Real-time Object Detection and Voice Labeling for Enhanced Accessibilityand Visual Interaction

Matta Swathi[1], Ramala Supraja[2*], Malavathu Lakshmi Prasanna[3], Shaik Sameer[4], Guntaka RamaKrishna Reddy[5]

[1,2,3,4,5]Lakireddy Bali Reddy College of Engineering, Mylavaram, India

*supraja123.ramala@gmail.com

Abstract: This work introduces a new approach to real-time object recognition using YOLO Version 7, an advanced system capable of real-time object detection in images, videos, as well as live webcam feeds. Unlike traditional methods, this system verbally discusses everything it finds, including the object's name and the accuracy and confidence levels of the algorithm. Apart from enhancing accessibility, computers may also be leveraged to develop educational and engaging resources. Using the MS COCO dataset and a pre-trained model, YOLO Version 7 ensures accurate and speedy object recognition, even for sma ll objects. By using the speed and precision of the system, the initiative aims to make information less intimidating and engaging, particularly for individuals with visual impairments. The dataset ensures comprehensive evaluations with 118,287 training shots, 5,000 validating images, and 20,288 assessment images spanning 80 object classes. The following are the advantages of the proposed method: speed, accuracy, increased visual interactions, faster and less interference, flexibility in all situations, accurate and quick item recognition, and improved handling of small objects. The solution gathers data from several sources, including cameras and picture/video files, and recognizes objects using the YOLO Version 7 algorithm.

Keywords:YOLO Version 7, Voice Labeling, Neural Network, Visual Interaction, Small Object Handling, Pre-trained Model, Versatility, Fast Interference,.

1 Introduction

Object identification is a subfield of computer vision that locates excitatory objects in pictures with a backdrop by using an incentive strategy. Object detection [1] risks are eliminated by erecting firmly closed boxes around such things and modifying the proper object range for each closed box." Similarly to other computer vision tasks, object detection is a complicated problem that may be accomplished by intensive learning [2]. One of the main issues with object detectionis the possibility of fluctuations in the number of objects in the frontal [3]. But to understand how they function, let's first look at how object detection prohibits the concept by presuming that there is in fact only one item per picture. If every image has just one item, it is simple to locate an enclosed box [4] and handle the object. As the closed box contains four integers, finding its location may be readily expressed as a return issue.

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_70

This mainly discusses the need for easily accessible and instantaneous object identification systems [5], which is a significant problem in the domains of vision in computers and artificial intelligence [6]. As technology advances, there is an increasing need for speed and precise object detection in images and movies. Traditional object identification approaches [7] often fall short in terms of both accessibility and speed, especially for individuals with visual impairments. This work presents YOLO Version 7 [8], a state-of-the-art technique known for its remarkable speed and accuracy to real-time object recognition, in order to solve this problem. The objective of the research is to bridge the gap between the evolving needs of modern applications and conventional object detection methods.

1.1 Motivation

The goal of this work is to bridge a sizable gap in the current state of object identification and computer vision systems. Because traditional techniques cannot meet the demands of real-time object detection, their implementation in dynamic and hectic contexts is hindered. Furthermore, the limited functionality of these solutions for visually impaired individuals highlights the pressing need for a more thorough and efficient remedy. The motivation is the will to overcome these challenges and have a good influence on a digital world where speed is not as important as accessibility. The study attempts to enhance real-time object identification by using YOLO Version 7, which is widely known for its exceptional speed and precision. The main objective is to improve user experiences across a range of domains, including applications that are interactive, security for the public, and education. The research envisions a system that can quickly and accurately identify objects, coupled with voice defining for accessibility, and its potential impact on society. This research is motivated by the idea of an era in which technology seamlessly corresponds with the various requirements of users, irrespective of their abilities or the rate at what data is required. The ultimate objective of this study is to enhance the capabilities of object detection systems to make them more responsive, inclusive,and in step with users' evolving expectations in the rapidly advancing area of technological development that we live in.

