IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

REAL TIME OBJECT DETECTION USING YOLOV3

[1] Guide – Ms. B. Jyothi

Assistant Professor

[2] R. Pravalika [3] G. Srinidhi [4] P. Viswas

Department of Computer Science and Engineering

ANURAG GROUP OF INSTITUTIONS

ABSTRACT

Real-time object detection and people counting are critical tasks in various fields such as surveillance, crowd monitoring and industrial automation. The development of deep learning models has significantly improved the accuracy and efficiency of object detection and people counting systems. One of the most successful object detection models is You Only Look Once (YOLO), a one-step object detection model that predicts boundary conditions and class probabilities for multiple objects in an image in a single pass. In our project, we propose a real-time object detection and people counting system using YOLO. The system consists of a camera that captures the video stream, a pre-processing module that performs image normalization and resizing, and a YOLO model for object detection and people counting. YOLO models are trained on large image datasets containing various objects and people to accurately track and count them. The system implements image and video processing using Python programming language and OpenCV library. The system was tested in different scenarios, including crowded areas, indoor and outdoor environments, and different lighting conditions. The system is also highly scalable and adaptable to different environments and scenarios.

INTRODUCTION

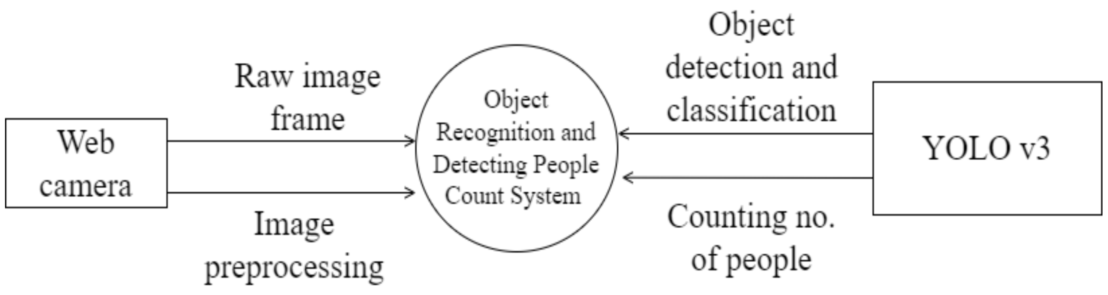
Real-time object detection is the process of identifying and locating objects in a real-time or near real-time video or image stream. The goal of real-time object detection is to accurately identify objects in a data stream and provide that information quickly and efficiently. Real-time object detection has many applications, including video surveillance, autonomous driving, robotics, and industrial automation. The main challenges of real-time object detection include managing large amounts of data, handling real-time constraints, handling complex scenes, and achieving high accuracy while maintaining low latency. To address these challenges, real-time object detection systems typically use deep learning algorithms such as convolutional neural networks (CNN) and

YOLO to perform object detection and classification. These algorithms are trained on large datasets of labelled images and can be fine-tuned for specific applications. Real-time object detection systems also typically use hardware acceleration, such as graphics processing units (GPUs) or field-programmable gate arrays (FPGAs), to achieve high performance and low latency.

PROPOSED SYSTEM

For real-time object detection, the proposed system is based on the YOLO (You Only Look Once) principle. The well-known YOLO deep learning technique uses a single neural network to predict the location, size and type of objects in an image or video.

