

# Live Feed Object Detection using YOLO Algorithm with Audio Feedback Mechanism

*by* SUBBA RAO MUPPALLA

---

**Submission date:** 30-Dec-2023 03:06PM (UTC+0530)

**Submission ID:** 2265628327

**File name:** ction\_using\_YOLO\_Algorithm\_with\_Audio\_Feedback\_Mechanism\_1.docx (2.64M)

**Word count:** 3637

**Character count:** 22445

# Live Feed Object Detection using YOLO Algorithm with Audio Feedback Mechanism

2<sup>1st</sup> Swathi Buragadda

Department of Computer Science and  
Engineering  
Lakireddy Bali Reddy College of Engineering  
(Autonomous)  
Mylavaram, India  
buragaddaswathi@gmail.com

2<sup>2nd</sup> Akuthota Kumar sai

Department of Computer Science and  
Engineering  
Lakireddy Bali Reddy College of Engineering  
(Autonomous)  
Mylavaram, India  
akuthotakumarsai@gmail.com

4<sup>3rd</sup> Sathuluri Keerthana

Department of Computer Science and  
Engineering  
Lakireddy Bali Reddy College of Engineering  
(Autonomous)  
Mylavaram, India  
keerthisathuluri715@gmail.com

6<sup>4th</sup> Jagarlamudi Akhil

Department of Computer Science and  
Engineering  
Lakireddy Bali Reddy College of  
Engineering (Autonomous)  
Mylavaram, India  
jagarlamudiakhil369@gmail.com

**Abstract**— In today's data-rich environment, the continuous generation of vast amounts of data necessitates advanced technologies and tools for effective rendering and insight extraction from metadata. Various domains, such as Self-Driving cars, Traffic Monitoring, Parking Occupancy, and Video Surveillance, rely on analyzing video feeds to extract valuable information. This data serves as a crucial resource for training models efficiently. This project introduces an innovative approach to object detection by integrating the YOLO (You Only Look Once) algorithm with a voice feedback system. YOLO's real-time capabilities make it ideal for quick and accurate object identification in images or video streams. Convolutional Neural Networks (CNNs) play a pivotal role in feature extraction, enhancing robust and accurate object detection. The YOLO algorithm splits the given input data into multiple parts using grid methodology and in each part, object will be identified via bound boxing mechanism. Convolutional Neural Networks (CNNs) enhance the system's capability to recognize detailed features. The methodology involves precise prediction of bounding box coordinates and identification of classes, enabling simultaneous real-time identification and localization of objects. The project extends beyond conventional object detection, exploring advanced techniques such as image segmentation, panoptic segmentation, dense pose estimation, and key point detection, enhancing the granularity of object identification. Additionally, the integration of a voice feedback system enhances accessibility, benefiting visually impaired individuals and scenarios with limited visual attention. This research contributes to the advancement of real-time object detection, with potential implications in diverse domains, including assistive technology, surveillance, and human-computer interaction.

**Keywords:** Object Detection, YOLO, Voice Feedback System, Neural Networks, Visual Impairment, Video Analytics, Bounding Boxes.

## I. INTRODUCTION

In our rapidly evolving world, an unprecedented surge in data creation occurs every second. The management and analysis of this colossal volume of information demand cutting-edge technologies and proficient tools to distill meaningful insights from metadata. This dynamic landscape has given rise to the indispensable role of sophisticated systems across various domains, ranging from the realms of self-driving cars to traffic monitoring and video surveillance. The essential task at hand involves the scrutiny of video

feeds, extracting valuable data that serves as a linchpin for the effective training of models.

In the rapidly evolving field of object detection, several models have been developed to address the challenges posed by the unprecedented volume of data generated in today's dynamic world. Each model comes with its unique characteristics, catering to specific requirements in various domains. Here, we briefly explore previous object detection models before delving into the specifics of the YOLO algorithm.

