SOCIAL DISTANCING DETECTION.

A Project Report

Submitted in partial fulfillment of the Requirements for the award of the Degree Of

MASTER OF SCIENCE (INFORMATION TECHNOLOGY)

By

MOHAMMED HUSSAIN IQBAL SHAIKH

1161

Under the esteemed Guidance of

Prof. SAIMA SHAIKH

DEPARTMENT OF INFORMATION TECHNOLOGY
MAHARASHTRA COLLEGE OF ARTS, SCIENCE & COMMERCE
(Affiliated to University of Mumbai)
MUMBAI - 400008.
MAHARASHTRA
2020-2021

INDEX

Maharashtra College Of Arts Science And Commerce

```
┌─────────────────────────────────────────┐
│                                           │
│          PROJECT OVERVIEW                 │
│                                           │
└─────────────────────────────────────────┘
```

INTRODUCTION

To limit the spread of an infectious disease, for instance, Covid-19, is to practice social distancing. This is not a new concept, as most societies have been aware of the value of keeping away from people who are

suffering from an infection for many generations. The objective is to reduce transmission, delaying the epidemic peak, reducing the size of the epidemic peak, and spreading cases over a longer time to relieve pressure on the healthcare system. It is an action taken to minimize contact with other individuals. It has been suggested that maintaining a distance of approximately 2 metres from another individual result in a marked reduction in transmission of most flu virus strains, including COVID-19.

In practice, this means that avoiding close proximity to other people will aid in slowing the spread of infectious diseases. Social distancing is one of the non-pharmaceutical infection control actions that can stop or slow down the spread of a highly contagious disease. The virus that causes COVID-19 is currently spreading easily from person-to-person. When a healthy person comes into contact with respiratory droplets from coughs or sneezes of an infected person, they are can catch the infection.

The World Health Organization (WHO) states that "COVID-19 is transmitted via droplets and fomites during close unprotected contact between an infector and infectee. A fomite is an object or material which is likely to carry infection, such as clothes, utensils, and furniture. Therefore, transmission of the infection can be avoided by staying away from other people as well as from touching infected fomites. Social distancing aims to decrease or interrupt transmission of COVID-19 in a population[1] by minimizing contact between potentially infected individuals and healthy individuals, or between population groups with high rates of transmission and population groups with no or low levels of transmission.

"The only way to prevent the spread of COVID-19 is Social Distancing. Keeping a safe distance from each other is the ultimate way to prevent the spread of this disease (at least until a vaccine is found)."

So this got me thinking – I want to build a tool that can potentially detect where each person is in real-time, and return a bounding box that turns red if the distance between two people is dangerously close. This can be used by governments to analyze the movement of people and alert them if the situation turns serious.

Objectives:

The main objectives are aims to decrease or interrupt transmission of COVID-19 in a population by minimising contact between potentially infected individuals and healthy individuals, or between population

Maharashtra College Of Arts Science And Commerce

Social Distancing Detection

groups with high rates of transmission and population groups with no or low levels of transmission. This study aim specifically the following listed below:

1) This study aims to support the reduction of the coronavirus spread and its economic costs by providing an AI-based solution to automatically monitor and detect violations of social distancing among individuals.

2) We develop a robust deep neural network (DNN) model for people detection, tracking, and distance estimation called DeepSOCIAL. In comparison with some recent works in this area, such as [15], we offer faster and more accurate results.

3) We perform a live and dynamic risk assessment, by statistical analysis of spatio-temporal data from the people movements at the scene. This will enable us to track the moving trajectory of people and their behaviours, to analyse the ratio of the social distancing violations to the total number of people in the scene, and to detect high-risk zones for short- and long-term periods.

4) We back up the validity of our experimental results by performing extensive tests and assessments in a diversity of indoor and outdoor datasets which outperform the state-of-the-arts.

5) The developed model can perform as a generic human detection and tracker system, not limited to social-distancing monitoring, and it can be applied for various real-world applications such as pedestrian detection in autonomous vehicles, human action recognition, anomaly detection, and security systems.

The virus that causes COVID-19 is currently spreading easily from person-to-person. When a healthy person comes into contact with respiratory droplets from coughs or sneezes of an infected person, they are can catch the infection.

"Social Distancing Detection uses existing IP cameras and CCTV cameras combined with Computer Vision to detect if people are at a specific distance and adhere to social distancing or not."

Purpose and scope:

- **Purpose:**

Maharashtra College Of Arts Science And Commerce

Social Distancing Detection

Social distancing is a term applied to certain actions that are taken by Public Health officials to stop or slow down the spread of a highly contagious disease. The Health Officer has the legal authority to carry out social distancing measures. Since these measures will have considerable impact on our community, any action to start social distancing measures would be coordinated with local agencies such as cities, police departments and schools, as well as with state and federal partners

The proposed methodology including the model architecture and our object detection techniques, tracking, and red-zone prediction algorithm will be proposed, experimental results and performance of the system will be investigated against the state-of-the-art. This project adopts a combination of lightweight neural network and Single Shot Detector(SSD) with transfer learning technique to achieve the balance of resource limitations and recognition accuracy so that it can be used on real-time video surveillance to monitor public places to detect if persons wearing maintaining safe social distancing.

