Ukrainian Journal
of Educational Studies and
Information Technology

ISSN: 2521-1234 online
Ukr. J. of Educ. Stud. and Inf. Technol.
11(3), 211 – 226
doi: 10.32919/uesit.2023.03.05

Vol. 11, Issue 3, 2023

Techniques for facial affective computing: A review

Bashir Eseyin Abdullahi, Emeka Ogbuju, Taiwo Abiodun, Francisca Oladipo

Submitted:
2023 – 08 – 18
Accepted:
2023 – 09 – 26
Published online:
2023 – 09 – 30

Abstract
Facial affective computing has gained popularity and become a progressive research area, as it plays a key role in human-computer interaction. However, many researchers lack the right technique to carry out a reliable facial affective computing effectively. To address this issue, we presented a review of the state-of-the-art artificial intelligence techniques that are being used for facial affective computing. Three research questions were answered by studying and analysing related papers collected from some well-established scientific databases based on some exclusion and inclusion criteria. The result presented the common artificial intelligence approaches for face detection, face recognition and emotion detection. The paper finds out that the haar-cascade algorithm has outperformed all the algorithms that have been used for face detection, the Convolutional Neural Network (CNN) based algorithms have performed best in face recognition, and the neural network algorithm with multiple layers has the best performance in emotion detection. A limitation of this research is the access to some research papers, as some documents require a high subscription cost.
Practice implication: The paper provides a comprehensive and unbiased analysis of existing literature, identifying knowledge gaps and future research direction and supports evidence-based decision-making. We considered articles and conference papers from well-established databases. The method presents a novel scope for facial affective computing and provides decision support for researchers when selecting plans for facial affective computing.
Keywords: facial affective computing; face recognition; face detection; emotion detection; artificial intelligence.

INTRODUCTION

The term affective computing was introduced by Rosalind Picard in 1995 as a type of computing that deliberately influences emotion" (Calvo et al., 2015). Affective computing is a branch of artificial intelligence that deals with the interaction between humans and computers and allows the computer to study and understand signs from humans, like gestures, facial expressions, body language and voice tones, to assess emotional state. Thus, affective computing is analysing facial looks in images and text videos using computer vision technology to analyse human emotional status (Gaudenz, 2022). Also, Affective computing can be explained as the interaction between humans and computers in which a computer can detect and respond to human emotions and other stimuli. Facial expressions, gestures, posture, speech and changes in the hand's temperature on the mouse can all indicate changes in the user's emotional state. The computer can identify and interpret all these (Flynn et al., 2020). According to Gu et al. (2019), we have seven ways emotional states can be classified: anger, fear, sadness, happiness, disgust, contempt and surprise. These emotional states can be examined by an affective computing system though there are variations in classification. There are many application areas of affective computing, such as in the health sector, to monitor the entry and exit of patience, staff records, patient diagnoses and premises security. Also, it can be applied in the sporting arena to check the activities of spectators and other areas.

Facial Emotion Recognition (FER) is a technology researchers use to analyse sentiments from different sources, like videos and pictures. FER belongs to the class of technologies, which is frequently called "affective computing", a multidisciplinary field of research on the ability of the computer to identify and interpret human emotions based on Artificial Intelligence technologies. The technology analyses facial expressions from static images and videos to reveal information on human emotional states. Facial expressions play an important role when recognising emotions and can be used in the non-verbal communication process and the identification of people (EDPS, 2021). The computer can recognise emotional states from people's faces the same way as humans; consequently, facial expression signs are often used in automatic systems (Samadiani et al., 2019). FER analysis has three steps which include (i) face detection, (ii) detection of expression, and (iii) classification of expression to emotional state. Emotion detection is based on studying facial landmark positions such as eyebrows, nose shape, lips position etc. Additionally, changes in video positions are also analysed in videos to identify shrinkages in a cluster of facial muscles (EDPS, 2021).

The remaining part of this paper consists of "the methodology", which has the method that was used to achieve the aim and objectives of this review, "The search process", which shows the process by which research papers were searched and retrieved for this research, "result and discussion" which shows and discuss the state-of-the-art artificial intelligence techniques that are being used for face detection, face recognition, and emotion detection as well as the state-of-the-art dataset that is being used for emotion detection, and finally "the conclusion" which shows the summary of the work that has been done in this research.

METHODOLOGY

The method used in this research is a systematic literature review method, which is known to be in-depth and reliable (Kitchenham et al., 2009). This systematic literature review guideline was proposed by (Jokonowo et al., 2018). The method involves formulating some research questions and then finding the solutions to the prepared questions by reviewing and analysing relevant studies based on the inclusion and exclusion criteria.

Inclusion criteria

Papers written in the English language and peer/blind-reviewed journals, workshops and conferences between the year 2018 and 2023 were included. Additionally, these papers must be relevant to artificial intelligence, face recognition, face detection and emotion detection to qualify for inclusion.

Exclusion criteria

Papers not written in English and unrelated to face recognition, face detection or emotion detection were excluded. Reports whose contributions to knowledge were not stated explicitly in the abstract were all excluded from the papers reviewed in this work.

Search string

We constructed a search phrase using keywords based on the stated research questions in section 4 to create a good search string. The databases presented in Table 1 were used to carry out our searches. The search strings are "Facial affective systems", "face recognition", "face detection", "emotion detection", "machine learning algorithms for face recognition", "machine learning algorithms for face detection", "deep learning algorithms for face recognition", "deep learning algorithms for face detection", and "review on facial affective systems".

