An Approach to Detect Violation of Social Distancing Norms in Public Domain

M.B.Chandak[1]

[1]Professor and Head, Department of Computer Science and Engineering
Shri Ramdeobaba College of Engineering and Management, NAGPUR

Abstract

The COVID-19 virus is an ongoing global crisis with more than 247,000 deaths globally. The lack of a vaccine makes precautions against the virus an essential step. One of the best ways to observe these precautions is Social Distancing. This article proposes a deep learning based framework for automating the task of monitoring social distancing using surveillance video, utilizing AWS Object Rekognition for its object detection models to detect humans. Amazon Rekognition makes it easy to add image and video analysis to your applications using proven, highly scalable, deep learning technology that requires no machine learning expertise to use. Distance is calculated between these objects to find the high-risk rate, low-risk rate and safe distance rate. An analysis of this data is presented.

Keywords

Social Distancing, Object Detection, Object Tracking, COVID-19, Amazon Rekognition, Video Surveillance

1. INTRODUCTION

The COVID-19 pandemic has forced the global community to look for alternate ways to stop the spread of this virus, one of the most prominent being Social Distancing. This project aims to propose a solution to detect the violation of social distancing, using evidence from a video. Social distancing aims at reducing the physical contact between possibly infected individuals and healthy persons. As per the WHO norms[1], it is prescribed that people should maintain at least 6 feet, or 2 meters, of distance among each other in order to follow social distancing.

A recent study[2] indicates that social distancing is an important containment measure and essential to prevent this coronavirus, because people with mild or no symptoms may carry corona infection and can infect others. Proper social distancing is the best way to reduce infectious physical contact, hence reducing the infection rate[3]. This helps "flatten the curve", meaning the reduced peak can match with the available healthcare infrastructure and offer better facilities to the patients. Motivated by this, in this present work, authors are attempting to check and compare the performance of popular object detection and tracking schemes in monitoring the social distancing.

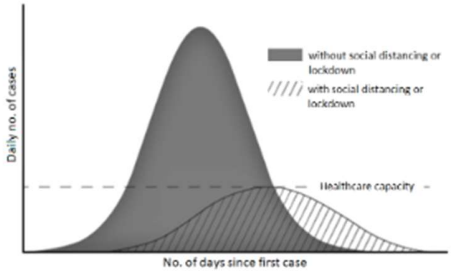

Fig. 1: An outcome of social distancing as the reduced peak of the epidemic and matching with available healthcare capacity.

Respiratory diseases are infectious where the rate and mode of transmission of the causing virus are the most critical factors to be considered for the treatment or ways to stop the spread of the virus in the community. Several medicine organizations and pandemic researchers are trying to develop vaccines for COVID-19, but still, there is no well-known medicine available for treatment. Hence, precautionary are taken by the whole world to restrict the spread of infection.

Motivated by this, in this present work people are attempting to check and compare the performance of popular object detection and tracking schemes in monitoring the social distancing. Rest of the paper structure is organized as follows: Section II presents the recent work proposed in this field of study, followed by the state-of-the-art object detection and tracking models in Section III. Later, in Section IV framework is proposed to monitor social distancing. In Section V experimentation and the corresponding results are discussed, accompanied by the outcome. In Section VI the use case is discussed. In Section VII the future scope and challenges are discussed and lastly Section VIII presents the conclusion of the present research work.

Manuscript received September 5, 2025
Manuscript revised September 20, 2025
https://doi.org/10.22937/IJCSNS.2025.25.9.5

2. Background Study And Related Work

Social distancing is surely the most trustworthy technique to stop the spreading of infectious disease, with this belief, in the background of December 2019, when COVID-19 emerged in Wuhan, China, it was opted as an unprecedented measure on January 23, 2020. As we all understand, social distancing though essential but economically painful measures to flatten the infection curve.

Since the pandemic began, many countries have been using technology-based solutions in various capacities to contain the outbreak[4],[5]. In India, the government is using the Arogya Setu App, which worked with the help of GPS and Bluetooth to locate the presence of COVID-19 patients in the vicinity area, helping others keep a safe distance from the infected people [6].

Human detection using visual surveillance system is an already established area of research, however, it has limited capabilities[7]. A variety of constraints, like low-resolution videos, varying articulated pose, clothing, lighting and background complexities, and limited machine vision capabilities, can be improved upon using prior information, thus improving the performance of the system[8].

