Emotion Detection Using Deep Learning

Shraddha Belhekar, Priya Patil, Snehal Hulule, Tanvi Ghare and Dhammjyoti Dhawase

June 19, 2023

\
International Conference on Communication and Information Processing
(ICCIP-2023)
Available on: Elsevier-SSRN

EMOTION DETECTION USING DEEP LEARNING

Shraddha Belhekar [1] , Priya Patil[2] , Snehal Hulule[3], Tanvi Ghare[4] ,Prof. Dhammjyoti Dhawase[5]*

aShraddha Belhekar, Nutan College of engineering & Research,Talegaon Dabhade,410507,India
bPriya Patil, Nutan College of engineering & Research,Talegaon Dabhade,410507,India
cSnehal Hulule, Nutan College of engineering & Research,Talegaon Dabhade,410507,India
dTanvi Ghare, Nutan College of engineering & Research,Talegaon Dabhade,410507,India
eProf. Dhammjyoti Dhawase, Nutan College of engineering & Research,Talegaon Dabhade,410507,India

Abstract

An important area of research is the detection and recognition of emotions using visual features extracted from facial expressions. This paper presents a project that focuses on developing an emotion recognition model using the ResNet50 deep learning architecture and training it on the AffectNet dataset. The achieved accuracy of 87% demonstrates the effectiveness of the proposed approach. The applications of this project are wide ranging from human-computer interaction to psychology, medicine, education and crime detection.

The paper also highlights future directions for improving model accuracy and performance. Suggestions include modifying the model layers and exploring larger and more diverse datasets to improve the training process. Additionally, the integration of multimodal data, such as combining facial expressions with voice analysis or physiological signals, holds promise for improving the robustness and accuracy of emotion recognition systems. Cultural and contextual factors that influence emotional expressions should be taken into account to develop more culturally sensitive models. Additionally, optimizing the model for real-time deployment on resource-constrained devices such as smartphones or wearables can expand the practical applications of emotion recognition systems.

By focusing on these future research directions, this paper aims to contribute to the development of emotion detection and recognition systems, which will ultimately lead to more accurate and versatile applications in various fields.

Keywords-: Emotion detection, Facial expression recognition, visual features

.

1. Introduction

Extensive research in digital imaging and image processing has led to the rapid development of deep learning, a versatile field with widespread applications. In the realm of deep learning, visual signals are both input and output, making it relevant to the field of signal processing. Facial emotions, expressed through gestures on the face, play a crucial role in interpersonal communication.

Automatic facial emotion recognition is essential in AI and robotics, with applications including individual identification, access control, video communication, human-computer collaboration, automated surveillance, and cosmetics. This analytical study aims to utilize facial expressions as input to recognize and classify emotions such as neutral, angry, disgusted, fearful, happy, sad, contempt, and surprise.

The objective is not only to develop an automated facial recognition system but also to improve its accuracy compared to existing systems. The human face is an expressive and communicative organ, capable of conveying a wide range of emotions without verbal communication. Facial gesture recognition enables the identification of emotional states from facial images, providing insights into a person's true character.

In this study, several tasks have been undertaken to enhance the accuracy of the facial recognition system. The goal is to create a robust framework that outperforms existing systems. By understanding and leveraging facial expressions, we can tap into a rich source of non-verbal communication, advancing the field of emotion recognition and its practical applications.

2. Motivation

The motivation behind this research lies in the increasing importance of digital imaging and deep learning. Facial expressions are vital in human communication, and automated facial emotion recognition systems have practical applications in human-computer interaction, psychology, medicine, education, and security. The goal is to develop an accurate and robust system that surpasses existing methods, leveraging deep learning and digital imaging to interpret and classify facial expressions. By understanding and recognizing emotions from facial gestures, we aim to enhance various domains and improve human-machine interactions, leading to personalized interfaces, tailored treatments, adaptive education, and enhanced security measures.

