| 2 | Jaskirat Kaur  Williamj eet  Singh  Multime dia  Tools and  Applicati ons  (2022)  81:38297  –38351  23 April2022 | Tools, techniques, datasets and application areas for object detection in an image: a review | This paper provides brief informatio n on various object detection techniques like RCNN, SPP, fast  RCNN,  Faster RCNN, YOLO, SSD,  RatinaNet, LADet etc. | The review highlights challenges in object detection, including improving video- based, 3D, and small object detection, and integrating one- stage and two- stage detectors. It suggests addressing domain adaptation challenges, promoting efficient automatic annotation methods, and emphasizing feature fusion and multi-task learning for better accuracy and speed... | This paper provides a comprehensive review of various datasets such as Pascal VOC,ILSVRC,MS  COCO,OID etc and various object detection techniques such as traditional object detectors like VJ detector,HOG,DPM and deep learning based object detectors like RCNN,SPP,Fast  RCNN,Faster RCNN,YOLO,SSD,RatinaNet,  LADet etc |

CMRCET B. Tech (CSE) Page No 10

Object Detection for Visually Impaired using YOLO v7

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.No** | **Authors,**  **Journal Name, YOP** | **Problem**  **Statement** | **Proposed solution** | **Solution** | **Remarks** |
| 3 | IEEE [2021](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [IEEE](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding)  [Internati](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [onal](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [Conferen](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [ce on](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [Consume](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [r](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [Electroni](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [cs and](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [Compute](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [r](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [Engineer](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [ing](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [(ICCEC](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) [E)](https://ieeexplore.ieee.org/xpl/conhome/9341295/proceeding) | Object Detection using deep learning | To overcome the limitations of traditional object detection methods, this paper advocates for the adoption of deep learning frameworks.  Leveraging convolutional neural networks (CNNs) and exploring modifications and optimization strategies, deep learning models can capture semantic, high- level features, enhancing object detection performance | The solution lies in embracing the advancements in deep learning tools, incorporating innovative architectures, and utilizing effective training strategies to achieve more robust and accurate object detection results. | Object detection using deep learning involves training neural networks to locate and classify objects within images or videos. Techniques like Faster R-CNN, YOLO, and SSD  have gained prominence for their ability to accurately identify objects, enabling applications in areas such as autonomous vehicles, surveillance, and image analysis |
| 4 | Desmond Elliott and Arjen P. de Vries 2020 | Describing Images using Inferred Visual Dependenc y Representat ions | The methodology used in this paper is as follows:The authors first used an object recognition model to identify the objects in an image.They then used a visual dependency parser to identify the relationships between the objects. | The paper introduces a method to train a Visual Dependency Representation (VDR) model with reduced human supervision. It uses an advanced object detector to identify objects in images, generating training data. | The paper introduces a method for generating Visual Dependency Representations (VDR) of images with minimal human supervision. It performs well in a large dataset of people in action but struggles with smaller, diverse datasets. However, transfer learning between datasets shows promise, emphasizing the importance of encoding object spatial relationships for action description tasks. |

CMRCET B. Tech (CSE) Page No 11

Object Detection for Visually Impaired using YOLO v7

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.No** | **Authors,**  **Journal Name, YOP** | **Problem**  **Statement** | **Proposed solution** | **Solution** | **Remarks** |
| 5 | Ashika K.,  Monika N.,  Nivetha M.,  Rohini A 2022 IJRTI |  Volume 7, Issue  9 | ISSN: 2456-  3315 | Object Detection for Visually Impaired People Using YOLO V4  Algorithm | A novel approach for blind navigation employs deep learning, utilizing a single-shot detector (SSD) for obstacle detection and distance estimation.  The prototype, deployed on a smartphone, processes real-time video streams to detect obstacles and communicates results through audio sequences. The SSD MobileNet Architecture achieved over 70% mean average precision, with a usability and efficiency exceeding 65% in evaluations. | The proposed system utilizes deep learning for real-time and accurate obstacle detection, employing the SSD MobileNet architecture for mobile deployment.  Training data is gathered through simulation, and the prototype comprises object detection, depth estimation, and feedback modules, integrating MonoDepth for depth estimation. | This innovative project addresses a crucial need in blind navigation by proposing a deep learning-based obstacle detection system. Utilizing simulation data, it tackles questions of data generation, mobile deployment, and obstacle distance estimation, enhancing the independence and safety of blind individuals through insightful navigation support. |
| 6 | Md. Atikur Rahman  , Muha mmad Sheikh Sadi [Comput](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [er](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [Methods](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [and](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [Program](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [s in](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [Biomedi](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [cine](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [Update](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update) [Volume](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update/vol/1/suppl/C) [1](https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update/vol/1/suppl/C), 2021,  100015 | IoT Enabled Automated Object Recognitio n for the Visually Impaired | The proposed system assists the visually impaired to recognize several objects and provides an audio message to aware the user. Four laser sensors are used in the system to detect the objects in the direction of the front, left, right and ground. The proposed system uses Single Shot Detector (SSD) model with MobileNet and Tensorflow-lite to recognize objects along with the currency note in the real-time scenario in both indoor and outdoor environments. | The proposed system is an IoT- enabled automated object recognition system that simplifies the mobility problems of the visually impaired in indoor and outdoor environments.  The overall accuracy of the proposed system in object detection and recognition is 99.31% and  98.43%  respectively. In addition, the proposed system sends all processed data to a remote server through IoT. | an IoT-enabled automated object recognition system that simplifies the mobility problems of the visually impaired in indoor and outdoor environments. The overall accuracy of the proposed system in object detection and recognition is 99.31% and 98.43%  respectively. In addition, the proposed system sends all processed data to a remote server through IoT. |

