COVIPROX: A Deep Learning Framework For Surveillance Based Contact Tracing

Angshuk Dutta
Nanyang Technological University
Singapore

Spriha Bankata Mishra
Nanyang Technological University
Singapore

Rishab Aaryan
Nanyang Technological University
Singapore

ABSTRACT

Mitigation of the COVID-19 pandemic requires a multi-pronged approach. One essential part of this approach is contact tracing. Contact tracing is the process of identifying close contacts of positive COVID-19 cases for isolation, either manually or digitally. Manual contact tracing is unreliable and incomprehensive. Digital contact tracing (DCT), on the other hand, could be a viable solution to minimize the proliferation of the virus. Binary Contact Tracing (BCT), a method under DCT, models infection as a binary event. This method has significant drawbacks. Digital contact tracing with Bluetooth signals is another method which is gaining popularity. However, it poses cybersecurity risks and raises privacy concerns. We pose an alternative solution that relies on surveillance images and CCTV footage to identify close contacts of COVID-19 patients by leveraging deep learning techniques. Our model, 'CoviProx', follows a structured pipeline: mask identification is performed on input images, followed by person reidentification to identify people across different frames. Then, distances between different individuals in the frame are calculated. Finally, an experimental formula is used to calculate the probability of an individual contracting COVID-19 based on the parameters outlined above. Different models, such as MobileNetV2 and Xception, are used during implementation. If deployed on a large scale, this system has the potential to identify all close contacts of a COVID-19 positive case accurately and efficiently.

KEYWORDS

Neural Networks, Person Reidentification, COVID-19, Contact Tracing

ACM Reference Format:
Angshuk Dutta, Spriha Bankata Mishra, and Rishab Aaryan. 2021. COVIPROX: A Deep Learning Framework For Surveillance Based Contact Tracing. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The spread of COVID-19 globally has been alarming and, in most countries, devastating. Many experts believe that COVID-19 will not be limited to being a pandemic but will instead transition into becoming an endemic disease in modern times. As many countries ramp up their vaccination efforts, the need of the hour is to create systems that will keep COVID-19 in check as economies inevitably open up and life settles into a new normal.

Safety management measures, such as social distancing and contact tracing, then, are paramount to these efforts [3]. Contact tracing, in particular, would enable governments to identify close contacts of COVID-19 cases and take the necessary measures (such as isolation and testing) immediately. Digital contact tracing (DCT) identifies possible exposures to COVID-19 by leveraging electronic information. Current DCT technology is mostly Bluetooth-signal based. Bluetooth signal strength is used to estimate proximity between devices, and by extension, between individuals who are close contacts of COVID-19 patients. This can be a very effective strategy, but its success is highly dependent on uptake rate [28]. Such apps have been deployed in Singapore and Australia to varied degrees of success [10] [35]. This strategy relies on people adhering to rules and using the apps diligently, most of which are deployed on an opt-in basis. As a result, enforcing Digital Contact Tracing is a difficult task in most countries.

An alternative solution is to use video analytics to perform digital contact tracing. In this method, surveillance footage can be used to identify contact between COVID-19 positive persons and the general population. Moreover, factors such as proximity and duration of contact can also be identified through this method. Some countries, such as China and South Korea, have adopted a multi-pronged DCT approach which includes security camera footage in addition to the usual resources (such as location data)[31]. However, this technique remains largely overlooked, despite its great potential.

We propose a Digital Contact Tracing methodology based on image recognition. 'CoviProx' is a robust system to predict individuals' chances of testing positive for COVID-19 by identifying individuals through facial recognition and generating probability scores of their likelihood of contracting COVID-19 based on proximity and duration of contact and usage of face masks.

2 BACKGROUND

Digital contact tracing technology is currently predominantly mobile-based and uses primarily GPS and Bluetooth data. It leverages the creation of digital trails which can then be used as a proxy for proximity between individuals [6].

Contact tracing can be invaluable in containing COVID-19 [2]. Manual contact tracing involves reviewing patient history on a case-by-case basis to identify affected individuals and isolate them. However, this technique often fails due to the difficulty associated with recalling all of one's contacts in the recent past. This technique also requires heavy resources in terms of specialised healthcare professionals who interview the patients.

