How to Cite:
Meena, S., Vats, K., & Kumar, D. . . (2022). Review of factors affecting facial recognition algorithms performance. *International Journal of Health Sciences*, *6*(S3), 9528–9541. https://doi.org/10.53730/ijhs.v6nS3.8253

Review of factors affecting facial recognition algorithms performance

Shweta Meena
Delhi Technological University
Email: shwetameena@dtu.ac.in

Kanishk Vats
Delhi Technological University
Email: 210600.kv@gmail.com

Deepak Kumar
Delhi Technological University
*Corresponding author email: deepak9013216128@gmail.com

***Abstract*---**The face is a significant part of the human body, recognizing people in large groups of individuals. Thus, because of its uniqueness and universality, it has turned into the most generally utilized and acknowledged biometric technique. Many algorithms have been used by various researchers for face detection and recognition. Research, innovation progression, and applications consolidating face recognition in the last twenty years have grown massively. In this paper, some essential existing approaches which are adjusted with dealing with the issues of face recognition have been presented close by their Face recognition accuracy and the variables capable of debasing the performance of the review. In the first section, various factors that decrease facial detection and recognition accuracy have been researched like posture variety, illumination, aging, facial expressions, etc. While in the second section of the paper, various methods have been examined that attempt to relieve the impact of discussed factors. Various algorithms give various exhibitions in various conditions like enlightenment, noise, posture, and mask change. All the previously mentioned methods are represented briefly to give an overall idea. The motive of the paper is to carry all the various methods to a similar spot and simplify it to review the paper.

***Keywords*---**ANN, facial recognition, facial detection and recognition, gabor wavelet, LDA, PCA.

Introduction

Facial Detection and Recognition (FDR) is one of the huge check structures reliant upon biometrics which can be used for the confirmation cycle just as reconnaissance. With the speedy extension in frauds step by step, Facial Recognition (FR) is transforming into a basic framework for us. Different uses of a capable FR framework are criminology, identification of hoodlums, extortion prohibition, surveillance, and so forth. A great deal of exploration has been done at every level, yet even after all that, no system is developed which can work in real-time under an unconstrained environment. FR has consistently been an exceptionally complicated task. Its genuine test lies in showing a computerized structure that matches the human capacity to see faces. Thus, a customized programmed electronic system with somewhat greater recognition accuracy, and quick handling is required. FDR process has steps as shown in Fig. 1:

- Face detection: To perceive the appearances in a picture or video using important features on the face like eyes, nose, mouth, and so on
- Feature extraction: Standardization, arrangement of the faces for stronger recognition exactitude.
- Face recognition: To recognize a specific individual in a picture or video by comparing the data set.

This paper is divided into five major sections. The first section introduced us to FDR. In second section, various factors that reduce FR accuracy have been discussed. The third section briefly describes some FDR methods and their importance. Section four gives a brief review of some algorithms used by researchers in recent years. Finally, the last section contains the conclusions we derived from this.

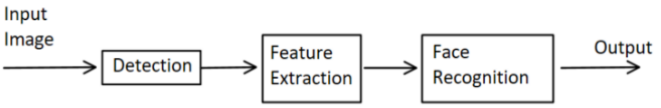

Fig 1. Steps of FDR

Factor degrading the FR performance

FR from pictures and recordings is a formidable task. Researchers have done a great deal of work trying to accomplish 100% precision, but it is not feasible due to several factors. It has been found that the components that spoil the precision of the FR systems are: noise, occlusion, illumination, pose variation, low resolution, aging, expressions, plastic surgery, and database. These elements can be requested into two classes: natural and outward factors. Characteristic parts incorporate the normal condition of the human face like the look, maturing, plastic surgery, and so on, impacting the framework, while extraneous elements include those which affect the face from the outside like noise, occlusion, illumination, low resolution, and pose variation. However, databases are also a major factor in deciding face recognition accuracy.

Noise

Computerized pictures are prone to various types of noise. This noise prompts low accuracy of FDR. Noise is present in pictures in various ways that are subject to picture creation. Dealing with these before proceeding further is a basic step in these common FDR frameworks [3]. Noise is visualized in Fig. 2.

