International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

A Survey of TensorFlow and CNN in Machine Learning for Recognizing Human Facial Expression

Dipika Gunvant Nasre
Student, MTech in Data Science
G. H. Raisoni University, Amravati

Dr. Rais Abdul Hamid Khan
Assist. Professor, Department of Computer Science & Engineering
G. H. Raisoni University, Amravati

---***---

***Abstract** – Emotion provides important details about human communication. During a conversation, it is customary to utilize facial expressions to express emotions. Furthermore, some interpersonal communication can be accomplished only through facial expressions. Some facial expressions are universal in that they convey the same feeling regardless of culture. If a machine could correctly perceive its user's facial expression, it might be able to assist them more quickly. This study creates a modular multi-channel deep convolutional neural network to improve face emotion recognition. A global average layer is used in the network output to avoid overfitting. The model's generalization ability can be improved by enhancing the dataset before training. Network offers a few advantages over other recognition algorithms. Finally, the trained recognition model is used to build a real-time face emotion recognition system. The results of the experiments will reveal that the system is capable of recognizing facial expressions in videos and photos.*

Keywords - convolutional neural network; network performance; expression recognition; machine learning; CNN; fer2013; TensorFlow;

1. INTRODUCTION

Understanding and recognizing emotions is aided by facial expressions. Even the term "interface" implies the centrality of the face in two-way communication. Reading facial expressions has been demonstrated in studies to significantly influence the interpretation of what is being said and control the flow of conversation. For effective communication, a person's capacity to interpret their emotions is critical. It would be great for machines to be able to understand human emotions in order to create optimum human-machine interfaces (HCI) [6].

The goal of this study is to figure out how computers can accurately detect the emotions of their numerous sensors. This experience was used to create a facial image that might be used to read people's emotions. Humans have seven fundamental emotions. These basic emotions include neutral, angry, disgusted, afraid, pleased, sad, and startled, and they may be identified by a person's facial expression. The neutral, happy, sad, and surprised frontal facial emotions are used in this study to provide an effective technique to recognize these four emotions. The research presented in this article looks into facial expression recognition [8].

This study makes use of a Kaggle dataset containing 48x48-pixel grayscale photos of faces. This study is primarily focused on enhancing the accuracy of earlier models. Few emotions can be included in grayscale photographs of the face pixels of the forehead. The model provided in this study is more accurate and faster than earlier models. TensorFlow is a powerful deep learning module in Python that can be used to run deep learning Convolutional Neural Networks for digits (handwritten) classification, image pre-processing and recognition, sequential models for translation, natural language processing (NLP), and partial differential equation (PDE) based tasks [10].

2. LITRATURE SERVEY

Several domains, such as machine learning, natural language, neurology, and others, contribute to the subject of emotion detection research. In prior research, they looked for universal indications of emotions in face expressions, voice traits, and textual data. Happiness, sadness, disgust, rage, fear, and surprise are some of the static categorizations for emotion. Later efforts improve the visual, sound, and textual data by merging them. The merging of this information yields the most accurate outcome. This fusion can be accomplished in three ways: early, late, or hybrid. Emotional aspects and cooperation between emotional processes and other intellectual procedures are featured in various philosophies.

2.1. Facial Expression Recognition on Video Data with Various Face Poses Using Deep Learning

(978-1-7281-1097-4/20/$31.00 ©2020 IEEE)

Human facial expressions provide nonverbal communication that conveys emotional states; as a result, they play an important part in human social interactions. With the passage of time, facial expression analysis research has progressed to include automatic facial expression identification by computers. Facial expression recognition is important in human-computer interactions, human behavior monitoring, educational practices, psychological research, and social robots. The creation of human face expression identification was carried out in this study utilizing a deep learning method based on Convolutional Neural Networks (CNN) with TensorFlow.

Humans are social creatures who can communicate verbally or nonverbally with other humans. Facial expression is a type of nonverbal communication in which the muscles of the face are used to convey emotional states in humans. It plays a crucial part in human social relationships.

2.2. Facial Emotion Recognition Using Deep Convolutional Neural Network

(978-1-7281-5197-7/20/$31.00 ©2020 IEEE)

Artificial intelligence's rapid development has made a significant contribution to the technological world. Machine learning and deep learning algorithms have achieved considerable success in various applications such as classification systems, recommendation systems, pattern recognition, and so on, as classical algorithms have failed to match human needs in real time. Emotion is crucial in determining a person's thoughts, behavior, and feelings. Using the benefits of deep learning, an emotion recognition system can be constructed, and various applications such as feedback analysis, face unlocking, and so on may be executed with high accuracy. This study's main goal is to develop a Deep Convolutional Neural Network (DCNN) model that can classify five different human face expressions. The manually gathered image dataset is used to train, test, and validate the model.