In the landscape of object detection models, the Faster R-CNN (Region-based Convolutional Neural Network) has gained prominence as a two-stage approach. This model incorporates the Region Proposal Network (RPN) to generate region proposals, achieving high accuracy through a region-based methodology. However, its computational complexity poses challenges for real-time applications.

Another notable model is the Single Shot MultiBox Detector (SSD), designed for faster processing speeds in real-time applications. Utilizing default bounding boxes at various scales and aspect ratios in each feature map, SSD strikes a balance between speed and accuracy. Its single-shot, multibox design makes it particularly suitable for scenarios that demand swift and efficient object detection.

In addressing the challenge of class imbalance, RetinaNet has introduced the Focal Loss mechanism. RetinaNet's architecture combines a feature pyramid network with the Focal Loss, allowing for effective detection of objects at different scales. This innovative approach enhances the model's adaptability to diverse scenarios, providing a solution to the class imbalance issue in object detection.

In the realm of object detection, the You Only Look Once (YOLO) algorithm stands out as a pioneering and efficient approach. YOLO adopts a single-shot approach to detect objects with bounding boxes and probabilities of various classes directly from the entire image [6]. In contrast to two-stage detectors, the YOLO algorithm divides the given input data feed into various pieces using grid methodology and, in each piece, object will be identified via bound boxing mechanism.

YOLO's strength lies in its ability to process images and video streams swiftly, making it well-suited for applications requiring rapid and accurate object identification. Its grid-based approach facilitates efficient handling of various object scales and categories, contributing to its versatility in domains such as self-driving cars, traffic monitoring, and video surveillance. In summary, the landscape of object detection models offers a spectrum of approaches, each with its strengths [11]. YOLO stands out for its real-time efficiency and accuracy, making it a compelling choice for applications demanding timely and precise analysis of visual data.

## II. LITERATURE SURVEY

As there is a lot of importance in generative AI research, the computer vision projects, particularly in the aspect of detection of multiple objects in an individual frame, has been involved lot of attention from fellow researchers. This paper, authored by Sanskruti Patel and Atul Patel, presents a comprehensive review of the recent advancements in object detection using Convolutional Neural Networks (CNNs). CNNs, equipped with a series of neural network layers, have demonstrated remarkable outcomes in visual imagery tasks [1]. The survey explores various applications of object detection, encompassing areas such as human-computer interaction, video surveillance, satellite imagery analysis, and transportation systems. It classifies CNN-based object detectors into single-stage or two-stage models and examines well-known architectures like R-CNN, Fast R-CNN, Faster R-CNN, Mask R-CNN, SSD, and YOLO. Additionally, the survey delves into benchmark datasets, detailing their features and significance, and summarizes research endeavors applying object detection models across diverse fields [7]. Given the growing demand for swift and precise object detection systems in applications such as face detection, video feed surveillance, and autonomous driving, this literature review offers a valuable overview of the current landscape and advancements in the field.

Ashwani Kumar and Sonam Srivastava propose a real-time object detection method designed to be adaptable to diverse environments and compatible with any device. The approach employs convolutional neural networks (CNNs) to construct a multi-layered model for classifying objects into predefined classes. By utilizing higher resolution feature maps, the model identifies and labels objects through the analysis of multiple images, potentially sourced from video frames. The system addresses variations in aspect ratio using separate filters and incorporates multi-scale feature maps to ensure robust object detection [3]. Training continues until the error rate is minimized, and the resulting model is tested on sample images. To enhance computational efficiency, the technique integrates the single-shot multi-box detector algorithm and the Faster Region Convolutional Neural Network architecture [9]. Object detection accuracy is assessed using parameters such as the loss function (LP), mean average precision (mAP), and frames per second (FPS). This work contributes to real-time object detection by amalgamating advanced deep learning techniques with image processing methodologies.

