International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Recent Apporaches for Facemask Forgery Detection: Review

Shivani Chaudhry[1], Dr. Krishna Kumar Singh[2]

Submitted: 29/12/2023 **Revised**: 05/02/2024 **Accepted**: 13/02/2024

Abstract: Facemasks are a challenging task to detect. It has attracted increasing attention in this era. Consequently, there is an essential and challenging difficulty in identifying face masks as it is easier for a mask to recognise the face, but mask recognition is critical since removing the mask face is very complex. A multi-stage procedure is used in traditional object detection. Over the last two decades, there has been an increase in the interest in virtual face stimulation. Today, advanced technology acts as a forgery for processing digital photographs and computer graphics. Using digital images to falsify is one of the most significant technological issues. However, law enforcement specialists are developing robust algorithms to eliminate counterfeiting systems. With this field contributing huge impetus, the latest advances in profound learning enable several new applications like the CNT to assist and utilise computer vision technology (CNNs). Face recognition has been one of the most significant research subjects in computer vision and biometrics in the recent decade. All Algorithms generally depend on elements, such as low resolution, illumination, expressions, which deteriorate their precision.

Keywords: *Facemasks Detection, Object Detection, Algorithms, Technologies, Forgery*

1. Introduction

Computer vision is an interdisciplinary scientific field that involves analysing digital images or movies in advanced ways for computers. Traditional tasks include vision processing, classification of images, object identification and picture recognition [1]. Object detection can detect instances of specified image class visual objects, which is the correct solution to the given problem. Mask recognition has thus become an important task for the worldwide social vision of computers [2]. Detecting faces masks is a difficult task. Due to the growth of coronavirus disease, it has received more and more attention in this era. The detection of facial masks is a critical safety concern and prevention of Covid-19. In medicine, the mask decreases a person with or without symptoms the possible risk of exposure to the infection. Detection of face masks is utilised in airports, hospitals, schools [3].

The identification of facial masks is thus becoming a highly important and difficult problem. Face recognition with a mask is easier, but face recognition with a mask is crucial since it is highly complex to remove a masked face There are many facial features in the masked face, like the nose, lips and chin. In the medical field, masks lower a person's potential danger of exposure, whether or not they have symptoms [4]. The initial step is facial recognition have to detect the face from a picture.

There is primarily a difficulty, such as detecting several masks and uncovered faces in the image. It can be solved by a typical way of object detection [5]. Viola-Jones Algorithm, Adaptive Boost and HOG Gradient Histogram are utilised for traditional face identification techniques. This classifies the object detection process as multi-phase and single-short detectors (SSD). Faster RCNN is featured in multi-stage detectors, including the single-stage detector (SSD), including YOLO (You Only Look Once). Mask detection techniques such as video analytics, image segmentation, fingerprints, DWT, and LBP are employed in several applications: (Local Binary Pattern). All these strategies are examined to verify whether a person wears a mask and identify an individual's face recognition [6].

A system must be devised to ensure that everyone follows this critical safety guideline. You can be certain of it by using a face mask detector. Face mask detection is the process of determining whether or not someone is wearing a mask. To determine whether a mask was present, the first step was to identify the face, which divides the procedure into two halves. Face detection is an object detection application used in conjunction with security, biometrics, and law enforcement. There are numerous detectable tor systems in use throughout the world. This science must be enhanced because the world cannot afford further technological advancement [7The detection of face attacks has become an essential facial identification challenge since biometric technologies have been applied in several mobile and control systems. An effective 3D face mask PAD technique must be developed [7]

Facial modulation is performed using the most advanced facial detection technique. These data sets are above all current video manipulation data sets in at least

[1]*Research Scholar, Department of Computer Science and Engineering, Starex University, Binola (Gurugaon) Haryana,122413*
[2]*Associate professor, Symbiosis Centre for Information Technology, Pune*
Ch.shivani24@gmail.com1, Krishankumar@scit.edu2

one magnitude. We present demands for the classification and segmentation of compact movies in different qualitative levels in the previous forensic video assignment style using the new data set. We also give a benchmark assessment to produce undefined forgings with known ground reality, for example, generative refinement models and halfway processes for separating high-resolution movies [8]. The field of digital forensic image analysis concentrates on identifying picture forgery for contents that have been misrepresented. Many methods have been used to detect picture forges, and most of them either evaluate abnormalities related to a standard camera pipeline or rely on sophisticated picture modifications. The noise of the picture was a useful indicator for splicing detection, among other elements [10].

The rare face descriptor from conventional face trackers generally cannot capture complete information on the forms of the facial components that effectively meet the identification of higher levels of features such as the face, feelings. In order to alleviate the limits of the sparse facial descriptor, we provide the ideas of a face mask, a thick facial descriptor comprising information on semantic facial regions, including the eyes and the lips. Studies suggest a new strategy for removing deep semantic image segmentation masks from the video series. Unlike semantics, facial mask extraction handles occlusion, such as the direction of the face sign [11].

Facial detection frames, including R-CNN, Quick R-CNN and Faster R-CNN, are also used to enhance facial detection. These approaches mainly perform facial detections and position boxes that might be adverse and unreliable, such as recovered facial features, background noise and raw spatial quantisation. Subsequent face demands such as facial recognition, facial expressions and facial orientation are particularly effective for such inconveniences. Therefore, it is necessary to check a facial and segmentation system [12]. Mask R-CNN has an excellent performance in several benchmarks in object detection, such as COO challenges or city-scapes. R-CNN is an upgraded R-CNN object detection model. Mask R-CNN adds a branch of a mask to predict segmentation masks of interest area for detection and segmentation, compared with the regular R-CNN sequence. The research suggests using the improved R-CNN mask as a face detection- and segmentation strategy for facing and imaging tasks to deal with the inadequacies of existing approaches. This approach is available in German. In particular, our technique is utilised to boost the sensory accuracy of the facial recognition of boundary-box regression in the Generalised Crossing over the Union [13].

Forensic image studies recently took an impetus to investigate photographers' authenticity. In this respect, we believe that progress in deep learning represents a rare opportunity to develop incredibly powerful imaging using neural networks. Unfortunately, these techniques depend on vast amounts of training data, and many previous forensic data sets are manually prepared on a small scale as a consequence. This absence of training information is an important obstacle to identifying the manipulation of train networks and makes it difficult to evaluate different techniques [14].

2. Object Detection Algorithms

This section will explore a range of object detecting approaches and their performance. This section will also explore classifications and applications.

As part of an object identification technique, computer vision can recognize things in still photos or movie video frames. The machine can determine if the image is of everything. Application possibilities for identification are rapidly expanding. As a result, scientists have spent over two decades addressing this issue. Identification of anything is a combination of object location and image classification methods.

There are significant downsides to using conventional object identification methods to recognize many faces in an image, such as ignoring occlusions. As a result, deep learning-based object detection has been built. This strategy focuses on acquiring complex characteristics, the management of occlusions, and contextual information. Processing time is halved, and as a result, accuracy is increased. Facemask detection has recently been a well-researched topic of object detection, owing to the Covid-19 outbreak. As a result, object detection algorithms employ the same techniques. The mask entirely obscures facial characteristics (such as the nose, mouth, and chin), making it impossible to determine the existence of a person's face. The first stage in most facemask detection approaches is to identify the face; the second step is to extract the mask's features. Facemask detection is facilitated by CNN-based algorithms that are capable of detecting faces.

In object detection, there are two broad groups of parent algorithms. There are two sorts of object detection algorithms: those based on algorithms and those based on deep learning. Fig. 1 illustrates object detection approaches hierarchically.

Traditional object detection uses a multi-stage process. The Viola-Joins sensor, which can be detected in real-time, is a popular detector. An integrated picture method extracts hair functional descriptors, selects useful functions and detects artefacts using a cascade detector [15]. Although the integrated image is used to simplify the algorithm, computation still costs a lot. In Sethi et al. [16], and efficient feature extractor, known as HOG, is proposed to detect humans that calculate directed gradient directions and magnitudes over the image cells. Later, the deformable part-based [DPM] model can detect and connect pieces to evaluate object classes [17].

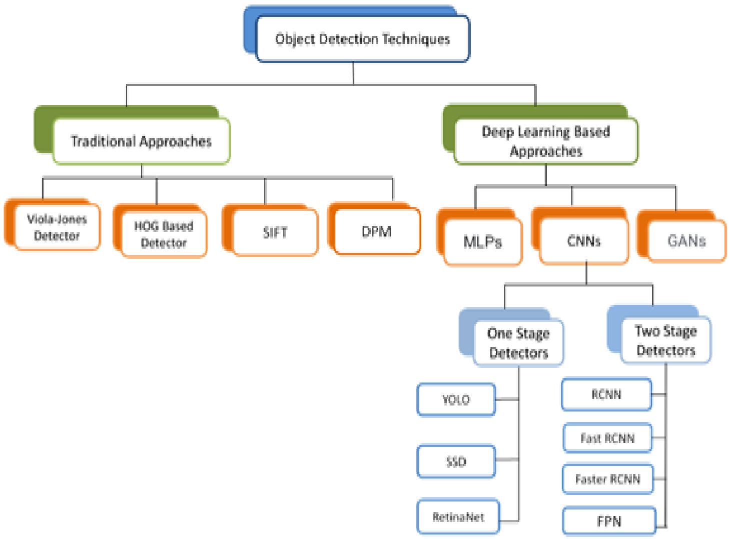


Fig 1 illustrates object detection approaches hierarchically

The deep learning sensor did not use hand-made features but recently demonstrated excellent efficiency because of its robustness and high ability to extract. Two common types, one-stage objects and two-stage objects, are available. A two-stage detector produces regional proposals in the first stage, followed by fine tones in the second stage [18]. The two-stage detector can achieve high but low-performance detection. R proposes R-CNN's seminal work. R-CNN selects candidate areas containing items using Selective Search [18]. Subsequently, the proposals are put in a CNN model for functions extraction and a [SVM] vector support machine is used for object type recognition. The second step of R-CNN is calculation-intensive since the network has to detect proposals one by one and uses a different SVM in the final classification [19].

