Proc. of the International Conference on Electrical, Computer and Energy Technologies (ICECET)
9-10 December 2021, Cape Town-South Africa

Imagery, Privacy and Ethics: An Overview of Partially Occluded Facial Biometric Analysis in the Era of Face Masks

Rajith Vidanaarachchi*†, Sachith Seneviratne†, Nuran Kasthuriarachchi‡, Jasper S. Wijnands†, Sanka Rasnayaka§
* *Faculty of Medicine and Health, University of New England*
† *Transport, Health and Urban Design, Melbourne School of Design, The University of Melbourne*
{rajith.vidanaarachchi, sachith.seneviratne, jasper.wijnands}@unimelb.edu.au
‡ *Department of Computer Science and Engineering, University of Moratuwa*
nuran.11@cse.mrt.ac.lk
§ *School of Computing, National University of Singapore*
sanka@comp.nus.edu.sg

***Abstract*—The COVID-19 pandemic has drastically changed human lifestyles, with implications on many aspects of human life. With the proliferation of masks to combat the spread of the virus, many computer vision workflows have been inadvertently affected to varying degrees. Consequently, many research articles have been dedicated to evaluating the impact to existing facial imagery recognition problems. Several works have attempted to either extend existing facial models or develop new techniques specific to masked faces. Many new benchmark tasks have also been introduced in this subdomain. However, a detailed review of such advancements is not available for perusal in this critical area for COVID-safe protocol development. In this work, we address this issue as the first review of masked facial recognition tasks and techniques robust to masked facial images. Our motivation is to provide a central reference for automated public health and COVID-safe identification protocols while also exploring the ethical aspects of further development of such techniques.**

I. Introduction

With the advent of the COVID-19 pandemic, many aspects of everyday life have changed. One of the main changes is the requirement to wear face masks in many jurisdictions globally. Studies have shown that face mask-wearing significantly reduces the chance of spreading respiratory illnesses such as COVID-19 [1]. Even with the COVID-19 pandemic settling down and mask-wearing requirements being lessened, we observe that from time to time, these requirements come back as a response to outbreaks. Apart from this, mask-wearing has been used to respond to other situations, such as air pollution incidents. Recent incidents of bushfires in Australia resulted in significant reductions in air quality to the point that the general population adapted to wearing face masks. Wearing masks is likely to become a regular part of life for many people.

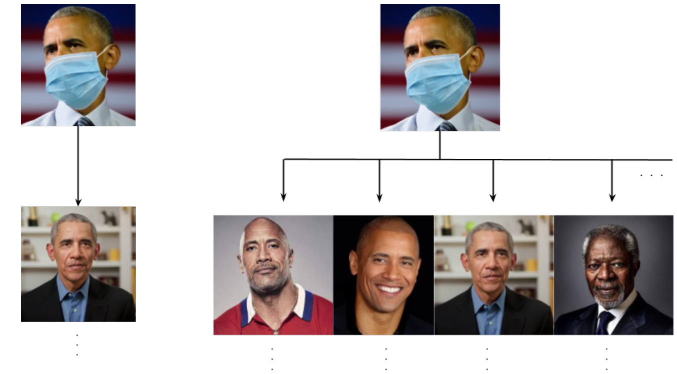
With this new lifestyle change affecting the global population, existing face recognition approaches have met newfound challenges. Many governments, administrations and jurisdictions around the world use biometrics for official purposes, including identification. One such example is airports using facial feature matching as a part of immigration checks [2]. However, with mask-wearing being mandatory at airports, facial recognition systems are not usable [3]. Another facial recognition utility hindered by masks is Apple face identification. The tech giant's smartphone users have found it inconvenient to use the face ID security functionality due to reduced recognition accuracy. The above examples illustrate the need for novel approaches to facial recognition in a mask-wearing society. We define this problem as the masked face recognition problem. Figure 1 illustrates two different facets of masked face recognition.

Another visual problem risen with the pandemic is the necessity for detecting face masks. Many jurisdictions around the globe have introduced laws and guidelines governing the use of face masks in public. In some jurisdictions, AI-based surveillance is used for face-mask detection. Some administrations use this for statistical data collection, while others use this for policing and enforcing reasons [4]. This problem is what we identify as face mask detection.

As mentioned in the previous paragraphs, the newly arising computer vision necessities due to the societal realities of the post-pandemic world can be categorised broadly into two main problems. Firstly, the problem of masked face recognition, and secondly, face mask detection. The former is an extension to current facial recognition problems, whilst the latter is a specific case of an object detection problem. Apart from this, novel research in masked face detection, masked face data collection and annotation, mask synthesis on unmasked faces have been identified as ongoing responses in computer vision to the COVID-19 pandemic.

The masked face recognition [5] problem carries additional challenges to existing facial recognition approaches. Key features of the face, such as nose, mouth, and chin, are entirely covered by masks. Several approaches have been proposed

978-1-6654-4231-2/21/$31.00 ©2021 IEEE



(a) Verification (b) Identification

Fig. 1: Masked face recognition may be performed either for a facial verification task, or a facial identification task.

