ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 2 | April 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

Emotion Detection from Face Recognition Using CNN

Atharv Indurkar[1], Uddhav Gole[2], Harshwardhan Baghel[3], Chetan Chacharkar[4], Sahil Bhuyar[5], Prof. Shweta Palaskar[6]

[1,2,3,4,5]*Undergraduate Student, Sipna College of Engineering and Technology, Amravati, (MS), India*

[6]*Assistant Professor, Sipna College of Engineering and Technology, Amravati, (MS), India*

Abstract: *Emotion detection through facial recognition represents a critical aspect of human-computer interaction, offering profound implications for areas ranging from security to customer service and healthcare. This paper presents a detailed exploration of the application of Convolutional Neural Networks (CNNs) in identifying and categorizing human emotions from static images and video inputs. We first outline the preprocessing steps necessary for preparing facial data for analysis, including face detection, alignment, and normalization. We then discuss the architecture of CNNs tailored for emotion recognition, highlighting the importance of convolutional layers, activation functions, and pooling layers in feature extraction from facial expressions. Our approach utilizes several advanced CNN models, which have been trained and validated on standard datasets like FER-2013. The performance of these models is evaluated based on their accuracy, precision, recall, and F1 score in recognizing basic emotional states such as happiness, sadness, anger, surprise, disgust, fear, and neutrality. The results demonstrate the effectiveness of CNNs in capturing subtle facial cues associated with different emotions, with significant accuracy improvements over traditional machine learning methods. The paper concludes with a discussion of the challenges in real-world applications and future directions for research in emotion detection using deep learning techniques.*

Keywords: *Emotion Detection, Facial Recognition, Convolutional Neural Networks, CNN, Machine Learning, Feature Extraction, Human-Computer Interaction, Deep Learning, Image Processing, etc.*

I. INTRODUCTION

Human-computer interaction technology refers to a kind of technology that takes computer equipment as the medium, and human-computer commerce technology refers to a kind of technology that takes computer outfit as the medium, to realize the commerce between humans and computers. A face recognition system (FRS) is a medium that allows cameras to automatically identify people. Because of the significance of correct and effective FRS, it drives the activeness of biometric exploration in the race to the digital world. In recent times, with the rapid-fire development of pattern recognition and artificial intelligence, more and more has been conducted in the field of mortal computer commerce technology. Facial Emotion Recognition (FER) is a flourishing study content in which numerous improvements are being made in diligence, like automatic restatement systems and machine-to-mortal contact. In discrepancy, the paper focuses on checking and reviewing.

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
Volume 3 | Issue 2 | April 2024
Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647
www.ijirid.in

Color facial birth features, emotional databases, classifier algorithms, and so on. The classical FER consists of two main ways feature extraction and emotion recognition. In addition, image pre-processing, includes face discovery, cropping, and resizing. Face discovery crops the facial region after removing the background and non-face areas. Eventually, the recaptured characteristics are used to classify feelings, which is generally done with the help of neural networks (NN) and other machine learning approaches.

The challenge of facial emotion recognition is to automatically fete facial emotion countries with high delicacy. thus, it is challenging to find the similarity of the same emotional state between different people since they may express the same emotional state in colorful ways. As an illustration, the expression may vary in different situations similar to the existent's mood, skin color, age, and the terrain girding.

Generally, FER is separated into three major stages as shown in Figure 1: (i) input image, (ii) pre-processing, (iii) Face Detection, (iv) Feature Extraction, and (v) Emotion Classification, (vi) output.

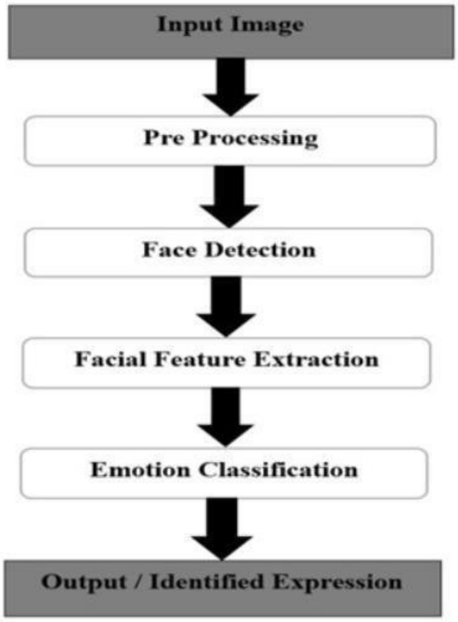

Figure 1: Classification Stages

1. **Input Images**

   This is the starting point where images of faces are provided as input to the system. These images can come from various sources such as cameras, video streams, or stored photographs.

