Deep Learning and Neural Network Implementation for Emotions Recognition

Renas R. Asaad

Department of Computer Science

Nawroz University

Kurdistan Region of Iraq

Abstract- Recently, we humans integrate into the world of our smart phones and our portable electronic devices to the point that that world can numb us in one way or another and separate us from the real world, so that we and the generations that come after us are fully adapted to dealing with huge amounts of digital information, so they are ready to be absorbed faster And ready to deal with it more efficiently than previous generations, but during this process of rapid adaptation and adapting to the digital age gradually begins to lose one thing, this thing is what the machine has not given it yet and it is human emotions. A key step in the humanization of robotics is the ability to classify the emotion of the human operator. In this paper we present the design of an artificially intelligent system capable of emotion recognition trough facial expressions. Three Promising neural network architectures are customized, trained. and subjected to various classification tasks, after which the best performing network is further optimized. The applicability of the final model is portrayed in a live video application that can instantaneously return the emotion of the user. Technology experts have found that we can create empathy through technology as well, which will consequently lead to what is known as "emotional intelligence." Instead of seeking about digital communication that is losing us to real communication, experts have found that we can employ technology in favor of that type of communication to restore Soul for social and emotional relationships that technology has lost its advantages for many years.

Keywords- Emotion Detection, Emotion Recognition, Deep Learning, Machine Learning, Neural Network.

I. INTRODUCTION

We now have a large number of advanced devices that are able to detect many of our biological functions. You may now wear one of them around your wrist that senses the pulse, temperature, level of sweating and the number of steps that you walked today, but what if we used these sensors to sense our emotions and help us to have Be able to express and recognize our emotions in a way that facilitates dealing with others, for example, why don't we use technology to read our mood for example? When we are exhausted and tired, there is a device that reads this and others know about it without the person having to speak [1].

Emotions are very important in our lives, they are the most influencing things on our memories, and on our decisions that we make on a daily basis, and since we now have devices that help us know the way in a strange country, or to translate languages instantly, and devices that predict when we arrive at home in order to save us energy consumption From heat and electricity, we will also have devices that help us better sense our emotions than we are now. Imagine, for example, that if you had a device that senses your bad mood at the beginning of the day in your car, it suggests a different and less crowded way than the one that you decided to take, so the device avoids you more troubles by simply sensing your mood on that day, or a device that suggests a specific radio station for you to listen to , Or a device that reads your bad mood and announces it to your co-workers, so they avoid getting into trouble with you so that it doesn't get any worse, and so humans are better able to communicate with each other emotionally smarter[2].

These devices can also help those who suffer from depression to notify those around them of their psychological state through devices that sense their mood and give notifications to their loved ones and friends about their condition, here we are talking about a software ability to adapt to the emotions of people, and have the ability to detect and sense those different emotions, In order for smart devices to have "soft" capabilities and also have capabilities that outweigh human intelligence, the mobile device can help you find your way in a strange country, and it can know if you are sad or angry and helps you and helps those around you to overcome that[3].

Figure.1: Emotion Recognition

II. Literature Review:

System development that is able to recognize emotions through facial expressions, previous research on the way humans reveal emotions as well as the theory of automatic image categorization is reviewed. In this study the role of interpreting facial expression in emotion recognition will be discussed[3].

Humans interact socially with the help of emotions, which are considered as a universal language. These emotions surpass cultural diversities and ethnicity. Facial expressions are responsible for conveying the information, which was difficult to perceive. It gives the mental state of a person that directly relates to his intentions or the physical efforts that he must be applying for performing tasks. As a result, automatic recognition of emotion with the help of high-quality sensors is quite useful in a variety of areas such as image processing, cybersecurity, robotics, psychological studies, and virtual reality applications to name a few. Efforts in this area are being made to gather information of high-quality to meet the demands of the system so that it can read, process and simulate human emotions. Geometric and machine learning based algorithms for an effective recognition are being refined, emphasizing emotion recognition in real-time and not just ideal laboratory conditions. Hence, building a system that is capable of both face detection and emotion recognition has been a crucial area of research[3].

