FACE EXPRESSION RECOGNITION USING OPEN CV AND CONVOLUTIONAL NEURAL NETWORK

Bangaru Lakshmi Mahanthi[1],Anuradha Sesetti [2], Vijaya Bharathi Manjeti[3]
[1,2,3] Department of CSE, GITAM deemed to be university, Visakhapatnam, India

Received on: 08/08/2023 Accepted on:05/09/2023

Keywords:
Deep learning Neural network, achine learning.

Abstract

The machines are constantly using for different purposes and the purposes are rising day by day. Perception for machines will help in performing a diversity of tasks even complex tasks. Machine perception involves machines to understand their physical environment and the motives of the conversation partner.We used some deep learning techniques in this project, such as Convolutionary neural networks on pictures, to categorize them into joy, sadness, rage, amazement, dislike and anxiety.CNN is giving better results than other statistical approaches so this approach is practiced. Utilizing CNN involves feature learning which is an important task to do. Moreover, the community turned into evaluated using two corpora, one becomes employed in the course of a community's education and it changed into also useful for a network's structure definition.The network providing first-class type accuracy results changed into examined against the second dataset. Even though the network became trained using the best one corpus, the network mentioned auspicious outcomes when examined on a special fact set, which displays facial emotions. While the consequences completed were no longer king of the network the proof gathered points out deep studying is probably suitable to classify facial emotion expressions. Thus deep mastering has the capacity to improve human system interaction because its capability to learn capabilities will permit machines to increase perception.

Corresponding Author: Email: lmahanth@gitam.edu

Introduction

The usage of machines in society has increased during the last year's especially.Present; they are utilized in many kinds of fields. As their exposure to human beings increases, interaction must also be better and more effortless.In achieving this, machines should have the capability to understand the encircling things A certain form of perception was created when machines would understand their environment. People use their senses to gain insights into their surroundings. Computer belief therefore attempts to mimic human senses to communicate with their world. We now have many ways to record their condition through cameras and sensors. Hence the usage of information with appropriate procedure permits to generate device belief. In recent years, it has proven to be a big success with the use of profound learning algorithms. For example, Jeremy Howard illustrated how TEDx could acquire brilliant tasks using profound study techniques[1-7]. This involves the possibility of interpreting language, recognizing

objects in photographs and assisting with clinical diagnosis. Effective computing claims that emotion analysis is significant for higher-quality of them. Across areas of older treatment across clinics. The use of robotics involves extensive awareness of the environment. Facial emotions are about the inner country of the subject. When a computer can obtain facial images, it can allow machines to understand their conversational partner mood by using DL strategies. In this sense, a deep understanding is likely to emerge as a fundamental problem for creating higher relationships between humans and machines, even as it gives machines with a little self-attention to their human peers and how it interacts more with their intellect [8-12].

2.1 Motivation and goals:

This mission is from the robotics laboratory's studies. The Visual Awareness and Interaction Division's Robotics Lab contributes to the design and creation of robots that can associate with humans in a social way. The aim is to support society by using robotics in fields such as e-learning, health and support technology [13].
The project is functionally a deep neural network of action with static pix of facial feeling. The network could then be real-time to detect the emotions of a program. Use this device; robots catch the inner state of their adversaries. This can be used for machines by having more appropriate answers in order to enhance their contact with people. The purpose and work of the Social Robotics Lab are therefore well-defined by this initiative [14].
Actually, this is an affective computing, machine-learning, computing vision multidisciplinary challenge. Another project goal is to know how these different fields are interconnected and how they can solve complex problems.

2.2 Methods and structure:

The project has been divided into two parts. The first step consisted of the usage of a data collection for fundamental instruction of facial expressions. Extended FER2019 is the picked data set-in addition; tests of its estimation accuracy were carried out on many network topologies. In recognition of its tremendous success on machine vision tasks, networks are based on topological topology. Google's library Tensor Flow wrapper Keras is used to execute the network and the training process [15-17].