In their literature survey, Bin Qasim Ahmad and Pettirsch Arnd explore object detection in images, with a focus on transitioning to video data, particularly in applications like autonomous driving. The study compares various methods, emphasizing Recurrent Neural Networks (RNNs) for object

detection in videos, covering feature-based and box-level methods [2]. The research highlights the importance of incorporating temporal context, with models leveraging multiple frames outperforming those on single frames. Operating on multiple frames simultaneously is recommended for improved mean Average Precision (mAP). To optimize computational speed, the study suggests avoiding excessively deep recurrent units and selectively working on keyframes. Additionally, advocating for different scales enhances detection quality and computational efficiency in video object detection networks.

Ashwani Kumar, Zuopeng Justin Zhang, and Hongbo Lyu conducted a study aimed at improving object detection techniques, particularly for real-time applications across diverse devices and environments. They enhanced the Single Shot Multi-Box Detector (SSD) algorithm, known for its speed, by incorporating depth-wise separable convolution and spatial dividable convolutions into multilayer convolutional neural networks. This improvement boosts classification accuracy without compromising speed [4]. The resulting algorithm utilizes more default boxes and multilayer networks, leading to increased accuracy in object detection. Evaluation metrics, including loss function, frames per second (FPS), mean average precision (mAP), and aspect ratio, validate the algorithm's high accuracy. The study attains noteworthy accuracy rates, exceeding 79.8%, and outperforms previous models in metrics like mAP, aspect ratio, and FPS using Pascal VOC and COCO datasets [10]. The paper highlights the algorithm's utilization of truth boxes for feature map extraction and suggests future research directions to extend the model to micro-object detection.

This research, conducted by K. Vaishnavi and G. Pranay Reddy, focuses on the development of a deep learning-based item recognizer for efficient object identification in images. Utilizing an enhanced Single Shot Multibox Detector (SSD) technique and a multilayer convolution network, the study aims to achieve rapid and accurate recognition of objects in both static and dynamic images [5]. The proposed model demonstrates a high accuracy rate, with over 80% of predictions being correct. By refining the SSD's object detection process through the selection of default boxes with optimal aspect ratios, the research contributes to improving the overall performance of object identification. Furthermore, the study emphasizes the potential impact of object identification technology in automating tasks traditionally performed by humans, akin to the transformative effects of the first Industrial Revolution. In contrast to earlier object recognition methods relying on handcrafted features and imprecise algorithms, this research introduces an end-to-end solution based on deep learning, specifically the SSD technique, which operates swiftly with a single layer of a convolution network [8]. The primary objective is to enhance the accuracy of the SSD method, addressing the limitations of previous object detection systems that depend on additional computer vision methods, resulting in slower and subpar performance. This research thus presents a comprehensive approach to advancing object detection through state-of-the-art deep learning techniques.

## III. EXISTING MODEL

10

### 3.1 REGION-BASED CONVOLUTIONAL NEURAL NETWORK

Region-based Convolutional Neural Network (R-CNN), revolutionized object detection by introducing a two-stage



approach. In the first stage, it employs a selective search algorithm to propose regions likely to contain objects. This algorithm, although effective, is computationally intensive. Subsequently, each proposed region undergoes feature extraction using a pre-trained CNN, transforming the region into a fixed-size feature vector [14]. Finally, these vectors are fed into support vector machines (SVMs) for classification and bounding box regression to precisely localize the detected objects. While accurate, R-CNN's major limitation lies in its sluggishness due to the exhaustive region proposal process.

### 3.2 CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks (CNNs) serve as the fundamental architecture in object detection systems, constituting the cornerstone for feature extraction from images [13]. Within the framework of R-CNN, CNNs assume a pivotal role in deciphering patterns within proposed regions. Engineered to capture spatial hierarchies, these networks enable the model to discern nuanced details and distinctive features of objects, bolstering the overall discriminative capacity of the object detection system [12]. The essence lies in the convolutional layers, where feature learning occurs, amplifying the system's capability to discern and characterize objects with heightened accuracy.

