JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Facial Expression Recognition using Deep Learning Algorithm

Poonam Gujre[1] Sudha Sharma[2] Trapti Sharma[3] Durgesh Mishra[4]
[1]M.Tech. Scholar , [2]Assistant Professor , [3]HoD, [4]Director
Department of Computer Science Engineering, Sri Aurobindo Institute of Technology, Indore

Abstract: Face expression recognition is essential in a wide variety of applications. The traditional feature extraction technique is used to complete a task of this kind, which needs a large amount of handling. Beforehand, several deep neural networks were utilized aimed at this purpose; however, several newer and probably superior techniques may now be used. Consequently, the purpose of this research is to provide a strategy for detecting faces that integrates the convolutional neural network (CNN) and image edge identification. The suggested method is divided into two stages: the first includes standardizing the facial expression in the image, and the second leverages convolution for image edge extraction. The procedure known as maximum pooling is then utilized to reduce the dimensionality. The expression is recognized once the Softmax classifier has completed categorizing it. The Fer-2013 dataset was used to complete the facial expression recognition study. The proposed technique can achieve an expression recognition rate of up to 99.68 percent aimed at the dataset being used. The presented approach takes fewer iterations to achieve recognition, and the system runs 5 to 15 times quicker than the SDSRN, AlexNet, and VGGnet algorithms.

Keywords*:* CNN ,AlexNet,VGGnet, Softmax classifier.

I. INTRODUCTION

Computer-human interaction is the product of technological developments in this arena. One of the most crucial aspects is having the computer behave like a person. Recent advances in this sector, such as pattern recognition and artificial intelligence, have increased the versatility of interaction [1]. Facial expression recognition is the most fundamental strategy for making any machine as intelligent as a human. The utility of this method may be observed in various security systems, gaming, distant education, and so on. This approach of facial expression identification takes expression characteristics from a predetermined input set and then teaches a computer to recognize facial expressions [2]. The photograph is subsequently sent to the computer as input for analyzing the facial expression. Neutral, delighted, mournful, surprise, and rejection facial expressions This expression recognition method is often employed in emotional quantification [3]. According to the findings, facial expression recognition technology is critical to explore and develop for several applications.
The classic CNN approach has limitations, such as the extensive training period. Therefore, there is a need to enhance the detection capabilities of ordinary CNN when the background is complex. We also demonstrated how to identify the facial emotions in any shot. The methodology given here combines the CNN method with picture edge detection. In this part, we will extract the boundaries of all layers included in the input photos. The max pooling approach was also utilized to lower the dimensionality of the inadvertently acquired data. In this study, the Fer-2013 face expression dataset was utilized.
People's facial expressions are critical components of interpersonal communication. Most of the time, in communication, words and actions go hand in hand [4]. Even in the absence of verbal language, facial expressions may transmit meaning. The subtle indicators that communicate the bigger message are a person's facial expressions. Humans and animals may communicate with one another without using words. They may do this using paralinguistic cues, gestures, facial expressions, body language, and eye contact.

Because it allows the transmission of information between persons, eye contact is a crucial component of communication. Maintaining eye contact with others allows you to connect with them and governs what and how they speak. There are joyful grins, sad grins, furious grins, disgusted grins, shocked grins, and fearful grins. When someone smiles, it shows that they are happy and causes the corners of their eyes to rise. The sensation of being loose refers to the look of grief, which is distinguished by its arched, crooked brows and frown [5]. Unpleasant or inconvenient things tend to send people off in a rage. When someone is unhappy, their brows are pulled together, and their eyelids appear thin and pinched [6].

The most important part of the ER process is the extraction and assembly of features. There are two forms of feature extraction: geometry-based and appearance-based [7]. The most prevalent kind is geometry-based feature extraction. One of the most important components of classifying facial emotions such as a smile, grief, rage, disdain, surprise, and fear into groups is classification. These are some examples of expressions [8].

The face will emit three unique sorts of signals most often: static, slow, and rapid [9]. Skin colour, including pigmentation and greasy deposits, face shape, bone structure, cartilage, and the form, location, and size of elements such as the eyes, nose, and mouth are static indicators [10]. Slow signals include wrinkles that never disappear and changes in the look of the face as time passes, such as alterations in muscle tone and skin texture [11].