Many of the research and activities in the past (before 2000) before Viola and Jones suggested a challenge were unsatisfactory. Kurien et al. [11] are the first to use rectangular facial boxes. However, it has many drawbacks since it has a wide range of features. There are 160,000 hair characteristics of a 24-digit photo, and even wild faces and front faces are not covered. The problem was observed, and people worked hard to incorporate complex features (HOG, SIFT, SURF and ACF). In order to differentiate the power of both pixels, the new NPD feature was introduced. Another well-known solution is Luo et al. [20], which categorised vector help as facial recognition. Another topic that is commonly studied is the increase in detection robustness. One easy technique is integrating several detections to be trained independently in distinct perspectives. Many deformable models are used to take faces from different perspectives. A basic model of

separate mixed learning is Hyun et al. [21]. These models must be planned and tested, and their performance is reduced over lengthy durations. In 2002 a neural network was used to identify semi-frontal human sides into complex images. In recent years, there have been many facial and face recognition. The best approach to identify a person is by making no human effort to recognise faces as numerous facial identification and facial detection techniques have been developed [7,38,56,58].

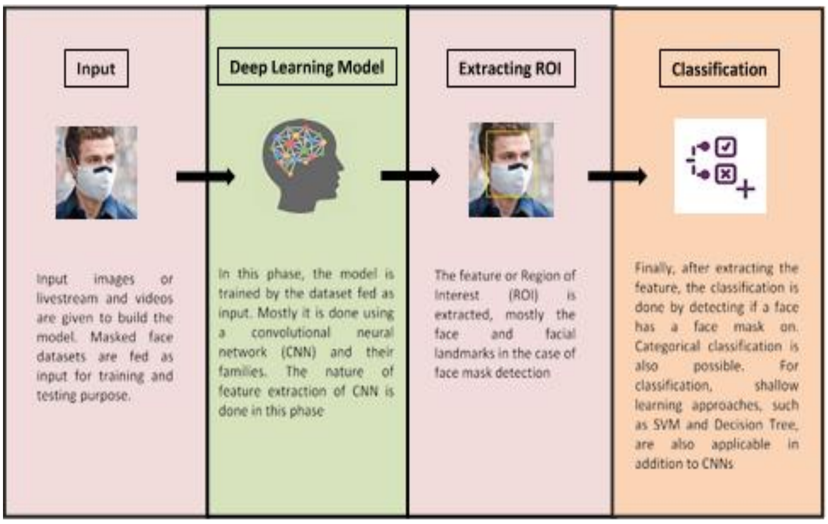


Fig 2 General infrastructure for how facemask detection algorithms work for almost all the techniques stated in Section III

2.1. Existing algorithms for facemask detection

For facemask forgery detection, Sapin and Kikan [22] developed a new anchor attention approach, which would improve the characteristics of face regions and increase their accuracy, combining the allocation of anchor point's strategy with data expansion. However, Ratta et al. [23] solution did not address the detection of masks. For real-time detection, however, the number of parameters was compressed, resulting in data set accuracy at a rate of just 89.6 percent. Chandani and Arora [24] created a WIDER Face-MAFA-SMD dataset using SSD. Chandani and Arora [24] proposed an FPN network, along with a Content Payment Mechanism introduced by RetinaMask, who picked ResNet or MobileNet as the backbone tool to use both cheap and high hardware. Heibbar and Kunte [25] used the model and hybrid learning methods to transfer and classify features. Absolute precision in Real-Time Face Dataset (RMFD) was 99.64%, the masked face data set was 99.49%and the labels in the wild (LFW) datasets were 100%.

There are two main methods of extraction, typically local and holistic. In a holistic approach based on features like Eigen, the features come from the whole face, which can also be affected by occlusion and speech changes [26].

ALGORITHMS	DESCRIPTION
SIFT	• DOG concatenated with images of different sizes. • Local extrema detection- Non-maxima suppression- Hessian matrix elimination of edge responses. • Images are sampled for gradient magnitudes and orientations around the keypoint location, and the key point's scale is used to select the degree of Gaussian blur.
SURF	• Convoluted box filter of different sizes and integral image. • The Hessian matrix and non-maximum suppression are used to identify potential key points in the data. • In a circular neighbourhood around the interesting point, a sliding orientation window of size pi/3 detects the dominant orientation of the Gaussian weighted Haar wavelet responses at each sample point.
PCA	• PCA is a technique for reducing the dimensionality • PCA aids in the classification of features. • The original data sets take on a new shape and location in a different space. • PCA calculates the best discriminating components without knowing the groups involved in the calculations.
LDA	• LDA is a more specialised form of generative analysis. • LDA allows for data classification to be performed. • There is no change in location, as LDA only attempts to separate classes and draw a decision region between them.

	• When a client defines groups, LDA calculates the best discriminating components.
LBP	• The LBP operator is used to specify a texture. • A square pixel matrix is used to generate the labels. • After thresholding, the produced labels are binary numbers. • The label histogram is used to describe the texture.
SVM	• SVM is a well-known binary classifier used extensively in research and engineering. • SVM performance is determined by the kernel function's Riemannian geometric structure. • The objective is to improve class separation by enhancing spatial resolution at the boundary.
MPCA	• Multilinear Principal Component Analysis (MPCA) can be used to learn about several factor interactions such as diverse viewpoints, lighting conditions, and expressions. • MPCA employs N transformation vectors to capture the face images' dimensionality accurately.
ANN	• A Neural Network (ANN) is a data processing paradigm inspired by organic nerve systems such as the brain. • It comprises many interconnected processing units (neurons) that solve issues. Lessons are acquired through observation. • An artificial neural network (ANN) recognises patterns or classifies data for a particular application. • Adjusting synaptic connections between neurons is a component of learning in biological systems.
FFNNs	• Signals in a feed-forward configuration ANNs can only go in one direction from input to output. • This indicates that one layer's output does not affect the output. • Pattern recognition is frequently utilised.
MLP	• The MLP feed-forward network is composed of many layers of neurons. • The MLP network comprises an input layer composed of sensory units (source nodes), a hidden layer composed of calculation nodes, and an output layer composed of computation nodes. • The input signal propagates left to right across the network, layer by layer. The objective is to train the network to respond appropriately to training input patterns while also responding appropriately to similar input patterns.
BP	• Back-propagation (BP) neural networks are multi-layer feed-forward networks trained using the error back-propagation algorithm. • The BP network can learn and store many input-output mapping relations without knowing the underlying mathematical equation. • Its learning rule is to use back-propagation to update the network's weight and threshold values until the error sum of squares is as reduced as possible.
CNN	• Localisation and normalisation of the face are critical in this difficult task that receives less attention in the literature. • The proposed CNN method locates and normalises human faces successfully in a fraction of the time required by previous approaches. • The proposed CNN technique locates eyes in any grayscale image by looking for eye and eye socket properties. • The technique works rapidly and reliably on a typical database. It makes a few percent of errors.

While those problems are solved in a local method, since only images' patches are considered, they are also invariant in scale and rotation. Several local extractors are effective, but they do not operate in certain environments, such as Gabor, LBP, SIFT, SURF, SIFT, for instance, is invariant in scale and rotation, since it is based on local features but has various lighting conditions [27]. Each method for extracting functions often fails to face more than one challenge; hybrid function extractor is therefore important. In the proposed work, the SIFT is being introduced and modified with retinal modelling of fixed labels and light adjustment filters to increase recognition rates when analysed in real-time video data sets and standard sets of data such as celebrity YouTube [28,29,30].

