WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Real Time Object Detection Using Deep Learning

B. Bhargavi[1], D. Hari Chandana[2], Md. Shabana[3], K. Govinda Raju[4]

[1,2,3]U.G. Student, CSE, B. Tech, Aditya Engineering College

[4]M. Tech, (Ph. D), Associate Professor, CSE, B. Tech, Aditya Engineering College

ABSTRACT—

Deep knowledge acquisition has had a wonderful influence on how the world is adjusting to artificial intelligence during the past few years. Single Shot Detector (SSD), Faster-RCNN, Region-based Convolutional Neural Networks (RCNN), and You Only Look Once are a few of the well-known object identification techniques (YOLO). Among these, Faster-RCNN and SSD perform better in terms of accuracy, but YOLO performs better when speed is prioritized over accuracy. Deep getting to be aware of combines SSD and Mobile Nets to feature surroundings pleasant implementation of detection and tracking. This algorithm performs surroundings pleasant object detection at the same time as no longer compromising on the overall performance. (YOLOv4) is a form of deep learning object detection algorithm. The YOLO detector has the capacity to predict an object's class, bounding box, and likelihood that that object's class will be discovered inside that bounding box. To extract more effective information, global features, channel attention and special attention are also applied. By using custom functions developed upon YOLov4 we get the count of the objects and a crop around the objects detected and move to individual folders of different classes. You solely seem to be as soon as (YOLO) is a household of speedy and specific one-stage object detectors. The majority of modern-day correct fashions require many GPUs to instruct with a huge mini batch size, attempting this with a single GPU outcomes in extraordinarily slow and unfeasible training. This issue is addressed by YOLO v4 by developing an object detector that can be trained on a single GPU with a smaller mini-batch size. YOLO are one-stage detectors, but there are also two-stage correct-but-slow detectors like R- CNN, quick R- CNN, and quicker R- CNN. We'll pay attention on the before ones. We additionally cowl the past, present, and future makes use of object detection in this learn about throughout a range of industries. This find out about proposes a customary trainable framework for object detection in images and videos, and whilst monitoring have a vary of applications.

Key Words: Object Detection, YOLOv4, Image Processing, Computer Vision, Tensor Flow, Training Models. Deep learning (DL).

Introduction

Real-time object detection extracts information from still or moving images to recognize and find at least one compelling target. Among the techniques discussed are image processing, pattern recognition, and AI. It has expansive applications likelihood in areas such as traffic prediction and control, accident anticipation, crowd counting, caution of hazardous products in factories, verification using face and iris code, object tracking and counting, military restricted region observing, spotting suspicious objects in crime detection and investigation. Due to the complexity and variability of detecting multiple targets application scenarios in balancing the ratio of precision to computing expenses in the actual world is tough. To beat this issue, an array of ways has been offered, most of them are based on PC vision and deep learning techniques.

YOLO is a real-time object recognition system that can identify many items at once. It also recognizes items more quickly and precisely than existing systems. It can estimate up to 80 and even more seen and unseen classes of objects [3-9]. The real-time attention device was once in a position to distinguish many matters from a single image, body a confined-edge field round close by objects and be educated and applied in a manufacturing machine quickly. It's also a breakthrough in object detection research that leads to improved, faster, and more adaptive computer vision algorithms. YOLOv4 greatly exceeds previous techniques in terms of detection, performance, and speed [10]. It's a "quickly operating" item detector that can be easily trained and employed in manufacturing system.

Along with the detection of various items, they are cropped and saved in their own folder. "To optimize neural networks detector for parallel calculations" was the major goal. It also includes a variety of possible designs and architectural choices after carefully examining the implications on the performance of a variety of detectors, as proposed by prior YOLO models [11-13]. Instead of selecting an interesting ROI in the image, this method predicts the classes and bounding boxes for the full picture at the same time and permit for speedy detection.