Table 1. **Academic database**

S/N	Database	Publisher	Link
1	Springer Link	Springer	https://link.springer.com/
2	IEEE Explore	IEEE	https://ieeexplore.ieee.org/Xplore/home/
3	Web of Science	Clarivate	https://www.sciencedirect.com/
4	Scopus	Elsevier	https://www.scopus.com/home/

Search process

1,267 papers were obtained from the database using the search strings; nevertheless, many were unrelated to the research, so the documents were first separated based on abstract and titles of papers, and 683 articles were removed, leaving 584 pieces. Furthermore, 38 relevant reports were extracted based on the inclusion and exclusion criteria.

Research questions

Three research questions were formed, and the answers to these questions were provided in this paper's result and discussion section. The research questions are;

RQ1: What AI techniques are being used for face detection, and how well have these techniques performed?

RQ2: What AI techniques are used for face recognition, and how well have these techniques performed?

RQ3: What AI techniques are being used for emotion detection, and how well have these techniques performed?

RQ4: What are the state-of-the-art dataset for emotion detection?

This literature review will answer these questions in the result and discussion part.

RESULT AND DISCUSSION

Answers to RQ1

AI techniques used for Face Detection

Gretchen et al. (2018) developed a system to track humans by detecting the head and face in real time. The researchers used computer vision techniques with the OpenCV library; the Python programming language was used for the coding. The archers' dataset contains images from webcam and Closed-Circuit Television (CCTV). The algorithms used for the implementation are the Haar-like classifier to detect the head and Consensus-based Matching and Tracking of keypoint for the object (CMT) object-tracking algorithm to track the head. During the system evaluation, the CMT that follows the charge was 68% accurate, while the Haar-like classifier that detects the human head was 83% accurate.

Lee et al. (2018) researched to solve face detection issues in different lighting states. The study creates an intelligent and effective face detection system in another lightning form. OpenCV and Visual Studio2015 were used for their implementation. The model was able to achieve an accuracy of 80%.

A framework that uses the Viola and Jones algorithm on employee facial database to identify personal employee statements was developed by Sriratana et al. (2018); the researchers used the OpenCV library, and the face detection model that was created was able to achieve an accuracy of 90%.

Gorbatsevich et al. (2019) worked on a system that detects human faces in 2019. The researchers adopted the Convolutiona Neural Network (CNN) based algorithm for face detection purposes, and a dataset called the WIDER FACE dataset was used to train the face detection model. The CNN algorithm was able to achieve an accuracy of 62.9%.

Seunghyun et al. (2019) proposed an approach to detect faces in real-time video. The study makes use of the Boundary-Salience-Multi-Branch Network (BSMNet), the Multi-Task-Cascaded Convolutional Neural Network (MTCNN) algorithms to see smaller fronts and the You-Only-Look-Once (YOLO) algorithm to detect larger faces. From the study, it was observed that optimisation of speed and accuracy is possible if some of the hidden layers in the neural network are removed.

A model that uses Red-Green-Blue (RGB) approach for facial and fundamental object recognition was developed by Ferdousi et al. (2019). The researchers used the YOLO and MTCNN algorithms for object detection and facial recognition. The OpenCV library and the researchers used datasets locally curated from the environment. An accuracy of 63-80% and 80-100% was achieved by the YOLO and the MTCNN algorithms, respectively.

Mehariya et al. (2020) developed a system that can take attendance of students attending a class or workers at work by detecting their faces. The researchers trained the model using images of students collected from a course as a dataset. The OpenCV, occupancy ratio, and see Multiscale techniques were adopted for the research. The researchers used the Haar-cascade algorithm to train the face detection model, which achieved an accuracy ranging between 90% and 100%.

Gupta et al. (2020) proposed a system based on image processing that can help to improve attendance systems in higher institutions. The system can also reduce the wastage rate when taking attendance. The attendance system aims to manage attendance students by estimating the times a student is present and the number of times the student is absent, then automatically producing a document that shows the report. The OpenCV library was used for the implementation. The Haar-cascade algorithm was used for face detection, and the LBPH was used for face recognition. From the evaluation, the model achieved an accuracy of 97%.

A system that can identify faces from video streams and images was developed by Hoque et al. (2020). The ATMega328p Microcontroller and the OpenCV were used for image processing. The researcher trained the model with an open-source dataset from Kaggle. The researchers used AdaBoost, Haar-cascade CAmshift, Viola Jonas and Housedorff Gap to identify human faces. And the Haar-cascade was used to classify faces. The face classification model was 83% accurate.

A system that detects a human face before extracting and using features from the detected face to recognise the beginning was developed by Singh & Jasmine (2019). The viola-jones detection algorithm, which uses the AdaBoost learning algorithm and integral image as classifiers, was used alongside the haar cascade and KLT algorithm for face detection. The viola-jones performed best in different lighting conditions compared to other algorithms used for face detection.

Table 2. **Summary of techniques used for face detection**

Papers	Techniques	Performance
Gretchen et al. (2018)	Haar-like classifier to detect the head and CMT object-tracking algorithm to track the head	CMT that tracks the head was 68% accurate, while the Haar-like classifier that detects the human's lead was 83% accurate
Lee et al. (2018)	Haar cascade algorithm	The model was able to achieve an accuracy of 80%.
Sriratana et al. (2018)	Viola and Jones's algorithm	The face detection model that was developed achieved an accuracy of 90%.
Gorbatsevich et al. (2019)	CNN algorithm	The CNN algorithm was able to achieve an accuracy of 62.9%
Lee et al. (2019)	BSMNet and MTCNN algorithms to detect smaller faces and the Yolo algorithm to detect larger faces	Not stated
Ferdousi et al. (2019)	YOLO and MTCNN algorithms for face detection and facial recognition, respectively	An accuracy of 63-80% and 80-100% was achieved by the YOLO and the MTCNN algorithms, respectively.
Mehariya et al. (2020)	Haar-cascade algorithm	The algorithm achieved an accuracy that ranges between 90% and 100%.
Gupta et al. (2020)	The Haar-cascade algorithm was used for the face detection process	The model achieved an accuracy of 97%.
Hoque et al. (2020)	Haar-cascade algorithm	The model was 83% accurate
Singh & Jasmine (2019)	Haar-cascade and KLT	Not stated