The second step concentrated on the model validation for a new collection of results. Test FER2019 and Live Camera. The goal is to compare all data sets and check the generalization efficiency of the network. A concentration

was also made on certain parameters and its effect on the approximation of accuracy of the models. These parameters were selected because they affect the actions of the networks:

- Network loss
- Learning rate
- Dropout
- Optimizers

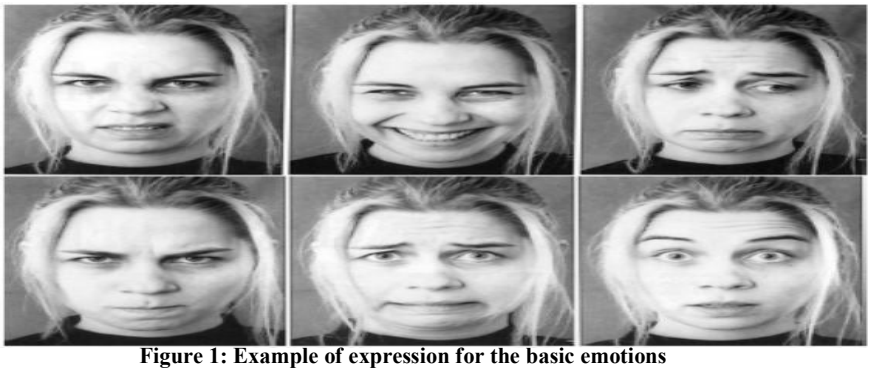

Figure 1: Example of expression for the basic emotions

3. LITERATURE SURVEY

Detailed research on the perception of facial emotions is discussed in which the pro parties of the classifier for study data set, facial emotion perception, are revealed. Visual image properties are analyzed and ranking approaches are addressed in help you better explore emotional perception processes. This research investigated of Potential picture reactions by using multiple types of classifiers. Few classifying algorithms are applied to classify feelings, such as K-Nearest Neighbor, Random Forest. A big neural network emerges and tries to solve data science problems. Different CNN, simulation and facial recognition qualified sets should be tested. Expression in the research field, identification has its meaning. In all areas of study, the perception of them is examined and analyzed. Emotion is detected from expressions with filter banks and CNN, with a strong accuracy rate that deep learning may even be used for the identification of emotion. Facial expression detection is also done with networks through image spectrogramming. All the above listed methods have used the standard function methods chosen from the MFCC, which are used for wave parameters including pitch. There are specific repositories used to define emotion on the face, features taken from facial expression pictures, classifications that distinguish numerous emotional groups. While a long short term memory is used as a facial emotional processing, when the volume of data collection is taken and the form of bottleneck is used. Although identification of speech emotion and optimal outcomes is shown, work on real-time facial expression continues. Facial emotional identification in real time is achieved by classifying RGB stimuli and implementing transition methodologies in which one question is solved and information is applied with another. The facial emotions using secret Markova models and deep learning networks with unweighted average alert (UAR) of around 56.36 percent have known feeling. For the identification of expressions from facial expressions utilizing various classifiers, various image forms and emotions have been examined. Aid computer vector research, local invariant function schooling, and strongly selective facial expression recognition study. Different critical features are analyzed and qualified to identify CNN emotions under which the data from different emotional repositories are collected [18].

4. PROBLEM IDENTIFICATION & OBJECTIVE

Emotions, and particularly those articulated in full by facial expressions, are the key to solving social interactions. David Mortensen shows that from his text and tone we understand only one third of other feelings, and other two-thirds come From peoples expression. Learning to perceive our feelings through device AI takes human-machine experiences to another level.The emotion detection field has evolved from a research project to an industry of billions. This is applicable to many processes:Research communication monitoring driver Impairment Testing the functioning of players with computer games helping medical professionals. We should also concentrate on identifying emotive feelings in images as it shows the human expression through action and gives more knowledge regarding our senses. Emotion identification is often extended to photos, photo and audio files [19].

4.1 Objectives:

Factors in culture, emotion are used. Affective provides artificial intelligence software that enables it to be effective in carrying out tasks previously performed manually by human beings mainly to obtain facial expressivity and vocal information related to particular circumstances in which audiences consent to share this knowledge. Certain applications of Affective include helping children with autism, helping blind people read face, helping robots communicate with people more intelligently and tracking signals of interest when driving to increase driver safety. A patent filed by Snap chat defines a method of gathering crowd data by conducting algorithmic emotion detection on geotagged selfies on users at public events. Emotion was an emotional analysis firm using artificial intelligence for the version of frowning, smiling, and other facial gestures, to predict "facially conveyed attitudes and behavior." Vision offers emotion detection via a real-time API for both online and Smartphone apps. Visage Technologies AB provides emotional calculation to marketing and science analysis as part of their Visage SDK.

