Face Detection and Expression Recognition Using Machine Learning: A Review

Kavita, Rajender Singh Chhillar
Department of Computer Science and Applications, Maharshi Dayanand University, Rohtak, Haryana, India

Abstract:

Face detection and recognition have become an emerging field in computer applications today. This is the main reason that the applications of face recognition are increasing frequently. There is a wide range of areas where it is used such as medicine, government, education, banking, security and surveillance, and many more. It is the process of identifying individual faces from faces. Besides, Face Expression Recognition (FER) helps to identify thebehavior of a person in online events. The biometric is the strongest part of face authentication therefore; the scope of innovation of new approaches develops the interest of researchers. Also, digital images experience different types of challenges during face detection due to the high degree of diversity in the appearance of faces in inconsistent environments which affects the quality and accuracy. Image preprocessing and machine learning take place an important role in it. This paper presents a survey of face detection and recognition applications. Moreover, demonstrate over-face expression recognition. The paper is an analysis of various models applied to different datasets. As seen results vary by employing a framework and applied on different datasets.

Keywords: face detection, face expression, face recognition, machine learning, dataset
DOI Number: 10.48047/nq.2021.19.12.NQ21232 **NeuroQuantology2021;19(12):394-400**

1. Introduction

Human evolution has advanced through several stages in a fast-forward mode. Real-time face recognition is an interesting field with continually growing challenges in the digital era environment. Theintrusive aspect of face recognition has garnered attention, as has the fact that it is the most reliable method of human identification among all biometric methods. The process of accurately identifying a person's face using a visual system and a pre-existing database is known as facial recognition. Because of its widespread implementation in security systems, video surveillance, commercial venues, and social networks like Instagram, Facebook, etc., it has become an indispensable instrument for human-computer interaction [1][2]. Due to its intrusive nature and primacy among biometric identifiers, face recognition has come under scrutiny after the rapid advancement of AI. Recognizing people in photos when their faces are showing a wide range of emotions requires a combination of ML & DL techniques [10].

2. Machine Learning

ML is computer programming with the ability to learn &develop autonomously, making decisions and modifying behavior appropriately. It gives computer systems the capacity to improve themselves based on their past performance, with no further programming required. Building intelligent, self-learning computer programs is the main goal. Machine learning not only employs a range of algorithms to repeatedly classify things to aid judgments in novel input settings or ambiguous real-world situations; but learns, explains, and upgrades data to predict better outcomes in face detection [2]. Both supervised and unsupervised models are used in machine learning. Using labeled examples, supervised machine learning algorithms may extrapolate previous knowledge to predict future outcomes in the form of fresh data. This algorithm is task driven and mainly used in Predictive Modelling. Regression and classification problems are solved with machine learning. Labeled data is used for the training set. Popular algorithms are Linear Regression (LR), Logistic Regression, DT, Random

Forest, KNN, Naive Bayes, KNN, SVM, and Neural Networks.

In addition, the machine learning technique known as "deep learning" allows computers to acquire knowledge about and insight into the world around them by mimicking human learning and cognition processes across several levels of abstraction. As the computer learns from its own mistakes, there is no need for a human operator. Deep Learning is predominantly a self-managing, self-teaching system that makes use of pre-existing data to train new algorithms to uncover new patterns. These patterns are used to make forecasts about new data that the computer has never seen before. In addition to being able to make judgments instantaneously, deep learning systems also can continually improve their performance [11]. The versatility of the neural network-based deep learning algorithm lies in its ability to perform both supervised and unsupervised learning.

3. Face Recognition

FR is a technique of pattern matching that can determine whether or not two faces in a database belong to the same person. FR plays a crucial part in the identification process because of how easy and effective it is to detect a human face to verify an individual's identity [1]. To begin, we use the same network to calculate the image's face embedding and then compare it to the other embeddings [2]. When comparing two photos for face recognition, if the created embedding is closer to or comparable to any other embedding, image1 will not preserve its embeddings but will instead save image2s. After comparing the two new embeddings with the old ones, found that the vector for image2 has more similar to the other face embeddings of image2, whereas the face embeddings of image1 have not more similar to any other embeddings [3][4].