Digital contact tracing, thus, is a more reliable alternative approach. Most DCT systems are cell phone based and use Bluetooth data to pinpoint close contacts of COVID-19 cases. Currently, many DCT applications use binary contact tracing (BCT). The idea behind this is to recommend self isolation to close contacts of recent COVID-19

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

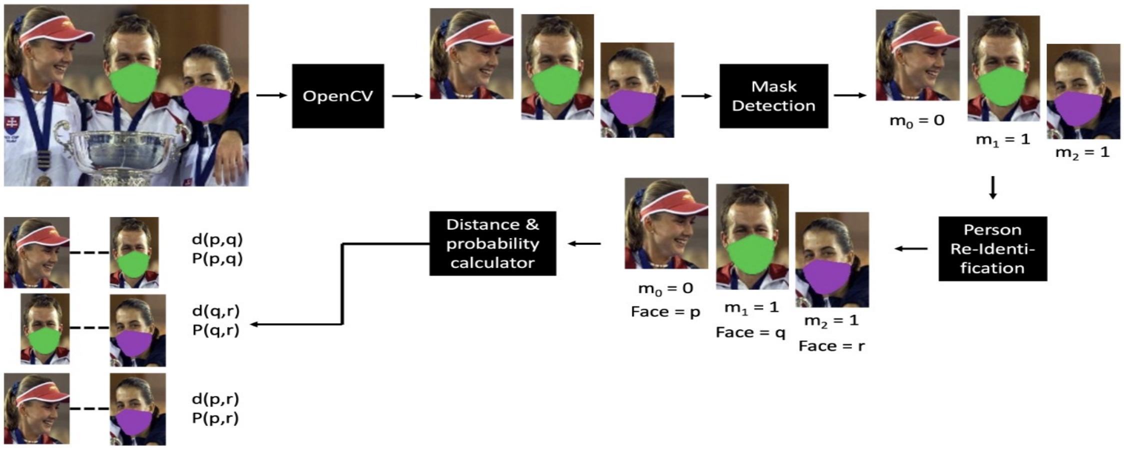

Figure 1: The diagram represents the flow of COVIPROX. The frame with masked and unmasked faces is passed to our opencv module which detects the faces. These faces are passed to the our mask detection module which assigns a label m ∈ 0,1. These collected faces are passed through our Person Reidentification module which retrieves the unique ID for each person. These faces are then passed to our distance and probability module to calculate the idstance and probabilities respectively. The final output is the probability.

patients. Abueg et al. [1] have noted that BCT can significantly decrease the proliferation of COVID-19. BCT's speed and simplicity are favourable, but the lack of a comprehensive approach may lead to oversight of some infected persons while quarantining healthy ones (Hinch et al. [19]).

Further, depending on positive COVID-19 tests may not be the most robust approach. Not only are tests often administered late since many people are asymptomatic when they first contract COVID-19 (Gandhi et al. [13]), but tests also have high false negative rates (Li et al. [25]). Moreover, we cannot ignore that some countries do not have an abundance of COVID-19 tests to rely on.

Machine learning and artificial intelligence for contact tracing applications offer some solutions to this problem [21]. The accuracy of DCT systems can be improved by integrating these technologies. The insights gleaned from collected data can aid experts in both the medical field and in government. ML techniques can analyze Bluetooth signal data effectively to create more accurate proximity detection systems. Similarly, CoviProx aims to leverage deep learning techniques to create a more robust DCT system.

3 RELATED WORKS

So far, digital contact tracing efforts have mostly concentrated around opt-in Bluetooth apps [36], [16] , [9]. These bring up a slate of issues [26]. Kleinman et al. [22] point out various limitations of such apps, the most salient of which is their dependency on adoption rate. In a population, identification of contacts by apps can vary with the square of the proportion of users.

Various apps for digital contact tracing have already been released globally [18] [8]. Singapore's TraceTogether app is one such app [23]. Released in March 2020, it was one of the first of its kind. The app is based on 'OpenTrace,' the reference implementation of an open source protocol for DCT known as BlueTrace [4]. Bay et al. [5] delineate how BlueTrace works by enabling two participating cell phones to exchange temporary identifiers, which avoid third party tracking via regular rotation. Authorities cannot directly access data in this model. One major limitation is that this model is susceptible to the leakage of metadata by the server (linkage attacks) [33]. Ahmed et al. [2] describe France's StopCovid app, which is based on a similar protocol known as 'ROBust and privacy-presERving proximity Tracing' protocol (ROBERT). BlueTrace and ROBERT store different kinds of user information on servers. ROBERT stores anonymous identifiers which afford higher privacy to users and are uploaded in a staggered manner. Many of these apps pose significant cybersecurity and privacy risks [24] [7].

Another important protocol for DCT is Private Automated Contact Tracing (PACT). The PACT Protocol Specification [30] outlines the working of the protocol, which estimates the distance between two devices based on the reception of Bluetooth Low Energy (BLE) packets. The protocol also attempts to determine the duration of contact, and optionally stores metadata which may be useful in estimating risk (such as the features of the place of contact, as, for example, restaurants and marketplaces may pose higher risk than open fields). However, factors such as infectiousness and mask usage are not taken into account in this protocol.