Based on our reviewed literature, we observed that the major techniques used for face detection are machine learning and deep learning techniques, which includes Haar-cascade, KLT, YOLO, BSMNet, Viola and Jones, CMT object-tracking, and MTCNN algorithms. The haar-cascade algorithm, which is a machine learning algorithm used to identify objects in images and video, is mostly used for face detection and has achieved high accuracy that ranges from 83% to 97% as used by researchers like

(Hoque et al. 2020; Gupta et al., 2020; Mehariya et al., 2020; Lee et al., 2018; Alcantara et al., 2018). The BSMNet, MTCNN, and YOLO algorithms have also been used for face detection and performed well by researchers like Gorbatsevich et al. (2019) and Lee et al. (2019). Other researchers like Ferdousi et al. (2019) used nonparametric algorithms for face detection.

Answer to RQ2

AI techniques used for face recognition

An attendance record system based on face recognition for Nigerian universities was proposed by Ogbuju et al. (2020). The method uses the Faster Region Convolutional Neural Network (FRCNN) algorithm for implantation. Google Colab was used to design and implement the model. The dataset used to train the model was collected locally by capturing the faces of students using the camera. The developed system, called "DLFacer, " has 97-99% accuracy.

A system that recognises human faces based on a computational-intelligent algorithm was developed by AbdELminaam et al. (2020). The researchers made use of. The model was created with the help of Deep Neural Networks (DNN). The result from the proposed method was compared with three different machine learning algorithms, which are Decision Tree (DT), K-Nearest Neighbor (KNN) and the Support Vector Machine (SVM). The DNN performed better than all the compared algorithms with the highest accuracy (99.06%), the highest recall (99.07), the highest precision (99.12%) and the highest specificity (99.10%).

A study on the application of facial recognition-based algorithms on examination officers' security systems at Omar Al-Mukhtar University was carried out by Bazama et al. (2021). The Independent Component Analysis (ICA) was utilised for face recognition. The data used to train the model consists of 100 images of people's faces at the school's College of Science. The ICA algorithm was able to achieve 86.7%.

A deep learning method was used for face identification and face recognition by Teoh et al. (2020). The CNN algorithm was used to train the face detection model. The researchers also utilised the use of the OpenCV library in Python. The model was able to achieve 91.7% in face recognition.

Wahana et al. (2020) designed a face recognition system that can be used for indoor applications. The face recognition process used the Local-None-Negative-Matrix-Factorization (LNMF) and the none-negative-matrix-factorisation-suppressed-carrier (NMFsc) methods. Videos of classroom students with 640X480 pixels were captured and used to train the face recognition model. the LNMF algorithm was 71.61% accurate with a computation time 152,634. In comparison, the NMFsc algorithm achieved an accuracy of 86.76% with a computation time of 152,785.

Singh & Jasmine (2019) introduced a system that automatically recognises people's faces. The system works by first detecting an individual's face, extracting the future from the face that has been seen and recognising the face. The Principal Component Analysis (PCA) algorithm was used for face recognition purposes, and the model outperformed other compared algorithms used in face recognition. The PCA algorithm achieved 77% accuracy.

A paper that proposes a robust and efficient method for face recognition in real-time was submitted by Vinay et al. (2018). The proposed method was based on Multilayer Perception (MLP), which uses the gradient-descent algorithm. The MLP was fed in with binary patterns extracted from the images. The method was tested on four different datasets, including the Grimace, FACES96, FACES95 and FACES94 datasets. The technique was able to achieve an accuracy of 91%.

Onyema et al. (2021) Presented a study on a deep learning-based method for facial recognition. The deep learning method presented for the research implementation is the Convolutional Neural Network (ConvNet). The dataset used for the implementation is the FER2013 dataset, which was gathered automatically using image search API from Google. The proposed model was able to achieve an accuracy of 70%.

An attendance system based on face recognition was developed by Bhagat et al. (2021). The system works by discovering and recognising faces from the classroom's live stream footage and then sending it to the appropriate faculty. The researchers generated a dataset from photos of kids in a class. The skin-splitting approach was used. The system was able to reach an accuracy of 82.3%

The research by Gupta et al. (2020) to improve the attendance system in higherm institutions also used the Local Binary Patterns Histogram (LBPH) for the face recognition process. The research used The OpenCV library for the implementation. The LBPH achieved an accuracy of 97%.