3. Related Work

Amir Hossein Farzaneh and Xiaojun Qi Department of Computer Science Utah State University Logan, UT 84322, USA[1] described FER in the wild via Deep attentive Center loss(2021).The proposed model works using a flexible method called Deep Attentive Center Loss (DACL) for Facial Expression Recognition (FER) under wild scenarios. Our hybrid approach takes advantage of a sparse reformulation of center loss to adaptively control the contribution of the deep feature representations in the Deep Metric Learning’s objective function. It achieved an accuracy of 65% on AffectNet dataset.

Andrey V. Savchenko, Lyudmila V. Savchenko, Ilya Makarov[2] described Classifying emotions and engagement in online learning based on a single facial expression recognition neural network.The proposed model is pre-trained on face identification and fine-tuned for facial expression recognition on static images from AffectNet using a specially developed robust optimization technique. The accuracy attained by this model was 63.03% on AffectNet dataset.

Andrey V. Savchenko[3] described Facial expression andattributes recognition based on multi-task learning of lightweight neural networks.The proposed model is characterized by the state-of-the-art emotion classification accuracy on AffectNet dataset and near state-ofthe-art results in age, gender and race recognition for UTK Face dataset. Moreover, it is shown that the usage of our neural network as a feature extractor of facial regions in video frames and concatenation of several statistical functions (mean, max, etc.) leads to 4.5% higher accuracy than the previously known state-ofthe-art single models for AFEW and VGAF datasets from the EmotiW challenges.

Hai-Duoung , Sun-Hee Kim, Guee-Sang Lee Facial Expression using Temporal ensemble of multilevel CNN[4] described The proposed model competitive performance in video based FER, temporal model also outperformed other single video mode but it required arge number of resources. It attained the accuracy of 49.3% on AFEW 7.0 dataset.

Shervin Minaee, Amirali Abdol rashidi[5] described Deep Emotion:-Facial Expression recognition using attentional CNN. The proposed model provided extensive experimental analysis on FER databases. However, the accuracy of the model decreases on occluded images.It attained the accuracy of 99.0% on FER dataset, 98.3% on FERG , 70.02% on FER2013 datasets.

Sunil Kumar, M K Bhuyan, Biplab Ketan Chakraborty[6] described Extraction of Informative regions of a face or FER The proposed model successfully estimated the importance of facial subregions. It attained the accuracy of 98.44% on MUG ,98,51% on JAFFEE, 97.01% on CK+ datasets.

Ebenzer Owusu, Justice Kwame , Appati Percy okae[7] described Robust FER system in Higher passes. The proposed model improves FER performance. Also the 2D phase conversions has been established to handle phase invariant FER problems successfully. It attained an accuracy of 98.90% on Bosphorus , 93.50% on BUD3FE, 97.20% on MMI, 98.20% on CK+ datasets.

Shrey Modi, Muhmmad Hussain Bohara[8] described Facial Expression Recognition Using CNN. This technology will provide a great boom to many things such as the robotics field, which will provide emotions to them and then to the blind community. It attained the accuracy of 73.5% on FER dataset.

Dimas Lima, Bin Li[9] described Facial Expression FER via Res Net-50 The proposed system focuses on FER dataset that achieved good results in multitasking classification.It attained the accuracy of 95.39+/-1.41 on the images that were captured by the canon digital camera.

CUIPING SHI, CONG TAN, LIGUO WANG [10] described Facial expressions recognition Method Based on a Multibranch CrossConnection. In the model, facial expressions were recognized quickly and accurately, which is helpful to realize the intelligent and real-time application of expression recognition .It attained an accuracy of 98.40 on CK+, 88.10 on FER+, 87.34 on RAF, 71.52 on FER2013 datasets.

