Deep Learning and Machine Learning based Facial Emotion Detection using CNN

Shubham Kumar Singh
ECE Department
DSEU, G.B.Pant Okhla-1 Campus
(formerly GBPEC)
New Delhi, India
shubhamkrsingh112000@gmail.com

Revant Kumar Thakur
ECE Department
DSEU, G.B.Pant Okhla-1 Campus
(formerly GBPEC)
New Delhi, India
revantthankur8@gmail.com

Satish Kumar
ECE Department
DSEU, G.B.Pant Okhla-1 Campus
(formerly GBPEC)
New Delhi, India
skysatish1999@gmail.com

Rohit Anand
ECE Department
DSEU, G.B.Pant Okhla-1 Campus
(formerly GBPEC)
New Delhi, India
roh_anand@rediffmail.com

Abstract— Human emotions are the psychological conditions of sentiments that are encountered impulsively which imply the expressions on the face. By using Machine Learning and Deep Learning in this paper, the emotions will be detected and human behavior will be extracted. The various body language approaches like idiomatic expressions, eye stirring and body movement are significant while applying for the association between machines and people. Out of these methods, facial emotion is most commonly employed as it refers to the mental sentiments and frame of person's mind. The identification of the various emotions is sometimes a very difficult job as no specified prototype or framework is there to differentiate the various kinds of sentiments and also there are various complications while recognizing the facial emotion expression. Facial emotions are so much crucial in non-verbal type communication which appears to the internal feelings of an individual that reflects on the faces. Machine Learning techniques, Deep learning and Neural Network algorithms are used for emotion recognition. This work is going to propose an efficient technique using Convolutional Neural Networks (CNNs) to detect anger, disgust, happiness, fear, sadness, calmness and surprisingness.

Keywords— Convolutional Neural Network, Machine Learning, Deep Learning, Facial Emotion Recognition, OpenCV, Deep Learning.

I. Introduction

Individuals connect with each other predominantly through voice and yet additionally by tokens of body, to feature specific pieces of their correspondence and to show emotions. One of the huge ways individuals show their emotions is by means of facial expressions [1] that are truly significant methods of connection. However nothing is affirmed verbally, hence there's much to be recognized with regards to the messages we send and get through the work of non-verbal cooperation. Facial expressions express non-verbal signs, and that they assume a vital part in relations [2]. Programmed recognition of facial expressions is regularly a vital part of normal human-machine interfaces; it ought to try and be utilized in conducting science and clinical practice. Despite the fact that people distinguish facial expressions successfully right away, steady expression recognition by machines keeps on being a test [3]. There are a few upgrades since a couple of years as far as face identification, including extraction components is concerned. Consequently, the methods utilized for expression order and mechanized framework that achieves this assignment are troublesome.

In this work, we present a methodology that is dependent on the Convolutional Neural Networks (CNN) [4][5][6] for identification of the various facial sentiments. The contribution to our framework is to use CNN to anticipate the face-name which ought to be one among these names: angry, happy, fear, sad, disgust, and neutral.

The paper has been organized as follows. Second section discusses the review in concerned with the related work followed by methodology section. Subsequently, analysis is done under the results section. Finally, a brief conclusion is drawn.

II. Literature Review

The authors in [7] discussed a facial sentiments identifier based on the movement of brow and eyes dependent on electrooculography (EOG) signals. Various developments of the eyes, to be specific, up, down, right, left, flicker and grimace, are distinguished and replicated in a symbol, planning to dissect how they possibly impart for the facial sentiments. The authors in [8] proposed a face detection strategy that gets information and consolidates thermal and apparent moving pictures. Utilizing this strategy, an itemized examination of the connection is directed between the power of emotions in intentional-facial-expressions and the facial surface temperature. Further, facial picture features have been extracted in [9] by utilizing Linear Discriminant Analysis (LDA) and Facial Landmark Detection after grayscale handling and editing, and afterward thinking about

Authorized licensed use limited to: Indian Institute of Technology Indore. Downloaded on April 25,2025 at 10:19:43 UTC from IEEE Xplore. Restrictions apply.

the precision after feeling acknowledgment and characterization to figure out which feature extraction technique is more viable. Further in [10], the authors presented the advancement of the E-Bot framework, that empowers human-robot communication dependent on the feeling recognized from the face identification. A portable application is created to manage the machine and escort the talk. The Google Cloud Vision API and Pre-trained Facial Expression Algorithm are investigated and hence the expectation precision is calculated. The authors in [11] used a Haar-cascade classifier executed by OpenCV. Results exhibited that the Artificial Intelligence (AI) framework based on frame error rate perceived the feelings more precisely and that the human beings contemplated the AI facial sentiments identification frameworks reliable and gainful. In [12], the authors stressed on the consequence of ageing on the facial identification process, which uses Gabor filters to extract facial characteristics and SVM for classification. While the processing time for the Gabor filter is longer, Log Gabor filter is more effective as mentioned in [13].

