International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 10-No.4 (2024) pp. 1015-1031
http://www.ijcesen.com

ISSN:2149-9144

Copyright © IJCESEN

Research Article

Revolutionizing Facial Recognition: A Dolphin Glowworm Hybrid Approach for Masked and Unmasked Scenarios

Naresh Babu KOSURI[1]*, Suneetha MANNE[2]

[1]Ph.D Scholar, CSE Department, JNTUK, Kakinada, Sr. Assistant Professor, Dept of CSE (AIML), Geethanjali college of Engineering and Technology, Hyderabad, India,
* **Corresponding Author Email:** naresh.kosuri@gmail.com - **ORCID:** 0009-0001-2215-345X

[2]Professor and Head of the Department, Dept of IT, Velagapudi Ramakrishna Siddhartha Engineering college, Vijayawada, India.
Email: suneethamanne74@gmail.com - **ORCID:** 0000-0002-8917-276X

Article Info:

DOI:10.22399/ijcesen.560
Received : 25 October 2024
Accepted : 28 October 2024

Keywords :

Facial recognition,
COVID break-out ,
Cropping,
glow-worm,
Dolphin glow worm

Abstract:

Machine learning has several essential applications, including classification and recognition. Both people and objects may be identified using the Machine learning technique. It is particularly important in the verification process since it recognizes the characteristics of human eyes, fingerprints, and facial patterns. With the advanced technology developments, nowadays, Facial recognition is used as one of the authentication processes by utilizing machine learning and deep learning algorithms and it has been the subject of several academic studies. These algorithms performed well on faces without masks, but not well on faces with masks. since the masks obscured the preponderance of the facial features. As a result, an improved algorithm for facial identification with and without masks is required. After the Covid-19 breakout, deep learning algorithms were utilized in research to recognize faces wearing masks. Those algorithms, however, were trained on both mask- and mask-free faces. Hence, in this, the cropped region for the faces is only used for facial recognition. Here, the features were extracted using the texture features, and the best-optimized features from the glow worm optimization algorithm are used in this paper. With these features set, the hybrid Dolphin glow worm optimization is used for finding the optimal features and spread function value for the neural network. The regression neural network is trained with the optimized feature set and spread function for the face recognition task. The performance of the suggested method will be compared to that of known approaches such as CNN-GSO and CNN for face recognition with and without masks using accuracy, sensitivity, and specificity will next be examined.

1. Introduction

The traditional biometric technologies relying on passwords or fingerprint are no longer secure since the COVID-19 may spread via touch and contaminated surfaces. It is safer to use face recognition without having to touch any hardware. Recent coronavirus research has shown that employing a face mask significantly lowers the spread of this virus among healthy and afflicted people. However, the following issues arise while using a face mask: (1) Scam artists and criminals take use of the mask to perpetrate theft and other crimes covertly. (2) When a large portion of the face is covered by a mask, community security systems and face identification are becoming particularly challenging jobs. (3) Whenever a mask is worn that prevents a whole facial picture from being provided for description, current face identification techniques are ineffective. (4) As it is needed for face normalization, position correction, and facial matching, revealing the nose area is crucial for the work of face identification. Facial masks have greatly hampered face recognition technologies because to these issues.
The effectiveness of facial-recognition algorithms is influenced by image quality. When comparison to a camera phone, the image appearance of scanned video is relatively poor. Even high-definition video is typically 720p, but it may be as high as 1080p (progressive scan). These numbers correspond to around 2MP and 0.9MP, although a

low-cost digital camera may capture 15MP. The distinction is clear to see.

Picture Size

How successfully a face will be identified depends on its relative size to the total picture size when a face-detection algorithm discovers a face in an image or in a still from a video clip. The recognised face is just 100 to 200 pixels wide due to the already modest picture size and the target's distance from the camera. Furthermore, it takes a lot of processing power to scan a picture for different face sizes. To reduce false positives during detection and hasten picture processing, the majority of algorithms allow for the choice of a face-size range.

Partially blocking

The process of face recognition is destroyed by object view elements. The major source of worry for the facial recognition procedure is that the accused individuals may deliberately conceal their faces in order to trick the security systems. By using part-based, feature-based, and fractals-based algorithms for effective face recognition, these issues may be resolved.

Pose diversity

The extraction of head characteristics becomes more difficult due to the position variation, which is one of the major obstacles for face recognition since the person's head may be rotated anywhere from 90 degrees to 45 or even 60 degrees. Mathematical normalisation is impracticable when the person's faces are inwards rotated. By choosing appearance-based strategies, geometric model-based methods, and multi-view-based methodologies, it may be overcome.

Lighting Situations

Face recognition software's performance is significantly impacted by lighting. The facial photos produce a shadow due to varying lighting conditions, making the identifying procedure difficult. Numerous methods utilised to address this problem and proposed by various scholars are not applicable in day-to-day activities.

Skin ageing

In general, as people age, their faces develop wrinkles, scars, moles, freckles, colour changes, texture changes, and shape changes, making them impossible to identify in the future and ultimately interfering with facial recognition technology. Present FR engine are not resistant to modifications brought on by age. For face recognition to be successful, an effective algorithm has to be implemented.

As a result of its important uses in surveillance systems, healthcare systems, and entertaining, facial recognition software is receiving increased attention in real world applications. Generally general, face emotion detection technology is created to foster human-system interaction. In this study, automated facial expression identification is created in order to recognise human face expressions for usage in real world applications. Numerous researchers have studied face recognition, but they still encounter certain restrictions. A new effective method is developed for a simpler automated recognition approach to address these problems. In order to address these issues, we separate the tasks of face mask identification and mask face identification. The first one determines whether or not the wearer is hiding their identity. This is applicable in public settings when wearing a mask is required. On the other hand, masked face recognition seeks to identify a face with a mask based on the areas around the eyebrows and nose. In this research, we use a deep learning-based approach to the second problem. We extract characteristics from the unmasked facial areas using a deep learning-based model that has already been trained (out of the mask region). It is important to note that the nose and mouth areas are the only predicted face locations where the occlusions in our scenario might occur; this can serve as a helpful guide to effectively tackle this issue. The remainder of this essay is structured as follows: The previous literature are presented in Sect. 2. We outline the inspiration for and significance of the work in Section 3. Sect. 4 provides specifics on the suggested approach. Empirical values are shown in Section 5. The article comes to a conclusion at the end.

2. Preliminary Knowledge

Through the use of a video camera, digital face recognition analyses a person's face pictures to determine its features. It measures the dimensions of the whole face structure, including those between the borders of the eyes, nose, mouth, and jaw. These measures are kept in a database and then compared and matched to provide an effective outcome. Although this biometric authentication is effective in identifying possible threats (terrorists, con artists, and criminals), it has several drawbacks that prevent it from being employed in high-level application. The biometric technology needs the person to physically present themselves for scanning, which may include putting their palm on a reader or their eyes in front of the scanning (Chen and Cheng). However, human assistance should not be necessary for the facial recognition systems. Unsurprisingly, there

is no delay or disruption, and most of the time, individuals are totally unaware that their faces are being photographed as they enter a designated area (under surveillance or that their privacy has been invade). There are around 80 nodal nodes on each human face, and they may be counted by creating a numerical code called a "face print" that symbolizes entire face images. The face characteristics vary from person to person due to age issues, recognisable landmark, and various peaks and troughs. The following list of characteristics used in face recognition includes some of them:

- The separation between the eyes
- the depth and breadth of the eye sockets
- the cheekbones' form
- and the thickness of the cheek bones

The system will find the patient's face and scans the images when the person (user) faces that camera while positioned around two feet away from the camera. These images are then utilised for further processing as well as for matching against by the authentication process or the facial databases. Based on his face position, the user might have to relocate and retry the authentication if it wasn't successfully detected the first time. In far less than 5 seconds, the algorithm often makes a wise choice.

