Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Ensuring Road Safety By Reading Driver's Facial Emotions Using Deep Learning

Sharon Sofia S S, S. Jasmine Mary, Dr. T. C. Subbu Lakshmi

PG scholar, Department of Information Technology, Francis Xavier Engineering College, Tirunelveli ,Tamil Nadu- 627 003

sharonsofia980@gmail.com

Research Scholar, Anna University Chennai.

jasmine.jessijas@gmail.com

Professor, Department of Information Technology, francis Xavier Engineering College. Tirunelveli

subbulakshmiit@francisxavier.ac.in

ARTICLE INFO

Received: 12 Nov 2024

Revised: 23 Dec 2024

Accepted: 20 Jan 2025

ABSTRACT

Researchers studying facial expression detection using deep neural networks have published findings in recent years that demonstrate how these methods get around the drawbacks of traditional machine learning techniques. In this paper, VGG-19 is used to monitor a driver's emotions and to increase its accuracy, which is trained using the similarity of the sample data. Data preparation comes initially, followed by the extraction of geometric characteristics and the detection of facial landmarks in input photos. The performance of the suggested VGG-19 classifier was compared to state-of-the-art techniques after these feature vectors were incorporated into it for facial expression classification. The findings demonstrate that our suggested approach performs comparably to deep learning techniques, which obtained 97.83% accuracy, 96.97% F1-score, 96.5% recall, and 97.73% precision.

Keywords: characteristics, emotions, demonstrate

1. INTRODUCTION

The intelligent car requires human-machine systems that are sensitive to emotions. Emotions have an impact on a driver's performance and are strongly linked to traffic accidents [1]. The number of fatalities from traffic accidents keeps increasing. One of the main causes deteriorating driving safety among these occurrences has been identified as the inability to control one's emotions [2]. Intelligent automobile human-machine systems are increasingly focusing on detecting and recognizing driver emotions.

Although EEG signals offer enhanced time resolution that allows researchers to study affective cues, their practicality in applications such as automobile human-machine systems is limited due to the need for numerous electrodes positioned at different points on the head[3]. Due to changes in sweat gland activity, GSR monitors emotions and stress. GSR doesn't require bulky, heavy equipment as EEG and MRI do; instead, it only requires sensors to be applied on the hands. However, even mild activity can significantly alter the GSR signal and make it useless for detecting driver emotions because drivers frequently move their hands while driving.

Because driving limits bodily movement, it is impractical to use gestures to gauge a driver's emotion. Affective contact also depends on speech, which is a basic human communication mechanism[5]. Feature extraction from unprocessed speech data is necessary for speech-based emotion identification. Its ability to identify highly emotive speech is not very accurate. Speech is insufficiently accurate to be the key pattern when it comes to the construction of the human-machine interface of an emotion-based vehicle. [6]The system cannot track feelings when there is no conversation, which usually happens when driving, because speech is not constant during the driving activity.

Computational models made up of several processing layers can learn data representations at various degrees of abstraction thanks to deep learning (DL). [7] Contemporary object identification, pattern recognition, speech recognition, and many other fields were significantly enhanced by these techniques. By employing the backpropagation approach to specify how a model should update its internal parameters to compute each layer's representation, DL reveals complex pattern structure in huge datasets. from the representation of the previous layer. [8]The enormous volume of data needed to optimize the model is one difficulty with DL models. Various methods use transfer-learning to solve this problem without requiring large datasets.

This research proposes a novel deep learning-based method for detecting drivers' emotions on the road from start to finish. To get around the previously mentioned dataset restriction, this study collected a dataset of driver facial expressions. The suggested approach can significantly increase driving safety by recognizing drivers' emotions. The influence of emotion on driving behavior has been the subject of numerous research. Sadness and anxiety impair driving focus, whereas anger produces road rage and raises driving risk.

CONTRIBUTIONS OF THIS WORK

A high-performing detection of driver’s facial emotion recognition using a VGG-19 Classifier system.

- To enhance the image quality, a modified median noise filter is used at the pre-processing stage.
- The feature will be extracted using DenseNet-121.
- A VGG-19 Classifier model built on the architecture and optimized for that images to be categorized as angry, happy and sad. The suggested approach is evaluated using the performance criteria Accuracy, Precision, Recall, and F1-score.

The following describes the way the paper is established: Section 1 illustrates the introduction. The literature survey are described in Section 2. Section 3 explains the suggested techniques and Section 4 displays the findings of the experiment.

2. LITERATURE REVIEW

Because of the continuous improvements in sensing technology and the topic of car emotion detection research is becoming more and more important due to its potential to make encounters safer, more

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

effective, and more enjoyable. The peer-reviewed literature tackles the issue of car emotion identification in order to encourage further research in this field. An overview of how these studies address the primary issues, such measuring emotions, determining relevant signal groups, recognizing emotional states, and altering interactions to enhance the driving experience is also given by Zepf and associates (2020) [9]. The emotional and cognitive analysis needed to construct ADAS, including the implementation of machine learning-based processing algorithms, psychological models, sensors that record physiological data, and distributed architectures needed for complicated interactions.