3. Object Detection and Technologies Used

A. Object Detection

Object Detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class, such as humans, buildings, or cars, in digital images and videos[9]. Every object class has some own special features, which are used in the classification of detected object into specific classes. For example, when the system wants to detect a circular-looking object, a center is calculated and objects from a particular distance from this center are selected. For face identification, features include eyes, nose, lips, skin colour, and distance between eyes. For human detection, this concept of face detection is extended, adding other features like hands, legs, torso, feet, the overall body shape of a human etc.

Object detection has applications in many areas of computer vision, including image retrieval and video surveillance.

B. Deep Learning

Deep Learning is an artificial intelligence function that imitates the workings of the human brain in processing data and creating patterns for use in decision making[10]. Deep learning is a subset of machine learning, with networks capable of learning unsupervised from unstructured data.

Deep learning uses a hierarchy of artificial neural networks to carry out the process of machine learning. These networks resemble a brain, with the connection of neuron nodes reminding one of a spider web. This hierarchical function enables processing of data with a nonlinear approach.

C. Amazon Rekognition

Object detection and tracking of model is done using Amazon Rekognition. Amazon Rekognition API, and the services it provides, can identify objects, people, text, scenes, and activities, as well as detect any inappropriate content. Amazon Rekognition also provides highly accurate facial analysis, face comparison, and face search capabilities. These include detection, analysis, and comparison of faces for a wide variety of use cases, including user verification, cataloguing, counting the number of objects present, and public safety scenarios like tracking an identified assailant.

Key Features

a. Labels

With Amazon Rekognition, you can identify thousands of objects (such as bike, telephone, building), and scenes (such as parking lot, beach, city). When analysing video, you can also identify specific activities such as "delivering a package" or "playing soccer".

b. Custom Labels

With Amazon Rekognition Custom Labels, you can extend the detection capabilities of Amazon Rekognition to extract information from images that is uniquely helpful to your business. For example, you can find your corporate logo in social media, identify your products on store shelves, classify your machine parts in an assembly line, or detect your animated characters in videos.

c. Content Moderation

Amazon Rekognition helps you identify potentially unsafe or inappropriate content across both image and video assets and provides you with detailed labels that allow you to accurately control what you want to allow based on your needs. Use Amazon A2I to enhance the accuracy of Amazon Rekognition image moderation predictions using human review.

d. Text Detection

In photos and videos, text appears very differently than neat words on a printed page. Amazon Rekognition can read skewed and distorted text to capture information like store names, forced narratives overlaid on media, street signs, and text on product packaging.

e. Celebrity Recognition
You can quickly identify well known people in your video and image libraries to catalog footage and photos for marketing, advertising, and media industry use cases.

f. Personal Protective Equipment (PPE) detection
With Amazon Rekognition, you can analyse images from your on-premises cameras at scale to automatically detect if persons in images are wearing Personal Protective Equipment (PPE) such as face covers (face masks), hand covers (gloves), and head covers (helmets) and whether the protective equipment covers the corresponding body part (nose for face covers, head for head covers, and hands for hand covers).

g. Face detection and Analysis
With Amazon Rekognition, you can easily detect when faces appear in images and videos and get attributes such as gender, age range, eyes open, glasses, facial hair for each. In video, you can also measure how these face attributes change over time, such as constructing a timeline of the emotions expressed by an actor.

h. Face search and verification
Amazon Rekognition provides fast and accurate face search, allowing you to identify a person in a photo or video using your private repository of face images. You can also verify identity by analyzing a face image against images you have stored for comparison.

Amazon Rekognition is based on the highly scalable, deep learning technology and includes an API for analyzing videos that can quickly analyze any video file stored in Amazon S3. Amazon Simple Storage Service (S3) is an object storage service that offers industry-leading scalability, data availability, security, and performance. It can be used to store and protect any amount of data for a range of use cases. It also provides management features that can be used for organising data, and also provides a durable system for worldwide use.

Fig. 1 A sample photo showing people detected and bounding boxes constructed

Amazon Rekognition Video can return the bounding box for common object labels such as people, and can be used to find the exact locations of objects in an image, count instances of detected objects, or to measure an object's size using bounding box dimensions.

