International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Review Article

A Review of Recent Advances Methodologies for Face Detection

Laxmi Narayan Soni* and Dr. Akhilesh A Waoo

Department of Computer Science and Engineering, AKS University, Satna, Madhya Pradesh, India

Received 20 March 2023, Accepted 04 April 2023, Available online 05 April 2023, **Vol.13, No.2 (March/April 2023)**

Abstract

In this work, we are focusing on the review of the field of face detection where the attention of the researchers has been very less. A lot of work has been done in face detection and face recognition, and many methods were used, but still, the facial recognition of infants, old age people and people with dark skin show this lack of research. This review paper will focus on all those who worked on face detection technology keeping in mind the right to equality and getting better results. The most popular algorithm and methodology for face detection is the Viola-Jones algorithm for many years. This algorithm is based on a combination of Haar-like features and AdaBoost learning. It is widely used in many applications such as facial recognition, security systems, and video surveillance. This paper provides a comprehensive and comparative overview of different methodologies of face detection and the current state-of-the-art in deep learning-based face recognition systems. We found the best AI method for face detection is Convolutional Neural Networks (CNNs). CNNs are a type of deep learning algorithm that can detect and classify objects in images. They are particularly well-suited for face detection because they can learn to recognize patterns in data that are not explicitly labelled.

Keywords: *Face Detection, Face recognition, Viola-jones, PCA, ANN, AlexNet, VGGNet, CNN, Deep learning.*

1. Introduction

Most research has been giving face detection rate results by testing face detection research only on a limited variety of people. Due to which they used to get high detection rate, but when that test was done on different origin of people, the detection rate used to decrease. There are so many methods and algorithms available to locate human face from the image with the complex background. There are a few methods of face detection and recognition, which we will do a comparative study here. Now a day's robustness and reliability are very important for security systems.

So far, a large number of face detection approaches have been developed in last few years. This research paper will discuss a brief description and comparison of these face detection approaches with its pros and cons.

Early algorithms used for face detection and recognition used mainly colour information and intensity of pixels. The colour information was used to distinguish the face area from the background and the intensity of the pixels was used to extract the features. With the advent of artificial intelligence, more sophisticated algorithms were developed which used the combination of colour and shape information to distinguish the face from the background.

*Corresponding author's ORCID ID: 0000-0002-9815-7342
DOI: https://doi.org/10.14741/ijcet/v.13.2.6

An AI system uses a combination of image processing techniques, such as histogram of oriented gradients, computer vision and machine learning to detect and recognise faces.

Deep learning algorithms use convolutional neural networks to detect and recognise faces. Deep learning algorithms are able to detect and recognise faces even in different lighting conditions, with different face colour, poses, sharpness of facial features, and facial expressions. The deep learning algorithms are able to detect the minute details, such as the wrinkles and other facial features, which are not possible with the earlier algorithms.

Other stages of face detection and recognition include feature-based face recognition, 3D face recognition, and facial biometrics. Feature-based face recognition uses a set of facial features such as eyes, nose, mouth and chin to detect and recognise faces. 3D face recognition uses 3D images of faces to detect and recognise faces.

In this time, lots of face recognition methodology have been implemented for different applications. Robustness and reliability are becoming more and more important for these applications, especially in security systems. Finally, we will introduce the research on face detection for real conditions and also define general evaluation criteria and use of general databases of face recognition methods.

2. The different stage of face detection and related methodology with pros and cons

Many methods of face detection and recognition are available in many research papers. Figure 1 shows several methods divided into feature based and image-based technology.