3. Related Work

Due to convolutional neural networks' (CNN) reliability in terms of data analysis and feature extraction due to their deep architecture, which enables them to perform a very large amount of computations on a single region of an image, most recent work in face recognition uses this deep learning technique. This technique is based on the architecture of convolutional neural networks (CNN), which was inspired by biological neural networks.The findings demonstrate that Mobile Net has the highest level of accuracy for both camera-input photos and video (real-time)[1-11] Oriented directional characteristics from a modular histogram are used to recognise faces: This study proposes a local directed pattern descriptor and modular histogram-based illumination-invariant face recognition system. An oriented local descriptor that may represent diverse patterns of face pictures under varied lighting circumstances is the proposed Modular Histogram of Oriented Directional Features (MHODF) [12]. Each sub-image texture is encoded using edge response values in various directions, and each picture is given a multi-region histogram. The accuracy of face recognition is greatly improved by the edge reactions, which are crucial. As a result, we demonstrate the usefulness of employing various directional masks, such as Prewitt kernels, Kirsch masks, Sobel kernels, and Gaussian derivative masks, for identifying the edge responses. The suggested MHODF algorithm's performance was assessed using a number of publically accessible datasets, and encouraging recognition rates were found. use cases using computing. Transfer learning to adopt weights from a related task, face recognition, which is learned on a large dataset, is one of the characteristics. To determine whether or not individuals were wearing face masks, we employed OpenCV, Tensor Flow, Keras, Pytorch, and CNN. Real-time video feeds and photos were used to evaluate the models. The model's precision has been attained, and it is continually being optimised. By adjusting the hyperparameters, we are creating a very accurate answer. This particular model may be used as an example of an edge analytics use case. On a public face mask dataset, the suggested approach also produces state-of-the-art results. as a result of the advancement of face mask detection [13]. a technique for identifying and alerting someone who misuses or doesn't use a face mask. The method makes use of the GoogleNet, AlexNet, and VGG16 models in deep learning. The bagging approach, an ensemble method, is used to synthesise the findings. According to the empirical results, the technique offers a face mask detection performance of more than 95% [14]. On the face mask detection dataset, the identification model is assessed. According to the experiment's findings, DENSENET201-SVM and EFFNET-LSVM had the greatest classification accuracy, with 0.9972. EFFNET-LSVM, meanwhile, offers the benefit of faster extracting features, classifier, and features size calculation [15]. the rise in cognitive bias brought on by masks, and they propose that individuals examining choices made by algorithms for facial recognition should indeed be notified of this possibility [16]. The purpose of this study is to identify a person wearing a mask by their face. When using a face mask and the Kaggle face database, the system's performance was assessed for COVID-19. The findings reveal that the suggested system performed well and outperformed the other face-recognition algorithms [17]. The bagging

approach, an ensemble method, is used to synthesise the findings. The technique provides a more then 95% efficiency of face mask identification, according to the experimental data [18]. The rate of change in the difficulty of face recognition throughout time was observed to be 80% in the normal participants and 34% in the RRMSp. When recognising faces with facial masks on, the RRMSp performed worse than healthy controls and often needed to remove the coverings. In contrast to healthy controls, RRMS participants did not demonstrate as much progress in the ability to identify masked images over time [19]. Whether the individual is wearing a disguise or not, the training system will still be able to identify them.

The suggested detection and recognition is also evaluated with current state-of-the-art methods, with the recognition accuracy for various classes averaging about 98.2% [20]. By additionally showing algorithm accuracy rates while wearing face masks, you may determine if this rise in cognitive bias was influenced by the perception of algorithm correctness. That the increases in motivated reasoning brought on the face masks was lessened by making individuals aware of the possibility for algorithmic mistakes. According to our research, people examining face recognition algorithms judgements should indeed be constantly reminded of the possibility of algorithmic faults to enhance the effectiveness of the human-algorithm collaboration These discoveries have significant ramifications for happiness and fulfilling social connections. The need for doctor-patient contacts in public general practice is even larger if some mental patients find it difficult to understand happy feelings. Additionally, masked face images were used to evaluate this approach, and the results indicate that it performs better than traditional CNN-GSO.

4. Proposed Work

The main advantages of this proposed approach is that the dataset contains both normal and masked images. This helps to create an automatic attendance system with face recognition procedure. This helps to minimize people interaction and maintain social distance between them. A summary of the proposed approach is shown in Fig. 1. There are four stages to it:

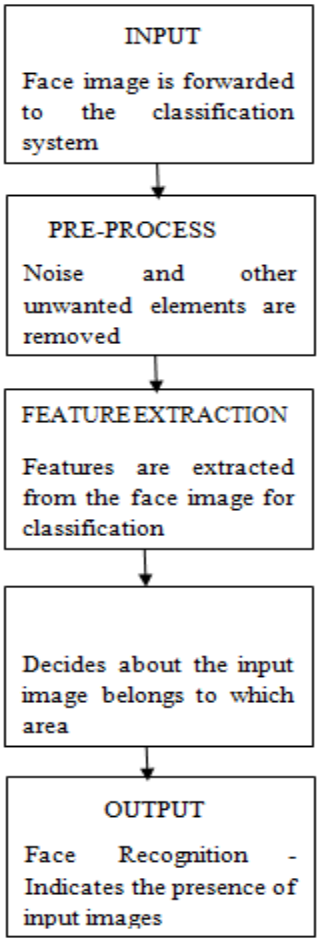

Figure 1. *Step-by-Step Processes for Face Recognition*

A. Image Acquisition

The first step in any classification process is image acquisition. For pre-processing, the input picture for the face image dataset is gathered from a searchable database. The image is reduced in colour and shrunk to 128 by 128 before being sent to the pre-processing step.

B. Pre-processing by wiener filter

Pre-processing is the method of reducing noise and enhancing the aesthetic appeal of a picture to make it more accurate for identification. The primary objective of the preprocessing method is to enhance the diagnosis potential for visual evaluation as well as computer-aided segmented. It is crucial to decrease noises by keeping the important image features when an image is acquired and transmitted since noisy corrupts the representation throughout both processes. In this study, the image quality is improved and removed using a wiener filter as preprocessing. The Wiener filter is often used in the

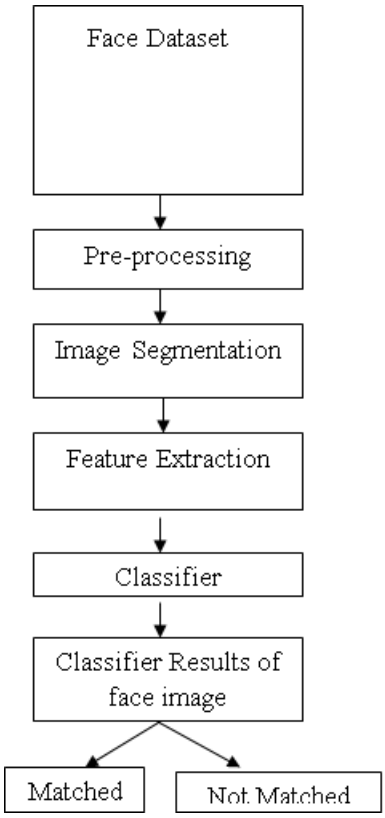

Figure 2. *Flowchart for Hybrid CNN*

frequency response to restore images that have been distorted by noise signal and white noise degradation caused by the Mean Squared (MSE) optimum and second stationarity filter. The impact of noise dispersion is not considered during inverted filter operation.The primary function of this filtering is to simultaneously reduce salt and white additive white gaussian noise, including such Gaussian noise, smoothing the input picture at low variance and smooth it more even when it is at high variance.The Wiener filtering is employed using two strategies, namely the continuous - time filter and the spectral properties of the input signal as well as the disturbance in the input picture (figure 2).