A notable development in computer vision and artificial intelligence is the YOLO Version 7 effort. Through its extraordinary speed and precision, it overcomes the constraints of existing approaches to provide real-time object identification. As a result, computer vision applications become more successful across a range of industries. To provide an additional inclusive user experience, the effort also integrates voice labeling into object recognition. This innovation opens up new ways for people with visual impairments in particular to engage with digital content and their surroundings. This emphasis on accessibility and speed makes important contributions to the development of assistive technology and computer vision, making it both more advanced and user-friendly. All things considered, the program makes a major contribution to the development of assistive technology and computer vision.

2 Related Work

Previous research has created the theoretical foundation for understanding the advancements in object recognition and computer vision methods and technologies. Early object recognition algorithms often made use of hand-crafted features and traditional machine learning techniques. These methods have problems with sustainability and the demands of real-time processing, despite the fact that they performed effectively in particular circumstances. As a result, the introduction of deep learning changed the paradigm for object detection. Region-based detection was the initial use of convolutional neural networks (CNNs), made possible by pioneering studies such as the R-CNN (Region-based Convolutional Neural Network). However, there were limitations in speed and computational cost with these approaches.

Kang examines the architecture and operation of the YOLO algorithm, an on-demand object detection and segmentation tool, using the COCO data set. Despite its constraints, the algorithm keeps improving and offers higher processing speeds. YOLOv7, an adaptable real-time object recognition technique that outperforms current detectorsboth in terms of speed and accuracy, with a maximum precision of 56.8% AP, presents IR Reasoner, a unique architecture for real-time object detection in thermal infrared images. It deals with problems including small target objects, picture noise, and background clutter. The method enhances feature maps and performs better than baseline models. Proposed a small object recognition technique for traffic signs based on the modified YOLOv7 model. It comprises a brief target identification layer, attention to oneself and convolutional mix modules, omni-dimensional dynamic convolution, and a standardized Gaussian Wasserstein distance. Using the YOLOV7 framework, enhances low-light fog image identification performance by combining photo defogging and enhancement modules. It yields better resultsin assessing neutral indices such as PSNR and SSIM. The robustness of the framework is enhanced and the perception of driving independently in unclear environments is improved by practicing using a real dataset.

Building on these advancements, recent research has focused on increasing the accuracy of object detection systems. One notable work is the ongoing improvement of the YOLO series, of which YOLO Version 7 is the most advanced in terms of real-time object detection. This version keeps YOLO's well-known speed while adding enhancements to accuracy and microscopic item recognition. Currently, object detection research emphasizes the importance of striking an equilibrium between speed, accuracy, and flexibility to assure applicability in a range of scenarios. In the area of accessibility, there is an increasing need to make systems for computer vision more inclusive. A few research have examined how voice interfaces might enhance user experience, particularly for individuals who is impaired.

3 Working Methodology

Data Collection

The type of data that is given into the model is one of the factors that learning algorithms must take into account. Larger data sets are more likely to be understood by an algorithm that uses machine learning which may also produce precise forecasts for data that was not previously known. We frequently need to perform attribute engineering on the data in order to develop more sections and features. Additionally, it's feasible that some data points have no values at all Therefore, in order to make the data really helpful for algorithmic prediction while creating estimations, we must exhaustively modify it.

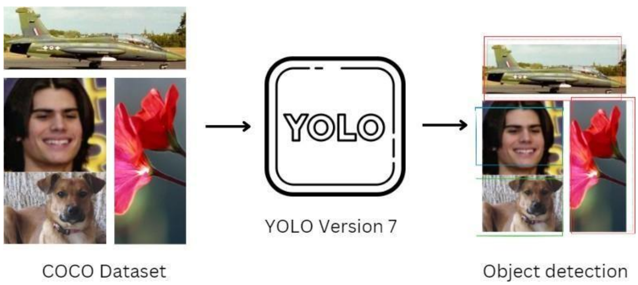

Fig.1 Data collection and object detection using YOLO

For this study, the set of Common Commodities in Context, or COCO 2017 dataset, was used. Because of its large and diverse image collection, this dataset is well respected in the field of artificial intelligence and is perfect for object detection challenges. A download command must be used to access and download the dataset in order to collect data . This ensures that anyone wishing to use the dataset for research or applications can do so without difficulty and makes it easier to locate. The YOLO Version 7 model (Fig. 1) is trained using the vast 118,287 pictures that comprise the training dataset. The training dataset's size plays a crucial role in helping the model identify patterns and features by efficiently capturing the subtleties and differences seen in real-world photos. To evaluate the performance of the trained model and modify its parameters, an alternative validation dataset is employed. This dataset, which contains 5,000 images, provides a representative sample of the whole data set to evaluate the accuracy and generalizability of the model to fresh data. The validation phase aids in adjusting the model's parameters to achieve optimal performance and guarantee that its predictions can generalize to a range of situations beyond the training set.