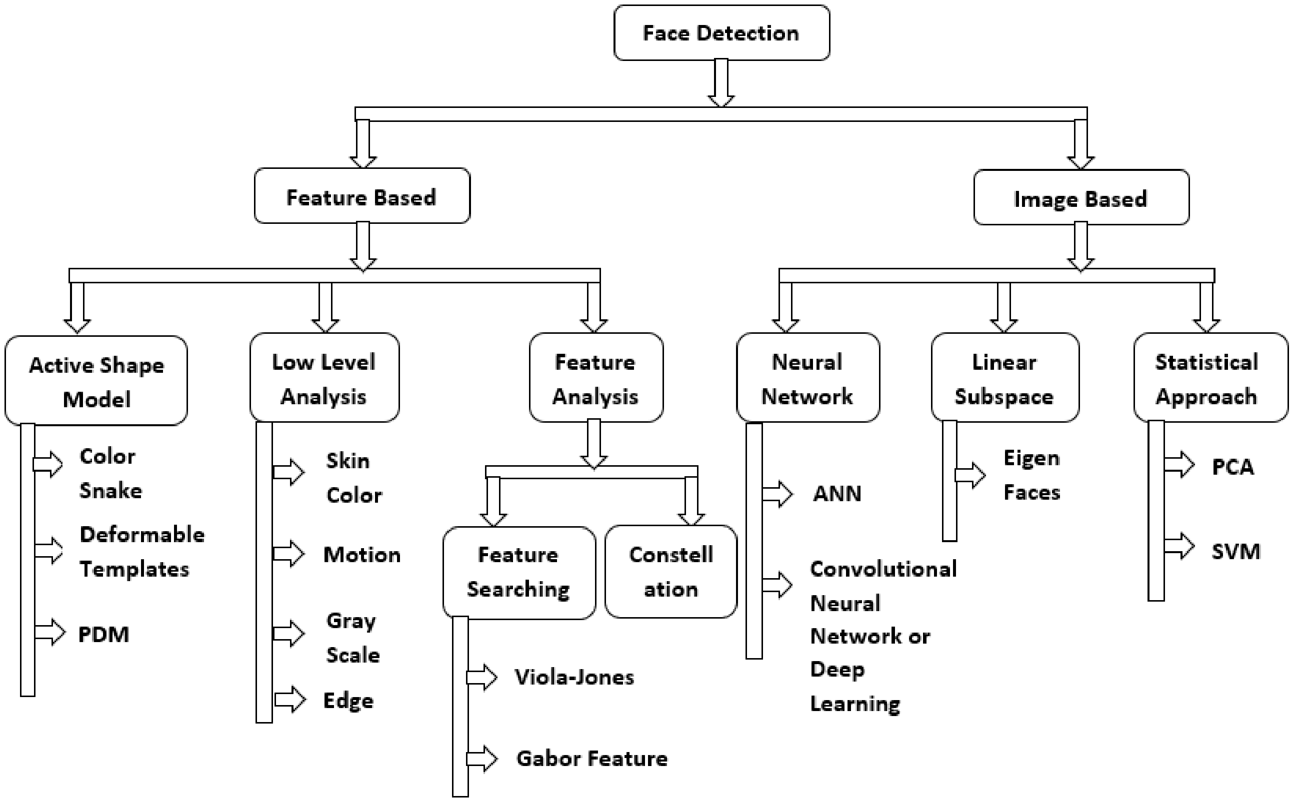

Figure 1. Different Methodology of Face detection

As we can see the figure 1 shows that the classification of all face detection and recognition methods. In this research paper we are focused on the all-important methodologies of face detection which gives the better result for face detection and recognition.

2.1 Color snakes

This method investigates the use of energy minimization as a framework within which to realize this goal. This technique seeks to design energy functions whose local minima comprise the set of alternative solutions available to higher-level processes [1].

The energy function, which is minimized, is a weighted combination of internal and external forces. The internal forces emanate from the shape of the snake, while the external forces come from the image and from higher-level image understanding processes [2]. The snake is defined parametrically as v(s) = [x(s), y(s)], where x(s), y(s) are x, y coordinates along the contour and s € [0, 1]. The energy functional to be minimized maybe has written as

$$E^*_{snake} = \int_0^1 E_{snake}\ (v(s))ds$$

The color snake algorithm is composed of three main parts: the first one is the face region estimation part, the second one is the detection part and the third one is the facial feature extraction part.

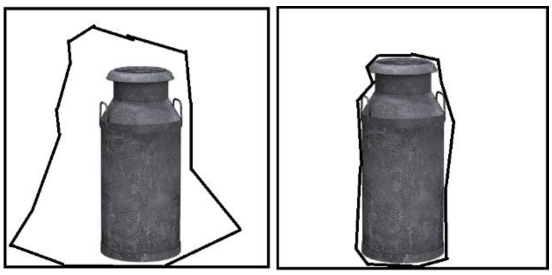

Figure 2. Initial and final position of a contour finding snake

Figure 2 shows a snake that find object contours in initial and the final position.

