# Multi-Facial Emotion Recognition Using Fusion CNN on Static and Real-Time Inputs: A Deep Learning Approach

Vineet Singh[1], Nitin Kumar Tripathi[2*], Rishi Jain[2]

[1]Badruka School of Management, Hyderabad, India
[2]Asian Institute of Technology, Thailand
*Corresponding email: nitinkt@ait.asia

**Abstract:**
Facial emotion recognition is a pivotal component in the domain of affective computing, aiming to bridge the gap between human emotional expression and machine interpretation. This study introduces a deep learning-driven framework for multi-facial emotion recognition, leveraging diverse data modalities including static images, video frames, and webcam inputs. The model was trained and evaluated using the CK+ dataset with a systematic data split for training, testing, and validation to ensure robustness. A Fusion Convolutional Neural Network (Fusion CNN) was proposed to optimize feature extraction and improve classification accuracy across heterogeneous input sources. The implementation was realized using Python with OpenCV and Keras libraries, while statistical validation, including chi-square tests and regression analysis, was conducted in R to assess model consistency and accuracy. Among the various models tested, the Fusion CNN demonstrated superior performance with an accuracy of 72.16%, surpassing traditional CNN and RNN architectures. The results underscore the potential of the proposed approach in advancing real-time emotion recognition systems, with future scope for integration into intelligent user interfaces and assistive applications.

**Keyword:** Facial Emotion Recognition, Fusion Convolutional Neural Network (Fusion CNN), Deep Learning and Multimedia Input Processing

## 1. Introduction

While language plays a crucial role in facilitating human contact, it is often accompanied by supplementary forms of expression, including gestures, posture, and vocal inflections, which enhance the whole communicative experience. The characteristics are commonly accompanied by physiological responses, such as elevated heart rate, and are contingent upon the circumstances of the interaction[1].

Despite the significant role that human computer interaction and human-mediated communication play in contemporary society, there remains a lack of essential tools for comprehending and addressing non-verbal cues related to attitudes, emotions, and mental states. These cues, which are commonly utilized in human communication and reasoning, are currently not adequately accounted for in these domains. According to the observation conducted by [2],individuals interact with computers in a manner similar to their interactions with other individuals. A few basic emotions can be seen in the Figure 1.


Figure 1: State of Emotion (source: facial expression recognition Archives - Sefik Ilkin Serengil (sefiks.com))

Since antiquity, facial expressions have been studied, in part because they are the most significant form of nonverbal communication. At first, renowned philosophers and thinkers like Aristotle and Stewart examined face expressions[3].

The study of face expressions evolved into an empirical field of research with Darwin. Researchers in the fields of psychology and cognitive science were very interested in Darwin's studies. Numerous studies linking facial expression to emotion and interpersonal communication were conducted in the 20th century. Most significantly, Paul Ekman reexamined Darwin's findings and asserted that there are six universal emotions that can be created and understood regardless of cultural background[4], [5], [6]
With the help of emotion human behavior can be recognized where the various data are available, and these can be used for the different computer-based environments to study or for the research purpose. All these complex's calculations take place in a 2D normalized emotional complex model valence scale can be seen in Figure 2.

Emotion is one of the most important human life to understand, but many of us are not able to understand the emotions of other people which can cause harm to him/her and later the consequences become very difficult, sometimes leading to suicide, depression and many other things[7].
While language helps to understand the tone of speech and classify it accordingly but when the new person enters other kingdom or other country then it becomes problem to talk and understand the situation so the emotion classifier might help at least to understand whether to talk or not talk based on the current mood or emotion of the respective person.

There are many studies conducted using the pre-approved dataset and real-time application not applicable many times, so in this part, after training the model, the real-time deployment and accuracy to the real world were generated. Few people who cannot speak (deaf-mute) it becomes very difficult to another person to adjust to their space, or they feel left out or improper treatment is given to them, but facial expression gives many understanding through which processing the facial complexity might help other understand the same.

Health is also one of the factors to be considered for emotion, because if the person is not well or any health issue by using the characterization or judging by character the output tends out to be in the result. Such as a person being ill from inside (not good health) and having smile on face, but it is clearly visible by body temperature or coughing so such character help to recognize the state of person, which is actually not happy, which in further state can be mental health issue, Trauma health issue, disorder health issue.

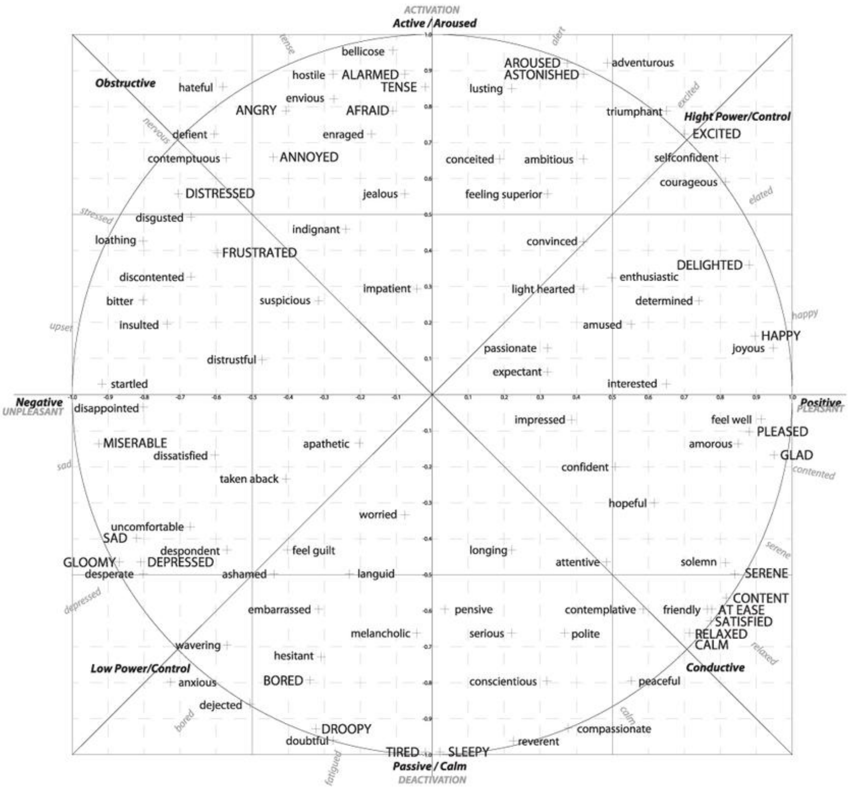



Figure 2: 2D normalized Emotion Complex (source: [The 2-D Emotion Wheel | Download Scientific Diagram (researchgate.net)](https://www.researchgate.net))

As due to different facial structure or facial complexion it is also a challenging part to say whether the person is happy or other things, because in Asian countries the complexity and also the facial structure, for example if we compare between the India and Thailand the facial structure ( like eyes) it seems more tedious to help to understand the emotion, so the dataset available are defined for specific part and specific place, because while training the model accuracy may change and which can result into wrong emotion detection.

As according to study the facial emotion study is being in continuous upgrade under biomedical as emerging because of the lack of understanding, but this is the new demand of era that emotion are taken upmost care and for which it is also stated that due to sad emotion the behavior on mental health is affected and due to which many disturbing scenario take place such as depression, suicide and other a lot on wellbeing too. To develop a model with high accuracy for multi facial emotional recognition for the image, video, and live webcam.

The main scope of this study is to bridge the model made for facial emotional recognition with the real world working and later test the efficiency of it and the accuracy of the same. The study is also done to serve the people in need or the hospital and the psychiatrist, the disabled people and last but not least the people facing problems might be true work done for them to save them from taking any further actions.

## 2. Literature review

Many research scientists have summed up a result and claim that wellie people live long, have a less chronic disease, which led to healthier lifestyle and low death rates is seen in such case[8]. So, emotion plays an important role in seeing the health and wellbeing both. Characterization which means the distinctive nature of a person or someone.

With which help the analyze the relationship. Health being an important aspect as the ratio of death seen which clearly shows healthy people live more than unhealthy people. WHO says 12.6 million are each year due to unhealthy environment. Emotion which can take control over a person should be also considered but many ignore such things and due to which lack of understanding creates the situation of vulnerability [4], [9], [10], [11].
One of the articles from medical news today says that major emotional distress can be seen at work because of primary reasons: concern about job security, long hours, low pay, poor working condition, increasing responsibility, a lack of control over work and relationship between colleagues and managers

Initial steps to understand is the emotion and understanding the emotion with the other behavior is the characterization. If the person is happy and he/she is automatically getting the good well and better health care is done[8], [9], [10], [12], [13], [14]., the objective is to find the perfect database as of which FER 2013 database is considered the one of the certified datasets among all. It is also said that this computer vision requires high-level image processing, and this type of project is the need of an era. The flow is to image acquisition, preprocessing of an image, face detection and feature extraction with extraction of image. The emotion which are taken into consideration are happy, sad, angry, disgust, surprise, fear which are accepted by many other researcher groups [11], [12], [13], [14], [15], [16]. Various emotions and their description are detailed in Table 1 below.

| EMOTION | Motion of Facial Parts |
|---|---|
| Happy | Open eyes, open mouth, lips corner pulled, cheeks raised |
| Sad | Outer eyebrow down, inner eyebrows raised, eyes closed, lip corner down |
| Surprise | Eyebrow up, open eyes, jaw dropped |
| Anger | Eyebrows pulled down, open eyes, lip tightened |
| Fear | Outer eyebrow down, inner eyebrows up, mouth open |
| Disgust | Lip corner depressor, lower lip depressor, eyebrows down, nose wrinkled |

Table 1: Emotion and facial Action Units (Saravanan et al., 2019)