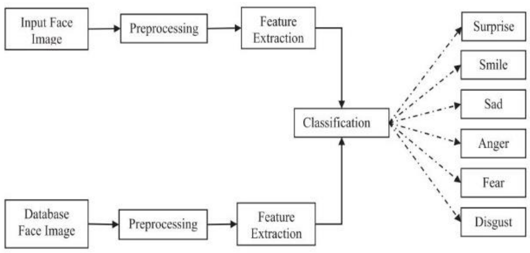

Figure 1. The architecture of facial expression recognition system.

Deep learning and, more specifically, convolutional neural networks (CNNs) have made it possible in recent years to extract and learn a significant amount of information that can be used in facial expression recognition systems. This information can be used to recognize a variety of facial expressions. Because of this, the quality of these systems has gotten a lot better. But it's important to remember that most clues come from just a few places, like the lips and the eyes, when it comes to facial expressions. Other parts of the face, like the ears and the hair, don't have much to do with the output. This means that the machine learning framework should focus only on the essential parts of the face and pay less attention to the other parts [12].

Human-machine contact has a lot of potential that can only be used to its fullest with clear and concise communication. Communication can happen both verbally and nonverbally. A critical example of the latter is figuring out how someone is feeling. Different ways, such as a person's speech, facial expression, body language, etc., can be used to figure out what they are thinking and feeling. [13] Some psychologists think about 55 percent of all communication is done through facial expressions. Even without words, a person's facial expressions can tell much about their feelings. Face expressions show how a person is feeling and what they are thinking. This is because facial muscles can move into different positions. Paying close attention to someone's facial expressions can help you figure out how they feel, what they plan to do, and other things. Emotion recognition through facial analysis is becoming more popular because it can be used in many places, like robotics, where the connection between people and machines is vital. Analysis of a person's facial expressions to figure out their feelings is also used for biometric security checks and surveillance, among other things. [14]

At this early stage in developing partnerships between humans and machines, communication is key to improving these interactions. It shouldn't be surprising that people prefer to use natural communication methods with technology, like the languages they have built up over time. Emotions, which create a logical way of passing on information, are another way that people naturally communicate with each other without using words. It would be helpful if robots could understand how people feel, making communication a step further. Emotion recognition can be done in several ways, but the most important ones are based on speech and facial expressions. Deep learning and more traditional machine learning are examples of how facial expressions and sounds can be used to figure out how someone is feeling [15].

II. BASIC TERMINOLOGIES

1. Face Detection: Determining whether or not a photograph contains a face is known as face detection. To do this, we must broadly understand what a face looks like. Seeing that people's faces aren't all that dissimilar is encouraging. The overall framework of a face is comprised of a person's mouth, nose, eyes, forehead, and chin. All of us share these facial features. It is a method for classifying individuals into two distinct categories: those with faces and those without. For instance, one may identify faces using a process known as object class detection. The process of determining the position of each item in an image, its size and whether or not it belongs to a particular class is referred to as object-class detection. It is possible to infer that from this [16]:

Figure 2. Detection of a face [16]

2. Face Identification: This part of the system compares the person being looked for to every other person in the database to get a list of possible matches to determine how likely they are [17].
3. Face Verification: In this step, the computer compares the person being checked with the person they say they are and decides if they are the right person or not.
4. Facial Expressions One way to describe a facial expression is as one or more moves or positions of the muscles just under the skin of the face. When other people see these movements, they can figure out how the person is feeling. It is a way to talk without using words. Communication plays a role in how people get along with each other. The ones that happen most often are:

Figure 3. Expressions of the face [17]

III. LITERATURE REVIEW

The last step is to discover facial landmarks unique to each human face using a facial landmark detection library. The human face is then divided into upper and lower halves, allowing the required traits to be extracted from any portion of the face. The given model takes into account both geometric and texture-based components. A vector of normalized features is obtained after feature extraction. These feature vectors are utilized for training a three-layer MLP, resulting in 96 percent test set accuracy [18].

The face identification accuracy of the proposed model is 92.46 percent with p less than 0.05, which is more than the 86 percent accuracy of the CNN Classifiers. The application of the Unsupervised Machine Learning algorithm to the proposed model was compared to that of the CNN Classifiers method. It was determined that the suggested model was more accurate than the Machine learning approach [19].