Given a corrupted input image *I (u,v)* transformed to Discrete Fourier Transform (DFT) to obtain *I(x,y).*Using the result of, the original picture spectrum is produced*I(x,y)* with the Wiener filter *F(x,y).*

$$S(x,y) = I(x,y)F(x,y) \qquad (1)$$

One estimate for the Wiener filter is,

$$F(x,y) = \frac{H^*(x,y)P_s(x,y)}{|H(x,y)|^2 P_s(x,y) + P_n(x,y)} \qquad (2$$

$H(x,y)$ denotes blurring filter and $P_s(x,y)P_n(x,y)$the energy spectra of the noise and the signal are indicated. The Wiener filter is used to clear the image and minimize distortions (filtering) (noise part).

$$F(x,y) = \frac{H^*(x,y)P_s(x,y)}{|H(x,y)|^2 + P_{sn}(x,y))} \qquad (3)$$

Fig.3 demonstrates the Wiener filter's spectral properties$P_{sn}(x,y)$ denotes noise-to-signal ratio. $P_{sn}(x,y)$ high frequency region is moderately huge as compared to $(P_{sn}(x,y) \gg |H(x,y)|)$and it has the effect of hiding the restoring

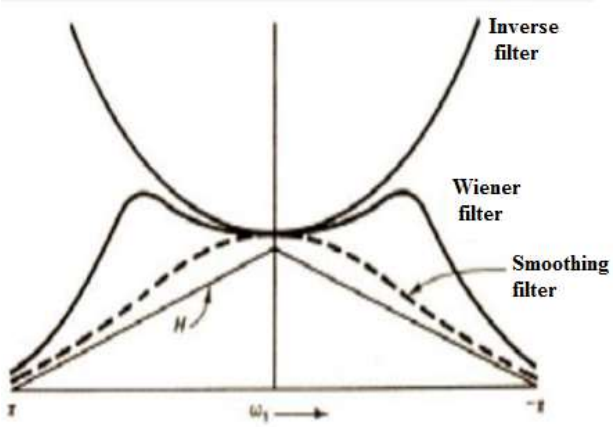

Figure 3. *Power Spectrum Characteristics of Wiener Filter*

filter's high frequency range. Wiener filter smoothens the input image so that the result is produced by multiplying each component by $P_s(x,y)$ is shown in the below equation

$$F(x,y) = \frac{H^*(x,y)}{|H(x,y)|^2 + \frac{P_n(x,y)}{P_s(x,y)})} \qquad (4)$$

$\frac{P_n}{P_s}$indicates the inverse of the signal-to-noise ratio, and the correlation between the noise and the signal is quite high$\frac{P_n}{P_s} \sim 0$ and the Wiener filter becomes $H^{-1}(x,y)$ denoted as inverse filter. Using this example, suppose the weak signal is$\frac{P_n}{P_s} \to \alpha$, $F(x,y) = 0$

The Wiener filter may be reduced to the following when there is no blur and additional white noise:

$$F(x,y) = \frac{P_s(x,y)}{P_s(x,y) + \sigma_n^2} \quad (5)$$

$\sigma_n^2$is a measure of noise power. If the picture is distorted by noise, the Wiener filter's frequency components cannot be rebuilt. By using the inverse wavelet transform, the denoised data may be recovered using the variance and wavelet coefficients harmonic values.

C. Feature Extraction

The goal of this technique is to retrieve the known feature that includes certain image-related data. To remove the identified skin area, use Gray Level Co-occurrence Matrix morphological operations. In order to simplify the process of the operation, precise information about the characteristics is provided by the skin area that has been excised. It is crucial to distinguish the texture of the image by estimating the pixel with particular values and stipulated geometric distribution of the image in order to obtain the quantitative measurements from this matrix. GLCM is a 2nd order way of measuring that measures the relationship between neighbouring pixels and considers the relationships predicated on the harvesting of a gray-scale image.

The following formula (Albregtsen) is used to extract the following GLCM features: Cluster Prominence, Cluster Shade, Inertia, Sum Entropy, Difference Entropy, Homogeneity, Angular Second Moment (ASM), Local Homogeneity, Inverse Difference Moment (IDM), Contrast, Entropy, Correlation, Sum of Squares, Variance, and Sum Average. Therefore, processing time will be reduced by extracting these properties. Consider P to be the input textured image's normalised GLCM. Here are a few of the calculated equations for certain characteristics.

$$Energy = \sum_{u,v} P(u,v)^2$$

$$Contrast = \sum_{u,v} |u-v|^2 P(u,v)^2$$

$$Homogeneity = \sum_{u,v} \frac{P(u,v)}{1+|u-v|}$$

$$Correlation = \sum_{u,v} \frac{P(u,v)}{1+|u-v|}$$

$$ClusterShade = \sum_{u,v=0}^{N-1} P_{uv}\,(u - M_i + v - M_j)^3$$

$$ClusterProminence = \sum_{u,v=0}^{N-1} P_{uv}\,(u - M_i + v - M_j)^4$$

The smooth of the image region is measured by the Inverse Distance Moment, which is the opposite of Contrast. The weight are equivalent to the square root of the inverse of the difference in grey levels the inverse of the Contrast weights.

It also goes by the name Uniformity and assesses how uniformly the picture area's texture is. When the region either has an existence value or even a recurring pattern of values, its value is at its peak.

The Generalized Glowworm Swarm Optimization (MGSO) algorithm selects the retrieved attributes to categorise the facial expression. The procedure of feature selection picks a portion of the extracted image's physical features while preserving the structural nature of the original features. The suitable candidates subset among the difficult 2N candidate subsets is selected using feature selection approaches based on a variety of estimation function. In order to identify the face expression that is adjusted by global search to increase the classification accuracy, glow-worm swarm optimization is used. GSO is a swarm - based programme that relies on glow worm behaviour. It is primarily chosen for the search of several optima, each with a distinct value for the optimization problem. A swarm must be able to split into several groups in order to achieve this goal; otherwise, just one global or local optimal will be attained. The behaviour of glowworms (to precisely locate the liver-affected area) is clearly characterised by their capacity to alter the strength of their luciferin emission, which causes them to seem to glow at various intensities.

