**J**ournal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 1: 2025
ISSN : **1906-9685**

# Deep Learning-Based YOLO Models for the Detection of People with Disabilities

**Mr. K. Uday Kiran[1], Dasari chenchu Mahendra[2]**
**#1 Assistant Professor Department of Master of Computer Applications**
**#2 Pursuing M.C.A QIS COLLEGE OF ENGINEERING & TECHNOLOGY**
**Vengamukkapalem(V),Ongole, Prakasam dist., Andhra Pradesh- 523272**

## ABSTRACT

A deep learning-based approach for detecting people with disabilities is crucial for enhancing accessibility and promoting inclusivity in various environments. By utilizing advanced object detection algorithms, including FastRCNN, FasterRCNN, and several YOLO models (v5s, v7-tiny, v8, v5x6, and v9), a robust framework is developed for accurate identification and recognition. These models leverage state-of-the-art architectures to improve detection capabilities, focusing on real-time performance and precision. The use of FasterRCNN alongside the latest versions of YOLO enables comprehensive analysis and detection, catering to diverse scenarios and ensuring reliable outcomes. The YOLO family of models is particularly effective in maintaining high accuracy while processing images rapidly, making it suitable for deployment in dynamic environments. A user-friendly front end will be developed using the Flask framework, incorporating authentication features for secure access. This implementation aims to facilitate the monitoring and support of individuals with disabilities, ultimately contributing to a more inclusive society by enabling better resource allocation and informed decision-making in accessibility initiatives.

## INTRODUCTION:

It is challenging for machines to distinguish and classify several things in an image. Within the realm of computer vision, object detection pertains to the procedure of locating and identifying an object present in an image or video. However, in recent years, there has been a notable effort in object detection. The fundamental components of object detection encompass feature extraction and processing, and object classification. Numerous techniques have been used including feature coding, feature aggregation, bottom feature extraction, and feature classification, object detection produced satisfactory results among all these techniques, feature extraction play a significant role in object detection and process recognition. Object detection plays a vital role in a wide range of applications, including but not limited to surveillance, cancer diagnosis, vehicle detection, and identification of objects in underwater environments. A variety of ways have been used to detect object accurately and effectively in different applications. However, these proposed methods continue to face challenges related to imprecision and ineffectiveness. In contrast, machine learning and deep neural network techniques prove to be more proficient in rectifying object detection shortcomings and mitigating these concerns.

**Objective:**

The objective is to develop a robust deep learning-based framework for

detecting people with disabilities using advanced object detection algorithms, including FastRCNN, FasterRCNN, and various YOLO models (v5s, v7-tiny, v8, v5x6, and v9). This framework aims to achieve high accuracy and real-time performance in diverse environments, facilitating accurate identification and recognition. Additionally, the project will create a user-friendly front end using the Flask framework, featuring secure authentication for users. Ultimately, the goal is to enhance accessibility, promote inclusivity, and support resource allocation for individuals with disabilities, fostering a more equitable society.

**Problem Statement:**

The lack of effective identification systems for people with disabilities leads to significant barriers in accessibility, limiting their participation in various social, economic, and educational environments due to insufficient recognition and support. Traditional methods of detecting individuals with disabilities are often manual, timeconsuming, and prone to errors, leading to inefficiencies in resource allocation and hindering timely interventions that could improve their quality of life.

Individuals with disabilities, including those with mobility, sensory, and cognitive impairments, are disproportionately affected, facing challenges in accessing essential services, navigating public spaces, and participating in community activities.

This lack of support contributes to social isolation, reduced independence, and negative mental health outcomes, further exacerbating the difficulties faced by individuals with disabilities in their daily lives.