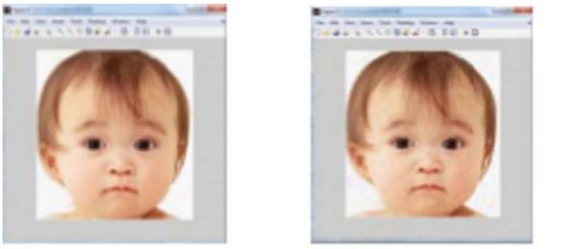

Fig. 2. Example of image having noise [3]

Fig. 3. Simulation of partial occlusion [23]

Fig. 4. Effect of illumination in face image [9]

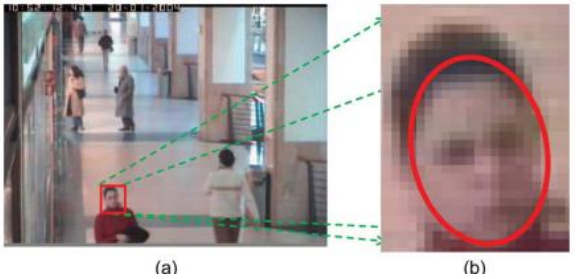

Fig. 5. A-frame from the CCTV video. (a) frame, (b) captured face [2]

Occlusion

One of the major challenges of FR is an incomplete face as it would be hard for even humans to recognize a person if a distinctive part of their face is hidden. For example, long hair and sunglasses can hide ears and eyes, masks or bushy beards can hide the lower half of the face, and so on. Partial occlusion is shown in Fig. 3. The accuracy of the FDR system is degraded by these factors. Various methodologies have been researched over the years to overcome these issues [1].

Illumination

The variations in illumination interfere with the identification of important parts of the face, thereby reducing the probability of FR. The main cause behind these varieties could be unbalanced light, shadow, contrast, darkness, etc. as shown in Fig. 4. Various methodologies related to illumination are being discussed in [4-9].

Low Resolution

The surveillance cameras are strict to look out for suspicious activities, hence they are not high quality enough to record high-quality videos leading to low-resolution faces as shown in Fig. 5. Comparing the low-resolution query picture with the high-resolution database picture is a difficult task. This degrades the performance of the FR system. Distinctive approaches are being researched to tackle this problem [1,2].

Pose Variation

Distinctive posture conveyance is another important issue of the FR system. Front-facing face recreation is needed to compare the query face with the database face. This is needed as the database contains the front-facing face. Comparing it with that different pose will yield unusable results. Various approaches were researched by the specialists to change the non-front-facing face over to the front-facing face to increase FR accuracy [10]. How different poses depreciate the performance of the FDR algorithm is being discussed in the proposed approaches [6, 11]. Pose variations are shown in Fig. 6.

Aging

The face is skin tissues and facial muscles. On Aging, these muscles contract or expand, leading to a change in facial highlights of an individual, as shown in Fig. 7. Although these are basic changes, they are still a factor that lowers the performance of FR systems [8, 12].

Expressions

We use facial expressions to express our feelings but it changes the shape of the human face due to the contraction of facial muscles. A small change like that can easily create ambiguity for the FR framework. Continuous research is being done for FR by considering facial expressions [4, 8]. Facial Expressions are shown in Fig. 8.

Plastic Surgery

One of the major concerning factors that affect FR accuracy. A small change will lead to tightening of facial muscles, while a major change may lead to an entirely new face, of which the database has no record. Many criminals or undercover agents go for plastic surgery to bury their identities. An example is shown in Fig. 9.

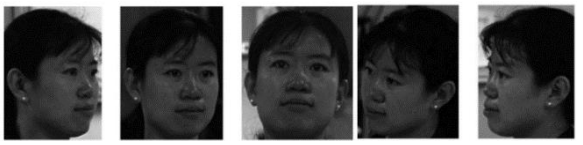

Fig. 6. Pose variations [24]

Fig. 7. Aging variations [12]

Fig. 8. Example of face having different expressions

Fig. 9. Example of Surgery effects

Databases

It is also a major factor in deciding FR accuracy. Databases used by researchers have different image quantities containing different combinations of the above factors, which leads to different accuracy results as their parameters are adjusted to them. Researchers only compare their results with those researches done on the same database, hence different results, since not all researchers take care of all the above factors, as it's not possible with current means. Some of the datasets used by researchers are CASIA WebFace, Colour FERET Database, LFWcrop Database, SCface - Surveillance Cameras Face Database, The Yale Face Database, and PIE Database, CMU.