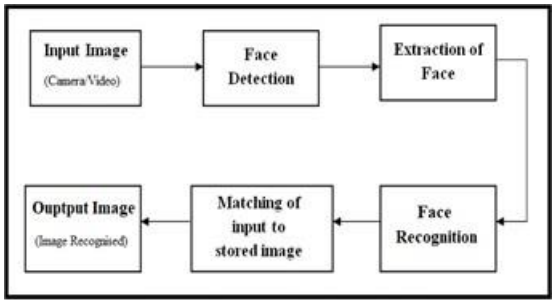

Fig 1. Face Recognition Block Diagram [3]

Hence steps will be taken for Face Recognition as:

i. Firstly, identify the human faces with a face detector.
ii. Predict face expressions after identification.
iii. Evaluate face encodings.
iv. Compare these face encodings of known faces with test images to find out the person in an image.

3.1 Face Detection

Humans can quickly detect faces in photographs, but computers have always struggled due to the dynamic aspect of faces. The issue can be defined as follows: Detect and localize an unknown face in a still or video picture [2]. The faces are identified using the Viola-Jones method. There are four stages:

i. Haar cascade features
ii. Integral image
iii. Adaboost training
iv. Cascading classifier

i. Haar Feature Selection

There are commonalities across all different types of human faces. The black-and-white image details may be picked out using Haar cascade features. Haar Features are used to identify these patterns. Figure 2 depicts several examples of feature expressions.

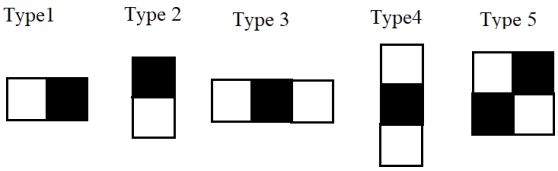

Fig 2. Different types of Haar features

Here, the squares are used to illustrate the Haar-like characteristic. The bridge of the nose tends to be brighter than the eyes, and vice versa, and the area around the eyes tends to be darker than the upper cheekbones. Matchable face features have a similar combinationof traits, which allows them to

- Location and size: eyes, mouth, bridge of the nose
- Value: oriented gradients of pixel intensities

Each Haar feature is a pixel-perfect representation of the window. Each tab in the secondary window corresponds to a feature. When the Haar characteristics are used to determine a value, the result is

Value = Σ (pixels in the black area) - Σ (pixels in the white area)

ii. Integral images

As illustrated in Fig. 3, the Viola-Jones face identification algorithm's second phase involves

transforming an input picture into an integral image. The pixels to the upper left and upper right of (x,y) are added together to form the integral picture at that point.

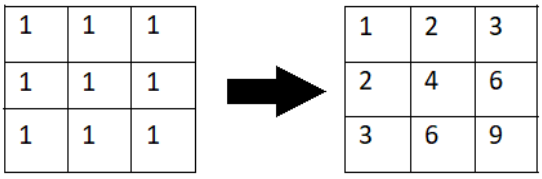

Fig 3. (a) input image **(b)** integral image

This simplifies the computation of the sum of all pixels inside a given rectangle down to just four numbers. These are the pixel values in the integral picture that aligns with the rectangle's edges in the input image [3].

iii. Adaboost training/machine learning method

Adaboost is an ML technique that helps choose the best features from a dataset containing 160,000 or more characteristics. To analyze all the features in a picture, the Viola-jones technique, for instance, would need a window size of 24x24, yielding 160,000+ features. Not every function has to be included. So we can get rid of them. Each of these characteristics is then assigned a weight in determining whether or not a specific window has a face. Weak classifiers are characteristics like these [4].