[15] the change in the hyperparameters setting of convolutional neural networks is needed to attain the higher efficiency. The hyperparameter setting are done after training the dataset and generating the new at same point. The FER2013 after using CNN gave the 72.16% efficiency. Number of kernel in first convolutional layer is 32 and max is 256 with step 32. The maximum number kernel is 512 steps with the step 64. The dropout value is 0.4 and 0.1 for the convolutional layers and for fully connected layer the dropout value is 0.1.

| **Network Type** | **Accuracy %** |
|---|---|
| Ensemble CNN | 75.8 |
| Fusion CNN + BOVW | 75.42 |
| Multitask CNN | 75.2 |
| Hybrid CNN | 73.73 |
| CNN | 72.7 |
| Residual | 72.4 |
| Deep CNN | 71.6 |

Table 2:Summary of results on FER2013

| | |
|---|---|
| [17] | Convolutional Layer:4<br>Kernel: 32, 64<br>Dropout: 0.2<br>Dense: 1024, 4096 |
| [18] | Convolutional Layer:3<br>Kernel: 32, 64<br>Kernel size: 4*4, 5*5<br>Dense: 1024 |
| .[19] | Convolutional Layer:6<br>Kernel: 32, 64, 128<br>Dropout: 0.1, 0.4, 0.5<br>Kernel size: 3*3<br>Dense: 2048 |
| [20] | Convolutional Layer:8, 16<br>Kernel: 32, 64, 128, 256, 512<br>Kernel size: 3*3, 5*5<br>Dropout: 0.5<br>Dense: 1000, 4096 |
| [21] | Convolutional Layer:3, 4, 5<br>Kernel: 32, 64, 128<br>Kernel size: 3*3<br>Dense: 2048 |
| [22] | Convolutional Layer:32<br>Kernel: 16 to 384<br>Dropout: 0.4 |

Table 3:Summary of Hyperparameter settings

[23] the author claims to use the Candide grid note to face landmark pattern as shown in Figure 3 . In this the grids are formed over and analyzed over it; geometric displacement shows the facial expression intensity frame which is then used by SVM. The database used while performing was Cohn-Kanade Database, which has multiclass SVM and which led to the accuracy of 99.7% and for proposed SVM it went to 95.1% for facial emotional recognition [24], [25].

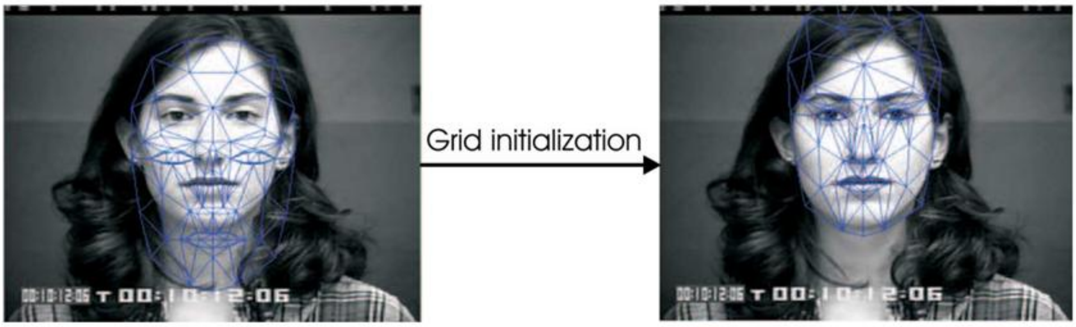



Figure 3: Grid Initialization (source: Bing)

[26] the author says the emotion is primary stage and also the emotion triggers the physical as well as mental fatigue and due to which problems start, later to which emotion are classified into 3 categories. (Low, medium, High) and to serve this new methodology is introduced DOG and HOG, where the pairs are formed of eyes, nose and lips and facial features are extracted along with it as shown in Figure 4.

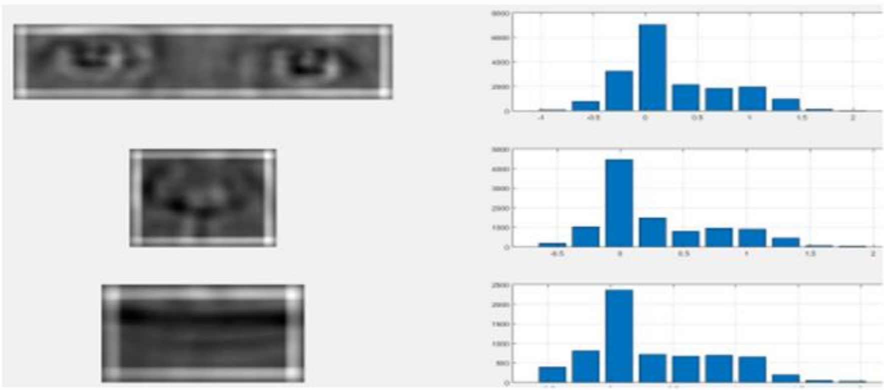


Figure 4:DOG and HOG (source: Bing)

| No | Name of paper | Year | Comments | Public. |
|---|---|---|---|---|
| 1 | A Multimodal Multi-Party Dataset for Emotion Recognition in Conversation [27] | 2019 | Lack of dataset, F and weighted average score was 65% for text+audio combination, Annotation is to be done, method used RNN and LSTM, RNN was | IEEE |

| | | | | |
|---|---|---|---|---|
| | | | useful and high accuracy, New model dataset defined | |
| 2 | Evaluation and Discussion of Multi-modal Emotion recognition.[28] | 2009 | Feedback is to be considered, DaFEx database used, Bi-Modal Performance was 78.17%. Recommendation for multi-cue fusion was given | ICCEE |
| 3 | Stress Recognition using Face Images and Facial Landmarks [29] | 2016 | 3 ways to detect stress, with facial landmarks and grey face image 64.63% was performance, own database used | IEEE |
| 4 | Modeling Stress using Thermal Facial Patterns: A spatio-temporal Approach [30] | 2013 | Svm technique was used and stress recognition were better with the HDTP features | IEEE |
| 5 | The Facial Stress Recognition Based on Multi-histogram features and Convolutional Neural Network[31] | 2018 | Neutral, low and high stress classes identified and HOG and DWT method was used for extraction , the database is FERET and k fold validation is 5. | IEEE |
| 6 | The Use of AI for Thermal Emotion Recognition [32] | 2020 | Thermal vs RGB images are compared and RGB gas good computer vision and even works on low data quality | IEEE |
| 7 | Emotional and Physical Stress Detection and Classification using Thermal Image Technique [33] | 2021 | The use of EEG, ECG and GSR is being used, deep learning techniques used are SVM and CNN with 89% accuracy. | Annals of RSCB |
| 8 | Cross Subject Multimodal Emotion Recognition Based on Hybrid Fusion [34] | 2020 | Only 3 classes to be used and the accuracy for hybrid fusion is 81.2% and mean accuracy is 74.2% , some limitation are covered. | IEEE |
| 9 | Convolutional MKL Based Multimodal Emotion Recognition and Sentiment Analysis [35] | 2016 | Temporal CNN is proposed, the audio visual and text are used with the good result and if only image is used about 71% is good probability. | IEEE |
| 10 | Computer Vision and Image Understanding [36] | 2016 | Multi modal video induced emotion recognition is being | ELSEVIER |




ISSN: 3088-2370 (Online) | © Geoinformatics International

| | | | | |
|---|---|---|---|---|
| | | | proposed based on 27 participants. | |
| 11 | A survey of Face Recognition Technqiues [37] | 2009 | Application of face recognition is stated, also the difficulties and feature based approach. | Journal of Information Processing Systems |
| 12 | A synthesis Based Approach for Thermal to Visible Face Verification [38] | 2021 | ARL_VTF and TUFTS multi spectral face dataset are used for which algorithm are created and lastly the strong performance on MILAB-VTF(B) dataset is recommended. | IEEE |
| 13 | Facial Emotion Recognition system through Machine Learning Approach. [39] | 2017 | These include image acquisition, pre-processing and feature extraction and ISED database is used | IEEE |
| 14 | A novel facial emotion recognition method for stress inference of facial nerve paralysis patient [40] | 2022 | Diagnosis of facial nerve paralysis is fone and accuracy is 66.59%, VGGNet model is used | IEEE |
| 15 | Mental stress recognition based on non-invasive and non contact measurement from stereo thermal and visible sensors [41] | 2015 | Internal emotion state, almost 98% of correct measurement of ROI and temperature was detected,SIFT 88.6% of correct matching. | IJAE |
| 16 | Automated classification and recognition of facial expression using infrared thermal imaging [42] | 2004 | Visible spectrum is used and IRTI was identified facial expressions were derived from it and the happiness, sadness and disgust were the results | IEEE |
| 17 | High resolution thermal face dataset for face and expression recognition [43] | 2020 | Established facial dataset(Equinox,Carl db), use of LBP pattern. | PAN M&MS |
| 18 | Video-based facial expression recognition using graph convolutional networks [44] | 2021 | The dynamic expression capture is the motive to be used in this and using the CNN, but the use of Graph CN is done, CASIA CK+ dataset is used | ICPR |