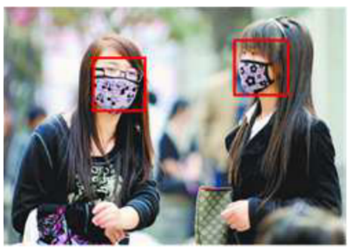


Fig. 2: Face mask bounding box detection [7]

to address this problem. However, in addressing these issues, two sub-problems have been identified. Firstly, the problem of matching a masked face to an unmasked reference can be identified. Secondly, the problem of matching a masked face to another masked face can also be considered. The solutions to this problem can be further extended to facial recognition problems associated with other face coverings such as religious face coverings, vocational face coverings, partially visible faces due to occlusions in the visual data, amongst other situations. Adjabi et al. [6] provide an updated review of face recognition under normal circumstances.

Face mask detection, on the other hand, is a specific sub-problem of object detection. As mentioned earlier, there is scientific and administrative interest in collecting data regarding face mask use. Figure 2 illustrates detected face masks in a real-world situation. Currently different types of face masks with different efficacy are in use amongst the public [8], and it has been identified that the manner of wearing the face mask also varies greatly [9]. Considering these variations, extensions to this problem of face mask detection include the face mask type classification as well as compliance with current mask-wearing regulations.

As an extension to the two main problems, we also identify privacy preservation with regards to the facial occlusions. Would face mask-wearing significantly decrease the likelihood of predicting private information regarding features such as age, sex, ethnicity? Is mask wearing not as effective as perceived in protecting privacy? We also identify the need of using privacy-preserving techniques when dealing with identity revealing information. Especially when AI-based methods are used in public settings solely for the use of detecting face masks, it is essential that the privacy of the individuals is not affected or that it does not become a public surveillance method [19].

II. Literature Review

This section will provide an overview of a selection of existing approaches in masked face recognition, face mask detection, mask synthesis, and an overview of the datasets available related to face masks.

A. Masked Face Recognition

As mentioned earlier, Masked Face Recognition can be divided into two subcategories — matching masked faces to unmasked faces and matching images of masked faces to different masked facial images. Most approaches are based on neural networks, except Ejaz et al. [12] and [20] which use principal component analysis (PCA). Table I provides an overview of the current masked face recognition approaches.

1) Principal Component Analysis based Masked Face Recognition [12]: Of the masked face recognition approaches we consider here, the approach of Ejaz et al. [12] is in the minority that do not use neural networks, but a PCA-based approach. Their analysis shows that the recognition capability is highly reduced in masked face datasets compared with that of non-masked datasets.

2) Neural Network-based Masked Face Recognition: The vast majority of Masked Face Recognition methods use Neural Network-based approaches. These approaches tend to use pre-trained neural networks with innovations for tackling masked faces. As mentioned earlier, the main problem with face recognition with masks is that some key facial features are missing in the images. Currently, two different approaches have been used to circumvent this issue. In restorative approaches, occluded parts of the images are restored using other facial images. In discarding approaches, the facial features typically occluded by a mask are discarded from all training images before neural network calibration.

a) Transfer Learning for Masked Face Recognition: Seneviratne et al. [3] propose a method which uses masked images to build a pretrained network and leverage aspects of transfer learning. They use a self-supervised method to first build a generic facial representation, which is then specialised for masked face recognition using a siamese network to perform transfer learning with a ResNet50 model. They achieve hold-out accuracies ranging between 91.6% to 100% across 7 different masked datasets, including a real-world masked face recognition dataset, on which they achieve 98.75%.

Mandal et al. [16] propose a transfer learning-based masked face recognition model to identify a person's masked face based on their non-masked face images. Their approach is based on a pre-trained ResNet-50 model and uses transfer learning methods where a network already trained on unmasked faces is fine-tuned on masked faces. Their method

Paper	Approach	Evaluation Dataset	Accuracy (%)
Hariri [10]	Occlusion removal approach and training on VGG16 architecture	RWMFD [11]	91.3%
Wang et. al [11]	Face-eye based multi-granularity model	In-house dataset (RWMFD)	95%
Ejaz et. al [12]	PCA features and distance metric	In-house dataset (500 images)	73.75%
Ding et. al [13]	Latent part detection model for discriminative partial feature learning	In-house datasets MFV, MFI and Synth mask LFW	97.9%, 94.3%, 95.7% respectively
Montero et. al [14]	Multi-task ArcFace method	MFR2 [15]	99%
Seneviratne et. al [3]	Self-supervision and transfer learning with a Siamese network	6 synth mask datasets (LFW etc.) and 1 In-house dataset	91.6% - 100% across all 7 datasets
Mandal et. al [16]	Transfer learning approach with the ResNet-50 architecture	RWMFD [11]	47.91%
Li et. al [17]	De-occlusion and knowledge transfer to create unmasked images for recognition	AR dataset [18]	95.4%

TABLE I: An overview of current masked face recognition research.

achieves 89% validation accuracy on an in-house unmasked dataset, and 47.91% on real-world masked face recognition dataset (RMFRD) (masked) [11].

b) Masked Area Discarding Approaches: Hariri [10] uses pre-trained convolutional neural networks, ResNet-50, VGG-16 and AlexNet to obtain features from unmasked regions in the face (i.e. eyes and forehead). This approach can be identified as an occlusion removal approach, as, in essence, masked parts of the face were discarded, both in masked and unmasked imagery, before the training stage. They use the cropped unmasked regions to extract features for a deep bag-of-features approach to recognise faces. Their approach shows high accuracy when used with the Real World Masked Face (RWMF) dataset.