The authors in [14] used the local binary pattern operator to extract features and categorize data from a picture. A texture descriptor or texture classifier is what the LBP is called. The authors in [15] proposed using LBP and Gabor filters to create an emotion detecting algorithm. Further in [16] and [17], the authors proposed some well-known strategies for recognizing the emotions. They assessed five basic emotions or moods that are commonly captured in the images of human faces: wrath, joy, normalcy, somnolence, and automatic machinery surprise. Unlabeled data can be classified using KNN, with labeled data receiving the closest majority vote. The various emotions are identified using KNN and Hidden Markov models in [18]. Authors in [19] employed KNN as a classifier for identifying the facial expressions. Overall, the accuracy rate has been 82.87 percent. The authors in [20] again used CNN for emotion recognition. They proved that the occurrences of Facial Action Units (FAU) can be used to process the detection. The Facial Action Unit is one of the sub-divisions of FACS that refers to the facial action coding system. This approach may detect the emotions depending on the facial expression motions. The usage of CNN reduced overfitting, resulting in a accuracy rate of 92.81 percent.

III. Methodology

The entire system is the training element of this work which considers the major difficulties encountered by machine learning. The algorithm must be trained using real-world data of human facial reactions. If the system is to detect an angry face, for example, the first system must be familiar with the angry face. The re-training technique is utilized to antecedent the system with these feeling categories. The data for re-training is gathered in the real-world. There are numerous other components to the system. Machine learning improves the accuracy with which emotions may be detected. It provides real-time feedback. The system doesn’t wait for a future result, and the image doesn’t need to be saved. Neoteric approaches can examine thousands of data in a short amount of time using modern computers, thus saving hundreds of hours. Furthermore, the cost of utilizing and installing such products is substantially lower. These data mining tools, if correctly adjusted, can produce better results than a human being. This research shows a broad and realistic system for the sentimental data mining using machine learning to discover emotion patterns. This paper develops a program based on model of deep learning [21][22] and emotion recognition using image processing and computer vision [23][24]. For this work, the proposed method employs CNN algorithm. This deep learning-based emotion recognition system consists of four phases, as follows:

(1) The public face database is trained using CNN.
(2) Seven probabilities are derived for each face frame..
(3) Single-frame probability is aggregated into fixed-length image descriptors [25][26] for each picture in the collection.
(4) All photos are classified using a support vector machine (SVM) [27][28] based on the learning of image descriptors from the competition training set.

A. Emotion Database

This is utilized in both media in the real-world and on the internet to acquire as much data as possible during the data collection processes. Emotions come in many forms in photographs. They filter the data that has been saved for later analysis. The data set comes from kaggle.com and is gathered from internet media. This data set was uploaded to this server one year ago. This site has the most reliable set of emotional data. The data are turned into grayscale images of faces using a (48x48) pixel resolution. It is divided in two parts: pixels and feelings. A numeral code between zero and six is included in the sensation section. Furthermore, each picture's pixel part comprises of a string that is included in statements. Furthermore, the image should only be a portrait of a person's face. As a result, the gathered images are shrunk and cropped to create a portrait of a face. In addition, there is a distinct picture.

B. Training Phase using Deep Learning

Convolutional neural networks are a fantastic method to apply the concept of deep learning to categorize the photos. The Python Keras module makes creating neural networks a simple one. The computers employ pixels to display the images. For majority of the cases, pixels in an image are related. For example, a collection of pixels might be used to depict the edge of a picture or some design. The concept of convolution uses this concept to assist them recognize the pictures. A convolution combines the results of multiplying a pixel matrix by a mask matrix, or 'kernel'. The convolution then slides on to the following pixel, repeating the process until all of the picture elements are covered. We'll be employing a sequential model type. The simplest method for building a model in Keras is sequential. To superimpose the different layers to the proposed model, we utilize the 'add ()' method. Conv2D layers makes up our initial two layers. Our 2-D matrices will be handled by these convolution layers as input images. Each layer has 64 nodes in the first layer and 32 nodes in the second layer. This number may be changed to some other number that totally depends on the dataset dimensions. In the proposed work, 64 and 32 nodes operate nicely together. The kernel size for the proposed convolution is filter matrix size. A mask with size of 3 will be used. RELU activation function is very important in CNNs. As noted before, each input's format image is 28,28,1, the number 1 denotes that the photographs are gray scaled. To

