Indian Journal of Computer Graphics and Multimedia (IJCGM)
ISSN: 2582-8592 (Online), Volume-3 Issue-2 August 2023

Convolutional Neural Network Driven Computer Vision Based Facial Emotion Detection and Recognition

Tsega Asresa, Getahun Tigistu, Melaku Bayih

Abstract: ***Computer vision is a sub branch of artificial intelligence (AI) that enables computers and systems to derive substitutive information from digital images and Video. Artificial intelligence plays a significant role in the area of security and surveillance, image processing and machine learning. In computer vision and image processing object detection algorisms are used to detect objects from certain classes of images or video. There is a scope identification of human face emotion Facial emotion recognition is done using computer vision algorism whether the person's emotion is Happy, sad, fear, disgust, neutral and so on. Object detection algorism are used in deep learning used to classify the detected the regions. Facial emotion recognition is an emerging research area for improving human and computer interaction. It plays a crucial role in security, social communication commercial enterprise and law enforcement. In this research project CNN is used for training the data and predicting seven emotions such as anger, happy, sad, disgust, fear neutral and surprise. In this paper the experiment will be conduct using convolutional neural network as classifier, since it is multi class classification relu, softmax (activation function), categorical cross entropy(loss function) dropout max pooling conducted. The researcher tried to train the model by 80/20, 70/30, 90/10 train test split. However 70/30 train test split out performs over the other. The performance of the model is measured by using the epoch 10 and dropout 0.3. Totally the model is performed 93.8% in the training accuracy and it 75% for the testing.***

Key words: ***Artificial Intelligence, Convolutional Neural Network, Computer Vision, Facial Emotion, Object Detection***

I. INTRODUCTION

Facial expression recognition is an emerging research area which improves human and computer interaction. Facial expression can be defined as the facial changes in response to a person's internal emotional state, intentions, and/or social communications[1]. Face Recognition is the task of computer vision and applied in different fields like security and surveillance[2]. Artificial intelligence plays a vital role in the area of computer vision and Machine learning.

Manuscript received on 11 November 2021 | Revised Manuscript received on 02 August 2023 | Manuscript Accepted on 15 August 2023 | Manuscript published on 30 November 2023.

*****Correspondence Author(s)
Mr. Tsega Asresa*, Lecturer, Wolaita Sodo University, School of Informatics, Ethiopia E-mail: tsegavision@gmail.com, ORCID ID: 0000-0001-9876-6685
Mr. Getahun Tigistu, Assistant Professor Arbaminch University, Faculty of Computing and Software Engineering, Ethiopia. E-mail: gettigistu@gmail.com
Mr. Melaku Bayih, Lecturer, Wolaita Sodo University, School of Informatics, Ethiopia. E-mail: melaku.bayih@wsu.edu.et

Computer vision (CV) is a field of computer science and it is a complement of biological perceptions that works enabling computers to see identify and process images. Face emotion detection and recognition technology has been used in computer vision and pattern recognition. It is the most trending in artificial intelligence research. Numerous completely different techniques are developed because of the growing variety of globe applications. For service mechanism, face detection and recognition area unit extraordinarily vital, within which the stress should be placed on security, real-time, high magnitude relation of detection and recognition[3][6][7][8]. Emotion is a mental state associated with the nervous system associated with feeling perception, behavioral reactions and degree of gratification and displeasure. One of the current applications of artificial intelligence (AI) using neural networks is the recognition of faces in images and videos for various applications. Most techniques process visual data and search for general pattern present in human faces in images or videos. Face detection can be used for surveillance purposes by law enforcers as well as in crowd management. Psychological theory states that human emotions can be classified into six archetypal types: surprise, fear, disgust, anger, happiness, and sadness. To display these archetypal types, the muscles in a person's face can change, the tone of voice can be altered, and the energy of a person's voice can increase or decrease. The shape and formation of the lips can also greatly contribute to understanding of speech in a noisy environment. All of these play a role in the process of communicating different feelings. Even if these signals are subtly displayed, humans can recognize these signals by processing information from the ears and eyes. It has been said that individuals focus more attention on projecting their own facial expressions and perceiving others' facial expressions than they do other nonverbal channels and often more than verbal communication as well [4]. In this paper, I present a method for identifying seven emotions such as anger, disgust, neutral fear, happy, sad, and surprise using facial images. Previous research used deep-learning technology to create models of facial expressions based on emotions to identify emotions[5][9][10]. In the current research I present the detail of implementation a deep learning approach that incorporate facial emotion recognition and detection using CNN. Convolutional neural network is widely used in image related researches.