Our solution uses neural networking models to analyze Real-Time Streaming Protocol (RTSP) video streams using OpenCV and MobileNet V2.We mix the approach of modern-day deep learning and classic projective geometry techniques which not only helps to meet the real-time requirements, but also keeps high prediction accuracy. If the person detected as not following the covid-19 safety guidelines, violation alerts will be send to the control center at state police headquarters for taking further action. It allows automating the solution and enforces the wearing of the mask and follows the guidelines of social distancing. This project was created to run on python and the accuracy obtained was between 85% and 95%.

Human detection using visual surveillance system is an established area of research which is relying upon manual methods of identifying unusual activities, however, it has limited capabilities Although human detection is an ambitious goal, due to a variety of constraints such as low-resolution video, varying articulated pose, clothing, lighting and background complexities and limited machine vision capabilities, wherein prior knowledge on these challenges can improve the detection performance.

Detecting an object which is in motion, incorporates two stages: object detection and object classification. The primary stage of object detection could be achieved by using background subtraction, optical flow and spatiotemporal filtering techniques.

- **Scope:**

Maharashtra College Of Arts Science And Commerce

Social Distancing Detection

The above suggests that dealing with interpersonal distances means to deal with evolutionary, developmental and cultural forces that shape, to a significant extent, our everyday life. As a consequence, the role of technologies for the analysis of such distances becomes crucial during pandemics, given that they must mediate between the forces above, responsible for the human tendency to get too close to avoid contagion, and the pressure of prophylactic measures,artificially designed to fight a pathogen inaccessible to our senses and cognition.

In recent years, object detection techniques using deep models are potentially more capable than shallow models in handling complex tasks and they have achieved spectacular progress in computer vision. Deep models for person detection focus on feature learning contextual information learning, and occlusion handling.

Deep learning object detection models can now mainly be divided into two families:

a. two-stage detectors such as R-CNN[9], Fast R-CNN and Faster R-CNN and their variants

b. one-stage detectors such as YOLO and SSD. In two-stage detectors detection is performed in stages, in the first stage computed proposals and classified in the second stage into object categories. However, some methods, such as YOLO, SSD MultiBox, consider detection as a regression issue and look at the image once for detection.

we are using Single Shot Detector MultiBox(SSD) which seems to be a good choice for real-time object detection and the accuracy trade-off is also very little. SSD uses the VGG-16 model pre-trained on ImageNet as its basic model to extract useful image feature. At the top of VGG16, SSD adds several convolutional feature layers of decreasing sizes.

Therefore, many other technology-based solutions such as [11,12] and AI related research such as [13–15] have tried to step in to help the health and medical community in copping with COVID-19 challenges and successful social distancing practices.

In such situations, Artificial Intelligence can play an important role in facilitating social distancing monitoring. Computer Vision, as a sub-field of Artificial Intelligence, has been very successful in solving various complex health care problems and has shown its potential in chest CT-Scan or X-ray based COVID-19 recognition and can contribute to Social-distancing monitoring as well.Besides, deep neural networks enable us to extract complex features from the data so that we can provide a more accurate understanding of the images by analysing and classifying these features. Examples include diagnosis, clinical management and treatment, as well as the prevention and controlof COVID-19.-

Maharashtra College Of Arts Science And Commerce

System Analysis

Maharashtra College Of Arts Science And Commerce

System Analysis

EXISTING SYSTEM.

Estimating the SSD requires one to solve a few classical Computer Vision and Social Signal Processing tasks, namely, scene geometry understanding, person detection/body pose estimation and social distance characterization. Indeed, the geometry of the scene is important to define a local reference system for measuring inter-personal distances. Clearly, a second and important task is the detection of people in the scene in possibly crowded environments. Once the target people are correctly localised in a scene, their distance can be locally estimated in order to realize if the mutual distance is lower than a threshold (e.g. 1m or 2m). Afterwards, this metric information is analysed to output whether there is a violation of the protocol or the short distance is due to a legitimate situation, e.g., a family walking together.

- Although the planarity constraint might not hold for the entire image, SSD has to do a local estimation of proximity for which is safe to relax the scene being piece-wise planar.
- Self-calibration approaches highly rely on the existence of a Manhattan world (e.g. vanishing points are detectable) or pedestrian walking in straight trajectories, which limit the applicability of such methods. Estimating depth from single image might be a viable option, but a metric reference is still needed.
- Estimating a metric reference for precise SD measures from images is an issue. Such reference extracted from pedestrians might be unreliable given the variations in anthropometric characteristics. Reasoning on the geometrical context of the scene (e.g., object shapes) can lead to a more robust metric estimate.

Tracking Technologies:

Since the onset of coronavirus pandemic, many countries have used technology-based solutions, to inhibit the spread of the disease. For example, some of the developed countries, such as South Korea and India, use GPS data to monitor the movements of infected or suspected individuals to find any possible exposure among the healthy people.

This may also help other people to maintain a safe distance from the infected person. Some law enforcement agencies use drones and surveillance cameras to detect large-scale rallies and have carried out regulatory measures to disperse the population. human detection using wireless signals by identifying phase differences and change detection in amplitude wave-forms. However, this requires multiple receiving antennas and can not be easily integrated in all public places.

Maharashtra College Of Arts Science And Commerce

AI-Based Research:

The utilisation of Artificial Intelligence, Computer Vision, and Machine Learning, can help to discover the correlation of high-level features. For example, it may enable us to understand and predict pedestrian behaviours in traffic scenes, sports activities, medical imaging, or anomaly detection, by analysing spatio-temporal visual information and statistical data analysis of the images sequences.

