COVID-19 Contact Tracing Using Computer Vision

Jethva Utsav
17110064
utsav.jethva@iitgn.ac.in

Madhav Tiwari
17110079
madhav.tiwari@iitgn.ac.in

Shweta Pardeshi
17110097
pardeshi.shweta@iitgn.ac.in

Prakash R
17110109
prakash.r@iitgn.ac.in

Abstract

As we know, COVID 19 has affected 100 million people to date. A key strategy for interrupting chains of transmission of SARS-CoV-2 and reducing COVID-19-associated mortality is to use contact tracing. It's impossible for humans to monitor every camera to identify potential contact. Hence, we plan to develop automated software that can be used to identify potential contacts. Our model processes the video and detects the people in each frame. Then we perform tracking and re-identification(ReID) to identifies entities in different frames. To remove redundancy, we perform ReID again and cluster similar entities. With this, we generate a graph that represents entities that were in contact. From this, we can identify the potential carrier of COVID-19 if someone turns out to be COVID positive.

1. Problem Statement

The COVID-19 (Coronavirus) outbreak originated from China in December 2019 and it has been spreading quickly across numerous nations. This virus has been declared as a pandemic by WHO because of its rapid spreading worldwide and also WHO declared coronavirus as a global risk. As of now, countries like India have stuck hard by the second wave of coronavirus. Every day India is making new records of covid positive patients and the number of deaths due to covid. There is a huge shortage of oxygen cylinders and even beds and ventilators in hospitals. Various nations have also taken some remarkable measures to contain the virus and stop it from spreading. Various countries have imposed a lockdown to forestall social gatherings and to hinder the spread of the virus which spreads through respiratory droplets as well as from the air.

These "social distancing" norms are one of the effective ways through which the virus can be contained. Contact tracing has widely been adopted to control the spread of Coronavirus-2019 (COVID-19).[5] It enables to identify, assess, and manage people who have been exposed to COVID-19, thereby preventing its further transmission. Contact tracing apps (like "Aarogya Setu" in India) rely on documented patient data to identify the infected. They leverage Bluetooth to share anonymized IDs with neighboring devices running the same app. If it spots an "infected" ID in the vicinity, the user is notified and his/her contacts in the vicinity are recursively traced.

Considering the seriousness of the current situation we have come up with a new robust method of contract tracing through video. Our problem statement is to identify the people who have violated the social distancing norms and with whom they have interacted.

2. Work Contribution

Name	Contribution to the Project
Utsav Jethva	Curation of Test dataset and Evaluation metrics
Madhav Tiwari	Object detection and Graph implementation
Prakash R	Implementation of ReID Models
Shweta Pardeshi	Tracking implementation and distance estimation

3. Related Work

Contact Tracing have been implemented multiple times in the past for use cases such as surveillance from CCTV camera but there are not much implementation on contact tracing for a disease. Contact tracing is not a sin-