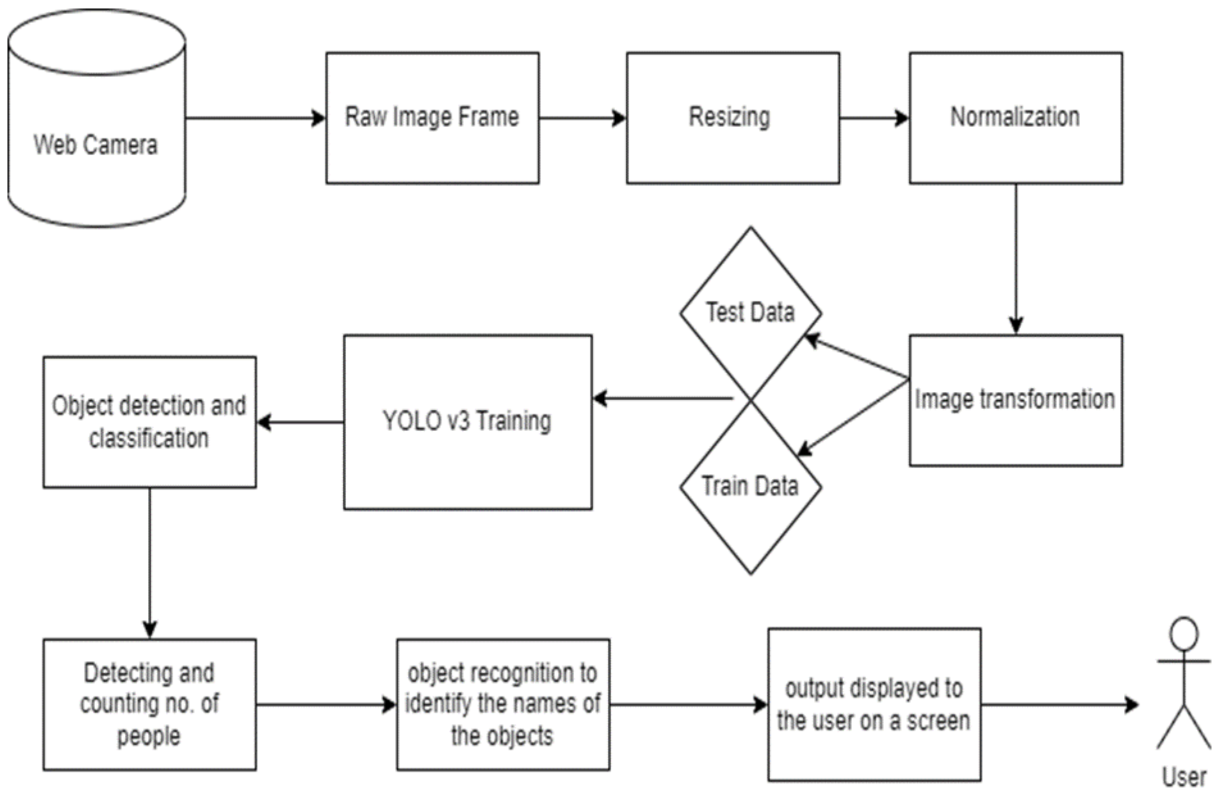
Flow of the project:


Web camera
Raw image frame
Image preprocessing
Object Recognition and Detecting People Count System
Object detection and classification
Counting no. of people
YOLO v3

ADVANTAGES:

- Faster: Compared to other algorithms, YOLO algorithm works relatively faster.
- It is a highly generalized network: YOLO can be said to be a highly generalized network due to its algorithm and training method.
- It processes each frame at a rate of 45 fps, which corresponds to a larger network, at 150 fps, which is better than real time for a smaller network.

ARCHITECTURE:


Web Camera
Raw Image Frame
Resizing
Normalization
Image transformation
Test Data
Train Data
YOLO v3 Training
Object detection and classification
Detecting and counting no. of people
object recognition to identify the names of the objects
output displayed to the user on a screen
User

IMPLEMENTATION

Module Description

A well-known computer vision technique for recognizing elements in images or movies is real-time object recognition using the YOLO (You Only Look Once) principle. YOLO is a deep learning algorithm that can quickly identify and classify features in images. The basic concept of YOLO is to divide an image into a grid of cells and predict the probability that an object is present in each cell, along with the bounding box coordinates of that object. YOLO uses a convolutional neural network (CNN) to learn these predictions.