II. RELATED WORKS

In recent years, researchers have created wide progress in developing automatic expression classifiers. Some expression recognition systems classify the face into a collection of prototypes feeling like happiness, unhappiness and anger. Others arrange to acknowledge the individual muscle movements that the face will turn out to supply 6objective description of the face. The simplest glorious psychological framework for describing nearly all of facial movements is that the Facial Action writing To Support this projected research study and deliver the goods the aim, many antecedent done connected analysis literature and Salient ideas (books, articles, and journals) as reviewed. The investigator consistently reviewed connected material therefore to grasp and analyze the data gap between the haves and have not's an answer and to elaborate the scientific ideas associated with this analysis and critically relate, and assess the researches done before this study. Investigator tried to search out a clear-cut analysis gap between existing face feeling recognition and detection system projected over CNN. The reviewed papers are valuated and prioritized supported some parameters; connection to the analysis domain, Quality of Journal, the year of their publication etc. Though Facial feeling recognition and detection is a stimulating and promising work with several applications. Though there have been few works in facial detection. I in short gift them as follows
In order to Support this projected {research project scientific analysis research study and deliver the goods the aim of this research, many antecedents done connected analysis literature and Salient ideas (articles, and journals) as reviewed.

III. METHODOLOGY

In order to do this research we follow the design science methodology. In this section we have focused of the proposed model of a research to classify the emotions as Happy, sad, fear, surprise, disgust, Anger and neutral. The overall process of emotion recognition using Deep CNN Architecture.
CNN is anural network which are sensitive to spatial information. They are cap able of recognizing complex shapes and patterns in a given image. In this project I built a CNN model which is capable of recognizing complex shapes and patterns in a given image and used for a wide variety of image classification algorisms. Classifying seven facial emotions happy, sad, fear, disgust, surprise. CNN is a feed forward multi-layer perception that is used for pattern classification

A. Image Preprocessing

Image sequencing is the process of determining the amount of overlap in succeeding image frames. This is essentially an image correspondence problem in which common scene points in two images are identified and matched. There are many methods available in the literature

B. Datasets

In this paper, we used a dataset FER, which consists of about 35889, well-structured 48 × 48 pixel gray-scale images of faces. The images are processed in such a way that the faces are almost centered and each face occupies about the same amount of space in each image. Each image has to be categorized into one of the seven classes that express different facial emotions. These facial emotions have been categorized as: 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise and 6=Neutral. Figure depicts one example for each facial expression category. In addition to the image class number (a number between 0 and 6), the given images are divided into three different sets which are training, validation, and test sets. There are about 29,000 training images, 4,000 validation images, and 4,000 images for testing. After reading the raw pixel data, we normalized them by subtracting the mean of the training images from each image including those in the validation and test sets. For the purpose of data augmentation, we produced mirrored images by flipping images in the training set horizontally. In order to classify the expressions, mainly we used the features generated by convolution layers using the raw pixel data. As an extra exploration, we developed learning models that concatenate the HOG features with those generated by convolutional layers and give them as input features into Fully Connected (FC) layers.

Out[7]:

	emotion	number
0	Angry	4953
1	Digust	547
2	Fear	5121
3	Happy	8989
4	Sad	6077
5	Surprise	4002
6	Neutral	6198

Figure 1: Sample Data Set

C. Data Preprocessing

The images in the FER 2013 dataset were scanned to extract the portion of the image which contains the face. The detected face is resized and the CNN contains 32 filters on the third layer.

IV. EXPERIMENTAL RESULTS

In order to verify the performance of the Convolutional neural network (CNN) in facial emotion detection and recognition the model uses a convolutional neural network for training the data.
FERC-2013 Database Experiment to contribute over existing efforts, the deep network was first trained using a database made available for Facial Expression Recognition.

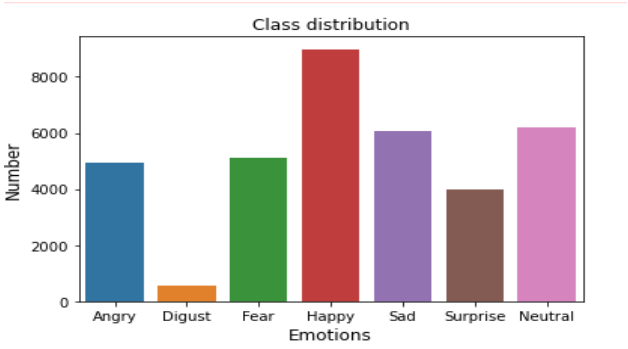

Figure 2: Classification Distribution

As we have illustrated in Figure 2 the class distribution between Datasets the 'Happy' class are more distributed over the other class and the 'Disgust' class are the lower distributed class.

