Journal of Theoretical and Applied Information Technology
15th November 2023. Vol.101. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MASKED FACE DETECTION: ADVANCEMENTS, CHALLENGES, AND RESEARCH DIRECTIONS

SANTO CHERIAN [1] , BINU THOMAS [2]

[1] Research Scholar, Mahatma Gandhi University, School of Computer Sciences, India

[2] Associate Professor. Marian College Kuttikkanam, Department of Computer Applications, India

E-mail: [1] santo.cherian@mariancollege.org, [2] binu.thomas@mariancollege.org

ABSTRACT

Facial identification has become indispensable in our regular lifestyle as a convenient and quick method of precisely validating identity. It has improved and expanded in lockstep with technological advances and intense learning. The most recent COVID-19 epidemic has highlighted the significance of hygienic and contactless identity validation. Individuals have been obliged to wear masks in order to limit the transmission of the coronavirus. However, this makes it impossible to monitor sizable crowds of mask-wearing people. The effect of using a mask on group facial identification can be a contentious subject that has yet to receive much attention. The paper overviews the various AI algorithms employed for masked face detection and their associated datasets. Current benchmarking initiatives that primarily focus on masked face identification algorithms are reviewed. Furthermore, this paper examines existing evaluation efforts that primarily concentrate on assessing the performance of algorithms specifically designed for identifying masked faces. The identified research directives could be an excellent starting field for researchers looking to create increasingly efficient and productive systems in masked face detection

Keywords: *Masked Facial Detection, Neural Network, Datasets, Face mask, Deep Learning*

1 INTRODUCTION

Recent decades have seen a global increase in research into facial recognition [1][2][3]. Faces rank among the significant environmental cues. They communicate details on a variety of fundamental characteristics, including identity, emotions, gender, as well as age. For societal perception and communication, faces must be processed typically. Substantial advancements have been achieved due to the emergence of artificial intelligence and the growth of technology [4][5]. Due to its ability to verify identity in various scenarios like surveillance cameras, security measures, and authorization processes without requiring personal participation, face recognition has emerged as the most extensively recognized form of biometric technology. During the last few years, face recognition has seen colossal advances in pushing progressive performance to close to human, typically surpassing human capabilities. To ascertain and manage user access, governmental and private organizations employ facial detection technologies [6][7]. Recent studies have demonstrated that sophisticated facial detection algorithms' accuracy is highly reliable and has reached 99.8% [8].

The COVID-19 outbreak has presented a severe threat to the entire humanity. From identifying masked faces [9][10], diagnosing COVID-19 patients, determining infection uncertainties, designing a disease surveillance and prediction system, and enhancing lesion profiling of COVID-19 CT scans, Artificial Intelligence (AI) strategies were indeed able to assist people in numerous ways in their struggle against the viral infection. This review principally centers on masked face detection techniques among these approaches. Before this epidemic, people frequently wore face masks to shield themselves against air pollution or paramedical staff inside health facilities or conceal their identities when committing crimes. However, in response to the COVID-19 outbreak, governmental organizations have adopted a plethora of biosafety laws to prevent infections [11][12][13]. One of these measures involves the compulsory utilization of face masks in public settings, as studies have demonstrated their effectiveness in protecting both wearers and individuals nearby.

The current pandemic situation presents a fresh and important obstacle for the current face recognition methods. These methods are now limited in their effectiveness as they can only analyze a small visible portion of the face. Consequently, it is necessary to verify if a person is wearing a mask and employ suitable techniques to accurately establish their identity. Therefore, it has become essential to incorporate innovative approaches to attain resilience in the existing systems.

Face masks cause occlusion and render specific facial characteristics unavailable during face detection. Additionally, masks invariably contribute incorrect features that prevent facial recognition systems from operating correctly. False characteristics taken from the lower portion of the face—the masked area, trick analysis algorithms, and hiding the fundamental attributes essential for face recognition, sharply increasing the rate of false positives. The proper identification of masked faces is a challenging research field since the masked features may have diverse inclinations, levels of occlusion, and diverse kinds of masks.

Face detection traditionally uses adaptive boosting and cascade classifiers for masked situations [14]. Detecting facial masks involves utilizing machine learning and deep learning methods, including those that rely on enhanced pre-trained and existing models known for their reliable performance. However, deep learning techniques have emerged at the forefront of masked face detection in recent years.

The use of deep neural networks, particularly convolution neural networks [15][16] that could effectively incorporate hierarchical features is generating much attention even though this methodology has been the subject of numerous suggestions. Studies have shown that these approaches outperform the other strategies because they automatically identify valuable features without requiring specially created attributes.

When conducting literature screening for masked face detection, it's essential to consider relevant research articles and studies that address the problem definition, datasets, methodology, evaluation metrics, real-time performance etc. The motivation behind selecting this topic can be driven by several factors and considerations. The widespread use of face masks, especially during global health crises like the COVID-19 pandemic, has highlighted the need for accurate and effective masked face detection systems.

Traditional face recognition and detection systems may struggle with masked faces due to the altered facial features. Accurate masked face detection is crucial in maintaining security measures in various public spaces. Detecting masked faces is a complex problem due to the varying types of masks, facial expressions, lighting conditions, and backgrounds. Masked face detection involves the intersection of computer vision, machine learning, image processing, and potentially medical imaging. Researching masked face detection can lead to the development of benchmark datasets, evaluation metrics, and standardized methodologies.

This article centers on a comprehensive examination of the latest advancements in facial recognition technology designed explicitly for identifying individuals wearing masks. And helps organizations and researchers in creating systems that may be increasingly effective and efficient. Section 2 provides the statistics of relevant literature and examines the most appropriate and effective methods for masked face detection and facial recognition published in literary works. Section 3 gives an overview of the features and the types of datasets used. Section 4 details the various methods and architectures used for masked facial detection. Finally, section 5 discusses approach limitations as well as prospective research approaches.

1.1. **Effect of face masks on facial detection and Recognition**

Several studies have recently been devoted to figuring out how face masks affect many facets of societal interaction and thinking [17][18]. The extensive research on face masks covers a wide range of topics. For example, research conducted before COVID-19 using masked and obscured faces have shown a decline in facial identification recognition [19]. According to a study [20], when different facial features were deleted, the accuracy rate of both recognized and unknown faces decreased. In a further investigation, the use of sunglasses reduced identification recognition's accuracy as well as speed [21].

Regardless of whether the faces were covered, surgical masks significantly negatively impacted the ability to match faces. The impairment was around the same size for both known and unknown faces. Additional research has demonstrated that faces decrease the accuracy of emotion detection or the severity of certain facial emotions, including happiness and rage [22].

In a significant study, Carbon [23] displayed faces featuring six important emotional reactions in two conditions: obvious and partially concealed (with masks). In the condition with the masked faces, they discovered lower levels of competence and confidence. Additionally, observers mistakenly regarded disgusted looks as appearing angry, whereas other emotions were judged to be neutral. These investigations all imply that face masks partially obstruct fundamental facial recognition mechanisms.

Prior research has mainly concentrated on the issue of how masks influence the recognition performance of facial detection or facial expression [24]. Even though most of these attempts highlighted the tremendous effects of masks on these factors, many real-world and theoretical questions have yet to be resolved. Measurement of categorization performance using masked faces might be considered a form of "reverse engineering. It can provide information on the significance of mouth region in classifying different facial dimensions.

The chart depicted in Figure 1 is produced using Biblioshiny and visualizes the annual production of scientific studies concerning face masks from 2012 to 2023, sourced from the Scopus database. It is worth mentioning that there was a significant increase in the number of publications between 2019 and 2022, mainly driven by the influence of the COVID-19 pandemic. Before COVID-19 pandemic, there were only a few studies focused on masked face detection, but it emerged as a significant and actively studied topic in that year and beyond. When it comes to the domain of 'masked face detection,' there exist only a handful of comprehensive reviews [29][52]. This comprehensive review on 'Masked face detection' intricately explores various facets, including a comparative study of two-stage and one-stage detectors, an investigation into the consequences of employing different backbone architectures alongside varying hyperparameters, an analysis of the impact of transfer learning, an examination of the significance of lightweight models, and an insightful exploration of the successful integration of Generative Adversarial Networks (GANs) into masked face recognition systems, among other crucial dimensions.