Authorized licensed use limited to: Indian Institute of Technology Indore. Downloaded on April 25,2025 at 10:19:43 UTC from IEEE Xplore. Restrictions apply.

link the convolutional and dense layers, 'Flatten' is utilized. After that, the model will provide a estimation based on the choice with the highest likelihood. After that, there will be a requirement to aggregate the model that is built using optimizer, loss, and metrics as three parameters. The optimizer tries to manage all things including the pace of learning. It will make use of the 'Adam' optimizer. The Adam optimizer adapts the learning rate throughout the entire training. The rate of learning indicates how soon the ideal weights for the model are computed. A slower learning rate refers to the precise weights (to some extent), but it will also take longer time to calculate the weights. For loss function, we'll employ 'categorical cross-entropy' which is the most widely used classification method. The model is performing better if the score is lower. To make things even easier to understand, while training the model, the 'accuracy' measure will be used to see the validation set's accuracy score. It will train using the 'fit ()' function training data (train X), target data (train y), validation data, and the number of epochs. It will validate the data using the test set supplied in its dataset. This has been broken into tests X and Y. The number of epochs determines that the proposed model cycles through the data how many times. As we run the additional epochs, the model will improve, but only to a degree. In our model, the number of epochs will be fixed to three. On that validation set, it has achieved 93 percent accuracy after three epochs.

C. Detection

In K-means clustering technique, the number of clusters is fixed to two. This is where the highest value in each row is located, and the average is calculated. Similarly, the average is calculated after determining the least value in each row. Pixels closer to the largest mean value are assembled in one cluster, while pixels closer to the lowest average value are clustered together in another, using these two places as a starting point. Depending upon the output of clustering, all the components in the image are computed. The bounding box function is used to segment the person's eyes first. Because the eyeballs are the first element found when traversing the pixel values column-wise, they are segmented first. The eye matrix is used to segment other facial features using a distance-based method. The image obtained after performing K-means clustering for the various expressions is exhibited. The major feature of this algorithm is that training takes a long time, although it takes a short time to detect. The Haar basis feature is used in this technique. The Haar characteristics are the ones that matter when it comes to detecting faces. There are many different types of characteristics, including Edge characteristics, Line Characteristics and Four Rectangle Characteristics. For example, if we want to identify the face of a person, we must first change the image to grayscale before moving on to image segmentation.

Assume that we need to identify the eyebrow. After that, we'll require edge features. If we wish to detect a nose, we'll need the black-white-black line features. Edge features are required if teeth are to be detected. The image moves on to the next feature after using these Haar features. Emotion recognition is based on the ratio of these identified features.

Using Fourier equation, we calculate the difference between dark (1 value) and white (0 value) shown in equation (1).

$$\Delta = \text{dark–white} = \frac{1}{n}\sum_{\text{dark}}^{n} I(x) - \frac{1}{n}\sum_{\text{white}}^{n} I(x) \qquad (1)$$

In an actual image, we obtain 0.74 when we compute the black region and average it, and 0.18 when we calculate the white region and average it. So, for an actual image, the difference between dark and white is 0.74-0.18=0.56.

