International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

FACIAL EXPRESSION ANALYSIS SYSTEM USING DEEP LEARNING

MATHAVAN K (1903035)
Department Of Computer Science And Engineering
PSN College Of Engineering And Technology
Tirunelveli.

MR.JOSEPH PUSHPARAJ,,M.E,
Assistant Professor,
Department Of Computer Science And Engineering,
PSN College Of Engineering And Technology,
TIrunelveli.

Abstract—**Emotion care for human well-being is important for all ages. Automatic emotion recognition plays a crucial role in various fields such as healthcare, human-computer interaction (HCI) and security and defense. Capturing the dynamics of facial expression progression in video is an essential and challenging task for facial expression recognition (FER). In this project, we propose a new low-cost and multi-user framework for emotion care system based on big data analysis for patient feelings, where emotion is detected in terms of facial expression. The system works with deep learning techniques on emotional big data to extract emotional features and recognize six kinds (e.g., angry, disgust, fear, happy, sad, surprise, and neutral) of facial expressions in real-time and offline. In addition, a new dataset for emotion recognition is collected to train the DCNN model. A deep convolutional neural network (DCNN) is further applied to the whole facial observation to learn the global characteristics of six different expressions.And the performed facial expression and predicted results can be saved into device to help cares to observe facial expressions of patients at any time and provide better suggestions to patients..**

INTRODUCTION

Affective computing is the study and development of systems and devices that can recognize, interpret, process, and simulate human affects. It is an interdisciplinary field spanning computer science, psychology, and cognitive science. While the origins of the field may be traced as far back as to early philosophical inquiries into emotion ("affect" is, basically, a synonym for "emotion."), the more modern branch of computer science originated with Rosalind Picard's 1995 paper on affective computing. A motivation for the research is the ability to simulate empathy. The machine should interpret the emotional state of humans and adapt its behavior to them, giving an appropriate response for those emotions.

Affective computing technologies sense the emotional state of a user (via sensors, microphone, cameras and/or software logic) and respond by performing specific, predefined product/service features, such as changing a quiz or recommending a set of videos to fit the mood of the learner.

affective computing is human-computer interaction in which a device has the ability to detect and appropriately respond to its user's emotions and other stimuli. A computing device with this capacity could gather clues to user emotion from a variety of sources. Facial expressions, posture, gestures, speech, the force or rhythm of key strokes and the temperature changes of the hand on a mouse can all signify changes in the user's emotional state, and these can all be detected and interpreted by a computer. A built-in camera captures images of the user and algorithms are used to process the data to yield meaningful information. Speech recognition and gesture recognition are among the other technologies being explored for affective computing applications.

Recognizing emotional information requires the extraction of meaningful patterns from the gathered data. This is done using machine learning techniques that process different modalities, such as speech recognition, natural language processing, or facial expression detection.

Emotion in Machines

A major area in affective computing is the design of computational devices proposed to exhibit either innate emotional capabilities or that are capable of convincingly simulating emotions. A more practical approach, based on current technological capabilities, is the simulation of emotions in conversational agents in order to enrich and facilitate interactivity between human and machine. While human emotions are often associated with surges in hormones and other neuropeptides, emotions in machines might be associated with abstract states associated with progress (or lack of progress) in autonomous learning systems. In this view, affective emotional states correspond to time-derivatives in the learning curve of an arbitrary learning system. Two major categories describing emotions in machines: Emotional speech and Facial affect detection.

Problem Statement

Human beings are mostly emotional, and our social interaction is measured by taking into consideration our ability to communicate emotions and to perceive the emotional states of others. Affective computing provides computing systems with mechanisms that emulate and/or interpret human emotions. Its main objective is to make communication with computing systems easier and more natural. combination, starting from facial expressions, oral intonation, psycho-physiological information, or even the texts used. To make them known to users, it is usual to employ avatars and speech synthesis, frequently combining the two. Although, in general, people are experts in recognizing and expressing emotions, sometimes there is misunderstanding when transmitting them. This may be caused by ambient issues (noise, lighting, or distance between interlocutors), or even personal issues (concentration or the behavior or confidence with the interlocutor).