Our objective is to demonstrate the differences and similarities between the methodologies above and the emerging hybrid and ensemble learning trend in FER systems. We aim to establish a broad structure for each kind of learning and look at the many technologies that may be utilized as components. We will conduct more extensive and exhaustive competitive performances and experimental comparisons of research on widely used datasets from 2014 to 2020. Our next step will be to broaden our survey's scope to cover our present application scenarios in Vietnam [20].

The ghost-module architecture utilizes a variety of fundamental linear transformations to minimize the number of parameters and give more feature maps. Due to its Ghost module-based design, GCNN can successfully extract and classify facial expression characteristics. Experiments have shown that this technique outperforms all other advanced methods when applied to three FER databases (RAF-DB, FER2013, and FERPlus databases) [21].

To the best of our knowledge, this is the first study to investigate the application of GAN to analyze manifold-valued representations to address the challenge of developing dynamic facial expressions. We conduct a quantitative and qualitative study of our proposed technique using two publicly accessible datasets, Oulu-CASIA and MUG Facial Expression. Our experiments show that our technology effectively generates realistic films with continuous action, a realistic aesthetic, and identity preservation. Furthermore, we show that our technique successfully generates dynamic facial expressions, transfers dynamic facial expressions, and enriches training data for improved emotion recognition models [22].

In this research project, we examined various techniques for determining the effectiveness of emotions up to and including correct observations. This study examines past studies on emotion recognition published between 2007 and 2021. Using an electroencephalogram (EEG) signal may achieve 95.20 percent accuracy, and combining statistical features and a neural network with EEG data can achieve 95 percent accuracy, according to the results of these studies. Using both the EEG data and facial expressions, the average accuracy ranges from 63 to 73 percent [23].

A simple way of automatically constructing a system's triplets allows users to mark correspondences readily. We constructed a neural network that transforms human expressions into avatar expressions. Extensive experimental findings and user evaluations suggest that nonprofessional users may produce high-quality facial expression retargeting results with our technology [24]. Our strategy saved time and effort, proving this point.

The model is trained using a convolutional neural network. Real-time testing using the Raspberry Pi 3B+ and Pi-Camera. PyQt5 creates the system's graphical user interface (GUI). The proposed approach has a 99.88% accuracy rate in testing [25].

At times, it may be difficult to recognize them due to the lack of a single model or framework capable of identifying the many types of sensations. Identifying emotional facial expressions might be difficult as well. Facial emotions are a kind of nonverbal communication that seems to represent a person's inner feelings. Face expressions are quite crucial in this kind of engagement. Machine learning methods and Deep Learning and Neural Network algorithms are employed in the emotion identification process. This article shows how Convolutional Neural Networks may be used to detect emotions (CNNs). Anger, disgust, pleasure, fear, sorrow, tranquilly, and surprise may all be detected using the presented approach [26].
In addition, we propose utilizing an exogenous dispelling loss to remove any external information from the endogenous representation. In a throwable strategy, the exogenous data is therefore utilized twice. It is first utilized as a conditioning variable for the objective task before being employed as an invariance generator for the endogenous representation. Both of these applications may be discarded. This approach is known as "throwable information networks" (THIN). We verify THIN experimentally in various contexts where exogenous information may be uncovered. Options include sensing fingers in extreme rotations and sizes and shapes. Using identity as the exogenous variable, we apply it to FER. We show [27] that THIN outperforms the most advanced methods on challenging datasets.

I am combining EDA and facial expression data improved this study's identification of clinically significant vs nonsignificant pain by 90.9%, yielding above-chance sensitivity and specificity. Multimodal pain treatment uses several pain response components [28].

IV. RECOMMENDED METHOD

4.1 The proposed strategy may be executed by pre-training the model and running the required trials. Here are the steps we need to do to finish our pre-processing:

1. Find the face in the photo and clip it out. This is done using the Haar classifier.
2. The image of the face has been scaled to a certain proportion.
3. Picture histogram equalization is performed to reduce image excess data. In this phase, the grey level is equalized.
4. Finally, the image's edge is removed. This job necessitates the usage of the Kirsch edge operator.

CNN-based facial expression recognition:

A feed-forward neural network is the CNN employed here for recognition. This might result in the feature extraction and network settings being optimized. The convolution, pooling, and connective layers are the three layers in question.