Among AI-Health related works, some researchers have tried to predict the sickness trend of specific areas, to develop crowd counting and density estimation methodologies in public places, or to determine the distance of individuals from the popular swarms using a combination of visual and geo-location cellular information. However, such research works suffer from challenges such as skilled labour or the cost of designing and implementing the infrastructures.

On the other hand, recent advances in Computer Vision, Deep Learning, and pattern recognition, as the sub-categories of the AI, enable the computers to understand and interpret the visual data from digital images or videos. It also allows computers to identify and classify different types of objects.

People Detection:

A Camera collects the input video sequences, and passes them to our Deep Neural Network model. The output of the model would be the detected people in the scene with their unique localisation bounding boxes. The objective is to develop a robust human (people) detection model, capable of dealing with various types of challenges such as variations in clothes, postures, at far and close distances, with/without occlusion, and under different lighting conditions.

Modern DNN-based object detectors (such as those listed in Figure 3) consist of three sections:input module and related operations such as augmentation, a backbone for extracting features and a head for predicting classes and location of objects in the output.

We have summarised a comprehensive list model design options including input augmentations, state-of-the-art core object detection modules (i.e., activation functions, backbone feature extractors, neck, and head). The table, offers a variety of possible choices for neck, head and other sub-modules (depending on the requirements of the model). However, we mainly focus on the requirement of this research.

Maharashtra College Of Arts Science And Commerce

Inputs and Training Datasets:

In order to have a robust detector, we would require a set of rich training datasets. This should include people with a variety in gender and age (man, women, boy, girl) with millions of accurate annotation and labelling. We selected two large datasets of MS COCO and Google Open Image dataset that satisfy the above-mentioed expectations, by providing more than 3.7 million annotated people.

On the other hand, in batch normalisation, the batch size reduction causes noisy estimation of mean and variance. To address this issue, we considered the normalised values of the previous k iterations instead of a single mini-batch. This is similar to Cross-Iteration Batch Normalisation (CBM).

A set of possible activation functions for BoF. We also investigated the performance of our model against ReLU, Leaky ReLU, SELU, Swish, Parametric RELU, and Mish. Our preliminary evaluations confirmed the same results provided by Misra for our human detection application. The Mish (Equation (1)) activation function converged towards the minimum loss, faster than Swish and ReLU, with higher accuracy. The result was consistent especially for diversity of parameter initialisers, regularisation methods, and lower learning rate values. Mish:

$$f(x) = x.\tanh(\text{softplus}(x))$$
$$= x.\tanh(ln(1 + e^x))$$

with derivations:

$$f'(x) = \frac{e^x \omega}{\delta^2}$$

as a self regularised non-monotonic activation function, where

$$\omega = 4(x + 1) + 4e^{2x} + e^{3x} + (4x + 6)$$

Maharashtra College Of Arts Science And Commerce

PROPOSED SYSTEM.

The proposed system focuses on how to identify the person on image/video stream whether the social distancing is maintained or not with the help of computer vision and deep learning algorithm by using the OpenCV, Tensor flow library.

Approach

1. Detect humans in the frame with yolov3.
2. Calculates the distance between every human who is detected in the frame.
3. Shows how many people are at High, Low and Not at risk.

Deep learning based framework is proposed that utilizes object detection and tracking models to aid in the social distancing remedy for dealing with the escalation of COVID-19 cases. In order to maintain the balance of speed and accuracy, YOLO v3. alongside the Deepsort are utilized as object detection and tracking approaches while surrounding each detected object with the bounding boxes. Later, these bounding boxes are utilized to compute the pairwise L2 norm with computationally efficient vectorized representation for identifying the clusters of people not obeying the order of social distancing.

Furthermore, to visualize the clusters in the live stream, each bounding box is color-coded based on its association with the group where people belonging to the same group are represented with the same color. Each surveillance frame is also accompanied with the streamline plot depicting the statistical count of the number of social groups and an index term (violation index) representing the ratio of the number of people to the number of groups. Furthermore, estimated violations can be computed by multiplying the violation index with the total number of social groups.

This system can be used in real-time applications requiring a secure monitoring of social distance between people and the detection of face masks for safety purposes due to the outbreak of Covid-19. Deploying our model to edge devices for automatic monitoring of public places could reduce the burden of physical monitoring, which is why we choose to use this architecture. This system can be integrated with edge device for use in airports, railway stations, offices, schools and public places to ensure that public safety guidelines are followed. It is also important to emphasizes here that VSD is a simpler problem than estimating every metric distances among people in any position in the image. Estimating social distance is necessary when two or more pedestrians get close enough for triggering the necessity of a measure. At this point, a local reference system can be estimated and metric references can be leveraged by using surrounding objects and the height of the local cluster of people.

Maharashtra College Of Arts Science And Commerce

SYSTEM DESIGN

Maharashtra College Of Arts Science And Commerce

SYSTEM DESIGN

DATA FLOW DIAGRAM

Activity diagram are a loosely defined diagram technique for showing workflow of stepwise and action, with support for choice, iteration and concurrency. An activity diagrams shows the over control. Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system.

The control flow is drawn from one operation to another. This flow can be sequential, branched, or concurrent. Activity diagrams deal with all type of flow control by using different elements such as fork, join, etc.