Layers are commonly used to organize the neural networks. Layers are made up of linked nodes that have an activation function. The network connects with some hidden layers through the input layer, that delivers the patterns to the network which performs the actual processing via a system of weighted connections. The Viola-Jones method is used to recognize faces and facial components, as well as to extract facial features and classify them using CNN. There are three stages of the facial emotion recognition system. Keras is a python-based open-source neural network that may be used for preprocessing, modeling, assessing, and optimizing the data. It is for high-level APIs, which are handled by the backend. It is made for creating the models with optimizer and loss functions, in addition to training with the fit function. It is meant for convolution and low-level computation in the backend using Tensors or TensorFlow. The preprocessing, modeling, optimization, and testing are all done with Python libraries, displaying the emotions with the highest proportion. It employs a sequential model with several layers, including image pre-processing, convolution, pooling, flattening, dense layers, RELU, and activation. The proposed system's first phase is image preprocessing, which includes face identification, detection and extraction. This algorithm recognizes the facial region regardless of the raw input image's size, brightness, background, or spatial transformation. The detection of face can be done by aggregating the classifiers in the complete structure, which increases detection efficiency while decreasing the cumbersome calculations. The final classifier is a linear mixture of all weak classifiers that distinguish between positive and negative weighted error (each learner's weight is proportional to its accuracy). Face parts (both eyes and mouth) are recognized, cropped, extracted, and normalized to a size of (64×64) pixels before being removed from the normalized face picture. The retrieved facial portions are reduced to (32x64) pixels in size. The lower image scale reduces the amount of data that the network must learn, and makes training faster with less intensive memory. For huge datasets, convolution layers will be added to improve accuracy. The data are gathered from a CSV file and translated into photos, which are then used to categorize emotions and their expressions. There are 34,488 photographs in the training dataset and 1,250 in the testing dataset. The emotions are categorized as joyful, surprise, angry sad, neutral, disgust, and fear. Distinct facial features, such as elevated cheeks, creases Others include the area surrounding the nose, wide-open eyes etc. The big dataset is trained for higher accuracy, and the outcome is the supplied image's object class. Pooling is a notion that goes hand-in-hand with convolution in deep learning visual object detection. A convolution maps a region of an image to a feature map, according to the theory. A (5x5) array of pixels,

Authorized licensed use limited to: Indian Institute of Technology Indore. Downloaded on April 25,2025 at 10:19:43 UTC from IEEE Xplore. Restrictions apply.

for example, might be translated to the oriented edge features. When we reduce all Photoshop layers to one backdrop layer, this is known as flattening. Layers can increase the size of a file while also using processing resources. We can merge several layers or even flatten the entire image to one background layer to save file size. The dense layer is a deep layer of a neural network that is closely linked. It is the most popular and often utilized layer. The dense layer does the following operation on the input and provides the result. Connection strengths (weights), inhibition or excitement, and transfer functions are used to move the activation value from node to node. Before changing the value according to its transfer function, each node adds up the activation values it receives. Dropout can be implemented in Keras by adding Dropout layers to our network architecture. In every batch, each Dropout layer will remove a user-defined hyper-parameter of units from the previous layer. Keep in mind that in Keras, the input layer is presumed to be the initial layer and is not added using the 'Add' command. After the convolutional layer and before max pooling, RELU is a sort of nonlinearity that is commonly used in neural networks. In the feature map, all the negative pixel values are replaced with zero. After the convolutional layer, it is usually employed.

A RELU example is the max function(x, 0) with input x, matrix from a convolved picture. RELU sets all the negative values in the matrix x to zero, while all other values remain intact. After convolution, RELU is generated, resulting in an activation function that isn't linear such as Sigmoid. Instead of using the usual stochastic gradient descent process, Adam is an optimization technique that helps to upgrade the network weights repeatedly depending on training data. It has been discovered that numerous automated systems for analyzing emotions have been used in a variety of projects. The majority of them, however, are found without any sort of framework or instructions on how to put them to good use. Law enforcement can benefit from understanding and maintaining emotion analysis abilities. Officials track and recognize the emotion trends more effectively using machine learning approaches.

Take the photograph from the user first, then clean it up. Then isolate a person's face and apply the Haar features to it. The image should then be compared to the prior training dataset. Python's Keras package. It employs a convolutional neural network (CNN). CNN employs a sequential model. Conv2D, AveragePooling2D, MaxPooling2D, Dropout, Activation, Dense, and Flatten are some of the other layers used. These layers select the emotion from the categorization set after the approach. This is the finished product. The normalized facial image is submitted to the feature extraction phase after some pre-processing (if necessary) to locate the essential features that will be used for classification. The complete flowchart of the methodology is shown in fig. 1.