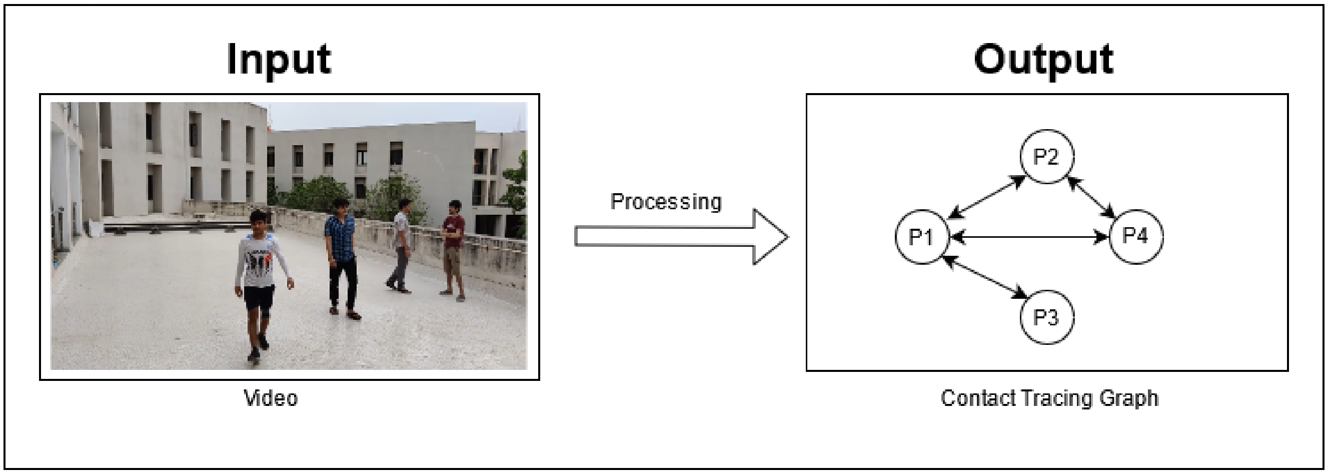

Figure 1. Problem Statement

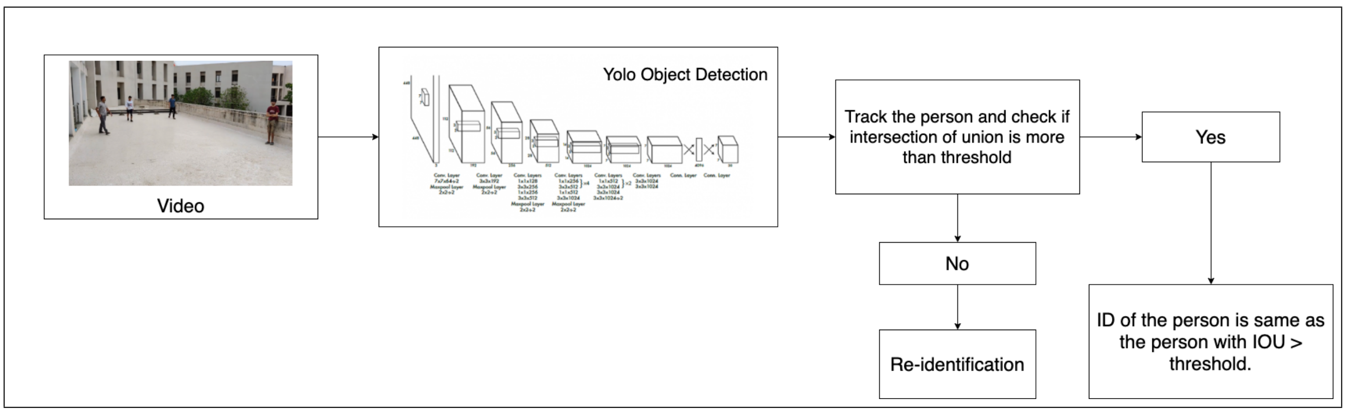

Figure 2. Algorithm Flowchart

gle problem, it has multiple sub problems such as person re-identification and distance estimation. Typically, methods for re-identification include two components: a method for extracting features from input images, and a metric for comparing those features across images. Research on re-identification usually focuses either on finding an improved set of features ([1], [2]), finding an improved similarity metric for comparing features ([3]), or a combination of both([6]).

4. Our Approach

4.1. Input

Any video can be provided as input to our algorithm. Primarily premise is that we are detecting people in one camera view. So only one video can be given as input. To make a case for re-identification, we have combined random parts of a 90-sec video to make a 15-sec video. In this way, we need to re-identify for better accuracy, and it allows us also to create a robust model.

4.2. YOLO

We have used a pre-trained YOLO model and weights for the detection of people in the video. YOLO takes a video as input and converts the video into a number of frames. Then on these frames, a single CNN is applied which divides the image into grids. Based on the grids, bounding boxes are predicted with a confidence score, and the boxes with a confidence score less than the threshold are eliminated. For each frame, the output of YOLO is the coordinates of the bounding box of detected people with a probability score.[7]

4.3. Tracking

After detecting the object as a person, we start tracking the object with respect to its boundary boxes. The YOLO Object detector recognizes all the human objects in the current frame and returns the boundary box of all of them. We then check if the Intersection of Union (IOU) is more than the given threshold. IOU is defined as the intersection of 2

