SOCIAL DISTANCING DETECTOR

Himanshu Singh , Mohit Jindal , Yash pal Singh , Harshit

DR. AKHILESH DAS GUPTA INSTITUTE OF TECHNOLOGY AND MANAGEMENT
(AFFILIATED TO GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI)

Course: - Information Technology

ABSTRACT

This project provides a deep learning-based system for detecting social distancing in order to reduce the consequences of the coronavirus outbreak. The detection model was created in order to check that peoples Keep a safe distance from by comparing and contrasting video feeds. The YOLOv3 algorithm-based open-source object identification pre-trained model and the COCO dataset takes the video frame as input. It is necessary to gauge the distance between people and a red frame and red line will be displayed around any non-compliant pair of persons in the display. Pedestrians strolling down the street in a pre-recorded video is used to validate this project.

INTRODUCTION

When a deadly coronavirus pandemic (Covid-19) comes out, if there is no effective treatment, The people will be concerned about the virus's proliferation. Due to an upsurge in the number of cases reported throughout the world, the World Health Organization (WHO) has designated Covid-19 a pandemic. Many countries have implemented quarantine measures in try to limit the epidemic, and during this critical period, the government has asked individuals to remain at home. As the number of patients continues to increase, there is currently no viable medicine or vaccine for the virus. Although scientists, medical organizations, as well as researchers have been working diligently to develop effective medications or vaccinations to combat this dangerous virus, there has not been a clear report of success at the time of this study, and there is no specific treatment or vaccine against the virus. Recommendations for the prevention or treatment of this new disease so that necessary preventive measures can be taken on a global scale to prevent illness from spreading. The community looks for new ideas on how to stop the virus from spreading. Public health officials, like the Centers for Disease Control and Prevention (CDC), should make it known that one of the most effective way to stop Covid-19 from transmitting is to prevent direct contact with other people. To flatten the curve, people all across the world are physically distancing themselves from the Covid-19 epidemic.

RELATED WORK

Many researchers in the medical and pharmaceutical areas have sought to find a cure for the COVID-19 infectious disease, but no definitive remedy has yet been discovered. Controlling the spread of such an unknown disease, on the other hand, is a different challenge. Controlling the prevalence might even be a contributing element, according to a slew of research using various implementation techniques, and social distance is an effective way to reduce transmission and limit the virus's spread in society. This section summarises the relevant research on human detection using deep learning. Deep learning is used in a large number of recent research on object categorization and detection. The state-of-the-art review focuses mostly on these machine learning-based object identification research projects. Within the computer vision job for categorization and localisation of its shape in video images, human detection is usually recognized as an object detection. Deep learning has demonstrated a strong trend in multi-class sensing and detection in computers, and has excelled on difficult datasets. Nguyen et al. offered a complete evaluation of ultra-modern on the latest improvement and demanding situations of human detection. Human descriptors, device learning algorithms, occlusion, and real-time detection are all included in the survey. Techniques that use a deep convolutional neural network (CNN) for visual recognition have been proven to outperform others on a variety of picture recognition benchmarks. Deep CNN is a multilayer perceptron neural network deep learning technique that contains many convolutional layers, sub-sampling layers, and fully linked layers. Later, for each object categorization supported by its dataset, the burden within the layers within the networks is trained. The CNN model was one of each of the categories in deep learning that are supervised feature learning approaches robust in detecting the thing in a variety of settings for object detection within the image. Because of the latest high-performance technology and enormous datasets like ImageNet, CNN has had a lot of success in large-scale image classification problems. In terms of spec, algorithms, and novel concepts, various CNN models for object detection with object localisation have been suggested. CNN models such as AlexNet, VGG16, InceptionV3, and ResNet-50 have been trained in recent years to understand outstanding finally lands up in beholding. Deep learning's effectiveness in vision is due to its neural network topology, which is capable of self-constructing the article descriptor and learning high-level attributes that aren't immediately provided within the dataset. In terms of accuracy and speed, these cutting-edge object detectors with deep learning have their advantages and disadvantages. Within the image, the item may have different spatial placements and aspect ratios. As a result, real-time object identification algorithms based on the CNN model, such as R-CNN and YOLO, are becoming increasingly popular, were created to recognise multi-classes in images over a wide range of regions. In terms of both speed and accuracy, YOLO (You Only Look Once) is the most popular strategy for deep CNN-based object recognition. Figure 1 depicts an example of the YOLO version.

