International Journal of Engineering in Computer Science

E-ISSN: 2663-3590
P-ISSN: 2663-3582
IJECS 2023; 5(1): 54-60
Received: 01-01-2023
Accepted: 05-03-2023

Sarthak Aggarwal
B.Tech, ECE, Department of Electronics and Communication, Bharati Vidyapeeth College of Engineering, New Delhi, India

Mohit Kumar Sharma
B.Tech, ECE, Department of Electronics and Communication, Bharati Vidyapeeth College of Engineering, New Delhi, India

Hari Om Kumar Jha
B.Tech, ECE, Department of Electronics and Communication, Bharati Vidyapeeth College of Engineering, New Delhi, India

Akhand Pratap Singh
B.Tech, ECE, Department of Electronics and Communication, Bharati Vidyapeeth College of Engineering, New Delhi, India

Dr. Ruchi Sharma
Assistant Professor, Department of ECE, Bharati Vidyapeeth College of Engineering, New Delhi, India

Gaurav Mitra
Assistant Professor, Department of ECE, Bharati Vidyapeeth College of Engineering, New Delhi, India

Correspondence
Sarthak Aggarwal
B.Tech, ECE, Department of Electronics and Communication, Bharati Vidyapeeth College of Engineering, New Delhi, India

Comparing pre-trained algorithms for facial emotion recognition: An analysis

Sarthak Aggarwal, Mohit Kumar Sharma, Hari Om Kumar Jha, Akhand Pratap Singh, Dr. Ruchi Sharma and Gaurav Mitra

DOI: https://doi.org/10.33545/26633582.2023.v5.i1a.90

Abstract
Facial emotion recognition is an evolving field that aims to detect and interpret human emotions based on facial expressions. This technology combines computer vision, machine learning, and artificial intelligence techniques to analyze facial features and accurately classify emotions such as happiness, sadness, anger, fear, surprise, and disgust. The human face is a powerful channel for emotional communication, and facial emotion recognition seeks to replicate this ability using computational algorithms. By capturing and analyzing facial cues such as eyebrow movements, eye widening, lip curvature, and changes in skin color, these systems can infer emotional states in real-time. The applications of facial emotion recognition are diverse and impactful. In psychology, it enables researchers to study emotional responses and behavior, providing insights into mental health conditions and aiding therapists in understanding their clients' emotional states. In human-computer interaction, this technology allows for more intuitive and personalized interactions, enhancing user experiences. In marketing and advertising, facial emotion recognition helps companies understand consumer emotional engagement and tailor their strategies accordingly. In healthcare, it can aid in the diagnosis and monitoring of mental health conditions, enabling personalized treatment options.

Keywords: Algorithms, facial emotion recognition, intelligence techniques

Introduction

Facial emotion recognition is a rapidly evolving field of research and technology that aims to detect and interpret human emotions based on facial expressions. It combines computer vision, machine learning, and artificial intelligence techniques to analyze facial features and infer emotional states accurately. This innovative technology has numerous applications in various domains, including psychology, human-computer interaction, marketing, entertainment, and healthcare.

The human face is a powerful channel for emotional communication. We instinctively rely on facial expressions to understand the emotions and intentions of others. Facial emotion recognition technology seeks to replicate this human capability using computational algorithms and models. By analyzing facial cues such as eyebrow movements, eye widening, lip curvature, and changes in skin color, these systems can identify and classify emotions such as happiness, sadness, anger, fear, surprise, and disgust.

One of the primary components of facial emotion recognition is computer vision, which involves capturing, processing, and analyzing visual information. With the advancements in camera technology, high-resolution images and videos of faces can be obtained with ease. Computer vision algorithms extract key facial landmarks and features, such as the position of the eyes, nose, and mouth, to form a representation of the face. This representation serves as the input for subsequent analysis and emotion classification.

Machine learning plays an important role in facial emotion recognition by training models to recognize and distinguish between various emotional states. These models are originally trained using enormous collections of labeled facial expressions that have been annotated with the emotions of human specialists. Through supervised learning, the models learn to associate distinct patterns in facial features with corresponding emotions. With time and exposure to different types of data, these models enhance their capacity to discern emotions from facial expressions.

