



Neural networks contribution in face mask detection to reduce the spread of COVID-19

Maminiaina Alphonse Rafidison¹ · Andry Harivony Rakotomihamina¹ · Sabine Harisoa Jacques Rafanantenana¹ · Rajaonarison Faniriharisoa Maxime Toky¹ · Mirado Mike Noé Raelina¹ · Hajasoa Malalatiana Ramafiarisona¹

Received: 27 October 2022 / Revised: 7 December 2022 / Accepted: 21 February 2023 /

Published online: 4 March 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

In front of COVID-19 propagation, we can protect our self by taking precautionary measures such as wearing face masks. It may be mandatory in particular public place although some persons ignore this rule. Several research in face mask detection area have emerged and most of studies are based on deep learning. In this paper, we present a method to detect whether person wear a mask or not to prevent the propagation of virus. The approach is based on combination of Pulse Couple Neural Network and Fully Connected Neural Network and the processing is divided in three steps: geometrical, feature extraction and decision. The geometrical module selects the Region of Interest for given image and the feature extraction module composed by Pulse Couple Neural Network extracts all pertinent information which will be used by the last module for decision. This decision module makes directly a decision in case of non-complex classification without neural network training otherwise the Fully Connected Neural Network continues the treatment. The input image may be captured from video surveillance sequence, the system triggers a signal alarm once a person doesn't wear face mask. Our proposed approach was tested with different datasets like Kaggle, AIZOO, Moxa3K, Real-World Masked Face Dataset, Medical Masks Dataset, Face Mask Dataset and the accuracy varies from 83.2% to 100% with minimum computation time.

Keywords Eyes detection · Face mask detection · Fully connected neural network · Orthogonal projection · Pulse couple neural network · Segmentation

1 Introduction

Before pandemic, people used a mask due of air pollution reason and currently, scientists have proved that it is one of best way to prevent COVID-19 transmission. According to World

✉ Maminiaina Alphonse Rafidison
maminyaina@gmail.com

Health Organization (WHO) report statement on a weekly epidemiological update on COVID-19 on 18th January 2022, edition 75: “the number of new COVID-19 cases increased in the past week (10-16 January 2022), while the number of new deaths remained similar to that of the previous week. Across the six WHO regions, over 18 million new cases were reported this week, a 20% increase, as compared to the previous week. Over 45 000 new deaths were also reported. As of 16 January, over 323 million confirmed cases and over 5.5 million deaths have been reported worldwide. Despite a slowdown of the increase in case incidence at the global level, all regions reported an increase in the incidence of weekly cases with the exception of the African Region, which reported a 27% decrease. The South-East Asia region reported the largest increase in new cases last week (145%), followed by the Eastern Mediterranean Region (68%). New weekly deaths increased in the South-East Asia Region (12%) and Region of the Americas (7%) while remaining approximately the same as the previous week in the other regions” [28]. These statistics demonstrate the rapidity contamination of COVID-19, so it motivates us to orient our research in face mask detection.

Due of recent apparition of COVID-19 pandemic, different studies were done to detect human face mask and most of them focus in deep learning using Convolutional Neural Network (CNN) or MobileNet which needs a lot of parameters and composed by the succession of convolution and spooling layers. With this concept, the solution architecture is complex, some pertinent information may be lost during feature extraction phase and the computation time is considerable.

In this paper, we will propose an efficient approach to detect the presence of face mask in human visage based on Pulse Couple Neural Network (PCNN) and Fully Connected Neural Network (FCNN). To accomplish this, we process the input image by detecting the eyes and once we have eyes position, we are able to identify the region of interest (ROI) of the face. This ROI will be the input of PCNN and a statistic module handles the binary output image by calculating different parameters such as gray level uniformity (GU), gray level contrast (GC), entropy (E) and cross-entropy (CE) to help decision module for the verdict in case of classic image otherwise the Fully Connected Neural Network (FCNN) decides for a complex case. To validate our approach, we applied it into different datasets used by the recent related works cited in the next paragraph and we have an improvement compared with them.

The main contributions of this work are cited below:

1. This paper may help to reduce the fast spread of COVID-19.
2. The solution doesn't need a prior training for a simple image.
3. The method is simple with system architecture non-complex because there are just a few parameters required and for this reason, it allows us to have an optimal computation time. Such kind of solution may be supported by embedded systems.
4. With this approach, different type of facemask can be detected such us a standard facemask or medical facemask.
5. In the health crisis, police and gendarmes therefore had a fundamental role in ensuring compliance with the rules applicable in the public space. Facemask detection is part of this task and can be automated using this research.

The rest of the paper is organized as follows: paragraph 2 summarizes the recent works related to our proposed approach. The paragraph 3 described the proposed method followed by performance comparison and discussion in paragraph 4. Finally, the paragraph 5 concludes

the paper. To ensure a good understanding of this article, Table 1 presents the list of abbreviations and definitions.

2 Literature review

As searcher in image processing area, we fill having a responsibility to contribute on non-propagation of COVID-19 and face mask detection is one method that we can propose for humanity purpose. Most of new publication research on this domain is focusing on deep learning which had a complex architecture because they use a CNN to have a best accuracy rate, so in front of that we decide to make a challenge by suggesting an approach without machine learning for non-complex scenario to compete the existing studies. The complicated case is always managed by a light convolutional neural network.

In 2019, during the first apparition of COVID-19 in Madagascar, the government mobilizes the military in the aim to make an indirect pressure to the population in the road or public place to wear a face mask. The objective is to consider it as normal gesture and part of habitual life. This military activity can be automated with intelligent system and they can focus on their main occupation. It is a benefit for the government to avoid resource wastage.

We opt PCNN for two reasons: first, it extracts easily the important information present in image without training or succession of convolution/spooling operations and a decision can be taken directly if needed. Secondly, it has an ability to form the image signature. Note that PCNN neural network has a static architecture independently of type of image to be processed.

In this section, we list the work jugged pertinent and related to this approach by ascending chronological order.

Table 1 Abbreviations table

Abbreviation	Definition
CE	Cross Entropy
CNN	Convolutional Neural Network
COVID-19	Coronavirus Disease 2019
E	Entropy
FCNN	Fully Connected Neural Network
FMD	Face Mask Dataset
GC	Graylevel Contrast
GU	Graylevel Uniformity
ICM	Intersecting Cortical Model
LFW	Labelled Faces in the Wild
MAFA	Masked FAcEs
MMD	Medical Masks Dataset
OM	Overall Merit
PCNN	Pulse Coupled Neural Network
Resnet50	Convolutional neural network that is 50 layers deep
RMFD	Real-World Masked Face Dataset
ROI	Region of Interest
SMFD	Simulated Masked Face Dataset
SVM	Support Vector Machine
VGG-16	Variant of VGG model with 16 convolution layers
WHO	World Health Organization
YOLO-v2	You Only Look Once - Version 2

Sammy V. Militante, Nanette V. Dionisio [22] (2020) published a paper entitled “Real-Time Facemask Recognition with Alarm System using Deep Learning”. This work introduces deep learning techniques VGG-16 CNN to distinguish facial recognition and recognize if a person puts or not a facemask. The accuracy result is around 96% trained with dataset which contains 25,000 images using 224×224 pixel resolution. The system develops a Raspberry Pi-based real-time facemask recognition that alarms and captures the facial image if the person detected is not wearing a facemask. A part of camera and speaker, the Raspberry Pi block is composed by facemask detection module, loading model, classification and google speech.