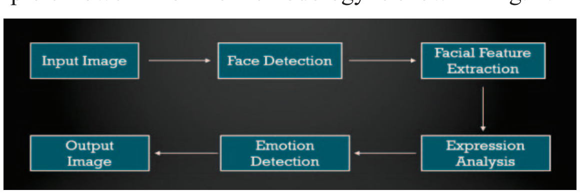

Fig. 1. Flowchart for the proposed work

IV. Results

The limited amount of data available for constructing a comprehensive framework was the first big issue in nature, which must be defeated. The most common response to this is movement learning. This process started with a pre-planned strategy and hence we calibrated the model with data acquired from the real-world. A series of preliminary studies confirms the hypothesis that face recognition would be more effective in highlighting the extraction. There are models that show how such systems can be used efficiently. On the datasets of a few hundred highlights or segments, machine learning methods operate very well. The system successfully classifies an image, classifies its sentiment, and chooses the image's matching emotion. The choice of profound understanding classifier is based on the fact that it processes the data across multiple layers. On the other hand, a deep learning algorithm can be most effective for less unestimated difficulties because large amount of data can be ingressed through it. For photos, the standard benchmark for developing deep learning models for broad picture recognition is more than 14 million images. A decision tree was used in the process to perfectly visualize the emotion detection pattern analysis. The character is represented by the nodes and layers in the experiment's decision tree, and conclusion is represented by the branch. If the data have been classified based on their movement, reactions, and order, it ideally corresponds to different sorts of emotions. This has also been categorized into trees and sub-trees, indicating if the person is sad, angry, or pleased, among other things. The various kinds of emotions are shown in fig. 2. To do this, a retrain method was utilized, which memorizes the pattern and satisfies the criterion. All the way to the top of the tree is continued when any of the conditions is met. It will cease checking and report "The emotion cannot be identified" if none of the conditions satisfy the intermediate condition. The emotion is unspecified if the emotions are not easy to apprehend. For the same emotion, various people have different ways of expressing it. Advanced ML technology can assist law enforcement authorities in estimating emotion, allowing machines to understand human emotion and behave and act more like humans.

The Keras library was used to classify and evaluate the emotions and obtain the data. The emotion is then identified with the use of Haar features and Numpy. With the help of Anaconda's platform, it creates an output from raw data, and the outcome is displayed in real-time. The decision tree is a systematic data mining process that aids in the generation of probability decisions by computing several features that are initially utilized to define the emotion pattern. It also created an effective field study to gather more people and diverse human beings, as well as varied emotional deferent expressions and a variety of faces, in addition to offline and online data collecting. The data set for online data gathering comes from kaggel.com. They deliver high-standard data sets. They used the numerical number of the photographs to transform the photos to pixel grayscale. As a result, it provides high-quality data and superior results.

The comparison between training loss and validation loss is shown in fig. 3 while fig. 4 shows the comparison between training accuracy and validation accuracy. The training loss in fig. 3 indicates how perfectly the model is fitting the training data, while the validation loss indicates how well the

Authorized licensed use limited to: Indian Institute of Technology Indore. Downloaded on April 25,2025 at 10:19:43 UTC from IEEE Xplore. Restrictions apply.

model fits new data. It might be observed that initially there is not much difference between training loss and validation loss but later on validation loss is higher that might be the case of overfitting. The remedies to this are to reduce our network size or to increase dropout. The training set is used to train the model, while the validation set is only utilized to estimate the model's behavior, Training accuracy and validation accuracy shown in fig. 4 are measured to evaluate model fitting. When there is a significant difference between these two, our model is overfitting. Initially, there is not much difference between validation accuracy and training accuracy.