Each of the previously stated topics has its safety issues addressed, highlighting current flaws and challenges. For example, a number of issues, such as imprecise driver status assessment, defects implementing technologies installed in cars to track how well and attentively drivers are doing is extremely difficult due to factors that influence of in-car sensor parts and communication issues between the driver and the ADAS. According to Davoli et al. (2020), the NasNet big CNN model uses transfer learning to identify the different driving emotions [10]. As part of this study endeavor, Additionally, a custom driver emotion recognition image collection is being created. More precision than DER based on facial pictures could previously achieve is made possible by the suggested model, which blends transfer learning in the NasNet-Large CNN architecture with an enhanced faster R-CNN. The suggested model performs better than certain newly revised cutting-edge methods by Zaman et al. (2022) [11] in terms of accuracy.

The findings revealed notable distinctions in the facial expressions of the AUs in both static and dynamic driving situations, suggesting that driving behavior may have an impact on how drivers perceive their emotions and express them. A human emotion dataset is necessary to precisely identify drivers' feelings and create a trustworthy human-machine interaction system that will enhance driving comfort and safety customized for drivers must be made available. The suggested dataset will be made publicly available by Li et al. (2021)[12], allowing researchers from around the globe to test and improve their driver emotion analysis methods. Every second of driving data is assigned a safety score using row labeling, This is founded on guidelines created with the assistance of traffic safety specialists in Malaysia. The rows are then added together to make portions of the timeframe. Based on a set of criteria, segment labeling assigns a safety rating to temporal segments that are formed. The score assigned to the created temporal segment reflects the driver's behavior during that period. Al-Hussein et al. (2022)[13] then select the optimal algorithm for a proposed recognition system to classify recorded driving information using the specified profiling procedure.

3. PROPOSED METHOD

Three steps make up the suggested method for segmenting and classifying diabetic retinopathy: (1) image pre-processing; (2) Feature extraction and (3) classification. Figure 3.1 shows the overall flow diagram for this research project. The Database of Emotion detection Images provides the input data set, which is a component of the Kaggle Dataset, to improve the classification performance. Pre-processing is done using modified median noise filter. To extract features, one uses the DenseNet-121.

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The suggested VGG-19 Classifier model distinguishes between angry, happy, and sad of facial emotion subtypes.

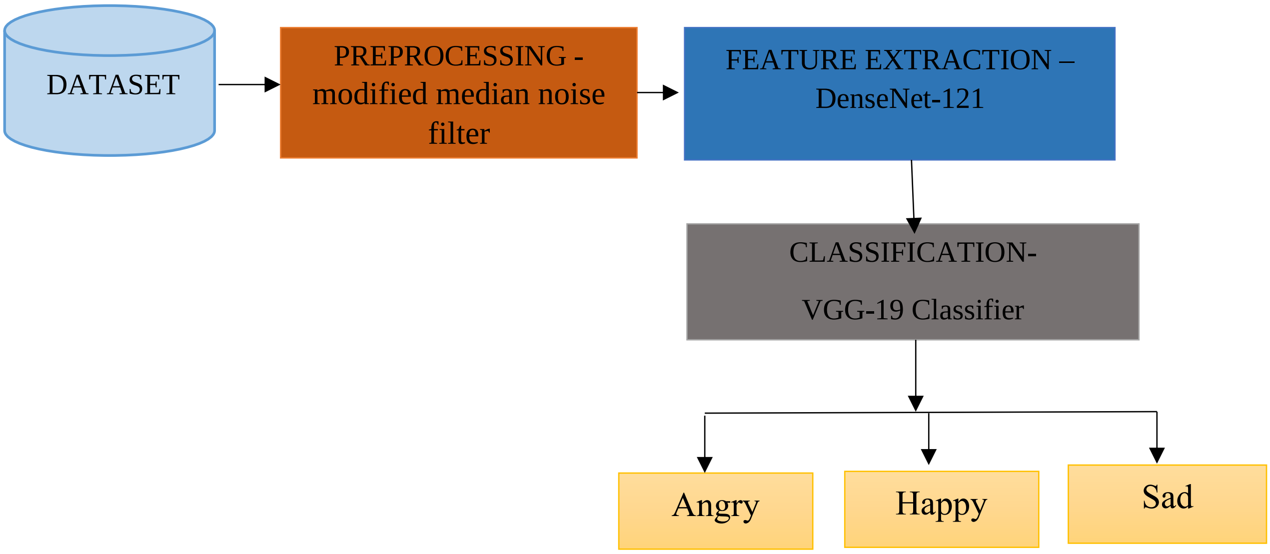

Fig 1: Proposed flow diagram

DATASET

This study employed the Facial emotion recognition dataset from (.https://www.kaggle.com/datasets/sujaykapadnis/emotion-recognition-dataset) to assess images of driver's emotion detection. The dataset consists of 6 distinct emotions: Happy, Angry, Sad, Neutral, Surprise and Ahegao. Images are RGB and presented as cropped faces with corresponding emotions. In this paper, the driver's facial recognition is done under the subtypes of angry, sad, and happy.

DATA PREPROCESSING

The first step in the diagnosing process is the preprocessing approach. Prior to being used in the feature extraction procedure, the input fundus images undergo preprocessing. First, the blurry photos are corrected as a preprocessing step.