For comprehensive testing and model performance evaluation, a dedicated testing dataset is employed. This dataset, which includes 20,288 pictures out of 40,670 unique images, offers a reliable standard by which to assess how well the model performs in real-world scenarios. The inclusion of a large testing dataset allows for a complete evaluation of the system's recall, accuracy, and precision and provides valuable information on the system's performance in actual time object detection.

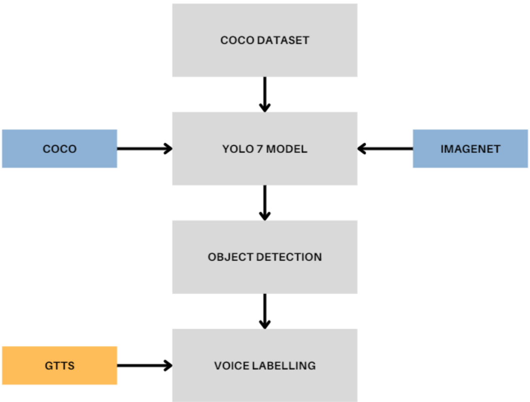

Fig.2 Working Methodology for Object Detection

After then, the YOLO Variant 7 algorithm processes the incoming data in many stages. As part of this, pictures must be prepared, which entails scaling and normalizing the frame of the input or image to facilitate analysis. The deep neural network component of YOLO version 7, that has been pre-trained on the COCO dataset, is then used to examine the processed input data. The network uses convolutional layers to recognize objects in the image by surrounding them in bounding boxes. Every object is assigned a class label, and the system produces confidence ratings that indicate the degree of accuracy of these detections (Fig. 2). In order to ensure accurate results and eliminate false positives, post-processing methods canbe applied to filter and adjust the elements that has spotted. The input picture or video frame with bounding boxes marked around the items that have been detected, class labels, and confidence ratings are all included in the system's final output. This methodology's capacity to handle data in real-time makes it distinctive and appropriate for applications that need accurate and timely object detection.

YOLO Version 7 Algorithm

With its remarkable speed and precision, the YOLO Version 7 (You Only Look Once) algorithm is a cutting-edge development in real-time object identification. Fundamentally, the technique uses a convolutional neural network (CNN), a type of deep neural network architecture [56], to evaluate input data and recognize objects in picture or video frames. The capacity of YOLO to anticipate many objects in a single network pass is one of its primary differentiators, which makes it incredibly effective for real-time applications. Preprocessing images is the first stage in the YOLO Version 7 algorithm. To maintain consistency and prepare the data for efficient analysis, the input picture or video frame is transformed, for example, by shrinking and normalizing the pixel values. Preprocessing is followed by the deep neural network's activation. YOLO Version 7 usually makes use of an intricate CNN architecture, which is frequently based on the most recent architectural advancements in the deep learning space. Because this network has been pre-trained on huge datasets, like the COCO dataset, it can identify a wide variety of objects and patterns.

Internal Object Detection Process

The YOLO Version 7 algorithm's core object identification mechanism consists of many important processes that work together to enable the algorithm to quickly and reliably identify objects in pictures or video frames (as shown in Fig.3). Preprocessing the input picture or video frame is the first step in the process. Standardizing the incoming data and getting it ready for useful analysis depend on this stage. Resizing the image to a specified dimension and normalizing the pixel values are common preparation operations. This guarantees that the data that is input into the deep neural network is consistent.