MODULE 1: DATA COLLECTION AND PREPROCESSING

The ability to recognize and classify elements of a photo or video in real time makes real-time object detection using YOLO (You Only Look Once) a well-known computer vision method. known. Data collection is a critical step when developing a real-time object tracking system based on YOLO. The performance of our models will be strongly influenced by the quality and quantity of data we collect. Specifying the objectives of our project is the first step in the data collection process. The capabilities of our real-time object detection system, the type of object you wish to detect, and the environment in which the detection will occur must all be determined. This will help you identify the data and sources you need to collect. After identifying your goals, you need to decide which data sources to use in your project. Depending on the type of things you want to track, you may need to collect information from a number of sources, including webcams, video, images, and other sensors.

Data cleaning is the removal of any false or meaningless information from a data set. This is a crucial step as it ensures that the model is trained on high-quality data, which improves its performance and accuracy. Data pre-processing involves putting the data into a suitable format for training the model. Photos should be scaled, grayscale and their pixel values normalized. As part of the pre-processing, it may be necessary to balance the dataset to ensure that each class contains an equal number of samples.

In conclusion, gathering data is an essential step in developing a successful YOLO real-time object detection system.

The steps in the process include defining goals, locating data sources, identifying data sources, recording data, adding data, and pre-processing data. Better object detection results can be obtained by using well-designed data collection methods to significantly improve model accuracy and performance.

MODULE 2: DATA TRAINING

Training a model according to the YOLO principle (You Only Look Once) requires several steps. Here, we will detail the steps required for training the model for real-time object detection using YOLO. Selecting a training dataset containing photos or videos of the item to be detected is the first step in training the model. Datasets should be diverse, including a variety of elements, backgrounds, and lighting. The dataset should contain a large amount of data that the model can learn from. In general, the more data a model contains, the better it will perform. Yet, collecting and annotating a large dataset is time-consuming and expensive.

Before training a model, the data must be pre-processed to make it suitable for the algorithm. Images need to be resized, pixel values need to be normalized, bounding boxes need to be encoded, etc. Resizing a photo to a specific size is called image resizing. This is necessary because YOLO works best with fixed size input photos. Pixel values must be scaled within a specific range, usually between 0 and 1, to be normalized. The coordinates of the objects in the photo are encoded as a series of parameters in the bounding box encoding. This allows the YOLO algorithm to calculate the position and size of objects in the photo.

Convolutional Neural Network (CNN) layers form the architecture of the YOLO model. The model design can be modified to meet unique project requirements, such as adding more layers or changing the activation feature. In a single pass through the neural network, the YOLO architecture is designed to predict the location and class of objects. Therefore, it outperforms other object detection algorithms in speed and efficiency. Use a loss function to measure the difference between the predicted output and the produced output. Location loss, classification loss, and confidence loss are all combined to create the YOLO loss function.

The difference between the ground truth bounding boxes and the expected bounding boxes is measured by the localization loss. The classification loss calculates the discrepancy between the actual class probabilities and those anticipated, or "ground truth," probabilities. The confidence loss calculates the discrepancy between the expected and actual objectness scores. Hyperparameters are parameters that regulate how the model behaves while being trained. They consist of regularization, batch size, and learning rate. To make sure the model is properly training and not overfitting to the training data, certain hyperparameters must be tuned.

The step size utilized to update the weights during training is controlled by the learning rate. The model may converge too soon in cases of high learning rates or too slowly in cases of low learning rates. The batch size regulates how many samples are utilized to update the weights throughout each training cycle. Although a higher batch size can speed up training, it can also cause overfitting and slower convergence. To avoid overfitting and enhance the model's generalizability, regularization approaches like L1 and L2 regularization can be applied.

The model can be trained on the training dataset once the hyperparameters have been established. In order to reduce the loss function and boost its ability to forecast object classes and bounding boxes, the model makes weight adjustments during training. Depending on the quantity of the dataset and the complexity of the model architecture, training a YOLO model may take hours or even days. The model should be trained until the accuracy on the validation dataset stops increasing and the loss function converges to a minimum. The model must be tested on a validation dataset after training to determine.