Eyries is an emotional reconnaissance firm that integrates its facial analytics and emotion analysis technologies, with the photo developers of embedded product producers, including carmakers, virtual robotics firms, to find the perceived utility of their creative portrait.

Many devices are now present to gather knowledge from feelings exchanged digitally and are primarily used for instructional purposes, as well as to allow players direct influence of their virtual reality, in the form of "like" button clicks and counts of positive and negative text phrases.

4.2 Sub fields of Emotion Recognition

Emotion Recognition can produce the best outcome by integrating multiple elements, including text, audio, video and physiology to interpret emotions using more than one method.

4.3 Emotion Recognition in Text

This is a favorite target of emotion identification research. Compared with other data forms, storing text data is better and easier to compress to maximum output due to the regular repetition of language words and characters. Emotions can be derived from two critical sources of text, written texts and conversations. Many people use it for written texts

FIGURE 2: Emotion Recognition in Text

Vocal signals is to recognize feelings that differ from emotional understanding in text

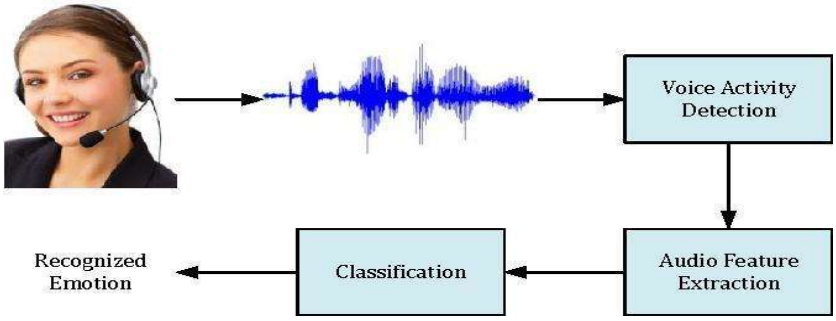

Figure 3:Emotion Recognition in Audio

Video data is an audio, image and text mix. Consequently, as more features are introduced into the identification process, the emotion identification in a video is typically greater than any other area of recognition. Nevertheless, the loading time can be longer.

4.4 FACIAL EMOTION RECOGNITION

Recognition of the emotions is the technique used to detect emotion. Human instinctively recognizes emotions but software cannot, present software has developed which can also recognize emotions. Over time, this has become reliable, and be able to determine emotions just as our brains do.

Governing institutions in the public domain should make effective use of the ability to identify emotions such as shame, anxiety and confusion. Industries often benefit from the visibility of passion to facilitate profit. Disney plans to use facial recognition to analyze the emotional reactions of the viewer with the introduction of a movie. Apple has also released a new app called Animoji, which will imitative your facial expressions using a computer-simulated emoji. This is not so late to say that such technologies will eventually be used in other applications.

5. Convolutional Neural Network

The visual cortex research is directly related to the development of neural networks in convolution. Hubel and Wiesel published in 1968 a report on monkey's visual cortex 'sensitive areas. This research was important in terms of the layout and structure of cells for the striate cortex (primary visual cortex).
Both forms of cells are basic and complex. The simple ones are based on edge forms, whereas the complicated ones are globally invariant and include a wider variety of artifacts. The different cell sets within the cortex will also map the whole visual field by taking advantage of the association between artifacts and shapes in local areas between visibility.Some of Hubel and Wiesel's first applications was dubbed the Neocognitron. Neocognitron was created by Kunihiko Fukushima in 1980 as a neural network model. The model's first layer is made up of basic units, while the second layer units are complex cells. The understanding of the visual cortex's spatial invariance function is the biggest accomplishment of Neocognitrons. In fact, the mapping of output is one by one. Every complex cell maps to a single template.However, the learning process is a big downside for Neocognitron. There was at the point no way of changing weight values for error behaviour across the entire network, such as backpacking. Throughout this period, few applications with back propagation were producedBack propagation conceptually tests the error variance in terms of the unit weight. Each time the weights are adjusted, the gradient increases. Gradient is then used to find weights that reduce the network error on the descent. When backpacking is used, network can automatically change its parameters with any optimizer such as Gradient Descent (GD). This finds a local minimum for a function that takes measures proportional to the gradient negative. Yann LeCun is a pioneer in CNN analysis, as has already been reported. LeCun's handwritten digit classifier (MNIST) was the first true backpack program. This was one of CNN's strongest programs as a vast deal of handwritten reviews was read at the point. LeCun's experiments also contributed to the usage of CNN topology to encourage potential scholars, one of LeNet-5 being one of the most common.
This research stresses that it may be easier to use machine learning methods rather than manually built approaches to solve pattern recognition problems. As it is very a difficult challenge to solve all the various cases of data input naturally, machine learning is the most suitable approach. Therefore, it defines a basic pattern recognition method. This framework comprises of two key modules: a function extractor that converts input data into vectors of small dimensions; and a classifier, which is commonly utilized and functional much of the time.The main components of CNN consist of local transmission areas, weight sharing, turnover process, sub-sampling, drop-out, and stochastic gradient desc.