Zi-Yu Huang1 , Chia-Chin Chiang1 , Jian-Hao Chen2 , Yi-Chian Chen3*, Hsin-Lung Chung1 , Yu-Ping Cai4 & Hsiu-Chuan Hsu2,5[11] described A study on computer vision for facial emotion recognition. It uses SE and ResNet50 algorithms and combine them called as SENet algorithm to recognize the human facial emotions . The authors trained their model of RAFDB and Affectnet dataset. It achieved an accuracy of 56.54% on Affectnet dataset and 65.67% on RAFDB dataset.

4. TECHNOLOGIES WE USED

4.1 Deep Learning

Deep learning is a subset of artificial intelligence. Deep learning is the next evolution of machine learning, allowing models to learn through artificial neural networks that mimic the human brain and analyze data in a structure that mimics that of humans.

Deep learning models do not require the intervention of a human programmer to tell them what to do with the data. It can independently learn from the vast amount of data available.

4.2 Tensorflow

TensorFlow is an stop-to-give up open supply python library getting to know platform. TensorFlow is a comprehensive system for coping with all elements of system learning systems. However, this path focuses on to develop and train machine learning models with help of the use of his particular TensorFlow API. We used Tensorflow 2.0 version for model training. TensorFlow 2.0 is a library that provides developers, researchers, and companies with a rich environment of gear.

4.3 KERAS

Keras is a deep learning API written in Python and running on the TensorFlow machine learning platform. De signed with a focus on enabling rapid experimentation. The key to good research is getting from idea to result s as quickly as possible. Keras is the high-level API for TensorFlow 2, an accessible and productive interface for solving modern deep learning-focused machine learning problems. It provides key abstractions and building blocks for developing and deplo ying machine learning solutions at high iteration rates. Keras allows engineers and researchers to take full adv antage of TensorFlow's scalability and cross-platform capabilities. Keras can run on TPU or large GPU clusters, and Keras models can be exported and run on browsers or mobile devices. The core data structures in Keras are layers and models.

4. System Overview

The project "Emotion Detection using Artificial Intelligence" is based on deep learning, a subdomain of Artificial Intelligence. The ResNet50 model, a type of deep learning ResNet architecture, is utilized. It incorporates skip connections to connect activations of a layer to further layers, forming residual blocks. These blocks are stacked to create ResNets. The implementation is done using Python, specifically TensorFlow 2.0, Keras API, and various libraries such as NumPy, Pandas, Seaborn, Matplotlib, etc. The system utilizes a Python GUI for deployment, with the front view implemented using libraries like tkinter, PIL, and cv2, and the back end utilizing TensorFlow and Keras for model implementation. The ResNet50 model ensures high cohesion, improved development efficiency, and accuracy due to the extensive dataset for classification. The system aims to achieve higher accuracy and efficiency in emotion detection

5.1 Proposed System

Our proposed system utilizes deep learning, specifically the ResNet-50 architecture, for accurate emotion detection from facial images. We acquire a labeled dataset, perform analysis and preprocessing, extract features, and train the model for emotion classification. The system demonstrates potential applications in human-computer interaction, psychology, medicine, and education. Future advancements include exploring larger datasets and optimizing real-time deployment on resource-constrained devices.

The system follows a series of steps to achieve accurate emotion detection.
1. Dataset:
- The AffectNet dataset is utilized, which provides a wide range of facial images with labeled emotions.
This dataset serves as the training and evaluation data for the system.

2. Image Resizing:
   - To ensure consistent processing, the input images are resized to a standard size. This step helps maintaining uniformity across different images.
3. Preprocessing:
   - Prior to further analysis, the images undergo preprocessing. Common preprocessing techniques such as normalization or histogram equalization can be applied to enhance the relevant features for emotion detection.