A Brief Explanation of FR Methods

It is not possible to recreate an experiment since parameters as well as algorithm

details are not made available. That is why researchers compare the database used and the results they obtained with the previous results. A typical and well-designed protocol can quickly benchmark various methods. The FR methods used by researchers can be divided into three groups [26]:

- Appearance-based Methods
- Feature-based Matching Methods
- Hybrid Methods

Appearance-Based Methods

These methods are encompassing techniques where the complete face is compared with the display face. The benefits of this methodology are:

1. They focus on only restricted regions or central focal points without annihilating any of the data in pictures.
2. Information on faces and geometry isn't required.
3. Recognition is straightforward as compared to other methodologies.
4. Quick and simple.

Yet, they are very inclined to facial varieties like illumination, posture, and articulations. Likewise, acknowledgment accuracy is low.

Feature-Based Matching Methods

These methods compare the test face with the gallery face by utilizing neighborhood highlights of the face like eyes, mouth, nose, etc. The area highlights have been taken care of into the essential classifier to perceive the face. In this method, recognition accuracy is better than previously discussed approaches. In any case, these techniques have exceptionally high memory usage and a lot of marked and clean datasets for preparing are required.

Hybrid Methods

It's a mixture of both methods. Both methods are merged to improve the recognition rate. However, hybrid methods are complicated and hard to implement. Research reviews of the above methods are listed in Table I

Table 1
Recent progress made on FDR

	Authors	Published in year	Dataset used	Algorithm used	Recognition accuracy	Review
Appearance-based methods	Mhadgut [31]	2021	Custom Dataset	YOLOv3	95.83%	(1) model detects the masked and non-masked face with a mean average precision of 95%

						(2) considered occlusion and pose variation (3) different YOLO models considered with the best results obtained in YOLOv3
	Ejaz et al. [28]	2019	ORL combined with a custom dataset	PCA, Viola-Jones	72-95%	(1) poor recognition rate on the masked face as the number of features obtained is less. (2) face without a mask gives a better recognition rate
	Khan et al. [19]	2018	NRC-IIT facial database and real-time video stream	PCA	69–86%	(1) it does not work well in an unconstrained environment (2) results are not good for low-resolution videos and pose variation
	Abdullah et al. [5]	2017	Real-time video stream	PCA	80%	(1) not useful for video data sets (2) good for front view faces (3) other factors are not considered (4) Performance evaluation of the algorithm is not done using an actual dataset (5) recognition accuracy is 80%
	Gao and Jong [11]	2015	FERET	Lucas Kanade, SIFT, two-phase alignment error	99.5%, ±15° 26.0%, ±60°	(1) this approach can deal with up to ±60° pose variation (2) beyond 40°, recognition accuracy drops drastically (3) time complexity is very high compared to other algorithms

TABLE 1 (continued)

	Authors	Published in year	Dataset used	Algorithm used	Recognition accuracy	Review
Feature-based Methods	Gupta et al. [32]	2021	ORL	SIFT, SURF, Decision Tree, Random Forest	95.5%	(1) different combinations of SIFT and SURF are used with 64 and 128 feature vector dimensions (2) test conducted on 5 datasets each containing different combinations of factors affecting FR (3) both classifiers were used to validate the performance of the proposed system. (4) system was computationally efficient
			FACE94		87.3%	
			M2VTS		98.8%	
			FERET		76.8%	
			YALE2B		99.7%	
	Fredj et al. [29]	2020	LFW, YTF, Training with CASIA WebFace	CNN based on GoogLeNet inception-v1	99.2%	(1) considered occlusion, illumination, pose variation, and noisy images (2) test performed on two datasets
					96.63%	
	Luo et al. [4]	2018	Training with CASIA WebFace	Unsupervised transfer learning, maximum mean discrepancy	98.70–98.93%	(1) Doesn't work well if the alignment of images compared is different (2) only image datasets are considered (3) It's Focus on variable illumination, facial expression, and age but does not consider poses variation and occlusion
			GBU, FERET			
	Fu et al. [1]	2017	CASIA WebFace	Guided CNN, loss function	91.9–97.1%	(1) A lot of training is required as the system is trained for every low resolution of each image (2) only partial occlusion and different resolution is considered
			LFW			
	Huang et al. [20]	2015	COX face	Point-to-set correlation	50.96–53.26%	(1) Aging, occlusion, and plastic surgery are not considered in the dataset