Ding et al. [13] employ a Latent Part Detection model to locate the latent facial part making it robust to masked facial data from both real-world and synthesised data sources. They show their method's effectiveness on three datasets; Masked Face dataset for Verification (MFV), Masked Face dataset for Identification (MFI), and Synthesised Labelled Faces in the Wild (LFW) datasets. They synthesise masks on unmasked image data as a data-augmentation technique and incorporate two knowledge sharing pipelines to detect global features without the masked area and the latent area.

Montero et al. [14] implement a masked face identification pipeline based on the ArcFace model, named MultiTask-ArcFace, which can recognise masked faces and classify mask usage. This model can also be used on unmasked faces without loss of accuracy. Their model is similar to [13], which also synthesise an artificial mask on unmasked facial images. Broadly we identify this as a discarding approach.

c) Masked Area Restoration Approaches: Li et al. [17] use a Generative Adversarial Network (GAN) approach to complete the masked portion of the face with a plausible substitute to assist in the task of masked face recognition. Their approach contains two steps, de-occlusion and distillation. The de-occlusion completes the face with the GAN, while the distillation step takes a pre-trained facial recognition model as a teacher and trains a student model for recognising synthesised faces. The facial restoration approach has specific challenges due to the difference between the generic synthesised faces and the natural face.

The method introduced by Bagchi et al. [20] — although not a Neural Network-based approach — is included under masked area restoration approaches as it is one of the earliest works in masked area restoration for facial recognition in a real-world setting. It uses a PCA-based method for reconstructing 3D features of a face for a varied range of occlusions, including occlusions of the mouth. This study does not consider face-mask specific occlusions as it was published before the everyday use of face masks. Their occlusion compensation scheme has achieved recognition accuracy of 91.3%.

B. Face Mask Detection

The problem of Face Mask Detection arises with the need to identify whether a particular image of a face has a face mask worn in it. This problem can be separated into three different categories from the way different methods have attempted to solve it.

Firstly, most methods follow a two-step process of locating a face and detecting a face mask on the face. Secondly, more nuanced approaches are found where face masks are categorised into different types, including the assessment of compliance with proper mask use guidelines. Thirdly, there are studies on face detection (just detecting the presence of a face, as opposed to recognising it), with masks on. These studies have been included in this section as well, as they handle a detection task instead of a recognition task. Table II provides an overview of current face mask detection research.

1) You Only Look Once Approaches: The majority of object detection pipelines are based on a variation of You Only Look Once (YOLO) [26]. YOLO is a fast unified approach to object detection using a single neural network. It employs a grid-based approach where a neural network determines bounding boxes and class probabilities in a single execution and then uses that information to detect the objects of interest. This approach makes YOLO based models suitable for real-time use, especially with video streams. Hence, it makes it suitable for mask detection tasks as well.

a) ResNet-50 and YOLO-v2 based Face Mask Detection: Loey et al. [21] uses this approach, where the ResNet-50 model is used to improve feature extraction with a deep transfer learning model. Then YOLO-v2 is used for the mask detection. They also introduce a new masked face dataset and use data augmentation approaches to circumvent the data scarcity problem often associated with face-mask related work.

Paper	Approach	Evaluation Dataset	Result
Loey et al. [21]	ResNet-50 and YOLO based Object Detection Apprach	Medical Masks Dataset, Face Masks Dataset [22]	81% (AP)
Cao et al. [23]	YOLO based Object Detection with Mosaic Data Augmentation	MAFA [7]	94% (acc)
Batagelj et al. [24]	Two step pipeline of Face Detection and Compliance check with off-the-shelf methods	MAFA [7] and Wider Faces [25]	95.72% (acc)
Ge et al. [7]	LLE CNNs	MAFA [7]	76.4% (AP)

TABLE II: An overview of current face mask detection research.