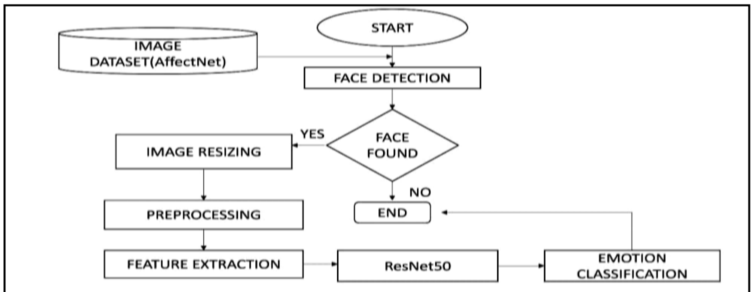

Fig 1 –Proposed Diagram Of Emotion Detection

Fig. 5.3.1

4. Face Detection:
   - A face detection algorithm is employed to detect and localize faces within the input images. This step is crucial as emotions are predominantly expressed through facial gestures.
   - If a face is successfully detected in an image, the system proceeds to the next step. If no face is detected, the image is skipped as it does not contain the necessary information for emotion classification.

5. Feature Extraction:
   - Here, relevant features are extracted from the localized face regions. This can be achieved using techniques such as landmark detection or deep learning-based feature extraction methods.
   - Feature extraction aims to capture discriminative information from the face that is indicative of different emotions.

6. ResNet50:
   - The ResNet50 architecture is employed as the core model for emotion classification. ResNet50 is a deep convolutional neural network that has shown impressive performance in various computer vision tasks.
   - Utilizing a pre-trained ResNet50 model allows the system to leverage the learned representations and hierarchical features to classify emotions accurately.

7. Emotion Classification:
   - The extracted features from the face regions are fed as input to the ResNet50 model for emotion classification.
   - The model predicts the emotions present in the input image based on the learned patterns and representations.
   - The classification results provide information about the dominant emotions expressed in the facial image.

This proposed system ensures a comprehensive pipeline for emotion detection from facial images. By utilizing the AffectNet dataset, resizing images, performing preprocessing, detecting faces, extracting features, and employing the ResNet50 model for classification, the system aims to achieve accurate and reliable emotion recognition.

5.2 Algorithm for ResNet50 Model

The algorithm for implementing emotion detection using deep learning with the ResNet-50 model is outline as follows:

1. Start the process.
2. Acquire the dataset containing facial images for emotion detection.
3. Perform dataset analysis, including image resizing to ensure compatibility.
   3.1 Resize the images to a standardized size for input to the model.
4. Extract features from the resized images using the ResNet-50 architecture.
   4.1 Utilize the ResNet-50 deep learning model to extract high-level features from facial expressions.
5. Train the model using the labeled dataset, optimizing its parameters for accurate emotion classification.
6. Test the module by inputting new facial images and evaluating the model's performance.
7. Provide the input image to the trained model for emotion classification.
8. Classify the emotion based on the output of the model.
9. Exit the process.

This algorithm outlines the sequential steps involved in implementing emotion detection using the ResNet-50 model. By following these steps, the system can effectively acquire and analyze the dataset, extract relevant features, classify emotions, and provide accurate results for facial images.

5. Testing

6.1 Model Testing :

We trained our recommendation model on AffectNet dataset following table shows the results and achieved training accuracy of 87.50% and validation accuracy of 87.48%.

Measure	Epochs	Model – Training Accuracy	Model – Validation Accuracy
Accuracy (%)	2	87.43	87.47
Accuracy (%)	10	87.50	87.49
Accuracy (%)	15	87.50	87.50

Table. 6.1.1

6.2Accuracy graph:

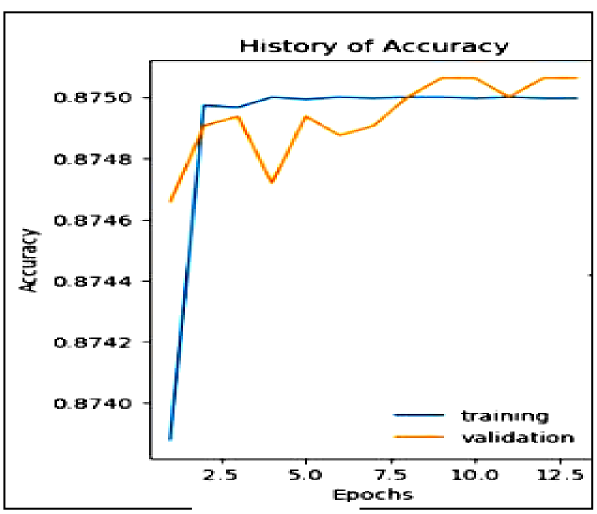

Fig. 6.2.1

7. Result and Discussion

Main Page: On Main Page ,it creates a Tkinter-based GUI application with two buttons for emotion detection and exiting the application.