For social distancing detection, Punn et al. [29] implemented person detection and tracking. People are first segregated from the model and then assigned IDs. Centroid coordinates and size of bounding box are used to create a violation index and track social distancing violations. This method does not measure the pairwise distance between people. Gupta et al. [15] leveraged Mask R-CNN to implement person detection, person tracking, estimation of distance from camera and estimation of pairwise social distancing.

4 PROBLEM FORMULATION

Given the surge of cases arising from the growing COVID pandemic, the importance of estimating the likelihood of a person contracting

Figure 2: Left image: Sample images of people (i) With Mask: Kaggle , (ii) With Mask: MAFA , (iii) Without Mask: FDDB , (iv) Without Mask: FDB
Right Image (i) Person without Mask , (ii) Person with Pseudo Mask

the virus swiftly, given a myriad of factors, has become a problem of paramount importance. In this paper we aim to infer that by using a deep learning framework 'COVIPROX'. We introduce the problem formally as follows:
Notation: Let p_i indicate the ith person present in a confined space where $i \in 1, ..., \mathcal{N}$ with $\mathcal{N}$ indicating the total number of people. Let $\mathcal{M} \sim Ber(p)$ indicate the possibility of a person wearing a mask. Furthermore, we assign $\mathcal{D} \in \mathbf{R}$ to be the set of the pairwise distance between all p_i and p_j $\forall i \in N$ and $\forall j \in N \backslash i$.

We consider the problem of estimating the conditional distribution of person i contracting covid $\mathcal{P}_i(\mathcal{Y}|\mathcal{M}_i, \mathcal{M}_j, d_{ij})$ where $\mathcal{Y}$ is a random variable indicating the possibility of COVID and $d_{ij} \in \mathcal{D}$. This probability indicates the likelihood of a person p_i contracting COVID from person p_j and vice versa given p_i or p_j get infected with COVID.

5 PROPOSED METHODOLOGY

In this section, we introduce our proposed architecture for the problem stated above. In order to estimate the conditional distribution of a person contracting COVID, we require necessary architectures to extract the required information. Thus we introduce an end to end deep learning framework to extract the necessary information required and estimate the likelihood. Before we elucidate on the architecture, we summarise the tasks as follows

- **Mask Detection:** Given a CCTV frame, we aim to infer whether the people present in the frame are wearing masks.
- **Person Reidentification:** To carry out automated contact tracing, it is of importance to identify people across different frames. In this step we infer the IDS of the people from a central Database utilising the faces detected in the mask identification step.
- **Distance Calculation:** It is well observed that the rate of transmission of COVID is inversely proportional to the physical distance between the individuals [14]. Hence, it must be taken into consideration for a robust estimation. Therefore, utilising the coordinates of the face detected by our mask identification model and the side length of the bounding boxes, we compute the distances between different people in a particular frame.
- **Probability Calculation:** Lastly, based on the Id', distances and the presence of a mask computed, we aim to infer the probability of person p_i contracting COVID from person p_j and vice versa . To accomplish this, we introduce an experimental formula to evaluate this conditional probability and update our database with the computed values.

5.1 Mask Detection

The problem of mask detection can be described as a binary classification problem. To address this, we utilise a pretrained VGG19 model[32] with batch normalisation. We employ batch normalisation [20] module in order to standardise the input across batches. Furthermore, We employ a transfer learning objective and retrain the model with fewer data points. This ensures that we achieve convergence with reduced number of epochs as the weights initialisation for the model is no longer random but obtained from pretraining. The input for this model is obtained from the frames of the CCTV footage.

5.2 Person Reidentification

The Person Re identification problem in our context is very significant. However, it can be notoriously difficult to implement due to low resolution, different camera angles, occlusions and noise. These issues make person reidentification an open research topic. In our context we have employed the works of He et.al [17] (i.e) a vision transformer based architecture for person re-id.
We utilise a vision transformer as our base architecture. Although the base architecture performs fairly well as stated by He et.al [17], the problems of misalignment and occlusion still persists. Accounting for latent variables like camera angles and occlusion, we finetune our module by utilising the jigsaws patch module. Given a global feature embedding f_g for an entire image the jigsaw patch module divides the image into k patches thus creating the embedding set $f_1, ..., f_k$. It then proceeds to shift a select m embedding. The extent of this shift is determined by a hyperparameter. Furthermore, it shuffles the embedding set. The shuffled embedding set is concatenated with the global embedding to create a robust representation. The entire model is then trained on a combination

of two losses: the ID loss and the triplet loss.

$$\mathcal{L} \ = \ \mathcal{L}_{ID}(f_g) \ + \ \mathcal{L}_T(f_g) + \frac{1}{k}\sum_{i=1}^{k}\mathcal{L}_{ID}(f_l^k) + \mathcal{L}_T(f_l^k)$$

where f_g is the global feature embedding and f_l^k is the patch embedding.