Loey, Mohamed, et al. (2021) [19] did a research in facemask detection “A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic”. The proposed study combines two methods: deep learning and classical machine learning. Normally, a standard CNN is composed by feature extraction layers and classification layer. Authors kept the feature extraction layers of Resnet50 Convolutional Neural Network and fully connected neural networks is swapped by the combination of decision trees, Support Vector Machine (SVM) and ensemble algorithm. The approach was validated with three datasets: Real-World Masked Face Dataset (RMFD), Simulated Masked Face Dataset (SMFD), and Labeled Faces in the Wild (LFW). The SVM classifier achieved 99.64% testing accuracy in RMFD. In SMFD, it achieved 99.49%, while in LFW, it achieved 100% testing accuracy.

Loey, Mohamed, et al. (2021) [20] same authors as previous work published a paper entitled “Fighting against COVID-19: A novel deep learning model based on YOLO-v2 (used for real time object detection) with ResNet-50 (50-layer convolutional neural network (48 convolutional layers, one MaxPool layer, and one average pool layer)) for medical face mask detection”. The concept remains the same that’s to say that the feature extraction is ensured by pooling and convolution layers of Resnet50 and the next step consists to detect medical face mask with YOLO-v2. They merge two medical face masks datasets to validate the method. To improve the object detection process, mean IoU has been used to estimate the best number of anchor boxes. The achieved results concluded that the adam optimizer achieved the highest average precision percentage of 81% as a detector. ResNet50 and YOLO-v2 take an important role in proposed detector model because after CNN feature extraction, YOLO-v2 alimnts the network training module in the aim to help evaluate detector module for decision.

Bhawna R., Udit U. (2021) [6] proposed a method face mask detection, “Face Mask Detection Using Convolutional Neural Network”. There are three steps to follow such as images preprocessing, extraction of pertinent information and classification. A standard tool for machine learning was used like Scikit-Learn, Keras, OpenCV and TensorFlow. The experience was performed with 1376 images taken from GitHub without and with mask. 90% of this dataset is used for training and 10% for validation. The model generates higher accuracy which was 99.1% whereas validation lost 87.5% and validation accuracy 97.3%.

Riya C., Rutva J. (2021) [25] published an article entitled “Detection of Face Mask using Convolution Neural Network”. They are using python libraries and MobileNet-v2 as convolutional neural network which is playing the role of transfer learning. The advantage is that there is some pretrained model for new model training. It can reduce the processing time. To train, validate and test the concept, they utilized the dataset that consisted of 993 masked faces pictures and 1918 exposed faces pictures. These pictures were taken from different assets like Kaggle and RMFD datasets. The success rate of person wearing mask detection is around 96% depending on the digital capabilities.

Gagandeep K., Ritesh S., et al. (2021) [12] present a research result entitled “Face mask recognition system using CNN model”. The study is based on fundamental machine learning using a known tool like Scikit-Learn, TensorFlow, OpenCV and Keras. Due of parameter value optimization of Convolutional Neural Network (CNN), the technique has a good performance. Five steps are followed to attain this result: data collection (collect 2 datasets with mask and no mask then labelling data), preprocessing (resize/convert image size, preprocess the data using mobilenet_v2 and perform hot encoding on labels), data splitting (split data for training and testing), building model (construct the raining image generator for augmentation based on mobilenet_v2 with “imagenet” weight, adding model parameters and saving/compiling model), testing model (make prediction on the testing set, plot training loss and accuracy and model evaluation), implementation model (read video by frame and resize, call preprocessing function and predict input data from saved model). The dataset used was downloaded from [Kaggle.com](https://www.kaggle.com) website containing 3832 images divided into two mask classes (1914 with mask and 1918 images without mask) and the performance judged high.

Goyal, H., Sidana, K., Singh, C. et al. (2022) [15] proposed an approach of face mask detection “A real time face mask detection system using convolutional neural network” based on Convolutional Neural Network (CNN). The proposed model has three modules: Data pre-processing (load the images, label images into its class, pre-process the image, split the dataset into training and testing and apply data augmentation to the images), CNN model training (RELU as activation function for hidden layers, add 5 conv. Layer/add 5 max pooling layers, insert flatten and dense layers, use Softmax for output layer, train model on training dataset and use adam optimizer, same the model) and Applying face mask detector (load the saved model, detect faces using face detector from OpenCV, pass as input to facemask detection model and decide if person wears a mask or not). This technique is validated using Kaggle dataset and the accuracy reaches 98%, a considerable improvement compared with MobileNet-V2, DenseNet-121, Inception-V3 and VGG-19.

To address the problem of face mask recognition in facial images, the authors (2022) [2] present another method of facemask detection using Depth-wise Separable Convolution Neural Network based on MobileNet (DWS-based MobileNet) instead of 2D convolution layers. The model decreases the number of parameters involved in training by adopting a lightweight network. The first step is to collect images with and without mask and process then in image processing module. Secondly, the depth-separable convolutional neural based on MobileNet takes over the flow and sending the output to model training/testing module before prediction. The overall accuracy is around 93% employing two different datasets: AIZOO and Moxa3K.

3 Proposed model

The system architecture of proposed model (Fig. 1) has three parts:

1. geometric module composed by eyes detection and mask filter. Once we have eyes position, we can do a projection to get the ROI of mask to form mask filter. This mask filter allows us to get the important region of human face which will be processed by the next module.
2. feature extraction module is the combination of PCNN which extracts the pertinent information and statistical analysis. This submodule introduces the notion of GU, GC,

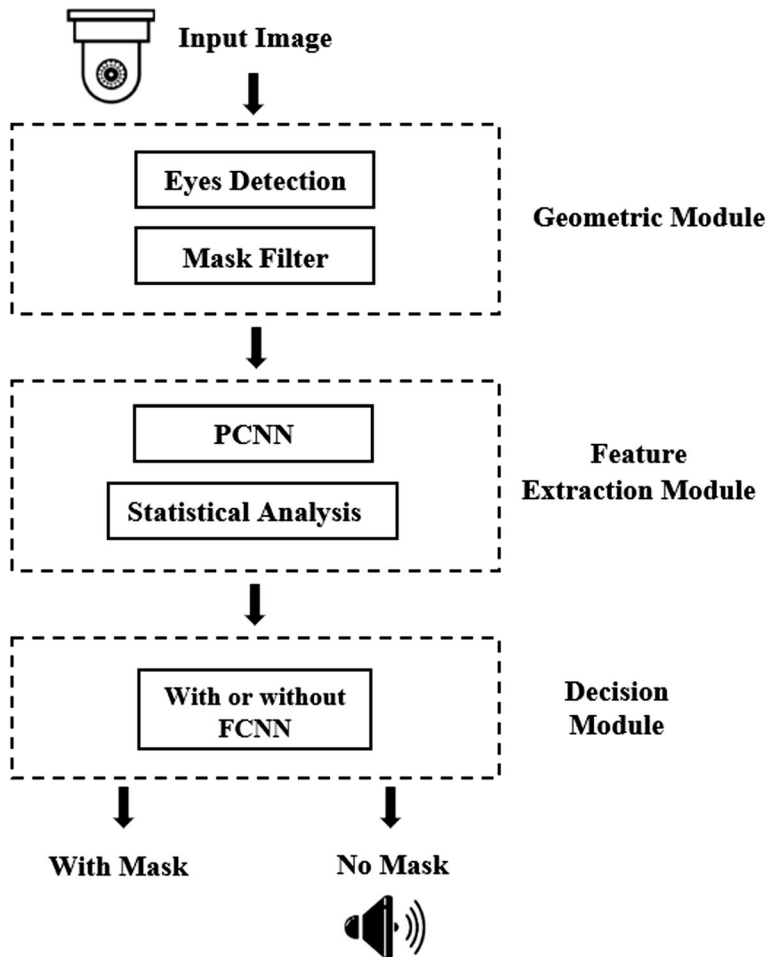


Fig. 1 Face mask detection method

E, CE and OM. The evaluation of these parameters helps decision module to determinate if the person is wearing mask or not.