Figure 1. Annual Scientific production in the research area masked face detection from 2012 to 2022

This article centers on a comprehensive examination of the latest advancements in facial recognition technology designed explicitly for identifying individuals wearing masks. We have summarized 12 open datasets and provided their available links that could help AI researchers to use them quickly. Section 2 provides the statistics of relevant literature and examines the most appropriate and effective methods for masked face detection and facial recognition published in literary works. Section 3 gives an overview of the features and the types of datasets used. Section 4 details the various methods and architectures used for masked facial detection. Finally, section 5 discusses approach limitations as well as prospective research approaches.

2 REVIEW OF RELATED WORK

2.1 Masked Face Detection

Masks have emerged as a typical method for containing the virus during the COVID-19 outbreak. Certain IT corporations suggested using automatic mask-detecting systems with the existing camera systems. Negi. Et al. [25] proposed CNN and VGG16-based architectures for mask detection. The models were trained on a dataset called the Simulated Masked Face Dataset, comprising 1509 images. The purpose of these models is to support the current identification system used for monitoring public areas, determining whether individuals are wearing masks. The results indicate that the CNN algorithm achieves an efficiency of 96.35% during the training, validation, and testing phases.

In order to prevent the spread of the virus amidst the pandemic, there was a growing requirement to recognize individuals who were wearing masks. To resolve this issue, a comprehensive training pipeline built on the ArcFace approach was suggested by Montero, D. et al. [26] with several modifications in the architecture and loss functions. The network was created using a ResNet-50 modification. The training was conducted on the MS1MV2 dataset comprising 5.8 million photos and encompassing 85,000 distinct identities. The template achieved 99.78% accuracy with mask-usage validation during the testing phase. The results were superior to all others, with a 12% improvement in accuracy over the original.

The authors in [27] introduced a two-stage approach for detecting individuals wearing masks by utilizing the transfer models of Faster-RCNN and InceptionV2. These models were trained on a dataset consisting of 7,804 images, with 26,403 instances of people wearing masks. The findings demonstrate that this method outperforms previous methods with a precision of 97.32% for simple situations and 91.13% for complicated scenarios. Face mask detection studies were primarily dominated by deep learning methods, with conventional ML-based approaches rarely utilized. The efficient detection of people wearing masks is a challenge for the current monitoring systems. Alguzo [28] proposed a deep learning framework using a multi-graph convolutional network. The model achieved an accuracy rate of 97.9% after being trained on a collection of over 7600 photos from real-world masked face datasets.

The best framework for detecting face masks was determined by comparing two conventional Machine Learning classifiers, KNN and SVM, with one DL algorithm, Mobilenet, by Vijitkunsawat and Chantngarm [29]. The findings demonstrated that MobileNet surpasses KNN and SVM in terms of performance. The importance of deep learning in computer vision has motivated researchers to utilize it for the purpose of detecting face masks. InceptionV3 deep learning framework was fine-tuned by Chowdary et al. [30] to automate the mask detection approach.

A hybrid method for face mask detection that derives from both the conventional ML and DL-based approaches was described by Loey et al. in 2021 [10]. In this hybrid approach, the Resnet50 approach was employed to bring out features from images, which were subsequently utilized for training three different algorithms - SVM, an ensemble algorithm, and a decision tree. This training aimed to classify pictures into two categories: masked or unmasked.

Bhattacharya [31] introduced the 'Hybrid FaceMaskNet' technique to detect face masks. This method utilized conventional machine learning techniques, personalized feature extractors, and deep learning as its foundation. However, owing to the sparse amount of data, various manual and deep learning feature extraction techniques were used to retrieve relatively robust features.

Using an automated system, Qin and Li [32] could determine mask compliance assessment. The scientists created the SRCNet approach, which describes three classification issues dependent on unconstrained 2D facial images, by integrating image super-resolution and classification frameworks.

2.2 Masked Facial Recognition

Assessments have advanced steadily over the past several years, and face recognition efficiency has increased. When facial photographs are examined in controlled situations, it has moved to a very satisfactory level. However, particularly in unrestricted contexts, both inter and intra facial variances can lead to erroneous identification of individuals by a face detection system. Deep learning is among the notable advancements in machine learning, accompanied by various other significant developments. Notably, convolutional neural networks are recognized for their instrumental role in the success of deep learning approaches, owing to their ability to extract data at a higher level of abstraction. The effective development of AlexNet architecture by Krizhevsky et al. [33] marked a significant advancement in facial recognition. After that, several popular deep CNN systems were presented, including VGGNet [34], GoogleNet [35], as well as ResNet [36]. A recent study [82] addresses the issue of objects with "different sizes" and "different aspect ratios" in a single framework that prevented or exceeded human-level accuracy in the representation of human faces, such as noise in face photographs.

Despite its importance in the current situation, there seems to be little article in the field of masked facial recognition. The methodologies widely utilized to resolve problems, including partial faces or occlusion, may be used to overcome challenges with masked face identification. The literature's most effective methods for treating occlusions adapt between recovery and reconstruction strategies.

Wang et al. [37]] introduced a Face Attention Network to address occlusion by employing a feature pyramid structure. This model utilized multiple layers to handle faces of different sizes, thus offering a range of attention central points across various feature layers to tackle the issue of occlusion. Robust LSTM Autoencoders were proposed by Zhao et al. [38] and comprised 2 LSTM elements. They were designed to convert substantial occlusion faces and recover the occluded region.

In the research by Gawali and Deshmukh [39], the obstructed regions were eliminated using the Iterative Closest Point Approach (ICP). To handle the blocked sections, they carried out the restoration using a statistical assessment of the contours. However, a masked face conceals all essential facial clues, making masked facial recognition difficult. Furthermore, occlusion and partial face do not

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

typically cover all facial characteristics available for feature extraction.

Duan and Zhang suggested [40] Boost GAN for obstructed face pictures using the Generative Adversarial System, which could successfully generate and detect faces exhibiting significant pose variations, including concurrently distorted regions. Ding et al. have developed a unique technique to detect the hidden facial portion that is difficult to disguise with a mask, and this area is used to retrieve crucial discriminative information. A unique end-to-end architecture for masked feature extraction was developed by Li et al. [41], which initially enforces face completeness explicitly before transferring rich domain information from a general face identification model that has already been trained.

Due to a variety of face masks as well as other diversified parameters, such as variabilities in face viewpoints, social background, age, skin color, illumination, and sexual identity, etc., existing research on face mask detection as well as recognition has not been able to achieve suitable identification and classification effectiveness. Hence, a robust, effective, and consistent dataset is of utmost importance to address the challenges posed by both face mask detection and recognition.

3 MASKED FACIAL DETECTION DATASETS

Datasets consist of instances that are used to develop learning models. Users can either generate them by extracting content from the web or by exploring other websites online. Researchers from all around the world have suggested a variety of datasets to train identification or classification methods to track the conditions surrounding mask usage. Most datasets are compiled by gathering photographs from the Web. A different approach to quickly generating the necessary dataset in practical applications is to combine various datasets. It is recommended that researchers develop their datasets in this manner. In the meantime, it helps to enhance the different datasets by downloading multiple images from the Internet.

Mask-to-face matching is required to create simulated samples. Since their boundaries can be identified accurately and assist in providing appropriate models, large faces are often chosen to synthesize masked features. However, it is challenging to realize a solid mapping for small-sized faces due to incorrect landmarks and varying head poses. Table 1 provides detailed information regarding the different datasets available in masked facial detection systems.

Table 1: Masked Face Detection Datasets

4 METHODS FOR DETECTING MASKED FACES

Face mask detection approaches can be classified as hand-crafted feature techniques, conventional methods, and neural networks, depending on how the features are manipulated [52]. Particularly, methods based on neural networks are emerging, producing remarkable and exceptional outcomes. Some of the various techniques for masked face detection are presented in Figure.2 [52]. As stated in [53], existing face detection models can be roughly categorized into boosting-based and deformable models, which describe the face by its parts.