MODULE 3: DETECTION OF OBJECTS

The technique of object detection in YOLO (You Only Look Once) entails estimating the location, size, and class of objects inside an input image or video. Unlike other object detection algorithms that scan photos or videos in pieces, the YOLO method processes the entire image or video at once, making it quicker and more effective. In order to forecast bounding boxes for objects in the input image or video, the YOLO technique employs a single neural network. Each bounding box has a class probability attached to it, which indicates the possibility that the object falls under a specific class. Convolutional neural networks (CNNs) are used by the YOLO method to extract information from the input picture or video. Several layers make up the CNN, which is able to recognize patterns and characteristics in images and videos.

Following that, a number of fully connected layers that forecast the bounding boxes and class probabilities process the CNN's output. The localization loss, classification loss, and confidence loss are all combined into one loss function that is used by the YOLO algorithm. The localization loss calculates the variation in bounding boxes between the predicted and actual ones. The classification loss calculates the discrepancy between the actual class probabilities and those anticipated. The difference between the expected and actual objectness scores is represented by the confidence loss. The output is post-processed after the YOLO algorithm has predicted the bounding boxes and class probabilities for the items within the input image or video. In post-processing, redundant bounding boxes are eliminated using non-maximum suppression, and low-confidence bounding boxes are filtered out using thresholding. The output of the YOLO method is visualized by defining bounding boxes around the identified items and assigning them a predicted class as the last step. As a result, real-time object detection and labeling is shown in an image or video.

Introduction to Technologies Used

Real-time object detection using YOLO (You Only Look Once) relies on several technologies to achieve efficient and accurate object detection in real-time or near real-time. These technologies include:

Deep learning: YOLO uses a deep convolutional neural network to detect objects in images or video frames. The network is trained on large datasets of labeled images to learn how to detect objects of interest.

Convolutional Neural Networks (CNNs): CNNs are a type of neural network that are commonly used in computer vision applications, including object detection. YOLO uses a CNN architecture to extract features from the input image and predict bounding boxes and class probabilities for each cell in the image grid.

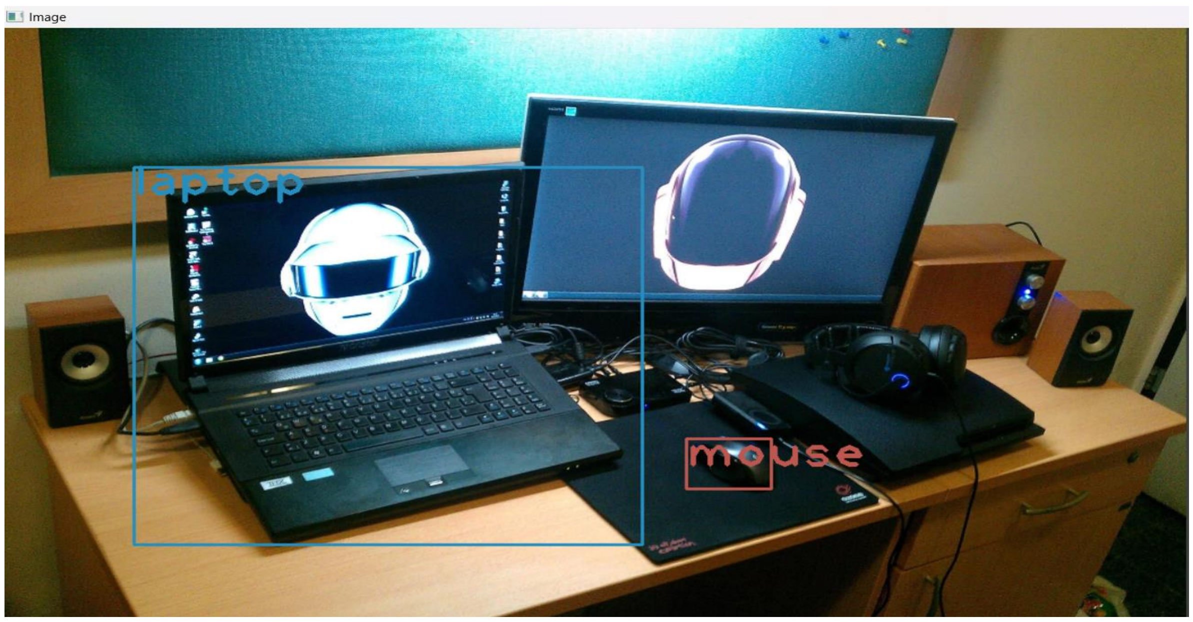
Hardware acceleration: To achieve high performance and low latency, real-time object detection using YOLO often relies on hardware acceleration, such as Graphics Processing Units (GPUs) or Field-Programmable Gate Arrays (FPGAs). These hardware accelerators can perform computations in parallel, enabling faster and more efficient object detection.