Iv. Classification

Extracted features are fed into the classifier. Classification can be done with a machine learning algorithm such as DT, RF, KNN, NB, SVM, and NN. For this, compare test image features with features stored in the dataset. If they matched then the image is detected, and after that recognizes the image.

3.2 Face Expression

Researchers frequently utilize facial expression recognition to analyze and identify emotions. Many research studies have focused on various emotional states, including happiness, surprise, anger, sadness, fear, and neutrality. It's important to use this technology because facial expression recognition could replicate human coding capabilities [2]. Many research fields may benefit from facial expression recognition (FER), including mental health diagnosis and detection of human social/physiological interactions [3][5][15]. The table represents a review of face expression detection in the form of a comparison of the feature table.

Table 1. Comparison of feature table

Citation	Face detection	Face recognition	Face expression
[1]	Yes	Yes	No
[2]	Yes	Yes	Yes
[3]	Yes	Yes	Yes
[5]	Yes	Yes	No
[6]	Yes	Yes	No
[7]	Yes	Yes	No
[8]	Yes	Yes	Yes
[9]	Yes	Yes	Yes
[11]	Yes	Yes	No
[13]	Yes	Yes	No
[14]	Yes	Yes	No
[15]	Yes	Yes	No

4. Literature Survey

The study analyzed several articles on the subject of facial identification, facial emotion recognition, and facial feature detection. This article covers the years 2017 through 2020. Certain challenges in the face recognition domain and facial traits are outlined in the study. The results might change when using other datasets. For researchers, the study reveals blind spots. The report covers strategies, outcomes, and restrictions [1]. The identification rate was broken down into its parts—the eyes, nose, jaw, and cheeks—as well as the influence of rotation on faces and zoom-out-style characteristics. The FET and LFW data sets, as well as the CS and linear SVM classifier, were utilized in conjunction with the VGG system's CNN model to identify facial recognition characteristics. The research observed that on the parts of faces, the recognition rate becomes low and when combined with these parts the rate becomes high. So, it was better that combine all the parts and impact of faces for good results and probes can use to increase performance with the same [2].

This study provided a comprehensive evaluation of the literature on facial expressions. For clarity, several algorithms and methods for detecting faces were studied and shortened for easy reference, including principal component analysis, linear bias pruning, optical flow, and Gabor filters. When the system's accuracy was put to the test, CK+ databases revealed trouble with correctness during cognition of facial expressions with surroundings and also when poses were modified. CNN, SVM, KNN, NB, HMM, DT, and many more are all examples of classification and

regression techniques [3]. Then a face recognition model was introduced with PCA; a statistical system to shorten the huge storage. PCA has the main feature of projection of self-space by forming a model of face recognition and then fixing the algorithm by doing programming on OpenCV and Python. When LBPH and HOG are used together, they outperform more established facial recognition algorithms like Eigenface and Fisher Face [4].

The Researchers presented a method for automatically recording attendance using advanced image processing techniques such as contrast modification, bilateral filtering, histogram equalization, and picture blending in conjunction with an LBP algorithm [6]. The conventional student attendance systems at different institutions were then recommended to be upgraded to prevent the waste of time and resources. The primary goals in the development of the system were to decrease the time burden, amount of labor, and the number of disposables needed over the long run, while simultaneously increasing the flexibility and performance of the attendance system method. Haar cascade was employed for detection, while the LBPH model and Python [7] were used for facial recognition.

The conventional face detection and identification system was introduced based on the CNN model. The usage of Python libraries after evaluating several articles on face detection, and recognition, including facial expression. When it comes to detecting and identifying human faces, the suggested CNN model demonstrates improved accuracy and prediction. The results demonstrate the limitations of state-of-art in face detection and identification. The suggested technique has proven effective in biometrics and surveillance [8]. When face detection, identification, and emotion categorization were all suggested, there were three goals. The viola jones algorithm was used for facial recognition. This paper presents the results of a face recognition and classification experiment using the CNN model and the KDEF dataset with VGG 16. There are several libraries, including dlib, that support the Python programming language. It's a good tool for finding out how people feel about E-learning. Using the Keras and Theano neural network modeling tools, a Python application was written to train a CNN and detect emotions for seven facial expressions using the FER2013 dataset [9].