Pros:

A. The feature can be found easily.

Cons:

A. Sensitive to light condition occlusion.
B. Energy minimization is a problem for higher-level processing.

2.2 Deformable templates

Deformable templates are a method of face detection that uses a deformable template to describe the shape of a face. This template can be altered in order to capture more detailed facial features and improve accuracy [3].

The deformable template works by first using a shape-based model to detect the most once likely facial features - such as eyes, nose, and mouth. Then, by using its deformable capabilities, it is able to refine the detected features in order to achieve better accuracy. This allows for improved facial recognition, as well as improved accuracy when detecting different types of faces.

The formula of the deformable template method for face detection is a combination of shape-based modelling, deformable capabilities, and classification algorithms. Finally, the classification algorithms are used to classify the faces into different categories.

Pros: A. The advantages of the deformable template method for face detection include increased accuracy, improved facial recognition, better detection of different face types, and increased speed.

B. Additionally, it is more reliable than conventional methods [4] and can be used in real-time applications.

Cons: A. The disadvantages of the deformable template method for face detection include increased computational cost and the potential for overfitting.

B. Additionally, the method requires a lot of training data in order to achieve good results and there may be difficulties in obtaining the necessary data.

2.3 PDM

PDM stands for Point Distribution Model and is a type of deformable template method used for face detection. This model creates an initial template using a shape-based model for facial features, and then refines this template using deformable capabilities. This results in improved accuracy when detecting different types of faces.

PDM works by first using a shape-based model to detect the most likely facial features - such as eyes, nose, and mouth. Then, by using its deformable capabilities, it is able to refine the detected features in order to achieve better accuracy. This allows for improved facial recognition, as well as improved accuracy when detecting different types of faces.

The results demonstrate the effectiveness of the proposed method in accurately estimating the pose of a face image sequence. [5] The pose estimates produced by the proposed method are significantly more accurate than those produced by other methods in both simulated and real data. Furthermore, the proposed method is able to accurately capture subtle changes in facial expression over time.

Pros

A. PDM method is a robust and efficient method for face detection. It is able to detect faces in different poses and lighting conditions.
B. It is a fast method and can detect faces in real-time.
C. It is a simple and easy to implement method.
D. It is a non-parametric method and does not require any training data.

Cons

A. PDM method is not suitable for detecting small faces.
B. It is not suitable for detecting faces with large variations in pose and lighting conditions.
C. It is not suitable for detecting faces with occlusions.

2.4 Skin color

The skin color method of face detection works by analyzing the color of an image to identify areas that are likely to be skin. It does this by comparing the color of each pixel in the image to a range of skin tones and then assigning a score to each pixel based on how closely it matches the skin tone range. The areas with the highest scores are then identified as potential faces. This method is often used in combination with other methods such as edge detection and shape analysis to improve accuracy.

The AdaBoost algorithm was used to learn a cascade of strong classifiers from a training set of face and non-face images. For face recognition, the research proposed a novel method based on sparse representation and discriminative learning. The algorithm is based on the assumption that the face images can be represented as sparse linear combinations of a set of basis images. The sparse representation is used to identify the most relevant features of a face image and the discriminative learning techniques are used to classify the face images. The proposed method has been tested on benchmark datasets and achieved promising results [6].

The results of this system showed a significant improvement in accuracy over existing methods. The combination of Gabor features extraction, PCA, and KNN-based classification has been found to be one of the most effective and efficient ways to achieve accurate face recognition.

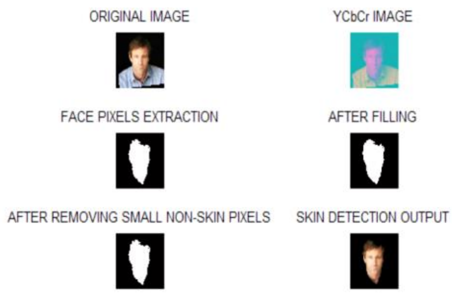

Figure 3. Skin color segmentation [6]

Figure 3 shows the segmentation of image for detecting the face. In order to improve the recognition rate, the proposed algorithm requires to be extended to detect and recognize faces in different orientations. It includes rotating the images to make them face normal, and using detection and recognition algorithms that can recognize faces in different orientations. Furthermore, the algorithm should be improved to detect and recognize faces in various lighting conditions, as well as different facial expressions. Additionally, the recognition rate can be improved by using more powerful deep learning models and better training techniques.