Artificial intelligence techniques are employed to enhance the performance of facial emotion recognition systems. Deep learning, a subfield of artificial intelligence, utilizes neural networks with multiple layers to extract complex features and learn hierarchical representations of facial expressions. Convolutional Neural Networks (CNNs) are commonly used in this context, as they excel at processing visual data. By combining these advanced techniques, facial emotion recognition systems can achieve impressive accuracy levels in emotion classification tasks.

Facial emotion recognition has a vast and varied range of applications. Researchers in the field of psychology study emotional reactions and behavior with this technology. Mental health issues like anxiety, depression, and autism spectrum disorders are discussed with valuable insights. Therapists and counselors can improve their understanding of their clients' emotional states by effectively detecting and evaluating emotions in real-time.. (O. Arriaga, 2017) [9].

Facial emotion recognition is also leveraged in human-computer interaction to create more intuitive and personalized user experiences. Intelligent systems can adapt their responses and behaviors based on the detected emotions of users. For instance, virtual assistants can adjust their tone and language to better match the user's emotional state, leading to more effective and empathetic interactions.

Moreover, facial emotion recognition finds applications in marketing and advertising. Companies can analyze the emotional responses of consumers to their products, advertisements, or brand experiences. This information helps them tailor their marketing strategies and create emotionally resonant content to engage and connect with their target audience.

In healthcare, facial emotion recognition can aid in the diagnosis and monitoring of mental health conditions. By analyzing the facial expressions of patients, healthcare professionals can detect subtle changes in emotional states, providing early intervention and personalized treatment options. (E. Jyoti and E. A. S. Walia, 2017) [10].

Facial emotion recognition is a fascinating and rapidly advancing field that combines computer vision, machine learning, and artificial intelligence techniques to interpret human emotions based on facial expressions. Its wide-ranging applications make it a valuable tool in psychology, human-computer interaction, marketing, and healthcare. As the technology continues to evolve, we can expect even greater accuracy and sophistication in detecting and understanding human emotions from facial cues.

Literature Review

Khaireddin, Y., & Chen, Z. (2021) [1]. proposed a Neural Network-based upright frontal face detection system. Small windows of the image are scanned by connected Neural Networks and the decision is made based on whether the face is contained in the window or not. The performance could be improved using multiple networks instead of a single network.

Xiaoxi, M *et al* (2017) [2] demonstrated a new method to detect faces in color images based on the fuzzy theory. Two fuzzy models were created to describe the skin and hair color, respectively. Uniform color space was used to describe the color information to increase the accuracy and stableness. The two models were used to extract the skin color and hair color regions and compare with the prebuilt head-shape models by using a Fuzzy Theory based Pattern-Matching Method to detect face candidates. The proposed method not only failed to detect the real face but also gave some false positives under some conditions.

Durand, K., *et al* (2007) [3] described a statistical method for 3D object detection. Product of histograms has been used to represent the statistics of both object and non-object appearance. Each histogram represents the joint statistics of a subset of wavelet coefficients and their position on the object. The algorithm developed is the first of its kind that can detect human faces with out-of-plane rotation.

Jain, D. K., (2019) [4] developed a combined approach using Adaptive Hough Transform, Template Matching, Active Contour Model, and Projective Geometry properties. Adaptive Hough transform has been used to detect the curves. Template Matching Technique has been used for locating the inner facial features and Active Contour Model for inner face contour detection. For accurately determining the pose, projective geometry properties have been utilized.

Surcinelli, P (2006) [5] proposed AdaBoost-based face detection technique. In this technique, a face detection framework that is capable of processing the images very quickly while achieving high detection rates was presented. For the fast computation of the features by the face detector, the concept of Integral Image was first introduced for image representation. A classifier has been developed using the AdaBoost Learning Algorithm. The classifiers are combined in cascade which allows the background regions of the image to be rejected while spending more computation on face-like regions.

Ryu and Oh (2016) [13] developed an algorithm for the extraction of eyes and mouth using the rectangular fitting from gray-level face images based on Eigenfeatures and Neural Networks. Since Eigenfeatures and sliding window were used, large training set was not required. The performance of the algorithm reduces when the face images are with glasses or beard.

