Video-Based Facial Recognition to Accurately Identify People Wearing Surgical Masks: Design and Architecture

Chanate Ratanaubol
Department of Business Administration Rajamangala University of Technology Rattanakosin, Thailand,
chanate.r@rmutr.ac.th

***Kongkiat Hirankerd**
Department of Architecture and Design, Rajamangala University of Technology Rattanakosin, Thailand,
kongkiat.hir@rmutr.ac.th

*Corresponding author E-mail: kongkiat.hir@rmutr.ac.th

Received: 14 November 2021 Accepted: 21 December 2021

ABSTRACT: *During the COVID-19 epidemic, wearing face masks is recommended or required by government in every country. Current, video-based facial recognition systems are ineffective at recognizing faces behind a surgical mask. The purposes of this study are 1) to design the architecture of a Video-Based Facial Recognition system developed to accurately identify people wearing a mask. 2) to develop a Video-Based Facial Recognition system for accurately identifying people wearing masks. This research describes a design and architecture that significantly improves face matching performance.*

Keywords: Video-Based Facial Recognition, COVID-19, surgical mask

1. Introduction

The use of facial recognition technology is prevalent in consumer applications which store images of the faces that appear in the images (Picassa; Corcoran & Costache, 2005). But obstacles to using the same technology from video sources are the disadvantages of video shooting techniques in most cases along with the inferior quality, resolution, and size of the input frames compared to still images.

Face detection and tracking elements are very important in the design of a recognition system. Features of the detection algorithm (detection rate, tolerance to speed changes, and memory requirements, etc.) will directly affect the properties of the overall perception system. It is clear, that unaltered, un-processed, unenhanced faces will likely not be recognized. Miyoung Cho and Youngsook Jeong (2017) commented that facial recognition is the most widely used biometric technology because it is both user friendly and more convenient than other biometric methods.
The detection algorithm should have a high detection rate and be resistant to variations such as changes in appearance, capture conditions and face size. The tracking algorithm should improve the tolerance for orientation, expression, and summation. All the above requirements are difficult to meet, especially for real-time situations.

Using artificial intelligence facial recognition methods, the system specified in this research can accept two kinds of inputs: still images and video. Accepting these input data types distinguishes the system from other digital footprint management systems used for office attendance in today's market. The system can also be used to remotely detect the faces of video conferencing participants when social distancing is required and there are no additional devices or technical support required for remote work unlike other, existing systems.

This research focuses on the use of digital footprint management systems for recording student participation and enrollment in school activities where most universities in Asia have difficulty assessing and documenting student participation (SEE Learning, 2019). The system we used in this research also added new input sources and methods: electronic video files, synchronized video, and social media video from Facebook.

The system in use before the prevention and mitigation efforts to limit the spread of COVID-19, included the face detection algorithm and techniques in this research using the OpenCV library and developed in Python Language, with a face detection accuracy of 93% (Ratanaubol and others,2021). This could also use recorded videos and photos from social media (Facebook).

2. Face Detection using OpenCV

This functionality performs face detection for each frame in a video. So, when it comes to detecting a face in a still image and detecting a face in the frames in a real-time video stream, there is not much difference between them. We will be using the Haar Cascade algorithm, also known as Voila-Jones algorithm to detect faces. It is basically a machine learning object detection algorithm which is used to identify objects in an image or video.

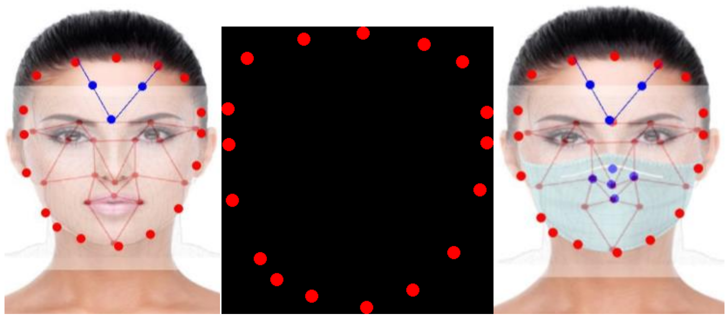

Figure 1. face and mask detector

After detecting a face from the webcam video stream, we are going to save the frames containing the face. Later we will pass these frames(images) to our mask detector classifier to find out if the person is wearing a mask or not.