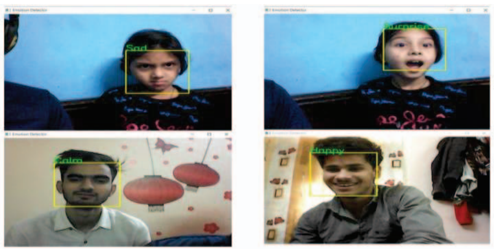

Fig. 2. Various kinds of emotions

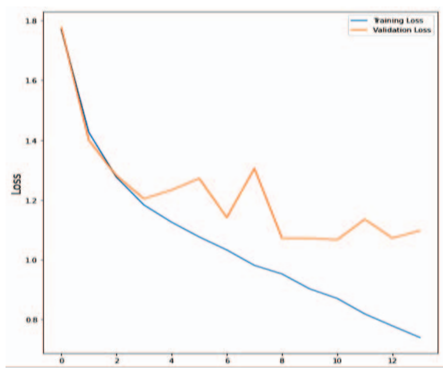

Fig. 3. Comparison of training loss & validation loss

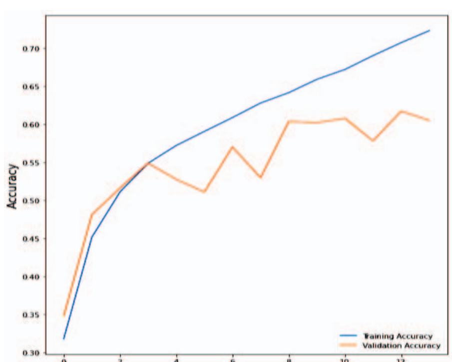

Fig. 4. Comparison of training accuracy & validation accuracy

The experts thought that using sentiment analysis to help identify emotion more correctly and perform accurate actions in response to accurate emotion identification could be beneficial. It would provide more information about the many types of expressions of their feelings as well as the percentage of each type of emotion that exists. We discovered while working on this work that if it wants to improve its accuracy, it will need a vast amount of test data and keywords. Extending the study activity also necessitates a lack of a sufficient number of raw data. If we wish to manage a large data of testing in the quickest amount of time, there is a need of a proper-layout graphics processing unit (GPU) qualified computer. So, if it receives large data in a rigid computer, it will be simple to raise the precision to more than 97 percent. It will also be able to apply that method to a different platform, achieve a different result and assist in determining the emotion expression pattern.

V. CONCLUSION

By evaluating and staring at another human, an experienced human may typically detect his or her emotions Currently, machines are tending to mimic human behavior. If the machine has been programmed to respond in the spirit of human sentiment at the moment, the machines can then act and behave like a person. On the other side, if the machine can recognize emotions, it can avoid a lot of things from happening. Data mining can promote precise expression frameworks by allowing machines to perform much like human beings with enhanced competency and error-free computing emotion. This framework was constructed or structured using comprehensive research and field studies to determine the emotion expression patterns. This method followed the framework layer by layer to achieve the desired result. It was feasible to identify emotions, as well as the type of emotion, in an actual photograph using these techniques. To better visualize the results and methods, decision tree approaches were implemented, which aid in determining which emotion percentages are large and which emotion percentages are small. The emotions with a low percentage have a low probability of surviving. It is now possible to discern precise emotions with obligation to this research. Machines, on the other hand, can more correctly discern emotion and, as a result, can provide appropriate responses. This machine has the potential to take the place of a human. In this work, the accuracy and loss have also been compared for the training and validation phase that show the nice results.

VI. FUTURE SCOPE

Facial emotion detection is a field arisen since past a few years. It may also be performed with the neural networks like Recurrent Neural Networks (RNNs) in near future. This concept may be then utilized in intelligence, military and forensic fields for identification purposes.

REFERENCES

[1] Ekman, P. (1993). Facial expression and emotion. *American psychologist*, *48*(4), 384.

[2] Fridlund, A. J. (2014). *Human facial expression: An evolutionary view*. Academic Press.

[3] O'Toole, A. J., & Castillo, C. D. (2021). Face Recognition by Humans and Machines: Three Fundamental Advances from Deep Learning. *Annual Review of Vision Science*, *7*, 543-570.

[4] Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). *SN Applied Sciences*, *2*(3), 1-8.

[5] Kamalraj, R., Neelakandan, S., Kumar, M. R., Rao, V. C. S., Anand, R., & Singh, H. (2021). Interpretable filter based convolutional neural network (IF-CNN) for glucose prediction and classification using PD-SS algorithm. *Measurement*, *183*, 109804.

[6] Bakshi, G., Shukla, R., Yadav, V., Dahiya, A., Anand, R., Sindhwani, N., & Singh, H. (2021). An Optimized Approach for Feature Extraction in Multi-Relational Statistical Learning. *Journal of Scientific and Industrial Research (JSIR)*, *80*(6), 537-542.

Authorized licensed use limited to: Indian Institute of Technology Indore. Downloaded on April 25,2025 at 10:19:43 UTC from IEEE Xplore. Restrictions apply.

[7] Cruz, A., Garcia, D., Pires, G., & Nunes, U. (2015, January). Facial Expression Recognition based on EOG toward Emotion Detection for Human-Robot Interaction. In BIOSIGNALS (pp. 31-37).

[8] Wang, S., Liu, Z., Lv, S., Lv, Y., Wu, G., Peng, P., ... & Wang, X. (2010). A natural visible and infrared facial expression database for expression recognition and emotion inference. IEEE Transactions on Multimedia, 12(7), 682-691

[9] Sun, L., Dai, J., & Shen, X. (2021, June). Facial emotion recognition based on LDA and Facial Landmark Detection. In 2021 2nd International Conference on Artificial Intelligence and Education (ICAIE) (pp. 64-67). IEEE.