				learning (PSCL)		(2) No effective results were obtained

	Authors	Published in year	Dataset used	Algorithm used	Recognition accuracy	Review
Hybrid Methods	Taskiran et al. [33]	2021	UvA-NEMO	VGGFace, VGGFace2, ArcFace	72.26 – 100%	(1) considered illumination, pose variation, noise, low-resolution, expressions (smile only) (2) different hybrid methods are considered, having promising results for each of the factors on both datasets (3) intensity of factors is varied to obtain a comprehensive result to find to which extent the said model can give acceptable accuracy
			FEEDTUM		71.76 – 100%	
	Aung et al. [30]	2021	FDDB	YOLOv2, VGG16 network model	93%	(1) considered pose variation, illumination, and skin color. (2) train on the high-resolution image. (3) has high computational complexity.
	Banerjee et al. [10]	2018	PaSC videos and CW images, CMU multi-PIE dataset	Supervised learning, generic 3D model, and Viola-Jones	88.45–97.28%	(1) Major failure beyond 40° pose variation (2) focused on pose and illumination only (3) Supervised Learning needs cleaned and labeled data
	Fathima et al. [21]	2015	AT&T, MIT-India, and Faces94 datasets	Linear discriminant analysis and Gabor wavelet	88–94.02%	(1) Not considered pose variation
	Lei et al. [23]	2009	CMU-MIT face	Modest AdaBoost,	95.20%	(1) a very huge number of images are needed to train

			database	improved independent component analysis, Hausdorff distance		the system to be capable of recognition

A Brief Explanation of FR Algorithms

There are many techniques used by researchers over the years for FDR. Some of them are machine learning-based approaches, such as Principal Component Analysis and Artificial Neural Network, while there are also feature-based approaches like Gabor Wavelet. This section briefly describes some of these techniques.

Principal Component Analysis (PCA)

PCA was a technique given by Turk and Pentland in 1991 [13]. From the given dataset of pictures, PCA identifies the important parts, which are referred to as eigenfaces by eliminating the redundancy. Then, all the eigenfaces obtained from the database are combined to form an average face to contain the significant features obtained, like eyes, nose, mouth, etc. This allows us to represent all images as a linear combination of eigenfaces, which is achieved by adding the mean face obtained to them. After which, images are projected on the eigenspace and the Euclidean distance between the target face and the eigenspace image is calculated. If the calculated distance is below a certain threshold, the target image is classified as belonging to a certain class k. PCA performance is affected by orientation, lightning, and the size of the image. It can be used in real-time under constrained conditions, which allows it to achieve a maximum of 96% accuracy. In this method, orientation, size, and lightning in the image degrade the accuracy. This framework can be utilized as a real-time application in a confined environment. The steps of the algorithm are explained below:

- The image from the database is taken and its eigenface is calculated.
- A set of weights is obtained, each associated with the image projected in the eigenspace
- Then, the Euclidean distance is calculated between the target image and the projected image to look for the face.
- If the distance is below a certain threshold of a class, it considers the face as recognized, otherwise, it is declared unknown.

Artificial Neural Network (ANN)

ANN is one of the most go-to methods in this field. [14] entails an approach for ANN in which the images are first processed with PCA, then are fed to ANN. It gives good accuracy. Similarly in another approach [15], to deal with noisy images, ANN is trained in three stages, first with the ideal images, then with noisy images in the second stage, and at last again with ideal images, which allows it to achieve high precision in ideal images and average in noisy images. In [16], CNN-

based FR algorithms were implemented in Google's open-source TensorFlow software. Although ANN gives high accuracy, it gets very complex in the training stage and can take too much time. It is also difficult to decide the number of neurons in the secret layer as there is no way to arrive at a particular number.

Gabor wavelet

It is a method that stimulates human eyes due to having similar characteristics and orientation representation. Its representation of the image is obtained through the convolutions of the target image and Gabor filters. After image-processing, n weight points are obtained corresponding to wavelets placed along the region of importance in a portable grey map. These weights along with their set of frequencies are the required features that are used for the training and testing of neural networks. It is a feed-forward neural network that uses backpropagation. This method remains unaffected due to glasses, facial hair, pose, and illumination.