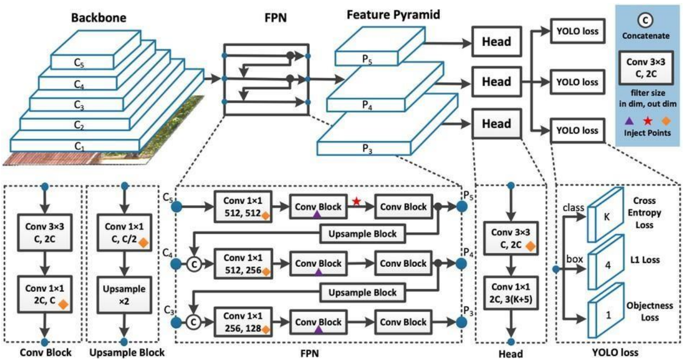

Fig.3 Steps in YOLO Version 7 Algorithm

A complex deep neural network architecture typically a convolutional neural network (CNN) is used by YOLO Version 7. Although the exact design may differ, it usually consists of many convolutional layers followed by fully linked layers. The network learns hierarchical features and patterns indicative of a wide variety of objects by pre-training on large-scale datasets like COCO or ImageNet. A grid is created from the input image, and the task of each grid cell is to anticipate the contents of the objects that lie inside it. YOLO predicts many bounding boxes, class names, and confidence ratings for every cell. The class labels provide the kind of item (person, automobile, etc.), the bounding boxes show the areas the algorithm thinks the object is in, and the confidence ratings show how confident the algorithm is in the correctness of its predictions.

The purpose of image preprocessing in the YOLO Version 7 method is to normalize the input data and improve the effectiveness of the deep neural network's subsequent analysis [66]. A number of methods (as shown in Fig.4) are used at this first stage to guarantee that the input photos or video frames have consistent pixel values, sizes, and formats.

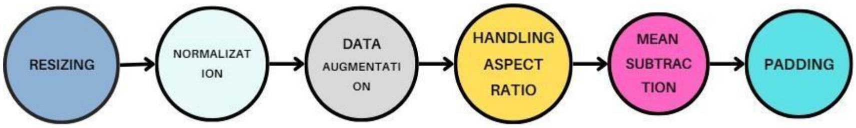

Fig.4 Image Preprocessing techniques.

Resizing is an essential step in the image preprocessing of Yolo Version 7 that aims to standardize the size of input photos or video frames. The method guarantees consistency in the input data by scaling the picture sizes to a constant dimension, which makes it easier for the deep neural network to process and analyze the data. In order to preserve consistency throughout the dataset and enable the model to efficiently learn features and patterns, resizing is necessary. In addition to simplifying the grid division procedure during object recognition, this uniform size allows YOLO Version 7 to forecast across a range of picture sizes and aspectratios quickly and accurately.

A vital step in YOLO Version 7's picture preparation procedure is normalization, which entails scaling pixel values to a defined range. Normalization is commonly used to bring pixel values between 0 and 1. This guarantees that all input photos or video frames have a uniform intensity range. This procedure is essential for encouraging convergence, stabilizing the deep neural network's training, and improving the model's capacity for efficient generalization. YOLO Version 7 reduces dynamic range fluctuations by normalizing pixel values which makes theobject identification algorithm more stable and dependable and ensures consistent performance on a variety of datasets and settings.The process of scaling includes resizing the input photographs at random, mimicking things at various distances, or altering sizes within the scene. This augmentation method helps train the model to recognize items of various sizes. Shearing adds distortions by tilting the picture along its axes, which helps the network identify objects from unusual perspectives. Scaling and shearing both improve the model's adaptability to various spatial arrangements. Color jittering is the process of introducing arbitrary changes to the input pictures' color channels. The model becomes more adaptable to changes in ambient circumstances or camera features thanks to this augmentation strategy, which helps it adjust to differences in color distribution. Color jittering helps to increase generalization by exposing the model to a wider range of color fluctuations. By randomly selecting regions from the input photos, random cropping can simulate situations in which items are partially obscured or appear at different points inside the frame. By encouraging the model to concentrate on various areas of interest, this method improves the model's capacity to manage intricate scenarios with a variety of item placements.

Handling Aspect Ratio

In order to ensure that the algorithm handles photos of different proportions efficiently, handling aspect ratio is a crucial part of image preprocessing in YOLO Version 7. Various techniques are used in this procedure to handle the variation in aspect ratios amongst the incoming photos. Because YOLO Version 7 is made to accommodate these variances, the neural network can effectively learn features and patterns regardless of the form of the source image. The method gains versatility and improves its ability to recognize objects in situations when the scene's form may vary by supporting different aspect ratios. By addressing aspect ratios, the model is made more resilient and operates consistently in a variety of input picture configurations.