Pros

A. Skin color method is a relatively simple and fast method of detecting face.
B. It is relatively accurate in cases where the face is well illuminated and the skin color is distinct from the background.
C. It requires less computing power than other face detection methods.

Cons

A. The accuracy of the skin color method is highly dependent on the lighting conditions and the color of the background.
B. It is not very reliable in detecting faces with darker skin tones.
C. It is also not very effective for detecting faces in complex backgrounds.
D. It can be easily fooled by false positives; such as objects with similar color to skin tone.

2.5 Gray scale

The algorithm begins by using skin color detection to identify the regions of the image which contain human skin. It then performs edge detection on the skin color regions to obtain a set of edges that define the boundaries between the skin color regions and other regions.

Next, it uses the Weber local descriptor method (WLD) to extract the local features of the detected edges. Finally, the algorithm uses entropy technique to analyze the extracted local features and identify the faces from the image.

The algorithm is advantageous because it combines several existing techniques in order to improve the accuracy and speed of face detection. Additionally, the algorithm is able to detect faces from different angles and in different lighting conditions, which makes it suitable for a wide range of applications.

The algorithm works well in detecting faces of fair complexion, but it may not work so well on faces with different skin tones, such as those with dark skin tones. To make the algorithm more robust and accurate, it is important to train the algorithm using image data sets with a variety of skin tones. Additionally, the algorithm should be tweaked to account for different lighting conditions and angles. Finally, the algorithm should be designed to account for facial features that may differ across skin tones, such as shape, size, and texture.

Pros

The proposed method has a few advantages over other methods [7].

A. Gray scale method is a simple and cost-effective way to detect faces in digital images.
B. It is a fast and accurate method for face detection.
C. It is a robust method that can detect faces in different lighting conditions.
D. It is a reliable method that can detect faces in different orientations.

Cons

A. It is not able to detect faces in low-light conditions [8].
B. It is not able to detect faces in different expressions.
C. It is not able to detect faces in different poses.
D. It is not able to detect faces in different sizes.

2.6 Viola-jones algorithm

The Viola Jones algorithm is a widely used face detection algorithm which works by scanning an image for features. The algorithm uses a cascade of classifiers, each trained to detect a specific feature in the image. Each classifier is constructed using a set of sample images of faces, and the algorithm compares the features of the test image with each of the sample images. If the features match those in the sample images, then the test image is classified as a face [9].

Each classifier produces an output score which indicates the probability that the test image contains a face. If none of the classifiers in the cascade produce a score above a certain threshold, then the test image is classified as not containing a face.

The main advantages of Viola-Jones algorithm are that it is fast, robust and accurate. It is also capable of detecting faces in real time. [10] The algorithm does not require any prior knowledge about the face being detected.

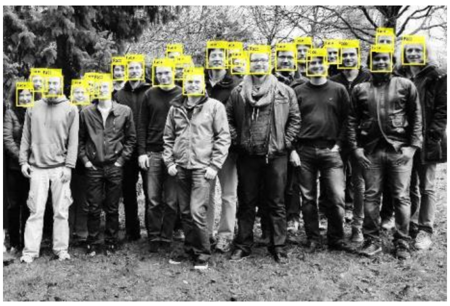

Figure 3. Best result using Viola-Jones algorithm [9]

It is also capable of detecting faces in different orientations, lighting conditions, and facial expressions. The algorithm is also able to learn from its mistakes and adapt itself to different conditions. The main disadvantage of Viola-Jones algorithm is that it may not be able to detect faces in low-light conditions or when the faces are partially obscured.

The four main points of the Viola-Jones face detection algorithm in detail are as follows:

1) Haar Feature Selection: This involves calculating the differences between different regions of an image to help identify objects within it. The algorithm has been trained to recognize the location of a face in an image by looking for specific patterns in the differences of light and dark areas.
2) Integral Image Concept: This involves pre-processing digital images, which is necessary to speed up face recognition. This technique helps to store image features in a single array for easier processing.
3) Adaboost Learning: This is a machine learning algorithm that works by combining a set of weak classifiers (or models) to build a strong classifier [10].
4) Cascade of Classifiers: The Cascade of Classifiers in the Viola-Jones algorithm is used to reduce the amount of computation time needed for face detection. It works by progressively eliminating false positives, or false alarm regions, from the image using multiple stages of classifiers. Each stage uses a classifier that is trained to recognize more complex patterns than the previous stage. This reduces the amount of false alarms and allows for faster processing.