| | | | | |
|---|---|---|---|---|
| 19 | Facial Expression Recognition in Image Sequences using Geometric Deformation Features and Support Vector Machines [45] | 2007 | Candide gride notes for face recognition is used, the grid tracking and deformation system used, multiclass SVM of classifiers 5 expressions. | IEEE |
| 20 | A multimodal features fusion framework for kinetic based facial expression recognition using dual kernel displacement analysis [46] | 2021 | Multi-modal feature fusion framework for kinetic, the performance of LDA and KDA is also done with average accuracy improving by 10%. | IEEE |
| 21 | Facial Emotion Recognition using Convolutional Neural Networks [47] | 2019 | Methods including decision trees, neural network are used, various hyperparameter tunning the finally accuracy was 0.60. | IEEE |
| 22 | Going Deeper in Facial Expression Recognition using Deep Neural Networks [48] | 2021 | Most of them are based on HOG,LBPH and Gabor where just changing hyperparameter are tuned to best accuracy. | IEEE |
| 23 | Hyperparameter optimization in CNN for learning centred emotion recognition for intelligent tutoring systems [49] | 2019 | CNN method is used , the problem of hyperparameter in CNN is defined, proposal of genetic algorithm for tuning the hyperparameter of CNN, for which the 8% growth was there. | IEEE |
| 24 | Convolution Neural Network Hyperparameters optimization for facial emotion recognition.[50] | 2021 | The optimal hyperparameter of network were determined by generating and training models based on Random search algorithm, FER2013 data with accuracy 72.16%. | ATEE |
| 25 | Oriented attention ensemble for accurate facial expression recognition [51] | 2020 | Weighted mask and correlation calculation with the fused to get the output. | ELSEVIER |
| 26 | Facial Expression recognition using dual dictionary learning.[52] | 2017 | Dual dictionary method is used for regression and SRC and CRC, database are Ck+, CK,MMI,JAFEE for input and training | ELSEVIER |
| 27 | Multimodal Emotional Recognition using Deep Learning [53] | 2021 | Multimodal are studied across unimodal as they offer high accuracy rate , | JASTT |

| | | | | |
|---|---|---|---|---|
| | | | emotional awareness problem is to be solved. Different dataset are studied with various method. | |
| 28 | Real time speech emotion recognition using RGB Image classification and Transfer Learning [54] | 2021 | AlexNET-SVM, FTAlexNet were investigated and Berlin Emotional Speech Database is used. Transfer learning is used for flexible model. | IEEE |
| 29 | Dual Stream Multi task Gender based Micro expression recognition [55] | 2020 | GEME is used to get the gender detetction, the unique micro expression is done using this and the age and gender can be commented or tell using expression. | ELSEVIER |
| 30 | Video and Image based Emotion Recognition Challenges [56] | 2015 | It consist of audio video with the AFEW database and this was conducted with the open problem | IEEE |
| 31 | A fully annotated Thermal face database and its application for thermal facial expression recognition [57] | 2018 | As there is not much data available and this field is vast growing in computer vision, so the author saw the short coming of data and have created database which is annotataed manually and process using the machine learning SVM.<br>Manual annotation is done and also this can be further used for medical parameters.<br>SVM gave the 75% efficiceny and the other gave low such as KNN BDT NB RF other models used for this. | IEEE |

Table 4:Summary of Literature review

The summary of the entire literature review is presented in Table 4.

## 3. Methodology

**Dataset**

There are already many pre-available datasets which has been worked on like some of the most common used datasets are.

- FER 2013 dataset
- Berlin Emotional Dataset
- Asian Character Dataset
- Music Mood Classification Dataset
- Cohn Kanade Dataset
- IEMOCAP dataset
- Dual Emotional dataset

But after listing out few of them, the most efficient dataset is Cohn Kanade, It is bit hard to pre-process, many researcher have used FER 2013 and Cohn Kanade dataset and efficiency with the accuracy of Cohn Kanade dataset gave the perfect result. The emotion and no. of instances related to it are mentioned in Table 5.

The model was trained for 200 epochs and achieved an accuracy measure of 71.1% on the FER2013 dataset, 99.31% on the CK+ database. Training data 20%, Testing 40% and Validation set 40%. [52], [53], [54], [55], [56], [57]

| **No.** | **Expression** | **No. of Instances** |
|---|---|---|
| **1** | Angry | 527 |
| **2** | Contempt | 47 |
| **3** | Disgust | 389 |
| **4** | Fear | 458 |
| **5** | Happy | 614 |
| **6** | Normal | 913 |
| **7** | Sad | 540 |
| **8** | Surprised | 602 |

Table 5 : CK+ dataset Instances

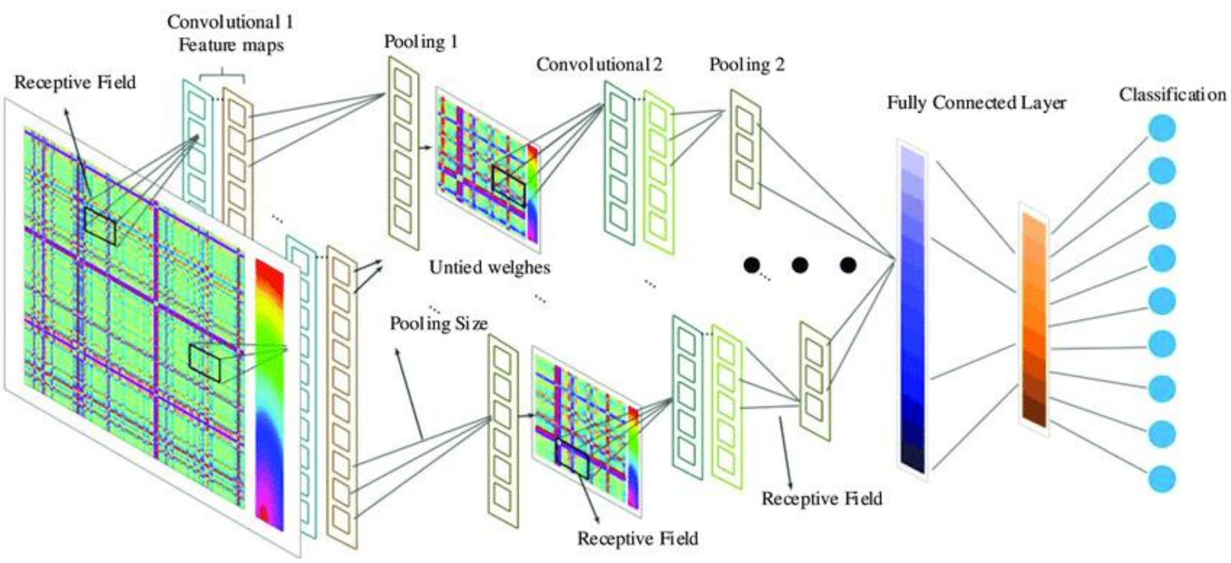


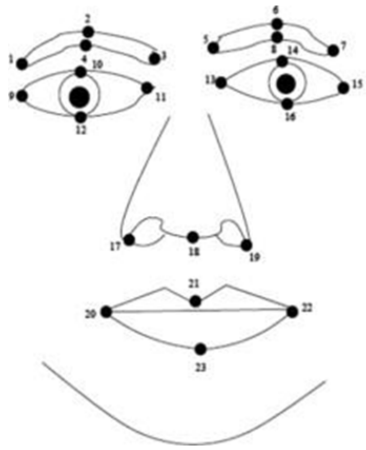

Figure 5: Convolutional Neural Network


International Journal of DeepTech in Medical Science and Technology,
Vol 1, No. 1, September-October 2025
ISSN: 3088-2370 (Online) | © Geoinformatics International

(source: https://www.bing.com/images/blob?bcid=S7F-N9gpFaoEcw )

***Geometrical and Grid Feature Extraction***

It is one of the easiest and primary extraction used for facial as this consume very less space and also it has been used by many researchers, the efficiency of geometrical feature extraction with cnn as shown in Figure 5 is 73.98% [56], [57] and this is the primary source of extraction and with the help of simple data process the features are extracted and labelled as below Figure 6.



| Geometric features | Number of features | Description |
|---|---|---|
| Eyebrow | 4x2 | Two extreme corners, upper and lower mid points |
| Eyes | 4x2 | Two extreme corners, upper and lower mid points |
| Nose | 3 | Two nostrils and nose tip |
| Lip | 4 | Two extreme corners, upper and lower mid points |

Figure 6: Geometrical Feature Extraction (source: literature review)

Grid feature extraction is one of the latest use extraction with use of 98% machine learning algorithm [58].In this extraction each nodes contains some weight and bias too. The grid is drawn over the face and analyses process take place and then processing is done using image processing and the facial feature extraction take place as shown in Figure 7.

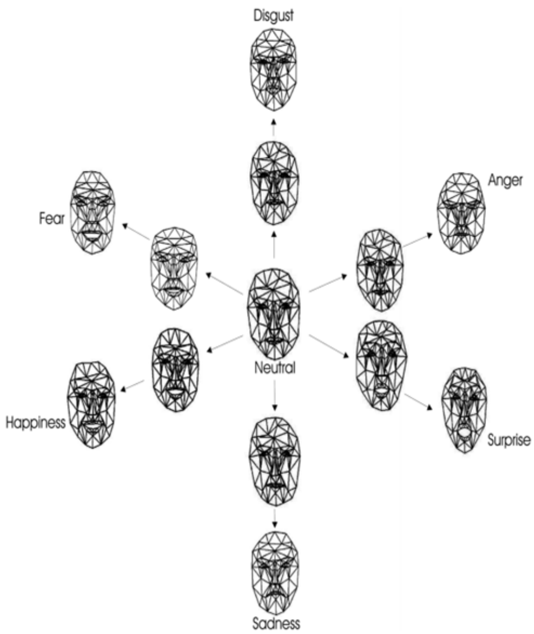


Figure 7: Grid feature Extraction ([56], [57])

**Facial Analysis**
In this the pattern of emotion to be analyzed with facial marking is being done as Table 6.

| Emotion | Motion of Facial parts |
|---|---|
| Happy | open eyes, open mouth,lip corner pulled, cheeks raised |
| Sad | Outer eyebrow down,inner eyebrows raised, eyes closed, lip corner down |
| Suprise | Eyebrow up, open eyes, jaw dropped |
| Anger | Eyebrow pulled down, open eyes, lip tightened |
| Fear | Outer eyebrow down, inner eyebrow up, mouth open |

Table 6: Facial Analysis

The proposed study was designed to develop and evaluate a multi-facial emotion recognition system utilizing diverse input modalities, including static images, video recordings, and real-time webcam feeds as shown in Figure 8. Initially, the publicly available CK+ dataset was employed for model training, validation, and testing, with the data split in a stratified manner to ensure balanced representation of all emotional categories. A custom Fusion Convolutional Neural Network (Fusion CNN) architecture was implemented to manage the high-dimensional feature extraction and classification tasks. This hybrid deep learning model leveraged the spatial capabilities of CNNs and the layered complexity of deep neural networks to enhance recognition accuracy.