However, the fact that the CNN architectures proposed include several million parameters does not make them ideal for use in low-power computers. Methods that rely on deformable component models, template comparison models, and 3D face models cannot operate with high definition video streams in real-time, not even in the quest on the front [31,32]. Detectors that use manuallydesigned features to define artefacts and classification cascades are still the best processing speed solution. Many different descriptors have been suggested to define facial characteristics. The most popular are rectangular hair-like characteristics, which are successful for building façade detectors and highly extracted with the integral image [33]. Textural features MCT and LBP, pixel intensity coded in the local domains, are monotonic light-changing invariants. Combined with HOG characteristics, LBP has shown strong generalising properties and can process nuanced, non-facial images in better terms than hair-like characteristics [34]

Quick R-CNN fixes this problem by adding the pooling layer area of interest [ROI] to input simultaneously in all proposed regions. Finally, for unreasonable searches that restrict the speed of these detectors, a regional proposal network [RPN] is proposed for faster R-CNN [33]. Faster R-CNN integrates all individual detector components into a neural network architecture end-to-end, including a regional proposal, extractor and detector feature. A single-stage detector detects only objects with a single neural network [23]. Certain anchor boxes that specify the width-height ratio of objects should be predefined to do this. The one-stage detector slightly scars the performance rather than the two-stage detector, significantly increasing detection speed [35]. YOLO split the image into several cells to meet the target and tried to fit the anchor boxes for objects per cell, but the method was unsuccessful for small objects. Researchers conclude that only the last feature output is a one-stage detector which is not fine since the last feature map has fixed receiving fields, which only certain areas can be displayed in the initial images [26]. Therefore, multi-scale detection in the SSD has been implemented, which detects faces in various sizes on several maps. Later, to improve detection performance, Katherineguo et al. [36] propose an architecture with hybrid SSD and FPN that includes a new focus loss feature to mitigate class imbalances.

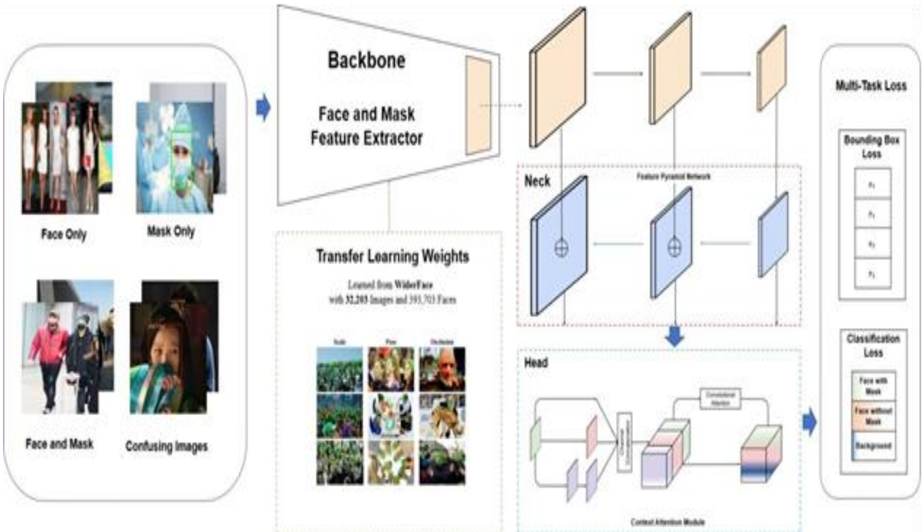


Fig 3 Retina Face Mask Architecture

2.2. Forgery Classification Task

Forgery in images can be termed as the any kind of conflicted or manipulation in images structure[38]. In respect to this the counterfeit functional classification with binary test issues. This research is carried according to these aspects but targeting the image forgery on front end basis, instead of this the research is carried in broad domain by using the advanced machine learning approach. The database thus selected is having the multiple frame for more than 700 dynamic images or simply saying videos[41].. Considering the approach of face2face different face are given to 150 frams. Different functioning fram on 128 x 128 pixel the contrast to the face pixel should be adjusted and quantified for compression [39]

2.3. Video Quality for Facemask Detection

The assessment of surveillance video quality (SVQA) is a subset of video quality assessment (VQA). Traditional video quality assurance approaches place a premium on the user's holistic impression of entertainment videos [40]. Typically, surveillance footage is used for recognition tasks such as pedestrian detection. Standard VQA methods such as SSIM may not yield accurate results when assessing video quality. Rather than entertainment value, the ideal SVQA technique should place a premium on the videos' recognition value [42].

Quality of recognition (QoR) is a research topic that examines the recognition quality of the video [43]. The Video Quality Expert Group (VQEG) recently announced the QART project, which would "study the quality of video used for task-based multimedia applications and recognition tasks." Several studies have been undertaken on car license plate recognition (LPR) to ascertain the relationship between recognition rate and video quality [44]. Typically, they evaluate only video compression and resolution in their tests. Apart from the target's size and the light's intensity, additional factors should be considered to ensure the method is reliable [45].

The et al. [46] investigated the objective algorithms are imperfect and cannot match human performance on benchmarks, reducing the precision of quality rating results. DFR is distinct from FR because it examines a distorted face's ability to be recognised compared to an undistorted face. The DFR challenge was chosen for SVQA because it directly compares the loss of semantic information in distorted surveillance film to reference videos. Its primary objective is to construct an objective DFR model. First, we generate a collection of faces from surveillance cameras [47].

PSNR and SSIM are the most often used SVQA approaches [48]. These techniques are incompatible with video surveillance that is task-based. In 2010, the Working Group on Video Quality in Public Safety (VQiPS) developed a guide for public safety [49]. Loy [50] used the VQiPS framework to outline definitions, research trials, and video quality evaluation trends. Numerous articles have utilised the LPR assignment as an illustration of SVQA. According to Loy [50], the pace of human recognition and the bit rate of video are interdependent. Loy [50] reported that human recognition outperformed ALPR software due to the software's lower recognition rate. Loy [50] investigated the relationship between ALPR recognition rates and bit rates/compression ratios using a logarithmic/logistic model. Before moving on to LPR, they initially assessed surveillance video quality using face pictures as the semantic carrier. They invented the concept of "critical video quality" to reduce bit rate without impairing accurate face detection, recognition, and tracking algorithms [51].

Facial quality assessment is a study topic comparable to SVQA based on face recognition. Compound features such as brightness are retrieved first and then blended using weights. Raut [51] have proposed a method for comparing a face image to a probabilistic face model based on patches.

The typical end beneficiaries of video and images are human beings. By using video analytical algorithms to automate their tasks. Video tracking, separate vehicles and small mobile applications are included in these systems [52].. A cell phone uses amateur cameras to tag a picture automatically or mark a current position based on an analysis of an area. The video can be analysed in two ways. The video will first be processed on the sensor, and the results can be sent to the server. The second is to send the video streams to the server and process the video on the server [53].

An analysis of a video sensor demands restricted bandwidth and computer complex sensor modules, increasing expenses and complicated equipment maintenance. Video server analysis means that sensor devices are less expensive and energy-efficient, but video streaming demands greater network capacity [30]. The cost of video sensors and the energy quality and bandwidth of the network is also agreed on.

The ultimate beneficiaries of videos are not human. Typical digital camera surveillance provides a maximum of 320-240, high SNR and 10-30 fps [54]. Usually, this level is the least appropriate standard for the human visual system. Although it illustrates how people have less quality work for video recognition, conventional video surveillance is recorded to enhance visual surveillance at the maximum cost [26].

2.4. Video based Face Recognition

As first suggested in the 1880s, face identification was one of the best domains for pattern recognition. For face photos like Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Elastische Graph Matching, the majority of the techniques are promising (EGM). More details than still photos can be provided, for example, space-time details. Video recognition has also recently become increasingly popular [4,9,55].

2.5. Face Detection and Recognition

The first stage of a facial recognition system is facial detection. Much research in this field has been conducted, which is useful and usable only for still images. Video loops may not be extended directly. Human faces have endless directions and places in the video sequences. Thus their identification offers academics various obstacles. [23,56–57]. Generally, there are three major video-based facial recognition methods.

However, the main downside to this method is to disregard the temporal details given by the video sequence. Second, the combination of detection and monitoring entails first frame facial detection and then sequence tracking. The information loss is conceivable because of the separation between the detection and tracking usage of information at once from one source [57]. Finally, a temporal technique employs a temporal connection between structures to detect several human faces in a video series rather than detecting each frame. This process typically consists of two parts, namely update detection, estimating and monitoring. It helps manage the detection and is less exposed to thresholds than the other two forms of detection. Lin and Cheng [15] present an iterative dynamic programming (DP) fast template matching method for detecting the front faces and monitoring the non-frontal faces using online adapted facing models. In the meantime, it was noticed that there were higher concentrations at the edge near facial characteristics but less concentration at the edge when slightly beyond facial characteristics appeared.

Rani and Sharma [27], Nazli and Maghani [26], Aljaberi [43] suggested a system for detecting and recording facial features in video sequences called edge counting pixels. Previous studies have categorised and monitored several unknown and varied objects using a graphical framework that maintains several hypotheses. Furthermore, automatic appearance models were produced based on acceptable video segment clusters. Some methods also combine border orientation characteristics to increase detection efficiency. Dayani et al. [58] proposed a facial detection process using the local histogram of wavelet coefficients represented about the object-fixed coordinate frame to exploit the temporal video information fully.

Moreover, to make a local decision, m and Khurshid [59] suggested Float boost dependent face detection and then used temporal information to confirm and validate the results. In video-based facial detection, real-time and multi-view facial detection is essential. There are two main methods:(A) The method consists of building a single sensor for all facet views;(B)the method concentrates on multiple detector responses for different views [60].