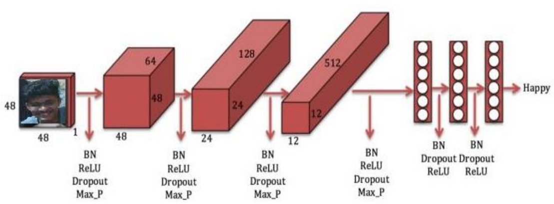

Figure 2. The deep network's design consists of four convolutional layers and two fully linked layers.

C1, C2, and C3 employ 32-bit, 64-bit, and 128-bit Conv nuclei. The Conv nuclei, in this case, are of the 5*5 type, while the S1 and S2 pooling layer is of the 2*2 type. The Softmax layer employed here has a total of 7 neurons. The connecting layer consists of 300 neurons linked to the S2 layer. This means that all of the levels are interconnected. We analyzed expressions that fell into the following categories: fear, rage, joy, disgusted, neutral, sad, and astonishment.

Convolution Layer: After applying the convolution filter to the picture, the value C1 is produced. Using feature mapping, we can get S2. Convolution is performed on S2 at this time, yielding C3 as a result. Using the same strategy as S4 was achieved. The following equations are used for computing:

$$y_j^l = \theta\left(\sum_{i=1}^{N_j^{l-1}} w_{i,j} \otimes x_i^{l=1} + b_j^l\right), j = 1,2, \dots\dots, M \quad (1)$$

l-1 represents the current and preceding layers, Yj represents the jth feature graph, Bj represents the jth feature graph bias, and M represents the total number of feature mappings. N represents the characteristic associated with the current layer.

Pooling Layer: The dimensions of a 2*2 pooling layer may be reduced by 50%. The pooling equation is provided beneath:

$$y_j^l = \theta(\beta_j^l down\left(y_j^{l-1}\right) + b_j^l) \quad (2)$$

There was no change in the values of j, l, or l-1 from the preceding calculation. Y represents down sampling for each layer and feature map, and B represents bias.

We require a 1D array, not a 2D one since the S2 layer produces a 2D array in this case. Input will then be in the form of 128 1D arrays. In this case, the feature vector spans the whole connection and has a length of 51200. As a consequence of the activation of each neuron:

$$h_{w,b}(x) = \theta(w^T x + b) \quad (3)$$

A neuron's output value is h, and its eigenvector is x. The letter b is used to represent bias. Theta serves as a measure of activation in neuroimaging.

In this case, the classifier is utilized as the last layer, called softmax. This has a variety of outcomes. Each neuron's output is between 0 and 1 for the input. The probability of belonging to the class varies between 0 and 1. As a consequence, the output with the highest likelihood is selected for classification.

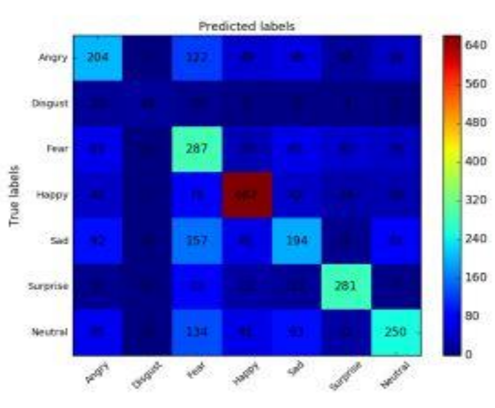

Figure 3. The shallow model's confusion matrix

V. OUTCOMES VALUATION

In this instance, the Fer-2013 database is utilized to interpret facial expressions. Here are particulars regarding the database.
• 35685 pictures • 48*48-pixel grayscale photos • These emotions include happiness, neutral, sorrow, anger, surprise, disgust, and fear.

This is the equipment used for the test:

• 64-bit operating system • Intel(R) Core I7 processor • 3.2 GHz primary CPU speed
• 8 Giga Byte RAM

The Sparse-Deep Simultaneous Recurrent Network algorithm (SDSRN), the AlexNet algorithm, and the VGGnet algorithm are contrasted with the newly suggested approach. During our experiment, we focused on three aspects: the amount of time spent in training, the amount of time spent in testing, and the recognition rate.