### 3.3 FAST R-CNN

Fast R-CNN addresses the inefficiencies of R-CNN by introducing the Region of Interest (ROI) pooling layer. This layer enables the extraction of fixed-sized feature maps from the proposed regions, eliminating the need for time-consuming computations in multiple stages. The integration of the ROI pooling layer allows for a single forward pass of the CNN for both feature extraction and subsequent classification. By consolidating these steps, Fast R-CNN achieves notable improvements in terms of both accuracy and computational efficiency compared to its predecessor.

### 3.4 MASK R-CNN

Mask R-CNN builds upon the foundation of Faster R-CNN by incorporating the capability of instance segmentation. It extends the bounding box predictions by introducing an additional branch that predicts segmentation masks for each object. The RoIAlign layer ensures precise mapping of regions, addressing misalignment issues common in earlier models. Mask R-CNN excels in tasks requiring detailed segmentation, as it simultaneously predicts object classes, bounding boxes, and pixel-level masks, providing a holistic understanding of object instances within an image.

### 3.5 SINGLE-SHOT DETECTOR

Single-Shot Detector (SSD) diverges from the two-stage paradigm by adopting a one-shot approach to object detection. SSD achieves this through the strategic placement of default boxes at multiple scales on feature maps. These default boxes are used to predict object categories and adjust bounding boxes [15]. Unlike its predecessors, SSD enables end-to-end training, optimizing both accuracy and speed. It excels in real-time applications by efficiently capturing objects of various scales and aspect ratios, making it a robust choice for dynamic environments.

## IV. PROPOSED MODEL

### 4.1 DATASET

The YOLO(You Only Look Once) algorithm can be trained on a variety of datasets depending on the application. There are various datasets such as : COCO(Common Objects in Context), VOC(Visual Object Classes), ImageNet and a lot more. In this project we used COCO(Common Objects in Context) dataset for Live Feed Object Detection. The COCO (Common Objects in Context) dataset encompasses a vast collection of more than 330,000 images.

|               |               |          |                |            |              |              |              |
|---------------|---------------|----------|----------------|------------|--------------|--------------|--------------|
| person        | fire hydrant  | elephant | skis           | wine glass | broccoli     | dining table | toaster      |
| bicycle       | stop sign     | bear     | snowboard      | cup        | carrot       | toilet       | sink         |
| car           | parking meter | zebra    | sports ball    | fork       | hot dog      | tv           | refrigerator |
| motorcycle    | bench         | giraffe  | kite           | knife      | pizza        | laptop       | book         |
| airplane      | bird          | backpack | baseball bat   | spoon      | donut        | mouse        | clock        |
| bus           | cat           | umbrella | baseball glove | bowl       | cake         | remote       | vase         |
| train         | dog           | handbag  | skateboard     | banana     | chair        | keyboard     | scissors     |
| truck         | horse         | tie      | surfboard      | apple      | couch        | cell phone   | teddy bear   |
| boat          | sheep         | suitcase | tennis racket  | sandwich   | potted plant | microwave    | hair drier   |
| traffic light | cow           | frisbee  | bottle         | orange     | bed          | oven         | toothbrush   |

Fig. 1. Classes in COCO Dataset

Each image is meticulously annotated with information about 80 distinct object categories, capturing a wide range of everyday items and scenarios.

```
{
  "info": {
    "description": "COCO 2017 Dataset",
    "licenses": [
      {
        "url": "http://creativecommons.org/licenses/by-nc-sa/2.0/",
      }
    ],
    "images": [
      {
        "license": 1,
      }
    ],
    "annotations": [
      {
        "segmentation": [ ]
      }
    ],
    "categories": [
      {
        "supercategory": "person"
      }
    ]
  }
}
```

Fig. 2. COCO JSON Format

Furthermore, the richness of the dataset extends beyond object annotations; each image is accompanied by five descriptive captions that provide context and detail about the depicted scenes. The computer vision field heavily depends on the COCO dataset which contains common objects in context as a fundamental resource for evaluating and benchmarking the performance of object detection models, including YOLO (You Only Look Once). The COCO dataset utilizes a JSON format to furnish details about each dataset and its associated images.