Related Work

[1] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, 2016:

They delivered YOLO, a novel object detecting strategy. Classifiers from earlier work on object detection are applied to attribute detection. The problem of item detection was instead presented as a regression problem to spatially separated bounding containers and associated variety probabilities. In one comparison, a single brain area can predict bounding boxes and form probabilities as quickly as it can from full snap photos. We introduce YOLO, a novel approach of object detection. Instead, we conceive object detection as a regression problem to spatially excellent bounding containers and corresponding classification probabilities. Bounding packing containers and category chances are without delay envisioned by means of a single neural community from whole snap shots in a single assessment. As the entire detection pipeline consists of a single network, detection performance can be tuned from beginning to end. Our integrated architecture is really quick. With a body price of forty-five frames per second, our crucial YOLO mannequin strategies photos in actual time. Fast YOLO, a scaled-down model of the network, tactics simply a hundred and fifty-five frames per 2d whilst nevertheless outperforming different real-time detectors in mAP by means of an issue of two. Modern detection methods generate less false positive predictions on background, but YOLO makes more localization errors. Last however no longer least, YOLO choices up very vast representations of objects. When generalizing from herbal photos, it outperforms other detection algorithms like DPM AND R CNN when used to other areas like art..

[2] Anurabha M. Roy, Jayabrata Bhaduri, 2022:

Our recent work suggests a real-time object identification framework called Dense-YOLOv4 that is especially built around an accelerated version of the YOLOv4 algorithm with DenseNet serving as the backbone to maximize attribute swapping and reuse. Additionally, to maintain fine-grain localised records, a modified direction aggregation nearby (PANet) has been used. In industrial orchards, one of the essential phases in yield estimation and shrewd spraying is real-time awareness of agricultural increase stages. Traditional detection techniques, on the other hand, are limited in their ability to precisely identify unique growth phases because of the high degree of occultation in surrounding leaves, the significant overlap between nearby fruits, variations in fruit size, color, cluster density, and unusual growth characteristics. The current work presents a real-time object identification framework known as Dense-YOLOv4 that is specifically built wholly on an extra high quality model of the YOLOv4 method in order to optimize attribute swap and reuse. To hold fine-grain localized information, a modified route aggregation neighborhood (PANet) has been put into vicinity. The model has been used to identify several mango growth phases in a high-occultation environment. The mannequin has been used to perceive a number of mango boom degrees in a problematic orchard state of affairs with a excessive diploma of occultation. The counseled model's imply common accuracy (mAP) and 1-score have multiplied to up to 96.20% and 93.61%, respectively, at detection charge of 44.2 FPS. With upgrades in precision, recall, 1-score, and mAP of 7.94%, 13.10%, 10.47%, and 4.73%, respectively, over the most latest YOLOv4, the counseled Dense-YOLOv4 has outperformed it. The modern-day work presents a framework that is advantageous and environment friendly for figuring out more than a few boom levels in a complicated orchard putting and may additionally be multiplied to pick out different fruits and crops, discover diseases, and utilize a number computerized agricultural applications.

[3] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, 2020:

A wide range of factors are mentioned to enhance Convolutional Neural Network (CNN) accuracy. It is necessary to test combinations of these features in practice using large datasets and to theoretically support the findings. While some characteristics, like batch-normalization and residual-connections, are pertinent to the majority of models, tasks, and datasets, others, like residual-connections, feature on best models extensively and for certain problems solely, or completely for small datasets. Many elements are purported to extend Convolutional Neural Network (CNN) accuracy. For a realistic check out of such distinctive pairings, large datasets must be employed, and the effects must be theoretically supported. Some qualities, such as batch-normalization and residual-connections, are applicable to the majority of models, tasks, and datasets, while others, such as precise models, issues, or small-scale datasets, completely benefit from other characteristics. We presume to include the following among these universal properties: Self-adversarial-training (SAT), Cross-Stage-Partial-connections (CSP), Cross mini-Batch Normalization (CmBN), Weighted-Residual-Connections (WRC), and Mish-activation. We combine a few of the new components, including WRC, CSP, CmBN, SAT, Mish activation, Mosaic information augmentation, CmBN, Drop Block regularisation, and CIoU loss, to obtain company-new outcomes. 43.5% AP (65.7% AP50) for the MS COA