This Page also uses Tkinter to build a GUI application that includes a text box for user input and a button to load a model and predict the output based on the entered input .It allows users to select an image, displays it within the GUI, and provides options to select an image and detect emotions from the image that we have taken input.

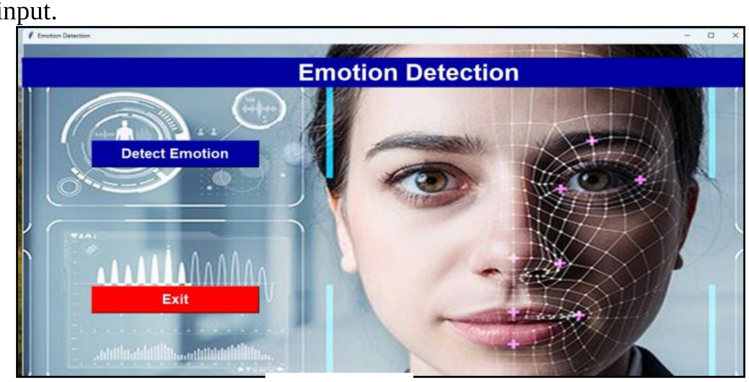

Fig. 7.1.1

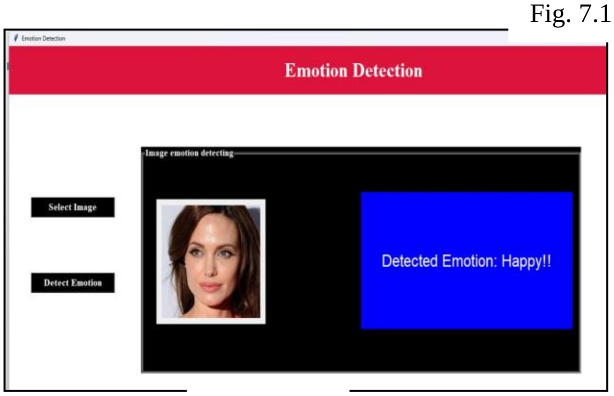

Fig. 7.1.2

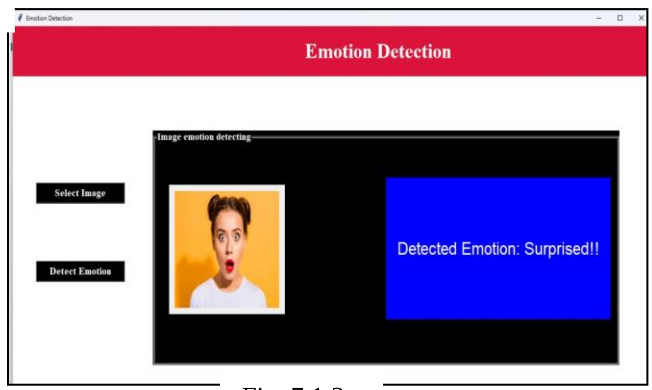

Fig. 7.1.3

In summary, this GUI application allows the user to select an image file, display the selected image, and display the emotion detected by the proposed model. This application is focused on emotion detection from images.