Typically, The noise removal stage consists of two phases. Noise detection comes first, followed by noise elimination.

$$N(m,n) = \begin{cases} 1, & if\ X(m,n) noisy \\ 0, & if\ X(m,n) noise\ free \end{cases} \tag{1}$$

Where $N(m,n)$ and $X(m,n)$ represents the input image as well as the noisy image. The lowest intensity value (i.e., 0) at the noise detection stage indicates pepper noise, whereas the highest intensity value (i.e., 255) indicates salt noise. Pixels with different brightness levels are considered noise-free. The highest and lowest intensity levels, which are typically not regarded as noise, are defined by the values of adjoining neighbors (the correlation of adjacent neighbors). Pixels are deemed noise-free when their intensity levels are 255 or nearly so. However, if neither the pixel intensity level nor the average neighbor level equals 255, it is considered noise. In a similar vein, pixels that have an intensity level of zero or nearly so are considered noise-free, if the normal neighborhood level.

Feature extraction

A 224 × 224 picture is used as the fixed-size input for DenseNet121. DenseNet121 includes over 8 million parameters and 121 layers. It is separated into dense chunks with different numbers of filters but the same feature map size. The layers that reside between the blocks are called transition layers, and they are in charge of down sampling and batch normalization.

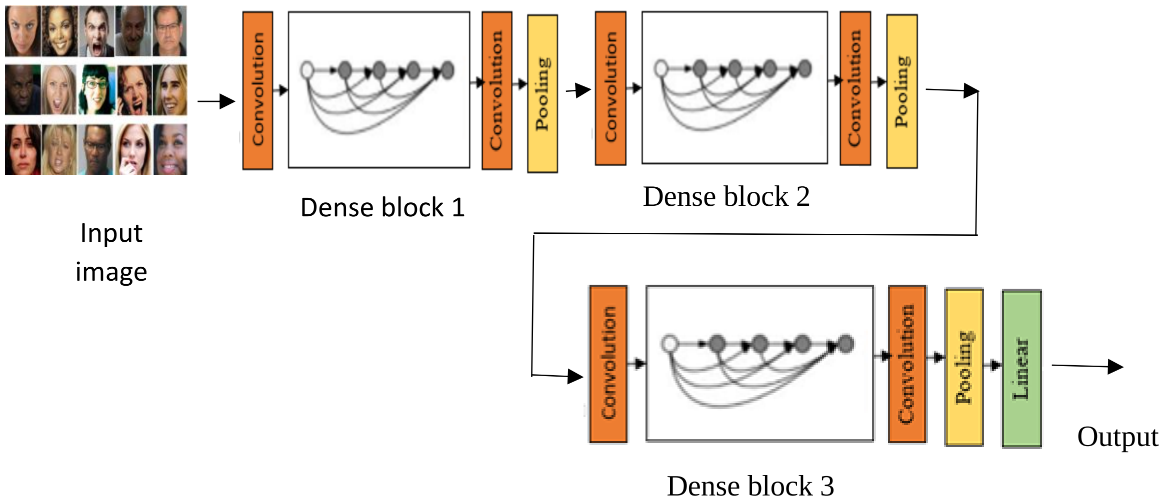

Fig 2: Feature extraction of DenseNet-121.

Examine an input image (a_o) that the suggested convolutional network processes. The nonlinear transformation f(n) is carried out by each NN layer in the network. Assume that layer n is made up of all feature maps from earlier convolutional layers. Input feature maps from Layers 0 through a–1 are merged. The nth layer's output is provided by

$$a_n = f([a_o, a \ldots \ldots, a - 1]) \quad (4)$$

The next steps in the transition layer are normalization, 3 × 3 convolution, and rectified linear units. The concatenation method becomes unfeasible if the feature map sizes are changed. Consequently, for the layers with different feature map sizes, down sampling is employed. The 1x1 convolution transition layer and 2x2 average pooling are positioned between two adjacent dense convolution blocks. The first Conv layer is made up of 7x7 Conv blocks with stride. Convolution learns the characteristics of the image while maintaining the relationships between the pixels.

$$f(a_o) = \max(0, a_o) \quad (5)$$

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

CLASSIFICATION

Among the CNN-based architectures under consideration is the multi-layered VGG-19 network architecture. It has 19 learnable weights for transfer learning and 16 convolutional layers for feature extraction during training. It used five fully connected (FC) layers and one output layer at the termination. The initial convolutional layer uses roughly 64 kernels (3 × 3 filter size) to extract features from the photographs that have been added. Between the convolution layers, a max-pooling layer has also been added.