D. MySql Database

A database is a collection of data, usually stored in electronic form. A database is typically designed so that it is easy to store and access information. A good database is crucial to any company or organisation. This is because the database stores all the pertinent details about the company. The various reasons why database are important are- manages large account of data, accurate, easy to update data, security of data, integrity of data, easy to search data.

The database used in this project is MySql. Distributed and supported by Oracle, MySql is a database system that runs on an online web server. It is free to download and use, fast, and reliable. It can work for both large scale and small scale applications. One of its biggest advantages is that it can compile on a number of platforms. It uses SQL, which is a standard language for storing, manipulating and retrieving data in databases.

The data is stored in form of tables. A table is a collection of related data, and it consists of columns and rows. Every data entry is listed in a row, and the table in relational databases is identified by a unique primary key. A row can have primary key set by one or more columns. The purpose of this key is to uniquely identify each row in the database. The practical result of this is that you can select a single row by just knowing its primary key. Columns are defined to hold a specific type of data, such as dates, numeric, or textual data. In the simplest of definitions a column is defined by its name and data type. The name is used in SQL statements when selecting and ordering data, and the data type is used to validate information stored.

The advantages of using a database system include reducing Data Redundancy, which removed duplicates, Data Sharing, where the users of the database can share the data among themselves and any remote users can also access the database simultaneously and share the data between themselves. Other advantages include Data Integrity, Data Security and Privacy, and also Backup and Recovery of data.

E. Flask

Flask is a lightweight WSGI web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. It has become one of the most popular Python web application frameworks. Flask offers suggestions, but doesn't enforce any dependencies or project layout. It is up to the developer to

choose the tools and libraries they want to use. There are many extensions provided by the community that make adding new functionality easy.

4. Proposed Idea

In the present article, an AWS 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, Amazon Rekognition is utilized as object detection and tracking approaches while surrounding each detected object with the bounding boxes. Later, these bounding boxes are utilized to compute single co-ordinate for identifying the clusters of people not obeying the order of social distancing.

A. Workflow

i. Upload

A video is uploaded, and stored in database.

ii. Process

a. The input video is converted into random frames.
b. Object detection algorithm is applied on each of these random frames.
c. Bounding boxes are created and their centre is calculated.
d. Distance algorithm is used to detect distance between detected objects. If the distance is less than the threshold value, violation is detected and the colour of the box changes to red. People in close range difference are marked as at-risk, and if the distance is higher than the threshold, safe distance is predicted.

iii. Analyze

a. These created frames are saved in the database, and accessed using the UI.
b. The count of all detected violations, risks and safe distance is also saved in the database.

B. Algorithm

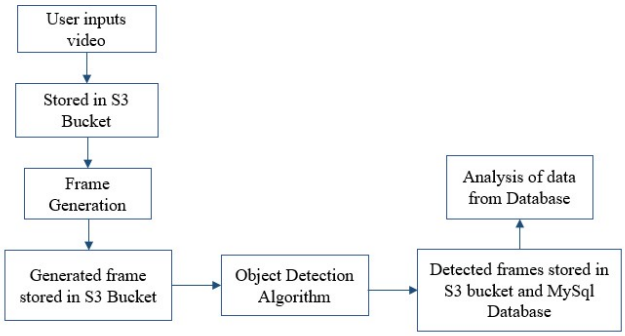

Fig. 2 Flowchart of Algorithm

The algorithm is as follows

1. User inputs video
2. This video is stored in Amazon S3 bucket (Fetch video). This is stored so that it can be easily accessed for frame generation algorithm.
3. Frame generation converts the video into frames.
4. Generated frames are stored in another S3 bucket (Frames undetected). These frames do not have object detection applied to them yet.
5. Object Detection algorithm applied on the frames stored in bucket frames undetected.
6. Frames with objects detected are stored in S3 bucket (frames detected) and MySql database.
7. Data is retrieved from database and analysis is performed.

C. User Interface

The system uses a WebApp for user interaction. It consists of a homepage, with miscellaneous information. It redirects to a page where user can upload their videos and view the violation detection. An analysis page is also provided.