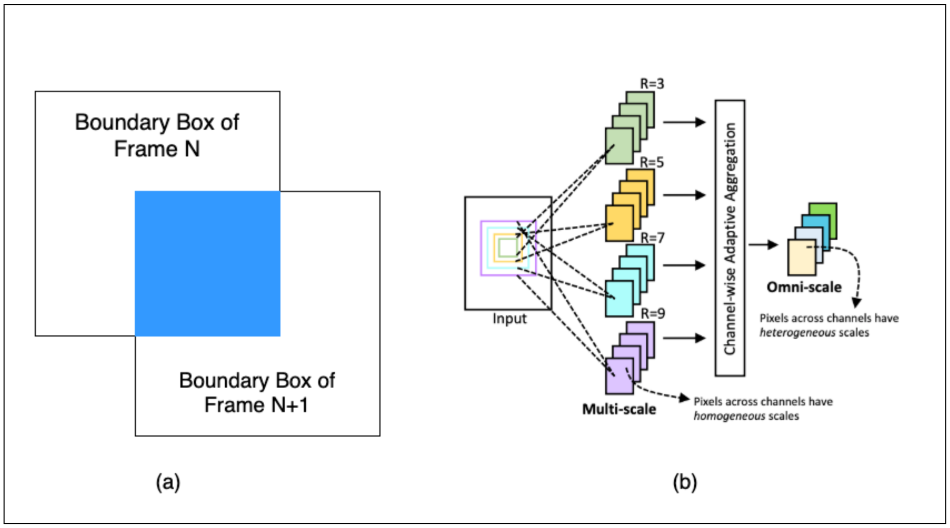

Figure 3. Model Architecture

Figure 4. Re-Identified Images

boundary boxes divided by the boundary box of the current frame. In Figure 3, as we can see the blue colour box is the intersection which is then divided by the area of the boundary box of N+1 frame. Basic intuition to using this metric is that the speed at which a person moves is reasonable; hence between consecutive frames, a person can't move very further. We find the best IOU of the present frames with the previous one, and if that is not larger than the threshold, we perform re-identification.

4.4. ReID

Person re-identification is the task of associating images of the same person taken from different cameras or from the same camera on different occasions. We are using re-identification here to develop a more sound and robust model. This helps us to generalize the model for any video. For this purpose, we use Torchreid [9], a software library built on PyTorch that allows fast development and end-to-end training and evaluation of deep re-ID models. We use cross-domain ReID models as there are generalized for any find of images not particular to a dataset. We achieve state-of-the-art results by using OSNet[10]. We scan the person who was not identified in tracking with some of the previously identified and sorted entities. If there is a match is above the threshold, then the person is tagged accordingly. If not, the person is assumed to be new.

4.5. Reducing redundancy

It may be the case that the two people identified before were the same and weren't identified because of some reasons. We sample out a few frames of each person identified and the run re-identification algorithm again to resolve this issue. If there the similarities are above the threshold, we start building a graph with edges between similar identities. From the graph, we perform clustering of connected components. This gives us a cluster of similar identities. Then, we identify all the similar entities by cluster id.

4.6. Graph

While parsing the frame for tracking and re-identification, we also calculate the distance between centroids of the bounding boxes. If it is less than 150 pixels, we assume that they are not following social distancing and add them to our violation list. We then re-tag them again using cluster id. Finally, we produce a graph with respect to time. The edges indicate that two identifies have been in contact.

5. Observations

Person Re-Id is a challenging task due to the similarity in objects. For example, people with identical features,clothes or colours. In public places, there are many different people wearing similar clothes, this makes it difficult to distinguish between them. Other major challenges of person re-identification in visual surveillance is that the appearance of the same person may change significantly across different camera views because of the great cross-camera variations in illumination, viewpoint, random inter-object occlusion and complex background clutter especially in crowded public spaces.[4] The videos are recorded by non-overlapping cameras under different environments hence using face detection is not useful for this task. Most existing person re-identification methods depend only on the spatial appearance information. The performance of state-of-the-art Person Re Identification models on benchmarks approaches saturation.[8]

We tried multiple supervised pre-trained models from torchreid library for person re-identification, 'osnet_x1_0' this worked the best on our test video. These models were using the entire appearance of the person instead of only using face data to distinguish between people. We observed that some models were performing better when we used euclidean distance metric to calculate distance between them while some performed better when used cosine similarity.