Yet another study introduced a deep-learning architecture for recognizing facial expressions, which made use of CNNs, RNN, and Attention Blocks. Better results may be seen on the FER dataset, although the model still has trouble differentiating between happy and sad states and between fear and surprise [10][17]. The ACNN model was presented for identifying pupils' emotions. The facial expression recognition system in the FER 2013 dataset is capable of reliably identifying a broad variety of student emotions since it is built on a framework of 4 convolutional layers, 4 max-pooling layers, and 2 fully connected layers. In a learning setting, the suggested system is doable [12].

In addition, the CNN model was used to implement a face recognition-based attendance tracking system. Throughout the whole process of collecting attendance, each person's face will be matched to the photo in the database to verify their identity[13]. The LFW dataset was then used to develop a unique end-to-end trainable CNN system for face detection and identification [14]. An effective deep convolutional network model was utilized, and it recognized the face in a variety of positions and extracted its most salient characteristics using a filter. With this method, On the LFW dataset for facial recognition, they achieve state-of-art performance while using just 64 bytes per face [15].

The study revealed a comparison of OpenCV's and a dlib efficiency, and it also tracked the two libraries' respective run times and levels of complexity. On the IOT platform, face detection and recognition apps developed using the OpenCV library were more efficient and had superior performance[16]. Two normalization methods were created for use on two levels of the modified Convolution Neural Network (CNN) architecture. The suggested technique improved accuracy by using batch normalization to both the first and last convolution layers. Facial features were extracted using CNN's architecture, and the fully connected layer employed a Softmax classifier to categorize faces. After testing the suggested method on the Face dataset, researchers found that it significantly boosted the efficiency of facial recognition [18].

5. Applications of Face Recognition

Applications for biometric-based security have drastically developed in recent years, particularly in face recognition. This software is a potent means for the government, smartphones, smart homes, and law enforcement to properly and effectively guarantee personal security [7][8][18]. The table illustrates thedomain and applications of face recognition.

Table2. Domain and Applications of Face Recognition

Domains	**Examples/ Applications**
Access Control	Devices access, vehicle access, home safety, secure physical locations, computer program access, certain IoT devices access
Immigration	Stricter border control, Terrorist alert, and secure flight boarding system
Education	Tracking student attendance, securing premises
Retail /Retailing	Secure payment processing system. online shopping, safe mobile payments, marketing targets, intelligent advertising
Healthcare	Prescriptions, medication adherence, analyze diseases
Human-Computer Interaction	Interactive gaming, proactive computing, cloud managing
Criminal Identification	IDs, driver's licenses
Security	Fraud detection, Shoplifting prevention, Master Card Identity Check, Automobile Security, Airport Safety, Computer Security
Surveillance	CCTV, and social media platforms (Facebook, Instagram, etc), perceive VIPs at sporting events
Law enforcement	Traffic management, forensic investigation, finding missing persons, entitlement programs

6. Methodology used

There are so many machine learning techniques present. From them, the convolutional neural network model gives better results. So, nowadays every researcher moves towards a deep learning model as well because it gives automatic results with more accuracy. A dataset is utilized for this purpose, with 80% of the data being used to train the model and 20% used for testing. Confusion matrices are used to measure a model's accuracy.

6.1 Datasets

There are several datasets are present for face recognition. FET, LFW, FER, CK+, and many more. The LFW dataset is utilized for face identification, recognition, and expression recognition, whereas the FER dataset is popular for the same purposes. The pictures used in FER are 48x48 pixels in size, and they depict a range of emotions, including anger, disdain, fear, pleasure, sadness, surprise, and neutrality. The dataset is comprised of 28709 training samples, 3589 public testing samples, and another 3589 private testing samples [7-10]. LFW is a picture collection of human faces that was put together to solve the issue of unconstrained facial recognition. Researchers from the University of Massachusetts Amherst created and are currently updating the database. The Viola-Jones face detector was fed 13,233 photos of 5,749 individuals collected from the internet. There are 1,680 identical pictures in the archive [12,13,15].