Table1. The obtained values for Experiment

METHOD	TRAINING TIME (hours)	TEST TIME (s)	RECOGNITION RATE (%)
Proposed CNN	2	1.4	99.68
CNN [19]	2.2	3.7	94.13
KNN [4]	4.3	8.5	89.37
AlexNet[29]	2.9	3.7	93.42
SDSRN [21]	3.2	3.9	90.67
VGGNet [30]	3.3	4.2	97.58

In the table above, you can see the evaluation's results. The findings show that the technique we're utilizing for expression recognition is working. The rate of recognition has risen dramatically in comparison to previous years. When compared to other algorithms like DSRN, the time needed for training and testing is significantly reduced. The overall time spent training has been reduced by 113 seconds. Additional information may be found in the graph that follows.

Figure 4. Facial expression training time

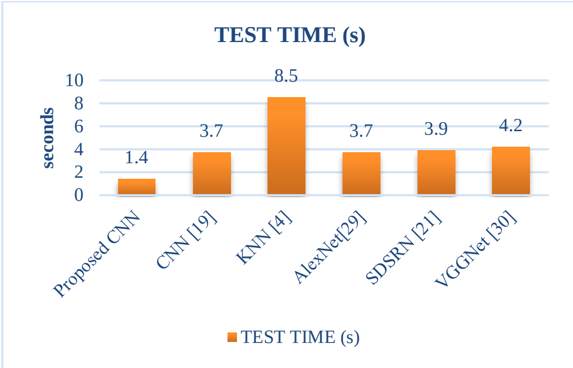

Figure 5. Facial expression testing time

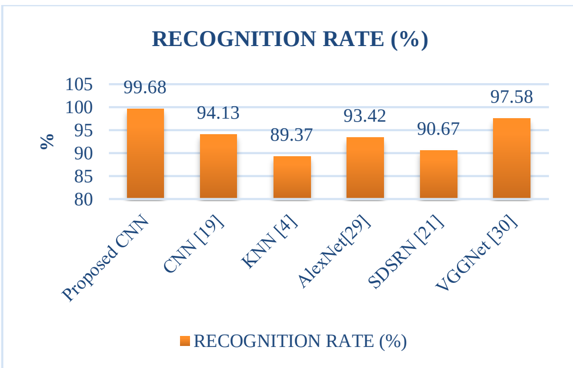

Figure 6. Facial expression rate

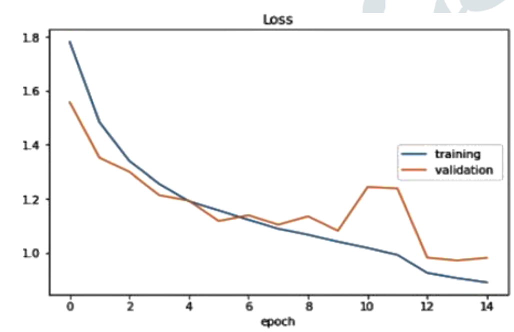

Figure 7. Graphical view of training and validation loss per epoch.

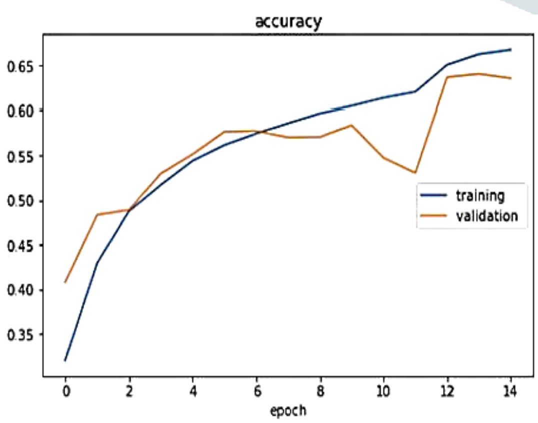

Figure 8. Graphical view of training and validation accuracy per epoch

It is much simpler to understand how the findings were obtained by looking at a graph. The time spent in training improves by 1.3 hours, while the time spent on the actual exam improves by 2.8 seconds. Additionally, there was a 10.03 percentage point increase in the recognition rate. The results of the test demonstrated that facial recognition is effective and reliable.