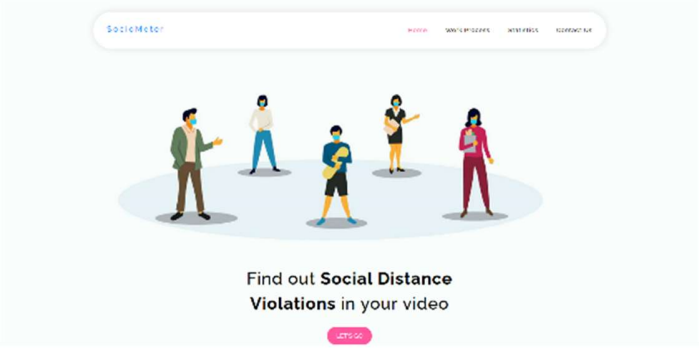

Fig. 3 Homepage of the WebApp

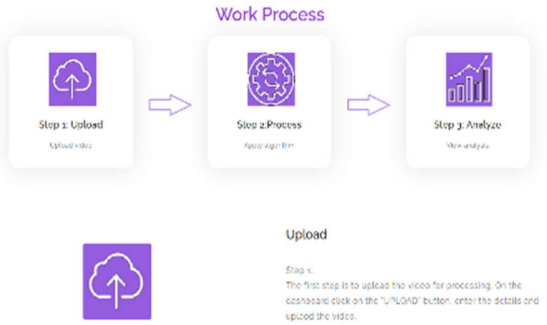

Fig. 4 Explanation of working of the project on homepage

D. Database Schema

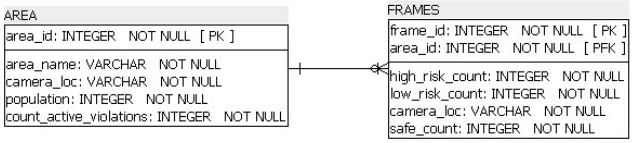

Fig. 5 Database Schema

The schema used for this project consists of two tables, Area and Frames
For Area, we have attributes,
area_id: primary key
area_name: name of area
camera_loc: location of camera (due to multiple cameras in each area, this gives the specific location of the video)
population: total population of the area
count_active_violations: total count of violations of the area till date
For Frames, we have attributes,
frame_id: primary key
area_id : foreign key
camera_loc: location of camera (common attribute)
high_risk_count: number of people under high risk due to social distancing violation
low_risk_count: number of people under low risk due to social distancing violation
safe_count: number of people not violating social distancing

5. OUTPUT

The input video is uploaded, processed and analysed. The output is displayed in the webapp. The database stores the frames and areas.

area_id	area_name	camera_loc	population	count_active_violation
100	Ramdaspeth	Cream Stone store	7000	19
101	Gittikhadan	Police Station	1500	88
102	Dhantoli	FSB	5000	47
103	Dharampeth	Haldiram	9000	181
104	Laxmi Nagar	Ashok Hotel	6000	100

Fig. 6 Database table Area showing area data

frame_id	area_id	camera_loc	high_risk_count	low_risk_count	safe_count
132	104	camera98	15	3	4
133	104	camera98	17	5	1
134	104	camera98	22	0	1
135	104	camera98	20	1	1
136	104	camera98	14	3	2
137	101	camera98	15	3	4
138	101	camera98	17	5	1
139	101	camera98	22	0	1
140	101	camera98	20	1	1
141	101	camera98	14	3	2
142	100	camera98	4	7	0
143	100	camera98	7	3	1
144	100	camera98	4	0	5
145	100	camera98	2	3	4
146	100	camera98	2	5	3
147	103	camera98	15	1	2
148	103	camera98	3	4	3
149	103	camera98	7	5	0

Fig. 7 Database table Frames showing all frames generated with detected objects along with the counts

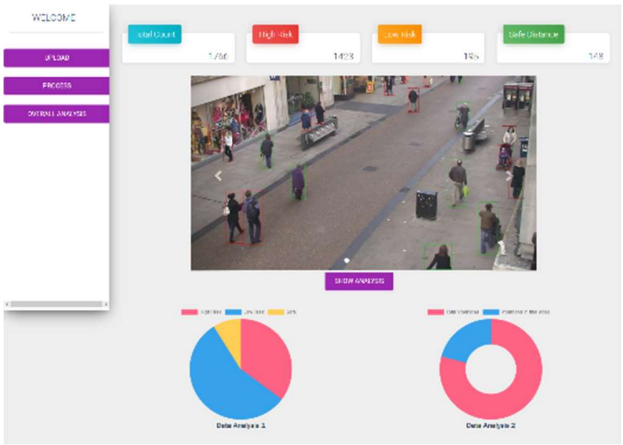

Fig. 8 Unique frames of the video, with violations, risks, and safe distance detected. The analysis of given video is also presented.