Changes based on linear PCA facial trajectories offered a useful method to investigate those changes. The design of the detector-pyramid was also seen in Kumar et al. [2], which followed an integrated strategy of ground-to-fine view decomposition and simple face-to-face or non-facial classification. Bhole [14] suggested an algorithm to achieve the lowest error rate by combining Cascade AdaBoost and the detector array. However, there is a significant problem in most of these methods since the variability of the multi-visual data set in the class is greater than that of the front-visual data set. Although Detector-Pyramid Architecture AdaBoost (DPAA) can solve these problems, the increased complexity leads to high mechanical loads and excess fitness in training. Kaur and Kaur [61] discussed over-fitting, but robust methods are still required.

The most crucial step of the entire process is face recognition. Some of the video algorithms are still tech. Videos can, however, convey more detail than images. The use of film offers four key advantages: Firstly, using video sequence continuity to improve still image effectiveness [53]. Secondly, new psychophysical and neurological research suggests that detailed information is crucial in recognising the human face. Thirdly, better pictures can be taken from the video stream, including a three-dimensional model or super-resolution images, to improve recognition effects [18]. Fourthly, apart from the above-listed reasons, video-based recognition enables the model of the subject to be learned or modified over time. Poor video quality, poor picture resolution, other factors, e.g. lighting, changes of poses, motion, occlusion, decoration, voice, broad camera distance. Due to these benefits and disadvantages, multiple aspects of video recognition are still present [25,41,62].

Most current methods use video face recognition Spatio-temporal information; some also use provisional voting to raise identification rates. Some algorithms also remove 2D or 3D facial structures. In addition to voting, Katherineguo et al. [36] proposed an approach based on the core features' kind, textures, and extraction. The knowledge of coherence in such a method is not fully included. Daniya et al. [58] proposed adding temporal information about human recognition in a video sequence. A state-space model was employed to monitor state vector and identity recognition variables.

Hyun et al. [21] obtained low-level statistical video models (e.g. PCA) to align the frame with a video stream or two video streams. Khan [53] examined two video sequences by selecting the pair of pictures closest to the same images. Some methods are used to form a statistical model face with video sequences. The video structure for each entity has been used for the mutual subspace approach to individually calculate distinct Eigen spaces, taking into account the angle between the input subspace and the reference subspaces formed by the primary image components. The main kernel angles were used to calculate the similarity between two video sequences in the original image and feature space. To enhance this, the author proposed a simple algorithm based on the characteristics and locations of the faces in order to select the frames and then used dimensional tests to transform them into new spaces [26]. The proposed system achieved better performance in online facial recognition through sparse video clips in a non-regulated environment. The latest classification algorithm (PCNSA) aims to resolve the problem of unequal, non-white covariance noise matrices between different groups. A moveable face has been recently modelled and observed in the Auto-Regressive and Moving Average (ARMA) model as a linear dynamic system. The commonly used Hidden Markov (HMM) models have also been used [1,28,58,75].

2.6. Convolutional Neural Network for Facemask Detection

There is a need to recognise individuals on millions of images submitted to social services every day, which has led to major steps towards addressing the issue of face detection. New techniques are invariant in poses and facial expressions [63]. In addition, they can also work in complex lighting and extreme occlusion. However, many algorithms that show excellent performance on facial recognition are highly computationally complex. This prohibits their use for video analysis. Although the advancement of modern facial detection methods leads to a rise in invariance in the head position and occlusion of the face, we only consider that frontal facial detection is a specific problem [64].

At the same time, the front location of the person with the camera is normal for many video analytics systems. That is why these detectors are so common in practical usage. Due to the broad interclass variance, the various ambient light conditions, and the complex structure of the background, a simple and fast detector with high precision and reaction to all possible face-images variations is not possible [60]. The traditional approach is to use different models for each head position. It has recently been shown that deep CNNs can research a whole range of two-dimensional projections within the boundaries of a single model thanks to the high capacity for generalisation [64].

The fact that several million parameters are found in the proposed CNN architectures renders it unable to be used in low power computing devices. Methods based on model-based component models, template-based models and 3D-face models cannot function in real-time with high-definition video streams to solve the problems of front search alone. Detectors that use manually built features to define objects and improve classifiers to classify them remain the best solution for processing speed [2].

2.7. Face Detection Ensemble

One of the most important and difficult things about computer vision and how humans and computers work together is face recognition, which helps people and computers figure out who is on each side of an image or video. It can be used for many things involving a human face, like alignment, face-identification or authentication, face-tracking and marking [59]. Face recognition is very important, and one needs to figure out who Face detection is hard because it cannot be assumed that all faces in an image are in the same place as the face location. People's faces can look very different based on their sex, age, face expressions, race, and lighting. They can also change a lot depending on their pose, lighting, orientation (in-plane rotation), size, level of occlusion, and background complexity. If there is a need to use these kinds of apps, one needs to quickly solve a problem with a good and stable facial detection method [65].

Diwan et al. [66] three key techniques are used in the well-known Viola-Jones (VJ) algorithm: an integrated picture strategy for hair extraction, a better algorithm for collecting weak classifications and careful waterfalls that quickly show negatives [67]. There are, however, several important limitations to VJ algorithms, such as the suboptimal cascades, the large pool size of hair-like features that make the training very easy, and the limited ability to represent hair characteristics like postal variations and light changes. Face Detection DataSet and Benchmark are two examples of software that does not work well in places that are not very nice (FDDB). Early hair extensions and improvements were made to help with many of these problems, such as hair-like rotated functions, sparse characteristics, and polygonal shapes [57].

Cao et al. [68] use old selection and filtering techniques that help speed up work and make a cascade paradigm boost algorithm even stronger. As a result, light invariance and speed were better when used with the updated census transform (MCT). A method is being proposed for analysing faces with arbitrary in-plane rotation and off-plane rotational angles in still and video images from before 2010. When Intel labs came up with a two-pronged strategy to speed up the SURF, they used multidimensional hair instead of single dimension hair [69]. This was a great 2D method that came out in the last decade.

Cao et al. [68] also have two simple methods that work better than commercial face detectors like Google Picasa. One is about rigid models, which are very similar to the VJ algorithm in structure. The other is an integral deformable partial model (DPM), a common way to detect objects. It combines the assessment of latent variables for alignment and clustering with multiple elements and decapitating components to manage internal variances.

Scholars use a system to model a face shape with a set of parameters so they can keep an eye on a face that looks distorted [1,10,19,38,75]. The facial monument algorithm has been used with the VJ method that has been changed. Also, Shah and Raskar [69] added to the original VJ tree classification by making it a deeper, two-leaf, quadratic structure. It is also important to use deep learning to deal with the difficulties of 2D face recognition. According to Cao et al. [68], they were the first to use a CNN and object detection zones. Its model, called the Region-CNN (R-CNN), has three parts. Some

2000 independent regional proposals are made in this stage. R-CNN then extracts the fixed-length deep vector function from each of these proposals. Instead, Cao et al. [68] deep-thick face detector (DDFD) does not need any posing or writing annotations. It can detect faces in a single, profound learning model with several rules. Scholars came up with a way to learn to get rid of small faces through a deep neural network. It would also be great if cheap profile cameras could be made so that they could get more information about a person's face than a monocular 2D approach could [73]. Using the uniform light principle, Kinect projects a pattern into a scene to measure the depth of each object and then uses a time-of-vol principle to figure out how far away an object moves when a light signal bounces back from it [65].

2.8. Forgery Detection Methods

Today, advanced technology serves as the go-to equipment forgeries to process digital images and computer graphics. In reality, the exploitation of digital pictures to produce falsifications is one of the largest technological problems [70]. However, law enforcement specialists develop systems that use powerful algorithms to remove the falsifications. What could surprise individuals who do not work in the industry is that very few digital papers (particularly those produced from health, law and government sources) nowadays do not contain any kind of falsification. It is feasible to detect forgery algorithms but almost completely depends on the image source [26].

Active and passive techniques were devised to prevent these falsifications. The active strategy is to make documents concisely and genuinely by using digital watermarking or signatures. All these approaches serve the same goals and make professional authentication work more visible and therefore help eliminate fraudulent works with succinct documentation and resilience in image processing [71]. The key foundations for detecting forgery detectors are fundamentally identical – the image details included or added. For example, a colour filter improved the image, the acquisition phase employed, or the camera lens's properties. The photo-response non-uniformity noise sensor can detect all this information professionally (PRNU). PRNU is a strong copy-move forgery detection algorithm unique to each camera [67]. Copy-move forgery can also be seen in other sophisticated algorithms. Those algorithms have three major processes: function extraction, correspondence and pixel post-processing to eliminate false alarms.

Furthermore, selecting invariant scale and rotation features is critical for emptying resilience. The Patch Match method is the greatest example of the high efficiency for forgery detection. The matching algorithm defaults more quickly than most other algorithms and employs a matching method that applies a dense matching field. The offset field of Patch Matching will make copy movement forgeries more efficient and flat [32].