Following model training on the curated dataset, the system was deployed in a real-time environment using OpenCV-integrated webcam input. The trained model was tested on live subjects to evaluate its practical performance in detecting emotions such as happiness, sadness, and disgust. To assess the perceptual validity of the model’s predictions, a user feedback mechanism was incorporated. Participants were prompted to confirm the correctness of the model's real-time emotional prediction through a structured questionnaire administered via Google Forms. This allowed for an empirical correlation between algorithmic output and subjective human feedback, providing a foundation for calculating real-world accuracy and system efficiency.

The entire implementation was carried out using Python, with supportive libraries such as OpenCV for image processing, Pandas for data handling, and custom kernel functions for model tuning. Furthermore, R programming language was employed for statistical modelling and hypothesis testing. Specifically, chi-square tests and regression analyses were conducted to evaluate the relationship between predicted and reported emotions, offering insights into model reliability and generalizability.




ISSN: 3088-2370 (Online) 

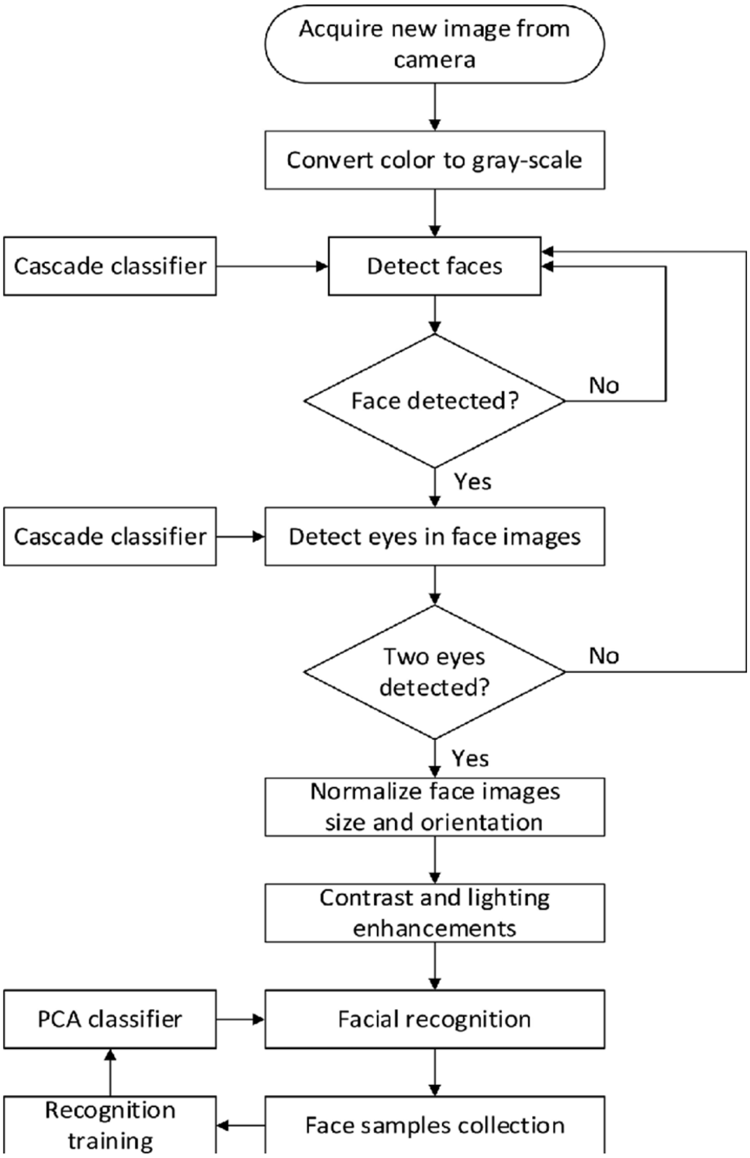



Figure 8: Flowchart of working Methodology

## 4. Results and Discussion

In the initial phase of this study, precise detection of facial landmark grid points was a critical step. The Canadian Grid Node system enabled accurate masking and identification of Facial Action Units (FAUs), as demonstrated in Figures 9 and 10. These figures illustrate the landmark points — such as the nose tip and eye centres — in a 2D grid format, effectively mapped on facial structures prior to preprocessing.

```
[7]:  train.isnull().sum()

[7]:  left_eye_center_x              10
      left_eye_center_y              10
      right_eye_center_x             13
      right_eye_center_y             13
      left_eye_inner_corner_x      4778
      left_eye_inner_corner_y      4778
      left_eye_outer_corner_x      4782
      left_eye_outer_corner_y      4782
      right_eye_inner_corner_x     4781
      right_eye_inner_corner_y     4781
      right_eye_outer_corner_x     4781
      right_eye_outer_corner_y     4781
      left_eyebrow_inner_end_x     4779
      left_eyebrow_inner_end_y     4779
      left_eyebrow_outer_end_x     4824
      left_eyebrow_outer_end_y     4824
      right_eyebrow_inner_end_x    4779
      right_eyebrow_inner_end_y    4779
      right_eyebrow_outer_end_x    4813
      right_eyebrow_outer_end_y    4813
      nose_tip_x                      0
      nose_tip_y                      0
      mouth_left_corner_x          4780
      mouth_left_corner_y          4780
      mouth_right_corner_x         4779
      mouth_right_corner_y         4779
      mouth_center_top_lip_x       4774
      mouth_center_top_lip_y       4774
      mouth_center_bottom_lip_x      33
      mouth_center_bottom_lip_y      33
      Image                           0
      dtype: int64
```

Figure 9: Code to get the FAU point. (The facial landmarks are pointed here as with each grid note, as we can see for ex. nose tip and left eye center, all the point are in 2D format)

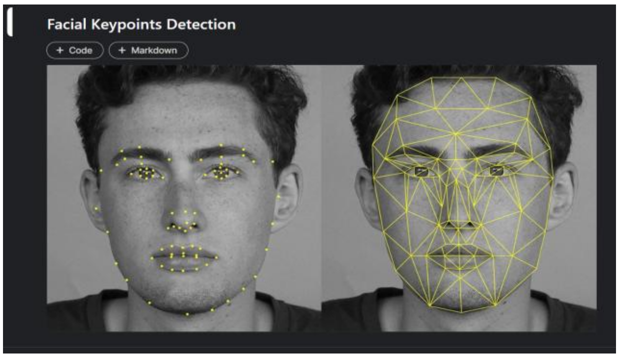



Figure 10: Grid Node Mask on Face with FAUs. (when the masking is done and point are identified and facial node are captured before pre processing)

Also, when the same was run over the dataset, the FAU was performed can be seen in the figure 11 below, this entire system is processed on the dataset considered.

Figure 11: FAUs over dataset (CK+)

Next, the hyperparameter tunning was one of the most important and crucial step because the difference in the accuracy was clearly been seen and also the selection of tunning parameter was one of the most care taken step, which can been seen on the figure 12.

```
%%time
history = model.fit(X_train,y_train,epochs = 10, batch_size = 32,validation_split = 0.2)

177/177 [==============================] - 185s 1s/step - loss: 3.5127 - root_mean_squared_error: 4.7714 - accuracy:
0.5767 - val_loss: 4.6713 - val_root_mean_squared_error: 5.6209 - val_accuracy: 0.6908
Epoch 8/10
177/177 [==============================] - 184s 1s/step - loss: 3.3821 - root_mean_squared_error: 4.6116 - accuracy:
0.6061 - val_loss: 2.3590 - val_root_mean_squared_error: 3.1863 - val_accuracy: 0.6929
Epoch 9/10
177/177 [==============================] - 185s 1s/step - loss: 3.3459 - root_mean_squared_error: 4.5604 - accuracy:
0.6219 - val_loss: 5.5018 - val_root_mean_squared_error: 6.1752 - val_accuracy: 0.6950
Epoch 10/10
177/177 [==============================] - 185s 1s/step - loss: 3.2291 - root_mean_squared_error: 4.4298 - accuracy:
0.6377 - val_loss: 4.4289 - val_root_mean_squared_error: 5.2131 - val_accuracy: 0.6716
CPU times: user 1h 57min 19s, sys: 14.2 s, total: 1h 57min 33s
Wall time: 31min
```

```
%%time
history = model.fit(X_train,y_train,epochs = 200)

221/221 [==============================] - 2s 7ms/step - loss: 0.9741 - root_mean_squared_error: 2.1510 - accuracy: 0.
7849
Epoch 198/200
221/221 [==============================] - 2s 7ms/step - loss: 0.9674 - root_mean_squared_error: 2.1311 - accuracy: 0.
7858
Epoch 199/200
221/221 [==============================] - 2s 7ms/step - loss: 0.9695 - root_mean_squared_error: 2.1370 - accuracy: 0.
7908
Epoch 200/200
221/221 [==============================] - 2s 7ms/step - loss: 0.9721 - root_mean_squared_error: 2.1354 - accuracy: 0.
7844
CPU times: user 11min 1s, sys: 48.9 s, total: 11min 50s
Wall time: 6min 23s
```

+ Code + Markdown

Figure 12: Hyperparameter tuning for accuracy.

When the model was being trained under the CNN model, the accuracy was obtained as 72.16% with the dataset used for this. The Figure 13 and 14 are the accuracy for the CNN model only.