Figure 2. Methods used for detecting masked faces [52]

4.1 Conventional Methods

Traditional face detection techniques have had excellent returns over the years [54][55]. Most traditional approaches for detecting masked faces are centered on the idea that if a mask is worn correctly, neither the nose nor the mouth can be seen, and vice versa. One or more detectors are created using custom datasets or OpenCV's resources. Conventional techniques fall into two categories: single-detector and multiple-detector approaches, depending on the number of detectors used.

4.1.1 Single Detector Method

Dewantara et al. [14] used a nose-to-mouth classifier to identify multi-pose concealed faces. Like LBP and HOG characteristics, Haar is used to train the models. The face region is referred to as "masked" if the nose and mouth are not detected. If not, "No mask" will be indicated.

4.1.2 Multiple Detector Method

Nieto-Rodriguez et al. [56] utilized two AdaBoost sensors to detect surgical masks. The initial sensor employed LogitBoost to identify faces, while the second one employed GentleAdaBoost to detect masks. Considering the overlapping zones, they developed a cross-class elimination approach to preserve areas with higher confidence. The technique is simple to use and has a 95% accuracy rate.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Fang et al. [57] developed a method that detects masked features in real-time. This approach utilizes haar-like components to detect both the face and the mouth. Initially, the algorithm identifies the facial region and then proceeds to detect the mouth, determining whether a mask is present. Tengjiao He [58] used skin color and eye recognition for mask detection. Locating the face image is done initially by using an elliptical skin model with a geometric relationship between the eyes and other facial features. Next, to determine the conditions for wearing a mask, the skin color coverage in the lower half of the face is estimated. But only certain situations can be used with this approach.

Conventional techniques for masked face detection often rely on the AdaBoost algorithm and Haar-like characteristics. As a result, they can be effective in close-up situations with clearly visible facial features. However, classifiers need help to adjust to complicated scenarios, such as far distances and changing lighting, because of their poor learning capacity. Methods based on neural networks are data-driven frameworks that could offer workable solutions.

4.2. Neural Network-based detection methods

4.2.1. Single-stage methods

A significant percentage of the methods are single-stage, deep learning-based approaches. They are typically employed when implementing a study quickly. This is due to their way of predicting boundary frames and classes, which only uses one deep neural model and does not include a separate stage for the recommendation of boundary boxes. As a result, they have several uses for real-time detection systems.

4.2.1.1 Faster Region-Based Convolutional Neural Network

With the R-CNN as its foundation, this is a widely used advanced algorithm. It takes the place of the selective search technique previously utilized to determine RoI. Figure 3 provides a detailed illustration that illustrates the concept. Furthermore, this technique is preferred when precision is an issue.

Figure 3. Faster R-CNN [59]

Faster R-CNN architecture was used by Razavi et al. [60] to identify people who were not wearing masks or not at a safe distance. This was used in various road repair projects to monitor the workers and ensure they wore masks and maintained the correct physical distances. Unfortunately, the dataset is constrained because it only includes scenes related to construction work.

4.2.1.2 Context-Attention R-CNN

Context Attention R-CNN is a new architecture proposed by Zhang et al. [61] for masked facial identification that combines multiple context feature extractor components, attention component, and several decoupling branches. Through the extraction of specific attributes, it can increase intra-class differences and decrease inter-class similarities. This study created a new dataset with 8635 faces with different conditions for experimental confirmation. There is, however, an imbalance between the classes in the dataset.

4.2.1.3 You Only Look Once (YOLO)

Wang et al. [62] suggested an all-encompassing edge-computing approach to identify masked faces. This serverless in-browser solution integrates YOLO and may be used on any device having a browser. A YOLOv2 with a ResNet-50 detector was developed by Loey et al. [63] to identify medical face masks. There are two steps in the procedure. The initial approach utilizes a deep transfer learning algorithm to extract features [63]. YOLOv2 implements the second component for masked face detection. This architecture has drawn attention due to its quickness. Jiang et al. [45] designed Squeeze and Excitation(SE) YOLOv3 to balance masked facial detection's effectiveness and running speed. SE is integrated into Darknet-53 to extract essential features, and GIoUloss, focal loss, is used to enhance balance and vibrancy [52]. A new face mask detection dataset of 52635 images was used by Kumar et al. [64] to assess both the original YOLO and tiny alternatives. Due to its enhanced feature extraction network, this improved small YOLOv4 is suggested as a successful and effective veiled masked face detector. Furthermore, it outperforms rival models because of its superior ability to learn even from generalized photos of the entities and make predictions with high precision.

4.2.1.4 Retina Face Mask

Retina Face Mask detector utilizes pixel-level facial localization to predict all key facial attributes in one step. Its core concept involves leveraging feature pyramid networks to merge rich semantic information. This detector incorporates an innovative context focus module to enhance focus on facial features and masks. With this method, face detection and 2D face alignment are carried out to the highest standards, and robust 3D face reconstruction is accomplished with a single-shot approach.[65].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.2.1.5 InceptionV3

Chowdary et al. [30] utilized the trained algorithm of InceptionV3 to ascertain the presence of a mask on an individual. This approach is classified as a transfer learning paradigm, as the final level of InceptionV3 is replaced with five layers. The Simulated Face Dataset was utilized for both training and testing. Image augmentation techniques were implemented to optimize the efficiency of model training and testing to overcome the scarcity of available data. It is reported to achieve a remarkable accuracy of 99.9% on a synthetic dataset.

4.2.1.6 MobileNet

To curb the transmission of SARS-COV-2, Dey et al. [48] developed a method called MobileNetMask, which utilizes in-depth training to identify face masks in multiple stages. The mask classifiers rely on SSD and ResNet-10 ROI recognition for their operation. MobileNet-V2 demonstrates its suitability for compact devices by demanding minimal computational power and adopting a mobile-centric approach, enabling efficient detection of face mask presence or absence in video streams. Comparative studies suggest that it outperforms alternative methods in terms of accuracy.

4.2.2 Two-Stage Method

Face pre-detection and face recognition (or face classification) are the two primary stages of two-stage techniques. Facial pre-detection is often performed by a large number of face detectors [66] [67] and, otherwise, object detectors [62] [68]. In the initial stage, object detectors may also offer facial feature descriptions. Several classifiers or algorithms form the second stage. Integrating an object detector with a classification algorithm can complete the task of detecting masked faces.

4.2.2.1 Neural Network + Neural Network

Wang et al. [27] introduced a two-step methodology for mask detection utilizing hybrid machine-learning techniques. In the initial step, they employed a deep training transfer framework called Faster R-CNN [69]. In the subsequent phase, they utilized a comprehensive learning system [70]. The input image is initially processed in the pre-detection phase, generating multiple potential regions of interest. These regions are then examined using a programmed Broad Learning System (BLS) model, effectively reducing false positives while accurately identifying masked faces. Ultimately, observed results are given labels for identification. Annotated data is essential for training the pre-detection algorithm. The faces identified and their corresponding masks can be utilized to create customized datasets for specific categorization purposes, distinguishing between masked and unmasked individuals and correct and incorrect mask usage. Regarding accuracy, faster R-CNN performs better than SSD and YOLO. In the validation phase, BLS is employed, utilizing a flat neural network topology known for its relatively high learning efficiency. Several versions of BLS have been introduced.

4.2.2.2 Neural Network + Hand Crafted

Zereen et al. [71] advanced a two-step technique to detect concealed faces and enforce compliance. The primary focus is on extracting facial features. The first step involves employing the MTCNN algorithm to identify whether the individual being analyzed is donning a face mask with multiple colors. It then proceeds to ascertain whether the person wears a mask matching their skin tone. This methodology aims to categorize face images into five distinct types. Remarkably, it attains a commendable accuracy rate of 97.13% while effectively detecting masks of different colors, especially when a skin-colored mask is involved. However, employing multiple methods results in higher computational costs, and incorporating an empirical threshold limits its adaptability.