[4] Malik Haris, Adam Glowacaz, 2022:

Object detection is chosen by automated using and vehicle safety designs. It is essential that object detection work in real time, be accurate every day, and be resilient to local environmental conditions. They need graphic processing algorithms to look at the contents of images as a result of this method. This article examines the precision of 5 essential image processing algorithms: RetinaNet, Single Shot Multi-Box Detector (SSD), Mask Region-based Convolutional Neural Networks (Mask R-CNN), and You Only Look Once v4 (YOLOv4). The automatic riding and auto protection systems require object detection. It is indispensable that object detection function in real-time, be typically accurate, and be resistant to climate and environmental conditions. To make the contents of the pictures appear to be there, this method requires the application of image processing techniques. In this article, RetinaNet, Single Shot Multi-Box Detector (SSD), Mask Region-based Convolutional Neural Networks (Mask R-CNN), and Region-based Fully Convolutional Network (R-FCN) accuracy of 5 well-known image processing methods are analyzed (YOLOv4). For this distinction research, we made advantage of a sizable dataset from Berkeley Deep Drive (BDD100K). Its advantages and risks are analyzed based on variables such as accuracy (with/without occlusion and truncation), calculation time, and precision-recall curve. This article's assessment of trendy deep learning-based algorithms helps readers recognize their benefits and hazards whilst working with real-time deployment constraints. We come to the conclusion that in a same trying out putting with complicated street situations and difficult climate circumstances, the YOLOv4 outperforms precisely in detecting difficult street goal items.

[5] Addie Ira Borja Parico & Tofael Ahamed, 2021:

This research sought to develop a reliable real-time pear fruit counting for mobile applications using just RGB data, variations of the modern object detection model YOLOv4, and a few object-tracking algorithms Deep SORT. This research also provided a methodical and practical approach for selecting the best model for a specified agricultural sciences software package. YOLOv4-CSP used to be regarded as the best model in terms of accuracy. This study's goal was to develop a dependable real-time pear fruit counting system for mobile applications using only RGB data, the contemporary object identification model YOLOv4 variations, and a few object-tracking techniques. This research about moreover furnished a methodical and sensible technique for deciding on the incredible model for an supposed utility in agricultural sciences. YOLOv4-CSP was once observed to be the satisfactory mannequin in phrases of accuracy, with an AP@0.50 of 98%. At a velocity of extra than 50 FPS and FLOPS of 6.8–14.5, YOLOv4-tiny was once located to be the perfect mannequin in phrases of velocity and computing value.

[6] Chris J. Maddison, Aja Huang, Ilya Sutskever, David Silver, 2015:

In this paper, the authors investigate whether or not deep convolutional networks can be effectively employed to represent and interpret this data. A large 12-layer convolutional neural network is trained using supervised learning from a library of games played by human experts. In 55% of cases, the nearby correctly anticipates the expert move, matching the accuracy of a 6 dan human player. Apart from any search, when the expert convolutional community was employed immediately to play games of Go, it outperformed the commonplace search engine GNU Go in 97% of games and was on par with a state-of-the-art Monte-Carlo tree search in terms of performance that replicates one million cross positions. Using a deep convolutional neural network, we clear up these quintessential troubles of Go know-how illustration and gaining information of in this context (CNN). While CNNs have been used to play Go in the past, with a range of levels of success (Schraudolph et al.,1994; Enzenberger, 1996; Sutskever &amp; Nair, 2008), previous architectures have often been constrained to one hidden layer of exceedingly small size, and have no longer taken advantage of modern-day enhancements in computational power. In this study, we describe and accumulate Go grasp the utilization of notably giant and deeper CNNs with 12 hidden layers and a range of billion connections.