This is why emotional resources are frequently validated by people, in order to ascertain whether they really express the correct emotion or if the interlocutors are able to perceive them adequately. Many times, resources are not very expressive or not correctly understood by humans; therefore, computing Facial expression is one of the most natural and immediate means for human beings to communicate their emotions, as the human face can express emotions sooner than people verbalize or even realize their feelings. Automatic facial expression recognition (FER) has become an increasingly important research area that involves computer vision, machine learning, and behavioral sciences. Much progress has been made in building computer systems to understand and use this natural form of human communication, although most of these systems attempt to recognize only a small set of prototypical emotional expressions. FER can be used for many applications, such as security, human-computer interaction, driver safety, and health care.Emotions are multimodal, and currently, they are recognized within the human-computer interaction (HCI) area using the following factors, both separately and in emotion plays an important role in human life and communication. In the daily life, emotion is an inextricable part of the interaction of human beings, which can be observed by the changes in physiological features and behaviors. Because emotion recognition has a great potential to improve our quality of life, in the past decades, emotion recognition has aroused a lot of attention of many researchers and has been a popular research topic in various fields such as robotics, human-computer interaction, and entertainment, to name a few. Meanwhile, emotion care can be very useful in medical applications when medical staff need to assess the patient's feeling and behavior during or after the surgery. With the development of big data and deep learning, huge amount of data including emotional data is generated in recent years, which cannot be handled with the traditional techniques.

Scope of the Project

Each person is unique and can express emotions in their own characteristic way, depending on their culture, age, gender or previous life experiences.Facial expressions represent one of the most important modes of communication through which people express their emotions and intentions.This project proposes a DL-FER based emotion care system, where emotion is detected in terms of facial expression.The application calculates a score for each facial expression and shows the highest score and emotion.And the highest score indicates the performed emotion meets the standard better. And the performed facial expression and predicted results can be saved into device to help cares to observe facial expressions ofpatients at any time and provide better suggestions to patients.The main contributions of this project are

1)A new low-cost and multi-user framework for emotion detection is proposed. The system is based on a Web application that uses CNN model to classify facial expressions and works in real time and online.

2)Using facial features extracted with publicly available facial landmarks, action unit detection tools, and emotional video databases, we show that the proposed method for categorizing emotional photograms allows a valid set of emotionally labeled photograms that can then be used for emotion recognition..

LITERATURE SURVEY

1.Spatial Augmented Reality Based Customer Satisfaction Enhancement and Monitoring System

Author: Udaya Dampage; D. A. Egodagamage; A. U. Waidyaratne; D. A. W. Dissanayaka; A. G. N. M. Senarathn,Year:2021

Overview:The customer-satisfaction analytical-model provides environmental knowledge-based dynamic capabilities to the human-centered dining environment, which is also utilized as an input to logistics 4.0 supply chain management, business-processes, and decision support-systems for future analytics. The perception of the quality of food was enhanced by the recreation of the live environment on the preparation of the menu ordered, utilizing the waiting time after each ordering session. The dishes are being portrayed in three-dimensional (3D) virtual menu also, adding SAR features in a special angle creating a 3D illusion to the naked eye. Menu suggestions are also proposed depending on the recommendations of the analytical model. The deep learning model monitors customer-satisfaction levels through emotion recognition.

On the analysis of the ordered menu and based on preferences available within the database, menu suggestions are also proposed depending on the recommendations of the customer satisfaction analytical model. The deep learning-based facial expression recognition system is running in the background to monitor customer satisfaction levels through facial expressions. The captured data will be processed and loaded into the just-in-time logistic management model; where, the post-processing results will be utilized as an input to logistics 4.0 supply chain management, business processes, and decision-support-systems used for future analytics such as determination of the profit margin values and also handle the replenishment of the supplies.