Table 3. **Summary of techniques used for face recognition**

Papers	Techniques	Performance
Ogbuju et al. (2020)	Faster Region Convolutional Neural Network (FRCNN)	The model has an accuracy of 97-99%
AbdELminaam et al. (2020)	deep-convolutional-neural-networks (DCNN)	The DCNN has an accuracy of 99.06%
Bazama et al. (2021)	The independent component analysis (ICA)	The ICA algorithm was able to achieve 86.7%.
Teoh et al. (2020)	Convolutional Neural Network (CNN)	The model was able to achieve 91.7% accuracy
Wahana et al. (2020)	The local-none-negative-matrix-factorization (LNMF) and the none-negative-matrix-factorization-suppressed-carrier (NMFsc)	The LNMF algorithm was 71.61% accurate, while the NMFsc algorithm achieved an accuracy of 86.76%.
Singh & Jasmine (2019)	PCA algorithm	The PCA algorithm achieved 77% accuracy
Vinay et al. (2018)	gradient-descent algorithm	The method was able to achieve an accuracy of 91%
Onyema et al. (2021)	convolutional-neural-network (ConvNet)	The model was able to achieve an accuracy of 70%.
Bhagat et al. (2021)	The skin-spliting approach	The model achieved 82.3%
Gupta et al. (2020)	LBPH algorithm	The model achieved 97% accuracy

From our reviewed paper on the commonly used techniques for face detection, we notice that CNN, RCNN, ICA, LNMF, NMFsc, PCA, ConvNet, gradient-descent, LBPH, and skin splitting algorithms have recently been used by researchers like (Bazama et al., 2021; Teoh et al., 2020; Wahana et al., 2020; Singh & Jasmine 2019; Vinay et al., 2018; Onyema et al., 2021; Bhagat et al., 2021; Gupta et al., 2020). It is also observed that most of these algorithms used in the reviewed papers are related to the deep learning CNN algorithm and have performed excellently when used for face detection purposes. The CNN-based algorithms have achieved high accuracies of 91.7% and 99.06%, except for the ConvNet, which earned a lesser accuracy of 70% due to the limited dataset the researchers used.

Answers to RQ3

AI techniques used for emotion detection

Ogbuju et al. (2021) presented research that detects emotions from online news/text. The authors used twelve (12) different online news channels in Nigeria as datasets to train the proposed model and a hybrid method of corpus-based lexicon and dictionary approaches to achieve the research aim.

Gupta (2018) proposed a method used to detect facial emotion. The detect emotions on faces in real-time; after the feeling has been seen, it will now be categorised according to the emotions category using the Support Vector Machine (SVM) algorithm. During the evaluation that was done on the model, it was observed that the SVM was able to achieve an accuracy of 93.7%. The researchers also state in their report that the facial landmarks may be tweaked to increase the model's accuracy.

Patel et al. (2018) developed a model that can detect a driver sleeping when driving based on facial expression, the research aimed to prevent road accidents. The researchers used the OpenCV library. Also, a vision device that monitors the face to detect the driver's facial emotion and eye blinking was used for the implementation. Artificial Neural Network (ANN) algorithm was used for the modelling, and at the evaluation stage, it was discovered that the algorithm was 90% accurate.

Jain et al. (2018) used a hybrid of CNN, and Recurrent Neural Networks (RNN) were used for face emotion detection since Deep Neural Networks (DNN) always perform better than traditional models when compared by researchers. Two publicly available datasets were used to train the hybrid of deep learning algorithms. The hybrid model was 90% accurate.

Benamara et al. (2019) developed a facial emotion detection system; the researchers used an ensemble of YOLO and CNN algorithms to train the emotion detection model. the model was evaluated in real-time, and the ensemble model attained an accuracy of 72.47%.

A real-time system that monitors elders' living conditions in case of emergency was developed by Punidha et al. (2020). The system monitors the elders and sends messages to the relatives to alert them during an emergency. The researcher used a locally curated dataset. The CNN algorithm was used for the implementation purpose, although the research's accuracy level was not stated in the report.

A study based on emotion recognition of speech and graphic visualisation of expression in an intelligent learning environment was carried out by Lu (2021). The model was trained using the Convolutional Neural Network Bi-Directional Long Short-Term Memory (CNN-BiLSTM) algorithm. The researcher conducted a simulation experiment to check the algorithm's performance. The result shows that the CNN-BiLSTM was 98.75% accurate.

Farhoumandi et al. (2021) tried to predict alexithymia using a facial emotion recognition model. Several students from the University of Tabriz were selected based on their scores on the Toronto Alexithymia scale, which was used as a dataset. The facial expression model was developed using SVM and Feed Forward Neural Network (FFNN). An accuracy ranging between 72.7 to 81.8 was obtained after evaluating the model's performance using the area under curve (AUC).

Research presented a real-time method for implementing a model that detects emission and deploying it in a robot conducted by Siam et al. (2022). Three datasets were used to develop the model: the Cohn-Kande dataset, which contains 593 video sequences; the JAFFE dataset, which contains facial expressions of female Japanese; and a dataset from the real-world face effective database. Machine learning algorithms such as SVM, KNN, Naïve Bayes (NB), Linear Regression (LR), Random Forest (RF) and MLP were used to train the model. At the end of the training, only MLP was deployed because it outperformed all the algorithms used. The MLP has the highest accuracy of 97%.

A facial emotion recognition model that can recognise emotions such as happiness, anger, neutrality, surprise, sadness, fear and disgust was proposed by Debnath et al. (2022). A deep learning algorithm called ConvNet was used to develop the model. The ConvNet algorithm used four layers for facial emotion detection. A dataset known as the JAFFE dataset was used to train the model. The ConvNet was able to achieve 98.13% accuracy.

Singh et al. (2022) Developed deep learning-based facial emotion detection using the CNN algorithm. The paper aims to detect emotions in humans and extract human behaviour. This model successfully saw emotions such as anger, happiness, fear, calmness, sadness, surprisingness and disgust. The model has an accuracy of 90-100.