8. Conclusion

We successfully implemented the image recognition the ResNet50 algorithm which is a pre-trained algorithm on AffectNet dataset. We attained an accuracy of 87.48% which is found to be much higher than previously trained models in the last years. The physiological characteristics of the human face that are relevant to various expressions such as happiness, sadness, fear, anger, surprise, and disgust were found to be differentiated from each other in our proposed model which is associated with geometrical structures that are restored as the recognition system's base matching template. This system's behavioural aspect relates the attitude behind various expressions as a property base. This research work contributes a resilient face emotion model based on the mapping of behavioural characteristics with physiological biometric characteristics.

8.1 Future Scope

The accuracy of this model can be increased by working on the different layers of the ResNet architechture. This model can further be utilized for Video based Emotion recognition. Also many new architechtures like ResNet101, ResNet52 can be trained on the AffectNet dataset. They can give better accuracy in terms of performance

9. References

1. Farzaneh, Amir Hossein & Qi, Xiaojun. (2021). Facial Expression Recognition in the Wild via Deep Attentive Center Loss. 2401-2410. 10.1109/WACV48630.2021.00245.

2. A. V. Savchenko, L. V. Savchenko and I. Makarov, "Classifying Emotions and Engagement in Online Learning Based on a Single Facial Expression Recognition Neural Network," in IEEE Transactions on Affective Computing, vol. 13, no. 4, pp. 2132-2143, 1 Oct.-Dec. 2022, doi: 10.1109/TAFFC.2022.3188390

3. A. V. Savchenko, "Facial expression and attributes recognition based on multi-task learning of lightweight neural networks," 2021 IEEE 19th International Symposium on Intelligent Systems and Informatics (SISY), 2021, pp. 119-124, doi: 10.1109/SISY52375.2021.9582508

4. H. -D. Nguyen, S. -H. Kim, G. -S. Lee, H. -J. Yang, I. -S. Na and S. -H. Kim, "Facial Expression Recognition Using a Temporal Ensemble of Multi-Level Convolutional Neural Networks," in IEEE Transactions on Affective Computing, vol. 13, no. 1, pp. 226-237, 1 Jan.-March 2022, doi: 10.1109/TAFFC.2019.2946540

5. Minaee, Shervin, Mehdi Minaei, and Amirali Abdolrashidi. 2021. "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network" Sensors 21, no. 9: 3046. https://doi.org/10.3390/s21093046

6. Sunil Kumar, M K Bhuyan, Biplab Ketan Chakraborty “Extraction of informative regions of a face for special expression recoginition” https://doi.org/10.1049/iet-cvi.2015.0273

7. Owusu, E., Appati, J.K. & Okae, P. Robust facial expression recognition system in higher poses. Vis. Comput. Ind. Biomed. Art 5, 14 (2022). https://doi.org/10.1186/s42492-022-00109-0

8. M. H. Siddiqi, R. Ali, A. M. Khan, Y. -T. Park and S. Lee, "Human Facial Expression Recognition Using Stepwise Linear Discriminant Analysis and Hidden Conditional Random Fields," in IEEE Transactions on Image Processing, vol. 24, no. 4, pp. 1386-1398, April 2015, doi: 10.1109/TIP.2015.2405346.

9. Bin Li, Dimas Lima, “Facial expression recognition via ResNet-50”,International Journal of Cognitive Computing in Engineering, Volume 2,2021Pages 57-64,ISSN 2666-3074,https://doi.org/10.1016/j.ijcce.2021.02.002.(https://www.sciencedirect.com/science/article/pii/S26 66307421000073)

10. C. Shi, C. Tan, and L. Wang, "A Facial Expression Recognition Method Based on a Multibranch Cross-Connection Convolutional Neural Network," in IEEE Access, vol. 9, pp. 39255-39274, 2021, doi:10.1109/ACCESS.2021.3063493

11. Zi-Yu Huang1 , Chia-Chin Chiang1 , Jian-Hao Chen2 , Yi-Chian Chen3*, Hsin-Lung Chung1 , Yu-Ping Cai4 & Hsiu-Chuan Hsu2, in National Library of Medicine doi: 10.1038/s41598-023-35446-4.