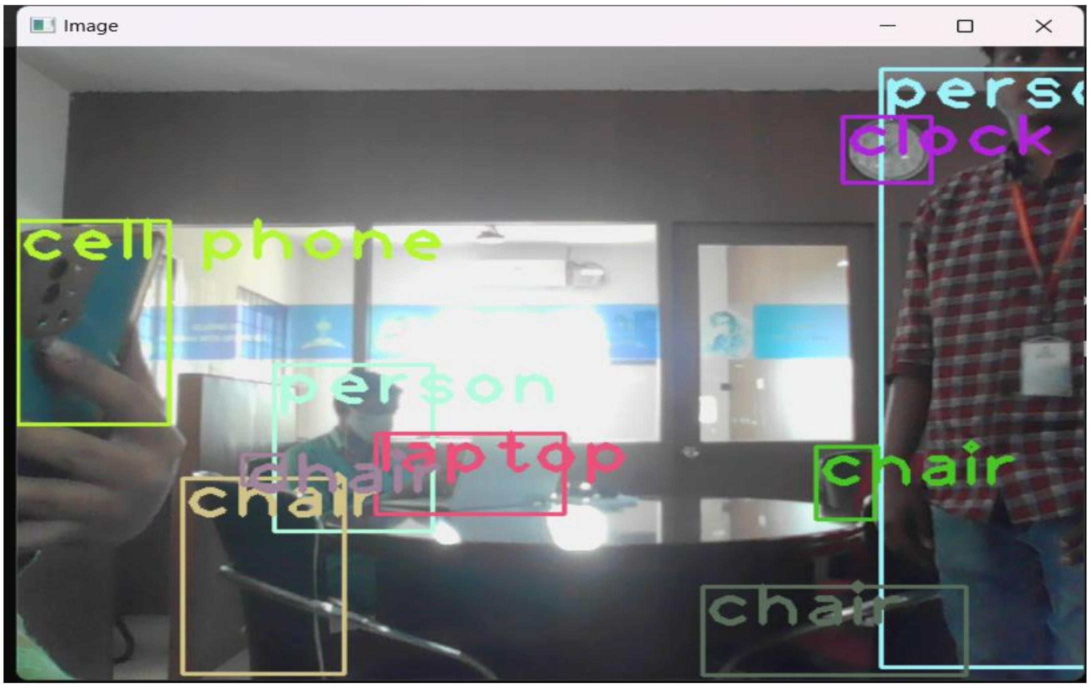
Image processing: Real-time object detection using YOLO may also rely on image processing techniques, such as image normalization and data augmentation, to improve the accuracy and robustness of the object detection algorithm.