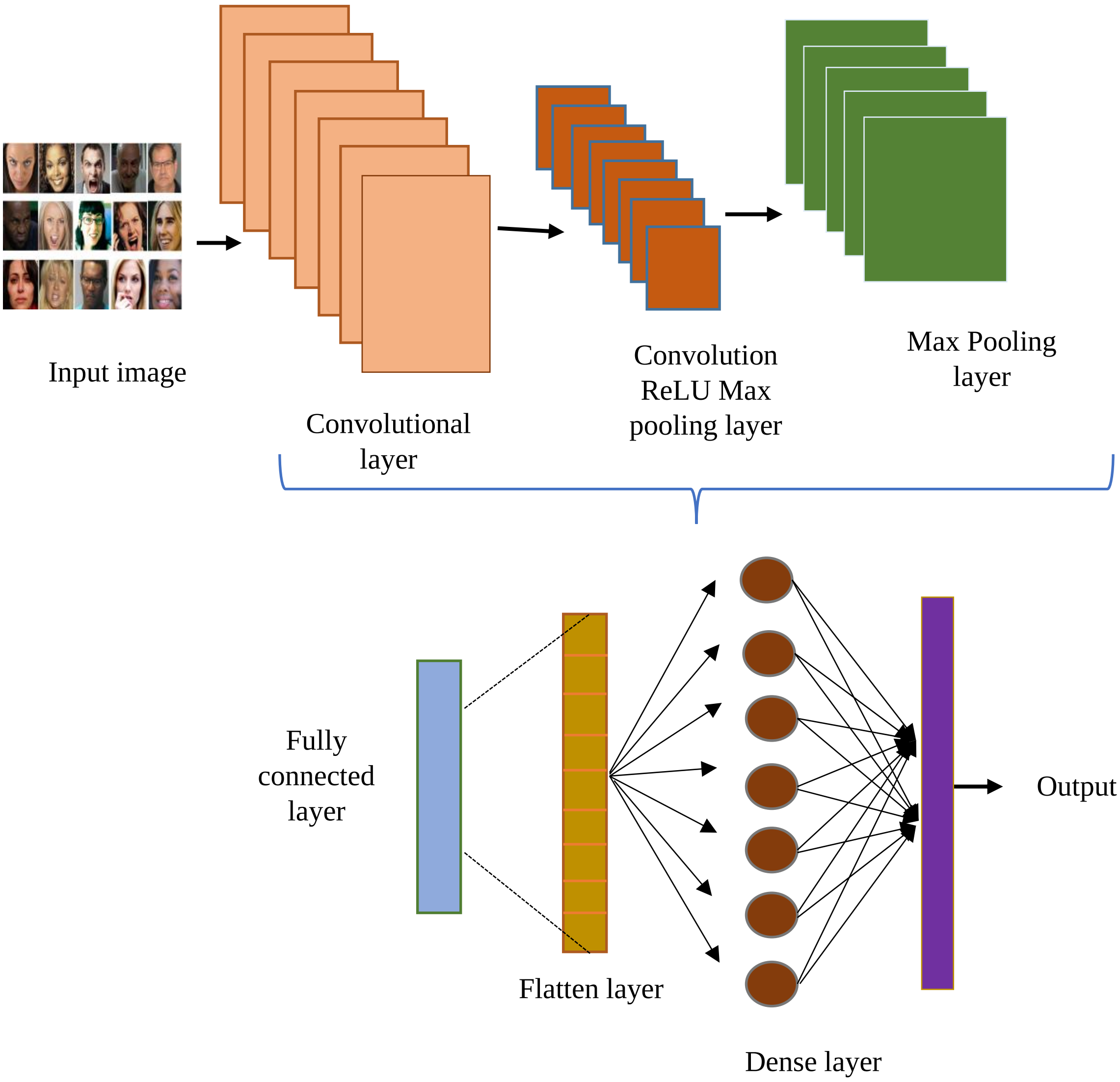

Fig 3: VGG-19 Architecture

PERFORMANCE MEASURES

Several evaluation measures, such as accuracy, precision, recall, and F1-measure, are used for the prediction and classification problems. The effectiveness of the proposed model is assessed using the assessment metrics listed below.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (24)$$

$$Precision = \frac{TP}{TP+FP} \quad (25)$$

$$Recall = \frac{TP}{TN+FN} \quad (26)$$

$$F1 - Measures = 2\ x\ \frac{Precision+Recall}{Percision+Recall} \quad (27)$$

4. EXPERIMENTAL RESULTS AND DISCUSSION

Three sets of tumor images are created: 20% are used for testing, while the other 80% are chosen for training from the dataset. Next, for additional processing, the original dataset is divided into training and testing datasets.

Table 1: Input images classification

Class	Total image	Training	Testing
Angry	1313	576	737
Happy	3740	1077	2663
Sad	3934	2153	1781

Three groups of the dataset were used in this investigation. Angry, Happy and Sad images were included in the training and testing set. Fig. 4 displays the images from the dataset that were used as data input.