2.Multimodal Emotion Recognition Fusion Analysis Adapting BERT With Heterogeneous Feature Unification

Author: Sanghyun Lee; David K. Han; Hanseok Ko,Year:2021

Overview:An effective communication among humans requires not only intellectual exchange but of sharing contextual emotions. While most humans are natural in perceiving others' emotional states, the sensitivities of recognizing key sentiments may not be even among us. When we look at Leonardo Da Vinci's 'Mona Lisa' many of us may judge that she is smiling and may conclude that her emotional state is positive. It turns out, however, that this seemingly prevalent observation may not necessarily be universal. In fact, an analysis of the painting has been undertaken to determine if the painting is really conveying a positive emotion. Perceiving other's emotional states is obviously an important factor in peer-to-peer human interactions.

The author proposed to integrate BERT in textual feature extraction in our architecture. In summary, propose to address these challenges of recognizing emotions from analyzing utterance-level multimodal input as follows. First, combine information from various unimodal features with relevant saliency, then efficaciously fuse them with appropriate placement of relative weights among the modalities for accurately recognizing emotions. Our proposed model, Heterogeneous Features Unification (HFU-BERT), integrates BERT into the architecture to effectively combine heterogeneous features extracted from both handcrafted and deep learning-based methods.

3.**Recognition of Teachers' Facial Expression Intensity Based on Convolutional Neural Network and Attention Mechanism**

Author:Kun Zheng; Dong Yang; Junhua Liu; Jinling Cui,Year:2020

Overview:Expert evaluation and sentiment analysis are used as feedback to teachers to improve teaching strategies, thereby increasing the learning efficiency of students. Therefore, it is not difficult for us to find that studying the changes of intensity and frequency of expressions could play an indispensable role in improving the quality of teaching. People have strong subjectivity about images or videos. Evaluating people's emotion with computer is able to avoid being subjective and one-sided. Facial expression recognition classifies expressions by extracting facial features in images or videos.The main works of this paper are shown as follows: 1) In order to perform the fine-grained division of expression data, 13 types of expression intensity dataset EIDB-13 are proposed and established. 2) A new convolutional neural network model called InceptionResNetV2+CBAM that can reflect the facial expression information of characters is proposed. Using the transfer learning method, the InceptionResNetV2 network model is used to extract the deep features of the facial expression pictures, and the attention module CBAM is inserted into the network to focus on details of the facial expression image. 3) Combine the face detection method in deep learning to detect the teacher's face and recognize his expression. The results can extract the teacher's transient expression strength information in the teaching video and provide data support for education researchers to study the influence of expression changes on the quality of teaching in the classroom.This paper designs an end-to-end facial expression intensity recognition system, which can detect and recognize the corresponding facial expressions of teachers in the teaching videos.

4.Multimodal Attention Network for Continuous-Time Emotion Recognition Using Video and EEG Signals

Author: Dong Yoon Choi; Deok-Hwan Kim; Byung Cheol Song,Year:2020

Overview: Recognition of human emotions is a key technology for ultimate human–robot interaction (HRI). In addition, emotion recognition has received much attention in the field of artificial intelligence. Conventional emotion recognition algorithms distinguished emotion categories by detecting changes in facial expressions.A few EEG-based algorithms employed the inherent asymmetry characteristics between EEG channels as salient features for deep learning-based emotion classification. However, the conventional techniques have a structure that recognizes only a single emotion per tens of seconds of video clip. So, it is hard to say that they can ultimately perceive emotional changes in the continuous-time domain. The proposed multimodal attention network analyzes intermediate features obtained from a video modality network and an EEG modality network, and determines the attention weight of each modality. As a result, the multimodal attention network contributes to improve the overall emotion recognition accuracy by selecting a more reliable one between video and EEG.Video modality and EEG modality have independent networks. The output features of the two networks are fused through the attention network, which calculates the attention weight of each modality. The weighted average of two modality outputs becomes the final emotion information.