ISSN: 3088-2370 (Online) | © Geoinformatics International

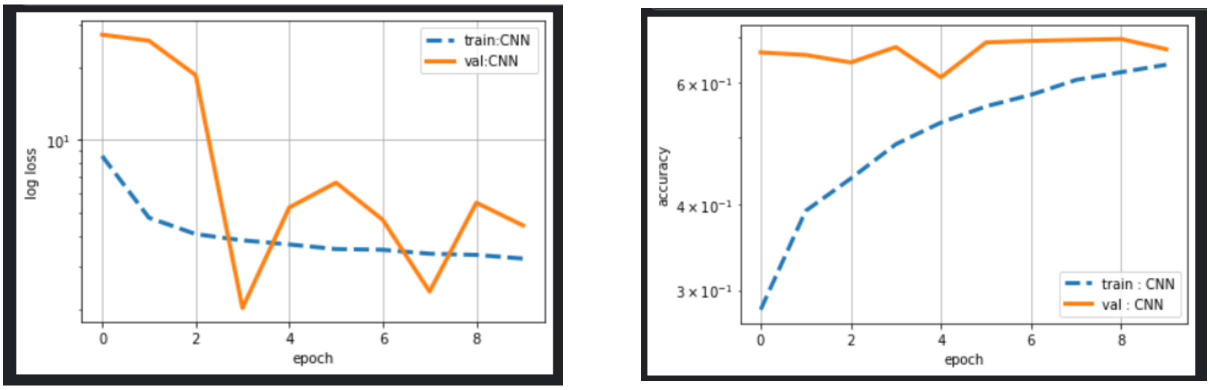



Figure 13: CNN model accuracy plot from the model prepared by changing the hyper parameter settings

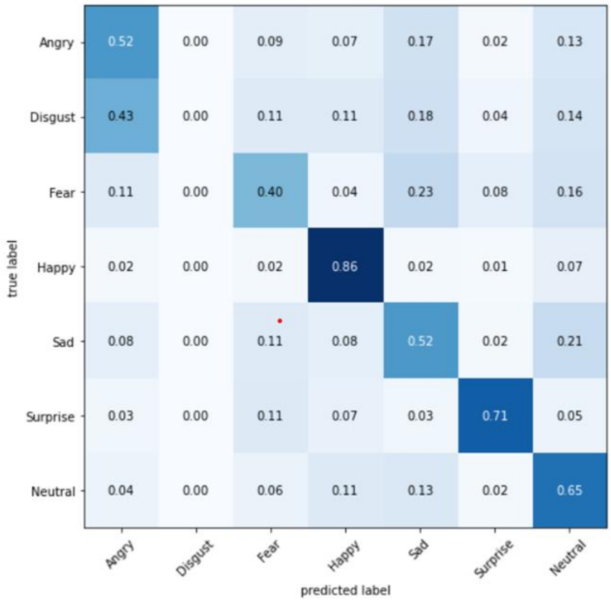



Figure 14: Confusion Matrix for CNN

The Convolutional Neural Network (CNN) was first employed for classification, yielding an accuracy of 72.16%. The accuracy trends, shown in Figure 4.1.5, demonstrate the impact of tuned hyperparameters on performance. Further validation through the confusion matrix in Figure 4.1.8 revealed that the model was most effective at recognizing the "Happy" class, while the "Disgust" class had negligible representation, possibly due to class imbalance in the dataset. To compare model performance, VGG-16 was also employed under similar experimental conditions. The accuracy achieved was 68.13%, as shown in Figure 15, with training dynamics presented in Figure 4.1.10. While CNN slightly outperformed VGG-16, the latter still exhibited competitive results, affirming its applicability for emotion recognition tasks.

```
fig = pyplot.figure(1, (14, 14))

k = 0
for label in sorted(df.emotion.unique()):
    for j in range(7):
        px = df[df.emotion==label].pixels.iloc[k]
        px = np.array(px.split(' ')).reshape(48, 48).astype('float32')

        k += 1
        ax = pyplot.subplot(15, 15, k)
        ax.imshow(px, cmap = 'gray')
        ax.set_xticks([])
        ax.set_yticks([])
        ax.set_title(emotion_label_to_text[label])
        pyplot.tight_layout()
```


Figure 15: Accuracy for VGG16.

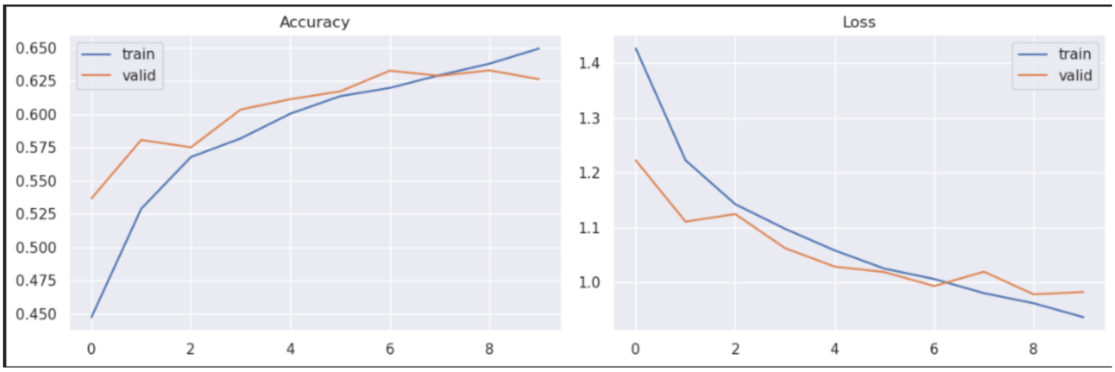



Figure 16: Accuracy for VGG16 plot

The third model implemented was a CNN built using the Keras framework, which achieved an accuracy of 62.5%. Although this was the lowest among the three, it still demonstrated acceptable performance and served as a baseline to validate the system pipeline. The training results for this model are presented in Figure 16.

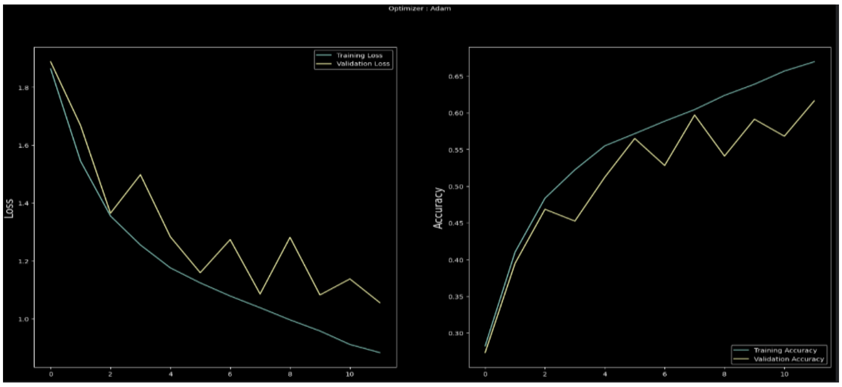



Figure 17: Plot for CNN with Keras

| Model | Accuracy (%) | Remarks |
|---|---|---|
| CNN (Custom) | 72.16 | Highest accuracy among the tested models after hyperparameter tuning. |
| VGG-16 | 68.13 | Slightly lower performance; still effective for emotion recognition. |
| CNN (Keras) | 62.5 | Moderate accuracy; performance lower than other models. |

Table 7: Accuracy Comparison of Different Models for Emotion Recognition

Following model evaluations, the system was tested in real-time applications using images, videos, and webcam input. Figure 17 displays the output for a sample dataset image. In video-based testing, the system captured sequential frames and processed them individually. For instance, a 3-minute video produced 7241 frames, all processed for emotion classification. Frame-wise emotional transitions were captured and visualised in Figure 18 and a summary of the video processing is shown in Figure 19.
*(The eyes have been masked because of data privacy with the user).*

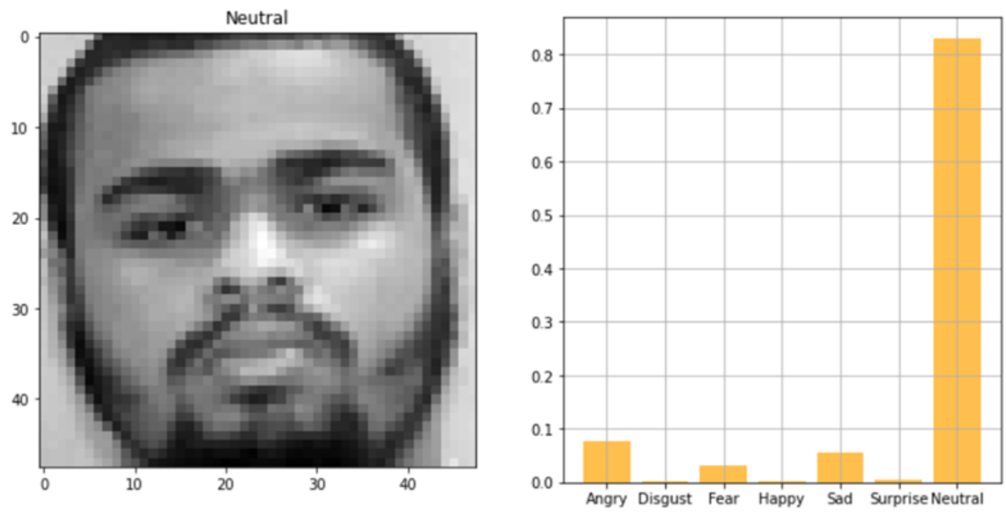



Figure 18: Result for Objective 1 using dataset image.