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

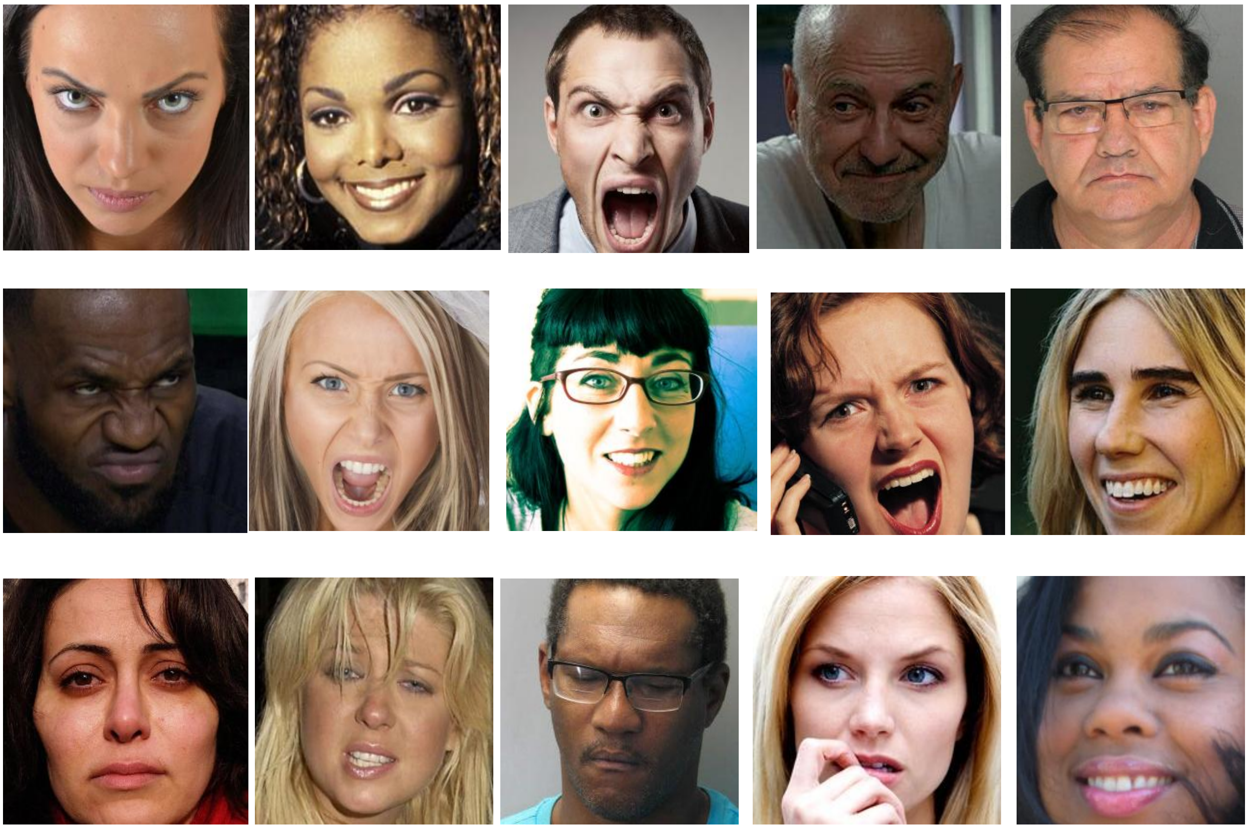

Fig 4: Input images

IMAGE PREPROCESSING

The next stage is to enhance the quality of the facial expression images using a modified median noise filter. A median filter was applied to a study on noise reduction in facial expression recognition images in order to improve contrast and clarity. The preprocessed images are displayed in Figure 7.

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

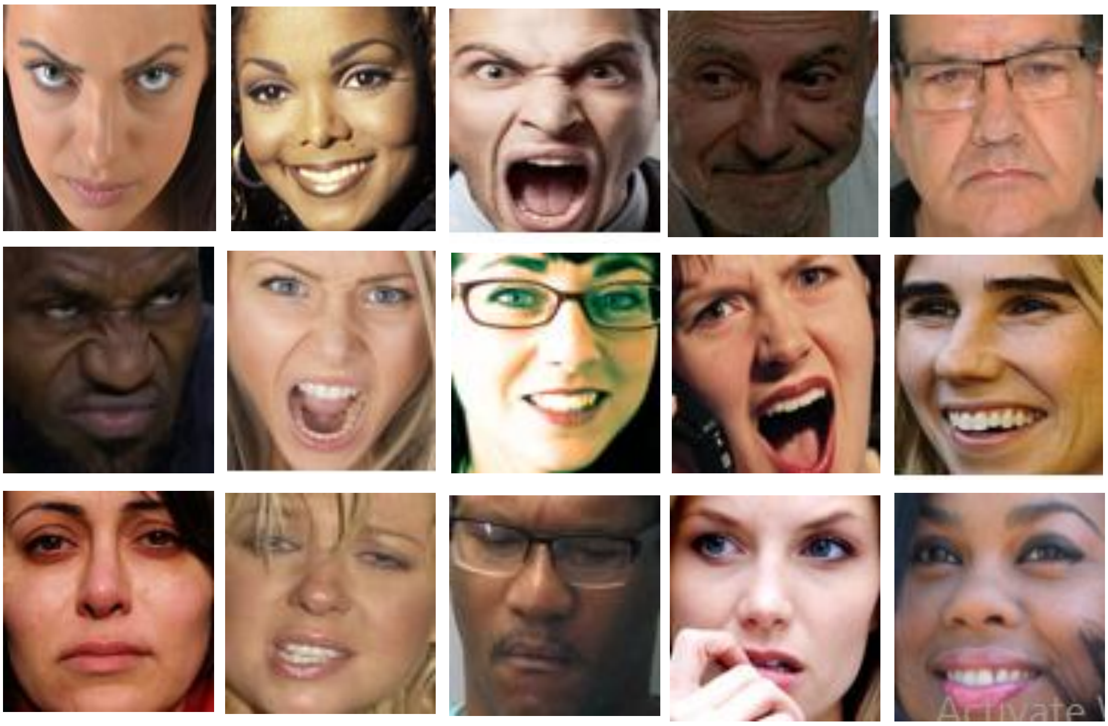

Fig 5: Preprocessed images

FEATURE EXTRACTION

After the features have been separated with the help of the preprocessed picture, the features are recovered using the DenseNet-121 approach.