RESULTS & DISCUSSION

A convolution process is described in mathematics to blend two functions. It process acts as a buffer. Everything that does not matter for the function map is to concentrate only on those basic facts. To execute this operation, two elements are needed

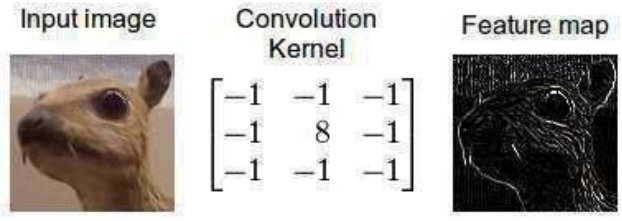

Figure 4: Convolution layer

The input data The convolution filter (kernel)This operation's feature map is the result. It offers the dynamics a schematic account of the convolutionary process. The amount of maps (output channels) helps the neural network to know features. Every channel is autonomous, so each channel needs to know a new feature from a new picture.
Ultimately, the padding form determines the algorithm for the convolution. Another form of padding dismisses the edge of the data as it cannot be checked for further information. The padding on the other side finishes the entry in 0. This is about raising and convoluting parameters.

Figure 5: Convolution layer

6. CONCLUSION & FUTURE SCOPE

In this study, a work was developed in order to identify face emotions by means of profound learning techniques. It's a complicated problem that has been discussed several times before with various approaches. The design development, which focuses on app testing, is one of the Deep Testing ventures, has produced successful success.
While the findings were not positive, they were much different than other methods, such as design engineering. This problem can be solved by deep learning methods, given a number of instances. Though feature design is not needed, image pre-processing increases the accuracy of classification. It decreases the input data.
The application of function development is now part of facial expression recognition applications. The biggest limitation is the absence of an detailed emotional dataset. The Image Net competition uses a list of 14 197 122 images, for example. A bigger dataset may be used to build networks with greater potential for learning functionality. Therefore, classifying feelings using deep learning strategies may be accomplished.th activities will help boost network performance and generalization. The first is to use the entire dataset through optimization. For broader databases, we use batch optimization. Another approach is one-on-one evaluation of feelings. They can track feelings that are difficult to identify. Finally, it seems useful to use a broader dataset to learn.
Optimizers are currently operating on large datasets. But for our project this was real. With a small dataset, attempting the entire dataset may have resulted in a stronger learning feature. In comparison, the application of other optimizers in this work may have been limited. It can be seen on a smoother type loss curve, or by stopping early convergence. Second, it is difficult to determine each emotion because of time constraints. Thus, you may determine the feelings are easier to identify and those are more difficult to classify. In fact, pre-training on each emotion will lead to a stronger learning result. This instruction (transference instruction) may have been given to the network afterwards. That may have lowered the preparation time and may have decreased the expense feature to a higher degree.A bigger sample will also contribute to a higher degree of instruction. Training in a wider input field and increasing network quality for longer. The network can learn more important functionality across a broader training size. When this is not done, this role also needs software engineering. Those details may not be accessible today, though. Several datasets may be a workaround, but it requires a careful process to normalize them.
Finally, it seems to be possible to enhance network output with complete data set for testing, pre-training on each emotion, and use a broad dataset. In future work on this subject, they should be discussed.