### 4.2 METHODOLOGY

In this project, at first the COCO (Common Objects in Context) dataset is imported to serve as a comprehensive and diverse collection of images for training and evaluation. Then Image processing techniques are applied to enhance the dataset, followed by augmentation procedures to increase the variety of objects and scenarios in both the training and testing datasets. The heart of the system lies in the training of the YOLO algorithm. During this phase, the model learns to detect and classify objects within the specified classes using the augmented dataset. Training involves optimizing parameters and adjusting weights to refine the algorithm's ability to make accurate predictions.

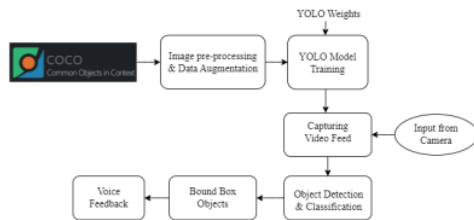


Fig. 3. Proposed model diagram

Convolutional Neural Networks (CNNs) are integral to the YOLO algorithm's feature extraction process, ensuring efficient real-time object detection. Utilizing bounding boxes, YOLO precisely defines the spatial extent of detected objects. CNNs' importance extends to image segmentation tasks, including semantic segmentation, panoptic segmentation, and dense pose estimation, offering detailed information on object boundaries, categories, and poses. CNNs excel at hierarchical feature extraction, automatically discerning features from low-level to abstract, enabling accurate identification of objects at different scales and orientations. Their spatial hierarchies preserve contextual information, facilitating precise object localization, while parameter-sharing reduces learnable parameters, enhancing efficiency in handling visual data. The translation invariance property further supports pattern recognition, making CNNs well-suited for varied object detection positions within the visual field.

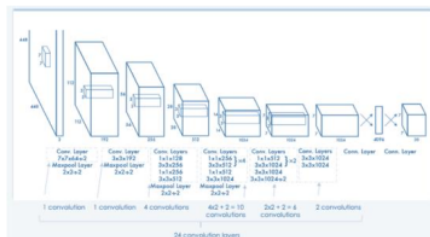


Fig.4. YOLO Algorithm Architecture

In the YOLO algorithm, the input image is partitioned into multiple parts, and each partition will undergo bound boxing mechanism. Anchor boxes, with predetermined shapes and aspect ratios, play a crucial role in this process, serving as references for predicting object dimensions and locations within each grid cell. For every cell, YOLO predicts bounding box coordinates (center's x and y coordinates, width, and height) and assigns probabilities to each class for the objects within the box, ensuring a comprehensive understanding of the detected entities.

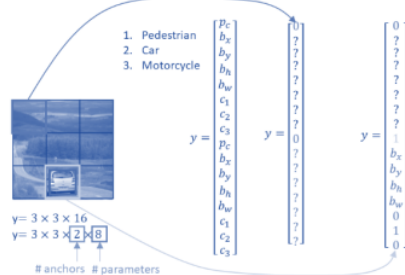


Fig.5. Anchor Boxes in YOLO

Bounding boxes are crucial elements in object detection, providing a structured framework for precisely delineating the spatial boundaries of detected objects within an image. Their primary role lies in localization, allowing the algorithm to pinpoint the exact position of objects by defining coordinates relative to the image frame. This localization, achieved through the top-left and bottom-right corners of the bounding box, ensures accurate identification of multiple objects within the visual field. The use of bounding boxes contributes to the overall precision and accuracy of object detection, enabling the system to make informed predictions about object shapes, sizes, and spatial relationships.

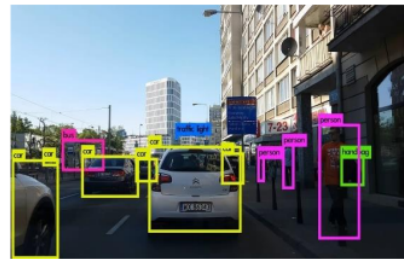


Fig.6. Bounding Boxes in YOLO

Once the YOLO model is trained, it is ready for real-time object detection in videos. A video capturing mechanism is implemented to feed frames into the algorithm sequentially. The YOLO algorithm analyzes each frame, predicting bounding boxes and class probabilities for detected objects. Simultaneously, a voice feedback system is integrated to communicate the identified objects to the user in real-time.