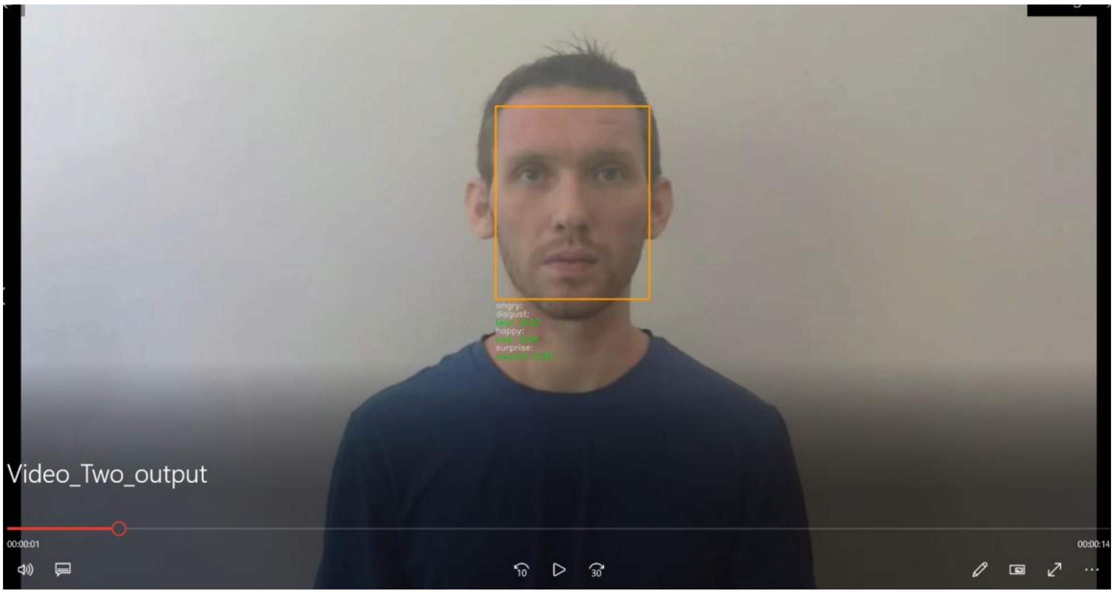



Figure 19: Video output being captured for Objective 1

Lastly, the system was deployed on a live webcam demonstration, as shown in Figure 20. The DeepFace package was integrated for simplified real-time facial expression classification. Due to privacy concerns, user eye regions were masked.

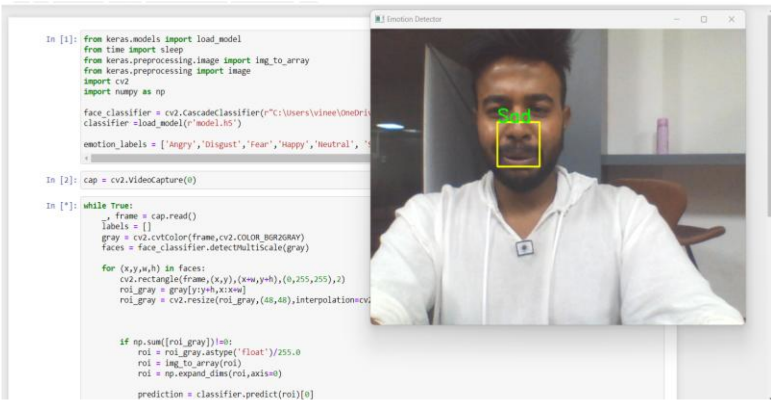



Figure 20: Webcam output for Obj 1

Overall, the proposed system demonstrated effective facial emotion recognition through accurate landmark detection and robust model training. The Canadian Grid Node enabled precise masking of Facial Action Units (FAUs), which was essential for consistent feature extraction across the dataset. Among the models tested, the custom CNN achieved the highest accuracy at 72.16%, followed by VGG-16 with 68.13%, and the Keras-based CNN with 62.5%. Hyperparameter tuning played a crucial role in enhancing model performance. The system's applicability was further validated through real-time emotion detection using images, videos, and webcam inputs, showcasing its potential for real-world emotion analysis tasks. These results affirm the feasibility of using deep learning frameworks for accurate and scalable facial emotion recognition[59], [60], [61].

## 5. Conclusion and Future Scope

The present study aimed to develop a robust and scalable deep learning-based framework for multi-facial emotion recognition across heterogeneous input formats including static images, video streams, and real-time webcam feeds. Through a comprehensive evaluation of various neural architectures, including a custom Convolutional Neural Network (CNN), a Recurrent Neural Network (RNN), and a Keras-implemented CNN, the study identified that the custom CNN outperformed the others, achieving an accuracy of 72.16%. This significant performance differential underscores the critical role of architecture-specific customization and pre-processing in enhancing classification efficacy for affective computing tasks. The results further demonstrate the feasibility of deploying such systems in real-world environments where emotion recognition must be performed with high reliability and across diverse visual inputs[62], [63], [64], [65].

A key contribution of this research lies in its methodological approach, which integrates efficient facial landmark extraction, noise reduction techniques, and structured training pipelines to enhance recognition stability and accuracy. The application of facial geometry analysis through the Canadian Grid Node System, combined with OpenCV-based alignment and normalization, facilitated more consistent emotional state detection across diverse lighting conditions, facial orientations, and subject demographics. This has practical implications for domains such as healthcare, e-learning, intelligent surveillance, and human-computer interaction, where emotional context significantly influences decision-making and user experience [66], [67], [68], [69], [70].

Moreover, the study highlights important considerations for ethical deployment, particularly in safeguarding against algorithmic bias and ensuring inclusivity across demographic variations. The findings contribute to the discourse on responsible AI design by emphasizing transparency, interpretability, and the ethical use of facial emotion recognition technologies. The model's performance, when benchmarked across different modalities, reinforces the potential of deep learning to decode human affect with a considerable degree of accuracy, while also pointing to limitations inherent in current datasets and evaluation frameworks [71], [72], [73], [74], [75].

Looking forward, future research may explore the integration of multimodal data sources, such as voice, text, and physiological signals, to augment the emotion classification framework. Additionally, the implementation of attention-based models or transformer architectures can be examined for capturing more nuanced affective states in real time. Expanding the training corpus to include cross-cultural, age-diverse, and expressionally varied datasets will enhance the generalizability and fairness of the system. Beyond technical improvements, clinical validation and interdisciplinary collaborations with psychologists, neuroscientists, and UX designers will be instrumental in translating this work into impactful applications. Ensuring adherence to ethical standards, particularly concerning privacy, consent, and fairness, will remain a foundational imperative as emotion recognition technologies advance toward widespread adoption [76], [77], [78], [79], [80].

## REFERENCES


[1] H. Magarde, K. Shokat, … R. J.-2024 F. I., and undefined 2024, “Driver Drowsiness Detection System and Techniques: Critical Analysis,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10698711/

[2] Z. Ghrban, N. E. A.-I. J. of Computing, and undefined 2023, “Gender and Age Estimation from Human Faces Based on Deep Learning Techniques: A Review,” *researchgate.net*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.researchgate.net/profile/Nidhal-El-Abbadi/publication/372116386_Gender_and_Age_Estimation_from_Human_Faces_Based_on_Deep_Learning_Techniques_A_Review/links/64dd244dad846e2882966b90/Gender-and-Age-Estimation-from-Human-Faces-Based-on-Deep-Learning-Techniques-A-Review.pdf

[3] M. B.-R. G. preprint and undefined 2022, “Review of the latest advanced driver assistance systems,” *researchgate.net*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.researchgate.net/profile/Mouadh-Bouitaoune/publication/364302770_REVIEW_OF_THE_LATEST_ADVANCED_DRIVER_ASSISTANCE_SYSTEMS/links/634448b476e39959d6b33a52/REVIEW-OF-THE-LATEST-ADVANCED-DRIVER-ASSISTANCE-SYSTEMS.pdf

[4] G. Sikander, S. Anwar, G. Husnain, … R. T.-I., and undefined 2023, “An adaptive snake based shadow segmentation for robust driver fatigue detection: A 3D facial feature based photometric stereo perspective,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10242059/

[5] S. Ali, M. Hanzla, A. R.-2022 24th I. Multitopic, and undefined 2022, “Vehicle detection and tracking from UAV imagery via Cascade classifier,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9972959/

[6] I. Nosheen, A. Naseer, A. J.-2024 5th International, and undefined 2024, “Efficient vehicle detection and tracking using blob detection and kernelized filter,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10473292/

[7] U. A.-2022 24th I. M. Conference and undefined 2022, “Human activity recognition via smartphone embedded sensor using multi-class SVM,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9972927/

[8] “Multi-Facial Emotion Recognition Using Fusion CNN... - Google Scholar.” Accessed: Sep. 13, 2025. [Online]. Available: https://scholar.google.com/scholar?start=50&q=Multi-Facial+Emotion+Recognition+Using+Fusion+CNN+on+Static+and+Real-

Time+Inputs:+A+Deep+Learning+Approach&hl=en&as_sdt=0,5&as_ylo=2020&as_yhi=2024

[9] M. Sindhuja, S. Nayak, A. K.-A. C. Proceedings, and undefined 2024, "Comparative study on existing driver drowsiness detection systems," *pubs.aip.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://pubs.aip.org/aip/acp/article-abstract/3075/1/020180/3305122

[10] U. Azmat, "Human Activity Recognition via Smartphone Embedded Sensor using Multi-Class SVM," *researchgate.net*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.researchgate.net/profile/Usman-Azmat/publication/363699434_Human_Activity_Recognition_via_Smartphone_Embedded_Sensor_using_Multi-Class_SVM/links/632ace4f071ea12e364c54e1/Human-Activity-Recognition-via-Smartphone-Embedded-Sensor-using-Multi-Class-SVM.pdf

[11] Y. Zhong, T. Liu, Y. Yang, Z. Shao, Y. Shang, and H. Ding, "A Novel Fatigue Detection Method Based on Video Transformer," *Springer*, vol. 1338 LNEE, pp. 495–505, 2025, doi: 10.1007/978-981-96-2204-7_47.