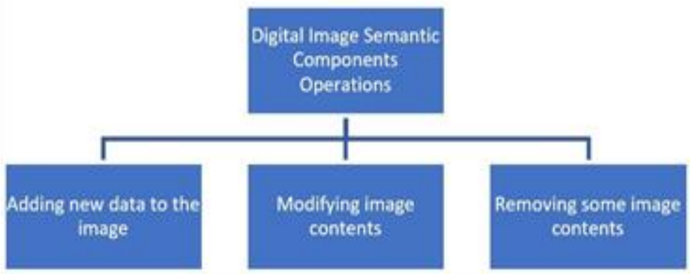


Fig 4 Various Types of Forgery

With this discipline adding enormous impetus, recent breakthroughs in deep learning offer numerous new applications, such as Convolutional Neural Networks, to support and use computer vision technology (CNNs) [72]. CNN's are mostly employed in visualisation applications and test raw input images with locally situated neighbourhood pooling and trainable filters.

Formerly, we offer new models for detecting and locating forgings via copy-move using these new methods [27]. Digital forgery is, in most cases, a simple improvement of the image but also can be completely manipulated. Retouching photographs entails restoring damaged images, whereas improving image is a transformation process [73]. This process has several beneficial effects, such as restoring missing data. However, it may also be prohibited in some situations, particularly official documents, academic papers, or medical material. The splicing of images consists of a copy and paste of more than one image in the image, whereas image cloning is another term for typical forgery for copy-move forge detecting kinds of digital images [73]. Developing a viable digital picture forgery detection system has been an active research topic in recent years. Many strategies for picture forgery have been used. On the other hand, processes for picture falsification detection might be divided into two main method categories: active and passive. Digital fingerprint, digital signature and digital watermarking [58] can be given as an active approach.

2.9. Algorithms for Facemask Recognition

Face recognition was one of the topics with the highest research in computer vision and biometrics over the last decade. Traditional techniques focused on hand-made characteristics, and The earliest stages to build the techniques of deeper neural networks are considered the backbone of all computer vision applications were traditional machine learning techniques like feature-based and geometry-based methods. This section explains five traditional algorithms for machine learning extensively employed for facial recognition [74].

Diwan et al. [65] suggest that the following can be formulated for difficulties with recognition 'It is clear from this point of view that if you have still or video photographs of a scene, identify or verify one or more people on the scene using the stored facial data base' (occlusions, poses, illumination, ageing, conditions, and facial expressions). For the past decade, academics have been focusing on using image processing to define the geometry of the face to match the exact faces of an image [75]. All algorithms are, in general, influenced by factors that degrade their accuracy, such as low resolution, lighting, expressions. Still, facial motion recognition is an enormous problem; many machine learning specialists have already achieved 100% accuracy in frontal facial images, and the accuracy is not trustworthy because of the previously described reasons [74].

3. Datasets

This section provides a detailed review of several of the datasets used in the facemask detection techniques presented in Section III. Additionally, this section discusses the datasets' unique properties, compositions, volumes, and limits.

A high number of training sets is required to apply deep learning algorithms successfully. Substantial datasets were required to perform mask detection tasks using deep learning-based algorithms after the Covid-19 burst outbreak. There are several face datasets available. However, the masked datasets are somewhat small in contrast.

The number of datasets, including masked faces, increased dramatically after the Covid-19 premium on mask detection algorithms. To summarize the information offered in this section, we have included a list of datasets in Table 3, along with an explanation of their characteristics, sources, and limitations. As a result, assessing the dataset's strengths and drawbacks is straightforward. The most frequently used datasets for facemask identification are depicted in Figure 12, which have been regularly used in the facemask detection technique. Due to the scarcity of datasets, most scientists constructed and trained models using numerous datasets rather than a single one. A significant reason is that particular datasets may result in erroneous judgments regarding skin tone, texture, and facial orientation. Occasionally, writers develop their unique datasets for training and testing their algorithms. Table 3 summarizes the datasets they used in their current facemask identification techniques, as well as the model performance

By utilizing many datasets rather than a single one for facemask detection, one can increase the diversity and length of the datasets while avoiding bias. Numerous efforts have been made to integrate diverse datasets and even create images independently, as illustrated in Table 4. Additionally, the table contains explanations of several datasets. Table 4 summarizes the data utilized in facemask detection techniques and their performance indicators to compare.

4. Performance Analysis

The next section is divided into two sub-sections. The first section discusses the important performance measures widely discussed in the literature. The next section analyzes the algorithms' performance, shortcomings, and future improvements in detail

MAJOR PERFORMANCE METRICS USED IN FACEMASKDETECTION

Several authors employed a range of performance measures to evaluate their models' performance. Although most research evaluated their performance using a variety of indicators, a small number also used a single statistic. This section distinguishes the two sub-sections.

Table 3: Comparisons of existing different methods of facemask recognition

Algorithms	Pros	Cons	References
Facemask recognition algorithms based on classical	Focuses on the manifold's local structure. The face is projected onto a linear subspace spanned by the eigenface images using these methods described below. It is easy to convert Mahalanobis distances with	Facial expressions and other factors, such as large variations in illumination, can make these methods ineffective. Compared to linear methods, kernel-based nonlinear methods do not significantly improve. When describing neighbouring changes	2,5,24,32

algorithms	probabilistic interpretation from the distance from face space to the plane of the mean image.	in face description, LLE, LLP, and LBP introduced a simple and effective way to describe them in terms of their proximity. DCV- and SVM-based methods used subspace-based approaches as part of their design. In order to preserve the local structure between samples, NPP and ONPP methods must be used. Uncertainty persists regarding neighbourhood size and optimal values.	
Artificial neural networks (ANN)	Unnaturally, artificial neural networks integrate non-negative matrix factorisation and radial basis functions. Process simplification can also be achieved by using ANNs' native linearisation feature and increasing computation speed. Ideal for recognising faces with partial distortion and occlusion, for example.	This approach has the disadvantage of requiring more training samples (instead of one or a limited number). Similar to other statistically based methods, it is inaccurate.	2,11,24, 48
Gabor wavelets	Using Gabor wavelets to capture salient visual properties like spatial localisation, orientation selectivity, and spatial frequency, the Gabor wavelets exhibit desirable properties. These methods are used in various types of biometric applications.	Face images are convolved with Gabor filters, which results in a high-dimensional Gabor feature space. For real-time applications, this approach is computationally intensive. Gabor features that have been simplified are also sensitive to variations in lightning.	1,25
The use of face descriptors	Develop image descriptors with the goal of learning. Image differences between images of a single person are minimised, while those between images of different people are maximised. Illumination and expression changes do not affect these methods' discrimination or robustness. Each of these descriptors is compact, easy to extract and highly discriminative.	The approach is computationally intensive during the descriptor extraction stage, but it is simple and performant for online applications.	6,48
Facemask recognition based on 3D	The traditional 2D capture process is extended, and accuracy is improved. In order to make the solution more robust, the depth information does not depend on pose or lighting conditions.	Be sure that all 3D face recognition system elements are calibrated and synchronised with existing 3D data before using them. Costly to compute and unsuitable for practical use.	4,11,21, 56
Recognition using video footage	The main advantage of this approach is that it allows video redundancy to enhance still image systems.	A relatively incomplete investigation. Problems with measuring similarity between two (or more) images are multiplied.	2,23,35

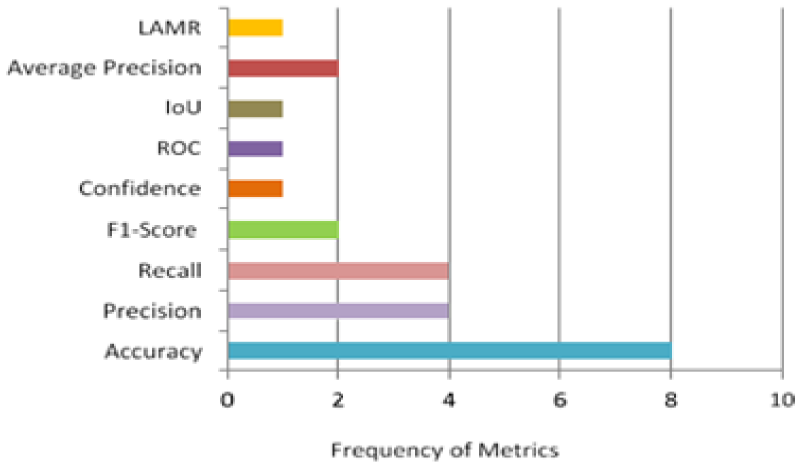
Table 3: summarizes the most frequently used datasets for facemask detection

Dataset	Composition	Characteristic	Limitations	References.
MFDD	Solely masked face images	Public dataset	Biased to chinese face and not so sufficient in number	76
RMFRD	Masked and unmasked face images of same subject	Effective in accueacy since the dataset is very large	Biased toward asian face images	77
MAFA	Contains masked face and	Categorical classification	Mostly preferable for occlusion	78

	any sort of occlusion on face	is easily deployable since mask type is specified	detection rather than physical mask detection	
SMFRD	Masked and unmasked face image of same person	Diverse and versatile	Training phase takes more memory and time	76
Masked Face-Net	Contains improperly worn masked face data along with masked faces	Benchmark dataset, categorical classification is easier than with other dataset	Biased toward surgical masks	79
LFW	Composed of celebrity images of different orientation	Benchmark dataset for face recognition	Does not contain any masked face image	80
FMDD	Contains only masked face images	Publicly available	Very limited in number	81