After obtaining the unique ID and information regarding the presence of a mask, we transition to calculating the pairwise distance between each ID present in the video frames.

5.3 Distance Calculation

In this section we introduce our function to calculate the pairwise distance between individuals. Obtaining the coordinates (x_{p_i}, y_{p_i}) and the side length of the bounding box b_{pi} for the face of person p_i and (x_{p_j}, y_{p_j}) and b_{p_j} for face of person p_j, the distance calculated is given by the following equation:

$$d(p_i, p_j) \ = \ \sqrt{(x_{p_i} - x_{p_j})^2 \ + \ (y_{p_i} - y_{p_j})^2} \cdot \frac{W}{\frac{b_{p_i} + b_{p_j}}{2}}$$

The W term represents the average size of the faces. This distance is then stored along with the unique IDs and mask information and is used for probability calculation.

5.4 Probability Calculation

In this section we formally introduce our empirical probability function. The base probability estimation is given using the following function:

$$I(p_i, p_j) = \begin{cases} \frac{100 - d(p_i, p_j)}{100} & d < 100 \\ 0 & d > 100 \end{cases}$$

Now, incorporating the Mask bernoulli variable $\mathcal{M}_i, \mathcal{M}_j$ we can estimate a more robust probability distribution. This estimation is given by the following formula.

$$\mathcal{P}(\mathcal{Y}|\mathcal{M}_i, \mathcal{M}_j, d_{ij}) \ = \ I(p_i, p_j) \cdot (1 \ - \ (0.2 \cdot (\mathcal{M}_i + \mathcal{M}_j)))$$

Algorithm 1: Algorithm For the Flow of **COVIPROX**

Result: The Condition distribution $\mathcal{P}(\mathcal{Y}|\mathcal{M}_i, \mathcal{M}_j, d_{ij})$
Input: Given a set of Frames $\mathcal{F}$ which consists of individual frames f_i
while $f_i \in \mathcal{F}$ **do**
  faces, coordinates, masks $\leftarrow$ $MaskDetection(\mathrm{f}_i)$
  ID $\leftarrow$ $PersonREID(faces, Database)$
  Distance $\leftarrow$ $GetDistance(faces, coordinates, UID)$
  $\mathcal{P} \leftarrow GetProb(UID, Distance, masks)$
  UpdateDatabse(ID, $\mathcal{P}$);
end

6 EXPERIMENTS

6.1 Datasets

Multiple datasets were carefully observed and selected based on their usefulness for training, and also the quality. Some datasets were merged for the model to observe different perspectives, and also have a bigger dataset for training.

6.1.1 Mask Detection Model. Four different datasets were used, two of which are datasets of images of people without masks, and two datasets with masks. The images were of different sizes, and also consisted of different numbers of people in each image. The first dataset with masks was obtained from kaggle. The dataset consisted of 1148 unlabelled images of people with masks. The next masked dataset was obtained from MAFA, which consisted of 25876 unlabelled images, giving us a total of 27024 images of people with masks. The first dataset of images of people without masks was obtained from FDDB, courtesy of University of Massachusetts, Amherst. The dataset consisted of 28204 images in total. The other dataset was the FDB dataset obtained from the Chinese University of Hong Kong, which consisted of 12880 images of people without masks.

6.1.2 Person Re-Identification Model. A set of 13190 images of people without masks was obtained, but labelled. The dataset, called the LFW dataset, was obtained courtesy of the University of Massachusetts, Amherst. Amongst these images, 1680 people had more than one image with the same label. Due to the lack of research in COVID-19, some challenges were faced to obtain a labelled dataset of people both with and without masks which is required for training the model to re-identify the person regardless of the mask's presence. On obtaining the dataset, OpenCV was then used to detect the facial features of each face in each image. OpenCV was used to draw polylines to connect the jawline with the centre of the nose, and fill the newly created shape with a random colour, and hence creating a pseudo mask, And this solved the problem of unavailability of masked images. The images were resized to a consistent size of 250x250 px. The masked images used the same labels and were used for training the Re-Identification model.