[12] L. Mou, Y. Zhao, C. Zhou, B. Yin, … W. G.-2022 I. 5th, and undefined 2022, "A review of personalized health navigation for drivers," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9874650/

[13] "Multi-Facial Emotion Recognition Using Fusion CNN... - Google Scholar." Accessed: Sep. 13, 2025. [Online]. Available: https://scholar.google.com/scholar?start=40&q=Multi-Facial+Emotion+Recognition+Using+Fusion+CNN+on+Static+and+Real-Time+Inputs:+A+Deep+Learning+Approach&hl=en&as_sdt=0,5&as_ylo=2020&as_yhi=2024

[14] Y.-H. Tsai, M.-C. Tseng, Chin-Yung, Wang, and Chiou-Shann, "Detection of Driver Drowsiness Using Multi-Task Learning," *csie.ntu.edu.tw*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.csie.ntu.edu.tw/~fuh/personal/DetectionofDriverDrowsinessUsingMulti-TaskLearning.pdf

[15] A. Reddy, I. Khan, L. Thampi, A. MA, and A. Kumar, "Smart Driver Assistance: Real-Time Drowsiness Detection Leveraging Facial Cues with MediaPipe and OpenCV," 2024, Accessed: Sep. 13, 2025. [Online]. Available: https://assets-eu.researchsquare.com/files/rs-4642662/v1_covered_e46bb385-d6d1-438d-bd77-6566362d2794.pdf

[16] X. Zeng, F. Wang, B. Wang, … C. W.-I. O. J. of, and undefined 2022, "In-vehicle sensing for smart cars," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9773999/

[17] A. Soultana, F. Benabbou, … N. S.-I. J. of, and undefined 2022, "A Systematic Literature Review of Driver Inattention Monitoring Systems for Smart Car.," *researchgate.net*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.researchgate.net/profile/Faouzia-Benabbou/publication/363155592_A_Systematic_Literature_Review_of_Driver_Inattention_Monitoring_Systems_for_Smart_Car/links/638079ffc2cb154d2925dbb4/A-Systematic-Literature-Review-of-Driver-Inattention-Monitoring-Systems-for-Smart-Car.pdf

[18] K. Xu, F. Li, D. Chen, L. Zhu, Q. W.-I. Access, and undefined 2024, "Fusion of lightweight networks and DeepSort for fatigue driving detection tracking algorithm," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10496102/

[19] C. Bisogni, L. Cascone, M. Nappi, and C. Pero, "Iot-enabled biometric security: enhancing smart car safety with depth-based head pose estimation," *dl.acm.org*, vol. 20, no. 6, Mar. 2024, doi: 10.1145/3639367.

[20] H. Lin, K. T.-I. C. on S. C. and Green, and undefined 2023, "A Technique for Authentic Fatigue Driving Detection Using Nighttime Infrared Images," *Springer*, vol. 1989 CCIS, pp. 123–145, 2025, doi: 10.1007/978-3-031-70966-1_6.

[21] N. N. Pandey and N. B. Muppalaneni, "A survey on visual and non-visual features in Driver's drowsiness detection," *Springer*, vol. 81, no. 26, pp. 38175–38215, Nov. 2022, doi: 10.1007/S11042-022-13150-1.

[22] M. Kamti, R. I.-T. research record, and undefined 2022, "Evolution of driver fatigue detection techniques—A review from 2007 to 2021," *journals.sagepub.com*, vol. 2676, no. 12, pp. 485–507, Dec. 2022, doi: 10.1177/03611981221096118.

[23] W. Xiao *et al.*, "Fatigue driving recognition method based on multi-scale facial landmark detector," *mdpi.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.mdpi.com/2079-9292/11/24/4103

[24] Z. Gao *et al.*, "Semantically-Enhanced Feature Extraction with CLIP and Transformer Networks for Driver Fatigue Detection," *mdpi.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/24/24/7948

[25] C. Yan, Y. Zhang, X. L.-2024 2nd I. C. on, and undefined 2024, "Fatigue Detection Method for UAV Remote Pilot based on YOLOv8," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10988450/

[26] S. Fu *et al.*, "Advancements in the intelligent detection of driver fatigue and distraction: A comprehensive review," *mdpi.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.mdpi.com/2076-3417/14/7/3016

[27] M. Arava, D. S.-P. C. Science, and undefined 2024, “Integrating lightweight YOLOv5s and facial 3D keypoints for enhanced fatigued-driving detection,” *peerj.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://peerj.com/articles/cs-2447/

[28] L. Lilly Thampi, K. Lata Kashyap, A. Karthikeya Reddy, I. M. Ahmed Khan IIIT Kottayam A Aswathy, A. Kumar, and S. Kumar, “Smart Driver Assistance: Real-Time Drowsiness Detection Using CNN and Computer Vision,” 2024, doi: 10.21203/rs.3.rs-4962655/v1.

[29] G. S. Commons, “Unobtrusive Assessment Of Student Engagement Levels In Online Classroom Environment Using Emotion Analysis,” 2021, Accessed: Sep. 13, 2025. [Online]. Available: https://digitalcommons.georgiasouthern.edu/etd/2265/

[30] Y. Albadawi, M. Takruri, M. A.- Sensors, and undefined 2022, “A review of recent developments in driver drowsiness detection systems,” *mdpi.com*, vol. 2022, p. 2069, 2022, doi: 10.3390/s22052069.

[31] “Multi-Facial Emotion Recognition Using Fusion CNN... - Google Scholar.” Accessed: Sep. 13, 2025. [Online]. Available: https://scholar.google.com/scholar?start=30&q=Multi-Facial+Emotion+Recognition+Using+Fusion+CNN+on+Static+and+Real-Time+Inputs:+A+Deep+Learning+Approach&hl=en&as_sdt=0,5&as_ylo=2020&as_yhi=2024

[32] S. Mandia, R. Mitharwal, and K. Singh, “Automatic student engagement measurement using machine learning techniques: A literature study of data and methods,” *Springer*, vol. 83, no. 16, pp. 49641–49672, May 2024, doi: 10.1007/S11042-023-17534-9.

[33] N. Noor, N. S.-B. of E. E. and, and undefined 2022, “Review on facial expression modeling,” *mail.beei.org*, vol. 11, no. 2, pp. 779–784, 2022, doi: 10.11591/eei.v11i2.3558.

[34] J. Hong, J. Shin, J. Choi, M. K.-P. of the IEEE/CVF, and undefined 2024, “Robust eye blink detection using dual embedding video vision transformer,” *openaccess.thecvf.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://openaccess.thecvf.com/content/WACV2024/html/Hong_Robust_Eye_Blink_Detection_Using_Dual_Embedding_Video_Vision_Transformer_WACV_2024_paper.html

[35] X. Wang, Z. Guo, H. Duan, and W. Chen, “RETRACTED CHAPTER: An Efficient Channel Attention CNN for Facial Expression Recognition,” *Springer*, vol. 808 LNEE, pp. 75–82, 2022, doi: 10.1007/978-981-16-6554-7_8.

[36] J. Zhao, W. Zhong, D. Zhang, … Y. W.-E. 2022; 4th, and undefined 2022, “EVAI: An approach for real-time video-based emotion recognition,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10048233/

[37] A. Srivastava, … S. B.-… C. on C., and undefined 2022, "Real-Time Based Driver's Drowsiness and Fatigue Detection System," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9996023/

[38] "Multi-Facial Emotion Recognition Using Fusion CNN... - Google Scholar." Accessed: Sep. 13, 2025. [Online]. Available: https://scholar.google.com/scholar?start=20&q=Multi-Facial+Emotion+Recognition+Using+Fusion+CNN+on+Static+and+Real-Time+Inputs:+A+Deep+Learning+Approach&hl=en&as_sdt=0,5&as_ylo=2020&as_yhi=2024

[39] M. Alkinani, W. Khan, Q. A.-I. Access, and undefined 2020, "Detecting human driver inattentive and aggressive driving behavior using deep learning: Recent advances, requirements and open challenges," *academia.edu*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.academia.edu/download/112298576/09107077.pdf

[40] Q. Abbas, A. A.-I. Access, and undefined 2021, "A methodological review on prediction of multi-stage hypovigilance detection systems using multimodal features," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9385066/

[41] X. Ran, S. He, R. L.- Sensors, and undefined 2023, "Research on fatigued-driving detection method by integrating lightweight yolov5s and facial 3d keypoints," *mdpi.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/23/19/8267

[42] Y. Ed-Doughmi, N. Idrissi, Y. H.-J. of imaging, and undefined 2020, "Real-time system for driver fatigue detection based on a recurrent neuronal network," *mdpi.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.mdpi.com/2313-433X/6/3/8

[43] X. Jin, X. Song, X. Wu, and W. Yan, "Transformer embedded spectral-based graph network for facial expression recognition," *Springer*, vol. 15, no. 6, pp. 2063–2077, Jun. 2024, doi: 10.1007/S13042-023-02016-Z.