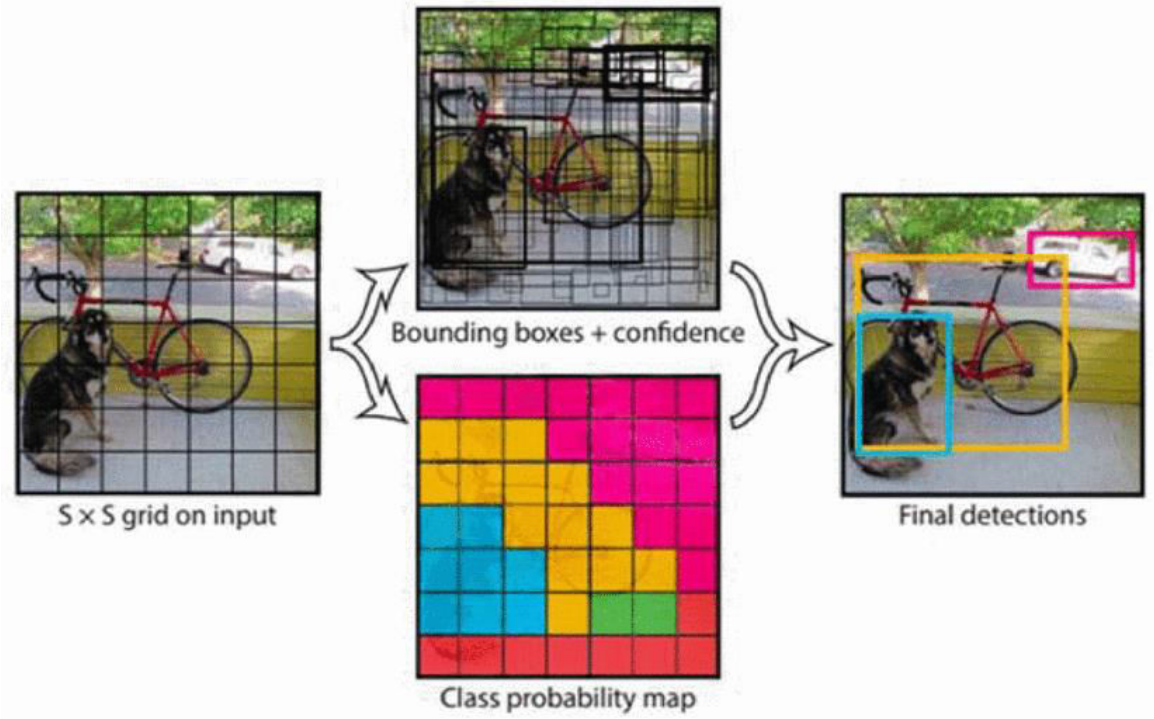

Figure 1. YOLO model for object detector pipeline illustration.

METHODOLOGY

This method for detecting social distancing was created to identify the security distance between persons in public areas.

During this research, the deep CNN approach and computer vision techniques were used. Initially, an open-source object identification network based on the YOLOv3 method was used to recognise the human in the video frame, with the model being trained using an open-source dataset called COCO Dataset. The COCO (Common Objects in Context) dataset might be used to do vast object recognition, segmentation, key-point identification, and labeling.

We use different tools to understand this:

1. **Jupyter Notebook: -** Jupyter Notebook is an open-source online tool that lets you record and share documents with realtime code, equations, visualisations, and narrative content. Data cleansing and transformation, numerical modelling, statistical modelling, data visualisation, machine learning, and other applications are only a few examples.
2. **Numpy: -** Python's NumPy module is the foundation for scientific computing. It's a Python package with a multi-dimensional array object, various derived objects (like arrays and masked arrays), and many procedures for fast array operations, such as mathematics, logic, form operations, sorting, selection, I/O, Discrete Fourier Transform, basic algebra, basic statistical operations, random modelling, and so on.
3. **Pandas:** - Pandas is an open-source data analysis and processing tool depending on the Python programming language that is quick, powerful, versatile, and simple to use.