6.1.3 Image Preprocessing. Image preprocessing contributes to increasing the accuracy, as well as the efficiency of the model.
For the purpose of mask detection, Multiple image processing techniques were used for training the model. The faces were cropped and resized to a consistent size of 200x200 px using OpenCV. The model generator was set to consider horizontal flips of the face, different rotations of the face as well as shear range preprocessing, where the faces are bent along one axis.
For Person Re-identification, Noise was added to the images in a random implementation to get a wider set of data, remove bias and make the model more real-time friendly. Random contrast brightness was also implemented to train the model to detect faces in different brightnesses and glares. For both models, grey-scaling was used to make the model more efficient, and this also removes color bias.

6.2 Metrics

6.2.1 Mean Average Precision. Mean average Precision can be defined as the Average of the Average Precision over all Queries. To elucidate in terms of information retrieval, Given a number of interested documents or queries n_q we define the AP to be as follows

$$AP@n_q \ = \ \frac{1}{n_T}\sum_{k}^{n} P@k \cdot rel@k$$

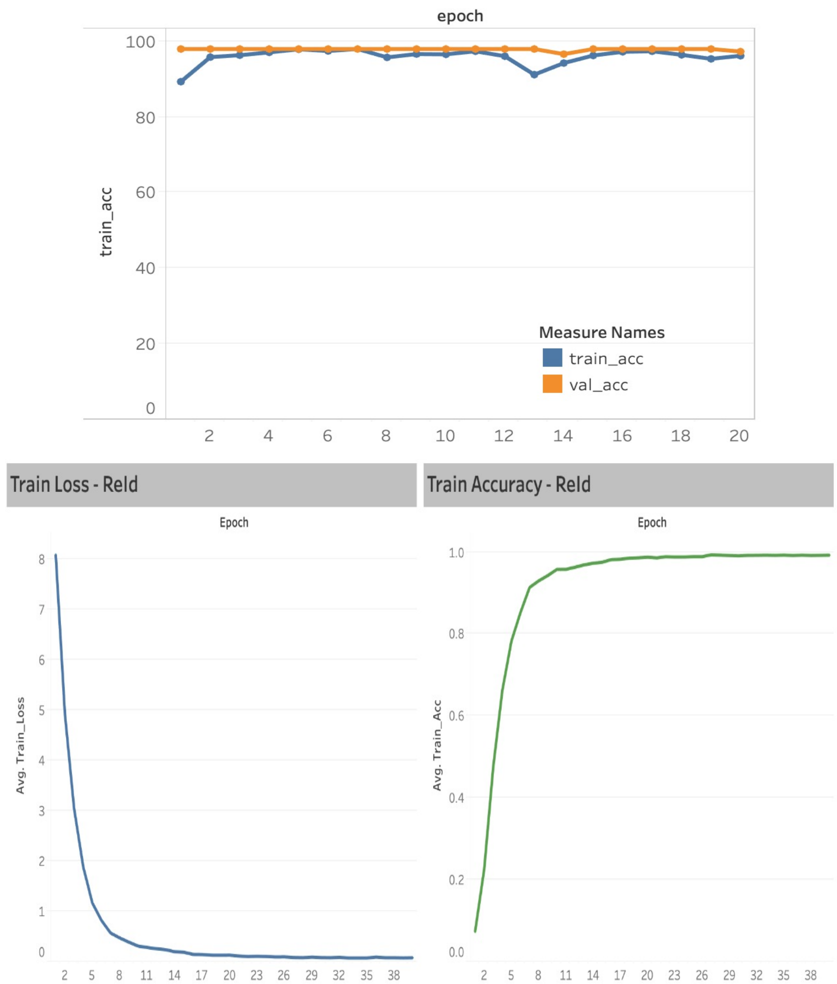

Figure 3: Top Left Fig: Train and Validation Accuracy for the face detection model. Top Right Fig: Train Loss and Train Accuracy for the person REID model Bottom Center Fig: Successful Application of the Face Detection Model

where n_T is the number of samples belonging to ground truth. rel is an indicator function whose value is based on the relevance of the item retrieved.
Based on this, the Mean Average Precision can be defined as follows:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$$

where i indicates the number of the query.

6.2.2 CMC Curves. The most commonly used performance measure for evaluating person re-id is known as a cumulative matching characteristic (CMC) curve, which is analogous to the ROC curve in detection problems. For re-identification, The Cumulative Matching Characteristic (CMC) curve is one of the most widely used inference measure and it is similar to the ROC curve in problems related to detection[27]. It is used to measure identification performance and is based on the relative ordering of the match scores corresponding to each sample. It can also be referred to as a rank-based metric[11]. The CMC curve represents results of an identification task by plotting the correctly identified probability versus the number of candidates returned. [27]

6.3 Model Architecture

6.3.1 Mask Detection. The implemented model is a vgg model with 19 layers with batch normalisation implemented. The batch normalisation is utilised to improve accuracy and standardise the pixel values across all batches. Furthermore the classifier block of the model is replaced with a custom classifier block consisting of 2 linear layers with a ReLU nonlinearlity and dropout. Due to resource limitations, we implement a batch size of 32 for training and 8 for testing. For dropout, a drop rate of 0.5 is discovered to be the optimal choice. Finally a learning rate of 10^{-4} is utilised.