- To address these issues, we aim to implement a deep learning-based framework utilizing advanced object detection algorithms, ensuring accurate detection and fostering an inclusive environment through improved accessibility solutions.
- The proposed system aims to develop an advanced deep learning-based framework for detecting people with disabilities, leveraging cutting-edge object detection algorithms such as FastRCNN, FasterRCNN, and various YOLO models, including YOLOv5s, YOLOv7-tiny, YOLOv8, YOLOv5x6, and YOLOv9. This framework is designed to provide accurate identification and recognition of individuals in real-time, ensuring effective monitoring and support. By utilizing the strengths of FasterRCNN and the YOLO family of models, the system will deliver reliable detection capabilities across diverse environments and scenarios. To enhance user experience, a front end will be developed using the Flask framework, which will include secure authentication features for safe access. This comprehensive approach not only facilitates the efficient identification of individuals with disabilities but also promotes inclusivity and accessibility, enabling better resource allocation and informed decision-making to improve the quality of life for those affected.
- Utilizing advanced YOLO models ensures high-speed, real-time detection capabilities, making it ideal for dynamic environments where timely identification is critical.
- Across Scenarios: By leveraging multiple object detection algorithms, including FasterRCNN and various YOLO versions, the proposed system achieves reliable performance across a wide range of conditions and environments.
- The implementation of a Flask-based front end with secure

authentication enhances user experience and ensures safe access, promoting wider adoption and usability.

- The advanced architectures employed in the proposed system provide improved accuracy in identifying individuals with disabilities, enabling better monitoring and support.
- By effectively detecting people with disabilities, the proposed system promotes inclusivity, ensuring that resources are allocated efficiently and improving the overall quality of life for affected individuals.

Literature Survey:

1. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS: https://www.mdpi.com/2504-4990/5/4/83

ABSTRACT: YOLO has become a central real-time object detection system for robotics, driverless cars, and video monitoring applications. We present a comprehensive analysis of YOLO's evolution, examining the innovations and contributions in each iteration from the original YOLO up to YOLOv8, YOLO-NAS, and YOLO with transformers. We start by describing the standard metrics and postprocessing; then, we discuss the major changes in network architecture and training tricks for each model. Finally, we summarize the essential lessons from YOLO's development and provide a perspective on its future, highlighting potential research directions to enhance real-time object detection systems.

2. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors: https://openaccess.thecvf.com/content/CVPR2023/html/Wang_YOLOv7_Trainable_Bag-ofFreebies_Sets_New_State-of-the-Art_for_Real-Time_Object_Detectors_CVPR_2023_paper.html

ABSTRACT: Real-time object detection is one of the most important research topics in computer vision. As new approaches regarding architecture optimization and training optimization are continually being developed, we have found two research topics that have spawned when dealing with these latest state-of-the-art methods. To address the topics, we propose a trainable bag-offreebies oriented solution. We combine the flexible and efficient training tools with the proposed architecture and the compound scaling method. YOLOv7 surpasses all known object detectors in both speed and accuracy in the range from 5 FPS to 120 FPS and has the highest accuracy 56.8% AP among all known realtime object detectors with 30 FPS or higher on GPU V100.

3. RTF-RCNN: An Architecture for Real-Time Tomato Plant Leaf Diseases Detection in Video Streaming Using Faster-RCNN: https://www.mdpi.com/2306-5354/9/10/565

ABSTRACT: In today's era, vegetables are considered a very important part of many foods. Even though every individual can harvest their vegetables in the home kitchen garden, in vegetable crops, Tomatoes are the most popular and can be used normally in every kind of food item. Tomato plants get affected by various diseases during their growing season, like many other crops. Normally, in tomato plants, 40–60% may be damaged due to leaf diseases in the field if the cultivators do not focus on control measures. In tomato production, these diseases can bring a great loss. Therefore, a proper mechanism is needed for the detection of these problems. Different techniques were proposed by researchers for detecting these plant diseases and these mechanisms are vector machines, artificial neural networks, and Convolutional Neural Network (CNN) models. In earlier times, a technique was used for detecting diseases called the benchmark feature extraction technique. In this area of study for

detecting tomato plant diseases, another model was proposed, which was known as the real-time faster region convolutional neural network (RTFRCNN) model, using both images and real-time video streaming. For the RTF-RCNN, we used different parameters like precision, accuracy, and recall while comparing them with the Alex net and CNN models. Hence the final result shows that the accuracy of the proposed RTF-RCNN is 97.42%, which is higher than the rate of the Alex net and CNN models, which were respectively 96.32% and 92.21%.