5.WGAN-Based Robust Occluded Facial Expression Recognition

Author: Yang Lu; Shigang Wang; Wenting Zhao; Yan Zhao,Year:2019

Overview:FER is a challenging subject because it is an interdisciplinary technology, and the research and development of FER can promote both the theoretical significance and life applications. Currently, most of the related works of this technology is to identify un-occluded facial expression images, and the excellent research results are endless. GAN can be used for classification in addition to generating visually realistic images. The success of deep neural networks in the fields of image recognition and object classification largely depend on a large number of manually labeled training datasets. However, in many applications, such labeled data volumes often fail to meet the requirements of deep model training and adding unlabeled sample data into training can perform semi-supervised classification. The method proposed in this study is based on the Wasserstein Generative Adversarial Network (WGAN) model, which consists of one generator and two discriminators. In this paper the proposed weighted reconstruction loss function and triplet loss function are used to constrain the generator together to achieve the purpose of optimizing image quality.Different from other studies, this study takes the de-occluded generated images, the original un-occluded images and the irrelevant-area occluded images together as the input of two discriminators.To recognize facial expressions, the computer needs a large number of labeled facial expression image training data, the difference between our training and the traditional training is that we take the irrelevant-area occluded images as the input of the classifier discriminator, it is mainly considered that when the occluded area is relatively small, the occluded area is in an insignificant area such as the forehead, chin, cheek, and only one eye is masked. The classifier trained by adding irrelevant area occluded images can extract more real expression features than the generated images, in this case, the separability of feature set is better than that of using only generated images and original un-occluded images. Finally, the parameters of generator and discriminators are determined by optimizing the whole network using the four loss functions constructed in this study.

Existing System

As an important way to emotion recognition, there are many diverse FER methods that achieve a good performance SVM, Linear Discriminant Analysis (LDA), Bayesian Network (BN), Neural Network (NN), Gaussian mixture model (GMM).AdaBoost,PCA.

SVM:SVMs are maximal margin hyperplane classifiers that exhibit high classification accuracy for small training sets and good generalization performance on very variable and difficult to separate data. The trained SVM model is subsequently used to dynamically classify unseen feature displacements and the result is then returned to the user. Classification thus identifies the feature.

LDA:LDA is easy to implement and no tuning parameters or adjustment required. LDA returned the prior probability of each expression class, the group means for each covariate, the coefficient for each linear discriminant (for the six classes, we have five linear discriminants) and the singular values that produced the ratio of the within-class and between-class standard deviation on the first two LDs variables returned the proportions of the variance by Stirling and by Bosphorus.

Neural Network.

GMM:A Gaussian Mixture Model (GMMs) is trained for each of the emotional states that we examine; angry (ANG), happy (HAP), neutral (NEU) and sad (SAD). The GMMs for each separate modality give us a limited picture of the overall emotional impression of a sentence. GMM for the total face (TOTAL) is trained using marker data from all facial regions. Neighboring markers are averaged in order to reduce the total number of markers. The choice of which markers to average is ad-hoc, however, the markers that are averaged belong to the same facial muscles and their movements are correlated.

ANN: An Artificial Neural Network is a nonlinear and adaptive mathematical model which is inspired by biological neural networks. The brain consists of larger amount of interconnected set of nerve cells called neurons. An artificial neural network consists of a smaller number of interconnected set of nerves or very simple processors, also called neurons, which are analogous to the biological neurons. It consists of an interconnected group of neurons which operating in parallel and communicating with each other through weighted interconnections. Artificial neural network changes its structure during a learning phase because in most cases it is adaptive system. It is used to model a complex relationship between inputs and outputs or to find data patterns.ANN is proposed to detect faces with the purpose of decreasing the performance time but still achieving the desired faces detecting rate.

Proposed System

This section presents the proposed deeply learned classifiers for facial emotion classification of unfiltered real-life face images

We propose a model that uses DCNN architecture to predict the emotion of human's faces from unfiltered real-world environments. The novel CNN approach addresses the Emotion labels as a set of discrete annotations and train the classifiers that predict the human's Expressions.

We design a quality and robust image pre-processing algorithm that prepare and pre-process the unfiltered images for the CNN model and this greatly has a very strong impact on the performance accuracy of our facial Emotion classifiers.

We demonstrate that pretraining on large-scale datasets allows an effective training of our Facial Emotion CNN model which enable the classifiers to generalize on the test images and then avoid overfitting.