6.3.2 Person Re-identification. We implement a vision transformer [12] as our base model. Utilising the research of He et.al [17] we implement the JPM (Jigsaw path module) with the shuffle constant to be 5. Furthermore we implement the same batch size for training and testing as done in Mask detection. A learning rate of 0.008 is

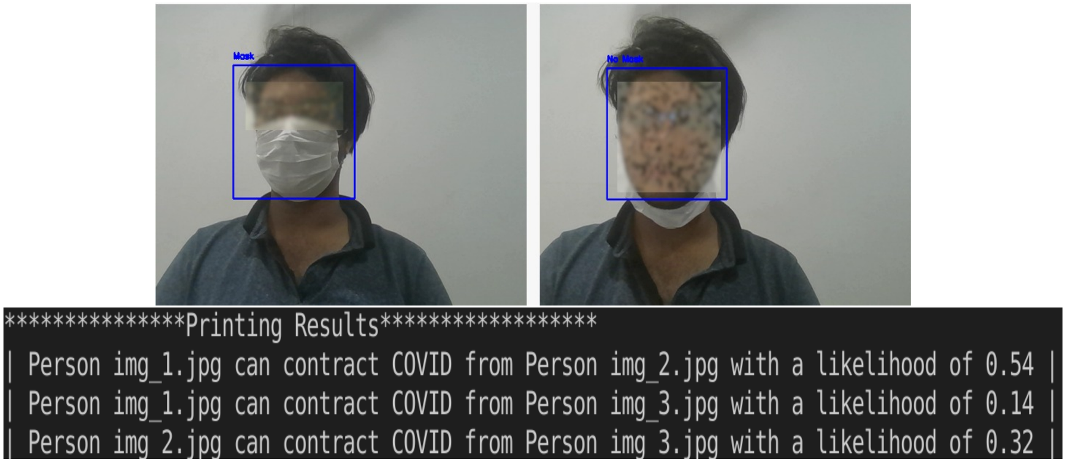

Figure 4: Top Figure: Successful application of the mask detection model bottom figure: Output of COVIPROX

utilised.
The reason for the selection of the vision transformer as our base model can be attributed to the global receptive field of the multihead attention mechanism [34]. In comparison CNN only possess a local receptive field due to downsampling operations like maxpooling. This leads to a loss of environment information which proves to be crucial for person ReID.

7 RESULTS

7.1 Mask Detection

To perform swift mask detection, a model with a relatively fewer number of trainable parameters and a high accuracy is required. However, for the tradeoff, we prioritise accuracy over number of parameters as the difference in parameters is found to be trivial. We select four different models for comparison: VGG19 with batch normalisation, Densenet, resnet50 and resnet 152. From our experiments, we observed VGG19 to be an optimal choice achieving a testing accuracy of **100%** and having fewer parameters with respect to the other variants.

Table 1: Table representing the accuracies of different models and our utilised model

Model	#Parameters	Test acc
DenseNet	7.544M	98.5
resnet50	26.147M	99.0
resnet150	60.783M	**100.0**
VGG19 with BN	**20.626M**	**100.0**

7.2 Person Re-Identification

For person reidentification we select our baseline to be the vision transformer model and observe the mAP achieved by both models to decide on the base architecture. Based on our experiments we observe that the transformer architecture with JPM performs better as compared to our baseline. For measurement metrics, mAP and CMC rank 1 are used. We observe a better accuracy utilising the JPM module. The results are detailed below.

Table 2: Table representing the mAP and CMC Rank 1 of our baseline model vit and our utilised model vit with JPM module for Person REID

Model	mAP	CMC Rank 1
Vit16	87.1	94.6
Vit16 with JPM	**89.0**	**95.1**

8 CONCLUSION

Robust contact tracing systems is important to mitigate the spread of the pandemic. Moreover, social distancing must be enforced and masks must be worn by the public. Self-isolation is important if there is a sizable chance of an individual having the virus, since it often presents as asymptomatic. Given these characteristics, it is the need of the hour to create a multi-pronged system that can identify flouting of social distancing or face mask rules, as well as utilise these use cases to calculate the probability of an individual contracting the virus from another individual they have been in contact with. Our project uses different models, such as VGG with batch normalisation and TRANSReID, to implement these separate processes smoothly and integrate them into one system for COVID-19 contact tracing. Future Work can include the deployment of this system using a distributed architecture.