The main element of an activity diagram is the activity itself. An activity is a function performed by the system. After identifying the activities, we need to understand how they are associated with constraints and conditions.

The basic purposes of activity diagrams is similar to other four diagrams. It captures the dynamic behavior of the system. Other four diagrams are used to show the message flow from one object to another but activity diagram is used to show message flow from one activity to another.

Activity is a particular operation of the system. Activity diagrams are not only used for visualizing the dynamic nature of a system, but they are also used to construct the executable system by using forward and reverse engineering techniques. The only missing thing in the activity diagram is the message part.

It does not show any message flow from one activity to another. Activity diagram is sometimes considered as the flowchart. Although the diagrams look like a flowchart, they are not. It shows different flows such as parallel, branched, concurrent, and single.

The basic usage of activity diagram is similar to other four UML diagrams. The specific usage is to model the control flow from one activity to another. This control flow does not include messages.

Activity diagram is suitable for modeling the activity flow of the system. An application can have multiple systems. Activity diagram also captures these systems and describes the flow from one system to another. This specific usage is not available in other diagrams. These systems can be database, external queues, or any other system.

Maharashtra College Of Arts Science And Commerce

Social Distancing Detection

We will now look into the practical applications of the activity diagram. From the above discussion, it is clear that an activity diagram is drawn from a very high level. So it gives high level view of a system. This high level view is mainly for business users or any other person who is not a technical person.

This diagram is used to model the activities which are nothing but business requirements. The diagram has more impact on business understanding rather than on implementation details.
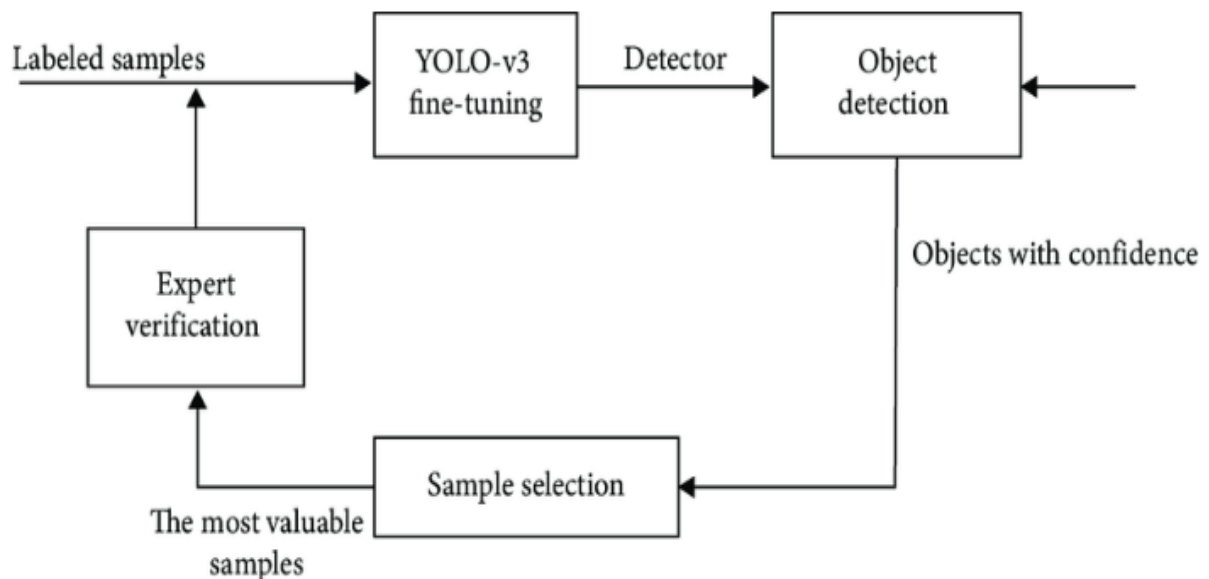

Activity diagram is basically a flow chart to represent the flow form one activity to another form .Activity can be described as an operation of the system.

So the control flow is drawn from one operation to another. This flow can be sequential, and concurrent. Activity diagram deals with all type of flow control by using different elements like fork.

Maharashtra College Of Arts Science And Commerce

SYSTEM ARCHITECTURE DESIGN

System Architecture Design is a major roll for any software development process.
The product will fail if Arts system architecture is not properly designed.

Maharashtra College Of Arts Science And Commerce
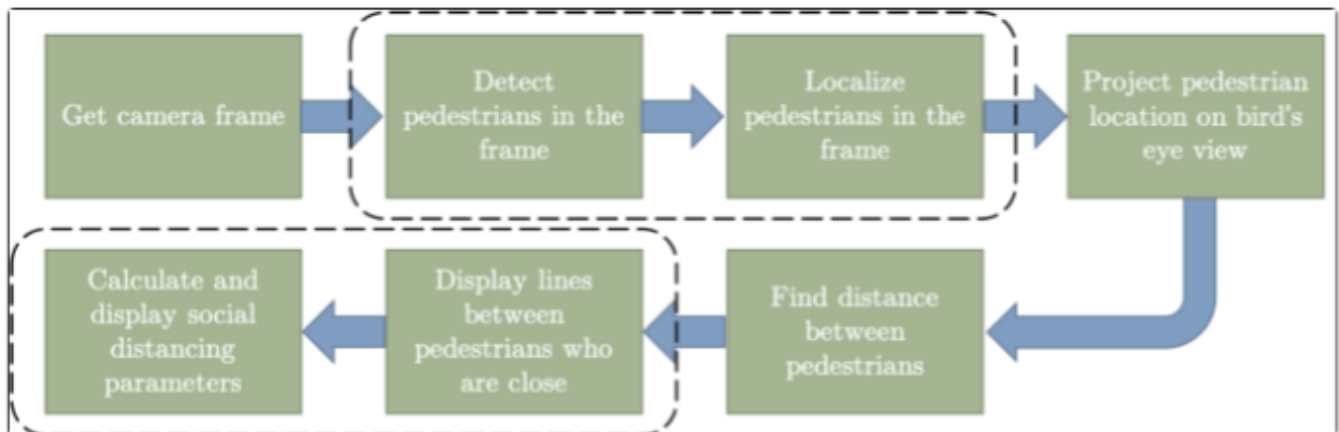
PROJECT IMPLEMENTATION