VI. CONCLUSION

Facial expression detection is gradually becoming one of the most important features that can be found in a range of mobile applications. Many algorithmic techniques and methodological frameworks have previously been worked on. However, there is still room for development, and in this piece, we will provide a method that uses CNN to recognize the expressions that are included inside an image. In addition to that, we utilized Softmax Classifier to organize the data appropriately. The Fer-2013 data set is being used here for the sake of the specific case. To get to these conclusions, we looked at three distinct times: the amount of time required for training, the amount of time required for testing, and the amount of time required for recognition. According to the experiment's findings, 99.68 percent of the people could accurately identify the persons that participated in the test. The approach that was explained may thus be used to recognize a variety of facial expressions.

Reference:

[1] X. Zhang, F. Zhang and C. Xu, "Joint Expression Synthesis and Representation Learning for Facial Expression Recognition," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 3, pp. 1681-1695, March 2022, doi: 10.1109/TCSVT.2021.3056098.

[2] Z. Hu and C. Yan, "Lightweight Multi-Scale Network with Attention for Facial Expression Recognition," 2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), 2021, pp. 695-698, doi: 10.1109/AEMCSE51986.2021.00143.

[3] Q. Xu and N. Zhao, "A Facial Expression Recognition Algorithm based on CNN and LBP Feature," 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2020, pp. 2304-2308, doi: 10.1109/ITNEC48623.2020.9084763.

[4] S. Roy Supta, M. Rifath Sahriar, M. G. Rashed, D. Das and R. Yasmin, "An Effective Facial Expression Recognition System," 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), 2020, pp. 66-69, doi: 10.1109/WIECON-ECE52138.2020.9397965.

[5] X. Xu, Z. Ruan and L. Yang, "Facial Expression Recognition Based on Graph Neural Network," 2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC), 2020, pp. 211-214, doi: 10.1109/ICIVC50857.2020.9177430.

[6] M. M. Kabir, T. A. Anik, M. S. Abid, M. F. Mridha and M. A. Hamid, "Facial Expression Recognition Using CNN-LSTM Approach," 2021 International Conference on Science & Contemporary Technologies (ICSCT), 2021, pp. 1-6, doi: 10.1109/ICSCT53883.2021.9642571.

[7] A. Kreinis, T. Damri, T. Leon, M. Litvak and I. Rabaev, "Telemoji: A video chat with automated recognition of facial expressions," 2021 International Conference on Visual Communications and Image Processing (VCIP), 2021, pp. 1-1, doi: 10.1109/VCIP53242.2021.9675330.

[8] L. Zhang, C. Xu and S. Li, "Facial Expression Recognition of Infants Based on Multi-Stream CNN Fusion Network," 2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP), 2020, pp. 37-41, doi: 10.1109/ICSIP49896.2020.9339265.

[9] H. Liu, J. Zeng and S. Shan, "Facial Expression Recognition for In-the-wild Videos," 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), 2020, pp. 615-618, doi: 10.1109/FG47880.2020.00102.

[10] J. L. Joseph and S. P. Mathew, "Facial Expression Recognition for the Blind Using Deep Learning," 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), 2021, pp. 1-5, doi: 10.1109/GUCON50781.2021.9574035.

[11] Y. Xia, W. Zheng, Y. Wang, H. Yu, J. Dong and F. -Y. Wang, "Local and Global Perception Generative Adversarial Network for Facial Expression Synthesis," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 3, pp. 1443-1452, March 2022, doi: 10.1109/TCSVT.2021.3074032.

[12] H. Zhang, W. Su, J. Yu and Z. Wang, "Identity–Expression Dual Branch Network for Facial Expression Recognition," in IEEE Transactions on Cognitive and Developmental Systems, vol. 13, no. 4, pp. 898-911, Dec. 2021, doi: 10.1109/TCDS.2020.3034807.

[13] D. Poux, B. Allaert, N. Ihaddadene, I. M. Bilasco, C. Djeraba and M. Bennamoun, "Dynamic Facial Expression Recognition Under Partial Occlusion With Optical Flow Reconstruction," in IEEE Transactions on Image Processing, vol. 31, pp. 446-457, 2022, doi: 10.1109/TIP.2021.3129120.

[14] C. Jia, C. L. Li and Z. Ying, "Facial expression recognition based on the ensemble learning of CNNs," 2020 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), 2020, pp. 1-5, doi: 10.1109/ICSPCC50002.2020.9259543.