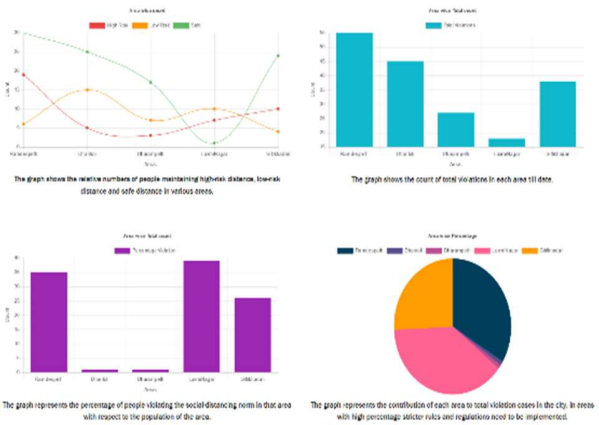

Fig. 9 Overall analysis of all the information in our database, like in different area, is presented.

6. USE CASES

Performing the detection of social distancing violation on multiple sample input videos, we get the use cases and their analysis.

TABLE I USE CASES

Sr.no	Video Length (in seconds)		Frames generated	Average Processing time (in seconds)
1	20		5	30
2	60		21	136

7. FUTURE SCOPE AND CHALLENGES

The project has large future scope, predominantly to be used as a system to check social distancing and apply it in a public, or office environment. The system can be modified to be used for various purposes, like invigilation during examinations, and even further for social purposes like illegal traffic violations.

Since this application is intended to be used in any working environment, accuracy and precision are highly desired to serve the purpose. Higher number of false positive may raise discomfort and panic situation among people being observed. There may also be genuinely raised concerns about privacy and individual rights which can be addressed with some additional measures such as prior consents for such working environments, and maintaining transparency about its fair uses within limited stakeholders.

8. CONCLUSIONS

The article proposes an efficient framework to automate the process of monitoring the social distancing via object detection and tracking approaches, where each individual is identified with the help of bounding boxes. The generated bounding boxes aid in calculation of distances and categorising them as high risk, low risk and safe, for every frame. Analysis of all these rates is provided.

REFERENCES

[1] L. Hensley, “Social distancing is out, physical distancing is inheres how to do it,” *Global News–Canada*, 27 March 2020.
[2] ECDPC, “Considerations relating to social distancing measures in response to COVID-19 second update. [Online]. Available: https://www.ecdc.europa.eu/en/publications-data/considerations
[3] M. W. Fong, H. Gao, J. Y. Wong, J. Xiao, E. Y. Shiu, S. Ryu, andB. J. Cowling, “Non Pharmaceutical measures for pandemic influenza in non healthcare settings social distancing measures”, 2020.
[4] S. K. Sonbhadra, S. Agarwal, and P. Nagabhushan, “Target specific meaning of covid-19 scholarly articles using one-class approach,” 2020.
[5] N. S. Punn and S. Agarwal, “Automated diagnosis of covid-19 with limited posteroanterior chest x-ray images using fine-tuned deep neural networks,” 2020.
[6] O. website of Indian Government, “Distribution of the novel coronavirus-infected pneumonia Aarogya Setu Mobile App,” https://www.mygov.in/aarogya-setu-app/, 2020.
[7] N. Sulman, T. Sanocki, D. Goldgof, and R. Kasturi, “How effective is human video surveillance performance?” in2008 19th InternationalConference on Pattern Recognition. IEEE, 2008, pp. 1–3.
[8] X. Wang, “Intelligent multi-camera video surveillance: A review,”Pat-tern recognition letters, vol. 34, no. 1, pp. 3–19, 2013.
[9] Dasiopoulou, Stamatia, et al. "Knowledge-assisted semantic video object detection." IEEE Transactions on Circuits and Systems for Video Technology 15.10 (2005): 1210–1224.
[10] Investopedia [Online]. Available: https://www.investopedia.com/terms/ d/deep-learning
[11] AWS User guide. [Online]. Available: https://docs.aws.amazon.com/ rekognition/latest/dg/rekognition-dg.pdf
[12] UI template [Online]. Available: https://www.creative-tim.com