4.1 Project Implementation Technology:

In this project we use Python language, CNN, Deep learning, YOLO v3. we discuss the steps taken to train our human detection model and the investigated datasets to train the model, followed by experimental results on people detection, social distancing measures, and risk infection assessment.
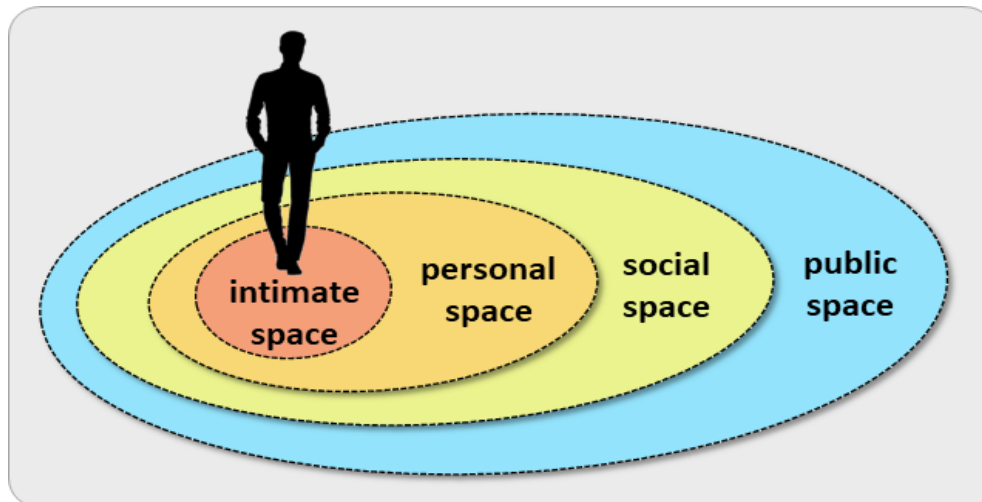
SCENE GEOMETRY UNDERSTANDING:

The task of measuring social distancing from images requires the definition of a (local) metric reference system. This problem is strongly related to the single view metrology topic as we consider the most common case of a fixed camera. An initial solution for estimating inter-personal distances requires the identification of the ground plane where people walk. Such ground plane serves in many video-surveillance systems to visualize the scene as a bird's eye view for ease of visualization and data statistics representation. Many works impose the assumption that the ground plane is planar. Then, the problem is to estimate a homography given some reference elements (e.g., known objects or manual measurements) extracted from the scene or using the information of detected vanishing points in the image

PERSON DETECTION AND POSE ESTIMATION:

Person detection has reached impressive performance in the last decade given the interest in automotive industry and other application fields Real-time approaches can now estimate people pose even in complex scenarios and even reconstruct the 3D mesh of the person body. The majority of the approaches estimate not only people location as a bounding box but also 2D stick-like figures, so conveying a schematic representation of the poseCapturing diffused small SDs with Computer Vision requires to localize multiple people, realizing the hardest scenario for pedestrian detection techniques. Specific pedestrian detection techniques have been designed to work in crowded scenes [29], [55], [97], [106], where saliency-based masks are often preferred to skeleton-based representations. When the image resolution becomes too low to spot single people, regression-based approaches are employed [12], [15], [54],[80], [92], [104], [111],

Maharashtra College Of Arts Science And Commerce

providing in some case density measures[73], [86], [87], [110]. This information, merged with a geometric model of the scene, will directly lead to a measure of the average SSD in the field of view. Obviously, regression or density-based approaches cannot provide additional cues on pose which are highly important for capturing human actions and interactions. To fill this gap, ad-hoc approaches individuate general crowd activities, classifying them as normal or not (e.g. a person collapsing and many people getting close).

VISUAL SOCIAL DISTANCE CHARACTERIZATION:

Social distances should be complemented with additional contextual information to understand whether social distancing rules are actually being broken or not, consequently suggesting the most proper reaction. a multi-layer pipeline, which will be detailed in the following, indicating which information can be accessed with the current Computer Vision technology. The deeper the layer (indicated by a darker color), the finer the visual analysis which is needed and the harder the corresponding request for Computer Vision.

In the former case, a high percentage of small SDs is characterizing the monitored area: This may occur at a crossing intersection or walking in a corridor of an airport. Here, many persons stand close, without an explicit will, possibly for a short (seconds) period. Computer Vision helps here providing robust approaches for pedestrian detection and counting.