Turetsky, B. I *et al* (2007) [6] introduced a technique for face detection and facial features extraction using Genetic Algorithm and Eigenfaces. Chen *et al* (2017) [14] presented an algorithm to detect multiple faces in a complex background. It was assumed that in the frontal-view face images, the centers of the two eyes and the center of the mouth forms an isosceles triangle, and in the side-view face images, the center of a eye, the center of a ear hole, and the center of the mouth forms a right triangle. However, the algorithm fails to perform well for darker images and when the eyes are occluded by hair. Consequently, many improved methods were investigated.

Collin, L., Bindra, (2013) [7] proposed a novel learning procedure called FloatBoost method for training the classifier. FloatBoost learning uses a backtrack mechanism after each iteration of AdaBoost learning to minimize the error rate. A method called Backtracking was employed to remove the unfavorable classifiers from the existing classifiers. A new statistical model has been introduced for learning best weak classifiers using a stage wise approximation of the posterior probability.

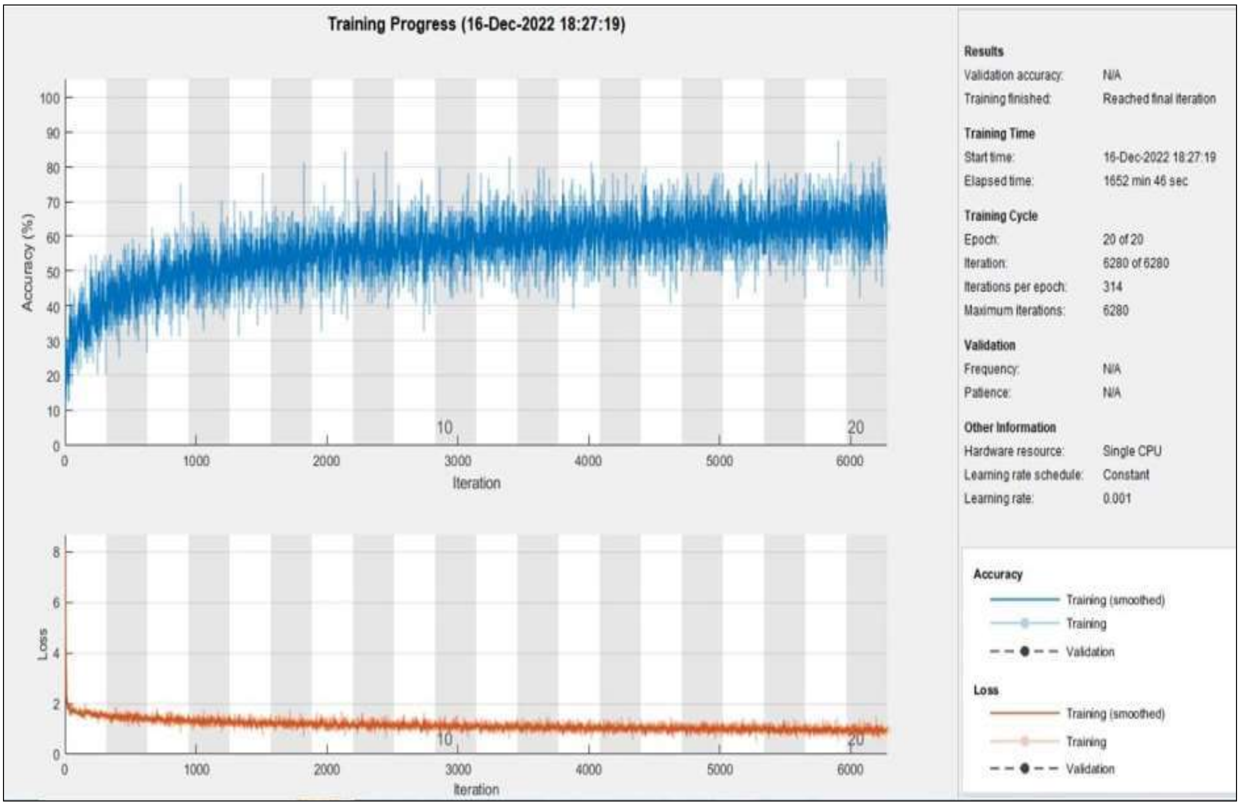
Methodology

Conventional approaches in facial emotion recognition (FER) refer to the traditional methods and techniques used to detect and classify emotions from facial expressions. These approaches were widely utilized before the advent of deep learning and convolutional neural networks (CNNs), which have significantly advanced the field. However, it is important to understand the conventional methods as they form the foundation for more recent developments. Here are some commonly used conventional FER approaches:

1. **Feature-Based Methods:** Feature-based methods involve extracting specific facial features or landmarks from facial images or videos and using them as input for emotion classification. Examples of facial features include the position and shape of the eyebrows, mouth, and eyes. Common feature extraction techniques include the use of geometric features, appearance-based features, or a combination of both. These features are often fed into machine learning algorithms such as Support Vector Machines (SVM), Hidden Markov Models (HMM), or decision trees to classify emotions.
2. **Facial Action Coding System (FACS):** FACS is a comprehensive system developed by Paul Ekman and Wallace Friesen for describing and classifying facial movements, known as action units (AUs). AUs are anatomically defined facial muscle movements associated with various emotional expressions. FACS provides a standardized framework for analyzing facial expressions and has been widely used in conventional FER. The FACS system assigns codes to different AUs, allowing for the quantification and interpretation of facial expressions.
3. **Appearance-Based Methods:** Appearance-based methods focus on analyzing the overall appearance of facial images without explicitly extracting facial features. These methods often involve techniques such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), or Local Binary Patterns (LBP). PCA and LDA aim to reduce the dimensionality of facial images while preserving discriminative information. LBP, on the other hand, captures local texture information by analyzing patterns in the image. These reduced-dimensional representations are then used as features for emotion classification.
4. **Gabor Wavelets:** Gabor wavelets are often employed in FER as they provide a multi-resolution analysis of facial images. These wavelets are designed to capture local variations in spatial frequency and orientation, mimicking the receptive fields of neurons in the visual cortex. Gabor wavelets can extract facial texture information that is informative for emotion recognition. These features can be fed into classifiers such as SVM or k-nearest neighbors (k-NN) for emotion classification.
5. **Decision Fusion:** Decision fusion methods combine the outputs of multiple classifiers or feature extraction techniques to improve the overall emotion classification performance. By fusing the decisions of different models or features, more robust and accurate results can be obtained. Common fusion strategies include majority voting, weighted voting, or using a fusion algorithm such as the Dempster-Shafer theory or Bayesian methods.

While conventional FER approaches have made significant contributions to the field, they often face challenges in handling complex and subtle variations in facial expressions. Deep learning and CNN-based approaches have shown superior performance in capturing intricate features and learning complex representations, leading to improved FER accuracy. However, understanding the conventional approaches is still valuable as they provide insights into the historical progression of the field and can serve as a baseline for comparison with more recent methods.

Results and Discussion

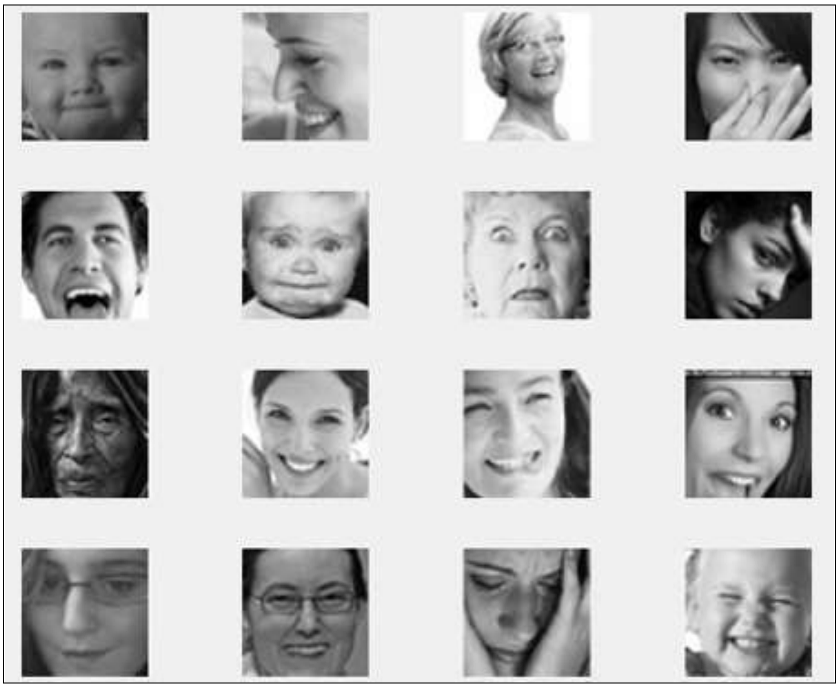


Result of Alexnet in the software MATLAB after 20 epochs with 63% accuracy and loss 0.08765