Table 4: Datasets and results from the available research on facemask detection

Dataset	Description	Performance in mentioned literature	References
MAFA	Dataset introduced by author himself	Average precision 74.6%	78
1.WiderFace (pretrained) 2. MASKED FACES	Trained with widerface and fine tuned with the latter	Accuracy is 86.6% and recall is 87.8%	82
MMD	Medical mask dataset	Accuracy is 98.70%	83
Customized	Dataset introduced by author himself	Accuracy is 98.65	84
1. RMFRD 2. FMDD 3. FFHQ	Combined three dataset to avoid bias and scarcity	Precision, recall , and FI-score for the facemask classifier of 98.28%, 100%, and 99.13%, respectively	85
Customized dataset	Author clamied to use customized dataset	Precision of 91.7% with a confidence score of 0.7	86
FMD	Facemask dataset combined by WiderFace and MAFA	-	87
Not mentioned		Accuracy is 96%	88
Customized	Customized images of factory worker	Accuracy is 97.6%	89
Not mentioned		Accuracy is 99.8%	90
1.FMDD 2. MMD	Combination of facemask dataset and medical mask dataset	Average Precision is 81%	91
Mask Augsburg Speech Corpus (MASC)	Voice recordings of 32 german speakers. Each samples contain one second, and sampling rate is 16khz	Accuracy is 74.6%	92
1. BAO 2. LFW	For traning phase LFW, For testing BAO	Recall above 95%, False positive rate below 5%	93
1. RMFRD 2. SMFD 3. LFW	LFW used for testing accuracy only	RMFD, SMFD, and LFW dataset showed 99.64%, 99.49% and 100% testing accuracy respectively	94



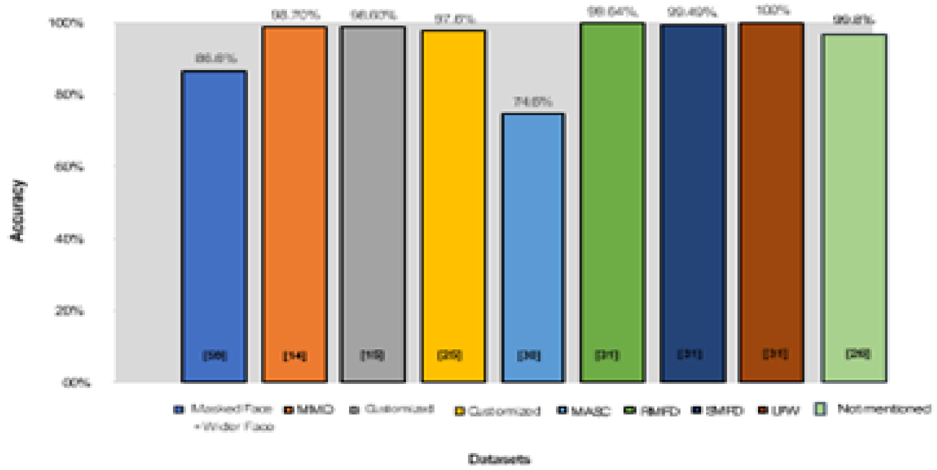


Fig.6: This graphic illustrates the accuracy of several facemask identification algorithms and their associated datasets**.**

Table 4. Reported performance of existing facemask detection methods summarized in Section III

Year	Performance Metric	Reported Performance	References
2015	Recall, False Positive	Rec. was above 95%, FP was below 5%	24
2017	Average Precision	AP was 76.4%	78
	Accuracy , Recall, IoU	Was 86.6%, Rec. was 87.8%	82
2020	Accuracy	Acc. Was 98.70%	83
	Accuracy	MaX Acc. 98.7%, Min Acc. 74.97%	84
	Precision, Recall, FI-Score	Prec. Was 98.28, Rec. was 100 and Conf.	85
	Confidence, Precision, Recall	was 99.13	86
	Precision, Recall	Prec. Score of 91.7% with Conf. score 0.7	87
	Accuracy	Acc. Was 96%	88
	AUROC	97.6%	89
	Accuracy	74.6%	92
	Recall, Precision, Testing Accuracy, FI-Score	Acc. Was 99.64%, 99.49%, and 100% for three different dataset	94 95
	Accuracy		90
	Train and Test Loss, Accuracy	Not Specified	
2021	Average Precision, Log Average Miss Rate	AP was 81%, LAMR was 0.4	91

Fig. 5 depicts the most often used and least frequently used performance indicators, as determined by various authors' assessment metrics. As shown in Fig. 15, accuracy is the most often used metric for evaluating most models; a column graph has been introduced to visualize the accuracy of different algorithms and their datasets in Fig. 6

Now, consider many of the most often used performance measures for evaluating facemask detection algorithms:

1) TRUE POSITIVE, TP

TP values relate to detection values when the model is wearing a facemask, which is visible in the image during mask detection.

2) FALSE POSITIVE, FP

In this instance, the model produces an error. The model detects the presence of a mask when none exists.

3) TRUE NEGATIVE, TN

To more accurately detect non-object zones, a genuine negative value rather than a false positive or false negative should be employed. Due to its negligible performance impact, this parameter is rarely used in mask detection.

4) FALSE NEGATIVE, FN

A false negative occurs when the model fails to recognize a facemask in a photograph.

5) ACCURACY

The accuracy of a facemask detection model is defined as the proportion of correctly detected objects relative to all possible predictions.

6) PRECISION

Precision is a term that refers to the probability of correctly predicting the bounding box of a mask using the ground truth. This is the number of authentic facemasks that the model properly detects in the image.

Precision = TP/TP + FP

7) RECALL

In the context of detection, recall refers to a detector's ability to locate and identify any possible facemask or actual value accurately.

Recall = TP/TP + FP

8) F1-SCORE

This is referred to as the harmonic mean since it incorporates precision and recall.

F1 – Score = 2 (Precision * Recall/Precision + Recall)

TABLE 4. Reported performance of existing facemask detection methods summarized in Section III.

9) CONFIDENCE

A model or algorithm can be characterized as a "mask" or "no mask" based on confidence in its skills.

10) INTERSECTION OVER UNION, IoU

The model uses this threshold to discriminate between True Positives and False Positives. It is determined as the bounding box's overlapping and union areas.

IOU = Area of Overlapping/Area of Union

5. Conclusion and Future Scope

5.1 Conclusion

In the realm of fraud detection, fraud detection is one of the fastest expanding research fields. We offered strategy comparisons, including accuracy, available algorithms, efficacy considerations, and evaluation technique and outcomes. Additionally, detailed descriptions of the datasets used in those procedures were provided. We reviewed existing algorithms for weaknesses and discovered and discussed potential challenges. Despite substantial research devoted to developing a successful facemask detection algorithm, other critical issues were mostly disregarded instead of focusing exclusively on the same set of challenges. This research identified several issues, including the requirement to maintain image resolution during the detection phase, the absence of a complete dataset, and categorical classifications.

Additionally, it specified future scopes, including a variety of datasets and facemask types and numerous scenarios of facemask usage and reconstruction of the masked face. This in-depth review would be extremely beneficial to help the scientific community better understand present facemask detection technologies by studying the field's limitations and potential challenges.

5.2 Future Scope

A future study will almost certainly gather and apply the results of emerging facemask identification algorithms to a benchmark dataset to compare their performance accurately. A single benchmark dataset will make it simple to compare the execution times, storage requirements, and accuracy of these methods.

References

[1] Tahaoğlu, G.; Ulutas, G.; Gençtürk, B. A New Robust Copy Move Forgery Detection Method.; October 5 2020; pp. 1–4.

[2] Kumar, V.; kansal, V.; Gaur, M. A Comprehensive Analysis on Video Forgery Detection Techniques. SSRN Electron. J. 2020, doi:10.2139/ssrn.3563382.

[3] Benhamza, H.; Djeffal, A.; Cheddad, A. A Review of Image Forgery Detection.; October 7 2020.

[4] Khudhair, Z.; Mohamed, F.; Abdulameer, K. A Review on Copy-Move Image Forgery Detection Techniques. J. Phys. Conf. Ser. **2021**, 1892, 012010, doi:10.1088/1742-6596/1892/1/012010.

[5] Rana, S. A Study on Image Forgery Detection. Int. J. Res. Appl. Sci. Eng. Technol. 2020, 8, 364–369, doi:10.22214/ijraset.2020.31882.

[6] Zhang, Y.; Gao, F.; Zhou, Z.; Guo, H. A Survey on Face Forgery Detection of Deepfake.; June 30 2021; p. 46.

[7] Gurunlu, B.; Ozturk, S. A Survey on Photo Forgery Detection Methods. ITM Web Conf. **2018**, 22, 01055, doi:10.1051/itmconf/20182201055.

[8] KN, S.; Chennamma, H. A Survey On Video Forgery Detection. 2015.

[9] Kumar, D.; Karthi, S.; Karthika, K.; R., C. A Systematic Study of Image Forgery Detection. J. Comput. Theor. Nanosci. 2018, 15, 2560–2564.