3. decision module decides immediately whether the input of feature extraction module is non-complex otherwise the FCNN does the computation.

Before going to first module, the input image must undergo to preprocessing by converting it to gray image then resize. Resizing is very important to guarantee the performance of geometric module.

```

7      % Eyes detection
8
9      % object detector using voila-jones face-detection algorithm
10     % the argument can be Eyes, Mouth, Nose, Upper body
11     detectorEyes = vision.CascadeObjectDetector('RightEye');
12
13     boundingboxEyes=step(detectorEyes,imgray);
14
15     detpicEyes=insertObjectAnnotation(imcolor,'Rectangle',boundingboxEyes,'EYE');

```

Fig. 2 Eyes detection code in Matlab

3.1 Geometric module

3.1.1 Eyes detection

In Matlab, a cascade object detector is available which uses the Viola-Jones algorithm. It has an ability to detect easily people's faces, noses, eyes, mouth or upper body. Our interest is in eyes detection and we share below the code in Fig. 2 that we use for this submodule.

Where “imgray” is the matrix of input gray image and “imcolor” is the resized original color image. Let's take two images from face mask data set shown in Fig. 3 available in data-fair.training website. Applying the code, we have the images in Fig. 4.

In general, the face mask should be under eyes region so we must find a method to isolate the region of interest (ROI) with mask filter.

3.1.2 Mask filter

Now we have eyes positions. Consider $A1(x1, y1)$ and $A2(x2, y2)$ the extreme points of eyes. First, we calculate the middle of the segment $[A1A2]$ and draw a line (Δ) which is parallel with $[A1A2]$ and perpendicular to a line pass in middle of $[A1A2]$, perpendicular to $[A1A2]$. Second, we calculate the orthogonal projection of the internal points of eyes region through (Δ) then we get two points $A3$ and $A4$. The mask interest region is the region R formed by $\{A1, A2, A3, A4\}$. Details presentation is in Fig. 5.

Applying this theory in our experimental images, we have the following results (Fig. 6).

Isolation of ROI in image gray original is the last step in mask filter. All regions outside of ROI are considered as black (gray level zero) then we have the output in Fig. 7.



(a) Original image without mask



(b) Original image with mask

Fig. 3 Sample experimental input image

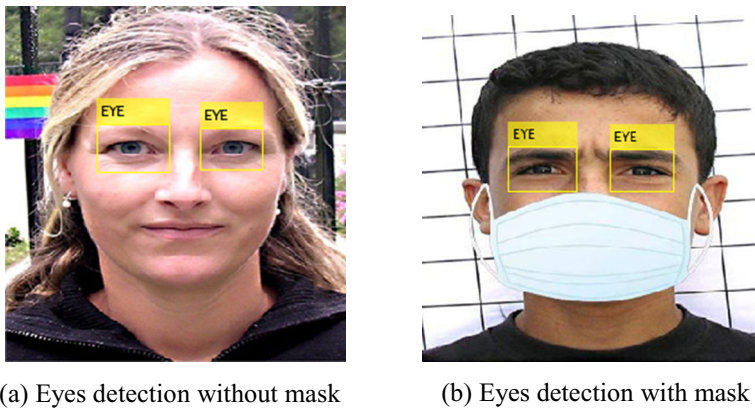


Fig. 4 Eyes detection using Viola-Jones method

3.2 Feature extraction module

3.2.1 Pulse couple neural network

The Pulse Couple Neural Network (PCNN) was invented in 1990 and recognized as powerful tool in image processing especially for image segmentation and image edge detection [7]. It has two types of input: feeding which is the gray level of a direct pixel, linking which is carrying the neighboring neuron information and the weights M and W are laterally connected. According to the Fig. 8, a PCNN is divided into three parts: inputs, linking and pulse generator which defines the threshold Θ in order to compare with the internal activity of neuron U to decide the value of the output Y (1 if neural activity is above the threshold and 0 otherwise).

The equation of each module is listed in Table 2 as following from Eqs. (1) to (5) and the advantage of PCNN is that a training is not required, only an iteration is enough.

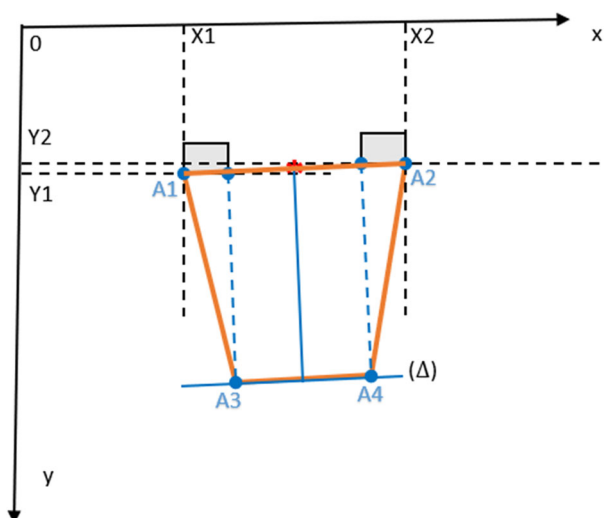


Fig. 5 ROI method detection

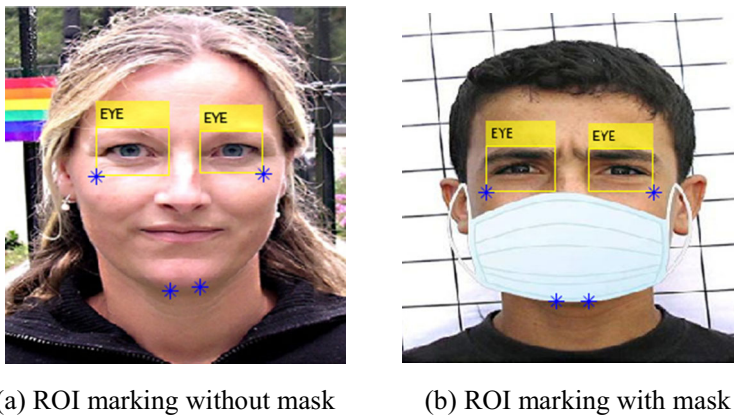


Fig. 6 ROI marking

Where S_{ij} , (α_L , α_F , α_θ , V_F , V_L , V_θ) are respectively the external stimulus and (normalized constants) [23]. The parameters value used during this study is given in Table 3 below and the constant synaptic weights M and W are defined in Eq. (6). They are chosen according to the research done by Xiangyu D. and Ma Y. [29] “how to choose PCNN parameter values to have a good performance” then we applied it to our technique.

$$M = W = \begin{bmatrix} 0.5 & 1.0 & 0.5 \\ 1.0 & 1.0 & 1.0 \\ 0.5 & 1.0 & 0.5 \end{bmatrix} \quad (6)$$

The input image is initial values of L , F and S matrix. The convolution between null matrix which has the same size as the input image R-by-C and weights matrix initiates the output value Y of PCNN. The first value of dynamic threshold θ is an R-by-C matrix of two.