REFERENCES

[1] Matthew Abueg, Robert Hinch, Neo Wu, Luyang Liu, William JM Probert, Austin Wu, Paul Eastham, Yusef Shafi, Matt Rosencrantz, Michael Dikovsky, et al. 2020. Modeling the combined effect of digital exposure notification and non-pharmaceutical interventions on the COVID-19 epidemic in Washington state. *medRxiv* (2020).
[2] Nadeem Ahmed, Regio A Michelin, Wanli Xue, Sushmita Ruj, Robert Malaney, Salil S Kanhere, Aruna Seneviratne, Wen Hu, Helge Janicke, and Sanjay K Jha.

2020. A survey of COVID-19 contact tracing apps. *IEEE access* 8 (2020), 134577–134601.

[3] Alberto Aleta, David Martin-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E Dean, M Elizabeth Halloran, Ira M Longini Jr, Stefano Merler, et al. 2020. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* 4, 9 (2020), 964–971.

[4] Elian Angius and René Witte. 2012. OpenTrace: an open source workbench for automatic software traceability link recovery. In *2012 19th Working Conference on Reverse Engineering*. IEEE, 507–508.

[5] Jason Bay, Joel Kek, Alvin Tan, Chai Sheng Hau, Lai Yongquan, Janice Tan, and Tang Anh Quy. 2020. BlueTrace: A privacy-preserving protocol for community-driven contact tracing across borders. *Government Technology Agency-Singapore, Tech. Rep* (2020).

[6] Yoshua Bengio, Richard Janda, Yun William Yu, Daphne Ippolito, Max Jarvie, Dan Pilat, Brooke Struck, Sekoul Krastev, and Abhinav Sharma. 2020. The need for privacy with public digital contact tracing during the COVID-19 pandemic. *The Lancet Digital Health* 2, 7 (2020), e342–e344.

[7] Andrea Cioffi, Cecilia Lugi, and Camilla Cecannecchia. 2020. Apps for COVID-19 contact-tracing: Too many questions and few answers. *Ethics, Medicine and Public Health* 15 (2020), 100575.

[8] Vittoria Colizza, Eva Grill, Rafael Mikolajczyk, Ciro Cattuto, Adam Kucharski, Steven Riley, Michelle Kendall, Katrina Lythgoe, David Bonsall, Chris Wymant, et al. 2021. Time to evaluate COVID-19 contact-tracing apps. *Nature Medicine* 27, 3 (2021), 361–362.

[9] Mathieu Cunche, Antoine Boutet, Claude Castelluccia, Cédric Lauradoux, and Vincent Roca. 2020. *On using Bluetooth-Low-Energy for contact tracing*. Ph.D. Dissertation. Inria Grenoble Rhône-Alpes; INSA de Lyon.

[10] Diganta Das and JJ Zhang. 2021. Pandemic in a smart city: Singapore's COVID-19 management through technology & society. *Urban Geography* 42, 3 (2021), 408–416.

[11] Brian DeCann and Arun Ross. 2013. Relating roc and cmc curves via the biometric menagerie. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. IEEE, 1–8.

[12] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020).

[13] Monica Gandhi, Deborah S Yokoe, and Diane V Havlir. 2020. Asymptomatic transmission, the Achilles' heel of current strategies to control Covid-19.

[14] Krzysztof Goniewicz and Amir Khorram-Manesh. 2021. Maintaining Social Distancing during the COVID-19 Outbreak. *Social Sciences* 10, 1 (2021), 14.

[15] Savyasachi Gupta, Rudraksh Kapil, Goutham Kanahasabai, Shreyas Srinivas Joshi, and Aniruddha Srinivas Joshi. 2020. SD-Measure: A Social Distancing Detector. In *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 306–311.

[16] Gary F Hatke, Monica Montanari, Swaroop Appadwedula, Michael Wentz, John Meklenburg, Louise Ivers, Jennifer Watson, and Paul Fiore. 2020. Using Bluetooth Low Energy (BLE) signal strength estimation to facilitate contact tracing for COVID-19. *arXiv preprint arXiv:2006.15711* (2020).

[17] Shuting He, Hao Luo, Pichao Wang, Fan Wang, Hao Li, and Wei Jiang. 2021. Transreid: Transformer-based object re-identification. *arXiv preprint arXiv:2102.04378* (2021).