2. **Preprocessing**

   In this stage, the input images are processed to facilitate better recognition and analysis. Common preprocessing steps include resizing and adjusting the image to a standard size.

 DOI: 10.5281/zenodo.10985172

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 2 | April 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

Grayscale Conversion: Transforming the image into grayscale to reduce computational complexity, as colors information might not be necessary. Normalization: Scaling pixel values to a certain range (e.g., 0 to 1) to stabilize learning.

3. **Face Detection**

   This step involves identifying and locating faces within the images. It may also involve Bounding Boxes: Drawing rectangles around detected faces to isolate them from the rest of the image. Region of Interest (ROI) Extraction: Cropping the detected faces for further processing.

4. **Face Feature Extraction**

   Description: In this crucial phase, the system extracts distinctive features from the detected faces that are necessary for recognizing emotions. Using CNNs, this process typically involves Convolutional Layers and applying filters to capture textures, patterns, and various facial features. Pooling Layers: Reducing the dimensions of the feature maps to condense the information. Activation Functions: Introducing non-linearity to help CNN learn complex patterns.

5. **Emotion Detection**

   This is the classification phase where the extracted features are used to predict the emotion of the detected face. It involves: Fully Connected Layers: These layers follow the feature extraction layers and are used to interpret the features and make a prediction. SoftMax Activation: In the final layer, a SoftMax activation function is often used to calculate the probability distribution over predefined emotion categories (e.g., happy, sad, angry, surprised).

6. **Output**

   Description: The final step is where the system outputs the detected emotion for each face. The output can be in various forms. Labels: Textual indication of the detected emotion (e.g., "Happy").

   Probabilities: A set of confidence scores corresponding to each emotion category, indicating the likelihood of each emotion.

II. LITERATURE REVIEW

Facial expression is the common signal for all humans to convey a mood. There are numerous attempts to make an automatic facial expression analysis tool as it has operations in numerous fields similar as robotics, drug, driving help systems, and lie sensors. [1]

Since the twentieth century, Ekman et al. defined seven introductory feelings, irrespective of culture in which a mortal grows with the seven expressions (anger, fear, happiness, sadness, contempt,

 DOI: 10.5281/zenodo.10985172

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 2 | April 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

disgust, and surprise). Discusses an expansive study on face emotion identification, including the dataset's features and the facial emotion recognition study classifier. [3]

Visual features of images are examined and some of the classifier ways are bandied in (6,7) which is helpful in the further examination of the styles of emotion recognition. This paper examined the vaticination of future responses from images grounded on the recognition of feelings, using different classes of classifiers. Some of the classification algorithms like Support vector machines, and Neural Networks similar to Convolution Neural networks. [5]

There are numerous issues like inordinate makeup disguise and expression which are answered using convolutional networks. The development of computer vision and machine literacy has made emotion recognition much more accurate and accessible to the public. As a result, facial expression discovery as a sub-field of image processing is rapidly expanding. Some of the possible operations are human-computer commerce, psychiatric compliances, drunk motorist recognition, and the most important is a taradiddle sensor. [6] [7]

III. PROPOSED METHODOLOGY

1. Data Collection

Dataset Selection: Choose a facial emotion dataset (e.g., FER-2013, CK+, AffectNet) that includes images labeled with emotions (anger, happiness, sadness, etc.). Ensure the dataset is diverse in terms of age, ethnicity, and lighting conditions for robust model performance.

Data Augmentation: To increase the diversity of the dataset and prevent overfitting, apply data augmentation techniques like flipping, rotation, scaling, and cropping on the images.

2. Preprocessing

Face Detection: Use a face detection algorithm (e.g., Haar Cascades, MTCNN) to locate and extract faces from the images.

Normalization: Scale pixel values to a range (e.g., 0-1) for consistent model input.

Resizing: Resize images to a uniform size (e.g., 48x48 pixels) to ensure they fit the input layer of the CNN.

Grayscale Conversion: Optionally, convert images to grayscale to reduce computational complexity, as emotion detection can often be performed without color information.

3. Model Design

Architecture Selection: Design a CNN architecture or select a pre-existing one (e.g., VGG16, ResNet) and modify it if necessary. A typical CNN for emotion detection might include several convolutional layers, activation functions (ReLU), pooling layers, dropout layers (for regularization), and fully connected layers.

Output Layer: The final layer should be a SoftMax layer that outputs the probability distribution over all possible emotion classes.

DOI: 10.5281/zenodo.10985172

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 2 | April 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

4. **Training**

Split the Data: Divide the dataset into training, validation, and test sets (e.g., 80% training, 10% validation, 10% test).