Linear Discriminant Analysis (LDA)

LDA is an algorithm that starts by reducing the dimensionality of pictures. It divides the pictures into various classes where each picture is represented as pixel vectors. By doing this, it finds a linear combination of images, which is later used to reduce dimensionality. After which, Euclidean distance-based classification is used.

3D model-based FR

In this method, the system takes images and measures the facial curves on a micrometer level. Then, a face template is created where each feature has a unique number. The same was done with database images before, allowing the system to directly compare their codes. Database images don't need to be in 3D as 2.5D scans can also be compared with 3D scans. The z-value is at max one for each (x, y) pair, making comparison possible. Constraints like illumination, lightening, and pose can be ignored while recognizing faces in this method

Conclusion

We have made a great deal of progress in the last two decades in the FDR region, but the problems faced have also increased, due to which the accuracy and results obtained so far are not acceptable. The primary factors which affect the results are pose variation, brightening, lightening, and environmental elements which cannot be controlled. Many face datasets have a variable composition that can be used in evaluating FR algorithms. Because of the distinction between these data sets, face datasets selection is a significant thought for research groups working at FR. They have to consider factors like dataset size, image resolution, and quality of images. Other factors are which of the above-mentioned factors are they considering and what factors that particular dataset covers. This is also the reason due to which researchers cannot actually compare their work, making it difficult to determine whose approach made more progress in FDR.

The algorithms can be classified into three categories, depending upon their approach to the image. These are appearance-based, feature-based, and the combination of two known as hybrid methods. Each category has its pros and cons depending upon the dataset used and the application of the area. Appearance-based approaches give faulty results for pose variation and low-resolution video, feature-based take too much memory and require clean and labeled datasets, while the combination of two, which is a hybrid, is very difficult to implement. The majority of the technique doesn't give satisfactory results for non-front facing faces, i.e., faces with various posture dissemination. Furthermore, the factors like plastic surgery, aging, and occlusion also influence the performance of the FR system. After this review, we have learned that the current results obtained are still not sufficient to build a dependable FR system that can work in unconstrained environment without being affected by above factors.

References

1. T. Fu, W. Chiu and Y. F. Wang, "Learning guided convolutional neural networks for cross-resolution face recognition," 2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP), 2017, pp. 1-5
2. W. W. W. Zou and P. C. Yuen, "Very Low Resolution Face Recognition Problem," in IEEE Transactions on Image Processing, vol. 21, no. 1, pp. 327-340, Jan. 2012
3. Tin, H.H.K.: Removal of noise by median filtering in image processing. In: 6th Parallel and Soft Computing (PSC 2011) Removal (2011)
4. Z. Luo, J. Hu, W. Deng and H. Shen, "Deep Unsupervised Domain Adaptation for Face Recognition," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 2018, pp. 453-457
5. N. A. Abdullah, M. J. Saidi, N. H. A. Rahman, C. C. Wen, and I. R. A. Hamid, "Face recognition for criminal identification: An implementation of principal component analysis for face recognition", AIP Conference Proceedings, vol. 1891, no. 1, pp. 020002, 2017.
6. C. Ding, C. Xu and D. Tao, "Multi-Task Pose-Invariant Face Recognition," in IEEE Transactions on Image Processing, vol. 24, no. 3, pp. 980-993, March 2015
7. P. Wang, W. Lin, K. Chao and C. Lo, "A Face-Recognition Approach Using Deep Reinforcement Learning Approach for User Authentication," 2017 IEEE 14th International Conference on e-Business Engineering (ICEBE), 2017, pp. 183-188
8. T. Ahonen, A. Hadid and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037-2041, Dec. 2006
9. S. Z. Li, R. Chu, S. Liao and L. Zhang, "Illumination Invariant Face Recognition Using Near-Infrared Images," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 4, pp. 627-639, April 2007
10. S. Banerjee et al., "To Frontalize or Not to Frontalize: Do We Really Need Elaborate Pre-processing to Improve Face Recognition?," 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 20-29