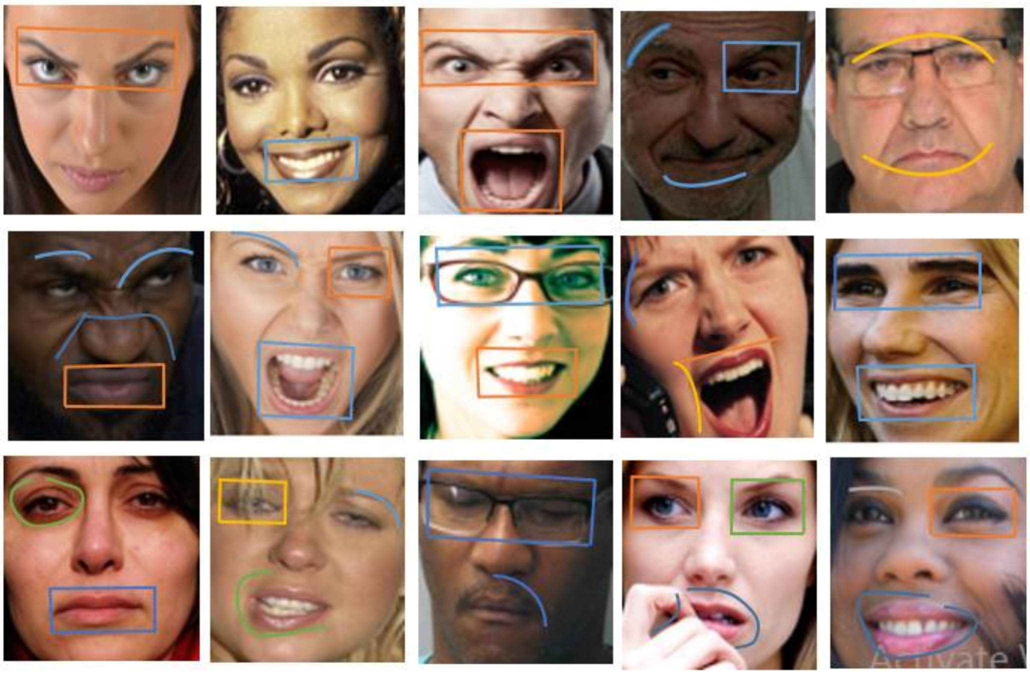

Fig 7: Feature extracted images

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

PERFORMANCE MEASURES

Here, the deep learning approach is used to assess the efficacy of the suggested methodology for both image feature extraction and classification. By applying DenseNet-121 models of feature exacted to classification using VGG-19 classifier to generate the accuracy, precision, recall, and F1-score metrics, the effectiveness of the suggested classifiers in this section is examined. Table 2 displays the results of the suggested strategy.

Table 2: The suggested VGG-19 model results for Facial emotion detection.

	Accuracy (%)	Precision (%)	Recall (%)	F1-measure (%)
Angry	98.1	97.4	95.8	97.7
Happy	98.3	97.5	97.3	97.4
Sad	97.1	98.3	96.4	95.8
Average	97.83	97.73	96.5	96.97

For Angry, the comparable figures were 98.1%, 97.4%, 95.8%, and 97.7% for accuracy, precision, recall, and F1-measure. The Happy was detected with 97.5% precision, 97.3% recall, 97.4% F1-measure, and 98.3% accuracy and for sad detection 97.1% of accuracy, 98.3% of precision, 96.4% of recall and 95.8% of F1-measure.

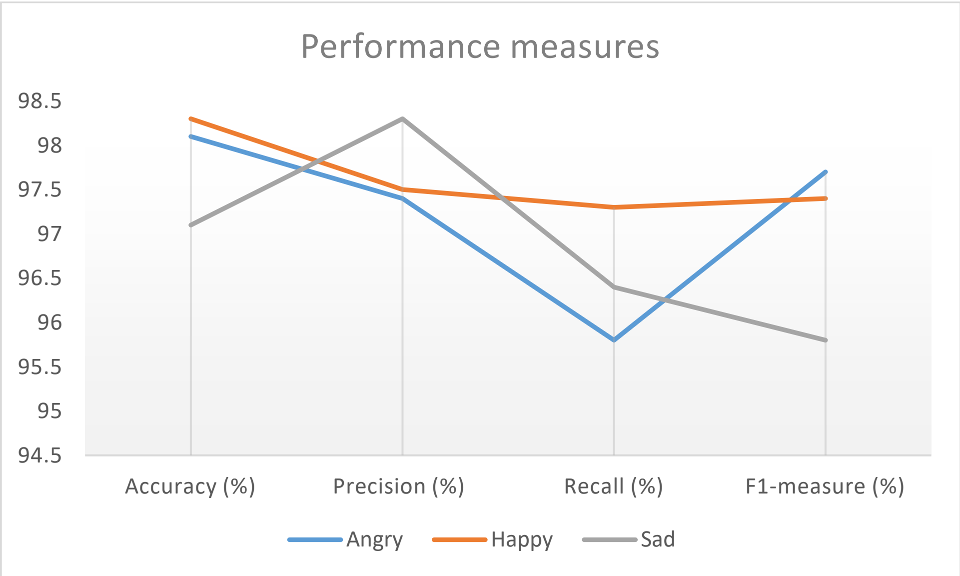

Fig 8:Performance measures

An analysis was conducted on the VGG-19 image classification data to evaluate the effectiveness of the proposed methodology. 97.73% precision, 96.97% F1-score, 96.5% recall, and 97.83% accuracy were obtained using the suggested approach. Figure 11 shows the confusion matrix, which shows the prediction performance generated by the VGG-19 classification model.