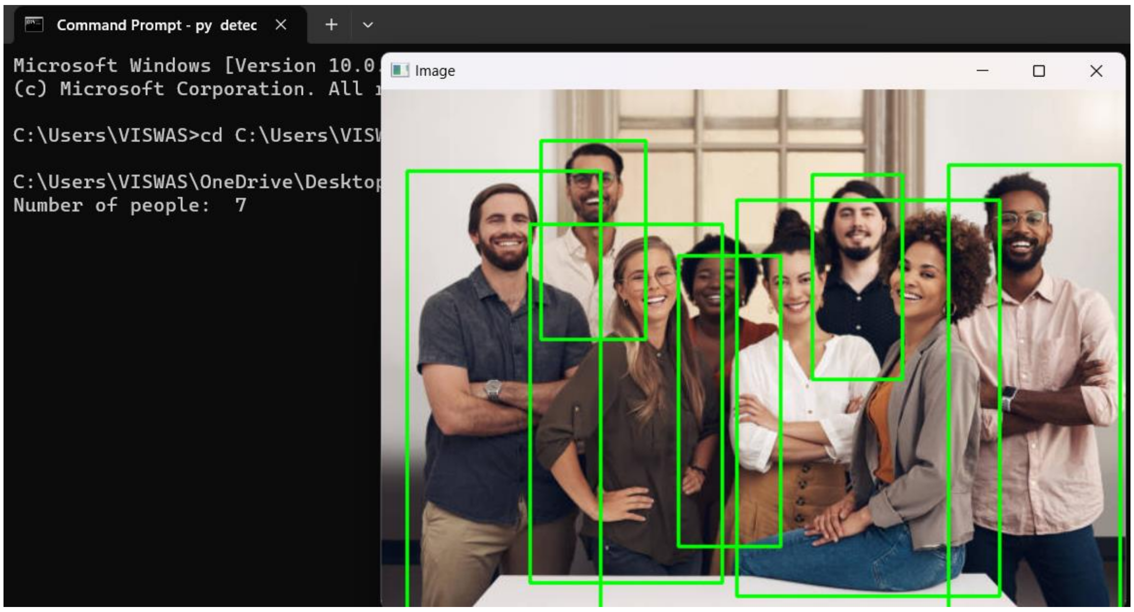
Optimization techniques: To achieve the best possible trade-off between accuracy and speed, YOLO uses several optimization techniques, including multi-scale training, anchor box clustering, and online hard example mining, to improve the performance of the algorithm.

Overall, real-time object detection using YOLO combines several advanced technologies to achieve efficient and accurate object detection in real-time or near real-time.

OUTPUT:



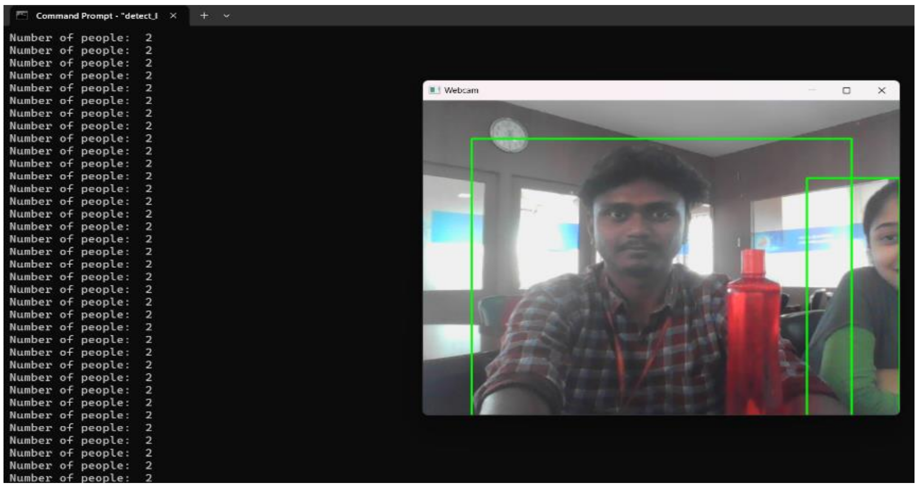




```
Microsoft Windows [Version 10.0
(c) Microsoft Corporation. All r

C:\Users\VISWAS>cd C:\Users\VISW

C:\Users\VISWAS\OneDrive\Desktop
Number of people:  7
```



```
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
Number of people:  2
```