3. PROPOSED WORK ON FACIAL EXPRESSION RECOGNITION MODEL

Figure 3.1: Types of Facial Expressions

The appearance feature extraction approach extracts the features of the overall face criteria, while the geometric feature extraction method extracts geometric components of the facial structure and motion of the facial muscles. The construction of patches surrounding the principally altering facial muscles when emotions are expressed, and the application of this information to the appearance-based feature extraction technique, have recently increased to contribute to the development of various deep learning algorithms [1].

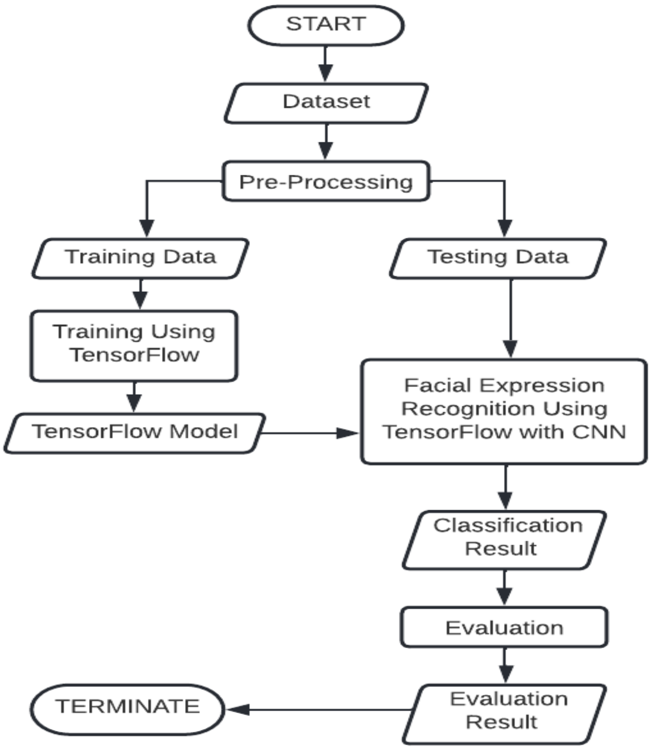

Figure 3.2: Workflow of the FER Model

Figure 3.2 illustrates this concept. The sum of all pixel values at the time of face detection is used to calculate the threshold value. It appears that the threshold value has been surpassed if the sum of the pixel values does not match the sum of pixel values in the FER-2013 data sheet. When the threshold value is exceeded, the algorithm repeats the detection procedure [5].

Pre-processing is useful for partitioning the dataset into training and testing data with the same number of photographs for each expression class so that the model is not biased towards one of the expression classes. This stage also equalizes the image size of each dataset to a preset size, ensuring a smooth training process [3].

4. PROBABLE IMPLIMENTATION

In the experiment, the facial expression dataset is FER-2013. In total, 35,887 pictures are included in the FER-2013 collection. There are three sections to the dataset: a training set (28,709 images), a test set (3,589 images), and a verification set (3,589 pictures). Each image is a grayscale image with a resolution of 48*48 pixels. Anger, disgust, fear, happiness, sorrow, surprise, and neutrality are the seven categories that make up the data collection. The dataset is difficult to recognize since it contains noise (all black photographs, cartoon pictures, non-expression images, and non-expression pictures). The FER-2013 dataset has a 65 percent eye recognition accuracy rating. A representative sample from the FER-2013 dataset. Table 4.1 shows the information and specification of the Dataset FER-2013 [4].

INFORMATION	SPECIFICATION
Size	48*48 (pixel)
Max_Batches (Iteration)	488
Batch	64
Subdivision	16
Color Channel	3 (RGB)
Learning Rate	0.01

Table 4.1: Information about Dataset

Static FER methods rely on static facial traits, but facial expressions are inherently dynamic. Saptio-temporal characteristics are used to capture the dynamics of face expressions. In, a new spatiotemporal feature representation learning for the FER system was proposed, which can endure expression intensity changes. They used typical expression states (such as onset, peak, and offset expressions) to specify facial sequences regardless of expression strength in this research. RGB photos are being replaced with depth images since they expose a person identify, which causes privacy issues. The top directional strengths with respect to the signs are considered using Modified Local Directional Patterns (MLDP). Depth photographs also eliminate the problem of pixel intensities [3].

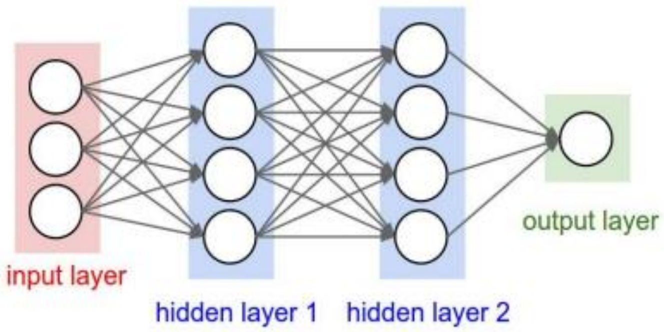

Figure 4.2: The Structure of CNN

Figure 4.2 Shows that the CNN and Special coefficients called action units (AU) were proposed for seven face expressions that were used as characteristics. To distinguish facial expression, a TensorFlow is developed. The extracted areas of facial expression traits are chosen as regions of interest (ROI) (mouth, eyes, and brows). Auto encoder, which can recover data and modify its representation to improve learning efficiency. Sparsity is used in the auto encoder to assist reduce computational complexity. To address the issues of multi-pose facial emotion detection, a pose-based Hierarchical Bayesian theme model is developed (FER). An intermediate face representation is learned using local appearance features and geometry information before an expression is recognized [9].