The voice feedback system serves as a crucial component for enhancing user interaction and accessibility. By seamlessly integrating voice communication, the system ensures an inclusive user experience, particularly beneficial for visually impaired individuals or situations where visual attention is limited.

## V. RESULTS

The integration of YOLO's swift and accurate object detection capabilities, coupled with the innovative addition of voice feedback, has yielded promising outcomes. The system demonstrated high accuracy in identifying and localizing objects in diverse scenarios, providing precise bounding box predictions. The voice feedback mechanism proved to be a valuable addition, enhancing user interaction and accessibility, especially for visually impaired individuals. The seamless integration of CNNs within the YOLO framework contributed to robust feature extraction, and it involves in the detailing of the bound boxing dimensions along with the predictions of the objects.

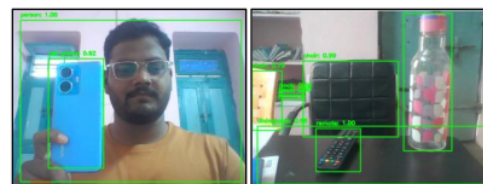


Fig.7. Detection of Objects within Bounding Boxes using YOLO

Here, in the left image the YOLO model predicts accurately that there is a person and phone with an accuracy

of 1.00 and 0.92, in the other image it predicts that there is bottle, remote, book, chair, table with an average accuracy of above 98% respectively along with voice feedback.



Fig.8. Object Detection with Live Video Input Feed

The results indicate the system's potential applicability across various domains, such as assistive technology and surveillance, where quick and accurate object identification is crucial. In this way one can utilize the true potential of YOLO algorithm for detecting the objects not only present in images but also in a live video surveillance.

## VI. CONCLUSION

This project achieves a harmonious synergy of advanced technologies, delivering a system that is both robust and user-centric. The successful implementation of this system involves several key steps, starting with the importation of the COCO dataset a diverse collection of images annotated with object labels. Image processing and augmentation techniques are then applied to enrich both the training and testing datasets, ensuring the robustness and generalization of the model. The YOLO algorithm is trained on the augmented dataset, leveraging its efficiency in real-time object detection. The training process involves the iterative refinement of the model's parameters to optimize its ability to accurately identify and localize objects. Post-training, the system can capture video input, running real-time object detection, and overlaying bounding boxes on detected objects. The real-time nature of the object detection, facilitated by the YOLO algorithm, ensures dynamic responsiveness to changing environments. The project's adaptability and versatility shine through, performing key role in various domains such as surveillance, accessibility for the visually impaired, and interactive installations. The integration of voice feedback adds an extra layer of inclusivity, enhancing the user experience by audibly communicating the identified objects. The introduction of voice feedback elevates the project's functionality, providing an auditory layer that not only enhances user engagement but also addresses scenarios where visual confirmation is challenging.