To train a face mask detector, we need to break our process into two distinct phases, each with its own respective sub-steps (as shown by Figure 1 above):

Training: Here we'll focus on loading our face mask detection dataset from data storage, training a model on this dataset, and then serializing the face mask detector to data storage,
Deployment: Once the face mask detector is trained, we can then move on to loading the mask detector, performing face detection, and then classifying each face as "with mask" or "without mask".

We will review each of these phases and associated subsets in detail in the remainder of this tutorial, but in the meantime, let's examine the dataset we'll be using to train our COVID-19 face mask detector.

Facial landmarks allow us to automatically infer the location of facial structures, including Eyes, Eyebrows, Nose, Mouth and Jawline

To use facial landmarks to build a dataset of faces wearing face masks, we need to first start with an image of a person not wearing a face mask. Implementing our COVID-19 face mask detector for images with OpenCV: Now that our face mask detector is trained, let's learn how we can: 1) Load an input image from disk 2) Detect faces in the image 3) Apply our face mask detector to classify the face as either "with mask" or "without mask". 4) We apply facial landmarks, allowing us to localize the eyes, nose, mouth, etc.

In this paper the authors have worked on masked face detection from a video source. The masked person is detected in this presented approach and 4 main steps are performed for the detection: these are; 1) estimation of distance between camera and person; 2) detection of eye line; 3) detection of parts of face; and, detection of eyes. This process and related algorithms have been analyzed using various video surveillance systems and achieved excellent accuracy in performance.

3. Data Flow Diagram

The proposed approach for the mask detection is stated here and the various implementation steps involved in this approach. The flowchart for this is presented that states the overall flow of the process. The model starts with the loading of the dataset for mask detection and the pre-processing of data is done with the help of OpenCV.

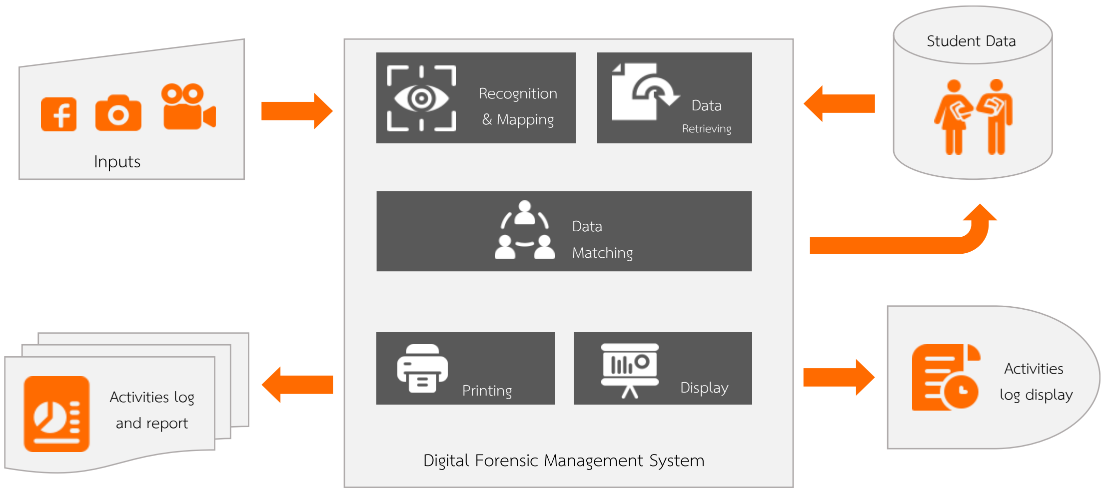

Figure 2. Digital Forensic Management System Data Flow Diagram

Data Flow Diagram. This shows system inputs from social media, still camera and video as well as student data which includes prerecorded stored images or video formatted for comparative analysis. System outputs are for reporting and logging. Management System processing of inputs show stages of image processing, data retrieval, matching as well as formatting for output for printing or digital display.

Facial Recognition and mapping are performed using a face shape algorithm, an accurate face landmark detector and face landmark identifier algorithm utilizing the concept of a deep, modular twist and interlocking blocks towards a total of 68 face landmarks. Landmarks and contours of the face processed by facial recognition and shape algorithms from the sources accurately identify a person wearing a mask.