Table 2: Training and Testing Set and Hyper Parameters

Data /	Training set	Testing set	Emotions
48X48,1	28709	3589	'Anger', 'Disgust', 'Fear', 'Happy', 'Sad', 'Surprise', 'Neutral
3x3 Filter	Image channel = RGB		Optimizer= ADAM
3 Convolution Layer			Activation function=Softmax

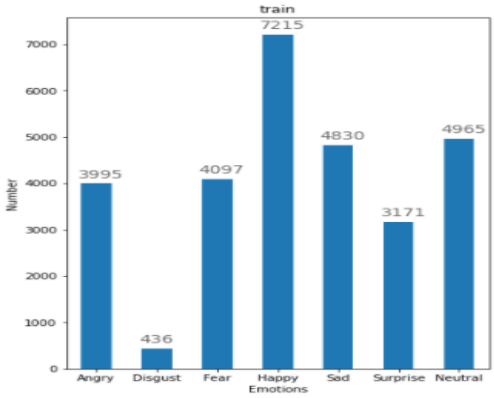

Figure 3: Training Data Distribution

As we have seen from the above Figure 4 the training data distribution between the seven class samples. The "Happy" class contain 7215 whereas the disgust class is 436.

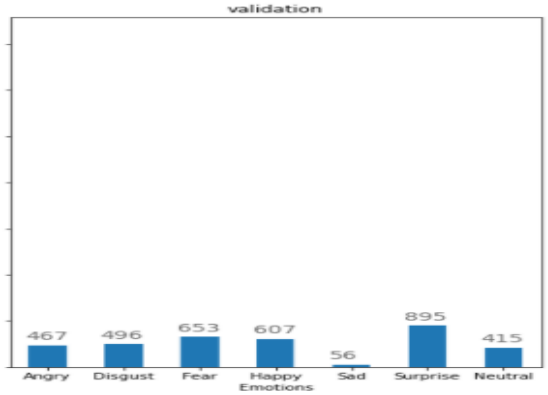

Figure 4: Validation Data Distribution

As we have seen from the above Figure 4 the Validation data distribution between the seven class samples. The "Happy" class contain 607 whereas the disgust class is 496.

V. RESULT AND DISCUSSION

The facial emotion recognition performance evaluation performed using the FER database the model is trained using convolutional neural network by epoch value ten on the FER-2013 dataset, Due to the nature of real-time classification, it is hard to get a definitive metric of our real-time system's accuracy. Our system reliably classified some emotions (the most reliable classification being happy), but struggled when lighting conditions changed, or backgrounds were noisy. Though we tried training on FER-2013 to better reflect real-time data, conditions like lighting in real-time differed from static images, making it hard for our program to transfer over skills learned on static databases.

Table 2: Overall Result of the Model

Accuracy	Loss	Val_voss	Val_ accuracy	Epoch
93.89%	0.1%	1.8%	78.6%	100

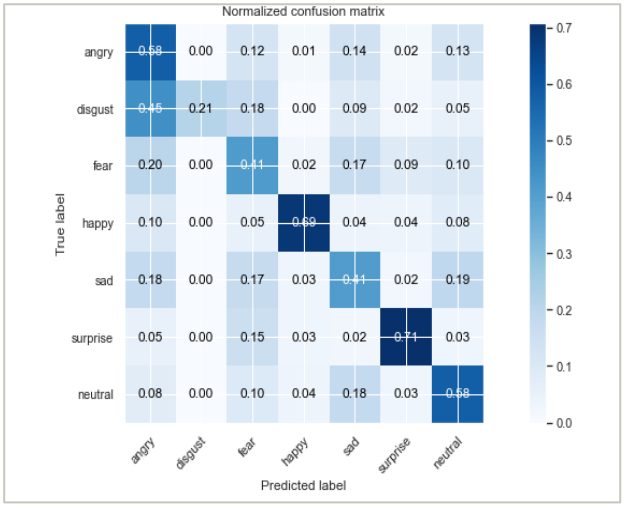

Figure 5: Confusion Matrix for Testing Set

This Figure 6 *shows the normalized confusion matrix true label and the predicted label.* The performance of the model the is very high for surprise, Happy, and disgust. Totally the model performs good accuracy in all.

VI. CONCLUSION

Development of facial emotion recognition system implementing the computer vision and enhancing the advanced feature extraction and classification. The purpose of the study is to classifying the person emotion in order to enhance the security, surveillance and promoting health of the individual's. It focuses on improving the performance of the system and driving more appropriate classification which may be useful in many real world applications. Facial recognition system can be used in digital in security system which can identify a person in any form of expression he presents himself.

In this research project, we have implemented a facial emotion recognition where we used a combination of specific image pre-processing steps and convolutional neural networks. We have used four totally different CNN structures in the datasets. The proposed approach is timely economical and higher accuracy compared to of the state of art approach. Different from previous researches, we combined multiple datasets to get a global picture. Network with large layers performed better over others in the first two experiments. The model performance is done using epoch 10 and dropout 0.3. The overall accuracy of the model is 93.8 in the training set and 75% in the testing set.