[18] Enrique Hernández-Orallo, Carlos T Calafate, Juan-Carlos Cano, and Pietro Manzoni. 2020. Evaluating the effectiveness of COVID-19 Bluetooth-Based smartphone contact tracing applications. *Applied Sciences* 10, 20 (2020), 7113.

[19] Robert Hinch, Will Probert, Anel Nurtay, Michelle Kendall, Chris Wymant, Matthew Hall, Katrina Lythgoe, Ana Bulas Cruz, Lele Zhao, Andrea Stewart, et al. 2020. Effective configurations of a digital contact tracing app: a report to NHSX. *Retrieved July* 23 (2020), 2020.

[20] Sergey Ioffe and Christian Szegedy. 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv:1502.03167 [cs.LG]

[21] Firuz Kamalov, Aswani Cherukuri, Hana Sulieman, Fadi Thabtah, and Akbar Hossain. 2021. Machine learning applications for COVID-19: a state-of-the-art review. *arXiv preprint arXiv:2101.07824* (2021).

[22] Robert A Kleinman and Colin Merkel. 2020. Digital contact tracing for COVID-19. *CMAJ* 192, 24 (2020), E653–E656.

[23] Sean Han Sheng Lai, Camelia Qian Ying Tang, Asok Kurup, and Gowreeson Thevendran. 2021. The experience of contact tracing in Singapore in the control of COVID-19: highlighting the use of digital technology. *International orthopaedics* 45, 1 (2021), 65–69.

[24] Franck Legendre, Mathias Humbert, Alain Mermoud, and Vincent Lenders. 2020. Contact tracing: An overview of technologies and cyber risks. *arXiv preprint arXiv:2007.02806* (2020).

[25] Dasheng Li, Dawei Wang, Jianping Dong, Nana Wang, He Huang, Haiwang Xu, and Chen Xia. 2020. False-negative results of real-time reverse-transcriptase polymerase chain reaction for severe acute respiratory syndrome coronavirus 2: role of deep-learning-based CT diagnosis and insights from two cases. *Korean journal of radiology* 21, 4 (2020), 505–508.

[26] Scott McLachlan, Peter Lucas, Kudakwashe Dube, Graham A Hitman, Magda Osman, Evangelia Kyrimi, Martin Neil, and Norman E Fenton. 2020. Bluetooth Smartphone Apps: Are they the most private and effective solution for COVID-19 contact tracing? *arXiv preprint arXiv:2005.06621* (2020).

[27] Sakrapee Paisitkriangkrai, Chunhua Shen, and Anton Van Den Hengel. 2015. Learning to rank in person re-identification with metric ensembles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1846–1855.

[28] Michael J Plank, Alex James, Audrey Lustig, Nicholas Steyn, Rachelle N Binny, and Shaun C Hendy. 2020. Potential reduction in transmission of COVID-19 by digital contact tracing systems. *medRxiv* (2020).

[29] Narinder Singh Punn, Sanjay Kumar Sonbhadra, Sonali Agarwal, and Gaurav Rai. 2020. Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques. *arXiv preprint arXiv:2005.01385* (2020).

[30] Ronald L Rivest, Jon Callas, Ran Canetti, Kevin Esvelt, Daniel Kahn Gillmor, Yael Tauman Kalai, Anna Lysyanskaya, Adam Norige, Ramesh Raskar, Adi Shamir, et al. 2020. The PACT protocol specification. *Private Automated Contact Tracing Team, MIT, Cambridge, MA, USA, Tech. Rep. 0.1* (2020).

[31] Mark Ryan. 2020. In defence of digital contact-tracing: human rights, South Korea and Covid-19. *International Journal of Pervasive Computing and Communications* (2020).

[32] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[33] Ruoxi Sun, Wei Wang, Minhui Xue, Gareth Tyson, Seyit Camtepe, and Damith Ranasinghe. 2020. Vetting security and privacy of global covid-19 contact tracing applications. *arXiv preprint arXiv:2006.10933* (2020).

[34] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*. 5998–6008.

[35] Seyed Shahim Vedaei, Amir Fotovvat, Mohammad Reza Mohebbian, Gazi ME Rahman, Khan A Wahid, Paul Babyn, Hamid Reza Marateb, Marjan Mansourian, and Ramin Sami. 2020. COVID-SAFE: an IoT-based system for automated health monitoring and surveillance in post-pandemic life. *IEEE Access* 8 (2020), 188538–188551.

[36] Qingchuan Zhao, Haohuang Wen, Zhiqiang Lin, Dong Xuan, and Ness Shroff. 2020. On the accuracy of measured proximity of Bluetooth-based contact tracing apps. In *International Conference on Security and Privacy in Communication Systems*. Springer, 49–60.