The individuals in the glow worm search agent contain a substance called luciferin that emits light, and each glow worm is drawn to the stronger glow of its nearby neighbours. When a glow worm is located inside its local decision domain, it recognises the other glow worm as a neighbour. The fitness of the present places is related to the luciferin intensity of the glow worms. The location of the glow worm in the search space is better the greater the luciferin intensity. Each glow worm's position will vary with

each iteration, and the amount of luciferin will also be updated.

The distribution of glow worms, luciferin updates, motion, and choice are the four stages that make up every repetition.

GSO implementation: In mGSO, every glow worm is distributed randomly at the start of the search space by adding dynamic tent map of chaotic n, a nonlinear phenomenon that improves periodicity and unpredictability and makes it easier for the glow worms to find the effective optimal value.

$$x_{k+1} = \begin{cases} 2x_k, 0 \le x_k \le 0.5 \\ 2(1-x_k), 0.5 \le x_k \le 1 \end{cases} \quad (7)$$

Update of Luciferins - In this stage, the sensing range and amount of each Luciferin are updated equally. It is assessed using optimization problem and is described as

$$l_i(t+1) = (1-\rho)l_i(t) + \gamma j_i(t+1) \quad (8)$$

Where ρ denotes luciferin decay constant and it range lies between $(0 < \rho < 1)$, $j_i(t)$ denotes the luciferin enhancement constant and $\gamma j_i(t)$denotes the objective function value at glow-worm *i*'s location at time *t*. When there are more than 10 iterations, the locations of the 5% of glow worms with the lowest fitness are swapped out for the average position of all glow worms to speed up the search process and arrive at the best result.

Movement step -The motion rule is used in mGSO to increase convergence time. The light worms constantly migrate in the direction of their neighbour j, who has a strong luciferin value that was produced via a probabilistic technique. The $p_j(t)$ represents the probability of glow-worms in movement phase which can be represented as

$$p_j(t) = \frac{(l_j(t) - l_i(t))}{\sum_{k \epsilon n_i(T)} (l_k(t) - l_i(t))} \quad (9)$$

$l_i(t)$ represents the luciferin value of glow-worm and the movement of glow-worms '*i*' can be evaluated as

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t) - x_i(t)}{\|x_j(t) - x_i(t)\|}\right) \quad (10)$$

Where *s* denotes step-size.

Decision Phase - In this phase the best value is decided andluciferin rule is updated, with the neighbour changes which can be represented as

$$r_d^i(t+1) = min\{r_s, max\{0, r_d^i(t) + \beta(n_t - |N_i(t$$

(11)

Where $r_d^i(t+1)$ denotes the local-decision domain of *'i'* at the *t+1* iteration, βrepresents a fixed variable that affects how quickly the neighbour region changes ,$'n_t'$is referred to be a boundary which regulates the number of clusters. As a consequence of this feature extraction, the most pertinent and advantageous characteristics are chosen and sent to a hybrid technique of the Dolphins as well as GSO for classifying facial gestures.The following steps are taken in order to build this experimental model:

- Create a keras sequentially linear stack architecture to make it easier to later add other neural network layers;
- Convolution layer 1 should be created, with the dimension of each face picture being 57 by 47, the number of build filters being 16, and the size of each filter being 3 by 3, with the same parameter being included to guarantee that the convolution will be produced.
- The max pooling 1 is created, the first complicated is carried out, and 16 5747 images are lowered to half, while the number remains unchanged. The fully connected layers image size remains constant, and ReLu is chosen as the learning algorithm of the component. As a result, 16 images are generated after the first convolution layer.
- Converting the original 16 images into 36 images requires creating two layers: a convolution layer 2 with a filter frequency of 36; a pooling layer 2 with a second downsampling that reduces the image of the pooled layer 1 by half more; and a layer with dropout. The benefit is that it will arbitrarily be included in each training iteration of the neural network. discard neurons in part to prevent overfitting
- Create a hidden state, choose ReLu as the layer's input signal, and then add the Drop - outs layer to the models once again
- Creating a flat level to transform the image of the pooling layer 2 to a one-dimensional vector.

- Create the convolutional layers and apply the activation function of softmax to transform and anticipate the probability
- Establishing the output

Auto update in database

A database containing the person names in columns and the date in rows will be built in order to carry out this procedure. The attendance will be marked as present in the excel sheet or database once the image has been identified through the categorization procedure. The automated attendance system that may be performed with the help of this system saves time.

5. Results & Discussion

We experimented on highly difficult masked viewing to assess the suggested technique. The dataset' contents and variances, experiment results employing the quantification of features extracted gleaned through several pre-trained algorithms, and a comparison with other state-of-the-arts are all covered in the sections that follow.

Dataset description

The primary goal of the Real-World-Masked-Face-Dataset is to enhance the effectiveness of the current face recognition algorithm on masked faces it during COVID-19 epidemic. It includes three different kinds of images: The Simulated Masked Face Recognition Dataset (SMFRD), the Real-World Masked Face Recognition Dataset (RMFRD), and also the Masked Face Detection Dataset (MFDD). In this study, we concentrate on the most recent datasets listed below.

- One of the most comprehensive real-world masked face datasets is RMFRD. It has 5,000 photos of 525 people with masks and 90,000 photographs of 525 subjects without mask. The important facial portions have been cropped using a semi-automatic annotating approach.
- The SMFRD includes 500,000 emulated mask faces of 10,000 participants drawn from the Labeled Faces in the Wild (LFW) and Webface datasets, two well-known datasets. Dlib library is used to perform the simulation. Although this dataset is comparable, it makes it more difficult because the generated coverings may not always be in the proper place.

Figure 4.*Sample Images from Dataset*

Figure 5. *Preprocessing Results (a). Input Image, (b). Gray Conversion and (c). Median Filtering*

Variables - X

X

150x12 single

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	5.1911	1.3708	-0.9600	3.3481	0.2994	7.1345	0.0977	0.9739	0.9531	0.2026	1.9418	38.1333	
2	4.6349	1.1542	-0.8864	3.3205	0.3162	6.9932	0.0737	0.9750	0.9641	0.2492	1.8741	35.2000	
3	5.6526	2.8026	-1.1134	3.3004	0.2450	7.4485	0.1148	0.9840	0.9461	0.1731	1.9985	31.0667	
4	5.2813	4.7656	-0.7377	2.1847	0.1889	7.6276	0.1718	0.9857	0.9264	0.1377	2.0401	29.3333	
5	4.3612	2.4544	-0.9177	2.7261	0.2338	7.0498	0.0704	0.9874	0.9653	0.2060	2.0144	20.4000	
6	5.2017	4.8678	-0.6723	2.0650	0.1837	7.6502	0.1687	0.9863	0.9261	0.1337	2.0371	28.9333	
7	5.8009	1.6713	-1.3609	4.5238	0.2844	6.8987	0.0868	0.9785	0.9595	0.2899	2.0873	35.7333	
8	4.0207	3.0833	-0.4442	1.8490	0.1863	7.4562	0.1016	0.9858	0.9501	0.1616	1.8845	35.3333	
9	5.7287	1.9326	-1.3744	4.4183	0.2749	6.9642	0.0845	0.9819	0.9606	0.2796	2.0917	30.5333	
10	4.7424	4.3592	-0.5837	1.9235	0.1772	7.5910	0.1312	0.9874	0.9350	0.1483	1.8273	32.5333	
11	4.6941	1.6027	-1.0221	3.5106	0.2686	7.1779	0.1301	0.9670	0.9368	0.2056	1.8673	33.8667	
12	4.7045	3.4959	-0.8011	2.3834	0.2085	7.4744	0.1041	0.9872	0.9495	0.1677	1.9398	30.2667	
13	4.4708	2.8786	-0.4659	2.2031	0.1778	7.6127	0.1075	0.9852	0.9467	0.1258	1.8040	34.6667	
14	4.8385	4.3990	-0.6741	2.1061	0.1814	7.5310	0.1328	0.9872	0.9349	0.1385	1.8912	30.6667	
15	4.6904	2.5562	-0.3790	1.9355	0.1937	7.5307	0.0960	0.9855	0.9522	0.1424	1.8484	50.4000	
16	4.5216	3.3700	-0.2715	2.0414	0.1530	7.7993	0.1367	0.9835	0.9353	0.1102	1.8291	48.4000	
17	5.0783	4.0367	-0.8914	2.4892	0.2051	7.3913	0.1250	0.9865	0.9393	0.1851	1.9454	33.8667	
18	3.7292	1.0233	0.6742	5.5624	0.3747	6.8142	0.0905	0.9641	0.9580	0.2844	2.0134	34.8000	
19	4.3567	2.3234	-0.3032	2.2591	0.1973	7.5019	0.1073	0.9821	0.9498	0.1521	1.9021	40.5333	
20	4.6909	4.1204	-0.2516	1.4903	0.1953	7.5622	0.1211	0.9891	0.9416	0.1325	1.9448	44.5333	