Table 4. **Summary of techniques used for emotion detection**

Authors	Techniques	Performance
Ogbuju et al. (2021)	corpus-based lexicon and dictionary approaches	Accuracies from 90-99%
Gupta (2018)	SVM algorithm	The algorithm achieved an accuracy of 93.7%.
Patel et al. (2018)	ANN algorithm	The algorithm was 90% accurate
Jain et al. (2018)	A hybrid of CNN and RNN	The hybrid model was 90% accurate
Benamara et al. (2019)	An essemble of YOLO and CNN algorithm	The ensemble model was 72.47% accurate
Punidha et al. (2020)	CNN algorithm	The model achieved 92% accuracy.
Lu (2021)	The CNN-BiLSTM algorithm	The CNN-BiLSTM was 98.75% accurate
Farhoumandi et al. (2021)	SVM and FFNN algorithms	An accuracy ranging from 72.7 to 81.8 was obtained after evaluating the model's performance using AUC.
Siam et al. (2022)	Machine learning algorithms such as SVM, KNN, NB, LR, RF and MLP were used to train the model.	MLP has performed best, with the highest accuracy of 97%.
Debnath et al. (2022)	Four-layer ConvNet deep learning algorithm	The ConvNet was able to achieve 98.13% accuracy
Singh et al. (2022)	CNN algorithm	The model has an accuracy of 90-97%

The techniques being used for emotion detection are ANN CNN, CovNet, FFNN, MLP, CNN-BiLSTM, YOLO, assemble of YOLO and CNN, Hybrid of CNN and RNN, SVM, KNN, NB, LR, and RF. We have also observed that many researchers have employed machine learning and deep learning techniques for emotion detection purposes. As reviewed in this research, one of the algorithms that were being used is the SVM which was used by researchers like (Gupta 2018; Farhoumandi et al., 2021; Siam et al., 2022) went ahead and compared the SVM with other algorithms which include KNN, NB, LR, RF and a multilayer perceptron (MLP) algorithm, eventually the MLP outperform the compared machine learning algorithm. Other algorithms used are deep learning algorithms which include the CNN algorithm, which was used by (Punidha et al., 2020; Singh et al.2022). The ANN algorithms were used by Patel et al. (2018), a four-layer ConvNet deep learning algorithm used by Debnath et al. (2022), CNN-BiLSTM used by Lu (2021), a hybrid of CNN and RNN used by Jain et al. (2018) and an ensemble of YOLO and CNN algorithm which was used by Benamara et al. (2019). From the reviewed literature, we observed that the CNN-BiLSTM performed the best compared to other studied algorithms in this paper and achieved an accuracy of 98.75%. The deep learning four layers ConvNet also achieved excellently with an accuracy of 98.13% which is very close to the performance of the CNN-BiLSTM. Other algorithms, like CNN, RNN, a hybrid of CNN and RNN, and the YOLO and CNN algorithm ensemble, have performed well in the reviewed papers.

Answers to RQ4

State-of-the-art emotion detection datasets.

- **Kaggle emotion detection dataset**: This is an emotion detection dataset from Kaggle, which contain about 35,685 images of faces in greyscale. The dataset is already divided into train data and test data. The images in the dataset are categorised into seven different emotions: happiness, anger, sadness, surprise, fear, disgust and neutral (Ares, 2021).
- **AffectNet**: This dataset is one of the largest facial emotion detection datasets. The dataset contains over 1 million images of different faces, which are collected from other sources from the Internet using about 1250 keywords related to emotion detection in six languages. The languages include English, Arabic, Portuguese, Farsi, German and Spanish (Mahoor, 2023).
- **Ascertain**: This dataset was introduced by Subramanian et al. (2018). The Ascertain database is a multimodal database for affect recognition and implicit personality using commercial sensors. Ascertain dataset has 58 users' emotional self-ratings with synchronously recorded Electrocardiogram (ECG), Electroencephalogram (EEG), facial activity data and Galvanic Skin Response, which were recorded using an off-the-shelf sensor.
- **Dreamer**: This is a multimodal dataset containing ECG and EEG signals documented during effect elicitation through audio-visual stimuli. The dataset includes calls from 23 participants and their emotional state after each stimulation (Katsigiannis & Ramzan, 2018).
- **EMOTIC Dataset**: This dataset contains real-time images of people with different emotions. The dataset images are annotated with 26 categories of emotions with three continuous dimensions: Arousal, Valence and Dominance. The dataset contains about 23,571 photos, and some ideas are collected manually from the Internet using the Google search engine (Kosti et al., 2019).

- **Google Facial Expression Comparison Dataset**: This dataset contains face images of people with different facial expressions; the dataset set was gathered to help researchers on topics that are related to facial expression analysis like expression-based photo album summarisation, expression-based picture retrieval, classification of emotion, expression synthesis etc. (Vemulapalli & Agarwala, 2018).
- **K-EmoCon**: This dataset is a multimodal dataset with comprehensive annotation of continuous emotions during a naturalistic conversation. The dataset contains multimodal measurements, including ECG, audio visual recordings and peripheral physiological signals, acquired during about 16 different 10-minute debates based on social issues (Park et al., 2020).
- **CLASS dataset**: This is a multimodal dataset that was majorly developed to support researchers consisting of synchronous recording of ECG, Plethysmography (PPG), and ElectroDermal Activities (EDA) of 62 volunteers (Markova et al., 2020).
- **JAFFE dataset**: The JAFFE dataset, which stands for the Japanese Female Facial Expression dataset, contains 213 images that show the facial expressions of ten different volunteer female Japanese, each of the female Japanese leading seven other facial expressions (Lyons et al., 2020).