Since, we don't have large computational power, we kept the frame rate lower. This resulted in a lower value of intersection over union (IOU) metric between two consecutive frames. While performing a person ReID using IOU, lower value of IOU resulted in detection of more number of distinct people than number of people present in the video. Hence, we were not able to perform person re identification using this method accurately.

	No. of Frames	No. of Objects
	36	4
	24	3
	13	2
Total	73	242

Table 1. Frame and Objects details

Person id	Number of Objects
0	14
1	10
2	9
3	44
7	10
48	11

Table 2. Re-identification using IOU

6. Results

We filmed a short video of around 30 seconds in IITGN hostel area which we used as a test dataset throughout this project. There are 4 people in the test dataset, some of them are walking around, some are standing, and throughout the video some of them indulge in talking with each other. Our goal is to create a contact tracing graph of people who might have violated social distancing norms and are potential COVID-19 spreaders. We can then use that graph to analyse the spread of COVID-19 in case any of the person gets infected with COVID-19. We have broken down our end goad into three sub problem. First we do re identification, then we analyse social distancing violations and then we create the contact tracing graph. We have evaluated each step with respect to different evaluation metrics.

6.1. Test Dataset details

In our test dataset we have total 73 frames with objects (person) appearing for a total of 242 times. From table 11, we can see 36 frames have four objects in it, 24 frames have 3 objects in it and 13 frames have 2 objects in it. Out of these 73 frame there are multiple cases of social distancing violations in many of the frames and we tried to detect it.

6.2. Evaluating Re-identification

Evaluation matrix for re-identification will the accuracy with which the model is able to correctly re-identify the person in each frames.

6.2.1 Intersection over Union (IOU)

Initially, We started detecting re-identification using basic intersection over Union method IOU between two consecutive frames. We kept threshold of 0.5 for IOU. We got a

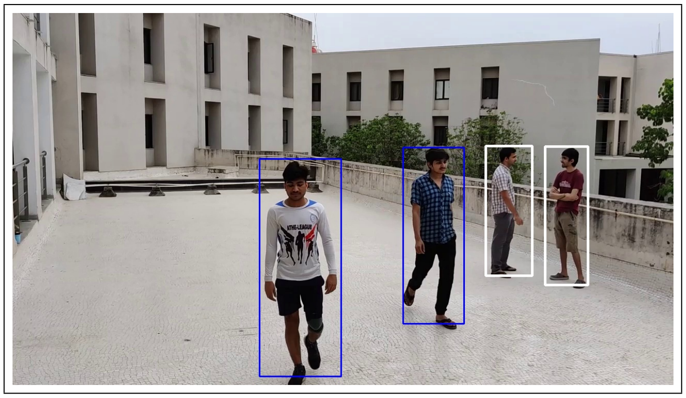

Figure 5. Evaluating Social distancing violations

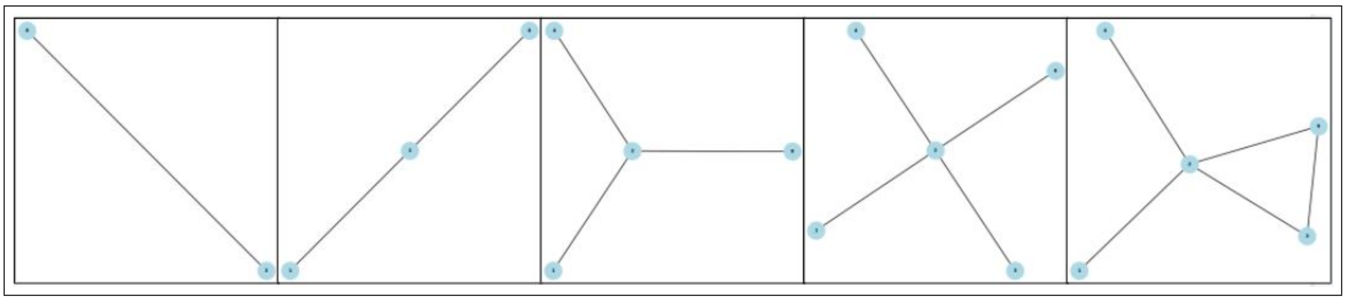

Figure 6. Contact tracing graph

Person ids	No. of Objects
1	31
2	21
3	40
4	25
5	11
6	16
7	4
8	2
9	4
10	12
11	7
12	26
13	16
14	5
15	1