Figure 6. *Feature Extraction Result*

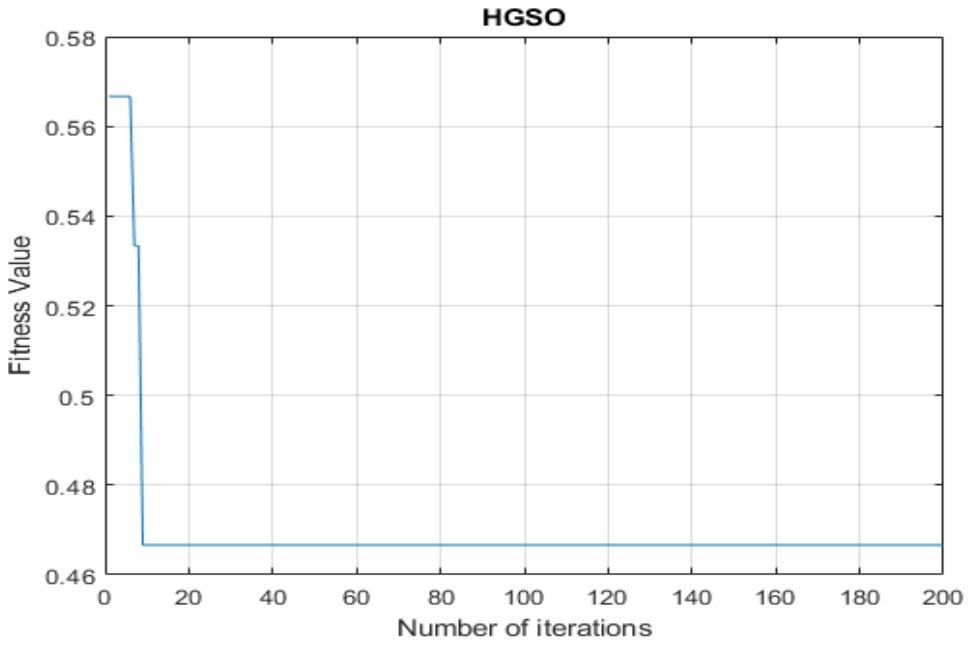

Figure 7. *Final optimal feature set*

Figure 8. *(1) Input Image, (2). Gray Conversion, and (3). Median Filtering*

Figure 9. *SURF Features*

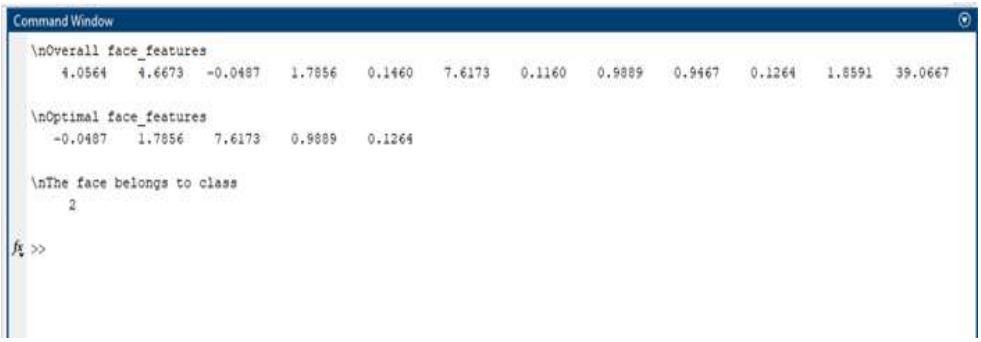

Figure 10. *Feature extraction and class*

Figure 11. *(1)Input Image, (2). Gray Conversion, and (3). Median Filtering*

Figure 12. *SURF Features*

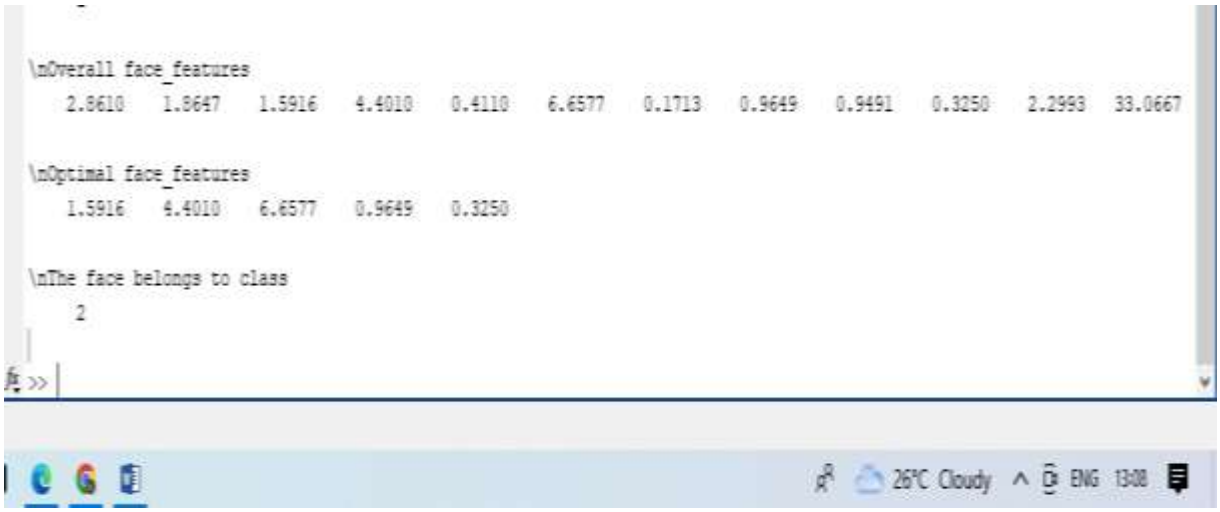

Figure 13. a*. Feature extraction and class*

S.No	Month	Date	Aidai	anhu	aixin	baibahe	baijin
1	Dec-23	29	present	present	present	absent	present
2	Dec-23	30	absent	present	absent	present	present
3	Dec-23	31	present	present	present	present	present