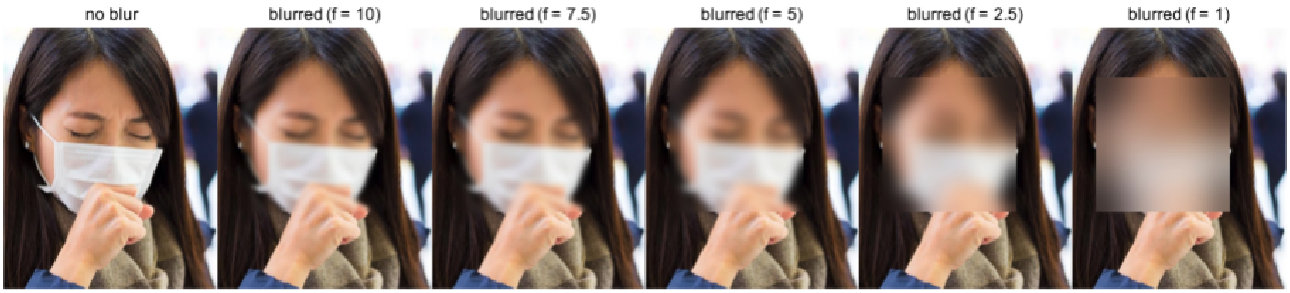


Fig. 3: A simple blurring operation would allow masks to be detected while protecting identity revealing features of a facial image [19].

b) Mask Hunter: YOLO-v4 based Face Mask Detection: Cao et al. [23] also introduces a YOLO based face mask detection pipeline, again with an object detection approach. They propose effective backbone, neck and prediction head structures based on the YOLO-v4 series and a novel improved Mosaic data augmentation. Compared to other methods, this method can maintain high accuracy at a faster frame rate, making it highly suitable for face-mask detection from real-time image streams.

2) Mask Detection and Compliance Check: Besides face mask detection, Batagelj et al. [24] also investigated mask type and compliance with current regulations. The rationale for their work is that for a COVID-19-like pandemic situation wearing a face mask in a compliant manner is significant. They propose a three-step pipeline, detecting the face location as the initial step, followed by detecting the presence of a face mask (if any) and classifying it as compliant or non-compliant. For each part of the pipeline, they evaluate existing algorithms' performance. As opposed to investigated pipelines, their pipeline achieves an accuracy that is at least 10% superior in accuracy.

3) Privacy Preserving Mask Detection: When certain jurisdictions started using mask detection systems based on AI for policing and statistical purposes, privacy concerns were highlighted. More awareness has been raised on social surveillance and privacy implications, especially in the wake of the mass surveillance related to the People's Republic of China's social credit system [27]. In fact, the European Union has introduced the General Data Protection Regulation (GDPR) [28], regulating any data containing private information — or identity revealing information. While many kinds of data collections can be interpreted as identity revealing information, in the EU, a ruling of the European Court of Justice [29] has determined that video streams containing facial images are identity revealing information. As such, GPDR is applicable in these cases.

Fig. 4: Non-occluded, mask-occluded, hand-occluded, non-mask accessory-occluded face boundary boxes detected in real-world images.

With France using CCTV streams from public transport to monitor face mask usage in public [4], the data analysis company associated with the surveillance [30] announced their adherence to the GDPR, albeit their methodology remains unknown.

Kühl et al. [19] provides an in-depth review on the privacy implications of mask detection and proposes a privacy-preserving mask detection pipeline. They present two versions, both adhering to the GDPR guidelines. However, the two systems have a trade-off between privacy preservation and accurate face mask detection. Overall, their method uses a blurring mechanism on edge, as shown in Figure 3, to conceal the identity of the person. Their methods achieve 95 - 99% accuracy in a privacy-preserved setting.

C. Detection of Masked Faces

Although similar to face mask detection or masked-face recognition, the detection of masked faces can be identified as a distinct task. It also is a novel research problem in the time of the COVID-19 pandemic. We include a discussion on masked face detection here as it is a detection problem instead of a recognition problem. In many use cases, straightforward de-

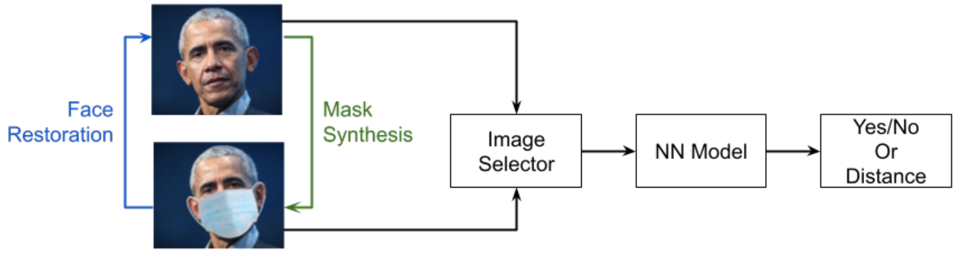


Fig. 5: A generalised pipeline for masked face recognition is depicted. Most methods use either a face restoration step or a mask synthesis approach to regularise the reference image and the testing image.

tection of a face is sufficient, rather than recognising the face. Some examples are facial filters in social media apps, face tracking for camera focusing, and face detection for counting purposes. Figure 4 illustrates different type of occlusions that affect face detection workflows. Most facial detection systems (such as social media apps), are already robust to some non-mask accessories. Masked face detection is a newly emerging challenge.

a) Detecting Faces in the Wild with LLE-CNNs: Ge et al. [7] introduces a Masked Face Detection approach. They utilise Local Linear Embedding Convolution Neural Networks (LLE-CNNs) to train their model, surpassing state-of-the-art methods by at least 15% accuracy when detecting masked faces. They further formalise the problem by defining the occlusion degree of the face. The eyes, nose, mouth and chin are considered to be landmark regions in the face. Occlusions of 1-2 regions is considered weak occlusions, while 3 and 4 occlusions are considered medium and heavy occlusions, respectively, for face detection. With a typical face mask worn in a compliant manner, nose, mouth and chin would be covered, resulting in a medium occlusion. However, with a significant part of the population regularly using glasses in combination with face masks, heavy occlusions make face detection challenging.