Table 5. **Summary of state of the art emotion detection datasets**

S/N	Authors	Dataset	Size	Category	Description
1	Ares (2021)	Kaggle emotion detection dataset	35,685	Seven categories of emotions	images of faces in greyscale
2	Mahoor (2023)	AffectNet dataset	1,000,000	Seven categories of emotions	images of different faces are collected from the Internet
3	Subramannian et al. (2018)	Ascertain dataset	58	Five categories of emotions	synchronously recorded ECG, EEG, facial activity data and Galvanic skin response
4	Katsigiannis & Ramzan (2018)	Dreamer dataset	23	Nine categories of emotions	ECG and EEG signals that were documented during effect elicitation through audio-visual-stimuli
5	Kosti et al. (2019)	EMOTIC dataset	23,571	26 categories of emotions	images of people collected manually from the Internet.
6	Vemulapalli & Agarwala (2018)	Google Facial Expression Comparison dataset	51,060	Three categories of emotions	recorded ECG, EEG, facial activity
7	Park et al. (2020)	K-EmoCon dataset	20	Nine categories of emotions	Measurement of ECG, audio-visual recordings and peripheral physiological signals
8	Markova,] et al. (2020)	CLAS dataset	62	Five categories of emotions	Synchronously recorded ECG, PPG, and EDA.
9	Lyons et al. (2020)	JAFFE dataset	213	Seven categories of emotions	Collection of images of female Japanese facial expressions

The state-of-the-art emotion dataset is Kaggle emotion detection, AffectNet, Ascertain, Dreamer, EMOTIC, Google Facial Expression Comparison, K-EmoCon, CLAS, and JAFFE datasets. Most of these dataset contains synchronously recorded ECG, EEG, EDA, facial activity data and Galvanic skin Response with different emotion categories. In contrast, others consist of facial images of people with different emotional states.

CONCLUSION

This paper has presented the AI techniques for facial affective computing by answering some research questions. Firstly the report showed the methods of face detection based on the reviewed paper to be Haar-cascade, KLT, YOLO, BSMNet, Viola and Jones, CMT object-tracking, and MTCNN algorithms. Secondly, papers on face recognition were reviewed, and the recently used techniques were found to be CNN, RCNN, ICA, LNMF, NMFsc, PCA, ConvNet, gradient-descent, LBPH and the skin splitting algorithms. Thirdly, we reviewed a paper on emotion detection and found out that ANN CNN, CovNet, FFNN, MLP, CNN-BiLSTM, YOLO, assemble of YOLO and CNN, Hybrid of CNN and RNN, SVM, KNN, NB, LR, and RF have been recently utilised for emotion detection. Finally, state of the art emotion detection dataset has been reviewed in this paper. The reviewed emotion detection datasets are Kaggle, AffectNet, Ascertain, Dreamer, EMOTIC, Google Facial Expression Comparison, K-EmoCon, CLAS, and JAFFE datasets. We observed from the checked state-of-the-art dataset that most datasets are gathered from different online platforms, and others contain images from foreign countries. For these reasons, we recommend a localised dataset that contains images of emotions from people from the locality.

RECOMMENDATION

It is observed that the haar-cascade algorithm performed all the algorithms used for face detection in the reviewed papers in terms of accuracy. Nevertheless, other deep learning algorithms like CNN will perform well for face detection. Hence, to improve accuracy and precision we recommend a hybrid of haar-cascade and CNN algorithms for face detection. Secondly, out of the recently used algorithms for face recognition, we observed that the CNN-based algorithms have also performed excellently as used by the researchers. However, other algorithms like the LBPH and gradient-descent algorithms have proven effective in recognising faces. Thus we recommend a hybrid of the CNN algorithm with either the LBPH algorithm or the gradient-descent algorithm. Also, based on our reviewed paper on emotion detection, we recommend neural network algorithms with multiple layers for emotion detection.

REFERENCES

AbdELminaam, D. S., Almansori, A. M., Taha, M., & Badr, E. (2020). A deep facial recognition system using computational intelligent algorithms. *PLoS ONE*, *15*(12), 1-27. https://doi.org/10.1371/journal.pone.0242269.

Alcantara, G. K. L., Evangelista, I. D. J., Malinao, J. V. B., Ong, O. B., Rivera, R. S. D., & Ambata, E. L. U. (2018). Head detection and tracking using OpenCV. In *2018 IEEE 10th*

International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM), Baguio City, Philippines, 2018 (pp. 1-5). https://doi.org/10.1109/HNICEM.2018.8666374.

Ares. (2021). *Emotion Detection*. Retrieved from: https://www.kaggle.com/datasets/ananthu017/emotion-detection-fer.

Bazama, A., Mansur, F., & Alsharef, N. (2021). Security System by face recognition. *AlQalam Journal of Medical and Applied Science*, *4*(2), 58-67. Retrieved from: https://www.ajol.info/index.php/ajmas/article/view/218189.

Benamara, N. K., Val-Calvo, M., Álvarez-Sánchez, J. R., Díaz-Morcillo, A., Vicente, J. M. F., Fernández-Jover, E., & Stambouli, T. B. (2019). Real-Time Emotional Recognition for Sociable Robotics Based on Deep Neural Networks Ensemble. In J. Ferrández Vicente, J. Álvarez-Sánchez, F. de la Paz López, J. Toledo Moreo, H. Adeli (Eds.), *Understanding the Brain Function and Emotions. IWINAC 2019. Lecture Notes in Computer Science*, vol. 11486 (pp. 171–180). Springer, Cham. https://doi.org/10.1007/978-3-030-19591-5_18.

Bhagat, S., Kashkari, V., Srivastava, S., & Sharma, A. (2022). Face recognition attendance system. *International Journal for Research in Applied Science and Engineering Technology*, *10*(1), 280-283. https://doi.org/10.22214/ijraset.2022.39702.