1. PP-YOLOE: An evolved version of YOLO:

https://arxiv.org/abs/2203.16250

ABSTRACT: In this report, we present PP-YOLOE, an industrial state-of-the-art object detector with high performance and friendly deployment. We optimize on the basis of the previous PPYOLOv2, using anchor-free paradigm, more powerful backbone and neck equipped with CSPRepResStage, ET-head and dynamic label assignment algorithm TAL. We provide s/m/l/x models for different practice scenarios. As a result, PP-YOLOE-l achieves 51.4 mAP on COCO test-dev and 78.1 FPS on Tesla V100, yielding a remarkable improvement of (+1.9 AP, +13.35% speed up) and (+1.3 AP, +24.96% speed up), compared to the previous state-of-the-art industrial models PP-YOLOv2 and YOLOX respectively. Further, PP-YOLOE inference speed achieves 149.2 FPS with TensorRT and FP16-precision. We also conduct extensive experiments to verify the effectiveness of our designs.

1. A Two-Stage Industrial Defect Detection Framework Based on Improved-YOLOv5 and Optimized-Inception-ResnetV2 Models:

https://www.mdpi.com/2076-3417/12/2/834

ABSTRACT: Aiming to address the currently low accuracy of domestic industrial defect detection, this paper proposes a Two-Stage Industrial Defect Detection Framework based on Improved-YOLOv5 and Optimized-Inception-ResnetV2, which completes positioning and classification tasks through two specific models. In order to make the first-stage recognition more effective at locating insignificant small defects with high similarity on the steel surface, we improve YOLOv5 from the backbone network, the feature scales of the feature fusion layer, and the multiscale detection layer. In order to enable second-stage recognition to better extract defect features and achieve accurate classification, we embed the convolutional block attention module (CBAM) attention mechanism module into the Inception-ResnetV2 model, then optimize the network architecture and loss function of the accurate model. Based on the Pascal Visual Object Classes 2007 (VOC2007) dataset, the public dataset NEU-DET, and the optimized dataset Enriched-NEU-DET, we conducted multiple sets of comparative experiments on the ImprovedYOLOv5 and Inception-ResnetV2. The testing results show that the improvement is obvious. In order to verify the superiority and adaptability of the two-stage framework, we first test based on the Enriched-NEU-DET dataset, and further use AUBO-i5 robot, Intel RealSense D435 camera, and other industrial steel equipment to build actual industrial scenes. In experiments, a two-stage framework achieves the best performance of 83.3% mean average precision (mAP), evaluated on the Enriched-NEU-DET dataset, and 91.0% on our built industrial defect environment.

System Architecture:


The proposed system is divided into several functional modules, each responsible for a specific part of the detection process. These modules work in coordination to ensure accurate, real-time detection of individuals with disabilities while maintaining system efficiency and usability.

You Only Look Once (YOLO) is one of the most popular model architectures and object detection algorithms. It uses one of the best neural network architectures to produce high accuracy and overall processing speed, which is the main reason for its popularity. If we search Google for object detection algorithms, the first result will be related to the YOLO model.

**Challenges in Object Detection:**

In object detection, the bounding boxes are always rectangular. As a result, if the object contains the curvature part, it does not help determine its shape. In order to find precisely the shape of the object, we should use some of the image segmentation techniques.

Some non-neural methods may not detect objects with high accuracy or may produce a large number of false-positive detections. Although neural network methods are more accurate, there are some drawbacks. For example, they require a large amount of annotated data for training. Training is often expensive in time and space and, as a result, prolonged on standard computers.

In order to solve these challenges, we can use the YOLO algorithm. Thanks to the transfer learning capabilities, we would be able to use already pre-trained models or spend some time fine-tuning models with our data. Furthermore, the YOLO algorithm is one of the most popular methods for performing object detection in real-time because it achieves high accuracy on most real-time processing tasks while maintaining a reasonable speed and frames per second, even on devices accessible to almost everyone.