The results of training the AlexNet model using the MATLAB software for a total of 20 epochs. The accuracy of the model is reported to be 63%. Accuracy is a measure of how well the model predicts the correct output compared to the actual output. In this case, the model correctly predicted the outcome for 63% of the instances. It indicates the overall performance of the model in terms of correctly classifying the data. The loss value is reported as 0.08765. Loss is a metric that quantifies the dissimilarity between the predicted output and the actual output. It represents the error or the difference between the predicted values and the ground truth values. A lower loss value indicates a better fit of the model to the data. these results suggest that after training the AlexNet model for 20 epochs using MATLAB, it achieved an accuracy of 63% and a relatively low loss of 0.08765.

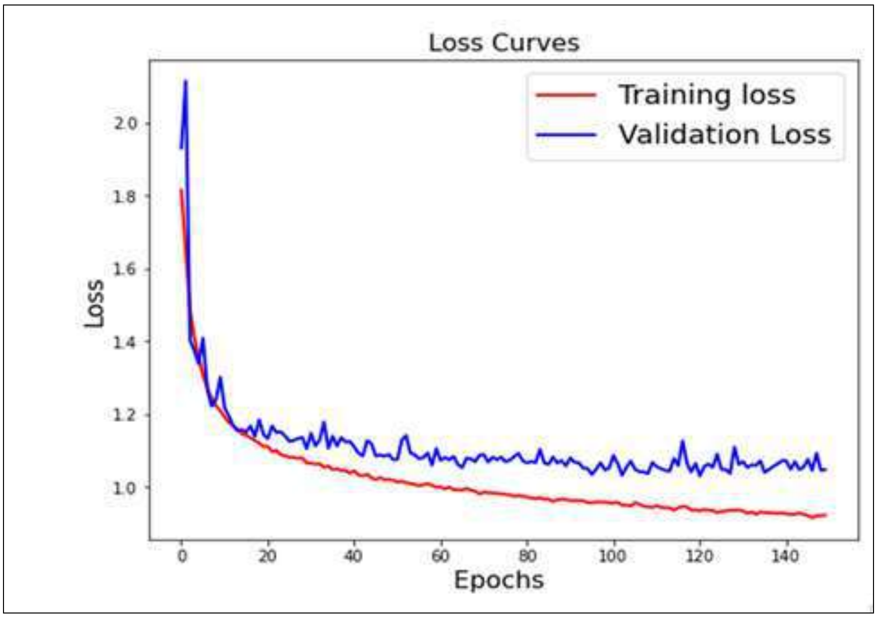
Predictions of FER 2013 Dataset Test Set Photos

The FER 2013 dataset is a widely used dataset in the field of facial expression recognition (FER). The predictions refer to the estimated facial expressions or emotions assigned to each test set photo by a trained model or algorithm. The FER 2013 dataset is a dataset that consists of facial images categorized into different emotion classes. It is often used to train and evaluate models for facial expression recognition tasks. The dataset may contain images of individuals displaying various facial expressions under controlled or natural conditions. The test set photos refer to a subset of the FER 2013 dataset that is specifically used for evaluating the performance of the trained model. The model is applied to these unseen images, and predictions are made based on the learned patterns and features from the training data.