[44] X. Jin, Z. Lai, Z. J.-I. T. on I. Processing, and undefined 2021, "Learning dynamic relationships for facial expression recognition based on graph convolutional network," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9509366/

[45] G. Du, L. Zhang, K. Su, X. Wang, … S. T.-I. transactions on, and undefined 2022, "A multimodal fusion fatigue driving detection method based on heart rate and PERCLOS," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9785381/

[46] S. Wu, L. Zhou, Z. Hu, … J. L. N. N. and L., and undefined 2022, "Hierarchical context-based emotion recognition with scene graphs," *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9868807/

[47] S. Das, R. B.-2023 15th I. C. on, and undefined 2023, “Vision-Based Fatigue Detection In Drivers Using Multi-Facial Feature Fusion,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10099741/

[48] B. L.-E. E. Trans. e Learn. and undefined 2021, “Facial expression recognition via transfer learning.,” *researchgate.net*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.researchgate.net/profile/Bin-Li-268/publication/350746086_Facial_expression_recognition_via_transfer_learning/links/61c3175ec48a3d26b747c138/Facial-expression-recognition-via-transfer-learning.pdf?origin=journalDetail&_tp=eyJwYWdlIjoiam91cm5hbERldGFpbCJ9

[49] X. Jin, Z. J.- Neurocomputing, and undefined 2021, “MiniExpNet: A small and effective facial expression recognition network based on facial local regions,” *Elsevier*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231221011619

[50] S. Gunanandhini, T. Thilagam, … G. M.-2024 I., and undefined 2024, “DrowsiGuard: Machine Learning-Based Driver Drowsiness Detection System for Enhanced Road Safety,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11026736/

[51] H. Yang, L. Liu, W. Min, … X. Y.-I. T. on, and undefined 2020, “Driver yawning detection based on subtle facial action recognition,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9057686/

[52] “Multi-Facial Emotion Recognition Using Fusion CNN... - Google Scholar.” Accessed: Sep. 13, 2025. [Online]. Available: https://scholar.google.com/scholar?start=10&q=Multi-Facial+Emotion+Recognition+Using+Fusion+CNN+on+Static+and+Real-Time+Inputs:+A+Deep+Learning+Approach&hl=en&as_sdt=0,5&as_ylo=2020&as_yhi=2024

[53] E. Civik, U. Y.-M. and Microsystems, and undefined 2023, “Real-time driver fatigue detection system with deep learning on a low-cost embedded system,” *Elsevier*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0141933123000972

[54] Z. Zhao, Q. Liu, S. W.-I. T. on Image, and undefined 2021, “Learning deep global multi-scale and local attention features for facial expression recognition in the wild,” *ieeexplore.ieee.org*, vol. 30, p. 2021, 2021, doi: 10.1109/TIP.2021.3093397.

[55] P. L. Mehta and A. K.-A. at S. 4030258, “Livai: A Novel Resource-Efficient Real-Time Facial Emotion Recognition System Based on a Custom Deep Cnn Model,” *papers.ssrn.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4030258

[56] M. Ramzan, A. Abid, M. Fayyaz, … T. A.-I., and undefined 2024, “A novel hybrid approach for driver drowsiness detection using a custom deep learning model,” *ieeexplore.ieee.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10623142/

[57] X. Liang *et al.*, “Patch attention layer of embedding handcrafted features in CNN for facial expression recognition,” *mdpi.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/21/3/833

[58] L. Kong, K. Xie, K. Niu, J. He, W. Z.- Sensors, and undefined 2024, “motion tracking convolutional neural network with bidirectional long short-term memory: Non-invasive fatigue detection method based on multi-modal fusion,” *mdpi.com*, Accessed: Sep. 13, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/24/2/455

[59] A. Rachid Hazourli, A. Djeghri, H. Salam, A. Othmani, H. Salam Emlyon, and F. Alice Othmani, “Deep multi-facial patches aggregation network for facial expression recognition,” *arxiv.org*, Accessed: Sep. 13, 2025. [Online]. Available: https://arxiv.org/abs/2002.09298

[60] “Multi-Facial Emotion Recognition Using Fusion CNN... - Google Scholar.” Accessed: Sep. 13, 2025. [Online]. Available: https://scholar.google.com/scholar?q=Multi-Facial+Emotion+Recognition+Using+Fusion+CNN+on+Static+and+Real-Time+Inputs%3A+A+Deep+Learning+Approach&hl=en&as_sdt=0%2C5&as_ylo=2024&as_yhi=2020

[61] R. Helaly, S. Messaoud, S. Bouaafia, M. A. Hajjaji, and A. Mtibaa, “DTL-I-ResNet18: facial emotion recognition based on deep transfer learning and improved ResNet18,” *Springer*, vol. 17, no. 6, pp. 2731–2744, Sep. 2023, doi: 10.1007/S11760-023-02490-6.

[62] K. N. Kumar Tataji, M. N. Kartheek, and M. V. N. K. Prasad, “CC-CNN: A cross connected convolutional neural network using feature level fusion for facial expression recognition,” *Springer*, vol. 83, no. 9, pp. 27619–27645, Mar. 2024, doi: 10.1007/S11042-023-16433-3.

[63] A. R. Hazourli, A. Djeghri, H. Salam, and A. Othmani, “Multi-facial patches aggregation network for facial expression recognition and facial regions contributions to emotion display,” *Springer*, vol. 80, no. 9, pp. 13639–13662, Apr. 2021, doi: 10.1007/S11042-020-10332-7.

[64] H. Boughanem, H. Ghazouani, and W. Barhoumi, “Facial emotion recognition in-the-wild using deep neural networks: a comprehensive review,” *Springer*, vol. 5, no. 1, Jan. 2024, doi: 10.1007/S42979-023-02423-7.

[65] S. Saleem, … S. Z.-J. of P., and undefined 2021, “Real-life dynamic facial expression recognition: a review,” *iopscience.iop.org*, doi: 10.1088/1742-6596/1963/1/012010/META.

[66] J. van Pelt, … A. van O.-N. I. 2007: P. of the, and undefined 2008, "Neuroinformatics in the Netherlands," *Springer*, pp. 673–677, 2008, doi: 10.1007/978-1-4020-8387-7_116.

[67] R. Abi, "Ethical and Explainable AI in Data Science for Transparent Decision-Making Across Critical Business Operations," *researchgate.net*, Accessed: Aug. 27, 2025. [Online]. Available: https://www.researchgate.net/profile/Roland-Abi/publication/393029564_Ethical_and_Explainable_AI_in_Data_Science_for_Transparent_Decision-Making_Across_Critical_Business_Operations/links/6861d4c6e4632b045dc87161/Ethical-and-Explainable-AI-in-Data-Science-for-Transparent-Decision-Making-Across-Critical-Business-Operations.pdf

[68] K. Dwivedi, B. Chugh, … J. C.-… on N. and, and undefined 2025, "A Systematic Review on Incorporation of Artificial Intelligence in Precision Healthcare," *ieeexplore.ieee.org*, Accessed: Aug. 27, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11102485/

[69] A. Kalejaiye, K. Shallom, E. C.-I. J. S. R. Arch, and undefined 2025, "Implementing federated learning with privacy-preserving encryption to secure patient-derived imaging and sequencing data from cyber intrusions," *academia.edu*, Accessed: Aug. 27, 2025. [Online]. Available: https://www.academia.edu/download/123798776/IJSRA_2025_2120.pdf

[70] A. D.-J. of S. and R. Archive and undefined 2025, "Artificial Intelligence and machine learning in fraud detection for digital payments," *eprint.scholarsrepository.com*, vol. 2025, no. 03, pp. 714–719, 2025, doi: 10.30574/ijsra.2025.15.3.1784.

[71] S. Karkhur, A. Beri, V. Verma, S. Gupta, P. S.- Cureus, and undefined 2025, "Artificial Intelligence in Neuro-Ophthalmology: Opportunities for the Diagnosis of Optic Neuropathies and Visual Pathway Disorders," *cureus.com*, 2025, doi: 10.7759/cureus.90142.

[72] J. Conde, C. Speelman, M. J.-A. and Ethics, and undefined 2025, "A new epoch of face analytics: technological evolution through ethical and legal challenges," *Springer*, Jun. 2025, doi: 10.1007/S43681-025-00678-9.

[73] M. Alqadi, S. V.-C. N. and N. Reports, and undefined 2025, "Artificial Intelligence in Vascular Neurology: Applications, Challenges, and a Review of AI Tools for Stroke Imaging, Clinical Decision Making, and Outcome Prediction," *Springer*, vol. 25, no. 1, Dec. 2025, doi: 10.1007/S11910-025-01422-W.

[74] M. D.-A. A. Directive and undefined 2025, "Regulating the Neural Frontier: Why Brain-Computer Interfaces (BCIs) Require a New Regulatory Model," *HeinOnline*, Accessed: Aug. 27, 2025. [Online]. Available: https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/anlsadced34&section=6

[75] A. Gera, A. Dhull, A. Singh, … K. S. E. A. I., and undefined 2025, "Legal and regulatory issues related to AI in healthcare," *Elsevier*, Accessed: Aug. 27, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780443247880000108

[76] S. Rezvany, "AI-Driven Risk Stratification with Privacy Data Partitioning," 2025, Accessed: Aug. 27, 2025. [Online]. Available: https://www.diva-portal.org/smash/record.jsf?pid=diva2:1967802

[77] A. Saikia, A. Kalita, & P. M.-R. T., and undefined 2025, "EXPLORING THE IMPACT OF AI ON PRIVACY AND ETHICAL CONSIDERATIONS: ANALYSING THE LEGAL AND REGULATORY FRAMEWORKS," *cyberleninka.ru*, Accessed: Aug. 27, 2025. [Online]. Available: https://cyberleninka.ru/article/n/exploring-the-impact-of-ai-on-privacy-and-ethical-considerations-analysing-the-legal-and-regulatory-frameworks

[78] H. E. Khodke, C. kumar, A. Podder, C. Jegadheesan, D. Parmar, and N. Yashan, "Artificial Neural Networks In Early Diagnosis Of Neurological Disorders: A Review Of Models, Biomarkers, And Clinical Integration," *theaspd.com*, vol. 11, no. 11, p. 2025, Accessed: Aug. 27, 2025. [Online]. Available: http://theaspd.com/index.php/ijes/article/view/1470

[79] Z. Ashraf, N. M.-I. of H. R. and A. in, and undefined 2025, "AI standards and regulations," *igi-global.com*, Accessed: Aug. 27, 2025. [Online]. Available: https://www.igi-global.com/chapter/ai-standards-and-regulations/365873

[80] Ł. Szoszkiewicz and R. Yuste, "Mental privacy: navigating risks, rights and regulation: Advances in neuroscience challenge contemporary legal frameworks to protect mental privacy," *embopress.org*, vol. 26, no. 14, pp. 3469–3473, Jul. 2025, doi: 10.1038/S44319-025-00505-6.


ISSN: 3088-2370 (Online) | © Geoinformatics International