[10] Sridhar, R., Wang, H., McAllister, P., & Zheng, H. (2018, July). E-Bot: a facial recognition based human-robot emotion detection system. In Proceedings of the 32nd International BCS Human Computer Interaction Conference 32 (pp. 1-5).

[11] George, A., Wimme, H., & Rebman, C. M. (2020). Artificial Intelligence Facial Expression Recognition Foremotion Detection: Performance and Acceptance. Issues in Information Systems, 21(4), 81-91.

[12] Kulkarni, K. R., & Bagal, S. B. (2015, December). Facial expression recognition. In *2015 Annual IEEE India Conference (INDICON)* (pp. 1-5). IEEE.

[13] **https://imotions.com/blog/facial-action-coding-system/**

[14] Zadeh, M. M. T., Imani, M., & Majidi, B. (2019, February). Fast facial emotion recognition using convolutional neural networks and Gabor filters. In *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI)* (pp. 577-581). IEEE.

[15] Lajevardi, S. M., & Hussain, Z. M. (2009, February). Facial expression recognition using log-Gabor filters and local binary pattern operators. In *Proceedings of the International Conference on Communication, Computer and Power* (pp. 349-353).

[16] Kundu, T., & Saravanan, C. (2017, December). Advancements and recent trends in emotion recognition using facial image analysis and machine learning models. In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* (pp. 1-6). IEEE.)

[17] Xiaoxi, M., Weisi, L., Dongyan, H., Minghui, D., & Li, H. (2017, August). Facial emotion recognition. In *2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)* (pp. 77-81). IEEE.

[18] Tanwar, P., Arora, D., & Anand, D. (2016). Facial Expression Detection using Hidden Markov model. *International Journal of Scientific & Engineering Research*, *7*(12), 25-28.

[19] Murugappan, M. (2011, June). Human emotion classification using wavelet transform and KNN. In *2011 International Conference on Pattern Analysis and Intelligence Robotics* (Vol. 1, pp. 148-153). IEEE.

[20] Liliana, D. Y. (2019, April). Emotion recognition from facial expression using deep convolutional neural network. In *Journal of physics: conference series* (Vol. 1193, No. 1, p. 012004). IOP Publishing.

[21] Shukla, R., Dubey, G., Malik, P., Sindhwani, N., Anand, R., Dahiya, A., & Yadav, V. (2021). Detecting crop health using machine learning techniques in smart agriculture system. *Journal of Scientific and Industrial Research (JSIR)*, *80*(08), 699-706.

[22] Singh, H., Rehman, T. B., Gangadhar, C., Anand, R., Sindhwani, N., & Babu, M. (2021). Accuracy detection of coronary artery disease using machine learning algorithms. *Applied Nanoscience*, 1-7.

[23] Saini, P., & Anand, M. R. (2014). Identification of Defects in Plastic Gears Using Image Processing and Computer Vision: A Review. *International Journal of Engineering Research*, *3*(2), 94-99.

[24] Sindhwani, N., Anand, R., Meivel, S., Shukla, R., Yadav, M. P., & Yadav, V. (2021). Performance Analysis of Deep Neural Networks Using Computer Vision. *machine learning*, *15*, 17.

[25] Kumar, R., Anand, R., & Kaushik, G. (2011). Image Compression Using Wavelet Method & SPIHT Algorithm. *Digital Image Processing*, *3*(2), 75-79.

[26] Choudhary, P., & Anand, M. R. (2015). Determination of Rate of Degradation of Iron Plates due to Rust using Image Processing. *International Journal of Engineering Research*, *4*(2), 76-84.

[27] Chandra, M. A., & Bedi, S. S. (2021). Survey on SVM and their application in image classification. *International Journal of Information Technology*, *13*(5), 1-11.

[28] Gopi, A. P., Jyothi, R. N. S., Narayana, V. L., & Sandeep, K. S. (2020). Classification of tweets data based on polarity using improved RBF kernel of SVM. *International Journal of Information Technology*, 1-16

Authorized licensed use limited to: Indian Institute of Technology Indore. Downloaded on April 25,2025 at 10:19:43 UTC from IEEE Xplore. Restrictions apply.