4. **Tensorflow:** - TensorFlow is an open-source machine learning platform with a lot of features. It features a large and flexible ecosystem of tools, libraries, and community resources, allowing academics to promote the most recent breakthroughs in machine learning while developers can quickly design and expand machine learning applications.
5. **Keras:** - Keras is a human-centric API, not a machine-centric one.
6. Keras reduces cognitive burden by providing a consistent and straightforward API, minimising the amount of user involvement necessary for typical use cases, and providing clear and responsive error signals. There's also a lot of documentation and developer instructions.
7. **Convolutional neural network (CNN): -** A convolutional neural community (CNN) is a form of artificial neural community utilized in photograph reputation and processing this is mainly designed to procedure pixel data. CNN has their "neurons" organized greater like the ones of the frontal lobe, the region chargeable for processing visible stimuli in people and different animals.
8. **Matplotlib:** - Matplotlib is a charting library that is part of the NumPy large data processing tool that may be used in the Python programming language. Matplotlib embeds graphics in Python programmes using object-oriented APIs.
9. **OpenCV:** - OpenCV is a free and open-source computer vision, machine learning, and image processing toolkit. Python, C++, Java, and a variety of additional programming languages are supported by OpenCV. It can analyse photos and video to identify items, people, and perhaps human handwriting. When it's combined with a variety of libraries, such as Numpy, a highly optimised library for numerical operations, the arsenal of weapons available to you expands, i.e., operations that you can do in Numpy can be combined with OpenCV.

All varieties of CCTV cameras, with resolutions ranging from VGA to Full HD, may be integrated and used with real-time performance. To adjust the camera, it records a video picture at a specified point and transforms the live footage from top to bottom into a two-dimensional view for more precise measurement estimate. This method assumes that pedestrians walk on the same plane in the video frame. Remove it from the frame and turn it into a top-down perspective. The position of each person is evaluated from the top view. Each and every distance shorter than the permitted distance between two persons in Los Angeles is shown by a red line as a warning, according to the prescribed minimum distance. Each and every distance shorter than the permitted distance between two persons in Los Angeles is shown by a red line as a warning, according to the prescribed minimum distance. Figure 2 shows the process flow of the social distance detecting tool.

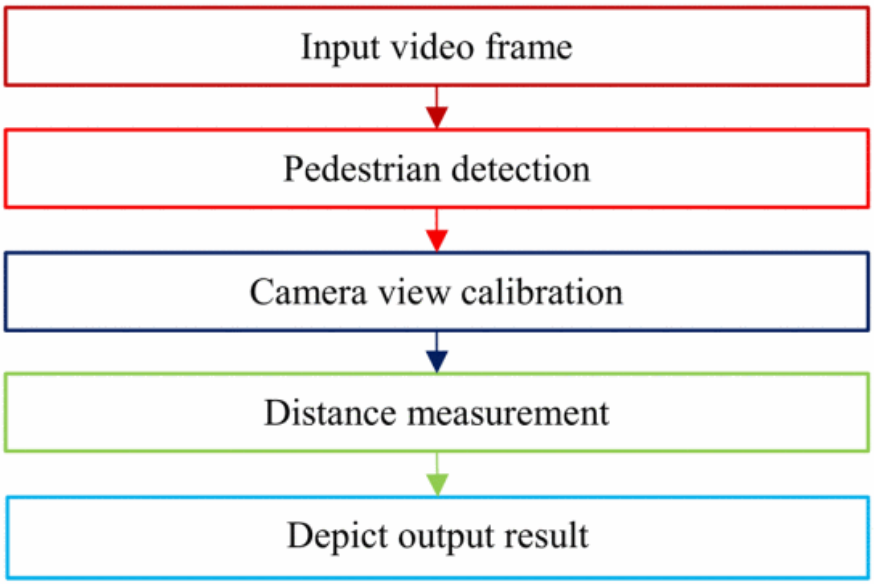

Figure 2. Pipeline for detecting social distance.

There are several steps to this model.

A. People Detection

The deep CNN model is a proposed object recognition method that reduces computational complexity by formulating recognition with a single regression problem. The goal is to develop a model to identify people (people) with different types of problems, such as differences in clothes. , Regardless of whether it is blocked or not, under different lighting conditions, you can pose from long distances and close distances. We rely on cutting-edge research; however, we have developed our own unique human classifier and trained our model on a wide variety of data sets. Before delving into the technical details, let us first understand the most advanced object recognition technology, and then introduce our human body recognition model. Human recognition framework.