D. Datasets

With the new reality of face-mask wearing in public places, novel methods were introduced, as we have described above, to circumvent different challenges in computer vision. However, these novel methods also required new datasets for training and validation. In this section, we will be looking at some of these datasets. Most datasets contain real-life masked facial images as well as images with masks synthesised upon them.

a) Real World Masked Faces Dataset (RWMF) [11]: Wang et al. [11] is one of the earliest masked face datasets released as a response to the COVID-19 pandemic. This dataset consists of three parts with real-life and simulated mask images geared towards mask detection and masked face recognition.

Firstly, they introduce the Masked Face Detection Dataset (MFDD), a dataset of crawled images of masked faces, and other data from similar research. This dataset is useful for training masked face detection and face mask detection models.

Secondly, they introduce the Real-world Masked Face Recognition Dataset (RMFRD), consisting of 5,000 masked images and 90,000 unmasked images of 525 subjects. This dataset is useful for training masked face recognition models.

Thirdly, they introduce the Simulated Masked Face Recognition Dataset (SMFRD), which incorporates simulated mask images. This dataset is more extensive than the RMFRD with 500,000 images of 10,000 subjects. It is, however, unclear how the mask simulation has been carried out.

b) Masked Face Segmentation and Recognition (MFSR) [31]: Geng et al. [31] introduces Masked Face Segmentation and Recognition (MFSR) dataset, with masked and unmasked faces for the same identities. Based on this dataset, they further introduce an Identity Aware Mask Generative Adversarial Network (IAMGAN) to generate or synthesise masks on non-masked facial images. Their dataset consists of two parts. Firstly, they present segmentation annotation for face masks in a set of 9742 images. Secondly, they present a recognition dataset with 11,615 images of 1,004 identities with paired masked and unmasked images for each identity. These images include pose variations, lighting changes, expression changes and diverse mask types.

c) Masked Faces (MAFA) [7]: This dataset by Ge et al. [7] contains 30,811 images with 35,806 masked faces (some images contain multiple faces). This dataset has diversified occlusion levels, classified as weak, medium, and heavy occlusions, presented as lateral, lateral-front, and front views. Further, it provides occlusion types such as simple, complex, hybrid masks, and human body occlusions (such as a hand covering the face). This dataset is helpful for masked face detection work as well as for face mask detection.

d) Face Mask Label Dataset (FMLD) [24]: This dataset of facial images with real masks annotates mask compliance, depending on whether the face mask covers the mouth and the nose adequately. The images are a collection from the MAFA dataset, as well as the Wider Face dataset.

E. Mask Synthesis

Mask synthesis is an additional computer vision task that has emerged while addressing the main masked face recognition and face mask detection problems. As described in the above sections, many novel approaches attempting to solve various recognition, detection and classification tasks required adequate training and validation data. However, public datasets

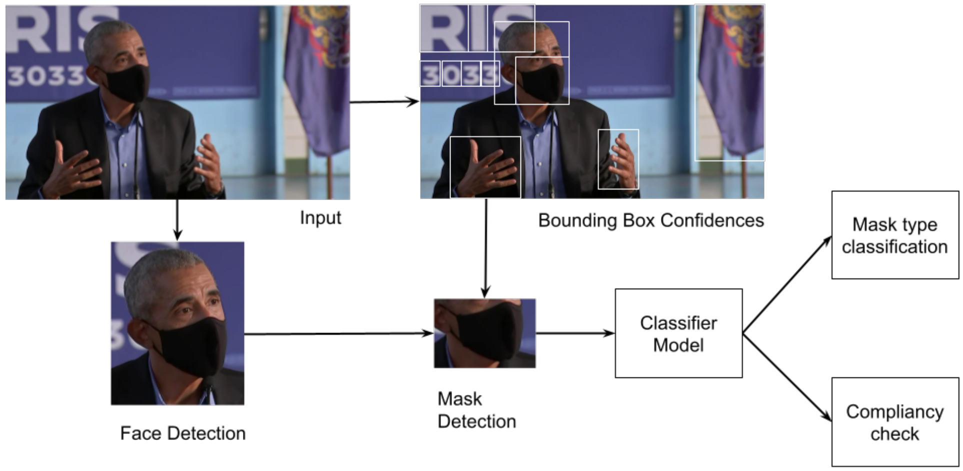


Fig. 6: A generalised pipeline for face mask detection is depicted. Either a direct face-mask detection or a mask-detection on a detected face is usually followed by a classifier for compliance or mask type.