Similar face mapping, contour and landmark facial recognition data previously stored with student identifying information is retrieved and compared to correctly identify a face in a video as a particular person and specific personal information is matched to the identified and recognized face.

A completed recognized face with associated desired personal information is then compiled and formatted for either a printed report or for output to a digital video display.

4. Data Relationship

The correlation of the data in this research consisted of a large component, namely student information, information of each type of activity in the university, Data of student faces with masks and without masks, and data that is an indicator of the volunteer spirit that arises in students arising from participating in university activities. A diagram showing the relationship of the data in the tables to each other is shown in Figure 3.

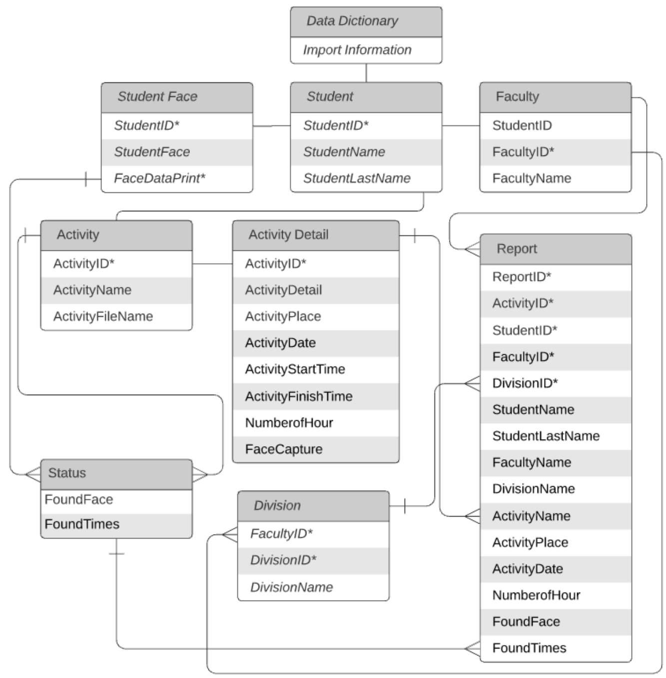

Figure 3. Correlation of the data Diagram

The diagram shows the **"**one**:**one**"** and **"**one**:** many**"** data relationships by showing the linking of the data in each table

5. Image Base Approach

In general, Image-based methods rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face and non-face images (Kumar and others, 2019). The learned characteristics are in the form of distribution models or discriminant functions that is consequently used for face detection.