SINGLE SHOT DETECTOR (SSD):

In this research, single shot detector (SSD) [63] is also used as another object identification method to detect people in realtime video surveillance system. As discussed earlier, faster R-CNN works on region proposals to create boundary boxes to indicate objects, shows better accuracy, but has slow processing of frames per second (FPS). For real-time processing, SSD further improves the accuracy and FPS by using multiscale features and default boxes in a single process. It follows the principle of the feed-forward convolution network which generates bounding boxes of fixed sizes along with a score.

YOLO v3:

For object detection, another competitor of SSD is YOLO [40]. This method can predict the type and location of an object by looking only once at the image. YOLO considers the object detection problem as a regression task instead of classification to assign class probabilities to the anchor boxes. A single convolutional network simultaneously predicts multiple bounding boxes and class probabilities.

Each prediction is monitored by computing objectness, boundary box regressor and classification scores. a schematic description of the YOLOv3 architecture is presented.
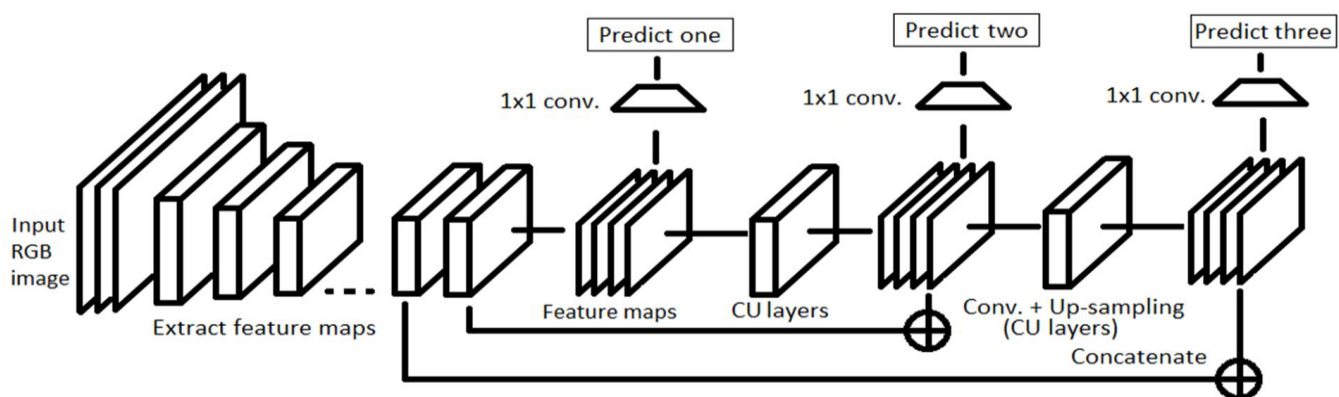
DISTANCE CALCULATION:

Now we have bounding box for each person in the frame. We need to estimate person location in frame. i.e we can take bottom center point of bounding box as person location in frame. Then we estimate (x,y) location in bird's eye view by applying transformation to the bottom center point of each person's bounding box, resulting in their position in the bird's eye view. Last step is to compute the bird's eye view distance between every pair of people and scale the distances by the scaling factor in horizontal and vertical direction estimated from calibration.

The second step to detect pedestrians and draw a bounding box around each pedestrian. To clean up the output bounding boxes, we apply minimal post-processing such as non-max suppression (NMS) and various rule-based heuristics, so as to minimize the risk of over fitting.

WORKING:

Running the program will give you frame (first frame) where you need to draw ROI and distance scale. To get ROI and distance scale points from first frame Code to transform perspective to Bird's eye view (Top view) and to calculate horizontal and vertical 180 cm distance in Bird's eye view.

ROI and Scale points' selection for first frame. The second step to detect pedestrians and draw a bounding box around each pedestrian. To detect humans in video and get bounding box details.

Now we have bounding box for each person in the frame. We need to estimate person location in frame. i.e we can take bottom center point of bounding box as person location in frame. Then we estimate (x,y) location in bird's eye view by applying transformation to the bottom center point of each person's bounding box, resulting in their position in the bird's eye view. To calculate bottom center point for all bounding boxes and projecting those points in Bird's eye view. Last step is to compute the bird's eye view distance between every pair of people (Point) and scale the distances by the scaling factor in horizontal and vertical direction estimated from calibration.

Maharashtra College Of Arts Science And Commerce

Social Distancing Detection

Lastly we can draw Bird's Eye View for region of interest (ROI) and draw bounding boxes according to risk factor for humans in a frame and draw lines between boxes according to risk factor between two humans. Red, Yellow, Green points represents risk to human in Bird's eye view. Red: High Risk, Yellow: Low Risk and Green: No Risk. Red, Yellow lines between two humans in output tell they are violating social distancing rules.

SOCIAL DISTANCING EVALUATIONS:

We considered the midpoint of the bottom edge of the detected bounding boxes as our reference points (i.e., shoes' location). After the IPM, we would expect to have the location of each person, in the homogeneous space of BEV with a linear distance representation. Any two people A &B with the Euclidean distance of smaller than r (i.e., the set restriction) in the BEV space were considered as contributors in social distancing violation:

Depending on the type of overlapping and the violation assessment criteria, we define a violation detection function V with the input parameters of a pixel metrics x, the set safe distance of r (e.g., 6 f t or 2 m), the position of the query human A, and the closest surrounding person B.