Figure 13.b. *Update the result in database*

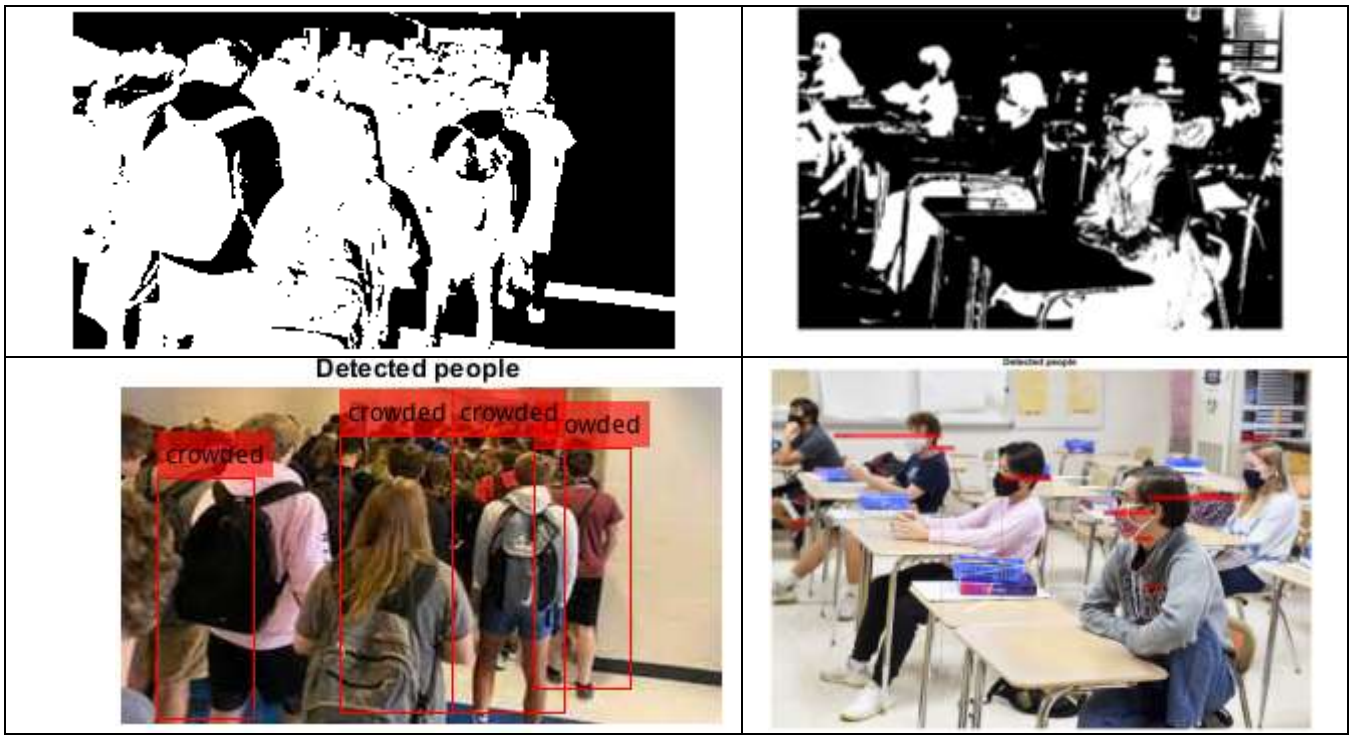

Figure 14. *Social distance*

Five different people images without mask at different illumination and angles were captured for feature extraction and training.

The above process was common for both training and test set with and without masks. Figs (4-12) illustrates the list of features extracted, row represents the images, column represents the features and feature selection is implemented using hybrid dolphin glowworm swarm optimization algorithms.

From figure 13. (a) the features for the test images will be extracted and the classes (that is the name of the person) will be determined. Then, the corresponding person attendance will be updated in the database as shown in the figure 13.b.Figure 14 shows the process in the social distance detection. In this, the left column shows the highly crowded images and the right column less crowded images. The first row is the input image, second row represents the gray scale image, third row represents the binarized images and final row represents the social distance results. The accuracy, sensitivity, and specificity measurements are used to assess the performance of the hybrid algorithm, as well as the results are taken from the binary classification. The temperature readings are assessed using

tr_p = True Positive Rate

tr_n = True Negative Rate

fa_p = False Positive Rate

fa_n = False Negative Rate

By using the method below, the specificity and the sensitivity are calculated.

$$sen = \left(\frac{tr_p}{tr_p + fa_n}\right) \quad (12)$$

$$spec = \left(\frac{tr_n}{tr_n + fa_p}\right) \quad (13)$$

ROC analysis is utilized to assess facial image classification accuracy. The x and y axes in this ROC are shown as

$$x = falsepositiverate(1 - spec) \quad (14)$$

$$y = truepositiverate\ (sen) \quad (15)$$

Table 1 shows the comparison evaluation of performance metrics such as accuracy, sensitivity and specificity for HYBRID-CNN and existing CNN and CNN-GSO. Because of high true positive rate the HYBRID-CNN attains efficient classification of facial recognition of accuracy 95.84%, sensitivity 90.12% and specificity 96.12%.

Table 1 *Comparison of Performance Metrics*

Performance metrics	HYBRID-CNN	CNN-GSO	CNN
Accuracy	96.74	95.12	88
Sensitivity	95.12	94.1	88.4
Specificity	95.78	94.2	88.1

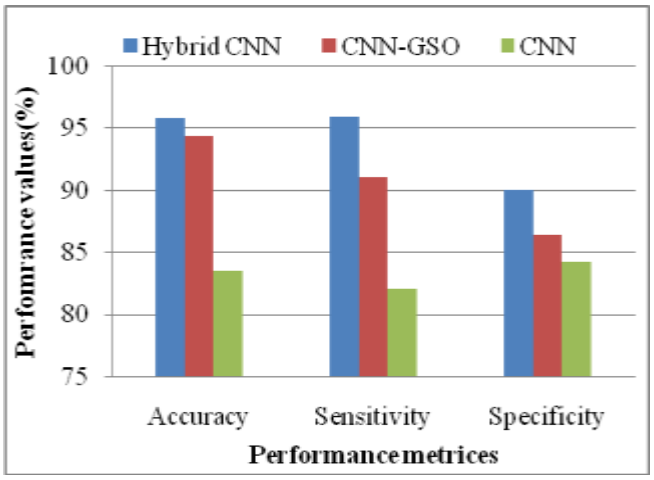

Figure 15*. Performance Comparisons of Various Classifiers*

Figure 6 shows the graphical representation of comparison of accuracy, sensitivity and specificity for HYBRID-CNN and existing CNN CNN-GSO. From the graph it is clear that the proposed HYBRID-CNN shows better results when compared to other algorithms.

Table 2 shows evaluation of ROC performance for facial expression recognition by using HYBRID-CNN and compared with CNN and CNN-GSO. The efficient pre-processing, feature extraction and feature selection with optimization techniques shows best values of specificity of 95.12%.