11. Y. Gao and H. J. Lee, 'Cross-pose face recognition based on multiple virtual views and alignment error', Pattern Recognition Letters, vol. 65, pp. 170–176, 2015.
12. Z. Li, U. Park and A. K. Jain, "A Discriminative Model for Age Invariant Face Recognition," in IEEE Transactions on Information Forensics and Security, vol. 6, no. 3, pp. 1028-1037, Sept. 2011
13. M. Turk and A. Pentland, "Eigenfaces for Recognition," in Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86
14. N. Jamil, S. Lqbal and N. Iqbal, "Face recognition using neural networks," Proceedings. IEEE International Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century., 2001, pp. 277-281
15. A. Reda and B. Aoued, "Artificial neural network-based face recognition," First International Symposium on Control, Communications and Signal Processing, 2004., 2004, pp. 439-442
16. L. Yuan, Z. Qu, Y. Zhao, H. Zhang and Q. Nian, "A convolutional neural network based on TensorFlow for face recognition," 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2017, pp. 525-529
17. B. S. Manjunath, R. Chellappa and C. von der Malsburg, "A feature based approach to face recognition," Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1992, pp. 373-378
18. P. Campadelli, R. Lanzarotti and C. Savazzi, "A feature-based face recognition system," 12th International Conference on Image Analysis and Processing, 2003.Proceedings., 2003, pp. 68-73
19. A. Khan et al., "Forensic Video Analysis: Passive Tracking System for Automated Person of Interest (POI) Localization," in IEEE Access, vol. 6, pp. 43392-43403, 2018
20. Z. Huang et al., "A Benchmark and Comparative Study of Video-Based Face Recognition on COX Face Database," in IEEE Transactions on Image Processing, vol. 24, no. 12, pp. 5967-5981, Dec. 2015
21. A. A. Fathima, S. Ajitha, V. Vaidehi, M. Hemalatha, R. Karthigaiveni and R. Kumar, "Hybrid approach for face recognition combining Gabor Wavelet and Linear Discriminant Analysis," 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS), 2015, pp. 220-225
22. Z. Lei, C. Wang, Q. Wang and Y. Huang, "Real-Time Face Detection and Recognition for Video Surveillance Applications," 2009 WRI World Congress on Computer Science and Information Engineering, 2009, pp. 168-172
23. Z. Lei, C. Wang, Q. Wang and Y. Huang, "Real-Time Face Detection and Recognition for Video Surveillance Applications," 2009 WRI World Congress on Computer Science and Information Engineering, 2009, pp. 168-172
24. R. Singh, M. Vatsa, A. Ross and A. Noore, "A Mosaicing Scheme for Pose-Invariant Face Recognition," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 37, no. 5, pp. 1212-1225, Oct. 2007
25. S. Anwarul and S. Dahiya, "A Comprehensive Review on Face Recognition Methods and Factors Affecting Facial Recognition Accuracy", Jan. 2020, pp. 495–514.
26. P. Borisagar, S. Jani, Y. Agrawal and R. Parekh, "An Efficient and Compact Review of Face Recognition Techniques," 2020 IEEE International Students'

Conference on Electrical,Electronics and Computer Science (SCEECS), 2020, pp. 1-5
27. G. Castaneda and T. M. Khoshgoftaar, "A Review of Performance Evaluation on 2D Face Databases," 2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService), 2017, pp. 218
28. M. S. Ejaz, M. R. Islam, M. Sifatullah and A. Sarker, "Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-5
29. H. Fredj, S. Bouguezzi, and C. Souani, "Face recognition in unconstrained environment with CNN", The Visual Computer, vol. 37, Feb. 2021.
30. H. Aung, A. V. Bobkov and N. L. Tun, "Face Detection in Real Time Live Video Using Yolo Algorithm Based on Vgg16 Convolutional Neural Network," 2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 2021, pp. 697-702
31. S. Mhadgut, "Masked Face Detection and Recognition System in Real Time using YOLOv3 to combat COVID-19," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 1-7
32. S. Gupta, K. Thakur, and M. Kumar, "2D-human face recognition using SIFT and SURF descriptors of face's feature regions", The Visual Computer, vol. 37, no. 3, pp. 447–456, Mar. 2021.
33. M. Taskiran, N. Kahraman, and C. Eroglu Erdem, "Hybrid face recognition under adverse conditions using appearance-based and dynamic features of smile expression", IET Biometrics, vol. 10, no. 1, pp. 99–115, 2021