[7] Akansha Bathija M.Tech Student, Dept of Computer Engineering K J Somaiya College of Engineering Mumbai, Maharashtra, India :

Neural networks are a family of algorithms that are frequently modelled after the human brain and are used to find patterns. They use a structure of uncooked enter system perception, such as marking or clustering, to view sensory data. Face detection, facial expression recognition, and humans cognizance in photographs. Identify the items in images or videos. Clustering and grouping are techniques for locating similarities. Using categorization, deep getting to know may also set up relationships between, for example, pixels in a photo and a person's title. One of the most popular machine learning techniques at the moment is neural networks.

Most research uses deep mastering algorithms to derive validated deep qualities. These include location, monitoring, identification, human crowd detection, self-stabilization, obstacle and crash avoidance, grasp of forested or mountainous trails, and object tracking. They have achieved excellent results in a variety of difficult tasks that have traditionally required the use of handmade features. As a result of the advancement of self-driving cars, intelligent video surveillance, facial recognition, and novel people counting applications, a need for quick and distinctive object detection constructs is emerging. These systems demand that each object in a photograph be named and classified, but also its placement by drawing a circle around the appropriate bounding box. Because of this, the problem of object identification in computer vision is far more difficult than that of image recognition.

[8] S. Moon, J. Lee, D. Nam, H. Kim and W. Kim, "A comparative study on multi-object tracking methods for sports events", 2017 19th International Conference on Advanced Communication Technology (ICACT), 2017.

Automatic object detection, object tracking are now difficult and time-consuming tasks due to recent advancements in underwater video surveillance systems. The method for such processing entails preprocessing, characteristic extraction, object classification, object detection, and object tracking. The potential to perceive shifting matters in underwater video has a range of feasible makes use of for remotely operated motors (ROVs) or self-reliant underwater vehicles (AUVs), together with the capacity to reveal fish and discover different underwater objects. The subject of underwater object detection is exacerbated by means of modifications in water structure, seasonal and climatic variations, temperature swings, and an inadequate, non-uniform supply of synthetic light.

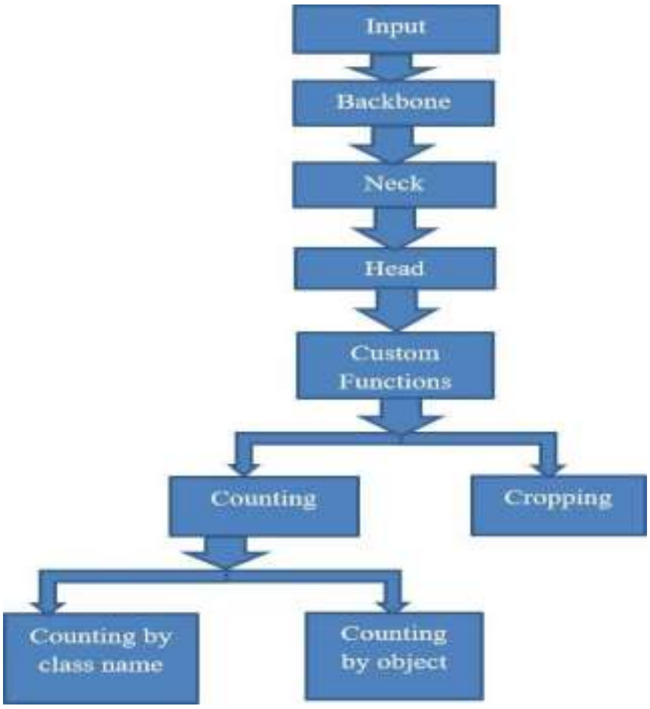
Methodology Used

A computer-based science known as object detection deals with finding instances of specific types of semantic objects (such people, buildings, cars, or animals) in digital photos and movies. It is related to laptop computer innovative and prescient and image digitally. It is associated to laptop imaginative and prescient and photograph digitally. It is associated to pc imaginative and prescient and picture processing. One of the real-time object detection models that is SOTA is the YOLOv4 (state of the art). YOLOv4 is the fourth entry in the YOLO collection. It demonstrated SOTA performance when applied to the COCO dataset, which consists of eighty different object categories. There is only one step in the YOLOv4 detector. The One-stage approach is one of the two primary cutting-edge methods for object detection that prioritizes inference speeds. The ROI (Region of Interest) is now not recognized in one-stage detector models, however the training and bounding packing containers for the complete photo are predicted. As a result, they are faster than detectors with two stages. Dataset: The MS COCO dataset, which is enormous scale objects detection, segmentation and inscribing dataset. The COCO dataset is generally used by AI and PC vision researchers for some PC vision projects.