6. System diagram

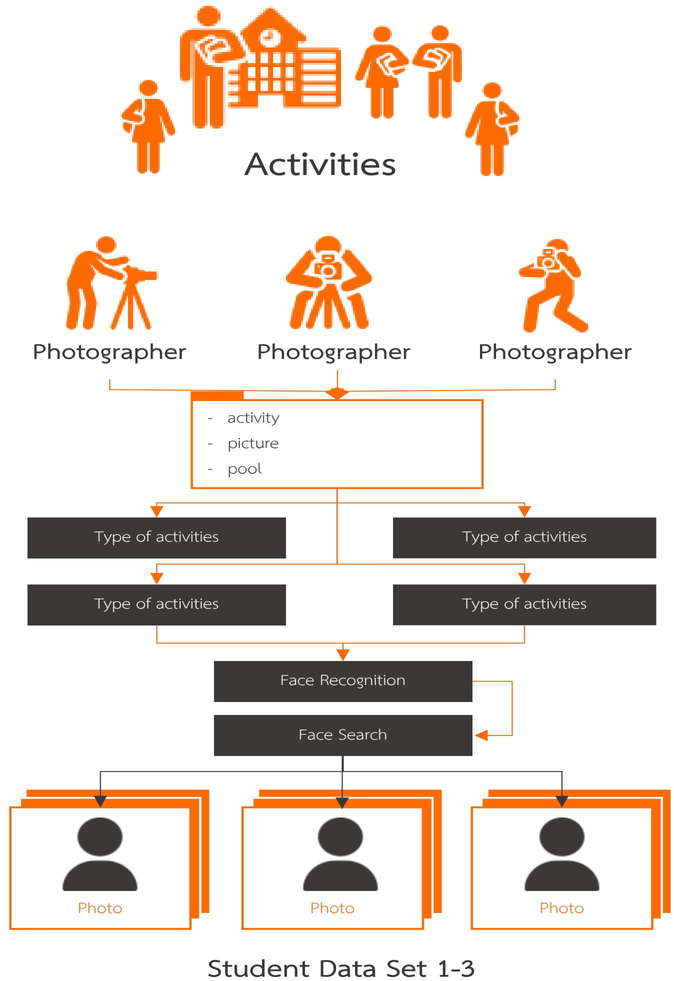

Figure 4. System diagram

System diagram shows logical sequence of operations of a student activity participation recording process with photographers creating images of students engaging in various activities, followed by facial recognition elementary processing then comparison to stored image and corresponding identifying information to product a final formatted output of an accurate log of student participation.

After facial recognition algorithms are applied to source the processed facial identifying data is compared to a database of facial images stored in a file format for face comparison. The data is matched and identified as a particular student through a face search procedure to produce a desired data set.

7. Implementation

The implementation of the proposed approach is presented here, and the various steps followed, and the data used for implementation is stated. Also, the selection of each library, framework, algorithm, and the dataset are stated in this section. Data at Source Here for out experiment we have downloaded the raw images by the OpenCV. "Mask" and "NoMask" are the tags given to the images at the start itself. The size and resolutions of the taken images are varied because they are taken from various devices having different configurations. Data Preprocessing Here the preprocessing steps have been stated which are performed for making the images noise free and make them clear for detection, and this preprocessed image could be given to the neural networks model as input. 1.Resizing the input image (256 x 256) 2. Applying the color filtering (RGB) over the channels 3. Scaling / Normalizing images 4. Center cropping the 5. Finally Converting them into tensors.

8. Result

Table 1. Example of table

Activities	Video recording	
	Face shape algorithm	
	Set1	**Set2**
1	81	92
2	89	94
3	94	98
4	94	95
5	96	97
6	97	99
7	89	91
8	95	89
9	97	96
10	90	94
Avg	92.2	94.5
T-test (1-Tails)	0.047893558 Sig*	

Note. $^{*}p < .05$.

There was a significant increase in the volume of face recognitions in the second set of photos taking using new system design and architectures compared to the old system design, $t(9) = 0.1711$, $p < .05$.

9. Conclusion

COVID-19 pandemic has come with various challenges to the world and the spread of this virus should be controlled as this virus has affected more than one crore peoples all over the world and the counting is still going on. One of the major precautions is to wear mask for the stopping of the spread of the respiratory droplets of infected peoples through cough or sneeze as well the healthy people should be covered with mask. So here we have presented an approach that uses deep learning algorithm, and the framework is used for implementation along with the OPENCV (python). The results state that the proposed model can detect the people with or without masks from the images as well from the video streams. The accuracy for the training and validation set is compared and found to be 0.048 Significantly increase.

10. Future work

Future work It is planned to improve Face Mask Detection to identify people who aren't wearing a mask and apply it on embedded devices like raspberry pi or jetson Nano. An alarm system can also be implemented to make a sound when someone without a mask enters the area or laser detectors that mark the individual who is not wearing a mask.

References

Corcoran, P., & Costache, G. (2005). Automated sorting of consumer image collections using face and peripheral region image classifiers. *IEEE Transactions on Consumer Electronics*, *51*(3), 747-754.

Cho, M. Y., & Jeong, Y. S. (2017, February). Human gesture recognition performance evaluation for service robots. In *2017 19th International Conference on Advanced Communication Technology (ICACT)* (pp. 847-851). IEEE.

SEE Learning. (2019). Social, Emotional, and Ethical Learning - An initiative for educating heart and minds. Atlanta: Emory University, Center for Contemplative Science and Compassion-based Ethics. https://seelearning.emory.edu/

Ratanaubol, C., Wannapiroon, P., & Nilsook, P. (2021, June). Video-Based Facial Recognition Develop for Accurately Identify People Wearing Surgical Masks. In *2021 3rd International Conference on Computer Communication and the Internet (ICCCI)* (pp. 19-22). IEEE.

Kumar, A., Kaur, A., & Kumar, M. (2019). Face detection techniques: a review. *Artificial Intelligence Review*, *52*(2), 927-948.