4.2.2.3 Hand-crafted Feature + Neural Network

Rudraraju et al. [72] used a combination of haar-like sequence classifiers plus two MobileNet algorithms to recognize face masks. First, a classifier detects face parts. The initial MobileNet algorithm is employed to distinguish between masks and without masks. The second version is utilized to differentiate between correct and wrong mask usage. In another investigation, using transfer learning algorithms by CNN are employed to evaluate whether an individual is wearing a mask [73]. Multiple models, such as VGG16, MobileNetV2, etc., are trained using a single dataset. By accumulating additional data, refining pre-existing models, and employing ensemble models, the accuracy and dependability of facial recognition systems can be substantially enhanced when identifying altered and profiled faces wearing masks.

Numerous face recognition methods employ a two-stage approach that combines a face recognition system with a classification algorithm. Usually, the pre-detection and classification models are trained separately, resulting in longer training durations when compared to single-stage methods. However, two-stage approaches offer several benefits, such as identifying small items, categorizing multiclass, and facilitating cross-class elimination.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.2.3 Multi-stage method

Multi-stage approaches typically include several processing phases. The primary idea behind methods [74] [75] is to estimate human posture. First, a set number of essential points in one individual is estimated. The ROI from the main image is then extracted using critical spots in the facial regions. ROI is then standardized and passed to a trained model to determine class. Some more procedures might be necessary to improve effectiveness.

At least two deep learning systems are used in multistage approaches. The design is more complex than one and two-stage techniques. Its primary goal is to increase the performance of masked face recognition. Experimental data from the original literature also support this argument. The disadvantage is obvious: numerous networks necessitate many calculations and expensive computing machines.

5. DISCUSSION AND FUTURE DIRECTIONS

Facial Detection can measure the strength and density of the crowd in an open space for crowd analysis. However, the effectiveness of the face detector can be affected by obstructions within the facial frame image, such as the presence of a mask. Before the onset of COVID-19, research studies were scarce and focused on identifying faces when wearing masks, primarily because there needed to be more datasets specifically tailored for masked facial recognition. However, the strategic review indicates that most datasets employed for face mask detection were artificially generated, resulting in a need for more accuracy in representing real-life situations. As a result, this disparity negatively affects the model's performance when applied to real-world scenarios.

In addition to the effective utilization of Deep Learning techniques for face detection, the practical application of AI methods for face mask detection is still in its early stages. Recognizing face masks has posed significant difficulties in image processing, especially during the COVID-19 pandemic, owing to various camera resolutions, obstacles, and variations in posture, lighting, angles, and other aspects. Most deep learning networks typically rely on input images that have a fixed resolution and size. However, when the resolutions of the input photographs differ, resizing them to a consistent size can lead to degraded image quality and distortion in facial regions, thereby reducing face detection performance. Moreover, nature-inspired computing models are comparatively less efficient than deep learning models when detecting face masks.

Most face mask detection models utilizing Deep Learning techniques were published in 2021. Research in this area indicates that the Super-resolution and classification networks (SRCNet) model [32] can accurately determine whether a person is wearing a face mask with a 98.70% accuracy rate. The combination of an SR network and a conventional facial recognition model accomplishes this. To evaluate this model, extensive facial image datasets, including the Medical Masks Dataset, were used. However, it should be noted that these datasets were relatively small, and the model did not account for variations in posture or dynamic environments when identifying face mask-wearing conditions. Therefore, a hybrid deep transfer learning model [10] was developed for face mask detection, incorporating SVM, decision trees, and ensemble methods. By leveraging three datasets, namely RMFD, SMFD, and LFW, the model attained an impressive testing accuracy of 99.64%. However, real-life video streaming was not included in this study. A different method, described in reference [30], utilizes transfer learning to classify individuals who do not have face masks. It employs Google's InceptionV3 pre-trained model, which consists of a convolutional neural network with 48 layers. To enhance the variety of training data, the technique applies picture augmentation. This method attained a training accuracy of 99.92% and a testing accuracy of 100% when applied to the SMFD dataset. Despite being the most precise model, its assessment was confined to a limited dataset and did not encompass a video streaming dataset.

One-stage detectors surpass two-stage detectors in speed, especially when prioritizing efficiency, although two-stage detectors demonstrate superior precision. Therefore, several algorithms have been developed specifically for devices with limited resources. A notable example within this category is MaskFAN [76], which utilizes a lightweight backbone, incorporates a modified loss function, and employs data augmentation techniques to improve the model's performance. Another novel approach [77] introduces a pose-specific categorization system that minimizes computational requirements. However, ensuring high performance with lightweight equipment remains a challenging problem for existing methodologies.

All neural network techniques are based on appearance, necessitating much-balanced information to train classifiers. Hence there is a need for creating higher balanced datasets. Generally, a benchmark dataset containing many data, many classes of mask-wearing settings, a variety of masked face types, and the correct ratio of realistic images is needed. To achieve incredible accuracy, many neural network systems can be integrated. Making a valid identification with a decent trade between efficiency and accuracy is an intriguing study path. Generally, resizing images with different resolutions is necessary; for most deep neural networks, the input size of the image is fixed. Attempting to process various image resolutions productively is an essential topic in future research. Future studies may implement deeper and broader deep learning models with excellent training parameters to develop more accurate face detection.

Exploring techniques for image editing and object removal to ensure overall coherence and restore extensive missing areas presents an intriguing problem. GAN-based approaches, known for their robust learning capabilities, are regarded as effective solutions. These approaches prove valuable in masked face detection[78] [79] and masked face expression detection. Additionally, there is a need to concentrate on face masking and related biometrics for multi-modal recognition [80]. By accurately predicting the positions of the landmarks, face alignment algorithms can help with a mixture of jobs, such as face recognition, expression analysis, and even virtual makeup or facial transformation applications. Overall, the challenge of face alignment with masks underscores the need for continued research and development in this area and the importance of considering real-world factors that can impact the performance of face alignment algorithms in practice [52]. Several solutions utilizing neural networks have been suggested by researchers [76] [81] to tackle the issue. However, there are still numerous valuable endeavors aimed at enhancing accuracy and minimizing inference time. Research in this area is considered promising.

Future research should focus on developing novel techniques to improve the accuracy of detection algorithms. This can involve exploring advanced deep learning architectures, incorporating contextual information, and leveraging additional modalities such as thermal imaging or 3D data to enhance detection performance. Another crucial challenge is addressing the wide range of mask variations, including different types of masks, colors, textures, and styles. Future algorithms should be designed to handle these variations effectively, ensuring robustness across different mask types and avoiding false positives or negatives. Future research should focus on developing efficient algorithms that can process video streams or large-scale datasets in real-time without sacrificing accuracy. Optimization techniques, parallel processing, and hardware acceleration can be explored to achieve the desired real-time performance. Transfer learning approaches can be investigated to leverage pre-trained models on large-scale face recognition datasets and fine-tune them for masked face detection tasks. This can help in improving the generalization capabilities of the algorithms.

By addressing these challenges and exploring the outlined future directions, the accuracy of masked face detection algorithms can be significantly improved, enabling their effective deployment in real-world applications and contributing to the development of more robust and reliable systems.

5.1 Observations

- Even though deep learning models performs efficiently at high accuracy, the use of various backbone architectures with different hyperparameters could give them even greater precision.
- Transfer learning, i.e., using pre-trained models like MobileNet, Inception V3, VGG-16, and others, is recommended because training a deep neural network is expensive due to the high computational complexity.
- Although two-stage detectors are more accurate, one-stage detectors surpass them in real-time applications.
- The model's performance has also been impacted by the poor snaps in the datasets, such as those with insufficient lighting, side viewpoints, etc.
- For real-time detection on low-resource devices, lightweight models are preferred.
- The successful integration of GANs into masked face recognition systems demands a thoughtful approach to design, training, and evaluation in order to ensure the realism and utility of the generated images for the recognition task. By producing a wide array of authentic masked face variations, GANs play a pivotal role in

bolstering the recognition models' ability to generalize effectively.

- Furthermore, dataset restrictions like inconsistent mask distribution and other biases highlight the ongoing work needed to generate representative and equitable training data for precise masked face identification algorithms.
- It is difficult to adapt traditional Face Recognition (FR) approaches to the unique requirements of Masked Face Recognition (MFR), as attempts to do so typically result in observable performance decrease.