[15] J. Zhang, X. Yao, Y. Zhang, N. Xu and X. Liu, "Spectral Norm Normalized Network for Facial Expression Generation," 2020 IEEE International Conference on Progress in Informatics and Computing (PIC), 2020, pp. 202-206, doi: 10.1109/PIC50277.2020.9350747.

[16] F. Zhang, T. Zhang, Q. Mao and C. Xu, "Geometry Guided Pose-Invariant Facial Expression Recognition," in IEEE Transactions on Image Processing, vol. 29, pp. 4445-4460, 2020, doi: 10.1109/TIP.2020.2972114.

[17] B. Zhang, D. Wei, Q. Zhang, W. Si, X. Li and Q. Zhu, "Classroom monitoring system based on facial expression recognition," 2021 20th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), 2021, pp. 108-111, doi: 10.1109/DCABES52998.2021.00034.

[18] M. A. Haghpanah, E. Saeedizade, M. T. Masouleh and A. Kalhor, "Real-Time Facial Expression Recognition using Facial Landmarks and Neural Networks," 2022 International Conference on Machine Vision and Image Processing (MVIP), 2022, pp. 1-7, doi: 10.1109/MVIP53647.2022.9738754.

[19] V. B and D. S. J. J. Thangaraj, "Innovative Facial Expression Identification for Criminal Identification using Unsupervised Machine Learning and Compare the Accuracy with CNN Classifiers," 2022 International Conference on Business Analytics for Technology and Security (ICBATS), 2022, pp. 1-4, doi: 10.1109/ICBATS54253.2022.9759043.

[20] V. -T. Dang, H. -Q. Do, V. -V. Vu and B. Yoon, "Facial Expression Recognition: A Survey and its Applications," 2022 24th International Conference on Advanced Communication Technology (ICACT), 2022, pp. 359-367, doi: 10.23919/ICACT53585.2022.9728930.

[21] B. Fang, G. Chen and J. He, "Ghost-based Convolutional Neural Network for Effective Facial Expression Recognition," 2022 International Conference on Machine Learning and Knowledge Engineering (MLKE), 2022, pp. 121-124, doi: 10.1109/MLKE55170.2022.00029.

[22] N. Otberdout, M. Daoudi, A. Kacem, L. Ballihi and S. Berretti, "Dynamic Facial Expression Generation on Hilbert Hypersphere With Conditional Wasserstein Generative Adversarial Nets," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 2, pp. 848-863, 1 Feb. 2022, doi: 10.1109/TPAMI.2020.3002500.

[23] N. Kumar and N. Gupta, "A Comprehensive Study on Machine Learning Approaches for Emotion Recognition," 2022 2nd International Conference on Artificial Intelligence and Signal Processing (AISP), 2022, pp. 1-5, doi: 10.1109/AISP53593.2022.9760652.

[24] J. Zhang, K. Chen and J. Zheng, "Facial Expression Retargeting From Human to Avatar Made Easy," in IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 2, pp. 1274-1287, 1 Feb. 2022, doi: 10.1109/TVCG.2020.3013876.

[25] P. Dhope and M. B. Neelagar, "Real-Time Emotion Recognition from Facial Expressions using Artificial Intelligence," 2022 2nd International Conference on Artificial Intelligence and Signal Processing (AISP), 2022, pp. 1-6, doi: 10.1109/AISP53593.2022.9760654.

[26] S. K. Singh, R. K. Thakur, S. Kumar and R. Anand, "Deep Learning and Machine Learning based Facial Emotion Detection using CNN," 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), 2022, pp. 530-535, doi: 10.23919/INDIACom54597.2022.9763165.

[27] E. Arnaud, A. Dapogny and K. Bailly, "THIN: THrowable Information Networks and Application for Facial Expression Recognition In The Wild," in IEEE Transactions on Affective Computing, doi: 10.1109/TAFFC.2022.3144439.

[28] B. T. Susam et al., "Automated Pain Assessment in Children Using Electrodermal Activity and Video Data Fusion via Machine Learning," in IEEE Transactions on Biomedical Engineering, vol. 69, no. 1, pp. 422-431, Jan. 2022, doi: 10.1109/TBME.2021.3096137.

[29] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton." Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.

[30] Simonyan, Karen, and Andrew Zisserman." Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).