Maharashtra College Of Arts Science And Commerce

ADVANTAGES, LIMITATIONS & APPLICATION

Maharashtra College Of Arts Science And Commerce

ADVANTAGES:

When there is no threat from a pandemic, there is no need for physical distancing, but some form of social distancing can still be useful for maintaining a scientist's intellectual independence either by developing a thick skin, or by avoiding the risk of bruises in the first place by avoiding social friction.

SDD helps to understand the reason why some people stand close, distinguishing whether they are socializing among themselves, or if they are interacting with the environment (as, for example, looking at a timetable in the airport), thus suggesting the most proper countermeasures to ensure SD (e.g., rising an audio alarm to discourage social interactions or putting markers into the floor so that people can watch the time table while keeping the right distances).

The advantages of SDD appears to be of particular importance since at the moment social distancing rules have to be expressed in simplistic terms (e.g., people have to be at least 2 meters far from one another) requiring one to distinguish between the intention (avoid contagion) and the rule (keep a minimum interpersonal distance). Such a distinction, evident to humans, poses a real and new challenge to a computational algorithm for SDD that could solve the problem by leveraging, for instance, the use of contextual information.

Differently, the number of false alarms would be so high that any benefit resulting from the use of technology would be canceled. In the following, we will discuss in detail the VSD problem and its connection to the Computer Vision and Social Signal Processing research domains. Starting from a geometrical point of view, i.e. estimating inter-personal distances between people from an image, we show that this first step does not take into account scene and social context.

For this reason, a further stage needs to elaborate the geometrical VSD in order to interpret whether the violation of the distance is a real cause of alert or an acceptable situation (e.g., a family, walking). Then, we contextualise the SDD in different application domains and we finally conclude with a description of the possible ethical shortcomings of SDD.

The statistical analysis showing the total number of social groups displayed by same color encoding and a violation index term computed as the ratio of the number of people to the number of groups.

Maharashtra College Of Arts Science And Commerce

LIMITATION:

Social distancing detector *did not* leverage a proper camera calibration, meaning that we could not (easily) map distances in pixels to actual measurable units (i.e., meters, feet, etc.).

Therefore, the first step to improving our social distancing detector is to utilize a proper camera calibration.

Doing so will yield better results and enable you to compute actual measurable units (rather than pixels).

Secondly, you should consider applying a top-down transformation of your viewing angle.

Applying a perspective transform or using stereo computer vision would allow you to get a more accurate representation of social distancing with OpenCV. While more accurate, the engineering involved in such a system is more complex and isn't always necessary. From there, you can apply the distance calculations to the top-down view of the pedestrians, leading to a better distance approximation.

My third recommendation is to improve the people detection process.

OpenCV's YOLO implementation is quite slow *not* because of the model itself but because of the additional post-processing required by the model.

To further speedup the pipeline, consider utilizing a Single Shot Detector (SSD) running on your GPU — that will improve frame throughput rate *considerably.*

To wrap up, I'd like to mention that there are a number of social distancing detector implementations you'll see online — the one I've covered here today should be considered a *template* and *starting point* that you can build off of.

APPLICATION:

Python:

Python is an interpreted, high-level, general-purpose programming language. Created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable use of significant whitespace.

Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically typed and garbage-collected.

Maharashtra College Of Arts Science And Commerce

Social Distancing Detection

Python is often described a language due to its comprehensive standard library. Python uses dynamic typing and a combination of reference counting and a cycle-detecting garbage collector for memory management. It also features dynamic name resolution, which binds method and variable names during program execution.

Artificial Intelligence:

Artificial intelligence (AI), sometimes called machine intelligence, is intelligence demonstrated by machines, Leading AI define the field as the study of "intelligent agents" any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals. Colloquially, the term "artificial intelligence" is often used to describe machines (or computers) that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving". As machines become increasingly capable, tasks considered to require "intelligence" are often removed from the definition of AI, a phenomenon known as the AI affect "AI is whatever hasn't been done yet. For instance, optical character recognition is frequently excluded from things considered to be AI, having become a routine technology. Modern machine capabilities generally classified as AI include understanding human speech, competing at the highest, autonomously operating cars, intelligent routing in content delivery networks, and military simulations.

OpenCV:

OpenCV Python is a library of Python bindings designed to solve computer vision problems. Python is a general purpose programming language started by Guido van Rossum that became very popular very quickly, mainly because of its simplicity and code readability. It enables the programmer to express ideas in fewer lines of code without reducing readability. Compared to languages like C/C++, Python is slower Compared to languages like C/C++, Python is slower. That said, Python can be easily extended with C/C++, which allows us to write computationally intensive code in C/C++ and create Python wrappers that can be used as Python modules.

This gives us two advantages: first, the code is as fast as the original C/C++ code (since it is the actual C++ code working in background) and second, it easier to code in Python than C/C++. OpenCV-Python makes use of Numpy, which is a highly optimized library for numerical operations with a MATLAB-style syntax. All the OpenCV array structures are converted to and from Numpy arrays.

This also makes it easier to integrate with other libraries that use Numpy such as SciPy and Matplotlib. OpenCV-Python is the Python API for OpenCV, combining the best qualities of the OpenCV C++ API and the Python language.