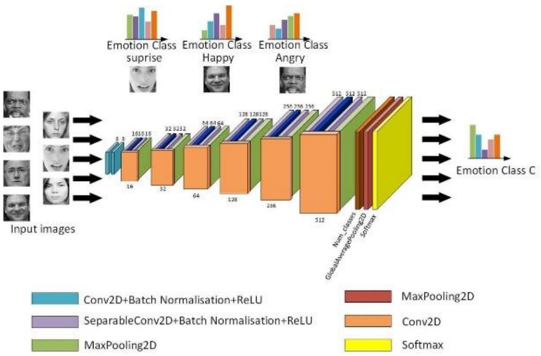
Health Care Recommendation System using predicted facial expressions.

Proposed System

For the system, image preprocessing is necessary beforean image is fed into the CNN model. The image preprocessingmainly consists of two stages: face detection, dataaugmentation. A face detector is adopted for face detectionin our system. If faces are detected, the four coordinatesof region of interest (ROI) of the faces would be returnedProject Flow Description to the system, the system would crop the faces and discardirrelevant background. Data augmentation is used to processthe detected face images and increase the quantity of data,because training process of deep learning model usuallyneeds huge amounts of data. The images are cropped by therandom bounding boxes that have different cropped rangesfrom 0.85 to 1. Then the data are randomly flipped androtated.After importing the haar cascade file we will have written a code to detect faces and classify the desired emotions. We have assigned the labels that will be different emotions like angry, happy, sad, surprise, neutral. As soon as you run the code a new window will pop up and your webcam will turn on. It will then detect the face of the person, draw a

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

bounding box over the detected person, and then convert the RGB image into grayscale & classify it in real-time. The DCNN model is developed using TensorFlow platform, which is an end-to-end open-source platform for machine learning.



1.Facial Data Set Annotated: Facial Expression Recognition 2013 (FER-2013) dataset was prepared in Challenges in Representation Learning: Facial Expression Recognition Challenge, which is hosted categories (e.g., angry, disgust, fear, happy, sad, surprise, and neutral) and three different sets such as training set (28.709 images), validation set (3.589 images), and test set (3.589 images). All images in this dataset are grayscale with 48 X 48 pixels, thus corresponding to faces with various poses and illumination, where several faces are covered by hand, hair, and scarves. Because of FER-2013 is collected from the Internet and has various real-world conditions, it becomes one of the largest and most challenging databases for facial expression recognition.

2.**Live Video Annotation:**Cameras should be deployed in critical areas to capture relevant video. Computer and camera are interfaced and here webcam is used.For every participant, one video with six kinds of facial expressions is collected and processed. A haar cascade classifier proposed by Viola and Jones is used to detect the face from video frame by frame. When a face is detected, the face image is saved into the database and labeled according to the facial expression the participant shows. Because the database has a lot of similar images due to the successive frames, we use difference hash (dhash) algorithm to select representative images from the dataset. The difference hash is one of image fingerprint algorithms, and it creates a unique hash value by calculating the difference between adjacent pixel values. To select images, form the dataset, we use difference hash to compute our image fingerprints because of its speed and accuracy.

3.**Preprocessing:**In that we will enhance the different features of images we get for example its intensity, contrast, saturation for different image processing. Low pass-filters a grayscale image that has been degraded by constant power additive noise. It uses a pixel wise adaptive Wiener method based on statistics estimated from a local neighborhood of each pixel.

4.**Face Detection:**The Background subtraction approach is mostly used when the background is static. The principle of this method is to use a model of the background and compare the current image with a reference. The foreground objects present in the scene are detected. It attempts to detect moving regions in an image by differencing between current image and a reference background image in a pixel-by-pixel fashion. We will use the static background for the image subtraction which will give us the human we have to track. This step detects objects of interest as they move about the scene. The action detection process is independently applied to all the static cameras present in the scene. For human recognition is feature extraction and representation where the important characteristics of image frames are extracted and represented in a systematically way as features.

5.Feature Engineering:Deep neural network (DNN) is a popular deep learning (DL) structure that consists of multiple-layered models of inputs. The DCNN architecture that we used to train and build the classifier model.The model can predict Facial Emotions.Hence, the DCNN model running at a fog node detects and labels the images with the name of the Facial Emotions having the highest probability, and saves those images.
6.Classification:In Classification stage, Convolutional neural networks algorithm is used for classification of Face Expression images. It is a non-parametric method which is used for both classification and regression.Deep Convolution Neural Network Classifier: The Deep Convolution Neural Network (CNN) classifier is used mainly for image and video recognition. The CNN is able for automatically learning the respective feature for data itself. The CNN follows few steps like receiving different inputs, calculating the sum of their weights, forward output to activation function and respond with the desired output.