## REFERENCES

- [1]. Patel, Sanskruti & Patel, Atul. (2020). Object Detection with Convolutional Neural Networks. 10.1007/978-981-15-7106-0\_52.
- [2]. B. Q. Ahmad and A. Pettirsch (2020), "Recurrent Neural Networks for Video Object Detection". <https://doi.org/10.48550/arXiv.2010.15740>.
- [3]. A. Kumar and S. Srivastava, "Object Detection System Based on Convolution Neural Networks Using Single Shot Multi-Box Detector," *Procedia Computer Science*, vol. 171, pp. 2610-2617, 2020, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2020.04.283>.
- [4]. Kumar, A., Zhang, Z.J. & Lyu, H. Object detection in real time based on improved single shot multi-box detector algorithm. *J Wireless Com Network* 2020, 204 (2020). <https://doi.org/10.1186/s13638-020-01826-x>.
- [5]. K. Vaishnavi, G. Pranay Reddy, T. Balaran Reddy, N. Ch. Srimannarayana Iyengar, and Subhani Shaik, "Real-time Object Detection Using Deep Learning," *Journal of Advances in Mathematics and Computer Science*, vol. 38, no. 8, pp. 24-32, 2023, Article no. JAMCS.101284, ISSN: 2456-9968.
- [6]. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779-788).
- [7]. Tejas Vedak , Devanshu Sharma , Vedang Koli, 2021, Object Detection based Attendance System, *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)* Volume 10, Issue 04 (April 2021).
- [8]. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2018). Focal Loss for Dense Object Detection. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 2980-2988).
- [9]. Smith, J., Williams, A., & Jones, B. (2021). "Semantic Segmentation for Urban Scene Understanding: A Comparative Study." *IEEE Transactions on Intelligent Transportation Systems*, 22(4), 1768-1780.
- [10]. Chen, L., Wang, Y., & Zhang, Q. (2019). "Facial Recognition in Surveillance Videos: A Deep Learning Approach." *Pattern Recognition Letters*, 120, 25-32.
- [11]. Rachita Byahatti , Dr. S. V. Viraktamath , Madhuri Yavagal, 2021, Object Detection and Classification using YOLOv3, *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)* Volume 10, Issue 02 (February 2021).
- [12]. Md. Mehedi Hasan Naim, Rohani Amrin , Md. Romzan Ali , Abdullah Al Zubaer , Md. Ariful Islam, 2020, Object Detection from Images using Convolutional Neural Network based on Deep Learning, *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)* Volume 09, Issue 09 (September 2020),
- [13]. Galvez, Reagan & Bandala, Argel & Dadios, Elmer & Vicerra, Ryan & Maningo, Jose Martin. (2018). Object Detection Using Convolutional Neural Networks. 2023-2027. 10.1109/TENCON.2018.8650517.
- [14]. Pradyuman Tomar , Sameer Haider , Sagar, 2022, A Study on Real Time Object Detection using Deep Learning, *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)* Volume 11, Issue 05 (May 2022),
- [15]. Liu, Wei & Anguelov, Dragomir & Erhan, Dumitru & Szegedy, Christian & Reed, Scott & Fu, Cheng-Yang & Berg, Alexander. (2016). SSD: Single Shot MultiBox Detector. 9905. 21-37. 10.1007/978-3-319-46448-0\_2. Chen, X., Zhao, L., & Jiang, Y. (2022).



# Live Feed Object Detection using YOLO Algorithm with Audio Feedback Mechanism

## ORIGINALITY REPORT

13%  
SIMILARITY INDEX

9%  
INTERNET SOURCES

11%  
PUBLICATIONS

5%  
STUDENT PAPERS

## PRIMARY SOURCES

1 jwcn-urasipjournals.springeropen.com 1%  
Internet Source

2 Submitted to National Institute of Technology Karnataka Surathkal 1%  
Student Paper

3 dokumen.pub 1%  
Internet Source

4 Kola Balavani, Dasari Sriram, Malla Bhavani Shankar, Devavarapu Sri Charan. "An Optimized Plant Disease Classification System Based on Resnet-50 Architecture and Transfer Learning", 2023 4th International Conference for Emerging Technology (INCET), 2023 1%  
Publication

5 Dhruthi L, Praveen K Megharaj, Pranav P, Niharika Kiran, Asha Rani K P, Gowrishankar S. "State-of-the-Art Object Detection: An Overview of YOLO Variants and their Performance", 2023 4th International 1%

# Conference on Smart Electronics and Communication (ICOSEC), 2023

Publication

6

K. Devi Priya, A. Sasi Samyogitha, A. Vamsi Krishna Reddy, B. Divya Sri. "ENSEMBLED CROPIFY – Crop & Fertilizer Recommender System with Leaf Disease Prediction", 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), 2023