DECLARATION STATEMENT

Funding/Grants/ Financial Support	No, I did not receive
Conflicts of Interest/ Competing Interests	No conflict of Interest to the best of my knowledge.
Ethical Approval and Consent to Participate	No, the Article does not require ethical approval and consent to participate with evidence.
Availability of Data and Material/ Data Access Statement	Not relevant,
Authors Contribution	All authors have equal contributions.

REFERENCE

1. A. Mohite, S. Gamare, K. More, and N. Patil, "Deep Learning based Card - Less Atm using Fingerprint and Face Recognition Techniques," Int. Res. J. Eng. Technol., vol. 6, no. 3, pp. 3504–3509, 2019.
2. H. R. Babaei, O. Molalapata, and A. A. Pandor, "Face Recognition Application for Automatic Teller Machines (ATM)," vol. 45, no. Icikm, pp. 211–216, 2012.
3. K. C. Kirana, S. Wibawanto, and H. W. Herwanto, "Facial Emotion Recognition based on Viola-Jones Algorithm in the Learning Environment," pp. 406–410, 2018. https://doi.org/10.1109/EECSI.2018.8752783
4. J. Haddad et al., "3D-CNN for Facial Emotion Recognition in Videos To cite this version : HAL Id : hal-02955528 3D-CNN for Facial Emotion Recognition in Videos," 2020. https://doi.org/10.1007/978-3-030-64559-5_23
5. A. Jaiswal, A. Krishnama Raju, and S. Deb, "Facial emotion detection using deep learning," 2020 Int. Conf. Emerg. Technol. INCET 2020, no. August, 2020, doi: 10.1109/INCET49848.2020.9154121. https://doi.org/10.1109/INCET49848.2020.9154121
6. Sharma, Dr. A., & Mansotra, Dr. V. (2019). Deep Learning based Student Emotion Recognition from Facial Expressions in Classrooms. In International Journal of Engineering and Advanced Technology (Vol. 8, Issue 6, pp. 4691–4699). https://doi.org/10.35940/ijeat.f9170.088619
7. Shanthi*, P., & Nickolas, S. (2019). Merged Local Neighborhood Difference Pattern for Facial Expression Recognition. In International Journal of Innovative Technology and Exploring Engineering (Vol. 9, Issue 2, pp. 4133–4141). https://doi.org/10.35940/ijitee.b7461.129219
8. Behaviour Based Data Dispatcher. (2019). In International Journal of Recent Technology and Engineering (Vol. 8, Issue 4, pp. 12940–12944). \https://doi.org/10.35940/ijrte.d8428.118419 https://doi.org/10.35940/ijrte.D8428.118419
9. Wanjau, S. K., Wambugu, G. M., & Oirere, A. M. (2022). Network Intrusion Detection Systems: A Systematic Literature Review of Hybrid Deep Learning Approaches. In International Journal of Emerging Science and Engineering (Vol. 10, Issue 7, pp. 1–16). https://doi.org/10.35940/ijese.f2530.0610722
10. Das, S., S, S., M, A., & Jayaram, S. (2021). Deep Learning Convolutional Neural Network for Defect Identification and Classification in Woven Fabric. In Indian Journal of Artificial Intelligence and Neural Networking (Vol. 1, Issue 2, pp. 9–13). Lattice Science Publication (LSP). https://doi.org/10.54105/ijainn.b1011.041221

AUTHORS PROFILE

Tsega Asresa is Senior Lecturer and Researcher In school of Informatics Department of computer Science Wolaita Sodo University Ethiopia. He was Received MSc Degree from Arbaminch University Faculty of Computing and software Engineering Department in Computer Science (2022) and His BSc from Wolaita Sodo University in Computer Science (2018). His Current Research interest Is Computer Vision, Deep Learning, Machine Learning, Reinforcement Learning and imaging.

Getahun Tigistu (Assistant Professor) a senior is faculty of Computing and Software Engineering, Arba Minch University, Ethiopia. He was Received MSc Degree AdisAbeba University in Computer Science and His BSc from Debre Birhan University in Computer Science. He is Faculty Dean of Computing and Software Engineering, Arba Minch University, Ethiopia. His Current Research interest Is Image processing, Artificial Neural Network and computer security.

Melaku Bayih a senior Lecturer in Wolaita Sodo University computer Science Department. He is Computer Scientists, graduated second degree from university of Debrecen, Hungary and completed BSc in computer science Haramaya University, Ethiopia. His Current Research interest Is Computer Vision, Image processing, Data science, Machine Learning and Deep Learning.