As PCNN input image is simple which is the ROI then 20 iterations are enough to obtain a good segmented image. For each iteration, PCNN generates a binary image. Johnson [17, 18] constructed such a transform, a kind of sum operation for each output binary image of the

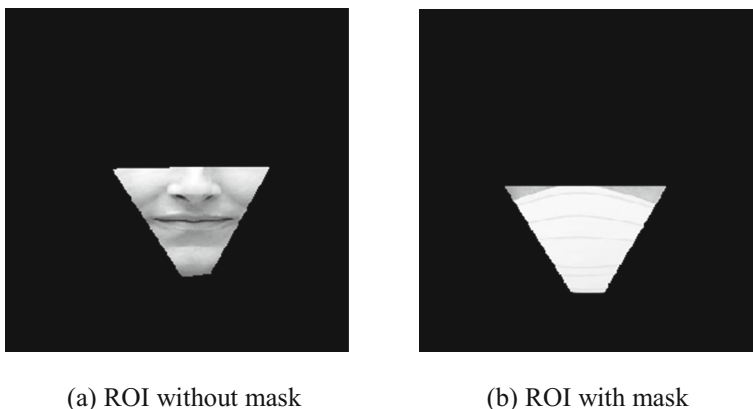


Fig. 7 ROI isolation

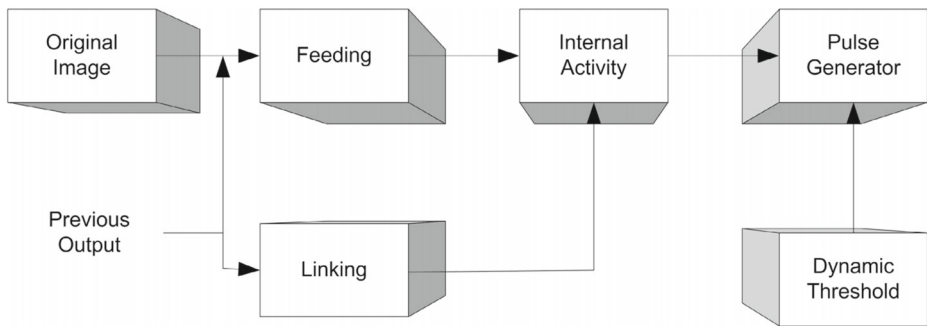


Fig. 8 PCNN diagram

PCNN, and then a one-dimensional time series $G[n]$ is obtained using Eq. (7), where $Y[n]$ is the output binary image of the PCNN at the n^{th} iteration.

$$G[n] = \sum_{i,j} Y_{i,j}[n] \quad (7)$$

From time series graph in Fig. 9 for each iteration, visually we are able to classify an image with mask and no mask. The time signature of no mask is less than with mask. We cannot stay on this statement, we should find a solution to make all processing automatic instead of interpreting. Both figures below (Figs. 10 and 11) represent the image segmented for each iteration. The problem that we solve is to select one image in which contains an important information.

3.2.2 Statistical analysis

We need to identify the best segmented image during iteration. For that, four parameters are considered to evaluate the quality of image segmentation: Gray-level Uniformity (GU), Gray-level Contrast (GC), Entropy (E) and Cross-Entropy (CE).

Table 2 Equation list of each PCNN component

Module	Equation	
Feeding	$F_{ij}(n) = S_{ij} + F_{ij}(n-1) \cdot \exp(-\alpha_F) + V_F \cdot (M * Y(n-1))_{ij}$	(1)
Linking	$L_{ij}(n) = L_{ij}(n-1) \cdot \exp(-\alpha_L) + V_L \cdot (W * Y(n-1))_{ij}$	(2)
Internal activity	$U_{ij}(n) = F_{ij}(n) \cdot (1 + \beta \cdot L_{ij}(n))$	(3)
Outputs	$Y_{ij}(n) = \begin{cases} 1, & \text{if } U_{ij}(n) > \Theta_{ij}(n) \\ 0, & \text{else} \end{cases}$	(4)
Dynamic threshold	$\Theta_{ij}(n) = \Theta_{ij}(n-1) \cdot \exp(-\alpha_\Theta) + V_\Theta \cdot Y_{ij}(n)$	(5)