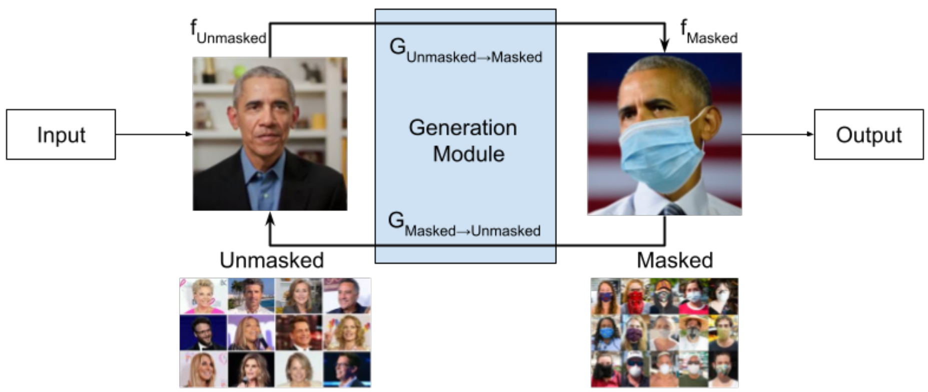


Fig. 7: A generalised pipeline for GAN-based approach to mask synthesis. An input image $f_{Unmasked} \in Unmasked$ is converted into a masked face $f_{Masked} \in Masked$ using domain transfer methods.

of people in real-life environments wearing masks are scarce, especially annotated ones. As a solution to this, many datasets used mask synthesis to overcome the scarcity of data. In this section, various mask synthesis approaches will be explored. Figure 8 illustrates the mask synthesis results with two different approaches.

1) GAN Based Synthesis: A variety of tools exist based on GANs to perform domain transfers on facial images. Most of these methods have been previously used for domain transfers from young to old, male to female, and vice-versa [32]. However, similar techniques can be used for domain transfers from unmasked to masked and vice versa. Many GAN-based generic domain transfer methods for facial images already exist [33]–[36].

a) IAMGAN [31]: Geng et al. [31] present a GAN-based method to create identity aware masking of standard unmasked full face images. Their method consists of two modules: a cyclic generator that converts standard faces into corresponding masked faces; and a multi-level identity preservation module that uses a semantic region guided approach.

2) Keypoint-based Mask Synthesis: Keypoint-based mask synthesis methods use image processing techniques to detect facial landmarks determining where a face mask would fit. Then they paste an annotated face mask onto the area of the face identified through the key points. A disadvantage of this method is that masks look more unnatural than with GAN-based systems. However, these methods have the added advantage of being able to adopt different kinds of face masks.

a) MaskTheFace Tool [15]: This tool, introduced by Anwar and Roychowdhry [15], uses a facial landmark-based approach for synthesising masks on unmasked faces. Their approach works on a range of poses as they use a facial tilt estimator in combination with multiple mask templates to match the estimated facial tilt. Then the mask template is warped to match the identified facial features. This tool allows applying multiple mask types as well as variations of the mask surface. Their method also works on images with multiple faces.

b) Ding et al. [13]: This research introduces a keypoint masking mechanism as part of their facial recognition pipeline.

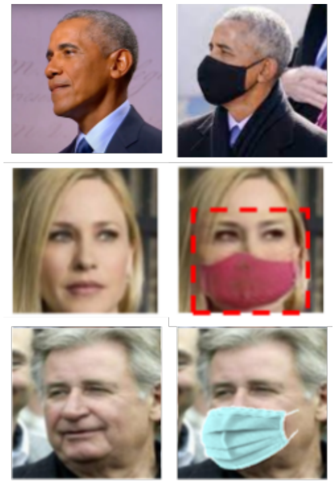
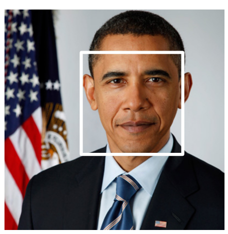
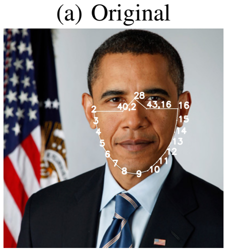


Fig. 8: Top row: Image of an unmasked face (L) and the same person's masked face (R) in a real-life scenario. Middle row: An unmasked face (L) and the same image with a GAN synthesised mask (R) [31]. Bottom row: An unmasked face (L) and the same image with a key-point based synthesised mask (R) [15]

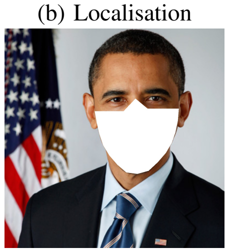
(a) Original



(b) Localisation



(c) Key points



(d) Masked

Fig. 9: Keypoint-based synthetic mask creation pipeline is illustrated here. Firstly, from the original facial imagery (a), masking area is localised (b). Then the key-points are identified for fitting a mask (c). This area is masked using suitable mask imagery to achieve a synthesised mask (d).