Mean Subtraction

In the YOLO Version 7 image preprocessing pipeline, mean subtraction is an optional step that involves subtracting the average pixel value of the whole dataset from each pixel in a picture. During training, this method helps the deep neural network to converge by centering the pixel values around zero. Subtracting the mean helps stabilize the training process by making the model less susceptible to changes in the dataset's overall brightness. Mean subtraction improves the model's capacity to identify object characteristics by focusing on relative pixel intensity differences rather than absolute values. It is especially helpful in situations when constant lighting conditions are necessary for successful object recognition.

Padding

In YOLO Version 7, padding is a complementary technique that involves adding additional pixels to a picture's borders during image preparation. Padding [77] is primarily used to handle size disparity problems and preserve grid division uniformity in the event of future object identification. YOLO Version 7 ensures that every picture input has consistent dimensio ns and prevents information loss while scaling by adding padding to ensure that all input photos have the same size. This stage helps the algorithm identify objects more successfully, especially when precise and dependable outcomes depend on preserving spatial connections and avoiding distortion.

Deep Neural Networks

A key element of the algorithm, the Deep Neural Networks (DNNs) utilized in YOLO Version 7 are essential for the precise and effective identification of objects within picture or video frames. Modern CNN architecture [78], which is well known for its capacity to extract hierarchical features from unprocessed pixel data, is usually utilized by Yolo Version 7. This deep neural network can recognize a wide variety of visual patterns and object classes since it is frequently pre-trained on large-scale datasets like COCO or ImageNet [79]. YOLO Version 7's CNN architecture [80] has been meticulously designed to strike a compromise between computational performance and model complexity. It is made up of several convolutional layers that work together to extract features from the input data at various levels, from low to high. Fully connected layers come after these convolutional layers, which help the model generate predictions about object bounding boxes, class labels, and confidence scores by further refining the learnt features.

An essential component of YOLO Version 7's deep neural network training process is the ImageNet dataset. ImageNet is a large picture database with millions of tagged photos over a wide range of object classifications. To provide the model a deep knowledge of the visual characteristics, patterns, and hierarchies seen in real-world photos, the deep neural network is first pre-trained on ImageNet. From commonplace items and animals to more intricate and specialized entities, ImageNet offers a wide range of object types [81]. The deep neural network can learn a broad variety of characteristics because of its diversity, which also helps it generalize well to various object identification tasks. The model learns to identify complex visual signals and changes that it could meet in real-world circumstances by being exposed to such an extensive dataset during pre-training. Pre-training on ImageNet functions [82] as a type of transfer learning, whereby the skills learned from a large dataset are applied to a job that is more specialized, such as object recognition in YOLO Version 7. This transfer learning technique is essential for assisting the model and reducing the constraints of small task-specific datasets.

One of the key objectives in computer vision is object detection, which is locating and identifying many items inside an image or video frame. You Only Look Once, or YOLO Version 7, is a cutting-edge method of object recognition. Using a deep neural network, the system analyzes the full image in a single pass and predicts bounding boxes, class labels, and confidence ratings all at once. YOLO differs from conventionaltwo-step approaches with its unified approach, which does away with the requirement for region proposal networks and greatly increases efficiency.

YOLO Version 7's adaptability in managing various object types and circumstances is one of its standout features. The system, which was trained on large-scale datasets such as COCO, is capable of identifying a broad variety of things, ranging from commonplace objects to highly specialized ones. Its efficacy in a variety of applications, including surveillance, driverless cars, and interactive user interfaces, stems from its capacity to identify small objects with accuracy and manage changes in scale and aspect ratio. YOLO Version 7 is an effective tool in the field of computer vision because of its ongoing evolution, which is a reflection of a dedication to pushing the limits of object identification.