Table 3 Parameters of PCNN in the experiments

Parameters	α_L	α_F	α_Θ	V_L	V_F	V_Θ
Values	1.0	0.1	1.0	0.29	0.88	20

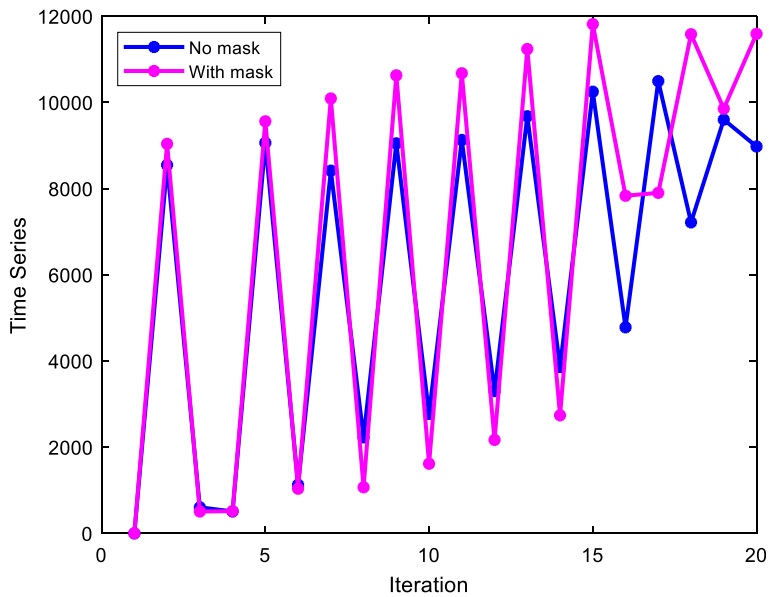


Fig. 9 Time series graph of experimented images in Fig. 3

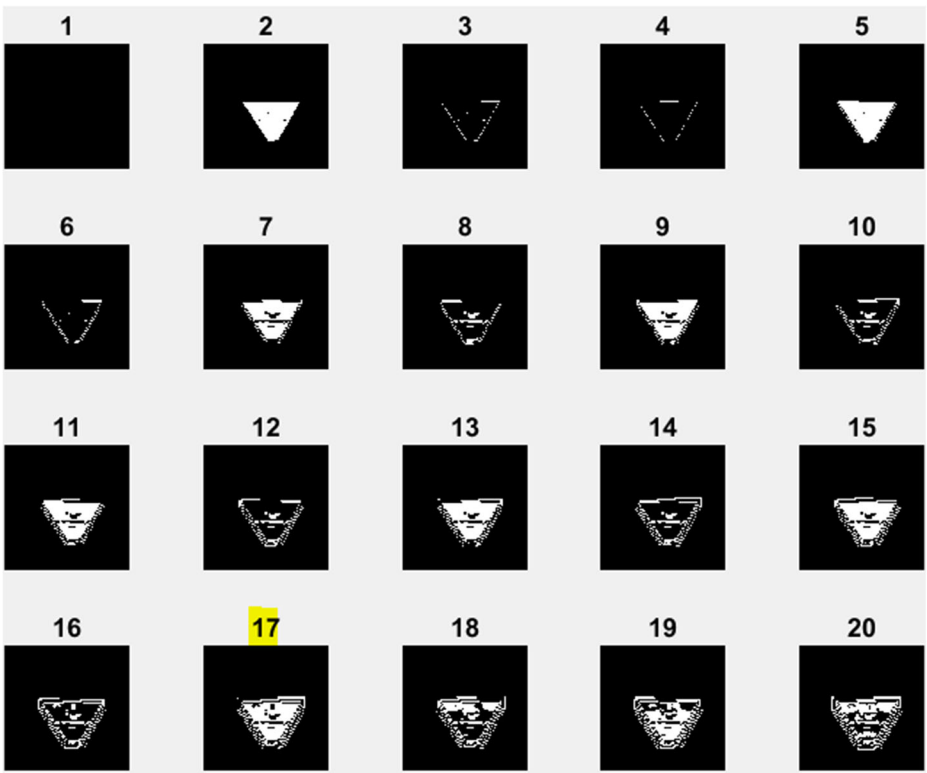


Fig. 10 PCNN output series without mask

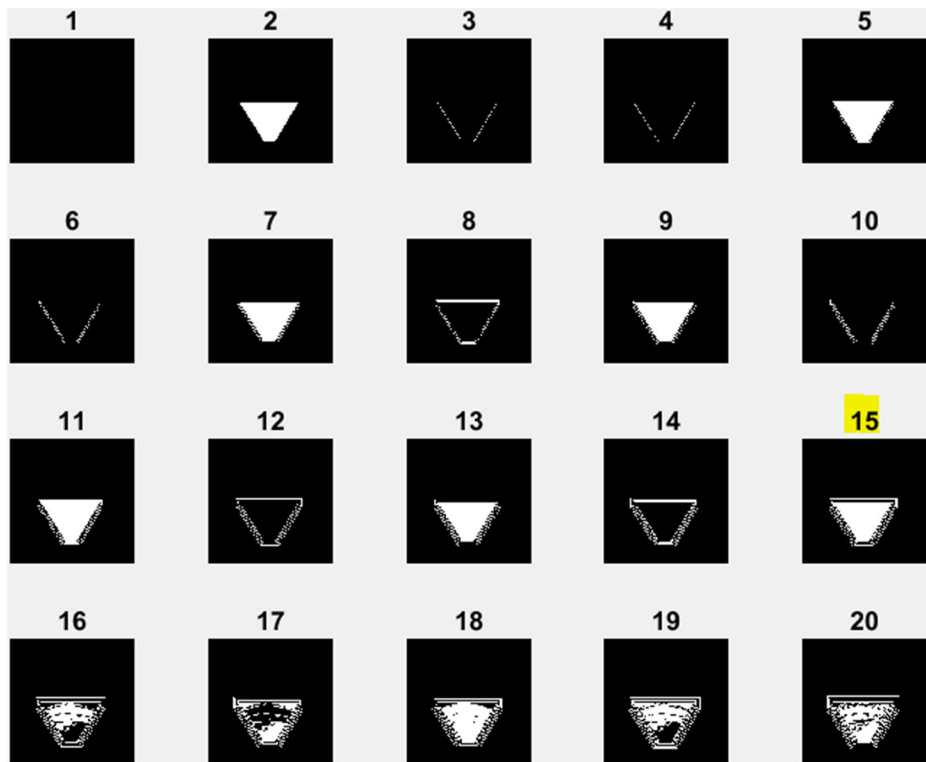


Fig. 11 PCNN output series with mask

Gray-level Uniformity

For a gray-level image (x, y) , let R_i be i th segmented region, A_i be the area of R_i (that is the number of pixels contained in region R_i) and C be the normalization factor, then the gray-level uniformity measure of $f(x, y)$ [24] is defined as in Eq. (8).

$$GU = 1 - \frac{1}{C} \sum_i \sum_{(x,y) \in R_i} \left[f(x,y) - \frac{1}{A_i} \sum_{(x,y) \in R_i} f(x,y) \right]^2 \quad (8)$$

Gray-level Contrast

In a gray-level image $f(x, y)$ consisting of the object with the average graylevel f_o and the background with the average gray-level f_b [24], a gray-level contrast measure can be calculated using Eq. (9).

$$GC = \frac{|f_o - f_b|}{f_o + f_b} \quad (9)$$

Entropy

E which is presented in Eq. (10), is the information entropy of binary image; P_1 is the probability of 1's in a binary image, and P_0 is that of 0's in the binary image [10].

$$E = -P_1 \log_2(P_1) - P_0 \log_2(P_0) \quad (10)$$

Cross-Entropy

Cross-Entropy in Eq. (12) is a measure of information difference between two probability distributions (Eq. (11)). The minimum cross entropy principle is used in image segmentation to search for the threshold which results in the least information difference between the two images before and after segmentation [24].

For two probability distributions $P = \{p_1, p_2, \dots, p_N\}$ and $Q = \{q_1, q_2, \dots, q_N\}$, their cross-entropy is defined as

$$D(P, Q) = \sum_{i=1}^N p_i \ln \frac{p_i}{q_i} + \sum_{i=1}^N q_i \ln \frac{q_i}{p_i} \quad (11)$$

In image segmentation, P and Q represent the original image and the segmented one respectively. Then

$$C(P, Q, : t) = \sum_{f=1}^t \left[f \cdot h(f) \cdot \ln \frac{f}{\mu_1(t)} + \mu_1(t) \cdot h(f) \cdot \ln \frac{\mu_1(t)}{f} \right] \\ + \sum_{f=t+1}^F \left[f \cdot h(f) \cdot \ln \frac{f}{\mu_2(t)} + \mu_2(t) \cdot h(f) \cdot \ln \frac{\mu_2(t)}{f} \right] \quad (12)$$

where, f is the gray-level, F is the maximal value of f , t is the threshold of segmentation and $h(f)$ is the histogram of original image. μ_1 and μ_2 are the average gray-level of objects and a background, respectively. They can be calculated by $h(f)$, shown as in Eqs. (13) and (14) [24].

$$\mu_1(t) = \frac{1}{\sum_{f=0}^t h(f)} \sum_{f=0}^t f \cdot h(f) \quad (13)$$

$$\mu_2(t) = \frac{1}{\sum_{f=t+1}^F h(f)} \sum_{f=t+1}^F f \cdot h(f) \quad (14)$$

We calculate GU , GC , CE , E and OM in Eq. (15) for both experimental images. The result for each iteration is shown in Tables 4 and 5.

$$OM = [GU + GC + (1-CE) + E]/4 \quad (15)$$

To select an excellent segmented image, the “Entropy must be maximum” or the “Cross Entropy must be minimum” or “Overall Merit must be maximum”. We check them from tables and if an iteration respects more than one criteria, the corresponding image is the image that we use for the next processing step. If we have three image candidates, we choose directly the one which has a maximum OM.

Four images meet the condition during our experience as per the information in Table 6. So, we select the image n°17 in Fig. 9 and image n°15 in Fig. 10. In summary, these measurements are used to find a best quality of image segmentation and edge detection.