Publication

1 %

7

Srishti Vashishtha, Suraj Kumar, Vishakha Bothra, Vishal Singhal, Aaryan Sharma. "Vehicle Detection System using YOLOv4", 2022 4th International Conference on Artificial Intelligence and Speech Technology (AIST), 2022

Publication

1 %

8

Submitted to University College London

Student Paper

1 %

9

[www.researchgate.net](http://www.researchgate.net)

Internet Source

1 %

10

[ijece.iaescore.com](http://ijece.iaescore.com)

Internet Source

<1 %

11

Submitted to University of Lincoln

Student Paper

<1 %

Submitted to University of North Texas



12

Student Paper

&lt;1 %

13

[archiv.ub.uni-heidelberg.de](http://archiv.ub.uni-heidelberg.de)

Internet Source

&lt;1 %

14

[www.coursehero.com](http://www.coursehero.com)

Internet Source

&lt;1 %

15

[www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)

Internet Source

&lt;1 %

16

Submitted to Liverpool John Moores University

Student Paper

&lt;1 %

17

Submitted to University of Hull

Student Paper

&lt;1 %

18

Yanming Hui, Jue Wang, Bo Li. "DSAA-YOLO: UAV remote sensing small target recognition algorithm for YOLOV7 based on dense residual super-resolution and anchor frame adaptive regression strategy", Journal of King Saud University - Computer and Information Sciences, 2023

Publication

&lt;1 %

19

[medium.com](https://medium.com)

Internet Source

&lt;1 %

20

"Image and Graphics", Springer Science and Business Media LLC, 2019

Publication

&lt;1 %

|    |   |      |
|----|---|------|
| 21 | "Medical Image Computing and Computer Assisted Intervention – MICCAI 2018", Springer Nature America, Inc, 2018<br>Publication   | <1 % |
| 22 | "Medical Image Computing and Computer Assisted Intervention – MICCAI 2018", Springer Science and Business Media LLC, 2018<br>Publication  | <1 % |
| 23 | Sepp Hochreiter, Jürgen Schmidhuber. "Long Short-Term Memory", Neural Computation, 1997<br>Publication  | <1 % |
| 24 | Vladimir Vrbaski, Slobodan Josic, Vuk Vranjkovic, Predrag Teodorovic, Rastislav Struharik. "Puppis: Hardware Accelerator of Single-Shot Multibox Detectors for Edge-Based Applications", Electronics, 2023<br>Publication | <1 % |
| 25 | <a href="https://export.arxiv.org/">export.arxiv.org</a><br>Internet Source   | <1 % |
| 26 | <a href="https://journaljamcs.com/">journaljamcs.com</a><br>Internet Source   | <1 % |
| 27 | <a href="https://thesai.org">thesai.org</a><br>Internet Source  | <1 % |
| 28 | <a href="https://www.diva-portal.se">www.diva-portal.se</a><br>Internet Source  | <1 % |

30

Ashwani Kumar, Sonam Srivastava. "Object Detection System Based on Convolution Neural Networks Using Single Shot Multi-Box Detector", Procedia Computer Science, 2020

Publication

&lt;1 %

31

Ashwani Kumar, Zuopeng Justin Zhang, Hongbo Lyu. "Object detection in real time based on improved single shot multi-box detector algorithm", EURASIP Journal on Wireless Communications and Networking, 2020

Publication

&lt;1 %

32

Paulo A. S. Mendes, António Paulo Coimbra, Aníbal T. de Almeida. "Forest Vegetation Detection Using Deep Learning Object Detection Models", Forests, 2023

Publication

&lt;1 %

33

Rawan Aamir Mushabab AlShehri, Abdul Khader Jilani Saudagar. "Detecting Threats from Live Videos using Deep Learning Algorithms", International Journal of Advanced Computer Science and Applications, 2023

Publication

&lt;1 %

---

Exclude quotes      On

Exclude matches      Off

Exclude bibliography      On