Table 3. Re-identification using torchreid 'osnet_x1_0' without redundancy

Cluster ids	person ids	No. of Objects
0	0, 4, 15	47
1	1, 6, 14	52
2	2, 5, 7, 9, 10, 12	78
3	3, 11	47
4	8	2
5	13	16

Table 4. Re-identification using torchreid 'osnet_x1_0' with redundancy

Cluster id 1	Cluster id 2	No. of violations detected	False violations
0	2	7	3
0	3	3	0
1	2	15	7
2	3	2	0
2	4	5	0

Table 5. Violations stats

total of 73 distinct person ids using these method. person id with no. of times it detected that person (No. of Objects) 10 or greater than 10 are mentioned in Table 2 2. Remaining distinct person ids have an average of 2 objects in it, with most them having either 1 or 2.

6.2.2 torchreid

We used a 'osnet_x1_0' pretrained model from torchreid implementation. It was the best performing model among the models that we tested and Table 3 and Table 4 contains the results for it. Initially while running the model we got 15 different person ids with most of the objects being in first 4 person ids as shown in Table3 3.

We then used a redundancy method to further improve the results by created clusters ids of person ids that seems similar as shown in table 4

6.3. Evaluating Social distancing violations

We then evaluated social distancing violations. We have set social distance violation threshold at 150 pixels. Table 5 5 shows the violations detected using the algorithm. We see that total 32 violations were detected. 10 Out of 32 violations were detected incorrectly.

6.4. Contact tracing graph

The graph in the figure 6, is across different time. We can look at the graph to understand who has been in contact. If a person has acquired COVID, then we can isolate all the connected entities. As we can see, the graph grows with time.

7. Our Contribution

Although extensive research has been going on for person re identification task, it's not been used for contact tracing problems. We implemented entire algorithm from scratch which is using a pre-trained model to create a software for contact tracing.

8. Implementation

You can find the code implementation here.

9. Future Work

To improve upon existing results, we can fine-tune the pre-trained models on our dataset. Fine-tuning improves the robustness of any pre-trained network without needing to train the model from scratch. Furthermore, we can perform depth estimation on the video to get the third co-ordinate of an object. This will give more accurate distance between two objects. Depth estimation can be done using deep neural networks trained in a fully supervised manner with the RGB images as input and the estimated depth as output.

References

[1] E. Ahmed, M. Jones, and T. K. Marks. An improved deep learning architecture for person re-identification. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3908–3916, 2015.

[2] N. K. Bansode, V. Manhas, R. Kumar, S. Shubham, and V. Nayal. Person identification using deep learning. 7:5384–5393, 2020. International Research Journal of Engineering and Technology (IRJET).

[3] S. Jaiswal and D. K. Vishwakarma. State-of-the-arts person re-identification using deep learning. In *2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)*, pages 238–243, 2019.

[4] X. Ma, X. Zhu, S. Gong, X. Xie, J. Hu, K.-M. Lam, and Y. Zhong. Person re-identification by unsupervised video matching, 2016. Article in Pattern Recognition.

[5] M. Qian and J. Jiang. Covid-19 and social distancing, 2020. Journal of public health.

[6] K. R. E. R. Person reidentification and recognition in video. 8827, 2014. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications.

[7] S. Shindea, A. Kotharia, and V. Gupta. Yolo based human action recognition and localization, 2018. International Conference on Robotics and Smart Manufacturing.

[8] T. Yu, D. Li, Y. Yang, T. Hospedales, and T. Xiang. Robust person re-identification by modelling feature uncertainty, 2019. IEEE/CVF International Conference on Computer Vision (ICCV).

[9] K. Zhou and T. Xiang. Torchreid: A library for deep learning person re-identification in pytorch, 2019.

[10] K. Zhou, Y. Yang, A. Cavallaro, and T. Xiang. Omni-scale feature learning for person re-identification, 2019.