Choose Loss Function and Optimizer: Use categorical cross-entropy as the loss function since this is a multi-class classification problem.

Hyperparameter Tuning: Adjust parameters such as learning rate, batch size, and number of epochs based on performance on the validation set.

Regularization: Apply techniques like dropout and data augmentation to prevent overfitting.

5. **Evaluation**

Metrics: Use accuracy, precision, recall, and F1 score to evaluate the model's performance on the test set. Confusion matrices can also provide insight into specific areas of strength and weakness in emotion classification.

Fine-tuning: Based on the evaluation, you may need to go back and adjust the model architecture, hyperparameters, or preprocessing steps to improve performance.

6. **Deployment**

Integration: Integrate the trained model into an application or system that can perform real-time emotion detection from camera feeds or static images.

Optimization for Inference: Optimize the model for faster inference if necessary, using techniques such as model quantization or pruning.

7. **Continuous Improvement**

Feedback Loop: Collect feedback and additional data from users to continuously train and improve the model.

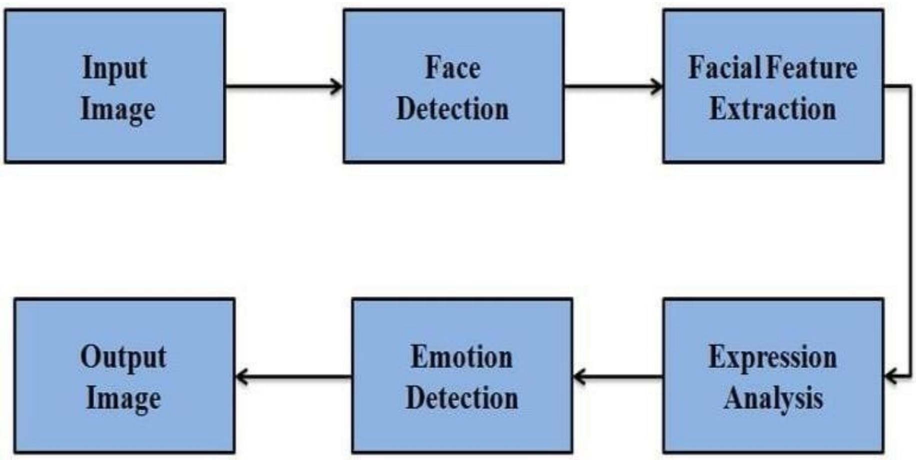

Figure 2: Emotion Detection Process

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
Volume 3 | Issue 2 | April 2024
Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647
www.ijirid.in