[10] Kumar, D.; R., C. A Systematic Study of Image Forgery Detection; 2018; Vol. 15;.

[11] Kurien, N.; Danya, S.; Ninan, D.; Raju, C.; David, J. Accurate And Efficient Copy-Move Forgery Detection.; November 1 2019; pp. 130–135.

[12] Tassia, R.; Kurdy, M.-B. Adaptive and Improved Approach for Image Forgery Detection. **2020**.

[13] Kaur, C.; Kanwal, N. An Analysis of Image Forgery Detection Techniques. Stat. Optim. Inf. Comput. **2019**, 7, doi:10.19139/soic.v7i2.542.

[14] Bhole, P. An Image Forgery Detection Using SIFT-PCA. Int. J. Eng. Res. **2020**, V9, doi:10.17577/IJERTV9IS060127.

[15] Lin, Y.-K.; Chang, C.-Y. Face Forgery Detection Based on Deep Learning. In; 2020; pp. 270–278 ISBN 978-3-030-46827-9.

[16] Sethi, S.; Kathuria, M.; Kaushik, T. Face Mask Detection Using Deep Learning: An Approach to Reduce Risk of Coronavirus Spread. J. Biomed. Inform. 2021, 120, 103848, doi:10.1016/j.jbi.2021.103848.

[17] Issa, H.; Saleh, I.; Hadri, S.; Abdulrahman, Y. Forgery Detection and Localisation in Scanned Documents.; December 28 2020.

[18] Khan, S.; Khan, K.; Ali, F.; Kwak, K. Forgery Detection and Localisation of Modifications at the Pixel Level. Symmetry 2020, 12, doi:10.3390/sym12010137.

[19] Mahfoudi, G.; Morain-Nicolier, F.; Restraint, F.; Pic, M. Object-Removal Forgery Detection Through Reflectance Analysis.; December 9 2020.

[20] Luo, Y.; Zhang, Y.; Yan, J.; Liu, W. Generalizing Face Forgery Detection with High-Frequency Features; 2021;

[21] Hyun, D.-K.; Lee, M.-J.; Ryu, S.-J.; Lee, H.-Y.; Lee, H.-K. Forgery Detection for Surveillance Video. Era Interact. Media **2013**, 25–36, doi:10.1007/978-1-4614-3501-3_3.

[22] Sapra, P.; Kikan, V. Analysis of Various Copy-Move Forgery Detection Techniques. In; 2021; pp. 400–405 ISBN 978-1-00-319383-8.

[23] Ratta, P.; Kaur, A.; Sharma, S.; Shabaz, M.; Dhiman, G. Application of Blockchain and Internet of Things in Healthcare and Medical Sector: Applications, Challenges, and Future Perspectives. J. Food Qual. **2021**, 2021, e7608296, doi:10.1155/2021/7608296.

[24] Chandani, K.; Arora, M. Automatic Facial Forgery Detection Using Deep Neural Networks. In; 2021; pp. 205–214 ISBN 9789811599552.

[25] Hebbar, N.; Kunte, A. Blind Forgery Detection in Digital Images: An Analysis. In; 2020; pp. 604–614 ISBN 978-3-030-30464-5.

[26] Nazli, M.; Maghari, A. Comparison between Image Forgery Detection Algorithms.; May 1 2017; pp. 442–445.

[27] Rani, H.; Sharma, N. Copy Move Forgery Detection with Hybrid Approach. **2020**, 9, 2585.

[28] Roy, A.; Dixit, R.; Naskar, R.; Chakraborty, R. Copy-Move Forgery Detection in Transform Domain. In Studies in Computational Intelligence; 2020; pp. 79–86 ISBN 978-981-10-7643-5.

[29] Kharanghar, M.; Doegar, A. Copy-Move Forgery Detection Methods: A Critique. In; 2021; pp. 501–523 ISBN 9789811554209.

[30] Mandankandy, A.; Shanmugam, P. Dbelm for Image Forgery Detection. **2020**, 7, 2348–6090.

[31] Khan, J. Copy-Move Forgery Detection Using Deep Learning; 2021;

[32] Dogonadze, N.; Obernosterer, J.; Hou, J. Deep Face Forgery Detection; 2020;

[33] Shah, Y.; Shah, P.; Patel, M.; Khamkar, C.; Kanani, P. Deep Learning Model-Based Multimedia Forgery Detection. **2020**, doi:10.1109/I-SMAC49090.2020.9243530.

[34] Wang, Y.; Hu, Y.; Liew, A.W.-C.; Li, C.-T. ENF Based Video Forgery Detection Algorithm. Int. J. Digit. Crime Forensics **2020**, 12, 131–156, doi:10.4018/IJDCF.2020010107.

[35] Dang, T.; Beghdadi, A.; Larabi, C. Blind Inpainting Forgery Detection.; December 5 2014.

[36] Katherineguo, J.; Daviddoermann; Azrielrosenfeld FORGERY DETECTION BY LOCAL CORRESPONDENCE. Int. J. Pattern Recognit. Artif. Intell. **2011**, 15, doi:10.1142/S0218001401001088.

[37] Lyu, Q.; Luo, J.; Liu, K.; Yin, X.; Liu, J.; Wei, Z. Copy Move Forgery Detection Based on Double Matching. J. Vis. Commun. Image Represent. **2021**, 76, 103057, doi:10.1016/j.jvcir.2021.103057.

[38] Sharma, H.; Kanwal, N. Video Interframe Forgery Detection: Classification, Technique

[39] Aljaberi, A. Topological Data Analysis for Image Forgery Detection. Indian J. Forensic Med. Toxicol. **2020**, 14, 1745–1751, doi:10.37506/ijfmt.v14i3.

[40] Menezes, H.; Ferreira, A.; Pereira, E.; Gomes, H. Bias and Fairness in Face Detection.; October 21 2021.

[41] Zecheng, T.; Xinyuan, W.; Hongli, Y.; Yansong, Z. U 2 -Net for Image Forgery Detection and Localisation.; April 1 2021; pp. 166–172.

[42] Keimel, C. Video Quality. In; 2016; pp. 11–29 ISBN 978-981-10-0268-7.

[43] Kong, R.; Zhang, B. An Effective New Algorithm for Face Recognition.; January 1 2015.

[44] Jhansi, G.; Rama, G.; Ranganath, K.; Juluri, T.; Vinay, C.; Reddy, K.; Chalamalla, V. Face Detection Authentication Analysis on Smartphones Face Detection Authentication Analysis on Smartphones. Mater. Sci. Eng. **2021**, 981, 2026.

[45] Mufti, T. FACE DETECTION AND RECOGNITION; 2020;

[46] Tathe, V.; Narote, A.; Narote, S. Face Detection and Recognition in Videos.; December 1 2016; pp. 1–6.

[47] Parekh, D. Face Detection and Recognition; 2020;

[48] TARIQ, I.; Mufti, T. FACE DETECTION AND RECOGNITION; 2020;

[49] Nobahar Sadeghi Nam, A. Face Detection. Int. J. Innov. Sci. Res. Technol. **2020**, 5, 688–692, doi:10.38124/IJISRT20SEP391.

[50] Loy, C.C. Face Detection. In; 2021; pp. 429–434 ISBN 978-3-030-63415-5.

[51] Raut, J.; Patil, S.; Gawade, S.; Meena, M. Face Detection and Recognition in Video.; March 15 2015.

[52] Daniya, T.; Gladiss, N.; R., C. Study on Digital Image Forgery Detection. **2020**, 29, 6851–6856.

[53] Khan, J. Random Forest-Based Copy-Move Forgery Detection; 2020;

[54] Jin, X.; He, Z.; Xu, J.; Wang, Y.; Su, Y. Object-Based Video Forgery Detection via Dual-Stream Networks.; July 5 2021; pp. 1–6.

[55] Uma, R.; Sathya, A. Soccer Game Optimization Based Forgery Detection of Digital Images. Forensic Imaging **2021**, 25, 200453, doi:10.1016/j.fri.2021.200453.

[56] Pavlovic, A.; Milosavljević, N.; Gavrovska, A.; Reljin, I. Copy-Move Forgery Detection Based on Multifractals. Multimed. Tools Appl. **2019**, 78, doi:10.1007/s11042-019-7277-1.

[57] Mandankandy, A.; Shanmugam, P. Random Region Rotation Based Image Forgery Detection 2021.

[58] Daniya, T.; Thirukrishna, J.T.; Kumar, D.; Kumar, M. ICSA-ECNN Based Image Forgery Detection in Face Images.; April 22 2021.

[59] M, M.; Khurshid, F. Digital Image Forgery Detection Approaches: A Review. In; 2021; pp. 863–882 ISBN 978-981-334-603-1.

[60] Kumar, V.; Kansal, V.; Gaur, M. A Comprehensive Survey on Passive Video Forgery Detection Techniques. In; 2021; pp. 39–57 ISBN 9789811584688.

[61] Kaur, S.; Kaur, M. Novel Method for Copy-Move Forgery Detection. Int. J. Comput. Appl. **2021**, 174, 10–14, doi:10.5120/ijca2021921064.

[62] AlZahir, S.; Hammad, R. Image Forgery Detection Using Image Similarity. Multimed. Tools Appl. **2020**, 79, doi:10.1007/s11042-020-09502-4.