CONCLUSION

Using YOLO for real-time object tracking and people counting is an efficient and accurate way to track and locate objects in real time. YOLO is a state-of-the-art deep learning model capable of detecting objects in images and videos with high accuracy and speed. It uses a unique neural network architecture to predict boundary blocks and class probabilities for multiple objects in an image, which makes it very efficient and fast. In short, using YOLO for real-time target tracking and people counting is an efficient and accurate method for real-time target tracking and location. Its high accuracy and speed make it suitable for a wide range of applications including surveillance systems, traffic monitoring, crowd analysis and industrial automation. By providing real-time information, YOLO enables proactive decision-making and improves operational efficiency.

REFERENCE

[1] M. H. Nugraha and D. Chahyati, "Tourism Object Detection Around Monumen Nasional (Monas) Using YOLO and RetinaNet," 2020 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 2020, pp. 317-322, doi: 10.1109/ICACSIS51025.2020.9263240.

[2] S. Kuan-Ying, C. Ming-Fei, T. Po-Cheng and T. Cheng-Han, "Establish a Dynamic Detection System for Metal Bicycle Frame Defects Based on YOLO Object Detection," 2022 IET International Conference on Engineering Technologies and Applications (IET-ICETA), Changhua, Taiwan, 2022, pp. 1-2, doi: 10.1109/IET-ICETA56553.2022.9971568.

[3] J. Ge, B. Zhang, C. Wang, C. Xu, Z. Tian and L. Xu, "Azimuth-Sensitive Object Detection in Sar Images Using Improved Yolo V5 Model," IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022, pp. 2171-2174, doi: 10.1109/IGARSS46834.2022.9883072.

[4] A. Sarda, S. Dixit and A. Bhan, "Object Detection for Autonomous Driving using YOLO [You Only Look Once] algorithm," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 2021, pp. 1370-1374, doi: 10.1109/ICICV50876.2021.9388577.

[5] H. Yun and D. Park, "Yolo-based Realtime Object Detection using Interleaved Redirection of Time-Multiplexed Streamline of Vision Snapshot for Lightweighted Embedded Processors," 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), Hualien City, Taiwan, 2021, pp. 1-2, doi: 10.1109/ISPACS51563.2021.9651042.

[6] R. L. Galvez, E. P. Dadios, A. A. Bandala and R. R. P. Vicerra, "YOLO-based Threat Object Detection in X-ray Images," 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Laoag, Philippines, 2019, pp. 1-5, doi: 10.1109/HNICEM48295.2019.9073599.

[7] I. V. S. L. Haritha, M. Harshini, S. Patil and J. Philip, "Real Time Object Detection using YOLO Algorithm," 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2022, pp. 1465-1468, doi: 10.1109/ICECA55336.2022.10009184.

[8] R. L. Galvez, E. P. Dadios, A. A. Bandala and R. R. P. Vicerra, "YOLO-based Threat Object Detection in X-ray Images," 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management(HNICEM),Laoag,Philippines,2019,pp.1-5,doi:10.1109/HNICEM48295.2019.9073599.

[9] J. Fan, J. Lee, I. Jung and Y. Lee, "Improvement of Object Detection Based on Faster R-CNN and YOLO," 2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), Jeju, Korea (South), 2021, pp. 1-4, doi: 10.1109/ITC-CSCC52171.2021.9501480.

[10] C. J. Huang, T. -W. Chen and Y. -C. Fan, "Chip and System Design of Real Time Object Detection Based on LS-R-YOLO Network," 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), Kobe, Japan, 2020, pp. 158-159, doi: 10.1109/GCCE50665.2020.9291775.