Calvo, R., D'Mello, S., & Gratch, J. (2014). *The Oxford handbook of affective computing*. UK: Oxford University Press.

Debnath, T., Reza, M., Rahman, A., Beheshti, A., Band, S. S., & Alinejad-Rokny, H. (2022). Four-layer ConvNet to facial emotion recognition with minimal epochs and the significance of data diversity. *Scientific reports*, *12*, 6991. https://doi.org/10.1038/s41598-022-11173-0.

EDPS. (2021). EDPS TechDispatch: Facial emotion recognition. *Publication office of the European Union*, *1*(2021), 1-5. Retrieved from: https://edps.europa.eu/data-protection/our-work/publications/techdispatch/techdispatch-12021-facial-emotion-recognition_en.

Farhoumandi, N., Mollaey, S., Heysieattalab, S., Zarean, M., & Eyvazpour, R. (2021). Facial emotion recognition predicts alexithymia using machine learning. *Computational Intelligence and Neuroscience*, *2021*(1), 2053795. https://doi.org/10.1155/2021/2053795.

Ferdousi, R., Israt, J. R., Nafisa, F., & Jia, U. (2019). An assistive model for visually impaired people using YOLO and MTCNN. In *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy (ICCSP '19)*. Association for Computing Machinery, New York, NY, USA (pp. 225–230). https://doi.org/10.1145/3309074.3309114.

Flynn, M., Effraimidis, D., Angelopoulou, A., Kapetanios, E., Williams, D., Hemanth, J., & Towell, T. (2020). Assessing the effectiveness of automated emotion recognition in adults and children for clinical investigation. *Frontiers in Human Neuroscience*, *14*(1), 1-12. https://doi.org/10.3389/fnhum.2020.00070.

Gaudenz, B. (2022). *Application of deep learning in computer vision*. Retrieved from: https://viso.ai/applications/computer-vision-in-healthcare/.

Gorbatsevich, V. S., Moiseenko, A. S., & Vizilter, Yu. V. (2019). Face DetectNet: Face detection via fully-convolutional network. *Computer Optics*, *43*(1), 63-71. https://doi.org/10.18287/2412-6179-2019-43-1-63-71.

Gu, S., Wang, F., Patel, N. P., Bourgeois, J. A., & Huang, J. H. (2019). A model for basic emotions using observations of behavior in Drosophila. *Frontiers in Psychology*, *10*, 781. https://doi.org/10.3389/fpsyg.2019.00781.

Gupta, N., Sharma, P., Deep, V., & Shukla, V. K. (2020). Automated attendance system using OpenCV. In *2020 8th International Conference on Reliability, Infocom Technologies and Optimisation (Trends and Future Directions) (ICRITO)* (pp. 1226-1230). IEEE. https://doi.org/10.1109/ICRITO48877.2020.9197936.

Gupta, S. (2018). Facial emotion recognition in real-time and static images. In *2018 2nd international conference on inventive systems and control (ICISC)* (pp. 553-560). IEEE. https://doi.org/10.1109/ICISC.2018.8398861.

Hoque, M. A., Islam, T., Ahmed, T., & Amin, A. (2020). Autonomous face detection system from real-time video streaming for ensuring the intelligence security system. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 261-265). IEEE. https://doi.org/10.1109/ICACCS48705.2020.9074260.

Jain, N., Kumar, S., Kumar, A., Shamsolmoali, P., & Zareapoor, M. (2018). Hybrid deep neural networks for face emotion recognition. *Pattern Recognition Letters*, *115*, 101-106. https://doi.org/10.1016/j.patrec.2018.04.010.

Jokonowo, B., Claes, J., Sarno, R., & Rochimah, S. (2018). Process mining in supply chains: A systematic literature review. *International Journal of Electrical and Computer Engineering*, *8*(6), 30-41. http://doi.org/10.11591/ijece.v8i6.pp4626-4636.

Katsigiannis, S. & Ramzan, N. (2018). Dreamer: A database for emotion recognation through EEG and ECG signals from wireless Low-cost off-the-shelf devices. *IEEE Journal of Biomedical and Health Informatics*, *22*(1), 98-107. https://doi.org/10.1109/JBHI.2017.2688239.

Kitchenham, B., Pearl, B. O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, *51*(1). 7-15. https://doi.org/10.1016/j.infsof.2008.09.009.

Kosti, R., Alvarez, J. M., Recasens, A., & Lapedriza, A. (2019). Context based emotion recognition using emotic dataset. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(11), 2755-2766. https://doi.org/10.1109/TPAMI.2019.2916866.

Lee, H.-W., Peng, F.-F., Lee, X.-Y., Dai, H.-N., & Zhu, Y. (2018). Research on face detection under different lighting. In *IEEE International Conference on Applied System Invention (ICASI)* (pp. 1145-1148). IEEE. https://doi.org/10.1109/ICASI.2018.8394486.

Lee, S., Kim, M., & Joe, I. (2019). SGNet: Design of Optimized DCNN for Real-Time Face Detection. In J. Park, H. Shen, Y. Sung, & H. Tian (Eds.), *Parallel and Distributed Computing, Applications and Technologies. PDCAT 2018. Communications in Computer and Information Science*, vol. 931. Springer, Singapore. https://doi.org/10.1007/978-981-13-5907-1_21.

Lu, X. (2021). Deep learning based emotion recognition and visualisation of figural representation. *Frontiers in Psychology*, *12*, 818833. https://doi.org/10.3389/fpsyg.2021.818833.

Lyons, M., Akamatsu, S., Kamachi, M., & Gyoba, J. (2020). Coding Facial Expression with Gabor Wavelets. In *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, 1998 (pp. 200-205). IEEE. https://doi.org/10.1109/AFGR.1998.670949.