Table 4 Experimental values of $GU/GC/CE/E/OM$ of Fig. 9 (without mask)

Iteration n°	GU	GC	CE	Entropy	OM
1	1.0000	0	0	0	0.5000
2	1.0000	0.0104	0.0000	0.4143	0.6062
3	1.0000	0.0007	0.2068	0.0521	0.4615
4	1.0000	0.0006	0.2226	0.0451	0.4558
5	1.0000	0.0111	0.0005	0.4315	0.6105
6	1.0000	0.0014	0.1481	0.0874	0.4852
7	1.0000	0.0103	0.0000	0.4098	0.6050
8	1.0000	0.0027	0.0840	0.1511	0.5175
9	1.0000	0.0110	0.0005	0.4310	0.6104
10	1.0000	0.0034	0.0643	0.1793	0.5296
11	1.0000	0.0111	0.0006	0.4336	0.6110
12	1.0000	0.0040	0.0492	0.2054	0.5401
13	1.0000	0.0118	0.0019	0.4514	0.6153
14	1.0000	0.0047	0.0367	0.2315	0.5499
15	1.0000	0.0125	0.0039	0.4692	0.6195
16	1.0000	0.0058	0.0213	0.2721	0.5642
17	1.0000	0.0128	0.0050	0.4768	0.6212
18	1.0000	0.0088	0.0019	0.3677	0.5936
19	1.0000	0.0117	0.0017	0.4489	0.6147
20	1.0000	0.0110	0.0004	0.4286	0.6098

Table 5 Experimental values of GU/GC/CE/E/OM of Fig. 10 (with mask)

Iteration n°	GU	GC	CE	Entropy	OM
1	1.0000	0.0000	0.0000	0.0000	0.5000
2	1.0000	0.0110	0.0000	0.4307	0.6104
3	1.0000	0.0006	0.2475	0.0450	0.4495
4	1.0000	0.0006	0.2460	0.0456	0.4501
5	1.0000	0.0117	0.0003	0.4476	0.6147
6	1.0000	0.0013	0.1757	0.0816	0.4768
7	1.0000	0.0123	0.0012	0.4644	0.6189
8	1.0000	0.0013	0.1726	0.0836	0.4781
9	1.0000	0.0130	0.0028	0.4809	0.6228
10	1.0000	0.0020	0.1304	0.1170	0.4971
11	1.0000	0.0130	0.0029	0.4824	0.6231
12	1.0000	0.0026	0.1006	0.1480	0.5125
13	1.0000	0.0137	0.0052	0.4992	0.6269
14	1.0000	0.0033	0.0775	0.1778	0.5259
15	1.0000	0.0144	0.0081	0.5160	0.6306
16	1.0000	0.0096	0.0018	0.3897	0.5994
17	1.0000	0.0096	0.0016	0.3921	0.6000
18	1.0000	0.0141	0.0068	0.5092	0.6291
19	1.0000	0.0120	0.0008	0.4569	0.6170
20	1.0000	0.0142	0.0069	0.5095	0.6292

3.3 Decision

In normal case, mouth and nose region should be present whether people don't wear a facemask and with mask, in maximum one region is there depending on mask's type. For standard face mask, there is no special region but directly the background. Depending on the number of regions present in ROI, the decision is made. We count the number of regions present and if it is equal or more than two, the processing flow continues with deep learning otherwise we judge that the person has a mask.

In event of deep learning, the FCNN manage the treatment. At least, five hidden layers are required and, on this paper, we fix it to five. The activation function for them is sigmoid function and all weights are initialized randomly. Concerning, output layer, the number of neurons is two (with mask and without mask). The neuron which has a high probability value determinate the image with facemask. The activation function Softmax ensures this probability format. We have ten neurons in input layer composed by $\{GU_1, GC_1, CE_1, Entropy_1, OM_1, GU_2, GC_2, CE_2, Entropy_2, OM_2\}$. The Fig. 12 helps us to understand this explanation. The percentage of image allocated for testing depends on the searcher choice but it is important to have a percentage training dataset more than testing images. The following paragraph details the training and the datasets description which are publicly available.

Table 6 Experimental choice of segmented images

Iteration n°	Number of criteria	Image type	Selected due of
2	1	With mask	Minimum CE
7	1	Without mask	Minimum CE
15	2	With mask	Maximum E and maximum OM
17	2	Without mask	Maximum E and maximum OM

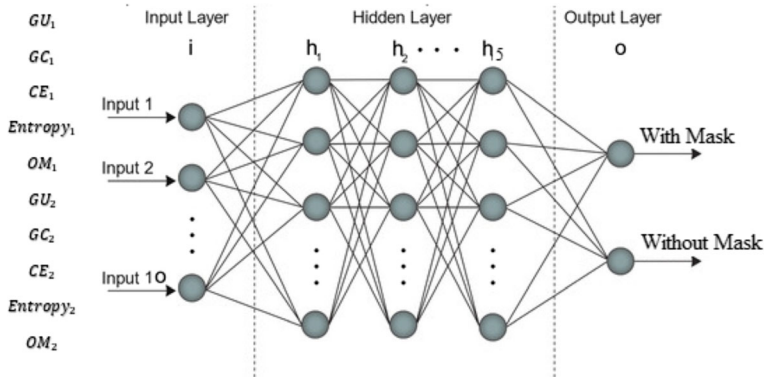


Fig. 12 FCNN architecture

3.4 Datasets training and testing

For dataset training/testing, six type of dataset which are used by authors in [2, 6, 12, 15, 20, 25] are the image database that we test our own method. The reason of this choice is to compare related works with same reference data. The Table 7 describes the content of all dataset used during this work and a sample of image is presented in Fig. 13.

The number of epochs for each dataset is limited to 80 and we remark that from 65th epoch, the result commences acceptable except for dataset5. The trend of training and validation is shown in Fig. 14.

Table 7 Dataset description

Dataset	Description
Dataset1	The first dataset is available in https://data-flair.training/ . The dataset consists of 1376 images with 690 images containing images of people wearing masks and 686 images with people without masks. It is an excellent dataset for people who want to try learning techniques of deep learning for face mask detection [7, 10].
Dataset2	The second dataset is the combination of sample from Kaggle and RMFD (Real-World Masked Face Dataset) datasets: 993 masked faces pictures and 1918 exposed faces pictures. RMFD was introduced by Wang et Al. in masked face recognition and dataset from Kaggle has 853 images belonging to three classes: with mask, without mask and mask worn incorrectly. A picture with mask and mask worn incorrectly are merged in same class of with mask [4, 8].
Dataset3	The third dataset is the combination of Medical Masks Dataset (MMD) and Face Mask Dataset (FMD). MMD was published by Mikolaj Witkowski and has 682 pictures with over 3 k medical masked faces wearing masks. The FMD dataset consists of 853 images. For this combination, we have 1415 images and the images in bad quality and redundancy are removed [5, 16].
Dataset4	The fourth dataset contains 3832 images divided into two classes: 1914 with mask and 1918 without facemask (https://drive.google.com/drive/folders/1Dm2sV8UrMd6OKzjVkw859WznhfSXFZF8)
Dataset5	The fifth dataset is AIZOO. It is fully accessible and generated by AIZOOTech [9] by integrating roughly 8000 photos from the WIDER FACE [30] and Masked FAcEs (MAFA) [13] datasets and re-annotating them to meet the face mask recognition scenario. A reasonable balance was achieved when Wider Face (50%) contributed the majority of regular faces and MAFA (50%) contributed the majority of mask-wearing faces, ensuring a reasonable balance between the two. For the purposes of testing, just a sample of 1839 photographs had been selected [26].
Dataset6	The last dataset is Moxa3K facial masking identification dataset [27]. It has 3000 photos, 2800 of which are for training and 200 for testing. In order to build the dataset, Kaggle photos and pictures from the Web were combined. A downside of the database is that it only comprises small face images that are not covered by masks.



Fig. 13 Preview datasets

4 Performance comparison and discussion

The proposed method can work without deep learning in case of simple image however a neural network training is required for complex situation. It means that the system that we propose here is a lite version.