Post-processing

The output of object identification algorithms must be refined by post-processing techniques, and YOLO Version 7 uses a number of tactics to improve the precision and dependability of its predictions. Non-Maximum Suppression (NMS) is a crucial post-processing technique . NMS is used to filter redundant or overlapping predictions after the first identification of objects with bounding boxes. This ensures that only the most certain and non-overlapping detections are kept in the final output. The identified bounding boxes are sorted from highest to lowest confidence score in accordance with Non-Maximum Suppression. Beginning with the bounding box with the highest degree of confidence, the algorithm iterates over the sorted list. It compares the overlap (IoU, or Intersection over Union) of each succeeding bounding box with the boxes that were previously chosen. The overlapping bounding boxes are deemed redundantif the IoU rises beyond a particular threshold; only the bounding box with the higher confidence score is kept. Until all bounding boxes have been inspected, this process is repeated.

Voice Labeling

The gTTS (Google Text-to-Speech) library is used for voice labeling, which improves the accessibility of object identification results produced by YOLO Version 7. This adds another level of information for users, particularly those who are visually impaired. Text descriptions, including item class names and confidence scores, may be rendered into natural-sounding speech using the gTTS toolkit. Real-time spoken communication of the discovered items is made possible by this text-to-speech translation, making the user experience more inclusive. In the YOLO Version 7, the detected item labels and related data are transformed into a textual format in order to integrate voice labeling with gTTS. The gTTS library receives this text after which it creates an audio file with the spoken representation of the things it has recognized. By integrating with Google's Text-to-Speech API, the gTTS librarygenerates human-sounding voices that narrate the object identification findings in a comprehensible manner.

4 Results

The object identification provides an extensive analysis that demonstrates the real-world object recognition performance of YOLO Version 7 using both image submissions and live streaming through a camera. The objective is to showcase the method's durability and real-time performance across a range of scenarios. First, as part of the evaluation process, static images of various real-world scenarios are submitted (as seen in Fig. 5). The YOLO Version 7 algorithmis tested by analyzing these images and correctly detecting objects inside them. The results show how effectively the system performs when dealing with a variety of object sizes, complex backgrounds, and lighting variations.

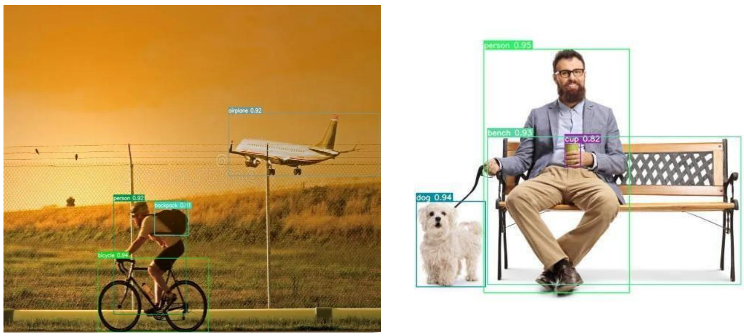

Fig.5 Detection of objects by uploading an Image.

Fig.6 Detection of objects by live streaming.

The next section (as seen in Fig.6) discusses the algorithm's performance in identifying objects in real time utilizing live camera streaming. This evaluation mimics dynamic environments with ever-moving elements. The efficacy of YOLO Version 7 in scenarios such as tracking traffic, monitoring, and user interfaces that are interactive is demon strated by the speed and accuracy with which it can recognize objects in a live video feed. Measures such as frame-per-second (FPS) rates and latency are demonstrated to evaluate the algorithm's efficacy in real-time applications. An evaluation study, testing YOLO Version 7 against current object identification techniques or earlier iterations of the YOLO algorithm, is included in the results section. This comparison sheds light on the algorithm's improvements and demonstrates how fast and accurate it is. Visual aids like bounding box overlays on photos, precision-recall curves, and sample frames from live streaming help to show off the algorithm's capabilities and performance in a variety of real-world applications.

Conclusion

Thus, YOLO Version 7, a cutting-edge object recognition system that uses both actual footage from a camera and image uploads to exhibit astounding capabilities in real-world conditions, is up and running thanks to this project. The speed and accuracy with which the algorithm can identify objects in a range of conditions, taking into consideration variations in scale, lighting, and dynamic environments, serves as an example of its performance. The addition of the gTTS collection for voice labeling improves accessibility even further and increases user engagement—particularly for those who are visually impaired. This comprehensive evaluation demonstrates how much more responsive and precise YOLO Version 7 is, with comparative studies and precision recall data included. This not only pushes the limits of object recognition technology but also broadens its applications to enhance accessibility and interactivity in a variety of sectors, such as traffic or surveillance and instructional materials. The results confirm that YOLO Version 7 is a versatile and powerful tool that may significantly impact applications for computer vision in real-world scenarios.