6.2 Confusion matrix

In many cases, the nature of the issue being addressed will dictate the optimal performance statistic to use. A table called a confusion matrix may be used to trace the projected classification against the actual classification.

Table 3. Confusion matrix used for datasets

Predicted Label	**Real Label**	
	Positive	**Negative**
Positive	TP	FP
Negative	FN	TN

Accuracy reveals how well a model is being taught any its potential future performance [12]. However, it does not provide specifics on how to apply this to the issue at hand. But, when there is a large discrepancy between social classes, accuracy fails to perform effectively. We can tell which transactions are legitimate and which are fraudulent by utilizing the confusion matrix and the dataset. The accuracy evaluation equation [8] may be written as

$$Accuracy = \frac{\sum TP+TN}{\sum TP+TN+FP+FN} \quad (1)$$

7. Results and discussion

For face Recognition the machine learning model's performance analysis is done firstly with datasets, represents in the table 4, and then the corresponding graph of models with accuracy is also generated, shown in Fig. 4.

Table 4. Performance of Machine Learning Models with datasets

Citation	Machine Learning Models	Datasets	Accuracy %
Sharmila et al. [5]	LBPH	Self	80
Suresh et al. [6]	LBPH	Self	99
Bah et al [7]	CNN	FER	70
Ahsan et al. [8]	CNN	FER	69
Ivanova et al. [9]	CNN	FER	64
Gupta et al. [10]	CNN	FER	70
N. Gupta et al. [12]	CNN+STN	LFW	86.3
Zhang et al. [13]	Deep CNN	LFW	99.7
Boyko et al. [15]	PSI- CNN	LFW	88.7
Nam et al. [16]	Deep coupled ResNet	LFW	99

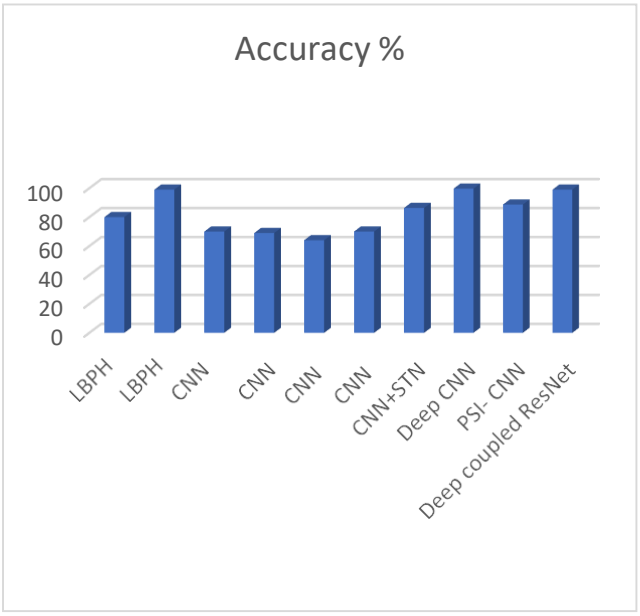

Fig 4. Performance of Machine Learning Models

8. Face Recognition Challenges

As a whole, face identification in digital images has to contend with a wide range of challenges, such as inconsistent illumination, large angular shifts, partial occlusions, altered faces, and the effects of time and wear. Similar to photos, films are often captured under inconsistent circumstances or from moving camera situations like visual surveillance [8]. There are numerous challenges as well as important elements that can significantly affect matching scores and facial recognition performance. Therefore; facing several difficulties. Facial appearance differences may be attributed to a wide range of causes, some of which are:

- Occlusion
- Illumination
- Pose variation
- Facial expressions
- Low Resolution
- Aging
- Plastic Surgery

9. Conclusion

To survey current progress in the area of face recognition, this work synthesizes the findings of a large number of other studies. This shows why facial recognition is so useful in many contexts. CNN frameworks have improved the presentation of face recognition procedures more than traditional ones, and these frameworks may be categorized as either

appearance-based, model-based, or hybrid. We still have a way to go when it comes to resolving issues like face expression, position variety, lighting, and occlusion. The paper directs to the analysis of models applied to the datasets. The findings that the accuracy and performance may certainly be improved by the researchers and provide a future aspect.