Pros

A. Viola-Jones algorithm is a fast and efficient algorithm for face detection. It can detect faces from a variety of angles and scales.
B. It is a robust algorithm that can detect faces in various lighting conditions, including low light and backlight.
C. The algorithm is computationally efficient and can be used in real-time applications.
D. It is a simple algorithm and can be easily implemented and gives better response in different type of faces.

Cons

A. Viola-Jones algorithm is not very accurate and may miss some faces.
B. It is sensitive to changes in the face, such as facial expressions and hairstyles.

2.7 ANN algorithm

Neural networks (ANNs) are used in face detection algorithms to detect and classify faces in images.

As shows in figure 4, ANNs analyse facial features within an image, such as the eyes and mouth, by extracting features from the image using convolutional layers and then learning a classifier that identifies whether or not the image contains a face [11].

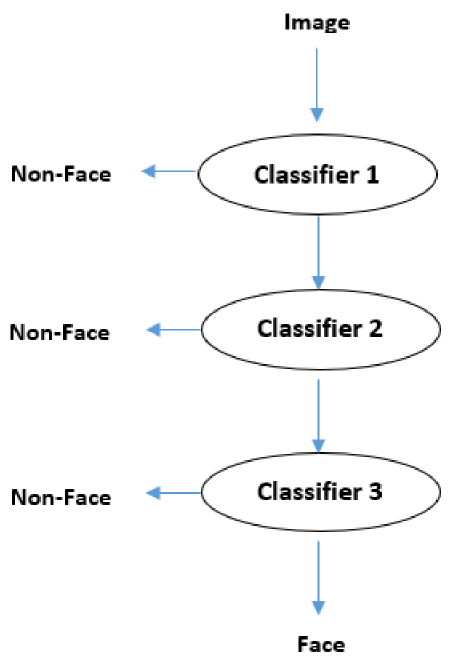

Figure 4. Attentional cascade representations

The benefits and results of using ANNs for face detection include improved accuracy and speed over traditional methods. ANNs are able to detect and classify faces in images more accurately and quickly than other algorithms, and can also detect features in an image that may be too small or subtle to be detected by traditional methods. Additionally, ANNs offer a higher degree of scalability, making them suitable for large-scale facial recognition applications.

ANNs for face detection work by analysing facial features within an image. This is done by extracting features from the image using convolutional layers and then learning a classifier that identifies whether or not the image contains a face [12].

The classifier is trained to recognize specific patterns in the differences of light and dark areas within the image. Once the classifier is trained, it can then be used to detect faces in other images with a high degree of accuracy.

Pros

It is a fast and accurate algorithm for face detection.
It is relatively simple to implement.
It is robust to changes in illumination and pose.
It is computationally efficient.

Cons

It is not suitable for detecting faces in low-resolution images.
It is limited to detecting frontal faces.

It is sensitive to false positives.
It is not suitable for detecting small faces.

2.8 Recent advances in deep learning techniques

The research has also demonstrated that Deep learning-based face recognition systems have shown excellent performance in recent decades. Different types of datasets such as still image-based, heterogeneous face image-based, video-based, and occlusion-based datasets have been used to train and evaluate the performance of the systems. The most commonly used datasets are LFR, IJB, YTF, and Ms-celeb-1M [13].

Despite the impressive performance of these systems, occlusion-based challenges still remain in the face recognition task. To address this issue, more datasets and novel algorithms are needed. Additionally, research should focus on developing systems that are robust to variations in pose, expression, age, and ethnicity.

Recent advances in deep learning techniques have enabled the development of powerful facial recognition systems. Different types of datasets such as still image-based, heterogeneous face image-based, video-based, and occlusion-based datasets have been used to train and evaluate facial recognition systems.

In particular, the Labeled Faces in the Wild (LFW) dataset, the IARPA Janus Benchmark (IJB) dataset, the YouTube Faces (YTF) dataset, and the Microsoft Celebrities (Ms-Celeb-1M) dataset have been shown to achieve near-perfect performance in various facial recognition tasks. However, occlusion-based challenges still remain a major challenge in the facial recognition task.