References

1. Zou, Zhengxia, et al. "Object detection in 20 years: A survey." *Proceedings of the IEEE* (2023).
2. Sinha, Rajat Kumar, Ruchi Pandey, and Rohan Pattnaik. "Deep learning for computer vision tasks: a review." *arXiv preprint arXiv:1804.03928* (2018).
3. Zhao, Zhong-Qiu, et al. "Object detection with deep learning: A review." *IEEE transactions on neuralnetworks and learning systems* 30.11 (2019): 3212-3232.
4. Wang, Yanghao, et al. "Random boxes are open-world object detectors." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023.
5. Amit, Yali, Pedro Felzenszwalb, and Ross Girshick. "Object detection." *Computer Vision: A ReferenceGuide*. Cham: Springer International Publishing, 2021. 875-883.
6. Kumar, DNS Ravi, N. Praveen, Hari Hara P. Kumar, Ganganagunta Srinivas, and M. V. Raju. "Acoustic Feedback Noise Cancellation in Hearing Aids Using Adaptive Filter." International Journal of Integrated Engineering 14, no. 7 (2022): 45-55.
7. Zou, Xinrui. "A review of object detection techniques." *2019 International conference on smart grid and electrical automation (ICSGEA)*. IEEE, 2019.
8. Cao, Lianjun, Xinyu Zheng, and Luming Fang. " The Semantic Segmentation of Standing Tree Images Basedon the Yolo V7 Deep Learning Algorithm." *Electronics* 12.4 (2023): 929.
9. Li, Kunyi, and Lu Cao. "A review of object detection techniques." *2020 5th International Conference on Electromechanical Control Technology and Transportation (ICECTT)*. IEEE, 2020.
10. Avanija, J., G. Sunitha, and K. Reddy Madhavi. "Semantic Similarity based Web Document Clustering Using Hybrid Swarm Intelligence and FuzzyC-Means." Helix 7, no. 5 (2017): 2007-2012.
11. Wang, Yang, Hongyuan Wang, and Zihao Xin. "Efficient detection model of steel strip surface defects based on YOLO-V7." *IEEE Access* 10 (2022): 133936-133944.
12. Yu, Wai, et al. "A novel multimodal interface for improving visually impaired people's web accessibility." *Virtual Reality* 9 (2006): 133-148.
13. Aron, Laudan, and Pamela Loprest. " Disability and the education system." *The future of Children* (2012): 97-122.
14. Halili, Siti Hajar. "Technological advancements in education 4.0." *The Online Journal of Distance Education and e-Learning* 7.1 (2019): 63-69.
15. Prasad, Dilip K., et al. " Challenges in video based object detection in maritime scenario using computer vision." arXiv preprint arXiv:1608.01079 (2016).
16. Kumar, Voruganti Naresh, U. Sivaji, Gunipati Kanishka, B. Rupa Devi, A. Suresh, K. Reddy Madhavi, and Syed Thouheed Ahmed. "A FRAMEWORK FOR TWEET CLASSIFICATION AND ANALYSIS ON SOCIAL MEDIA PLATFORM USING FEDERATED LEARNING." Malaysian Journal of Computer Science (2023): 90-98.
17. Van Swaaningen, René. " Public safety and the management of fear." *Theoretical criminology* 9.3 (2005): 289-305.
18. Coughlan, James M., Huiying Shen, and Brandon Biggs. " Towards accessible audio labeling of 3D objects." *Journal on technology and persons with disabilities:... Annual International Technology and Persons with Disabilities Conference*. Vol. 8. NIH Public Access, 2020.
19. Li, Xin, and Yiliang Shi. " Computer vision imaging based on artificial intelligence." *2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)*. IEEE, 2018.
20. KR, Sarath Chandran. " Real time object identification using deep convolutional neural networks." *2017 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2017.