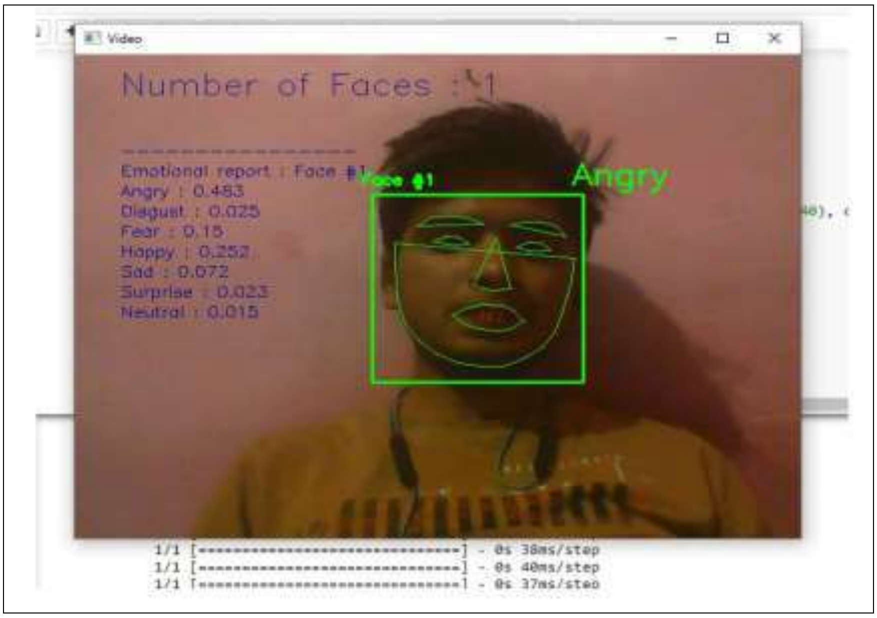
Dataset CK Plus Images Screenshot

The provided data mentions "Dataset CK Plus Images Screenshot," which suggests that it includes screenshots or images from the CK Plus dataset. The CK Plus dataset is a widely used dataset in the field of facial expression analysis and recognition. It contains a collection of facial images representing various facial expressions. The dataset is often used for training and evaluating models and algorithms in the field of emotion recognition. The term "images screenshot" implies that the data consists of screenshots taken from the CK Plus dataset. These screenshots could be visual representations of the facial images present in the CK Plus dataset. It's important to note that while the CK Plus dataset typically consists of labeled facial images with corresponding emotion labels, the term "screenshot" suggests that the provided data might be a visual representation or snapshot of a subset of images from the CK Plus dataset.



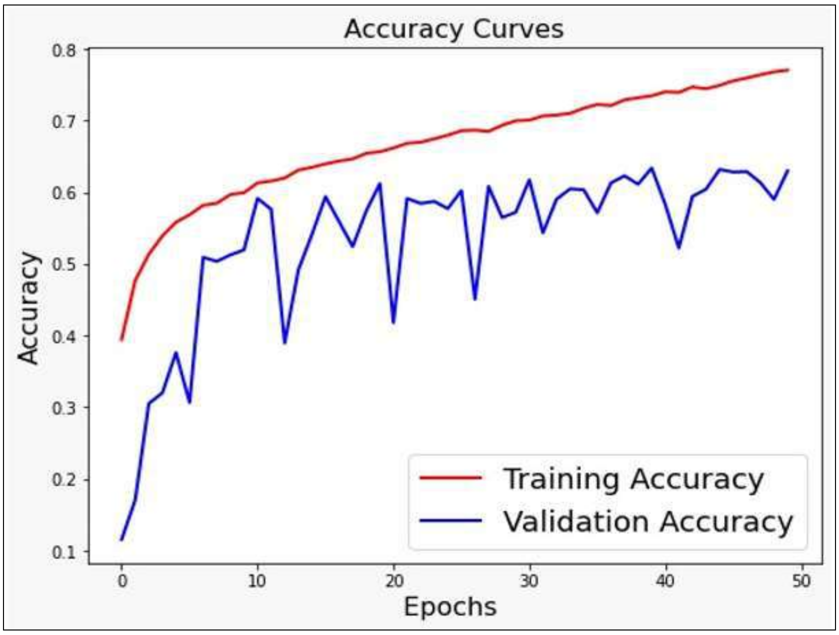
Loss Curve of Inception NET on Python Jupyter Notebook for FER 2013

The data provided refers to the loss curve of the Inception NET model trained on the FER 2013 dataset using a Python Jupyter Notebook. The loss curve illustrates the changes in the loss function's value over the course of training. The loss function measures the discrepancy between the predicted output of the model and the actual ground truth labels. Inception NET is a convolutional neural network architecture commonly used for image classification tasks. It is known for its ability to efficiently process images and capture intricate patterns and features. In this case, the Inception NET model was employed to train and classify facial expressions in the FER 2013 dataset. The FER 2013 dataset is a dataset consisting of facial images categorized into different emotion classes. It is widely used for training and evaluating models in the field of facial expression recognition. The dataset may contain images of individuals displaying various emotions such as happiness, sadness, anger, surprise, etc.