It also identifies multiple facial landmarks to locate a bounding box for the mask. Then, the Delaunay triangulation algorithm [37] is used to divide the mask into many small triangles, which are affined into the facial image. This method, however, only works for front-facing facial images of individuals. This method also allows a variety of mask types to be synthesised.

III. Overview of Computer Vision Pipelines

In this section, generalised computer vision pipelines for a selection of tasks are presented.

Figure 5 illustrates a generalised pipeline for masked face recognition. Face restoration and mask synthesis processes are significant in a facial recognition task where an unmasked face image is matched to a masked image, or vice-versa. Upon selecting the image, a neural network model is typically trained, which outputs a distance metric or a matched / un-matched status. The processes used for verification and identification does not portray major differences in the pipeline.

Figure 6 illustrates two common approaches to mask detection. Some methods directly seek the detection of masks using a bounding box confidences approach. Others use a similar process to detect a face before localising the masked area. Usually, this step is followed by a classifier model to check compliance or mask type.

Figures 7 and 9 illustrate GAN- and keypoint-based mask synthesis approaches respectively. In a GAN-based approach, an unmasked input image is converted into a masked image by using domain transfer techniques using a GAN. Two generation functions, $G_Unmasked \rightarrow Masked$ and $G_Masked \rightarrow Unmasked$, are used for the domain transfer. Keypoint-based methods are more straightforward and use a facial keypoint detection method following a facial localisation step to identify landmark features of the face. A pre-arranged mask image is mapped onto the key points of the face to synthesise the mask.

IV. Discussion and Conclusion

Based on our exploration of academic literature in the area of masked facial recognition, we find that some areas such as Masked Facial Recognition are currently under-explored, indicating significant potential for future research work. Additionally, very few works evaluate on multiple datasets, which is an important consideration for real world deployment and use. Due to the current ongoing COVID19 crisis, research in this area has considerable merit with regard to the development of health related applications.

Our findings indicate that many work utilises synthetic masking as a means to generate data. However, due to the domain gap between masked and synthetic images, it is important that work is evaluated on work datasets of both types. This highlights the need to have more reproducible results in this area, including the need for more standard benchmarks and large scale labelled masked datasets.

We find that most works in this area can be summarised into several pipelines, which provide useful generalizations for reasoning about work in this area. We have presented these pipelines and introduced them alongside work which utilises them so that future work may effectively build upon them and have more relevant benchmarks or prior work to compare against. Due to the relevant recency of this area, we believe our work will help guide the development of more robust future technical work.

However, a key issue in this regard are the ethical considerations of such work. We posit that any work in this area needs to weight the potential benefits with regards to ethical use cases such as health related applications against potential misuses in automated surveillance workflows. We find that there are many instances both in academic literature and in main-stream media where such concerns are raised. This highlights the importance of careful ethical evaluation of developed methods by the biometrics and computer vision research communities prior to developing such work, especially as the implications of such technologies can have far reaching consequences beyond the context of the COVID-19 pandemic.

References

[1] H. Ueki, Y. Furusawa, K. Iwatsuki-Horimoto, M. Imai, H. Kabata, H. Nishimura, and Y. Kawaoka, "Effectiveness of Face Masks in Preventing Airborne Transmission of SARS-CoV-2," *mSphere*, vol. 5, no. 5, 10 2020.

[2] S. Gold, "Border control biometrics and surveillance," *Biometric Technology Today*, vol. 2012, no. 7, pp. 9–11, 7 2012.

[3] S. Seneviratne, N. Kasthuriaarachchi, and S. Rasnayaka, "Multi-Dataset Benchmarks for Masked Identification using Contrastive Representation Learning," *arXiv preprint arXiv:2106.05596*, 2021.

[4] H. Fouquet, "Paris Tests Face-Mask Recognition Software on Metro Riders," 2020. [Online]. Available: https://www.bloomberg.com/news/articles/2020-05-07/paris-tests-face-mask-recognition-software-on-metro-riders?sref=C3P1bRLC

[5] F. Boutros, N. Damer, J. N. Kolf, K. Raja, F. Kirchbuchner, R. Ramachandra, A. Kuijper, P. Fang, C. Zhang, F. Wang, and others, "MFR 2021: Masked Face Recognition Competition," in *2021 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2021, pp. 1–10.

[6] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," *Electronics (Switzerland)*, vol. 9, no. 8, pp. 1–53, 7 2020.

[7] S. Ge, J. Li, Q. Ye, and Z. Luo, "Detecting masked faces in the wild with LLE-CNNs," in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua. Institute of Electrical and Electronics Engineers Inc., 2017, pp. 426–434.

[8] A. Kumar, B. Bhattacharjee, D. Sangeetha, V. Subramanin, and B. Venkatraman, "Evaluation of filtration efficacy of various types of facemasks using ambient and PAO aerosols following with different sterilization methods," *medRxiv : the preprint server for health sciences*, p. 152808372110287, 7 2020.

[9] T. L. D. Huynh, ""If You Wear a Mask, Then You Must Know How to Use It and Dispose of It Properly!": A Survey Study in Vietnam," *Review of Behavioral Economics*, vol. 7, no. 2, pp. 145–158, 2020.