**Applications:**

YOLO has a wide range of real-world applications because it solves one of the most common computer vision problems, object detection. Some of them are:

- Autonomous driving – YOLO can be used in autonomous vehicles for detecting objects such as cars, people, traffic signs, and similar
- Security – for example, detecting people in restricted areas
- Manufactory – detecting anomalies in production
- Sport – tracking players and many others

How Does YOLO Work?

Let’s say that we have an image with two bounding boxes representing a cat and dog. The first step that YOLO does is dividing the image into a grid. For example, a 3×3 grid as below:

With the existence of a grid, it’s possible to detect one object per grid cell instead of one object per image. For each grid cell, we can encode a vector that will describe the cell. For instance, the first cell from the

top-left doesn't have any object, and we describe it as:

$$C_{1,1} = (P_c, B_x, B_y, B_w, B_h, C_1, C_2) = ($$

(1)

where is the probability of the object class, and are coordinates of the center of the bounding box, relative to the cell, and are width and height of the bounding box relative to the whole image, and are 0 or 1 depending on which class represents the bounding box ( for cat and for dog). Vector consists of symbols because if the first component is equal to zero, then the rest of the components can have random numbers are they are not taken into consideration.

**IMPLEMENTATION:**

**EXECUTION PROCEDURE**

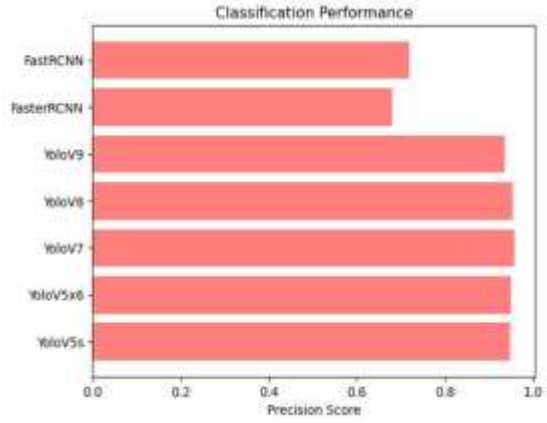
**Tables:**

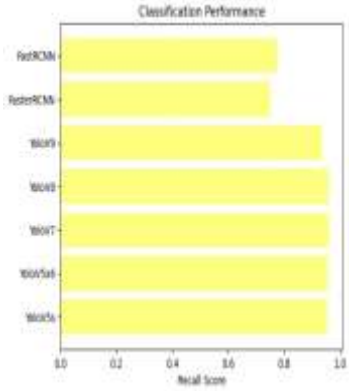
**Performance Evaluation – Detection**

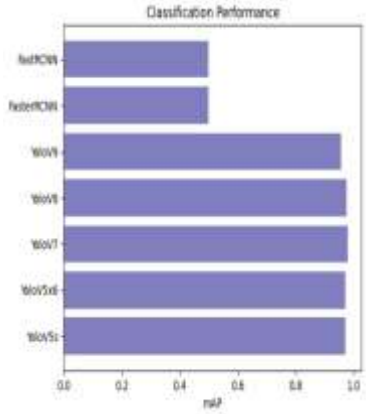
| ML Model | Precision | Recall | mAP |
|---|---|---|---|
| YoloV5s | 0.948 | 0.955 | 0.970 |
| Extension YoloV5s6 | 0.950 | 0.958 | 0.970 |
| YoloV7 | 0.958 | 0.962 | 0.970 |
| YoloV8 | 0.955 | 0.961 | 0.976 |
| Extension YoloV9 | 0.936 | 0.936 | 0.957 |
| Faster RCNN | 0.681 | 0.750 | 0.498 |
| Fast RCNN | 0.720 | 0.774 | 0.498 |