[63] Kumar, B.; Kumar, K. Blockchain Solution for Evidence Forgery Detection. J. Comput. Theor. Nanosci. **2020**, 17, 5570–5576, doi:10.1166/jctn.2020.9454.

[64] V, B.; P, A.; A, K.; N, G.; S, C. Forgery Detection in Documents. Int. J. Electron. Commun. Eng. **2018**, 5, 15–20, doi:10.14445/23488549/IJECE-V5I5P104.

[65] Peplow, M. Holograms Aid Forgery Detection. Nature **2004**, doi:10.1038/news040809-1.

[66] Diwan, A.; Sharma, R.; Roy, A.; Mitra, S. Keypoint Based Comprehensive Copy-move Forgery Detection. IET Image Process. **2020**, 15, doi:10.1049/ipr2.12105.

[67] Alenezi, A.; Sapar, S.; Rakhimov, I. FORGERY DETECTION USING KRAWTCHOUK MOMENTS. Far East J. Math. Sci. **2019**, 111, 181–194, doi:10.17654/MS111020181.

[68] Cao, S.; Zou, Q.; Mao, X.; Wang, Z. Metric Learning for Anti-Compression Facial Forgery Detection; 2021;

[69] Shah, S.; Raskar, P. METHODS FOR FORGERY DETECTION IN DIGITAL FORENSICS. Int. J. Electron. Secur. Digit. Forensics **2021**, 1, 1, doi:10.1504/IJESDF.2021.10035056.

[70] Privman-Horesh, N.; Haider, A.; Hel-Or, H. Forgery Detection in 3D-Sensor Images.; June 1 2018; pp. 1642–16428.

[71] Satpute, R.; Shubham, P.; Mansi, P.; Shah, S.; Aditi, S. Image Forgery Detection & Localization. SSRN Electron. J. **2021**, doi:10.2139/ssrn.3838598.

[72] Wang, Xiang yang; Wang, C.; Wang, L.; Yang, H.; Niu, P. Robust and Effective Multiple Copy-Move Forgeries Detection and Localisation. Pattern Anal. Appl. **2021**, doi:10.1007/s10044-021-00968-y.

[73] Ojeda-Aciego, M.; Rodríguez-Jiménez, J. Formal Concept Analysis with Negative Attributes for Forgery Detection. Comput. Math. Methods **2020**, doi:10.1002/cmm4.1124.

[74] Chen, S.; Yao, T.; Chen, Y.; Ding, S.; Li, J.; Ji, R. Local Relation Learning for Face Forgery Detection; 2021;

[75] Jaffar, R.; Rasool, Z.; Hanon AlAsadi, A. New Copy-Move Forgery Detection Algorithm.; September 1 2019; pp. 1–5.

[76] 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,’’ 2020.

[77] Real World Masked Face Recognition Dataset. Accessed: Jun. 17, 2021.

[78] S. Ge, J. Li, Q. Ye, and Z. Luo, ‘‘Detecting masked faces in the wild with LLE-CNNs,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 2682–2690

[79] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, ‘‘MaskedFace1Net—A dataset of correctly/incorrectly masked face images in the context of COVID-19,’’ Smart Health, vol. 19, Mar. 2021, Art. no. 1001.

[80] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, ‘‘Labeled faces in the wild: A database forstudying face recognition in unconstrained environments,’’ in Proc. Workshop Faces ‘Real-Life’ Images, Detection, Alignment, Recognit. Marseille, France: Erik Learned-Miller and Andras Ferencz and Frédéric Jurie, Oct. 2008, pp. 1–15.

[81] Larxel’s Face Mask Detection Dataset. Accessed: Jun. 17, 2021.

[82] W. Bu, J. Xiao, C. Zhou, M. Yang, and C. Peng, ‘‘A cascade framework for masked face detection,’’ in Proc. IEEE Int. Conf. Cybern. Intell. Syst. (CIS), IEEE Conf. Robot., Autom. Mechatronics (RAM), Nov. 2017, pp. 458–462

[83] B. Qin and D. Li, ‘‘Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19,’’ Sen1sors, vol. 20, no. 18, p. 5236, Sep. 2020.

[84] M. Inamdar and N. Mehendale, ‘‘Real-time face mask identification using facemasknet deep learning network,’’ India, Jul. 2020.

[85] A. Chavda, J. Dsouza, S. Badgujar, and A. Damani, ‘‘Multi-stage CNN architecture for face mask detection,’’ 2020, arXiv:2009.07627.

[86] S. Yadav, ‘‘Deep learning based safe social distancing and face mask detection in public areas for COVID-19 safety guidelines adherence,’’ Int. J. Res. Appl. Sci. Eng. Technol., vol. 8, no. 7, pp. 1368–1375, Jul. 2020.

[87] M. Jiang, X. Fan, and H. Yan, ‘‘RetinaMask: A face mask detector,’’ 2020, arXiv:2005.03950.

[88] S. V. Militante and N. V. Dionisio, ‘‘Real-time facemask recognition with alarm system using deep learning,’’ in Proc. 11th IEEE Control Syst. Graduate Res. Colloq. (ICSGRC), Aug. 2020, pp. 106–110

[89] P. Khandelwal, A. Khandelwal, S. Agarwal, D. Thomas, N. Xavier, and A. Raghuraman, ‘‘Using computer vision to enhance safety of workforce in manufacturing in a post COVID world,’’ 2020, arXiv:2005.05287.

[90] M. T. C. Jagadeeswari, ‘‘Performance evaluation of intelligent face mask detection system with various deep learning classifiers,’’ Int. J. Adv. Sci. Technol., vol. 29, no. 11, pp. 3083–3087, May 2020.

[91] 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, Feb. 2021, Art. no. 102600.

[92] N.-C. Ristea and R. T. Ionescu, ‘‘Are you wearing a mask? Improv1ing mask detection from speech using augmentation by cycle-consistent GANs,’’ 2020, arXiv:2006.10147

[93] A. Nieto-Rodríguez, M. Mucientes, and V. M. Brea, ‘‘System for med1ical mask detection in the operating room through facial attributes,’’ in Proc. Iberian Conf. Pattern Recognit. Image Anal. Spain: Springer, 2015, pp. 138–145.

[94] 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, Jan. 2021, Art. no. 108288

[95] V. V. V. Vinitha, ‘‘COVID-19 facemask detection with deep learning and computer vision,’’ Int. Res. J. Eng. Technol., vol. 7, no. 8, pp. 3127–3132, 2020

Author Profile

Dr. Krishna Kumar Singh is an educator, research analyst, writer and trainer with more than 20 years of experience. He is certified professional of SAP S/4HANA, SAP Cloud analytics, SAP ERPsim, SAS analytics, Salesforce, Cortex games, Google and big data analytics. He is an associate professor at Symbiosis centre for information technology, Symbiosis International University, Pune and providing I.T training to personnel of different Industrial/corporate houses like Steel authority of India Ltd, Infosys, TCS, Oriental Bank, Axis Bank (Axis Bank Young Banker’s program), Indian army, Navy, air force and many more. He has provided training to the faculty members during different FDP’s (faculty development program) on data science at Lebanese French University, Iraq, Erbil, AKGIM Ghaziabad, U.P.College Varanasi, AGBS Pune, Graphic era University, IEC group of institutions Greater Noida, GNOIT Greater Noida, NMIMS Mumbai, Amity University Noida and many more. He has associated with National Institute for Entrepreneurship for Small Business Development (NIESBUD) a ministry of Skill initiative as trainer in the area of data analytics and conducted various training programs in information technology. He is also associated with various prestigious B-Schools in Delhi, NCR as guest or visiting faculty AIMA, ICFAI and Sharda University to name a few. Previously he was associated with Amity University, Noida, Graphic era Institute of Technology, Dehradun, College of engineering IILM academy and Galgotias College of engineering and technology as a faculty member. He is an alumnus of Banaras Hindu University and G. B. Pant engineering college, Uttarakhand. He had cleared GATE examination twice. Dr. Krishna had received Scholarship from World Bank (under TEQIP) for Ph.D and from Ministry of HRD for M.Tech. His area of research is machine learning, business analytics, financial BIG data analytics, cloud computing and green computing. There are 06 students pursuing Ph.d under his guidance. There are 05 patent, more than 30 research papers and more than 65 citations in SCI or SCOPUS indexed journals/conferences to his credit. He has written six books on Big data analytics, fraud detection analytics, MASEEI (management in action social, economic and ethical issues), Cloud computing, Mobile marketing and financial market analytics. He is an editorial board member of different SCOPUS indexed research journal and IEEE conferences. He can be contacted at singhkrishna80@gmail.com, krishnakumar@scit.edu, +91-9810866492.

Shivani Chaudhry is pursuing PHD (Computer Science & Engineering) from STAREX University , Gurugram, Haryana. She has completed her M.tech (Computer Science & Engineering) from Deenbandhu Chhotu Ram University of Science & Technology (DCRUST), Murthal ,Sonipat, Haryana, India in 2018 and She has completed her B.tech (Computer Science & Engineering) from SITM, Sonipat, India in 2016.