6. CONCLUSION

To cope with the epidemic effectively, establishing centralized mechanisms capable of autonomously determining whether an individual is using a facial mask or otherwise has emerged as an intriguing subject for those specializing in this field. A plethora of studies has recently commenced in this field. However, the motivation of this work is to present a comprehensive evaluation of the many options for implementing such an effective system. In this paper, a variety of masked face detection approaches are studied, and it is determined that deep learning techniques are the most promising and often used ones because of their versatility, representation learning capabilities, scalability, etc. This study examines noteworthy advancements in the field of detecting masked faces. Various datasets were evaluated based on their source of images, image realism, class distribution, and experimental outcomes. Despite many datasets being available, the RMFD dataset is widely used because of its popularity which can be attributed to its realistic and diverse nature, its balanced class distribution, and its availability to the research community. The optimization of hyperparameters within existing frameworks is a promising avenue for future research aimed at achieving quicker and enhanced results. Additionally, machine learning techniques can be employed to explore innovative strategies for extracting features.

REFERENCES:

[1] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, “Past, Present, and Future of Face Recognition: A Review,” Electronics, vol. 9, no. 8, p. 1188, Jul. 2020, doi: 10.3390/electronics9081188.

[2] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, “Deep Learning for Computer Vision: A Brief Review,” Comput. Intell. Neurosci., vol. 2018, pp. 1–13, 2018, doi: 10.1155/2018/7068349.

[3] B. K. Chakraborty, D. Sarma, M. K. Bhuyan, and K. F. MacDorman, “Review of constraints on vision-based gesture recognition for human–computer interaction,” IET Comput. Vis., vol. 12, no. 1, pp. 3–15, 2018.

[4] S. Qiu, Q. Liu, S. Zhou, and C. Wu, “Review of Artificial Intelligence Adversarial Attack and Defense Technologies,” Appl. Sci., vol. 9, no. 5, p. 909, Mar. 2019, doi: 10.3390/app9050909.

[5] K. Dang and S. Sharma, “Review and comparison of face detection algorithms,” in 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, Noida, India: IEEE, Jan. 2017, pp. 629–633. doi: 10.1109/CONFLUENCE.2017.7943228.

[6] C. M. Cook, J. J. Howard, Y. B. Sirotin, J. L. Tipton, and A. R. Vemury, “Demographic Effects in Facial Recognition and Their Dependence on Image Acquisition: An Evaluation of Eleven Commercial Systems,” IEEE Trans. Biom. Behav. Identity Sci., vol. 1, no. 1, pp. 32–41, Jan. 2019, doi: 10.1109/TBIOM.2019.2897801.

[7] S. Dargan and M. Kumar, “A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities,” Expert Syst. Appl., vol. 143, p. 113114, Apr. 2020, doi: 10.1016/j.eswa.2019.113114.

[8] M. Wang and W. Deng, “Deep Face Recognition: A Survey,” Neurocomputing, vol. 429, pp. 215–244, Mar. 2021, doi: 10.1016/j.neucom.2020.10.081.

[9] X. Fan and M. Jiang, “RetinaFaceMask: A Single Stage Face Mask Detector for Assisting Control of the COVID-19 Pandemic.” arXiv, Dec. 15, 2021. Accessed: Jan. 04, 2023. [Online]. Available: http://arxiv.org/abs/2005.03950

[10] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, “A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic,” Measurement, vol. 167, p. 108288, Jan. 2021, doi: 10.1016/j.measurement.2020.108288.

[11] K. Karthik et al., “Biosafety Concerns During the Collection, Transportation, and Processing

of COVID-19 Samples for Diagnosis," Arch. Med. Res., vol. 51, no. 7, pp. 623–630, Oct. 2020, doi: 10.1016/j.arcmed.2020.08.007.

[12] T. M. L. Souza and C. M. Morel, "The COVID-19 pandemics and the relevance of biosafety facilities for metagenomics surveillance, structured disease prevention and control," Biosaf. Health, vol. 3, no. 1, pp. 1–3, Feb. 2021, doi: 10.1016/j.bsheal.2020.11.007.

[13] K. Simonyan, O. Parkhi, A. Vedaldi, and A. Zisserman, "Fisher Vector Faces in the Wild," in Procedings of the British Machine Vision Conference 2013, Bristol: British Machine Vision Association, 2013, p. 8.1-8.11. doi: 10.5244/C.27.8.

[14] B. S. Bayu Dewantara and D. Twinda Rhamadhaningrum, "Detecting Multi-Pose Masked Face Using Adaptive Boosting and Cascade Classifier," in 2020 International Electronics Symposium (IES), Surabaya, Indonesia: IEEE, Sep. 2020, pp. 436–441. doi: 10.1109/IES50839.2020.9231934.

[15] C. Peng, N. Wang, J. Li, and X. Gao, "DLFace: Deep local descriptor for cross-modality face recognition," Pattern Recognit., vol. 90, pp. 161–171, Jun. 2019, doi: 10.1016/j.patcog.2019.01.041.

[16] C. Li, Y. Huang, W. Huang, and F. Qin, "Learning features from covariance matrix of gabor wavelet for face recognition under adverse conditions," Pattern Recognit., vol. 119, p. 108085, Nov. 2021, doi: 10.1016/j.patcog.2021.108085.

[17] I. Molnar-Szakacs, L. Q. Uddin, and M. B. Heffernan, "The face behind the mask: The future of interpersonal interaction," Neuron, vol. 109, no. 12, pp. 1918–1920, Jun. 2021, doi: 10.1016/j.neuron.2021.05.030.

[18] E. Noyes, J. P. Davis, N. Petrov, K. L. H. Gray, and K. L. Ritchie, "The effect of face masks and sunglasses on identity and expression recognition with super-recognizers and typical observers," R. Soc. Open Sci., vol. 8, no. 3, p. rsos.201169, 201169, Mar. 2021, doi: 10.1098/rsos.201169.

[19] T. I. Dhamecha, R. Singh, M. Vatsa, and A. Kumar, "Recognizing Disguised Faces: Human and Machine Evaluation," PLoS ONE, vol. 9, no. 7, p. e99212, Jul. 2014, doi: 10.1371/journal.pone.0099212.

[20] B. C. M. Stephan and D. Caine, "What is in a View? The Role of Featural Information in the Recognition of Unfamiliar Faces across Viewpoint Transformation," Perception, vol. 36, no. 2, pp. 189–198, Feb. 2007, doi: 10.1068/p5627.

[21] D. Fitousi and M. J. Wenger, "Variants of independence in the perception of facial identity and expression.," J. Exp. Psychol. Hum. Percept. Perform., vol. 39, no. 1, pp. 133–155, Feb. 2013, doi: 10.1037/a0028001.

[22] M. Calbi et al., "The consequences of COVID-19 on social interactions: an online study on face covering," Sci. Rep., vol. 11, no. 1, p. 2601, Jan. 2021, doi: 10.1038/s41598-021-81780-w.

[23] C.-C. Carbon, "Wearing Face Masks Strongly Confuses Counterparts in Reading Emotions," Front. Psychol., vol. 11, p. 566886, Sep. 2020, doi: 10.3389/fpsyg.2020.566886.

[24] F. Grundmann, K. Epstude, and S. Scheibe, "Face masks reduce emotion-recognition accuracy and perceived closeness," PsyArXiv, preprint, Oct. 2020. doi: 10.31234/osf.io/xpzs3.

[25] A. Negi, K. Kumar, P. Chauhan, and R. S. Rajput, "Deep Neural Architecture for Face mask Detection on Simulated Masked Face Dataset against Covid-19 Pandemic," in 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India: IEEE, Feb. 2021, pp. 595–600. doi: 10.1109/ICCCIS51004.2021.9397196.

[26] D. Montero, M. Nieto, P. Leskovsky, and N. Aginako, "Boosting Masked Face Recognition with Multi-Task ArcFace." arXiv, Apr. 21, 2021. Accessed: Feb. 05, 2023. [Online]. Available: http://arxiv.org/abs/2104.09874

[27] B. Wang, Y. Zhao, and C. L. P. Chen, "Hybrid Transfer Learning and Broad Learning System for Wearing Mask Detection in the COVID-19 Era," IEEE Trans. Instrum. Meas., vol. 70, pp. 1–12, 2021, doi: 10.1109/TIM.2021.3069844.