REFERENCES

1. Rasha Atallah, Amirrudin Kamsin, Maizatul Akmar Ismail, Sherin Ali Abdelrahman, and Saber Zerdoumi, "Face Recognition and Age Estimation Implications of Changes in Facial Features: A Critical Review Study", IEEE 2018.
2. Daniel Canedo and António J. R. Neves, "Facial Expression Recognition Using Computer Vision: A Systematic Review", doi: 0.3390/app9214678, MDPI 2019.
3. Rajesh K M, Naveenkumar M, "A Robust Method for Face Recognition and FaceEmotion Detection System using Support Vector Machines", IEEE 2017.
4. Maliha Khan, Sudeshna, Rani Astya, Shaveta Khepra, "Face Detection and Recognition Using OpenCV", ISBN: 978-1-7281- 4826-7/19/, IEEE 2019.
5. Sharmila, Raman Sharma, Dhanajay Kumar, Vaishali Puranik, Kritika Gautham "Performance Analysis of Human Face Recognition Techniques", IEEE 2019.
6. V. Suresh, S. Chakravarthi Dumpa, C. Deepak Vankayala, and J. Rapa, "Facial Recognition Attendance System Using Python and OpenCV," Quest Journals Journals of Software Engineering and Simulation, vol. 5, no. 2, 2019, pp. 2321–3809.
7. S. M. Bah, F. Ming, "An improved face recognition algorithm and its application in attendance management system", Array Elsevier Volume 5, March 2020, pp. 100014.
8. M. K. Ahsan, M., Newaz, S. H., "Human Face Detection Techniques: A Comprehensive Review and Future Research Directions", MDPI 2021.
9. E. Ivanova and G. Borzunov, "Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions", Elsevier 2020.
10. A. Gupta, S. Arunachalam, R. Balakrishnan, "Deep self-attention network for facial emotion recognition", Elsevier 2020.
11. I. Lasri, A.R. Solh, and M. El Belkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network", IEEE 2019.
12. N. Gupta, P. Sharma, V. Deep, and V. K. Shukla., "Automated Attendance System Using OpenCV",2020 8th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), IEEE 2020.
13. H. Zhang, and L. CHI, "End-to-end spatial transform face detection and recognition", Virtual Reality & Intelligent Hardware, Beijing Zhongke Journal, Elsevier 2020.Vol 2 Issue 2, pp119-131.
14. M. Alghaili, Z. Li and Hamdi A. R. Ali, "FaceFilter: Face Identification with Deep Learning and Filter Algorithm", Hindawi 2020, Article ID 7846264, 9 pages.
15. N. Boyko, O. Basystiuk, and N. Shakhovska, "Performance Evaluation and Comparison of Software for Face Recognition, based on Dlib and Opencv Library", IEEE 2018.
16. G. P., Nam, H. Choi, J., & Kim, I. J., "A Pyramid-Based Scale-Invariant CNN Architecture for Face Recognition Robust to Various Image Resolutions", Appl. Sci. 2018.
17. Lu, Ze, Xudong Jiang, and Alex Kot, "Deep Coupled ResNet for Low-Resolution Face Recognition", IEEE" IEEE Signal Processing, Letters 25.4, 2018, pp 526-530.
18. I. R. Pandey, M. Raj, K. K. Sah, Tojo Mathew, and M. S. Padmini, "Face Recognition Using Machine Learning", Volume: 06 Issue: 04, International Research Journal of Engineering& Technology, IRJET 2019.