**Graphs:**

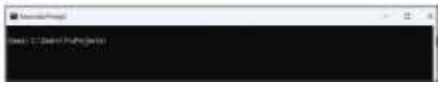
**Comparison Graphs**




**Screens:**

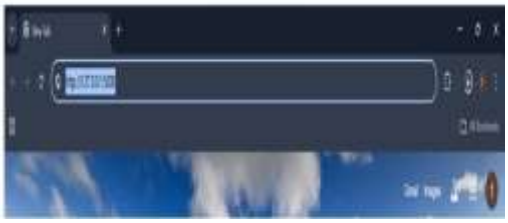


Step -4



Step - 7

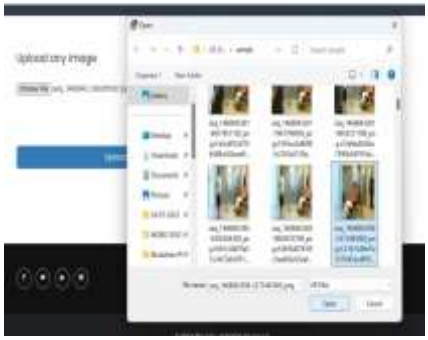

Step -5



User Dashboard


Result of Step – 9
Test case 1


Step – 9
test case 1



**Result of Step – 9**
**Test case 3**

## CONCLUSION

In conclusion, the development of a deep learning-based framework for detecting individuals with disabilities marks a significant step toward enhancing accessibility and fostering inclusivity in society. By integrating advanced object detection algorithms such as FastRCNN, FasterRCNN, and various YOLO models, this system is capable of providing accurate and realtime identification of individuals based on their unique needs. The comprehensive approach ensures that diverse environments and scenarios are effectively addressed, promoting better resource allocation and informed decision-making regarding accessibility initiatives. The userfriendly front end, built using the Flask framework, further enhances the system's usability by incorporating secure authentication features, ensuring safe access for users. Ultimately, this innovative solution not only facilitates improved monitoring and support for individuals with disabilities but also contributes to a more inclusive society. By enabling better understanding and recognition of the challenges faced by differently-abled individuals, this framework aims to empower stakeholders, including policymakers and caregivers, to implement strategies that foster inclusivity and address discrimination. Through the application of cutting-edge technology and a commitment to social responsibility, the project aspires to significantly improve the quality of life for individuals with disabilities, paving the way for a more equitable future.

**Future Scope:**

The future scope of this deep learning-based framework for detecting individuals with disabilities is promising. Enhancements may include integrating additional algorithms and refining existing models to boost accuracy and efficiency. Expanding the system to recognize a wider range of assistive devices will improve its utility. Incorporating real-time data analytics can facilitate personalized support for users, while collaboration with healthcare providers may lead to better resource allocation and informed decision-making. Furthermore, exploring edge computing solutions will enable deployment in various settings, ensuring seamless accessibility. Ultimately, continued advancements in technology will enhance the framework's adaptability and effectiveness in promoting inclusivity and accessibility.

**REFERENCES**

[1] J. Terven and D. Cordova-Esparza, ‘‘A comprehensive review of YOLO: From YOLOv1 and beyond,’’ 2023, arXiv:2304.00501.

[2] C.-Y. Wang, A. Bochkovskiy, and H.-Y.-M. Liao, ‘‘YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,’’ in Proc. IEEE/CVF Conf. Comput. Vis.
Pattern Recognit. (CVPR), Vancouver, BC, Canada, Jun. 2023, pp. 7464–7475.

[3] M. Alruwaili, M. H. Siddiqi, A. Khan, M. Azad, A. Khan, and S. Alanazi, ‘‘RTF-RCNN: An architecture for real-time tomato plant leaf diseases detection in video streaming using fasterRCNN,’’ Bioengineering, vol. 9, no. 10, p. 565, Oct. 2022.

[4] S. Xu, X. Wang, W. Lv, Q. Chang, C. Cui, K. Deng, G. Wang, Q. Dang, S. Wei, Y. Du, and B. Lai, ‘‘PP-YOLOE: An evolved version of YOLO,’’ 2022, arXiv:2203.16250.

[5] Z. Li, X. Tian, X. Liu, Y. Liu, and X. Shi, ‘‘A two-stage indus?trial defect detection framework based on improved-YOLOv5 and Optimized-Inception-ResnetV2 models,’’ Appl.
Sci., vol. 12, no. 2, p. 834, Jan. 2022.