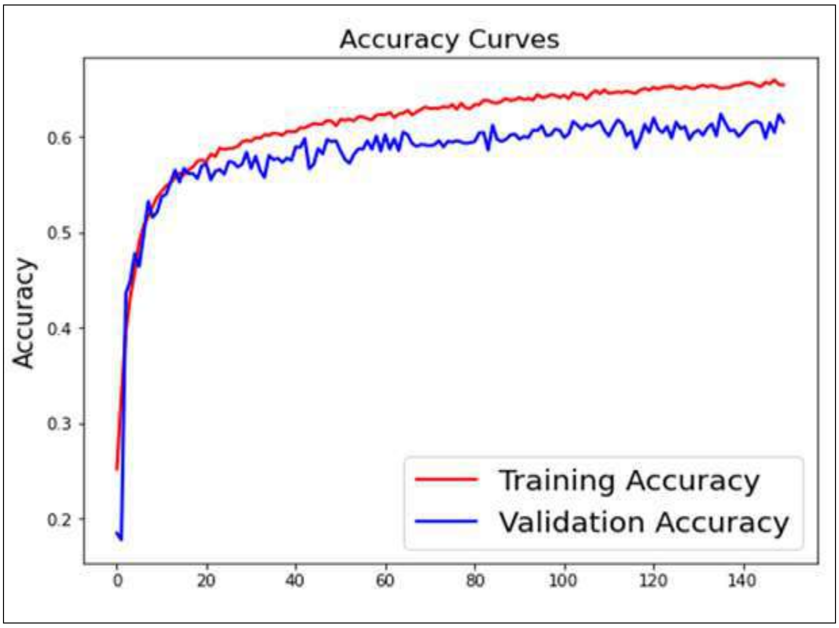
Real Time Prediction of Emotion using Haar Cascade Technique

Real-time prediction means making predictions or inferences in real-time as data is being received or processed. In this context, it implies that emotions are being predicted instantly or continuously as new data or input is provided. Emotion prediction involves determining or classifying the emotions expressed by individuals based on their facial expressions. This could include identifying emotions such as happiness, sadness, anger, surprise, fear, etc. The Haar Cascade technique is a popular object detection algorithm used for identifying specific objects or patterns within images or video streams. In the context of emotion prediction, the Haar Cascade technique is likely employed to detect and recognize facial features and expressions associated with different emotions. It utilizes a trained model, known as a Haar Cascade classifier, to identify specific patterns or features in the face that correspond to different emotions.



Accuracy Curve of FER 2013 Inception NET on CK+ Dataset

The FER 2013 dataset is a well-known dataset for facial expression recognition. It consists of facial images categorized into different emotion classes, such as happiness, sadness, anger, surprise, etc. In this case, the Inception NET model is being trained and evaluated on the FER 2013 dataset to recognize and classify facial expressions. Inception NET is a convolutional neural network architecture commonly used for image classification tasks. It is designed to efficiently process and extract features from images, enabling accurate classification. In this context, the Inception NET model is utilized to train and evaluate the CK+ dataset for facial expression recognition. The CK+ dataset is a widely used dataset for facial expression analysis. It consists of labeled facial images depicting various facial expressions. The dataset is often used for training and evaluating models in the field of facial expression recognition.



Accuracy Curve of Inception NET on FER 2013

The accuracy curve represents the performance of a model over the course of training or evaluation in terms of its accuracy. It illustrates how well the Inception NET model is able to correctly classify or predict facial expressions in the FER 2013 dataset. The accuracy curve is typically plotted with the number of training iterations or epochs on the x-axis and the corresponding accuracy values on the y-axis. Inception NET is a convolutional neural network architecture commonly used for image classification tasks. It is designed to process and extract features from images effectively, allowing for accurate classification. In this context, the Inception NET model is employed to train and evaluate the FER 2013 dataset for facial expression recognition. The FER 2013 dataset is a widely used dataset for facial expression recognition. It consists of facial images categorized into different emotion classes, such as happiness, sadness, anger, surprise, etc. The Inception NET model is trained on this dataset to learn the patterns and features associated with each facial expression, enabling accurate classification.

Conclusion

In conclusion, facial emotion recognition technology has emerged as a powerful tool for detecting and interpreting human emotions based on facial expressions. It combines computer vision, machine learning, and artificial intelligence techniques to accurately analyze facial features and classify emotions such as happiness, sadness, anger, fear, surprise, and disgust. The applications of this technology span across various domains, including psychology, human-computer interaction, marketing, and healthcare.