[10] W. Hariri, "Efficient Masked Face Recognition Method during the COVID-19 Pandemic," 5 2021.

[11] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Pei, H. Chen, Y. Miao, Z. Huang, and J. Liang, "Masked Face Recognition Dataset and Application," *arXiv*, 2020.

[12] M. S. Ejaz, M. R. Islam, M. Sifatullah, and A. Sarker, "Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition," in *1st International Conference on Advances in Science, Engineering and Robotics Technology 2019, ICASERT 2019*. Institute of Electrical and Electronics Engineers Inc., 2019.

[13] F. Ding, P. Peng, Y. Huang, M. Geng, and Y. Tian, "Masked Face Recognition with Latent Part Detection," in *MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia*. Association for Computing Machinery, Inc, 10 2020, pp. 2281–2289.

[14] D. Montero, M. Nieto, P. Leskovsky, and N. Aginako, "Boosting Masked Face Recognition with Multi-Task ArcFace," 4 2021.

[15] A. Anwar and A. Raychowdhury, "Masked Face Recognition for Secure Authentication," 2020.

[16] B. Mandal, A. Okeukwu, and Y. Theis, "Masked Face Recognition using ResNet-50," 2021.

[17] C. Li, S. Ge, D. Zhang, and J. Li, "Look through Masks: Towards Masked Face Recognition with De-Occlusion Distillation," in *MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia*. Association for Computing Machinery, Inc, 10 2020, pp. 3016–3024.

[18] A. M. Martinez and R. Benavente, "The AR face database," *CVC Technical Report 24*, 1998.

[19] N. Kühl, D. Martin, C. Wolff, and M. Volkamer, ""Healthy surveillance": Designing a concept for privacy-preserving mask recognition AI in the age of pandemics," *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2020.

[20] P. Bagchi, D. Bhattacharjee, and M. Nasipuri, "Robust 3D Face Recognition in Presence of Pose and Partial Occlusions or Missing Parts," *International Journal in Foundations of Computer Science & Technology*, vol. 4, no. 4, pp. 21–35, 2014.

[21] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection," *Sustainable Cities and Society*, vol. 65, p. 102600, 2021.

[22] Larxel, "Face Mask Detection — Kaggle," 2020. [Online]. Available: https://www.kaggle.com/andrewmvd/face-mask-detection

[23] Z. Cao, M. Shao, L. Xu, S. Mu, and H. Qu, "Maskhunter: Real-time object detection of face masks during the COVID-19 pandemic," *IET Image Processing*, vol. 14, no. 16, pp. 4359–4367, 2020.

[24] B. Batagelj, P. Peer, V. Štruc, and S. Dobrišek, "How to Correctly Detect Face-Masks for COVID-19 from Visual Information?" *Applied Sciences*, vol. 11, no. 5, p. 2070, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/5/2070/htm https://www.mdpi.com/2076-3417/11/5/2070

[25] S. Yang, P. Luo, C. C. Loy, and X. Tang, "WIDER FACE: A face detection benchmark," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, 2016, pp. 5525–5533.

[26] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem. IEEE Computer Society, 2016, pp. 779–788.

[27] F. Liang, V. Das, N. Kostyuk, and M. M. Hussain, "Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure," *Policy and Internet*, vol. 10, no. 4, pp. 415–453, 12 2018.

[28] The European Parliament and the Council of the European Union, "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data," pp. 16–32, 2016.

[29] Case C-212/13, "František Ryneš v Úřad pro ochranu osobních údajů."

[30] "Computer vision — Datakalab — France." [Online]. Available: https://www.datakalab.com/?lang=en

[31] M. Geng, P. Peng, Y. Huang, and Y. Tian, "Masked Face Recognition with Generative Data Augmentation and Domain Constrained Ranking," in *MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia*. Association for Computing Machinery, Inc, 2020, pp. 2246–2254.

[32] Y. Liu, Q. Li, Z. Sun, and T. Tan, "A3GAN: An Attribute-Aware Attentive Generative Adversarial Network for Face Aging," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2776–2790, 2021.

[33] Y. Choi, M. Choi, M. Kim, J. W. Ha, S. Kim, and J. Choo, "StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-Image Translation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8789–8797.

[34] Z. He, W. Zuo, M. Kan, S. Shan, and X. Chen, "AttGAN: Facial Attribute Editing by only Changing What You Want," *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5464–5478, 2019.

[35] G. Zhang, M. Kan, S. Shan, and X. Chen, "Generative adversarial network with spatial attention for face attribute editing," in *Lecture Notes in Computer Science*, vol. 11210 LNCS, 2018, pp. 422–437.

[36] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2017-Octob. Institute of Electrical and Electronics Engineers Inc., 2017, pp. 2242–2251.

[37] D. T. Lee and B. J. Schachter, "Two algorithms for constructing a Delaunay triangulation," *International Journal of Computer & Information Sciences*, vol. 9, no. 3, pp. 219–242, 6 1980. [Online]. Available: https://link.springer.com/article/10.1007/BF00977785