Table 2 *Results of the ROC Mathematical Evaluation for All Classification models*

(1-spec)	HYBRID-CNN	CNN-GSO	CNN
0.1	0.25	0.17	0.11
0.2	0.31	0.21	0.15
0.3	0.39	0.25	0.15
0.4	0.48	0.33	0.22
0.5	0.62	0.45	0.3
0.6	0.73	0.57	0.31
0.7	0.8	0.7	0.43
0.8	0.96	0.78	0.6
0.9	1.01	0.94	0.67

The evaluation of ROC effectiveness for HYBRID-CNN, current CNN, and CNN-GSO is shown in Figure 7. The chart makes it evident that HYBRID-CNN outperforms other approaches in terms of effectiveness.

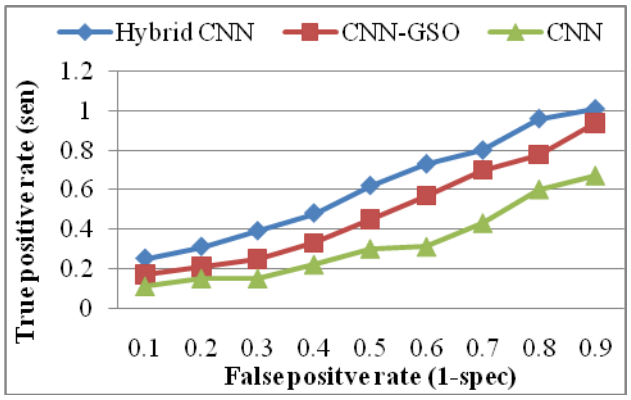

Figure 16*. compares the properties and performance of several classifications using ROC curves*

Table 3 compares CNN-GSO with CNN when evaluating processing speed. The planned system was completed in less time due to the shorter training period and effective pre-processing. Figure 7 illustrates the comparative comparisons of processing performance between the proposed HYBRID-CNN and the current CNN-GSO and CNN. It is evident that the HYBRID-CNN classifies facial images from huge datasets in a shorter amount of time. The accuracy, recall, and F-measure evaluation results of the HYBRID-CNN are shown numerically in Table 4. The suggested HYBRID-CNN obtained an accuracy, recall, and F-measure performance values of 96.1%, 97.14%, and 96.45% due to a low negatives rate and a significant positive prediction rate. The proposed HYBRID-CNN and the current CNN-GSO and CNN are compared in Fig along with their

Table 3 *Processing times for all classifications methods*

No of images	HYBRID-CNN	CNN-GSO	CNN
10	8.012	10.745	16.417
20	16.14	18.147	22.417
30	21.524	24.254	28.154
40	24.47	27.147	33.177
50	29.147	32.2574	39.5412

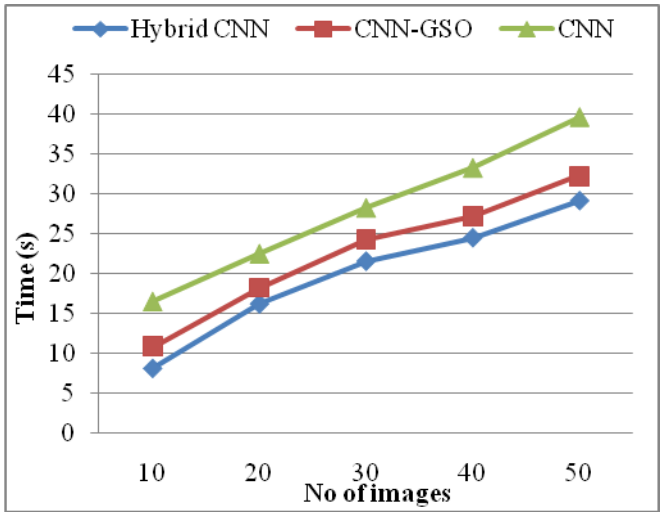

Figure 17. *Comparison of process times for different classifications*

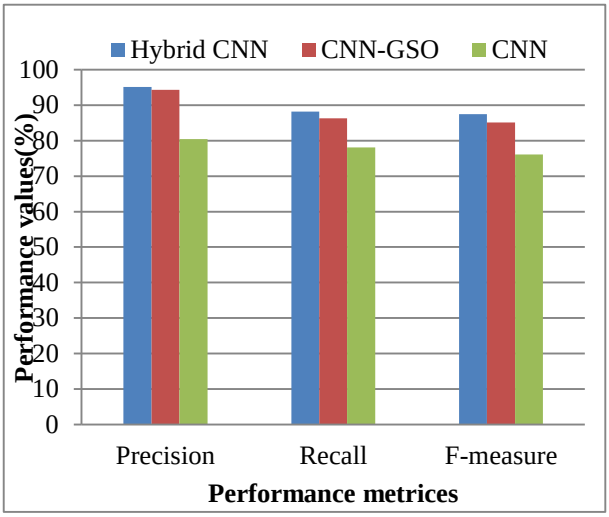

Figure 18. *Comparison of Precision, Recall, and F-measure for Different Classification models*

Table 4 *The empirical evaluations for all classifications' precision, recall, and F-measure*

Performance metrics	HYBRID-CNN	CNN-GSO	CNN
Precision	95.47	90.12	88.25
Recall	95.24	90.25	88.15
F-measure	95.36	90.52	89.14

F-measure performances. It demonstrates that when compared with current methods, the proposed Method achieves good accuracy, recall, and F-measure results. Again, for proposed HYBRID-CNN and the current CNN-GSO and CNN classifications, the outcomes of the evaluation of performance parameters including sensitivity, specificity, accuracy, recall, F-Measure, accuracy, and processing time are shown in Tables 5, 6, and 7. To assess the suggested effectiveness, each parameter is examined using 100 total images from combination dataset for both Indian and JAFFE. The figure 17 make it evident that the proposed Method achieves the greatest outcomes compared to other classifications already in use, thanks to appropriate feature identification, a quick training procedure, and optimal kernel parameter settings

6. Conclusion and Future work

By using deep and machine-learning learning algorithms, face detection is employed as a part of the authentication mechanisms, which has been the focus of several academic investigations. On features without mask, those methods worked well, but poorly on those with masks. since the majority of the face characteristics were hidden by the masks. A better method is thus needed for face recognition both when using and without masks. Deep learning techniques were used in studies to identify faces using masks during the Covid-19 outbreak. However, those algorithms were developed on features with and without masks. As a result, in this case, face recognition is the sole use made of the clipped area for the faces. The features in this case were retrieved to use the texture characteristics, and this work uses the glow worm optimization algorithm's best-optimized features. The hybrid Dolphin glow worm optimization is applied with these features set to determine the neural network's ideal features and spreading value of the objective function. The spread functions and optimized feature set for the face recognition problem are used to train the regression neural network. The effectiveness of the proposed strategy will then be compared to well-established techniques as CNN-GSO and CNN for facial recognition software with and without masks, looking at accuracy, sensitivity, and specificity.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.