YOLOv4 architecture: The YOLOv4 architecture des describes the basic working of the yolov4 system. We can see that the YOLOv4 approach is used in the From the picture of the architecture.

- Input
- Backbone
- Neck

- Head
- Custom Functions



Counting all of the objects in the picture:

Within the file core/functions.py, a customized characteristic is constructed It may also be used to be counted and tune the rely of objects identified in every photograph or video at any given time. It can hold song of the complete wide variety of objects determined or the variety of objects detected through class.

Counting objects per class:

Change one line in the detect.py or realize video.py script to matter quantity of objects for every classification of your object detector the use of the customized flag "—count". The default fee for the by way of type parameter in the count number objects approach is False. When this parameter is set to True, the be counted is executed per class.

Crop Detections and save them as new images:

In the file core/functions.py, a customized characteristic is constructed which can be utilized to any detect.py or discover video.py command to crop the YOLOv4 detections and retailer them as new image. All you have to do is add the —crop flag to any command to crop detections. The detections that are cropped are saved in the detections/crop/ folder.

SYSTEM CONFIGURATION

1.Git Bash:

Git can be thought of as a collection of command-line tools created specifically for Windows environments. UNIX command line terminals are included in many operating systems, including Linux and macOS. It makes Linux and macOS complementary working structures when working with Git. Windows do no longer have the UNIX trend command interface. Instead, Home Windows Command Prompt, a non-UNIX terminal, is used by Microsoft Windows. In order to execute Git from the command line, Git for Windows provides a Bash emulation. To put it another way, Git Bash is a tool that provides a layer of emulation for Microsoft Windows systems to use the Git command-line interface. It is just a package that installs some commonly used bash programmers on a PC running Windows. It let us use all the Git factors as properly as most of the general UNIX instructions in a command-line interface on Windows.

2. Visual Studio Code:

The Visual Studio IDE serves as a creative platform from which you can edit, debug, and create code, as well as develop and publish apps. The software development process is improved by Visual Studio's inclusion of compilers, code completion tools, graphical designers, and many other features in addition to the conventional editor and debugger that are offered by the majority of IDEs. The most complete IDE for C++ and.NET developers running on Windows. A nice assortment of tools and features that will elevate and improve each stage of software development are included. Windows, macOS, and Linux-compatible standalone source code editor. With extensions to support almost any programming language, it is the preferred choice for JavaScript and web developers.

3. Windows or Linux Operating system:

Microsoft Windows is a personal computer operating system (OS) created by the Microsoft Corporation. It is also known as Windows and Windows OS (PCs). The Windows OS quickly took control of the PC market because it offered the first graphical user interface (GUI) for IBM-compatible Computers. Most PCs (around 90%) use some version of Windows. The initial release of Windows, in 1985, was merely an addition to Microsoft's already-existing MS-DOS disc operating system. Windows was the first operating system to let DOS users visually navigate a virtual desktop by opening graphical "windows" that displayed the contents of electronic folders and files with the click of a mouse, as opposed to typing commands and directory paths at a text prompt. Windows was based in part on licensed concepts that Apple Inc. had used for its Macintosh System Software.

4.Open CV:

OpenCV is the large open-source library for the laptop vision, desktop learning, and photo processing and now it performs a main function in real-time operation which is very essential in today’s systems. By the use of it, one can manner photographs and movies to discover objects, faces, or even handwriting of a human. When it built-in with a number of libraries, such as NumPy, python is successful of processing the OpenCV array shape for analysis. To Identify picture sample and its a number of elements we use vector area and operate mathematical operations on these features. OpenCV is a cross-platform library the usage of which we can increase real-time laptop imaginative and prescient applications. It usually focuses on photo processing; video seize and evaluation such as aspects like face detection and object detection.