[28] A. Alguzo, A. Alzu'bi, and F. Albalas, "Masked Face Detection using Multi-Graph Convolutional Networks," in 2021 12th International Conference on Information and Communication Systems (ICICS), Valencia, Spain: IEEE, May 2021, pp. 385–391. doi: 10.1109/ICICS52457.2021.9464553.

[29] W. Vijitkunsawat and P. Chantngarm, "Study of the Performance of Machine Learning Algorithms for Face Mask Detection," in 2020 - 5th International Conference on Information Technology (InCIT), Chonburi, Thailand: IEEE, Oct. 2020, pp. 39–43. doi: 10.1109/InCIT50588.2020.9310963.

[30] G. J. Chowdary, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, “Face Mask Detection using Transfer Learning of InceptionV3,” 2020, pp. 81–90. doi: 10.1007/978-3-030-66665-1_6.

[31] M. sadhukhan and I. Bhattacharya, “HybridFaceMaskNet: A Novel Face-Mask Detection Framework Using Hybrid Approach,” In Review, preprint, Jul. 2021. doi: 10.21203/rs.3.rs-476241/v1.

[32] B. Qin and D. Li, “Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19,” Sensors, vol. 20, no. 18, p. 5236, Sep. 2020, doi: 10.3390/s20185236.

[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[34] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition.” arXiv, Apr. 10, 2015. Accessed: Feb. 05, 2023. [Online]. Available: http://arxiv.org/abs/1409.1556

[35] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[36] O. Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge.” arXiv, Jan. 29, 2015. Accessed: Feb. 05, 2023. [Online]. Available: http://arxiv.org/abs/1409.0575

[37] J. Wang, Y. Yuan, and G. Yu, “Face Attention Network: An Effective Face Detector for the Occluded Faces.” arXiv, Nov. 22, 2017. Accessed: Feb. 06, 2023. [Online]. Available: http://arxiv.org/abs/1711.07246

[38] F. Zhao, J. Feng, J. Zhao, W. Yang, and S. Yan, “Robust LSTM-Autoencoders for Face De-Occlusion in the Wild,” IEEE Trans. Image Process., vol. 27, no. 2, pp. 778–790, 2018, doi: 10.1109/TIP.2017.2771408.

[39] S. Gawali and R. Deshmukh, “3D Face Recognition Using Geodesic Facial Curves to Handle Expression, Occlusion and Pose Variations,” Int. J. Comput. Sci. IT, vol. 5, pp. 4284–4287, May 2014.

[40] Q. Duan and L. Zhang, “BoostGAN for Occlusive Profile Face Frontalization and Recognition.” arXiv, Feb. 26, 2019. Accessed: Feb. 12, 2023. [Online]. Available: http://arxiv.org/abs/1902.09782

[41] S. Ge, J. Li, Q. Ye, and Z. Luo, “Detecting Masked Faces in the Wild with LLE-CNNs,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 426–434. doi: 10.1109/CVPR.2017.53.

[42] Z. Wang et al., “Masked Face Recognition Dataset and Application.” arXiv, Mar. 23, 2020. Accessed: Feb. 06, 2023. [Online]. Available: http://arxiv.org/abs/2003.09093

[43] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, “MaskedFace-Net – A dataset of correctly/incorrectly masked face images in the context of COVID-19,” Smart Health, vol. 19, p. 100144, Mar. 2021, doi: 10.1016/j.smhl.2020.100144.

[44] B. Roy, S. Nandy, D. Ghosh, D. Dutta, P. Biswas, and T. Das, “MOXA: A Deep Learning Based Unmanned Approach For Real-Time Monitoring of People Wearing Medical Masks,” Trans. Indian Natl. Acad. Eng., vol. 5, no. 3, pp. 509–518, Sep. 2020, doi: 10.1007/s41403-020-00157-z.

[45] X. Jiang, T. Gao, Z. Zhu, and Y. Zhao, “Real-Time Face Mask Detection Method Based on YOLOv3,” Electronics, vol. 10, no. 7, p. 837, Apr. 2021, doi: 10.3390/electronics10070837.

[46] F. I. Eyiokur, H. K. Ekenel, and A. Waibel, “Unconstrained Face-Mask & Face-Hand Datasets: Building a Computer Vision System to Help Prevent the Transmission of COVID-19,” Signal Image Video Process., Jul. 2022, doi: 10.1007/s11760-022-02308-x.

[47] B. Batagelj, P. Peer, V. Štruc, and S. Dobrišek, “How to Correctly Detect Face-Masks for COVID-19 from Visual Information?,” Appl. Sci., vol. 11, no. 5, p. 2070, Feb. 2021, doi: 10.3390/app11052070.

[48] S. K. Dey, A. Howlader, and C. Deb, “MobileNet Mask: A Multi-phase Face Mask Detection Model to Prevent Person-To-Person Transmission of SARS-CoV-2,” in Proceedings of International Conference on Trends in Computational and Cognitive Engineering, M. S. Kaiser, A. Bandyopadhyay, M. Mahmud, and K. Ray, Eds., in Advances in Intelligent Systems and Computing, vol. 1309. Singapore: Springer Singapore, 2021, pp. 603–613. doi: 10.1007/978-981-33-4673-4_49.

[49] S. Singh, U. Ahuja, M. Kumar, K. Kumar, and M. Sachdeva, “Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment,” Multimed. Tools Appl., vol.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

80, no. 13, pp. 19753–19768, May 2021, doi: 10.1007/s11042-021-10711-8.

[50] D. L. de Medeiros, "AIZOO Dataset Dataset," Roboflow Universe. Roboflow, Apr. 2022. [Online].Available: https://universe.roboflow.com/diego-lucena-de-medeiros/aizoo-dataset

[51] "Mask Dataset," Make ML. [Online]. Available: https://makeml.app/datasets/mask

[52] B. Wang, J. Zheng, and C. L. P. Chen, "A Survey on Masked Facial Detection Methods and Datasets for Fighting Against COVID-19," IEEE Trans. Artif. Intell., vol. 3, no. 3, pp. 323–343, Jun. 2022, doi: 10.1109/TAI.2021.3139058.

[53] S. Zafeiriou, C. Zhang, and Z. Zhang, "A survey on face detection in the wild: Past, present and future," Comput. Vis. Image Underst., vol. 138, pp. 1–24, Sep. 2015, doi: 10.1016/j.cviu.2015.03.015.

[54] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA: IEEE Comput. Soc, 2001, p. I-511-I–518. doi: 10.1109/CVPR.2001.990517.

[55] P. Viola and M. Jones, "Robust Real-time Object Detection".

[56] A. Nieto-Rodríguez, M. Mucientes, and V. M. Brea, "System for Medical Mask Detection in the Operating Room Through Facial Attributes," in Pattern Recognition and Image Analysis, R. Paredes, J. S. Cardoso, and X. M. Pardo, Eds., in Lecture Notes in Computer Science, vol. 9117. Cham: Springer International Publishing, 2015, pp. 138–145. doi: 10.1007/978-3-319-19390-8_16.

[57] T. Fang, X. Huang, and J. Saniie, "Design Flow for Real-Time Face Mask Detection Using PYNQ System-on-Chip Platform," in 2021 IEEE International Conference on Electro Information Technology (EIT), Mt. Pleasant, MI, USA: IEEE, May 2021, pp. 1–5. doi: 10.1109/EIT51626.2021.9491842.

[58] T. He, "Mask Wearing Detection Method Based on the Skin Color and Eyes Detection," J. Phys. Conf. Ser., vol. 1906, no. 1, p. 012012, May 2021, doi: 10.1088/1742-6596/1906/1/012012.

[59] "Faster R-CNN | ML," GeeksforGeeks, Feb. 27, 2020. https://www.geeksforgeeks.org/faster-r-cnn-ml/ (accessed Feb. 13, 2023).

[60] M. Razavi, H. Alikhani, V. Janfaza, B. Sadeghi, and E. Alikhani, "An Automatic System to Monitor the Physical Distance and Face Mask Wearing of Construction Workers in COVID-19 Pandemic," SN Comput. Sci., vol. 3, no. 1, p. 27, Jan. 2022, doi: 10.1007/s42979-021-00894-0.