Maharashtra College Of Arts Science And Commerce

TENSORFLOW:

Tensor Flow is an open source library for fast numerical computing. It was created and is maintained by Google and released under the Apache 2.0 open source license. The API is nominally for the Python programming language, although there is access to the underlying C++ API. Unlike other numerical libraries intended for use in Deep Learning like Theano, Tensor Flow was designed for use both in research and development and in production systems, It can run on single CPU systems, GPUs as well as mobile devices and large scale distributed systems of hundreds of machines. Computation is described in terms of data flow and operations in the structure of a directed graph.

Nodes: Nodes perform computation and have zero or more inputs and outputs. Data that moves between nodes are known as tensors, which are multi-dimensional arrays of real values.

Edges: The graph defines the flow of data, branching, looping and updates to state. Special edges can be used to synchronize behavior within the graph, for example waiting for computation on a number of inputs to complete.

Operation: An operation is a named abstract computation which can take input attributes and produce output attributes. For example, you could define an add or multiply operation.

YOLOV3:

YOLOv3 is the latest variant of a popular object detection algorithm YOLO – You Only Look Once. The published model recognizes 80 different objects in images and videos, but most importantly it is super fast and nearly as accurate as Single Shot MultiBox (SSD). First, it divides the image into a 13×13 grid of cells.

The size of these 169 cells varies depending on the size of the input. For a 416×416 input size that we used in our experiments, the cell size was 32×32. Each cell is then responsible for predicting a number of boxes in the image. For each bounding box, the network also predicts the confidence that the bounding box actually encloses an object, and the probability of the enclosed object being a particular class.

Most of these bounding boxes are eliminated because their confidence is low or because they are enclosing the same object as another bounding box with very high confidence score. This technique is called non-maximum suppression.

Easy integration with an OpenCV application: If your application already uses OpenCV and you simply want to use YOLOv3, you don't have to worry about compiling and building the extra Dark net code.

OpenCV CPU version is 9x faster: OpenCV's CPU implementation of the DNN module is astonishingly fast. For example, Dark net when used with OpenMP takes about 2 seconds on a CPU for inference on a single image. In contrast, OpenCV's implementation runs in a mere 0.22 seconds! Check out table below.

Python support: Dark net is written in C, and it does not officially support Python. In contrast, OpenCV does.

CONCLUSION AND FUTURE WORK

Maharashtra College Of Arts Science And Commerce

CONCLUSION AND FUTURE WORK:

The emerging trends and the availability of intelligent technologies make us to develop new models that help to satisfy the needs of emerging world. So we have developed a novel social distancing detector which can possibly contribute to public healthcare. The model proposes an efficient real-time deep learning based framework to automate the process of monitoring the social distancing via object detection and tracking approaches, where each individual is identified in the real-time with the help of bounding boxes. Identifying the clusters or groups of people satisfying the closeness property computed with the help of Bird's eye view approach.

The number of violations is confirmed by computing the number of groups formed and violation index term computed as the ratio of the number of people to the number of groups. The extensive trials were conducted with popular state-of-the-art object detection models Faster RCNN, SSD, and YOLO v3, since this approach is highly sensitive to the spatial location of the camera, the same approach can be fine tuned to better adjust with the corresponding field of view.

This system works very effectively and efficiently in identifying the social distancing between the people and generating the alert that can be handled and monitored. we have presented the VSD problem as the estimation and characterization of inter-personal distances from images. Solving such problems allows a quick screening of the population for detecting potential behaviours that can cause a health risk, especially related to recent pandemic outbreaks.

We pointed out that VSD is not only a Computer Vision problem related to geometrical proxemics since people distancing has to be weighted given the social context in the current scene. Close relationships can allow closer interpersonal distances as well as being a caregiver of individuals with fragile conditions. We have shown that understanding such social context is a compelling problem in the literature of signal social processing that requires further research efforts for a reliable solution.

As the solution is intertwined with the decoding of social relationships from images, there are strong ethical and privacy concerns that need to be addressed with novel privacy-by-design solutions. Past this grievous global crisis, VSD has still an important role in several application fields thus providing a continuous source of interest in this new problem.

The solution has the potential to significantly reduce violations by real-time interventions, so the proposed system would improve public safety through saving time and helping to reduce the spread of coronavirus. This solution can be used in places like temples, shopping complex, metro stations, airports, etc.

Maharashtra College Of Arts Science And Commerce

Social Distancing Detection

The above mentioned use cases are only some of the many features that were incorporated as part of this solution. We assume there are several other cases of usage that can be included in this solution to offer a more detailed sense of safety. Several of the currently under development features are listed below in brief:

- **Coughing and Sneezing Detection:** Chronic coughing and sneezing is one of the key symptoms of COVID-19 infection as per WHO guidelines and also one of the major route of disease spread to non-infected public. Deep learning based approach can be proved handy here to detect & limit the disease spread by enhancing our proposed solution with body gesture analysis to understand if an individual is coughing and sneezing in public places while breaching facial mask and social distancing guidelines and based on outcome enforcement agencies can be alerted.

- **Temperature Screening:** Elevated body temperature is another key symptom of COVID-19 infection, at present scenario thermal screening is done using handheld contactless IR thermometers where health worker need to come in close proximity with the person need to be screened which makes the health workers vulnerable to get infected and also its practically impossible to capture temperature for each and every person in public places, the proposed use-case can be equipped with thermal cameras based screening to analyze body temperature of the peoples in public places that can add another helping hand to enforcement agencies to tackle the pandemic effectively.

Maharashtra College Of Arts Science And Commerce

Social Distancing Detection

Maharashtra College Of Arts Science And Commerce