Emotion detection systems have two drawbacks: translation of facial images might degrade recognition performance, and there are no robust classifiers. To remedy this, features are extracted using stationary wavelet entropy. Face expression recognition employs the local directional ternary pattern (LDTP) as a face descriptor. LDTP uses ternary patterns and directional information to encode information about emotion-related characteristics (eyes, eyebrows, upper-nose, and mouth) [11].

5. CONCLUSION

TensorFlow is utilized to train the model for facial expression recognition in this paper. The FER2013 Dataset is subjected to the Convolutional Neural Network (CNN) Machine Learning technique. The TensorFlow technique works well for classifying real-life data movies with a variety of face postures, lighting, and backgrounds. Suggestions for future work are the subject of the dataset, which must include explicit face expressions that can be differentiated from one another.

REFERENCES

[1] N. B. Kar, K. S. Babu, A. K. Sangaiah, and S. Bakshi, “Face expression recognition system based on ripplet transform type II and least square SVM,” Multimed. Tools Appl., vol. 78, no. 4, pp. 4789– 4812, 2019, doi: 10.1007/s11042-017-5485-0. 2019

[2] N. K. Benamara et al., “Real-Time Emotional Recognition for Sociable Robotics Based on Deep Neural Networks Ensemble,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11486 LNCS, 2019, pp. 171–180. 2019

[3] X uLL, Zhang S M, Zhao J L. Expression recognition algorithm for parallel convolutional neural networks [J]. Journal of Image and Graphics, 2019, 24(02): 0227-0236. 2019

[4] Ravindran, A., Gokul Krishna, G., Sagara, Sarath, S.” A comparative analysis of machine learning algorithms in detecting deceptive behavior in humans using thermal images”, IEEE International Conference on Communication and Signal Processing, ICCSP 2019, art. no. 8697911, pp. 310-314. 2019

[5] G. Cao, Y. Ma, X. Meng, Y. Gao and M. Meng, "Emotion Recognition Based On CNN," 2019 Chinese Control Conference (CCC), Guangzhou, China, 2019, pp. 8627-8630.doi: 10.23919/ChiCC.2019.8866540

[6] Lu Lingling liu, "Human Face Expression Recognition Based on Deep Learning-Deep Convolutional Neural Network", 2019 International Conference on Smart Grid and Electrical Automation (ICSGEA)

[7] S. Suresh, H. T. P Mithun and M. H. Supriya, "Sign Language Recognition System Using Deep Neural Network," 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 2019, pp. 614-618. doi: 10.1109/ICACCS.2019.8728411

[8] "Facial Action Coding System - Paul Ekman Group", Paul Ekman Group, 2020. [Online]. Available: https://www.paulekman.com/facial- action-coding-system/. [Accessed: 20- Oct- 2019].

[9] A. Shah, "Through The Eyes of Gabor Filter", Medium, 2018. [Online]. Available: https://medium.com/@anuj_shah/through-the-eyes-of-gabor-filter-17d1fdb3ac97. [Accessed: 15- Oct- 2019].

[10] "Challenges in Representation Learning: Facial Expression Recognition Challenge", Kaggle.com, 2013. [Online]. Available: https://www.kaggle.com/c/challenges-in-representation-learning- facial-expression-recognition-challenge. [Accessed: 17- Oct- 2019].

[11] A. Chandra, "McCulloch-Pitts Neuron — Mankind's First Mathematical Model Of A Biological Neuron", Medium, 2018. [Online]. Available: https://towardsdatascience.com/mcculloch-pitts- model-5fdf65ac5dd1. [Accessed: 20- Oct- 2019].

[12] "Callbacks - Keras Documentation", Keras.io, 2020. [Online]. Available: https://keras.io/callbacks/. [Accessed: 16- Oct- 2019].

[13] T. Nguyen, A. Fuchs and F. Pernkopf, "Acoustic Scene Classification Using Deep Mixtures of Pre-trained Convolutional Neural Networks," 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA, 2019, pp. 871-875

[14] Lei, X., Pan, H., Huang, X.: A dilated CNN model for image classification. IEEE Access 7, 124087–124095 (2019)

[15] Lu, L., Yi, Y., Huang, F., Wang, K., Wang, Q.: Integrating local CNN and global CNN for script identification in natural scene images. IEEE Access 7, 52669–52679 (2019)