The accuracy of the model is reported to be 63%. Accuracy is a measure of how well the model predicts the correct output compared to the actual output. In this case, the model correctly predicted the outcome for 63% of the instances. It indicates the overall performance of the model in terms of correctly classifying the data.

It represents the error or the difference between the predicted values and the ground truth values. A lower loss value indicates a better fit of the model to the data. these results suggest that after training the AlexNet model for 20 epochs using MATLAB, it achieved an accuracy of 63% and a relatively low loss of 0.08765.

Facial emotion recognition has the potential to revolutionize fields such as psychology by providing deeper insights into emotional responses and behavior. It can aid therapists and counselors in understanding their clients' emotional states and designing more effective interventions. In human-computer interaction, this technology enables more intuitive and personalized interactions, as systems can adapt their responses based on the detected emotions of users. This leads to more engaging and empathetic user experiences.

In marketing and advertising, facial emotion recognition allows companies to analyze the emotional responses of consumers and tailor their strategies accordingly. By understanding how their target audience emotionally engages with their products or brand experiences, companies can create more resonant and impactful content.

In healthcare, facial emotion recognition has the potential to assist in the diagnosis and monitoring of mental health conditions. It can provide valuable information to healthcare professionals, enabling early intervention and personalized treatment options.

References

1. Khaireddin Y, Chen Z. Facial emotion recognition: State of the art performance on FER2013. arXiv preprint arXiv:2105.03588.
2. Xiaoxi M, Weisi L, Dongyan H, Minghui D, Li H. Facial emotion recognition. In 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP). IEEE, 2017 August, 77-81.
3. Durand K, Gallay M, Seigneuric A, Robichon F, Baudouin JY. The development of facial emotion recognition: The role of configural information. Journal of experimental child psychology. 2007;97(1):14-27.
4. Jain DK, Shamsolmoali P, Sehdev P. Extended deep neural network for facial emotion recognition. Pattern Recognition Letters. 2019;120:69-74.
5. Surcinelli P, Codispoti M, Montebarocci O, Rossi N, Baldaro B. Facial emotion recognition in trait anxiety. Journal of anxiety disorders. 2006;20(1):110-117.
6. Turetsky BI, Kohler CG, Indersmitten T, Bhati MT, Charbonnier D, Gur RC. Facial emotion recognition in schizophrenia: when and why does it go awry? Schizophrenia research. 2007;94(1-3):253-263.
7. Collin L, Bindra J, Raju M, Gillberg C, Minnis H. Facial emotion recognition in child psychiatry: a systematic review. Research in developmental disabilities. 2013;34(5):1505-1520.
8. Munim KM, Islam I, Khatun M, Karim MM, Islam MN. Towards developing a tool for UX evaluation using facial expression, in 2017 3rd International Conference on Electrical Information and Communication Technology (EICT); c2017. p. 1-6.
9. Arriaga O, Valdenegro-Toro M, Plöger P. Real-time convolutional neural networks for emotion and gender classification, ArXiv Prepr. ArXiv171007557; c2017.
10. Jyoti E, Walia EAS. A Review on Recommendation System and Web Usage Data Mining using K-Nearest Neighbor (KNN) method, Int. Res. J. Eng. Technol. IRJET. 2017;4(4):2931-2934.
11. Andra Adams, Marwa Mahmoud, Tadas Baltruˇsaitis and Peter Robinson, Decoupling facial expressions and head motions in complex emotions, 2015 International Conference on Affective Computing and Intelligent Interaction (ACII).
12. Yeasin M, Bullot B. From Facial Expression to Level of Interest: A Spatio-Temporal Approach, IEEE Conf. Computer Vision and Pattern Recognition. 2004;2:922-927.
13. Kim MJ, Bae SH, Ryu JC, Kwon Y, Oh JH, Kwon J, *et al*. SESN2/sestrin2 suppresses sepsis by inducing mitophagy and inhibiting NLRP3 activation in macrophages. Autophagy. 2016 Aug 2;12(8):1272-1291.
14. Chen HX, Chen W, Liu X, Liu YR, Zhu SL. A review of the open charm and open bottom systems. Reports on Progress in Physics. 2017 May 2;80(7):076201.