Comparing with related works in [6, 20, 25], we have an accuracy better. We have a bit difficulty to decide when the image test is from MMD because some medical face mask has a

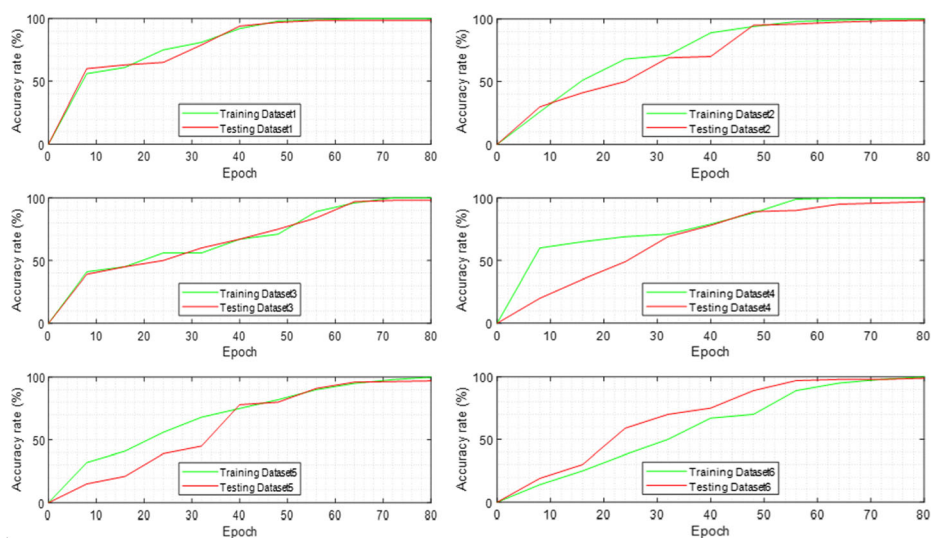


Fig. 14 Accuracy graph

Table 8 Accuracy comparison

Authors	Method	Dataset Name	Accuracy (%)	vs Proposed Method (Accuracy %)
[19] Loey et al. (2020)	A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the covid-19 pandemic	RMFD / LFW / SMFD	99.64% using SVM on RMFD 100% on LFW dataset	–
[20] Loey et al. (2020)	Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection.	Dataset3	99.49% on SMFD dataset 81.0	83.2
[6] Bhawna R. & Udit U. (2021)	Face Mask Detection Using Convolutional Neural Network (CNN)	Dataset1: 90% for training 10% for testing	99.1	Test with training images without FCNN: 99.6 Test done with testing images without FCNN: 100
[25] Riya C. & Rutva J. (2021)	Face mask detection using convolution neural network	Dataset2	99.0	99.2
[12] Gagandeep K. et al. (2021)	Face mask recognition system using CNN model	Dataset4	Good performance ^a	98.7
[15] Goyal H et al. (2022)	A real time face mask detection system using convolutional neural network	Dataset4	98.0	98.3
[2] Asghar M.Z et al. (2022)	Facemask detection using Depth wise Separable Convolution Neural Network	Dataset5: 80% for training 20% for testing	93.16	94.3
		Dataset6: 80% for training 20% for testing	92	93.8
[11] Fan X. et al. (2021)	A deep learning-based light-weight face mask detector with residual context attention and Gaussian heatmap	Dataset5	92	94.3

^a at the time of comparing the results, the author of this reference did not give an exact percentage of precision, he only mentioned that it was good

complicated form and some image has a multiple face without and with mask. Due of this reason, the accuracy is less than 84%. The accuracy comparison is presented in Table 8.

We have a best performance with dataset1 because the face mask worn has a similar model. According to the research done by Bhawna R. & Udit U. [6], they need 90% of dataset for training and 10% for testing. The accuracy measurement is done with 10% and they got 99.1%. By applying our algorithm in the same dataset without deep learning, we obtain 99.6% for base training and 100% for test which is perfect.

For dataset2, the experience is the same as with dataset1. The bit difficulty we were facing is about mask face multicolor which create multiple region. About computation time, this approach needs around 2.36 milliseconds to decide for one image however 2.5 second is the classification time of 1020 images for Asghar M.Z et al. (2022) [2] model (2.45 milliseconds per image).

Before closing this paragraph, we share in the Table 9 below the advantage/inconvenient of each method and in Table 10, we illustrate another performance indicator like precision, recall and F1 score [1, 3, 14, 21] which are defined in Eqs. (16)–(18).

$$Precision = \frac{Tp}{Tp + Fp} \quad (16)$$

$$Recall = \frac{Tp}{Tp + Fn} \quad (17)$$

Table 9 Advantage and Inconvenient of each technique

Work	Advantage	Inconvenient
[22]	Gives a precise and speedily results for facemask detection	Accuracy degradation with noising images
[19]	The SVM classifier achieved the highest accuracy possible with the least time consumed in the training process	High consume time to get highest accuracy
[20]	High precision with medical face mask detection	- Highest computation time - Weakness with simple face mask detection
[6]	Simple and speed method and able to detect easily a simple face mask	- The performance is impacted once the percentage of training dataset is reduced. - Difficulties to detect a complex face mask or medical face mask
[25]	High capabilities with live video streaming	Average success rate
[15]	High precision, excellent in real time monitoring minimum processing time	Too much parameters, High number of epochs
[2]	Number of parameters minimized High accuracy	Necessity of high capacity of hardware
Proposed Method	- Minimum computation time - High accuracy for detecting a standard face mask	- Cannot work properly if the eyes are hidden - Detection success rate reduced with in image which presents a multiple face

Table 10 Another performance metrics comparison

Method	Dataset n°	Precision	Recall	F1
Proposed Method	1	0.9979	0.9979	0.9979
	2	0.9936	0.9987	0.9961
	3	0.9826	0.8125	0.8895
	4	0.9979	0.9890	0.9934
	5	0.9634	0.9780	0.9707
	6	0.9594	0.9742	0.9667
[15]	4	0.9750	0.9750	0.9800
[12]	4	0.8900	0.8950	0.8900
[11]	5	0.9050	0.9050	0.9050
[2]	5	0.9400	0.9250	0.9300

$$F1 = \frac{2 * (Precision * Recall)}{Precision + Recall} \quad (18)$$

Where Tp , Tn , Fp , Fn are respectively True positive, True negative, False positive and False negative.

True positives are accurately predicted as being in a positive class, whereas false positives are images that were incorrectly predicted as being in a positive class. *True negatives* are accurately predicted to be in the negative class, whereas false negatives are incorrectly predicted to be in the negative class [15].

We can see that our method leads the performance especially with dataset 4.

5 Conclusion

In this research paper, we proposed a method to detect whether people wear face mask or expose face. The objective is to slow down the spread of COVID-19 and we contribute to elaborate a system to trigger an alarm if a person is coming without mask in particular public place. The technical is based on combination of PCNN and FCNN neural networks and divided into three modules: geometric module which detects the eyes and from eyes position, we deduct the mask filter to get the region of interest; feature extraction module in which the PCNN extracts the pertinent information by selecting one image well segmented and one image with edge detection and applies some criteria to select the final image candidate that we count the number of region; decision module analyses the number of region and if it is equal to one, the FCNN doesn't need to intervene otherwise the decision is taken by pretrain neural network. Due of this smart conception, the time response is minimal. Concerning the accuracy, we validate our research with different dataset such as Kaggle, AIZOO, Moxa3K, Real-World Masked Face Dataset, Medical Masks Dataset, Face Mask Dataset and the average accuracy is around 86.68% which varies from 83.2% to 100%. Due to this performance, the approach can detect the presence or absence of mask in human visage and contributes on spread reduction of COVID-19 because once people break the imposed rule, an alarm is triggered and these access to a specific place are blocked. Therefore, the presence of many security agent or military is not required and they can handle other task. Meanwhile, the suggested technique has a weakness when a several visages appear on single image and it may be improved by detecting

the upper body then eyes instead of eyes detection directly. In future work, we would like to build a system able to check whether a facemask is worn incorrectly and taking care the physical distance between people.

Data availability Data is contained within the article.

Declarations

Conflict of interests The authors declare that they have no conflict of interest.