5.GPU (Graphics Processing Unit):

Graphics processing technological know-how has advanced to supply special advantages in the world of computing. The state-of-the-art snap shots processing devices (GPUs) unencumber new chances in gaming, content material creation, computer learning, and more. Designed for parallel processing, the GPU is used in a substantial vary of applications, which includes portraits and video rendering. Although they’re great diagnosed for their abilities in gaming, GPUs are turning into greater famous for use in innovative manufacturing and artificial Genius (AI). They grew to become greater bendy and programmable, bettering their capabilities. This allowed photos programmers to create greater fascinating visible results and sensible scenes with top-quality lighting fixtures and shadowing techniques. Other builders additionally started to faucet the energy of GPUs to dramatically pace up extra workloads in excessive overall performance computing (HPC), deep learning, and greater.

6.CSPDarknet53:

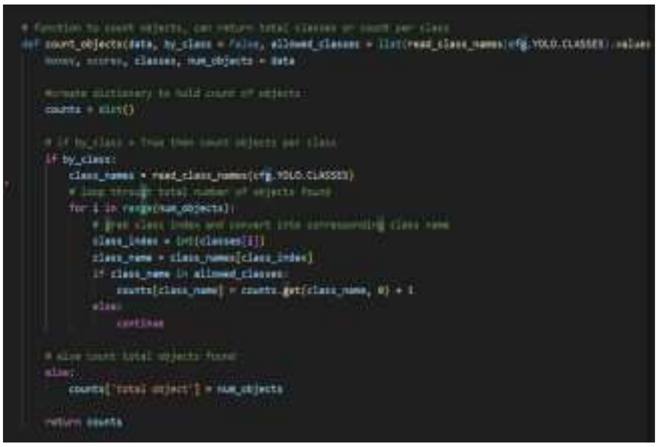
CSPDarknet53 is a convolutional neural community and spine for object detection that makes use of DarkNet-53. The function map of the base layer is divided into two components using a CSPNet technique, and they are then combined using a cross-stage hierarchy. The use of a breakup and merge method permits for greater gradient float via the network. This CNN is used as the spine for Yolov4. DarkNet-53 is regularly used as the basis for object detection issues and YOLO workflows.

IMPLEMENTATION

The YOLOv4 for the computer vision enhancement mainly concentrates on three functioning’s. They are Counting all of the objects in the picture, number of objects per class and cropping the detections from the image.

Counting total number of objects:

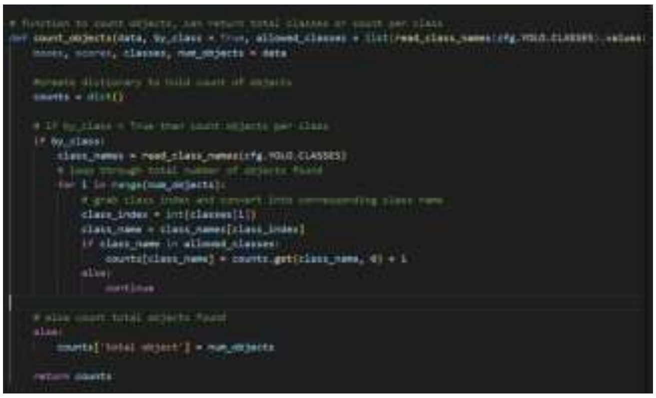
Here, the function works on the various features of the image and detects the objects present in them. It finally returns the total count of the number of objects present given input image. The total count is displayed on the top left corner of the output image.



```
# function to count objects, can return total classes or count per class
def count_objects(data, by_class = False, allowed_classes = list(read_class_names(cfg.YOLO.CLASSES).values
    boxes, scores, classes, num_objects = data

    #create dictionary to hold count of objects
    counts = dict()

    # if by_class = True then count objects per class
    if by_class:
        class_names = read_class_names(cfg.YOLO.CLASSES)
        # loop through total number of objects found
        for i in range(num_objects):
            # grab class index and convert into corresponding class name
            class_index = int(classes[i])
            class_name = class_names[class_index]
            if class_name in allowed_classes:
                counts[class_name] = counts.get(class_name, 0) + 1
            else:
                continue

    # else count total objects found
    else:
        counts['total object'] = num_objects

    return counts
```