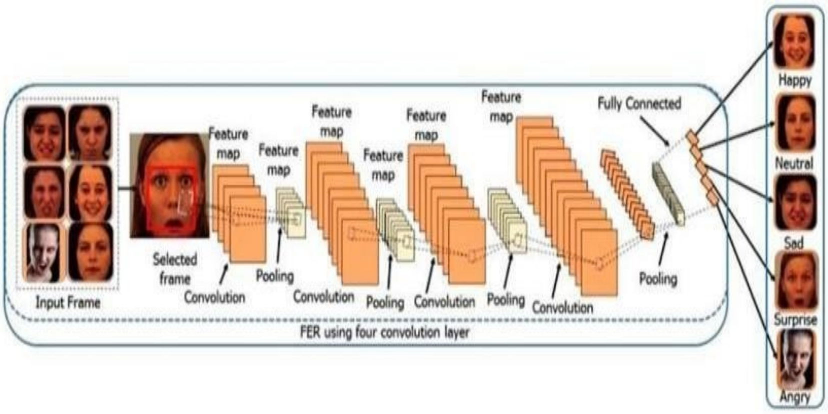

Figure 3: Image Classification

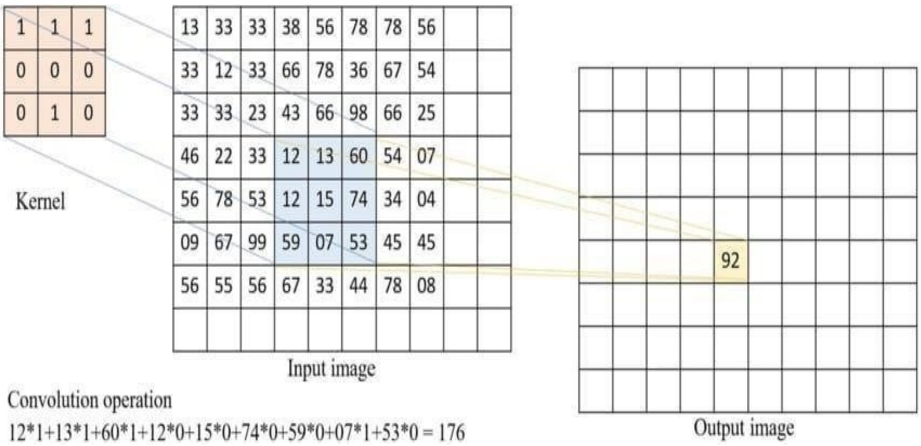

Figure 4: Convolution Filter

IV. DATASETS

In projects focused on detecting emotions through facial recognition using machine learning and Convolutional Neural Networks (CNNs), the importance of datasets is paramount. These datasets are typically comprised of extensive collections of facial photographs, each tagged with labels denoting the emotional expression being portrayed, such as joy, sorrow, rage, astonishment, revulsion, terror, and calmness. Through processing these images across its layered architecture, the CNN learns to identify patterns linked to these emotional states. This process includes the automated identification of facial characteristics such as the contours and placement of the eyes, lips, eyebrows, and other significant facial points, and understanding how these characteristics alter with varying emotions.

A dataset that is thoroughly labeled is vital for the effective training of the CNN, enabling the algorithm to grasp the diverse manners in which emotions are exhibited across various individuals, under different lighting scenarios, and against assorted backdrops. Notable datasets employed in such endeavors encompass the Facial Expression Recognition 2013 (FER-2013). dataset, the Extended Cohn-Kanade (CK+) dataset, and the Real- world Affective Faces Database (RAF-DB), among others.

 DOI: 10.5281/zenodo.10985172

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 2 | April 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

V. RESULT

In evaluating the algorithm's efficiency, the initial trials utilized the FEB-2013 emotion dataset. This dataset comprised merely 7178 images featuring 412 models, which limited the accuracy to a peak of 55%. To address the issue of reduced performance, additional datasets were sourced online, and the author incorporated personal photographs showcasing various expressions. With the expansion of the dataset to include over 11,000 images, a noticeable improvement in accuracy was observed. The dataset was divided, allocating 70% for training and the remaining 30% for testing purposes.

The configuration for both the background subtraction CNN (the initial CNN phase) and the facial features extraction CNN (the secondary phase) remained consistent in terms of the number of layers and filters. The experimental range for the number of layers was between one and eight, revealing optimal accuracy at four layers. This finding suggested an unexpected relationship where the number of layers was directly related to accuracy but inversely related to execution speed. Therefore, four layers were chosen based on achieving the highest accuracy.

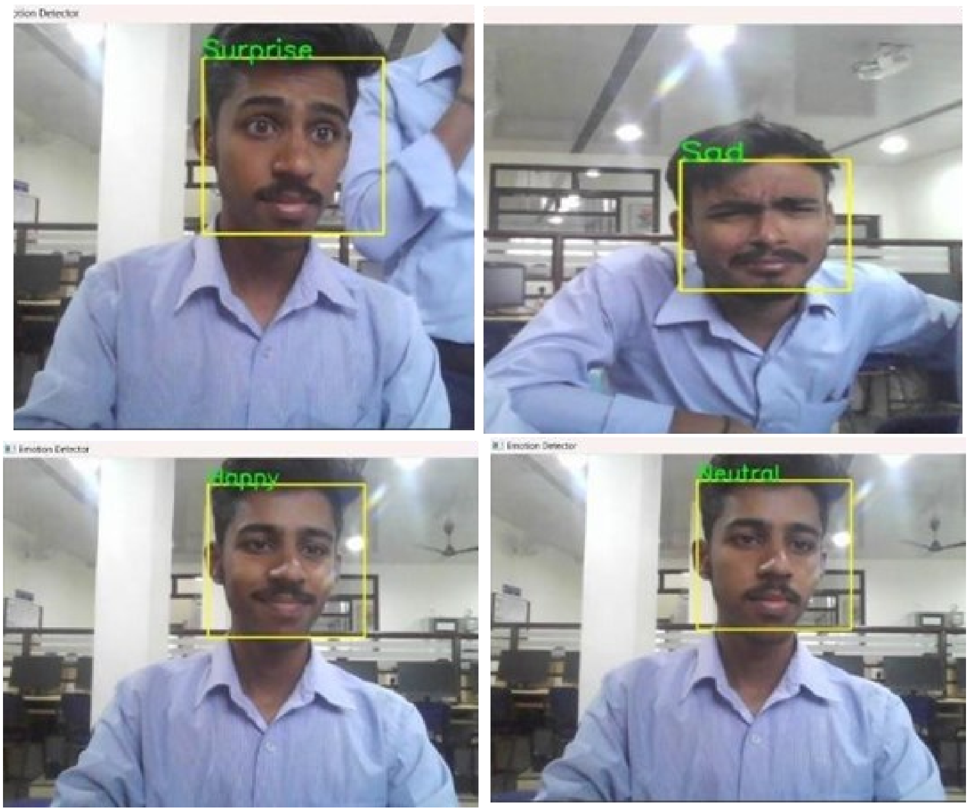

Figure 5: Emotion Detection

 DOI: 10.5281/zenodo.10985172

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 2 | April 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