[61] J. Zhang, F. Han, Y. Chun, and W. Chen, "A Novel Detection Framework About Conditions of Wearing Face Mask for Helping Control the Spread of COVID-19," IEEE Access, vol. 9, pp. 42975–42984, 2021, doi: 10.1109/ACCESS.2021.3066538.

[62] Z. Wang, P. Wang, P. C. Louis, L. E. Wheless, and Y. Huo, "WearMask: Fast in-browser face mask detection with serverless edge computing for COVID-19," Electron. Imaging, vol. 35, no. 11, pp. 229-1-229–1, 2023, doi: 10.2352/EI.2023.35.11.HPCI-229.

[63] 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," Sustain. Cities Soc., vol. 65, p. 102600, Feb. 2021, doi: 10.1016/j.scs.2020.102600.

[64] A. Kumar, A. Kalia, K. Verma, A. Sharma, and M. Kaushal, "Scaling up face masks detection with YOLO on a novel dataset," Optik, vol. 239, p. 166744, Aug. 2021, doi: 10.1016/j.ijleo.2021.166744.

[65] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA: IEEE, Jun. 2020, pp. 5202–5211. doi: 10.1109/CVPR42600.2020.00525.

[66] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," Artif. Intell. Rev., vol. 52, no. 2, pp. 927–948, Aug. 2019, doi: 10.1007/s10462-018-9650-2.

[67] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," IEEE Signal Process. Lett., vol. 23, no. 10, pp. 1499–1503, Oct. 2016, doi: 10.1109/LSP.2016.2603342.

[68] W. Liu et al., "SSD: Single Shot MultiBox Detector," 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.

[69] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans.

Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[70] C. L. P. Chen and Z. Liu, "Broad Learning System: An Effective and Efficient Incremental Learning System Without the Need for Deep Architecture," IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 1, pp. 10–24, 2018, doi: 10.1109/TNNLS.2017.2716952.

[71] A. N. Zereen, S. Corraya, M. N. Dailey, and M. Ekpanyapong, "Two-Stage Facial Mask Detection Model for Indoor Environments," in Proceedings of International Conference on Trends in Computational and Cognitive Engineering, M. S. Kaiser, A. Bandyopadhyay, M. Mahmud, and K. Ray, Eds., Singapore: Springer Singapore, 2021, pp. 591–601.

[72] S. R. Rudraraju, N. K. Suryadevara, and A. Negi, "Face Mask Detection at the Fog Computing Gateway," presented at the 2020 Federated Conference on Computer Science and Information Systems, Sep. 2020, pp. 521–524. doi: 10.15439/2020F143.

[73] J. Tomás, A. Rego, S. Viciano-Tudela, and J. Lloret, "Incorrect Facemask-Wearing Detection Using Convolutional Neural Networks with Transfer Learning," Healthcare, vol. 9, no. 8, p. 1050, Aug. 2021, doi: 10.3390/healthcare9081050.

[74] H. C. V. Lopes and P. Garza, "Monitoring COVID-19 prevention measures on CCTV cameras using Deep Learning".

[75] H. Lin, R. Tse, S.-K. Tang, Y. Chen, W. Ke, and G. Pau, "Near-Realtime Face Mask Wearing Recognition Based on Deep Learning," in 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA: IEEE, Jan. 2021, pp. 1–7. doi: 10.1109/CCNC49032.2021.9369493.

[76] Y. Sha, J. Zhang, X. Liu, Z. Wu, and S. Shan, "Efficient Face Alignment Network For Masked Face," in 2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2021, pp. 1–6. doi: 10.1109/ICMEW53276.2021.9456012.

[77] M. Asad, A. Hussain, and U. Mir, "Low complexity hybrid holistic–landmark based approach for face recognition," Multimed. Tools Appl., vol. 80, pp. 30199–30212, 2021.

[78] M. Geng, P. Peng, Y. Huang, and Y. Tian, "Masked Face Recognition with Generative Data Augmentation and Domain Constrained Ranking," in Proceedings of the 28th ACM International Conference on Multimedia, Seattle WA USA: ACM, Oct. 2020, pp. 2246–2254. doi: 10.1145/3394171.3413723.

[79] H. Deng, Z. Feng, G. Qian, X. Lv, H. Li, and G. Li, "MFCosface: A Masked-Face Recognition Algorithm Based on Large Margin Cosine Loss," Appl. Sci., vol. 11, no. 16, p. 7310, Aug. 2021, doi: 10.3390/app11167310.

[80] M. I. Amin, M. Adeel Hafeez, R. Touseef, and Q. Awais, "Person Identification with Masked Face and Thumb Images under Pandemic of COVID-19," in 2021 7th International Conference on Control, Instrumentation and Automation (ICCIA), Tabriz, Iran: IEEE, Feb. 2021, pp. 1–4. doi: 10.1109/ICCIA52082.2021.9403577.

[81] T. Wen, Z. Ding, Y. Yao, WeiZhang, Y. Ge, and X. Qian, "Towards Efficient Masked-Face Alignment Via Cascaded Regression," in 2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Jul. 2021, pp. 1–5. doi: 10.1109/ICMEW53276.2021.9455962.

[82] P. V. Lal, U. Srilakshmi, and D. Venkateswarlu, "FACE RECOGNITION USING DEEP LEARNINGXCEPTION CNN METHOD," Journal of Theoretical and Applied Information Technology, vol. 100, no. 2, 2022.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Masked Face Detection Datasets

S.NO	Name of the Dataset	Description	Annotation	Link
1.	MAFA [41]	* The images are all sourced from the Web. * Every face region has six properties that are individually annotated. * It is more of an occluded face dataset. * Detecting obscured faces.	Yes	https://drive.google.com/drive/folders/1nbtM1n0–iZ3VVbNGhocxbnBGhMau OG
2.	Masked Face Detection Dataset - MFDD [42]	* The photographs are obtained through the Internet. * The dataset solely focuses on the masked face class. * It provides the capability to train detection models that can ascertain whether someone is wearing a mask or not.	No	https://github.com/X-zhangyang/ Real-World-Masked-Face-Dataset
3.	Masked Face Net Image Dataset -MFNID [43]	*The framework consists of four stages: face recognition, facial features identification, mask-to-facial mapping as well as manual picture filtering * It includes a single variety of modeled masks. * To produce simulated accurate or inaccurate masked faces.	No	https://github.com/cabani/MaskedFace-Net
4.	Moxa3K [44]	* The dataset carefully takes into account boundary conditions. *It has a range of samples, including those with rotation, congestion, blurriness, and various lighting circumstances. * Researchers are provided with additional choices to train different machine learning algorithms, thereby enhancing the stability of masked face detection.	Yes	https://shitty-bots-inc.github.io/MOXA/index.html
5.	Properly Wearing Masked Face Detection-PWMFD [45]	*The images are sourced from datasets such as MAFA MFDD and Wider Face. in addition to the internet * The provided dataset enables the creation of a model capable of identifying individuals based on their adherence to mask-wearing, distinguishing between correct usage, incorrect usage, or the absence of masks altogether.	Yes	https://github.com/ethancvaa/Properly-Wearing-Masked-Detect-Dataset
6.	Unconstrained Face Mask Dataset-UFMD [46]	* A large dataset that includes information about gender, age, nationality, and interior and exterior settings. *UFMD considers a significant number of head posture variations, which enhances the reliability of masked facial detectors. * Measuring social distance, face-to-hand interactions, and face mask recognition	Yes	https://github.com/iremeyiokur/COVID-19-Preventions-Control-System
7.	Face Mask Label Dataset-FMLD [47]	FMLD considers realistic conditions like head posture, lighting, and picture quality.	Yes	https://github.com/borutb-fri/FMLD
8.	Dey Dataset [48]	* Images from MFDD, as well as SMFD, are chosen. *The position of the head might be frontal or profile. *Since face regions comprise a substantial portion of an image, most scenarios are simple.	No	https://github.com/chandrikadeb7/Face-Mask-Detection