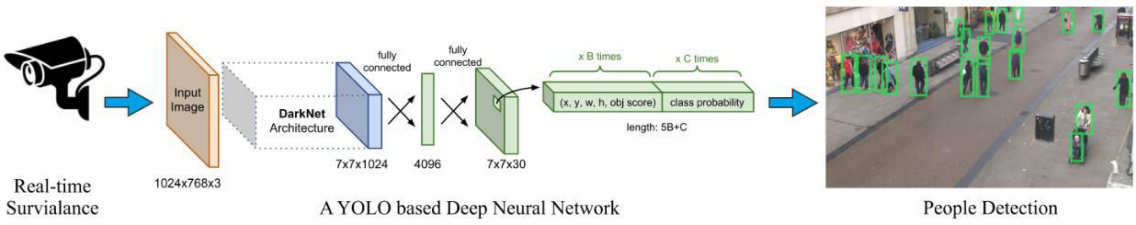

Figure 3. The overall structure of the people detection module.

B. People Tracking

Following the detecting phase, the next stage is to track down and identify everyone. We use simple online and real-time tracking technology (SORT) as the basis of the Kalman filter and Hungarian

optimization technology for tracking people. Based on the simultaneous current measurement and a mathematical model of human movement, The Kalman filter estimates a person's position at time t + 1.

C. Distance Measurement

The position of every individual's bounding box (x, y, w, h) in the specific viewpoint is calculated and transformed to a top-down view at this stage of the pipeline. The lower centre point of the bounding box is used to assess the bottom view. The top view may be used to compute the distance between each pair of individuals, and these distances are scaled using the zoom factor obtained from the camera view calibration. (x1, y1) or (x2, y2) are the two individuals in the image. The distance between the two individuals is denoted by the letter d. can be calculated using the following formula:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

We use an Open-source model to compare our model. These models are:

1. EfficientDet: - EfficientDet is a Google brain team-created object recognition model, and the research article describing the technique was published on July 27, 2020.
It's the successor of EfficientNet, and with a brand-new neural network design option for an object recognition test, it already outperforms RetinaNet, Mask R-CNN, and YOLOv3 architectures. We utilise EfficientDet D5, EfficientDet D6, and EfficientDet D7 for this.
2. Faster R-CNN with Resnet-101 v1: - Faster R-CNN with Resnet-101 (v1) is initialized by Imagenet sorting checkpoints. Trained on the COCO 2017 dataset.
3. **Mask R-CNN Inception V2: -** On the COCO dataset, Mask R-CNN Inception V2 was trained. This model is used to segment instances of objects.
4. **CenterNet HourGlass 104: -** The CenterNet object recognition model with hourglass spine, trained on the COCO 2017 dataset and training images.
5. **YOLO: -** The YOLO algorithm provides better performance in all aspects discussed, as well as a high frame rate for real-time use. The YOLO algorithm is a regression-based algorithm. It does not select interesting parts of the image, but predicts the category and frame. Separate the entire image in one pass of the algorithm.

RESULT AND DISCUSSION

We tried comparing a number of models (including a custom-made model) for detecting objects in a given image (or a frame in case of videos) and the comparison of the models is as follows (mAP = mean Average Precision): -

Model	COCO mAP	Speed(ms)
YOLO	44.2	51
EfficientDet D5	49.7	222
EfficientDet D6	50.5	268
EfficientDet D7	51.2	325
Faster RCNN Resnet 101 v1	31.8	55

Custom Model	20.2	122
CenterNet HourGlass 104	44.5	197

From our results, we observe the accuracy in identifying objects using YOLO is somewhat lesser than EfficientDet and CenterNet but the speed at which YOLO identifies (or classifies) the objects make it a much better and practical model.

So, for the final detector, we used the weights we get by using the YOLO model for detecting the social distancing in any given video/image.

REFERENCES

1. World Health Organisation. WHO Corona-viruses Disease Dashboard (August 2020). Available athttps://covid19.who.int/table.
2. Adlhoch, C.et al. Considerations relating to social distancing measures in response to the corvid-19 epidemic (2020). European Centre for Disease Prevention and control, technical report.
3. Implementation of Mitigation Strategies for Communities with Local COVID-19, [online] Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019.
4. Implementation of Mitigation Strategies for Communities with Local COVID-19 Transmission, [online] Available: https://www.cdc.gov/coronavirus/2019-ncov/downloads/community-mitigation-strategy.pdf.
5. D.T. Nguyen, W. Li and P.O. Ogunbona, "Human detection from images and videos: A survey", Pattern Recognition, vol. 51, pp. 148-75, 2016.
6. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified real-time object detection", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016.
7. Fong, M. W.et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—social distancing measures. Emerge. infectious diseases26, 976, DOI: 10.3201/eid2605.190995(2020).
8. J. Redmon and A. Farhadi, Yolov3: An Incremental improvement, 2018.
9. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified real-time object detection", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016.
10. R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587, 2014.
11. K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014.