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

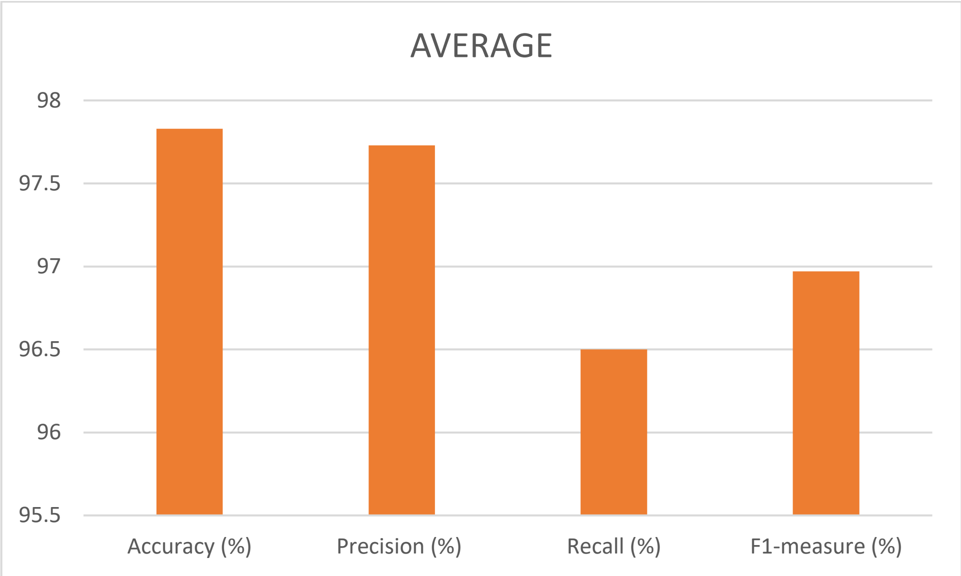

Fig 9: Average of Performance measures

The proposed method has a 97.83% success rate in correctly categorizing tumors as Angry, happy and sad. The overall effectiveness of facial expression recognition is shown in Figure 10. The comparison between the suggested model and references is shown in Table 3 and which implies that proposed method shows better results.

Table 3: Comparison with other suggested methods

Reference	**Method**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-measure (%)**
Sukhavasi et al. (2022) [14]	multi-task cascaded neural networks	96.1	NA	NA	NA
Abosaq, et al.(2022) [15]	VGG-16	95	96	92	94
Proposed method	VGG-19	97.83	97.73	96.5	96.97

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

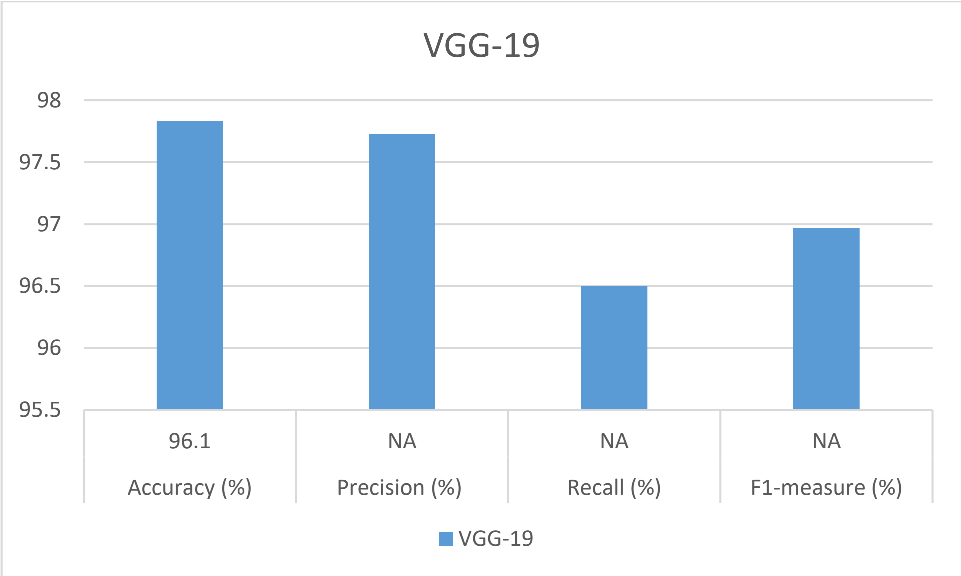

Fig 10: Graphical representation of proposed method

CONFUSION MATRIX

Figure 11 shows the confusion matrix, which shows the facial emotion detection performance generated by the VGG-19 classification model. The predicted accuracy of the model is evaluated using this sample, which includes 1313 Angry images out of 737 were tested and 737 for training and 3740 Happy images out of 1077 images for testing and 2663 for training that are accurately classified. For sad images there were 3934 in total,2153 were detected for testing and the remaining for the testing process.

Angry	1313 7.3%	576 5.76%	737 1.61%	**96.7%** **3.3%**
Happy	3740 2.67%	1077 1.1%	2663 1.58%	**95.4%** **4.6%**
Sad	3934 1.79%	2153 2.3%	1781 3.72%	**97.7%** **2.3%**
	96.6% **3.4%**	**96.5%** **3.5%**	**96.7%** **3.3%**	**97.83%** **2.17%**
	Angry	Happy	Sad	

Fig 11: Confusion matrix

Journal of Information Systems Engineering and Management
2025, 10(3)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

CONCLUSION

In order to support sophisticated driver assistance systems in intelligent automobiles a real-time driving scenario, this study proposes novel deep network algorithms for facial expression detection that can recognize the emotions of the driver.. This is how the proposed approach works and uses the driver's facial expressions during various mood swings. There is a lot of contrast in images. To cut down on noise, images are preprocessed using a modified median noise filtering model. In order to extract features, images were preprocessed. The DenseNet-121 method is used to extract landmark characteristics from driver photos' facial emotions. The suggested approach achieved 97.83% accuracy, 96.97% F1-score, 96.5% recall, and 97.73% precision. The experiment's findings demonstrate that, in terms of accuracy and other performance metrics, the proposed VGG-19 algorithm performs better than cutting-edge techniques.