[6] X. Chen, K. Kundu, Y. Zhu, A. G. Berneshawi, H. Ma, S. Fidler, and R. Urtasun, ‘‘3D object proposals for accurate object class detection,’’ in Proc. Adv. Neural Inf. Process. Syst., C. Cortes,
N. D. Lawrence, D. D. Lee, M. Sugiyama, R. Garnett, Eds. New York, NY, USA: Curran Associates, 2015, pp. 424–432.

[7] H. Bilen and A. Vedaldi, ‘‘Weakly supervised deep detection networks,’’ in Proc. IEEE Conf.
Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 2846–2854. [8] S. Ren, K. He, R. Girshick, and J. Sun, ‘‘Faster R-CNN: Towards real-time object detection with region proposal networks,’’ in Proc. Adv. Neural Inf. Process. Syst., 2015, pp. 91–99. [9] X. Chen, S. Xiang, C.-L. Liu, and C.-H. Pan, ‘‘Vehicle detection in satellite images by hybrid deep convolutional neural networks,’’ IEEE Geosci. Remote Sens. Lett., vol. 11, no. 10, pp.
1797–1801, Oct. 2014.

[10] A. Mukhtar, M. J. Cree, J. B. Scott, and L. Streeter, ‘‘Mobility aids detection using convolution neural network (CNN),’’ in Proc. Int. Conf. Image Vis. Comput. New Zealand (IVCNZ), Auckland, New Zealand, Nov. 2018, pp. 1–5.

[11] A. Vasquez, M. Kollmitz, A. Eitel, and W. Burgard, ‘‘Deep detection of people and their mobility aids for a hospital robot,’’ in Proc. Eur. Conf. Mobile Robots (ECMR), Paris, France, Sep. 2017, pp. 1–7.

[12] M. Kollmitz, A. Eitel, A. Vasquez, and W. Burgard, ‘‘Deep 3D perception of people and their mobility aids,’’ Robot. Auto. Syst., vol. 114, pp. 29–40, Apr. 2019.

[13] T.Ahmad, Y. Ma, M. Yahya, B. Ahmad, S. Nazir, and A. U. Haq, ‘‘Object detection through modified YOLO neural network,’’ Sci. Program., vol. 2020, pp. 1–10, Jun. 2020.

[14] H. Law and J. Deng, ‘‘CornerNet: Detecting objects as paired keypoints,’’ in Proc. Eur.
Conf. Comput. Vis. (ECCV), Munich, Germany, 2018, pp. 734–750.

[15] Y. Liu, P. Sun, N. Wergeles, and Y. Shang, ‘‘A survey and performance evaluation of deep learning methods for small object detection,’’ Expert Syst. Appl., vol. 172, Jun. 2021, Art. no.
114602.

[16] J. Yan, Z. Lei, L. Wen, and S. Z. Li, ‘‘The fastest deformable part model for object detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Columbus, OH, USA, Jun.
2014, pp. 2497–2504.

[17] Y. Zheng, C. Zhu, K. Luu, C. Bhagavatula, T. H. N. Le, and M. Savvides, ‘‘Towards a deep learning

framework for unconstrained face detection,'' in Proc. IEEE 8th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS), Sep. 2016, pp. 1–8.

[18] R. Girshick, J. Donahue, T. Darrell, and J. Malik, ''Rich feature hierarchies for accurate object detection and semantic segmentation,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 580–587.

Authors:

**Authors:**

**Mr. K. Uday Kiran**
Mr. K. Uday Kiran is an Assistant Professor in the Department of Master of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. He earned his Master of Computer Applications (MCA) from Bapatla Engineering College, Bapatla. His research interests include Machine Learning,Programming Languages. He is committed to advancing research and fostering innovation while mentoring students to excel in both academic and professional pursuits.

**Dasari chenchu Mahendra[2]**is an MCA Scholar, Dept. of MCA, In QIS College of Engineering & Technology, Ongole. His areas of interest are Machine Learning, Deep Learning.