Pros

A. High accuracy: Deep learning methods for face detection are capable of achieving high accuracy in face detection tasks.
B. Robustness: Deep learning methods are robust to changes in illumination, pose, and facial expressions.
C. Speed: Deep learning methods are capable of detecting faces in real-time.
D. Scalability: Deep learning methods can be easily scaled to larger datasets.
E. Adaptability: Deep learning methods can be adapted to different types of data and tasks.

Cons

A. High computational cost: Deep learning methods require large amounts of data and computing power to train.
B. Limited interpretability: Deep learning methods are difficult to interpret and explain.
C. Lack of generalization: Deep learning methods can be prone to overfitting, resulting in poor generalization to unseen data.
D. Data bias: Deep learning methods can be biased towards certain types of data, leading to inaccurate results.

Conclusion

The first successful method used for face detection is Haar cascade classifier. This is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is a very fast and effective method for face detection. There is a table shows some method with its accuracy results.

Table 1. Attentional cascade representations

S. No.	Methods	Training Dataset	Accuracy In %
1	DeepID [14]	CalebFaces	92.2
2	DeepFace [15]		92.5
3	Range Loss [16]	VGGs	93.7
4	CNN [17]		94.9
5	FaceNet [18]		95.1
6	DeepVisage [19]	MS Celeb-1M	95.9
7	PRN [20]	VGGFace2	96.3
8	FSENet [21]	MS Celeb-1M	97.89
9	ArcFace [22]		98.02
10	BiometricNet [23]	Casia	98.6

The first successful method used for face detection is Haar cascade classifier. This is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is a very fast and effective method for face detection.

The second successful method is Local Binary Pattern (LBP). LBP is a non-parametric method for face detection. It is based on the concept of local texture analysis and uses a set of predefined binary patterns. The advantage of this method is that it is fast and reliable.

The third method is Histogram of Oriented Gradients (HOG). HOG is a feature descriptor used for object detection. It is based on the concept of local gradient orientation and uses a set of predefined orientation bins. The advantage of this method is that it can work with low resolution images and is more robust than the other methods.

Black skin face detection is still a challenge for many algorithms. The issue is that algorithms are often developed using datasets composed primarily of white faces, making them less accurate at detecting black faces. Additionally, lighting and other environmental factors can pose a challenge for face detection algorithms, which can be more difficult for darker skin tones. Researchers are currently working on improving the accuracy of face detection algorithms for black skin tones.

Finally, the fourth method is Deep Learning based Face Detection, which is the most accurate and reliable face detection method. It is based on convolutional neural networks and uses a set of layers of convolutional filters. The advantage of this method is that it is very accurate and can work with high resolution images.

In conclusion, according to the table 1, Deep learning-based face recognition systems have shown excellent performance in recent years. It is providing the better result for black skin face(s), and new born baby's also. Different types of datasets and architectures are used to improve the performance of the Facial Recognition systems.

Future work

The most commonly used datasets for face recognition are Labeled Faces in the Wild (LFW), Adience, and MegaFace. LFW is a dataset of 13,000 images of faces collected from the web. Occlusion-based challenges still appear in the FR task, but more datasets and novel algorithms may reduce these problems.

References

[1] M. Kass, A. Witkin, D. Terzopoulos et al., "Snakes: Active Contour Models," *International Journal of Computer Vision*, pp. 321-331, 1988.

[2] K. H. Seo, W. Kim, C. Oh, J. J. Lee, et al., "Face Detection and Facial Feature Extraction Using Color Snake", *Industrial Electronics Proceedings of the IEEE International Symposium*, 7803-7369, july 2002.

[3] A. L. Yuille, et al., "Deformable Templates for Face Recognition," *Journal of Cognitive Neuroscience,* Volume 3, Number 1, January 1991.

[4] A. Bennett & I. Craw, et al., "Finding Image Features Using Deformable Templates and Detailed Prior Statistical Knowledge," *Conference of Mowforth*, P. (eds) BMVC91. Springer, London. pp 233–239, 1991.

[5] C. M. Cheng, S. H. Lai, K. Y. Chang, et al., "A PDM based approach to recovering 3D face pose and structure from video," *International Conference on Information Technology: Research and Education*, ISBN:0-7803-7724-9, August 2003.

[6] B. Dhivakar; C. Sridevi; S. Selvakumar, P. Guhan, et al., "Face detection and recognition using skin color," *3rd International Conference on Signal Processing, Communication and Networking (ICSCN),* ISBN:978-1-4673-6823-0, March 2015.