***Table 5**: Outcomes of Mathematical Analysis of Different Parameters for Hybrid-CNN*

Dataset Images	tr_p	tr_n	fa_p	fa_n	sen (%)	$spec$ (%)	Accuracy (%)	F-measure (%)	Precision (%)	Recall (%)	Processing time (s)
5+5	6	2	0	2	90.12	98.05	96.14	0.9130	95.05	88.12	27.87
10+10	11	5	1	3	90.11	96.1	95.47	0.9130	94.98	88.09	28.14
15+15	18	7	1	3	90.12	94.21	95.14	0.9247	95.12	88.1	29.04
20+20	27	9	1	4	90.10	93.14	94.78	0.9400	94.97	88	29.21
Total=100	62	23	3	12	90.12	96.12	95.84	87.45	95.1	88.14	29.147

***Table 6:** Outcomes of Mathematical Analysis of Different Parameters for CNN-GSO*

Dataset Images	tr_p	tr_n	fa_p	fa_n	sen (%)	$spec$ (%)	Accuracy (%)	F-measure (%)	Precision (%)	Recall (%)	Processing time (s)
5+5	3	4	0	3	91.04	86.06	94.52	88.12	94.89	86.45	32.01
10+10	7	8	0	5	91.2	86.24	94.48	87.57	94.54	86.34	32.15
15+15	12	10	1	7	91.1	86.32	94.42	87.45	94.32	86.21	32.22
20+20	18	12	1	9	91.14	86.45	94.44	87.34	94.12	86.12	32.45
Total=100	40	34	2	24	91.1	86.4	94.45	87.45	94.35	86.32	32.25

***Table 7**: Outcomes of Mathematical Analysis of Different Parameters for CNN*

Dataset Images	tr_p	tr_n	fa_p	fa_n	sen (%)	$spec$ (%)	Accuracy (%)	F-measure (%)	Precision (%)	Recall (%)	Processing time (s)
5+5	1	4	1	4	82.43	84.45	84.14	76.4	80.54	78.24	38.54
10+10	4	10	0	6	82.21	84.31	84.1	76.24	80.34	78.16	39.12
15+15	9	13	0	8	82.11	84.15	83.87	76.1	80.28	78.1	39.54
20+20	14	15	1	10	82	84.1	83.78	76	80.32	78.04	39.78
Total=100	28	42	2	28	82.1	84.21	83.5	76.14	80.41	78.12	39.5412

- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1]Mohan, M., Sojasomanan, &Kaleeswari, M. (2021). Automatic Face Mask Detection Using Python.

[2]Hieu Luu, T., Nguyen Ky Phuc, P., Yu, Z., Dung Pham, D., & Trong Cao, H. (2022). Face Mask Recognition for Covid-19 Prevention. *Computers, Materials & Continua*.

[3]Razali, M.N., Shafie, A.S., &Hanapi, R. (2021). Performance Evaluation of Masked Face Recognition Using Deep Learning for Covid-19 Standard of Procedure (SOP) Compliance Monitoring. *2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE), 6*, 1-7.

[4]Barragán, D., Howard, J.J., Rabbitt, L.R., &Sirotin, Y.B. (2022). COVID-19 masks increase the influence of face recognition algorithm decisions on human decisions in unfamiliar face matching. *PLOS ONE, 17*.

[5]Alaluosi, W.M., & Mohammed, A.S. (2022). Biometrics Face Recognition Using Method of Wavelet and Curvelet Transforms with COVID-19. *Review of Computer Engineering Studies*. DOI:10.18280/rces.090207

[6]Hieu Luu, T., Nguyen Ky Phuc, P., Yu, Z., Dung Pham, D., & Trong Cao, H. (2022). Face Mask Recognition for Covid-19 Prevention. *Computers, Materials & Continua*. 73(2) doi: 10.32604/cmc.2022.029663

[7]Kuzu Kumcu, M., Tezcan Aydemir, S., Ölmez, B., Durmaz Çelik, N., &Yücesan, C. (2022). Masked face recognition in patients with relapsing–remitting multiple sclerosis during the ongoing COVID-19 pandemic. *Neurological Sciences, 43*, 1549 - 1556.

[8]Fatima, M., Ghauri, S.A., Mohammad, N.B., Adeel, H., & Sarfraz, M. (2022). Machine Learning for Masked Face Recognition in COVID-19 Pandemic Situation. *Mathematical Modelling of Engineering Problems*. DOI:10.18280/mmep.090135

[9]Barragán, D., Howard, J.J., Rabbitt, L.R., &Sirotin, Y.B. (2022). COVID-19 masks increase the influence of face recognition algorithm decisions on human decisions in unfamiliar face matching. *PLOS ONE, 17*.

[10]Escelsior, A., Amadeo, M.B., Esposito, D., Rosina, A., Trabucco, A., Inuggi, A., Pereira da Silva, B., Serafini, G., Gori, M., & Amore, M. (2022). COVID-19 and psychiatric disorders: The impact of face masks in emotion recognition face masks and emotion recognition in psychiatry. *Frontiers in Psychiatry*, 13:932791. doi: 10.3389/fpsyt.2022.932791

[11]You, C.E., Pang, W.L., & Chan, K.Y. (2022). AI-Based Low-Cost Real-Time Face Mask Detection and Health Status Monitoring System for COVID-19 Prevention. *WSEAS TRANSACTIONS ON INFORMATION SCIENCE AND APPLICATIONS*. 19:256-263 DOI:10.37394/23209.2022.19.26

[12]Chelbi, S., &Mekhmoukh, A. (2022). A practical implementation of mask detection for COVID-19 using face detection and histogram of oriented gradients. *Australian Journal of Electrical and Electronics Engineering, 19*, 129 - 136. https://doi.org/10.1080/1448837X.2021.2023071

[13]Min, D., Anandamurugan, S., Mohanasundaram, K., Pandiyan, P., Thangaraj, R., & Kaliappan, V.K. (2023). Real-time face mask position recognition system using YOLO models for preventing COVID-19 disease spread in public places. *International Journal of Ad Hoc and Ubiquitous Computing*. 42(2);73 – 82 DOI: 10.1504/IJAHUC.2023.128499

[14]Guerra, N.C., Pinto, R., Mendes, P.S., Rodrigues, P.F., & Albuquerque, P.B. (2022). The impact of COVID-19 on memory: Recognition for masked and unmasked faces. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.960941

[15]Castellano, G., De Carolis, B., &Macchiarulo, N. (2023). Automatic facial emotion recognition at the COVID-19 pandemic time. *Multimedia Tools and Applications*, 82, 12751–12769. https://doi.org/10.1007/s11042-022-14050-0

[16]Chester, M., Plate, R.C., Powell, T., Rodriguez, Y., Wagner, N.J., & Waller, R. (2022). The COVID-19 pandemic, mask-wearing, and emotion recognition during late-childhood. *Social Development (Oxford, England)*. https://doi.org/10.1111/sode.12631

[17]Bansal, A., Dhayal, S., Mishra, J., & Grover, J. (2022). COVID-19 Outbreak: Detecting face mask types in real time. *Journal of Information and Optimization Sciences, 43*, 357 - 370.

[18]Singh, S., Aggarwal, A., P, R., Nelson, L., Damodharan, P., & Pandian, M.T. (2022). COVID 19: Identification of Masked Face using CNN Architecture. *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 1045-1051.

[19]Mbombo, J.K. (2022). Peace in the Face of the COVID-19 Pandemic: Making Sense of the Paralysis at the UN Security Council. *Peace & Change*. https://doi.org/10.1111/pech.12512

[20]Dange, B.J., Khalate, S.S., Kshirsagar, D.B., Gunjal, S.N., Khodke, H.E., & Bhaskar, T. (2022). Face Mask Detection under the Threat of Covid-19 Virus. *2022 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS)*, 1-6.