Counting number of objects per class :

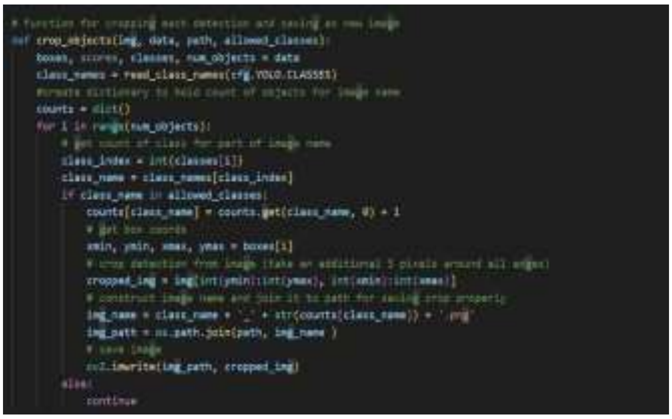
The defined function extracts the features of the image detects the objects based on the class name. It finds out whether the class is present in the 80 classes of present in the YOLOv4. The output is displayed on the top left corner of the screen addressing count of each class.



```
# function to count objects, can return total classes or count per class
def count_objects(data, by_class = True, allowed_classes = list(read_class_names(cfg.YOLO.CLASSES).values(
    boxes, scores, classes, num_objects = data

    #create dictionary to hold count of objects
    counts = dict()

    # if by_class = True then count objects per class
    if by_class:
        class_names = read_class_names(cfg.YOLO.CLASSES)
        # loop through total number of objects found
        for i in range(num_objects):
            # grab class index and convert into corresponding class name
            class_index = int(classes[i])
            class_name = class_names[class_index]
            if class_name in allowed_classes:
                counts[class_name] = counts.get(class_name, 0) + 1
            else:
                continue

    # else count total objects found
    else:
        counts['total object'] = num_objects

    return counts
```

Cropping the detections from the image :

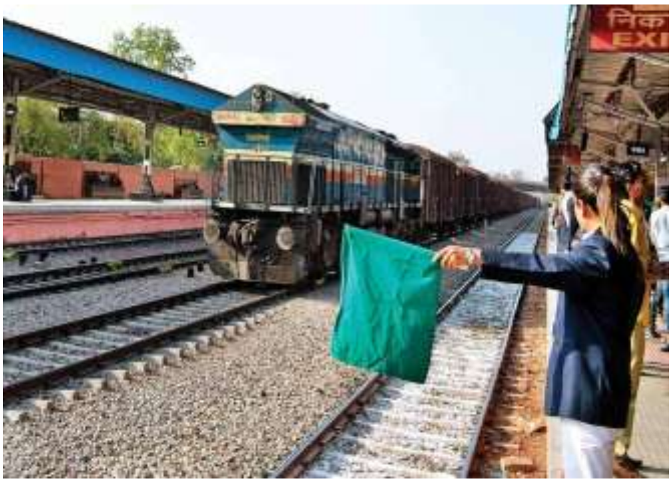
The cropping function defined mainly concentrates on the coordinates of the image. These coordinates of all the different objects present in the picture are extraceted and gives the result as cropped detections from the image.