References

1. Aldoski J (2022) Image classification accuracy assessment. https://www.researchgate.net/publication/359364984_Image_Classification_Accuracy_Assessment
2. Asghar MZ, Albogamy FR, Al-Rakhami MS, Asghar J, Rahmat MK, Alam MM, Lajis A, Nasir HM (2022) Facial mask detection using depthwise separable convolutional neural network model during COVID-19 pandemic. *Front Public Health* 10:855254. <https://doi.org/10.3389/fpubh.2022.855254>
3. Asgher U, Khalil K, Jawad M, Riaz A, Butt S, Ayaz Y, Naseer N, Nazir S (2020) Enhanced accuracy for multiclass mental workload detection using long short-term memory for brain–computer interface. *Front Neurosci* 14:584. <https://doi.org/10.3389/fnins.00584>
4. Basu A, Ali MF (2021) COVID-19 face mask recognition with advanced face cut algorithm for human safety measures. In: 2021 12th international conference on computing communication and networking technologies (ICCCNT), pp 1–5
5. Bhagyashree K, Suma V (2021). Face mask recognition using TensorFlow a deep learning framework. *Int J Eng Sci Comput* 11(05). <https://ijesc.org/upload/be315cc7134aa17060d718e91fd33896.Face%20Mask%20Recognition%20using%20TensorFlow%20a%20Deep%20Learning%20Framework.pdf>
6. Bhawna R, Udit U (2021) Face mask detection using convolutional neural network (CNN). https://easychair.org/publications/preprint_open/91Sc
7. Chacon MMI, Zimmerman SA, Rivas PP (2007) Image processing applications with a PCNN. In: Liu D, Fei S, Hou Z, Zhang H, Sun C (eds) *Advances in neural networks – ISNN 2007*. ISNN. Lecture notes in computer science, vol 4493. Springer, Berlin. https://doi.org/10.1007/978-3-540-72395-0_109
8. Chavda A, Dsouza J, Badgujar S, Damani A (2021) Multi-stage CNN architecture for face mask detection. In: 6th international conference for convergence in technology (I2CT), 2021, pp 1–8. <https://doi.org/10.1109/I2CT51068.2021.9418207>
9. Chiang D (2020) Detecting faces and determine whether people are wearing mask. Available online at: <https://github.com/AIZOOTech/FaceMaskDetection>
10. Deng H, Feng Z, Qian G, Lv X, Li H, Li G (2021) MFCosface: a masked-face recognition algorithm based on large margin cosine loss. *Appl Sci* 11(16):7310. <https://doi.org/10.3390/app11167310>
11. Fan X, Jiang M, Yan H (2021) A deep learning-based light-weight face mask detector with residual context attention and Gaussian heatmap to fight against COVID-19. *IEEE Access* 9:96964–96974. <https://doi.org/10.1109/ACCESS.2021.3095191>
12. Gagandeep K, Ritesh S, Puneet KT, Srijan KY, Prabhesh P, Rohit R, Anshu V, Manik R (2022) Face mask recognition system using CNN model. *Neurosci Inform* 2(3):100035, ISSN 2772-5286. <https://doi.org/10.1016/j.neuri.2021.100035>
13. Ge S, Li J, Ye Q, Luo Z (2017) Detecting masked faces in the wild with l1e-cnns. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, Honolulu, pp 2682–2690
14. Goutte C, Gaussier E (2005) A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. *Lect Notes Comput Sci* 3408:345–359. https://doi.org/10.1007/978-3-540-31865-1_25
15. Goyal H, Sidana K, Singh C, Jain A, Jindal S (2022) A real time face mask detection system using convolutional neural network. *Multimed Tools Appl* 81:14999–15015
16. Jakir Hussain GK et al (2021) The Face Mask Detection Technology for Image Analysis in the Covid-19 Surveillance System. *J Phys Conf Ser* 1916 012084. <https://doi.org/10.1088/1742-6596/1916/1/012084>
17. Johnson JL (1994) Pulse-coupled neural nets: translation, rotation, scale, distortion, and intensity signal invariance for images. *Appl Opt* 33:6239–6253 (J1)
18. Johnson JL, Padgett ML (1999) PCNN models and applications. *IEEE Trans Neural Netw* 10:480–498 (J2)

19. Loey M et al (2020) A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement* 167:108288
20. Loey M et al (2020) Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustain Cities Soc* 65:102600
21. Mahum R, Irtaza A, Nawaz M, Nazir T, Masood M, Mehmood A (2021) A generic framework for generation of summarized video clips using transfer learning (SumVClip), pp 1–8. <https://doi.org/10.1109/MAJICC53071.2021.9526264>
22. Militante SV, Dionisio NV. (2020). RealTime facemask recognition with alarm system using deep Learning. 2020 11th IEEE control and system graduate research colloquium (ICSGRC). IEEE
23. Rafidison MA, Ramafiarisona H (2021) Modified convolutional neural network for Ariary banknotes authentication. *Int J Innov Eng Res Technol* 8(1):62–69
24. Rafidison MA, Ramafiarisona H (2021) Automatic color image segmentation based on pulse coupled neural network. *Int J All Res Educ Sci Methods* 9(4):1031–1036 Retrieved <http://www.ijaresm.com/automatic-color-image-segmentation-based-on-pulse-coupled-neural-network>
25. Riya C, Rutva J (2021) Face mask detection using convolution neural network. *ArXiv*, abs/2106.05728
26. Rosebrock A (2020) COVID-19: face mask detector with OpenCV, Keras/TensorFlow, and deep learning. PyImageSearch. Available online at: <https://www.pyimagesearch.com/2020/05/04/covid-19-face-maskdetector-with-opencv-keras-tensorflow-and-deep-learning/>. Accessed 4 May 2020
27. Roy B, Nandy S, Ghosh D, Dutta D, Biswas P, Das T et al (2020) Moxa: A deep learning based unmanned approach for real-time monitoring of people wearing medical masks. *Trans Indian Natl Acad Eng* 5:509–518. <https://doi.org/10.1007/s41403-020-00157-z>
28. Son C, Hegde S, Smith A, Wang X, Sasangohar F (2021) Effects of COVID-19 on college students' mental health in the United States: interview survey study. *J Med Internet Res* 22(9):e21279. <https://doi.org/10.2196/21279>
29. Xiangyu D, Ma Y (2014) PCNN model analysis and its automatic parameters determination in image segmentation and edge detection. *Chin J Electron* 23:97–103
30. Yang S, Luo P, Loy CC, Tang X (2016) Wider face: a face detection benchmark. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, pp 5525–5533

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Affiliations

Maminaiaina Alphonse Rafidison¹  • Andry Harivony Rakotomihamina¹ • Sabine Harisoa Jacques Rafanantenana¹ • Rajaonarison Faniriharisoa Maxime Toky¹ • Mirado Mike Noé Raelina¹ • Hajasoa Malalatiana Ramafiarisona¹

Andry Harivony Rakotomihamina
rakotomihaminaandry@gmail.com

Sabine Harisoa Jacques Rafanantenana
rshjacques@gmail.com

Rajaonarison Faniriharisoa Maxime Toky
tfaniriharisoa@gmail.com

Mirado Mike Noé Raelina
raelina89@gmail.com

Hajasoa Malalatiana Ramafiarisona
mhramafiarisona@yahoo.fr

¹ Telecommunication-Automatic-Signal-Image-Research, Laboratory/Doctoral School in Science and Technology of Engineering and Innovation, University of Antananarivo - Ecole Supérieure Polytechnique d'Antananarivo, Ankatso, 101 Antananarivo, Madagascar