REFERENCES

[1] Eyben, Florian, Martin Wöllmer, Tony Poitschke, Björn Schuller, Christoph Blaschke, Berthold Färber, and Nhu Nguyen-Thien. "Emotion on the road—necessity, acceptance, and feasibility of affective computing in the car." *Advances in Human‐Computer Interaction* 2010, no. 1 (2010): 263593.

[2] Tango, Fabio, and Marco Botta. "Real-time detection system of driver distraction using machine learning." *IEEE Transactions on Intelligent Transportation Systems* 14, no. 2 (2013): 894-905.

[3] Vicente, Francisco, Zehua Huang, Xuehan Xiong, Fernando De la Torre, Wende Zhang, and Dan Levi. "Driver gaze tracking and eyes off the road detection system." *IEEE Transactions on Intelligent Transportation Systems* 16, no. 4 (2015): 2014-2027.

[4] Arumugam, Subramanian, and R. Bhargavi. "A survey on driving behavior analysis in usage based insurance using big data." *Journal of Big Data* 6, no. 1 (2019): 1-21.

[5] Bota, Patricia J., Chen Wang, Ana LN Fred, and Hugo Plácido Da Silva. "A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals." *IEEE access* 7 (2019): 140990-141020.

[6] Stilgoe, Jack. "Machine learning, social learning and the governance of self-driving cars." *Social studies of science* 48, no. 1 (2018): 25-56.

[7] Sohail, Anwar Mehmood, Ying Wah Teh, Nauman Khan, Zahid Khan, and Anis Koubaa. "A Comprehensive Review: Analysis of Machine Learning, Deep Learning, and Large Language Model Techniques for Revolutionizing Driver Identification." *Authorea Preprints* (2024).

[8] Hole, Y., Hole S., L. P. Leonardo Cavaliere, B. Nair, M. Hasyim and H. B. Bapat, (2023). "Blockchain Usages in Hospitality Management," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 2798-2801, doi: 10.1109/ICACITE57410.2023.10183291.

[9] Hole, Y., Hole S., A. A. Ayub Ahmed, E. Efendi, I. Ibrahim and M. Hasyim, (2023). "Internet of Things Issues and Challenges," 2023 3rd International Conference on Advance Computing and

Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 1370-1373, doi: 10.1109/ICACITE57410.2023.10183221.

[10] Bhattacharya, Sweta, Siva Rama Krishnan Somayaji, Thippa Reddy Gadekallu, Mamoun Alazab, and Praveen Kumar Reddy Maddikunta. "A review on deep learning for future smart cities." *Internet Technology Letters* 5, no. 1 (2022): e187.

[11] Zepf, Sebastian, Javier Hernandez, Alexander Schmitt, Wolfgang Minker, and Rosalind W. Picard. "Driver emotion recognition for intelligent vehicles: A survey." *ACM Computing Surveys (CSUR)* 53, no. 3 (2020): 1-30.

[12] Davoli, Luca, Marco Martalò, Antonio Cilfone, Laura Belli, Gianluigi Ferrari, Roberta Presta, Roberto Montanari et al. "On driver behavior recognition for increased safety: a roadmap." *Safety* 6, no. 4 (2020): 55.

[13] Zaman, Khalid, Zhaoyun Sun, Sayyed Mudassar Shah, Muhammad Shoaib, Lili Pei, and Altaf Hussain. "Driver emotions recognition based on improved faster R-CNN and neural architectural search network." *Symmetry* 14, no. 4 (2022): 687.

[14] Li, Wenbo, Yaodong Cui, Yintao Ma, Xingxin Chen, Guofa Li, Guanzhong Zeng, Gang Guo, and Dongpu Cao. "A spontaneous driver emotion facial expression (defe) dataset for intelligent vehicles: Emotions triggered by video-audio clips in driving scenarios." *IEEE Transactions on Affective Computing* 14, no. 1 (2021): 747-760.

[15] Al-Hussein, Ward Ahmed, Lip Yee Por, Miss Laiha Mat Kiah, and Bilal Bahaa Zaidan. "Driver behavior profiling and recognition using deep-learning methods: In accordance with traffic regulations and experts guidelines." *International journal of environmental research and public health* 19, no. 3 (2022): 1470.

[16] Sukhavasi, Susrutha Babu, Suparshya Babu Sukhavasi, Khaled Elleithy, Ahmed El-Sayed, and Abdelrahman Elleithy. "Deep neural network approach for pose, illumination, and occlusion invariant driver emotion detection." *International Journal of Environmental Research and Public Health* 19, no. 4 (2022): 2352.

[17] Abosaq, Hamad Ali, Muhammad Ramzan, Faisal Althobiani, Adnan Abid, Khalid Mahmood Aamir, Hesham Abdushkour, Muhammad Irfan et al. "Unusual driver behavior detection in videos using deep learning models." *Sensors* 23, no. 1 (2022): 311.