CMRCET B. Tech (CSE) Page No 12

Object Detection for Visually Impaired using YOLO v7

**CHAPTER 4**

**RESULTS AND DISCUSSION**

CMRCET B. Tech (CSE) Page No 13

Object Detection for Visually Impaired using YOLO v7

**CHAPTER 4**

**RESULTS AND DISCUSSION**

In the realm of object detection within computer vision, three prominent methodologies, namely Region-based Convolutional Neural Network (R-CNN), Spatial Pyramid Pooling (SPP), and Single Shot Multibox Detector (SSD), showcase distinct strengths and limitations. R-CNN achieves unparalleled accuracy in object localization, providing precise bounding box information and excelling in complex scenes. However, its computational intensity during training and the time-consuming nature of its multi-step process limit its real-time applicability. On the other hand, SPP introduces flexibility by handling variable-sized input images and preserving spatial information at multiple scales. Despite these advantages, the increased model complexity and training resource intensity pose challenges. SSD, with its real-time object detection capabilities and a streamlined single network architecture, demonstrates efficiency in applications like self-driving cars and surveillance. While it sacrifices some accuracy compared to two-stage methods like R-CNN and faces challenges with scale variations, its ability to strike a balance between speed and accuracy makes it a compelling choice, especially in dynamic scenarios.

Each methodology presents a trade-off between accuracy and efficiency, catering to specific use cases. R-CNN prioritizes accuracy but at the expense of computational resources, while SPP enhances flexibility but introduces complexity. SSD, excelling in real-time scenarios, faces accuracy limitations in certain contexts. The selection among these methods hinges on the task's requirements and the nuanced balance between precision and speed. As the field evolves, emerging architectures may offer refined solutions, pushing the boundaries of what's achievable in computer vision and object detection.

CMRCET B. Tech (CSE) Page No 14

Object Detection for Visually Impaired using YOLO v7

CHAPTER 5

**CONCLUSION**

CMRCET B. Tech (CSE) Page No 15

Object Detection for Visually Impaired using YOLO v7

**CHAPTER 5**

**CONCLUSION**

Object detection is a complex field with a range of approaches. Region-based Convolutional Neural Network (R-CNN), Spatial Pyramid Pooling (SPP), and Single Shot Multibox Detector (SSD) are three popular methods. Each has strengths and limitations that make them suited to different applications. R-CNN is highly accurate but computationally demanding and multi-step, making it unsuitable for real-time use. SPP can handle input images of varying sizes and preserve spatial information, but it is complex and computationally expensive. SSD is a good option for real-time use, with a streamlined architecture and a good balance of speed and accuracy.

Choosing between these methods depends on the specific needs of a task, taking into account factors such as precision, efficiency, and adaptability. As technology advances, these methods will continue to evolve, with new approaches emerging to shape the future of computer vision applications.

CMRCET B. Tech (CSE) Page No 16

Object Detection for Visually Impaired using YOLO v7

**REFERENCES**

CMRCET B. Tech (CSE) Page No 17

Object Detection for Visually Impaired using YOLO v7

## REFERENCES

<https://ieeexplore.ieee.org/document/9262344> [1]

[https://openaccess.thecvf.com/content\_iccv\_2015/papers/Li\_CategoryBlind\_Human\_Action\_ICCV\_2015\_paper.pdf](https://openaccess.thecvf.com/content_iccv_2015/papers/Li_CategoryBlind_Human_Action_ICCV_2015_paper.pdf%20)  [2]

<https://www.sciencedirect.com/science/article/pii/S2666990021000148?via%3Dihub> [3]

<https://iopscience.iop.org/article/10.1088/1757-899X/1085/1/012006/meta> [4]

[https://www.nature.com/articles/s41598-022-17863-z](https://www.nature.com/articles/s41598-022-17863-z%205)

<https://www.mdpi.com/1099-4300/25/4/553> [6]

<https://towardsdatascience.com/understanding-sppnet-for-object-detection-and-classification-682d6d2bdfb>

<https://towardsdatascience.com/understanding-and-implementing-faster-r-cnn-a-step-by-step-guide-11acfff216b0>

<https://www.researchgate.net/figure/A-CNN-with-a-spatial-pyramid-pooling-layer_fig3_342288794>

<https://towardsdatascience.com/implementing-ssd-in-keras-part-i-network-structure-da3323f11cff>

<https://www.researchgate.net/post/How-to-implement-Spatial-Pyramid-Pooling-layer-in-CNN>

<https://medium.com/analytics-vidhya/review-spatial-pyramid-pooling-1406-4729-bfc142988dd2#:~:text=Spatial%20pyramid%20pooling%20is%20applied,proposal%20to%20generate%20final%20predictions>.

<https://www.baeldung.com/cs/ssd>

<https://medium.com/@smallfishbigsea/understand-ssd-and-implement-your-own-caa3232cd6ad>

<https://link.springer.com/chapter/10.1007/978-3-319-46448-0_2>

<https://towardsdatascience.com/step-by-step-r-cnn-implementation-from-scratch-in-python-e97101ccde55>

CMRCET B. Tech (CSE) Page No 18