[7] L. Laur; M. Daneshmand; M. Agoyi; G. Anbarjafari, et al., "Robust grayscale watermarking technique based on face detection," *23nd Signal Processing and Communications Applications Conference (SIU),* SSN: 2165-0608, May 2015.

[8] J. Das; H. Roy, et al., "Human Face Detection in Color Images Using HSV Color Histogram and WLD," *International Conference on Computational Intelligence and Communication Networks*, ISBN:978-1-4799-6929-6, November 2014.

[9] L. N. Soni, A. Datar and S. Datar, et al., "Implementation of Viola-Jones Algorithm Based Approach for Human Face Detection," *International Journal of Current Engineering and Technology*, pp. 1819-1823, Vol.7, No.5, Sept/Oct 2017.

[10] L. N. Soni, A. Datar, S. Datar, "Viola-Jones Algorithm Based Approach for Face Detection of African Origin People and Newborn Infants," *International Journal of Computer Trends and Technology (IJCTT)*, Vol. 51, Number 2, September 2017.

[11] M. Owayjan, R. Achkar, M. Iskandar, et al., "Face Detection with Expression Recognition using Artificial Neural Networks," *3rd Middle East Conference on Biomedical Engineering (MECBME* ISSN: 2165-4255, October 2016.

[12] S. Q. Alhashmi; K. H. Thanoon; O. I. Alsaif, et al., "A Proposed Face Recognition based on Hybrid Algorithm for Features Extraction," *6th International Engineering Conference Sustainable Technology and Development (IEC),* ISBN:978-1-7281-5911-9, February 2020.

[13] M. T. H. Fuad, A. A. Fime, D. Sikder, M. A. R. Iftee, J. Rabbi, N. S. Al-rakhami, A. Gumaei, O. Sen, M. Fuad, M. N. Islam, et al., "Recent Advances in Deep Learning Techniques for Face Recognition," *Preparation of Papers for IEEE Transactions and Journal,* ISBN: 0-7803-7369-3, Vol. 9, July 2021.

[14] Y. Sun, X. Wang, X. Tang, et al., ''Deep learning face representation from predicting 10,000 classes,'' *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1891–1898, June 2014.

[15] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, et al., ''DeepFace: Closing the gap to human-level performance in face verification,'' *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1701–1708, June 2014.

[16] X. Zhang, Z. Fang, Y. Wen, Z. Li, Y. Qiao, et al., ''Range loss for deep face recognition with long-tailed training data,'' *IEEE International Conference on Computer Vision (ICCV),* pp. 5409–5418, Oct. 2017.

[17] C. Ding, D. Tao, et al., ''Trunk-branch ensemble convolutional neural networks for video-based face recognition,'' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 40, Issue. 4, pp. 1002 - 1014, April 2018.

[18] F. Schroff, D. Kalenichenko, J. Philbin, et al., ''FaceNet: A unified embedding for face recognition and clustering,'' *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823, June 2015

[19] A. Hasnat, J. Bohné, J. Milgram, S. Gentric, L. Chen, et al., ''DeepVisage: Making face recognition simple yet with powerful generalization skills,'' *IEEE International Conference of Computer Vision Workshops (ICCVW),* pp. 1682–1691, Oct. 2017.

[20] B.-N. Kang, Y. Kim, D. Kim, et al., ''Pairwise relational networks for face recognition,'' *European Conference on Computer Vision (ECCV),* vol. 11206, pp 646–663, October 2018.

[21] X. Cheng, J. Lu, B. Yuan, J. Zhou, et al., ''Face segmentor-enhanced deep feature learning for face recognition,'' *IEEE Transactions on Biometrics, Behavior, and Identity Science*, Vol, 1, Issue. 4, pp. 223 - 237, October 2019.

[22] J. Deng, J. Guo, N. Xue, S. Zafeiriou, et al., ''ArcFace: Additive angular margin loss for deep face recognition,'' *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 4690–4699, June 2019.

[23] A. Ali, M. Testa, T. Bianchi, E. Magli, et al., ''BioMetricNet: Deep nconstrained face verification through learning of metrics regularized onto Gaussian distributions,'' *European Conference on Computer Vision, Switzerland: Springer*, vol. 12370, pp. 133–149, November 2020.