Factor	FER-2013 dataset	Own dataset
Accuracy	62%	28%
Training Time	Longer due to Complexity	Shorter compared to Kaggle data
Inference Time	Fast with a trained model	Varies with implementation
Complexity	High with Deep Layers	Lower but depend on kernel
Dataset size	Large dataset size	Limited images
Diversity in images	High	Low
Data Requirement	Large Data Sets	Performs with less data

VI. CONCLUSION

In this study, we introduce a method for recognizing facial expressions using a CNN model that adeptly extracts features from faces. This innovative technique processes training images by directly utilizing the pixel values of the photos. The accuracy in identifying emotions saw significant improvement with the elimination of background distractions. Recognizing facial expressions plays a vital role in enhancing human communication, thus enriching interactions. Moreover, advancements in facial expression recognition research could soon enhance societal feedback mechanisms as well as human-robot interaction (HRI) dynamics.

The focus of emotion detection primarily lies on the facial geometry, such as the eyes, eyebrows, and mouth. This review considers various experiments conducted under controlled conditions, in real-time scenarios, and with unstructured, spontaneous images. Recent studies, especially those dealing with profile angles, show promise for a wide array of practical applications in the real world, like monitoring patients in healthcare settings or in security surveillance systems. Additionally, the scope of recognizing emotions could broaden to include deciphering feelings from vocal tones or body language, catering to the needs of new industrial challenges.

REFERENCES

[1] K. F. Azizan Illiana, "Facial Emotion Recognition: A Brief Review," in International Conference on Sustainable Engineering, Technology and Management 2018 (ICSETM-2018), 2020.

[2] R. Shyam, "Convolutional Neural Network and its Architectures.," Journal of Computer Technology & Applications, vol. 12, no. 2, pp. 6-14, 2021.

[3] R. Shyam, "Automatic Face Recognition in Digital World," Advances in Computer Science and Information Technology (ACSIT), vol. 2, no. 1, pp. 64-70, 2015.

[4] S. R. N. S. M. A. H. Akhand, "Facial Emotion Recognition Using Transfer Learning in the Deep CNN," MDPI, vol. 10, no. 9, 2021.

[5] N. Mehendale, "Facial emotion recognition using convolutional neural networks (FERC)," SN Applied Sciences, vol. 2, no. 3, 2020.

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 2 | April 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

[6] R. Shyam, "Enhanced Object Detection with Deep Convolutional Neural Networks," International Journal of All Research Education and Scientific Methods (IJARESM), vol. 9, no. 7, pp. 27-36, 2021.

[7] R. Shyam, "Machine Learning and Its Dominant Paradigms," Journal of Advancements in Robotics, vol. 8, no. 2, pp. 1-10, 2021.

[8] Facial Emotion Recognition of Students using Convolutional Neural Network", by Imane Lasri, Anouar riad solh, Mourad EL belkacemi .11-oct 2022.

[9] "Deep learning | Nature." [Online]. Available: https://www.nature.com/ articles/nature14539.

[10] B. Ko, "A Brief Review of Facial Emotion Recognition Based on Visual Information," Sensors, vol. 18, no. 2, p. 401, 2018.

[11] R. C. Chivers, V. V Ramalingam, and A. Pandian, "Facial Emotion Recognition System – A Machine Learning Approach Facial Emotion Recognition System ± A Machine Learning," 2018.

[12] M. H. Siddiqi, M. Alruwaili, J. Bang, and S. Lee, "Real Time Human Facial Expression Recognition System using Smartphone," vol. 17, no. 10, pp. 223–230, 2017.

[13] A. Savva, V. Stylianou, K. Kyriacou, and F. Domenach, "Recognizing Student Facial Expressions: A Web Application," IEEE Glob. Eng. Educ. Conf. EDUCON, vol. 2018–April, pp. 1459–1462, 2018.

[14] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie, "Facial expression recognition via learning deep sparse autoencoders," Neurocomputing, vol. 273, pp. 643–649, 2018