REFERENCES

[1]G. Hintin ,Greves and A. Mohemed, 2013 “Emotion Recognition with Deep Recurrent Neural Networks,” in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6645–6649.

[2]K.Weint and C.-W. Huaang, 2017 “Characterizing Types of Convolution in Deep Convolutional Recurrent Neural Networks for Robust Speech Emotion Recognition,”pp.1–19

[3M. Yusuf, M. B. Mustaafa, M. and M. Malekzedeh, 2018 “emotion recognition research: an analysis of research focus,” Int. J. Speech Technol., vol. 0, no. 0, pp. 1–20.

[4]Cavidon ,H. M. Fayak, M. Lich, 2017,“Evaluating deep learning architectures for Emotion Recognition,” Neural Networks, vol. 92, pp. 60–68

[5]A. M. Badshaah, J. Ahmed and S. W. Baek, 2017, “Emotion Recognition from Spectrograms with Deep Convolutional Neural Network,” 2017 Int. Conf. Platf. Technol. Serv., pp. 1– 5.

[6]Routrey, M. Swaen and P Kabisetpathy, 2018 ,“Database, features and classifiers for emotion recognition: a review,” Int. J. Speech Technol., vol. 0, no. 0, p. 0.

[7]K.-Y. Hueng, C.-H. Wiu, T.-H. Yieng, M.-H. Sha and J.-H. Chiu, 2016, “Emotion recognition using autoencoder bottleneck features and LSTM,” in 2016 International Conference on Orange Technologies (ICOT), pp. 1–4.

[8]M. N. Sttilar, M. Leich, R. S. Bolie, and M. Skinter, 2017 “Real time emotion recognition using RGB image classification and transfer learning,” 2017 11th Int. Conf. Signal Process. Commun. Syst., pp. 1–8.

[9]Schaller, G. Rigull, and M. Ling, 2004, “Emotion recognition combining acoustic features and information in a neural network - belief network architecture,” Acoust. Speech, Signal Process., vol. 1, pp. 577–580.

[10]H. Jiaang et al., 2017 ,“Investigation of emotions for identifying depression using various classifiers,” Expression Commun., vol. 90, pp. 39–46.

[11]Q. Maio, M. Diong, Z. Huaeng, and Y. Zhaen, 2014 “Learning Salient Features for Emotion Recognition Using Neural Networks,” IEEE Trans. Multimed., vol. 16, no. 8,.

[12]P. Hairár, R. Burgeet, and M. K. Duitta, 2017, “Facial Emotion Recognition with Deep Learning,” in Signal Processing and Integrated Networks (SPIN), pp. 4–7.

[13]with C++.Packt Publishing Ltd, 2018. Khan, Aftab, Sakhi Rehman, MuhammadWaleed, Ashfaq Khan, Umair Khan, Tariq Kamal, Shehryar Khan Afridi, and Safdar Nawaz Khan Marwat. "Forensic video analysis: passive tracking system for automated Person of Interest (POI) localization." IEEE Access 6 (2018): 43392-43403.

[14] Banerjee, Sandipan, Joel Brogan, JanezKrizaj, Aparna Bharati, Brandon Richard Webster,VitomirStruc, Patrick J. Flynn, and Walter J. Scheirer. "To frontalize or not to frontalize: Do we really need elaborate pre-processing to improve face recognition?." In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 20-29.IEEE, 2018.

[15]Villán, A. F. (2019). Mastering OpenCV with Python: A practical guide covering topics from image processing, augmented reality to deep learning with OpenCV 4 and Python 3.7. Packt Publishing Ltd.

[16] Gollapudi, S. (2019). Learn computer vision using OpenCV: with deep learning CNNs and RNNs. Apress.

[17]Khan, M., Chakraborty, S., Astya, R., &Khepra, S. (2019, October). Face Detection and Recognition Using OpenCV. In 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) (pp. 116-119).IEEE.

[18]agtap, A. M., Kangale, V., Unune, K., &Gosavi, P. (2019, February). A Study of LBPH, Eigenvalue, Fisher face and Haar-like features for Face recognition using OpenCV.

[19]Khan, Sikandar, AdeelAkram, and NighatUsman."Real Time Automatic Attendance System for Face Recognition Using Face API and OpenCV." Wireless Personal Communications 113.1 (2020): 469-480.