Mahoor, M. H. (2023). *AffectNet*. Retrieved from: http://mohammadmahoor.com/affectnet/.

Markova, V. I., Ganchev, T., & Kalinkov, K. (2020). CLAS: A database for cognitive load affect and stress recognition. In *2019 International Conference on Biomedical Innovations and Applications (BIA)*, Varna, Bulgaria (pp. 1-4). IEEE. https://doi.org/10.1109/BIA48344.2019.8967457.

Mehariya, J., Gupta, C., Pai, N., Koul, S., & Gadakh, P. (2020). Counting students using OpenCV and integration with firebase for classroom allocation. In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, 2020 (pp. 624-629). IEEE. https://doi.org/10.1109/ICESC48915.2020.9155825.

Ogbuju, E., Adetayo, A. P., & Oblikwu, P. (2020). A face recognition system for attendance record in a Nigeria university. *Journal of Scientific Research and Development*, *19*(2), 38-45.

Ogbuju, E., Yemi-Peters, V., Osumah, T., Agbogun, J., & Ejiofor, V. (2021). A frontpage online news analysis: an opinion mining approach. *Journal of Computer Science and its applications*, *28*(1), 107-120. https://doi.org/10.4314/jcsia.v28i1.9.

Onyema, E. M., Shukla, K. P., Dalal, S., Mathur, N. M., Zakariah, M., & Tiwari, B. (2021). Enhancement of patient facial recognition through deep learning algorithm: ConvNet. *Journal of Healthcare Engineering*, *2021*, 5196000. https://doi.org/10.1155/2021/5196000.

Park, C. Y., Cha, N., Kang, S., Kim, A., Khandoker, H., Hadijileontiadis, L., Oh, A., Jeong, Y., & Lee, U. (2020). K-EmoCon, a multimodal sensor dataset for continuous emotion recognition in naturalistic conversation. *Scientific Data*, *7*, 293. https://doi.org/10.1038/s41597-020-00630-y.

Patel, R., Patel, M., & Patel, J. (2018). Real time somnolence detection system in OpenCV environment for drivers. In *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)* (pp. 407-410). IEEE. https://doi.org/10.1109/ICICCT.2018.8473234.

Punidha, A., Inba, S., Pavithra, K. S., Ameer Shathali, M., & Athibarasakthi, M. (2020). Human emotion detection using machine learning techniques. *SSRN Electronic Journal*, 1-7. https://doi.org/10.2139/ssrn.3591060.

Saim, A. I., Soliman, N. F., Algarni, A. D., Abd El-samie, F. E., Sedik, A., & Ding, B. Y. (2022). Deploying machine learning techniques for human emotion detection. *Computational Intelligence and Neuroscience*, *2022*, 1-16. https://doi.org/10.1155/2022/8032673.

Samadiani, N., Huang, G., Cai, B., Luo, W., Chi, C.-H., Xiang, Y., & He, J. (2019). A review on automatic facial expression recognition systems assisted by multimodal sensor data. *Sensors*, *19*(8), 1863. https://doi.org/10.3390/s19081863.

Singh, S. K., Thakur, K. R., & Anand, R. (2022). Deep learning and machine learning based facial emotion detection using CNN. In *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2022 (pp. 530-535). IEEE. https://doi.org/10.23919/INDIACom54597.2022.9763165.

Singh, S. & Jasmine, G. (2019). Face recognition system. *International Journal of Engineering Research and Technology (IJERT)*, *8*(5), 263-266.

Sriratana, W., Mukma, S., Tammarugwattana, N., & Sirisantisamrid, K. (2018). Application of the OpenCV-Python for personal identifier statement. In *2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICEAST.2018.8434429.

Subramannian, R., Wache, J., Abadi, M. K., Vieriu, L. R., Winkler, S., & Sebe, N. (2018). ASCERTAIN: Emotion and personality recognition using commercial sensors. *IEEE transactions on Affective Computing*, *9*(2), 147-160. https://doi.org/10.1109/TAFFC.2016.2625250.

Teoh, K. H., Ismail, R. C., Naziri, S. Z. M., Hussin, R., Isa, M. N. M., & Basir, M. S. S. M. (2020). Face recognition and identification using deep learning approach. *Journal of Physics: Conference Series*, *1755*, 012006. https://doi.org/10.1088/1742-6596/1755/1/012006.

Vemulapalli, R. & Agarwala, A. (2018). *Google facial expression comparison dataset*. Google research.

Vinay, A., Gupta, A., Bharadwaj, A., Srinivasan, A., Balasubramanya, M., & Natarajan, S. (2018). Deep learning on binary patterns for face recognition. *Procedia Computer Science*, *132*, 76-83. https://doi.org/10.1016/j.procs.2018.05.164.

Wahana, D. G., Hidayat, B., Aulia, S., & Hadiyoso, S. (2020). Face recognition system for indoor applications based on video with LNMF and NMFsc methods. *Journal of Southwest Jiaotong University*, *55*(6), 1-10. https://doi.org/10.35741/issn.0258-2724.55.6.18.

About the authors:

Bashir Eseyin Abdullahi, Federal University Lokoja, Nigeria. bashir.abdullahi-pg@fulokoja.edu.ng

Emeka Ogbuju, Federal University Lokoja, Nigeria. emeka.ogbuju@fulokoja.edu.ng

Taiwo Abiodun, Federal University Lokoja, Nigeria. taiwo.abiodun-pg@fulokoja.edu.ng

Francisca Oladipo, Federal University Lokoja, Nigeria. francisca.oladipo@fulokoja.edu.ng