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

9.	Singh Dataset [49]	* The model is being trained to ascertain whether someone is wearing a mask or not. *. These pictures are from MAFA as well as Wider Face. * Analysis of the extent of crowing can be conducted using the detection findings.	Yes	https://drive.google.com/drive/folders/1pAxEBmfYLoVtZQlBT3doxmesAO7n3ES1?usp=sharing
10.	Wearing Mask Detection Dataset - WMD [27]	* Most of the photos in the database depict real-life scenarios from China's efforts to combat COVID-19, ensuring the inclusion of authentic situations. * Test dataset is split into three categories according to the complexity of the identification function and the percentage of masked individuals in each picture.	Yes	https://github.com/BingshuCV/WMD
11.	AIZOO -Tech [50]	*Data from Wider Face and MAFA *Medium-difficulty scenarios are represented in the images	Yes	https://github.com/AIZOOTech/FaceMaskDetection
12.	Kaggle [51]	* Faces sans masks, properly donning masks, and incorrectly wearing masks are the three categories in the dataset.	Yes	https://www.kaggle.com/andrewmvd/face-mask-detection

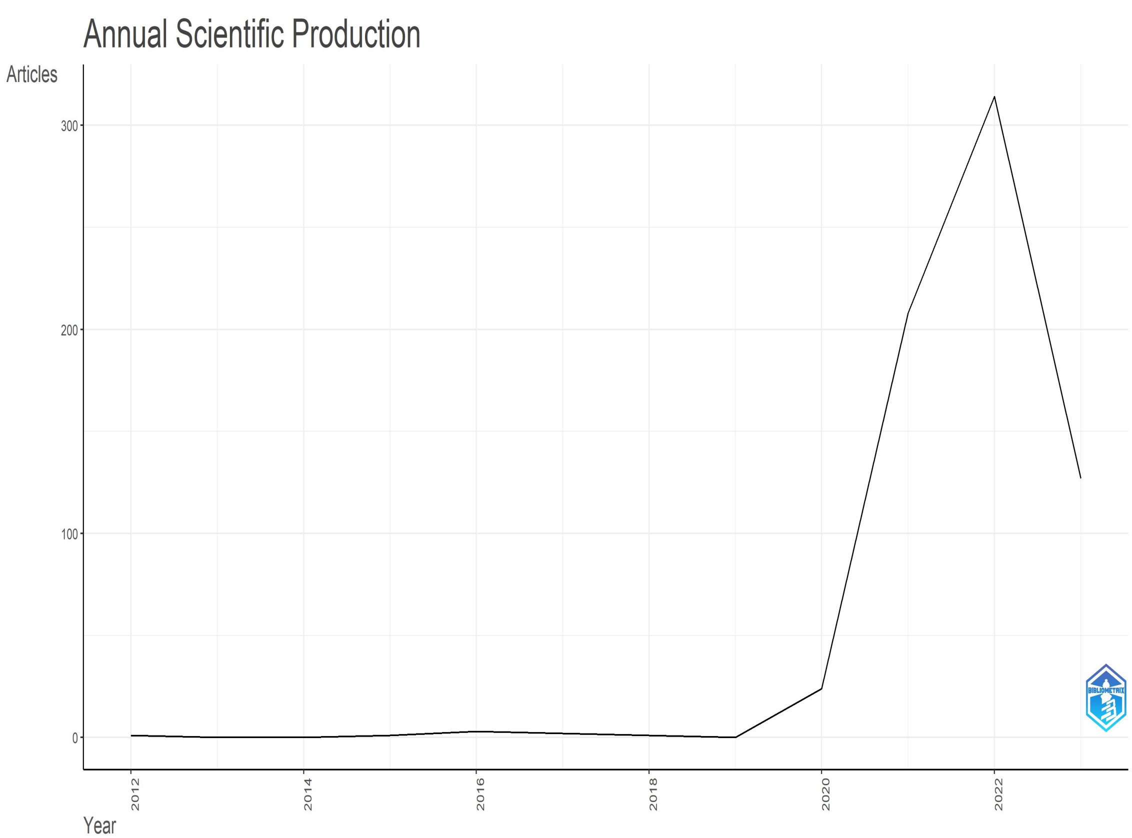

Figure 1: Annual Scientific production in the research area masked face detection from 2012 to 2022

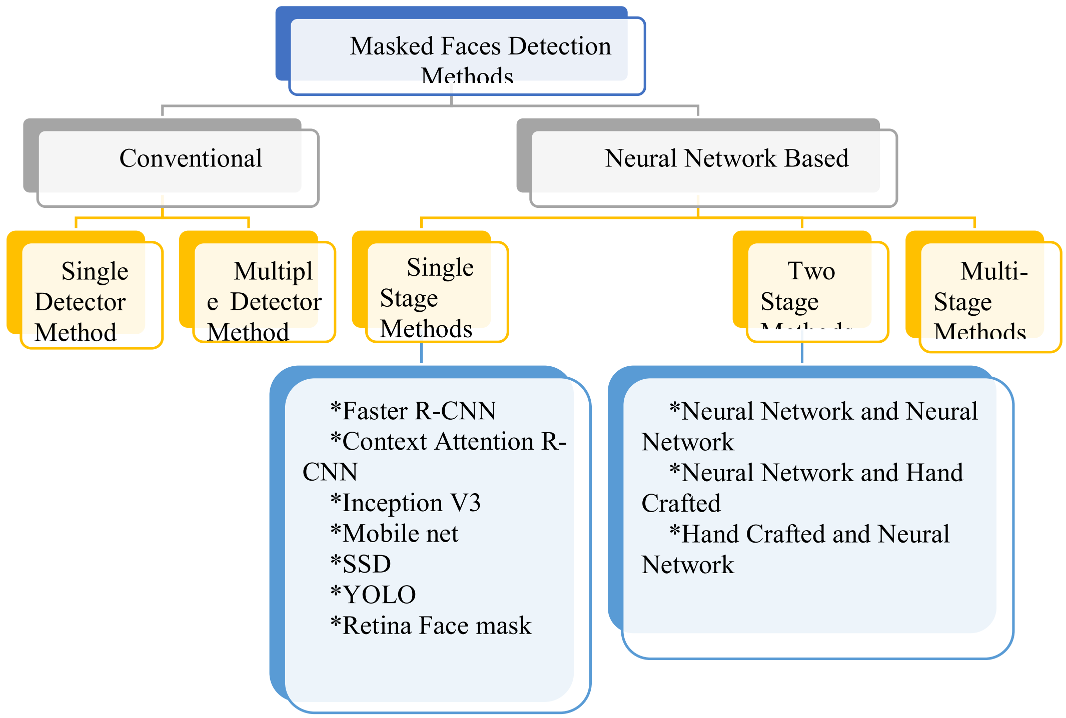

Figure 2: Methods used for detecting masked faces [52]

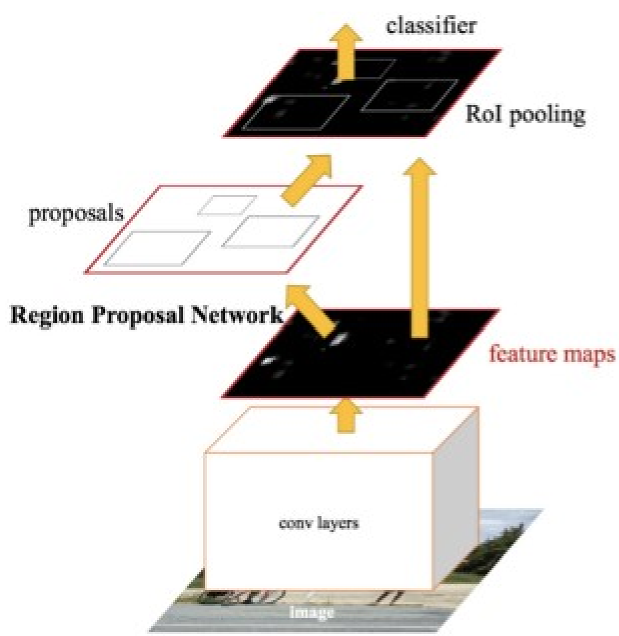

Figure 3: Faster R-CNN [59]