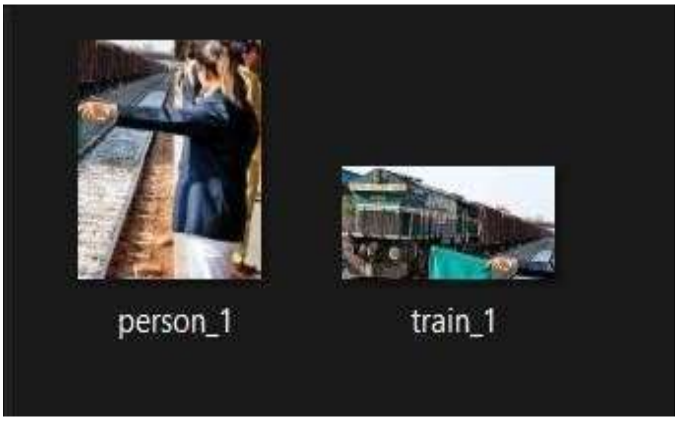
```
# function for cropping each detection and saving as new image
def crop_objects(img, data, path, allowed_classes):
    boxes, scores, classes, num_objects = data
    class_names = read_class_names(cfg.YOLO.CLASSES)
    #create dictionary to hold count of objects for image name
    counts = dict()
    for i in range(num_objects):
        # get count of class for part of image name
        class_index = int(classes[i])
        class_name = class_names[class_index]
        if class_name in allowed_classes:
            counts[class_name] = counts.get(class_name, 0) + 1
            # get box coords
            xmin, ymin, xmax, ymax = boxes[i]
            # crop detection from image (take an additional 5 pixels around all edges)
            cropped_img = img[int(ymin):int(ymax), int(xmin):int(xmax)]
            # construct image name and join it to path for saving crop properly
            img_name = class_name + '_' + str(counts[class_name]) + '.png'
            img_path = os.path.join(path, img_name )
            # save image
            cv2.imwrite(img_path, cropped_img)
        else:
            continue
```

Results and Discussions

Input Image 1:



Output Image1:



Conclusion

In this work, we have discussed the specifics of detecting the items that are present in image. In general, YOLOv4 is a scuttling-edge object identification model that detects the items present in the image. The custom functions allow us to analyze the data more efficiently. A confidence score can also be calculated while detecting the objects in the image which provides the user with the probability of the detected objects. This system can be used in surveillance. We can say the implemented system can also have the same levels of accuracy and more. We may conclude that YOLOv4 has higher accuracy and better features when compared to YOLOv3 and other real time object detection techniques. In this study, 800 photos from a proprietary dataset were used to train a detector for real-time object detection across 6 distinct classes. Using YOLOv4 to monitor the items over multiple frames, moving objects are detected. By fine-tuning the detector while training the system over more epochs, accuracy and precision can be improved.

FUTURE SCOPE

Detecting objects in a given picture or video body has been round for years, it is turning into greater huge throughout a vary of industries now extra than ever before. Object detection in pix and video has obtained loads of interest in the laptop imaginative and prescient and sample cognizance communities over current years. We have developed many strategies for object detection, however the software of deep studying guarantees greater accuracy for a wider range of object lessons. Similar methods for weakly supervised image segmentation will hopefully be used in subsequent research. When we apply weak labels to classification data during training, we also intend to enhance our detection performance by employing more potent matching algorithms. There is an abundance of tagged data for computer vision. In order to create more robust models of the visual world, we will keep looking for ways to combine various data sources and data structures.

References

[1] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, 2016:

[2] Anurabha M. Roy, Jayabrata Bhaduri, 2022:

[3] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, 2020:

[4] Malik Haris, Adam Glowacaz, 2022:

[5] Addie Ira Borja Parico & Tofael Ahamed, 2021:

[6] Chris J.Maddison, Aja Huang, Ilya Sutskever, David Silver, 2015:

[7] Akansha Bathija, 2017

[8] S. Moon, J. Lee ,D .Nam, H . Kim and W. Kim, "A comparative study on multi-object tracking methods for sports events", 2017.

Web Links:

- https://www.cvfoundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html
- https://www.sciencedirect.com/science/article/abs/pii/S0168169922000114
- https://arxiv.org/abs/2004.10934
- https://www.mdpi.com/2079-9292/10/16/1932
- https://www.mdpi.com/1424-8220/21/14/4803