7. **Prediction:**In this module the matching process is done with trained classified result and test Live Camera Captured Classified file. Hamming Distance is used to calculate the difference according to the result the prediction accuracy will be displayed.

8.Recommendation System: Facial Expression based recommend the health care systems where suggestions are based on an influence about a user's emotion and based on a degree of domain expertise and knowledge. Rules are defined that set context for each recommendation.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

CONCLUSION

In this project, we adopt a deep learning technique to process emotional big data and develop an emotion care system using facial expression recognition system. We propose an algorithm that recognizes emotional changes in continuous-time domain by using video. In addition, emotion recognition has received much attention in the field of artificial intelligence. Conventional emotion recognition algorithms distinguished emotion categories by detecting changes in facial expressions. Recently, various emotion recognition mechanisms based on convolutional neural network (CNN) which are trained in an end-to-end manner have been developed and showed reliable performance.The Graphical Web User Interface allows users to do Realtime validation of the system. We have considered seven discrete and unique emotion classes (angry, disgust, fear, happy, neutral, sad and surprise) for emotion classification. So, there is no overlapping among classes.

Future Enhancement

In the future work, we will try to implement our method on the videos for other applications and use different kinds of emotion representations (valence, arousal, etc.) for this task

References

[1] E. Ahmed, I. Yaqoob, I. A. T. Hashem, I. Khan, A. I. A. Ahmed, M. Imran, and A. V. Vasilakos, ``The role of big data analytics in Internet of Things,'' Comput. Netw., vol. 129, pp. 459 - 471, Dec. 2017.

[2] K. Lin, F. Xia, W. Wang, D. Tian, and J. Song, ``System design forbig data application in emotion-aware healthcare,'' IEEE Access, vol. 4,pp. 6901- 6909, 2016.

[3] Y. Jiang,W. Li, M. S. Hossain, M. Chen, A. Alelaiwi, and M. Al-Hammadi,``A snapshot research and implementation of multimodal informationfusion for data-driven emotion recognition,'' Inf. Fusion, vol. 53,pp. 209 - 221, Jan. 2020.

[4] A. Chatterjee, U. Gupta, M. K. Chinnakotla, R. Srikanth, M. Galley, andP. Agrawal, ``Understanding emotions in text using deep learning and bigdata,'' Comput. Hum. Behav., vol. 93, pp. 309317, Apr. 2019.

[5] Y. Ma, Y. Hao, M. Chen, J. Chen, P. Lu, and A. Ko²ir, ``Audio-visualemotion fusion (AVEF): A deep efficient weighted approach,'' Inf. Fusion,vol. 46, pp. 184 - 192, Mar. 2019.

[6] M. S. Hossain and G. Muhammad, ``Emotion recognition using deeplearning approach from audio-visual emotional big data,'' Inf. Fusion,vol. 49, pp. 69 - 78, Sep. 2019.

[7] Y. Miao, H. Dong, J. M. A. Jaam, and A. E. Saddik, ``A deep learningsystem for recognizing facial expression in real-time,'' ACM Trans. Multi-media Comput., Commun., Appl., vol. 15, no. 2, pp. 1 - 20, Jun. 2019, doi:10.1145/3311747.

[8] G. Muhammad, M. S. Hossain, and N. Kumar, ``EEG-based pathologydetection for home health monitoring,'' IEEE J. Sel. Areas Commun.,vol. 39, no. 2, pp. 603 - 610, Feb. 2021.

[9] A. E. Saddik, ``Digital twins: The convergence of multimedia technologies,''IEEE Multimedia, vol. 25, no. 2, pp. 8792, Apr./Jun. 2018.

[10] M. Chuah and M. Diblasio, ``Smartphone based autism social alertsystem,'' in Proc. 8th Int. Conf. Mobile Ad-hoc Sensor Netw. (MSN),Dec. 2012, pp. 6 - 13.