It is a well-established fact that human beings are responsible for the depiction of six basic emotions, namely happiness, anger, surprise, sadness, fear, and disgust [1]. These primary emotions form the primary classification of the study of human emotional responses. Apart from these basic emotions, several other emotions have been considered for research. These include contempt, envy, pain, drowsiness and various micro expressions. Facial expression is seen as the primary mode of recognition of human emotion. It works on facial motion and the deformations of

facial features to classify them into emotion categories. This classification is based on visual information and may not be the sole indicator of emotion. Other factors also contribute to the recognition of a person's emotional state such as voice, body language, gestures or even the direction of the gaze. Emotion recognition, therefore, demands a more precise knowledge of all these factors together with contextual information to convey more accurate results [1].

Facial emotional recognition is essentially pattern recognition and involves finding regularities in the set of data being analyzed. Using these regularities, faces, as well as emotions, can be recognized. Various techniques are followed to carry out the tasks that widely fall into two classes, method of parameterization, and the method of recognition. The method of parameterization includes segmentation, assigning binary labels to each pixel and detection where a boundary box is obtained when the face is located in the given data [2]. FACS (Facial Action Coding System) is an example of this method, where all facial emotions are considered and described by the contraction of facial muscles being considered as AUs (Action Units) [3]. These Geometric Feature-based techniques give importance to the structural shape of facial components such as nose, mouth, and eyes. he other method is appearance-based where attributes such as intensities, pixel values, and histograms are considered. After exhaustive training is done with the help of prelabeled datasets, machine learning techniques are applied to detect emotions [1].

III. Objectives:
- The method depends on computer and automation
- Detects high pressure, emotional volatility and internal conflict
- Estimates of deception through interview subjects via speech and / or text analyzes
- Uses multiple analyzes with support for algorithms

IV. Image Classification Techniques:
- Unsupervised Classification:

In unsupervised classification, it first groups pixels into "clusters" based on their properties. Then, you classify each cluster with a land cover class[4].

- Supervised Classification:
  In supervised classification, you select representative samples for each land cover class. The software then uses these "training sites" and applies them to the entire image[4].

V. FEELINGS RECOGNIZE THROUGH THE VOICE

The machine can identify a person by the tone of his voice, but can he know how he feels from the tone of his voice himself? Can she determine his feelings and his psychological and mental state from his way of speaking ?, and can she determine human feelings if she has enough information in her database about facial expressions and changes in heartbeat in the case of feeling different feelings, while you are reading these words now many programmers and engineers work To train different electronic devices to recognize facial expressions and how to analyze each expression individually[5].

One of the architects of what is known as" emotional programming, one of the types of programming for training a machine in identifying and analyzing human emotions, said in his talk at TED talks in Berlin that the machine must learn empathy, and for this he and a team work. Huge emotional programming engineers collect data from thousands of volunteers to conduct experiments that allow them to recognize their facial expressions if they are exposed to a specific situation or recognize their emotions towards each other so that the machine has a modern database based on real emotions and facial expressions[6].

Perhaps emotional technology helps you avoid many embarrassing situations, or helps you to be more intelligent in your field of work, or helps you to maintain your emotional relationship with your wife more understandably, this also means that this technology will change our interaction with smart devices, and from the way Our dependence on it and our addiction to it. Interactive programming engineers may think that this will improve the quality of social relationships, but at the same time it will support the idea that came out of the womb of Silicon Valley in the first place, which is the idea of individuality[9].

Imagine that there are electronic devices dedicated to reading your emotions and mood on a daily or instantaneous basis, devices that are able to make your mood better, and devices that help you find the right person for you based on your emotional and mental state and the characteristics that suit your psychological state, this does not seem very far from reality, as we have Now websites and apps that are primarily concerned with finding the best people for you based on the information that you give yourself for the algorithms such as the information on your personal account and your personal image like the "tinder" or "crown" application, the applications that turned the sacred link into just a digital game[6].
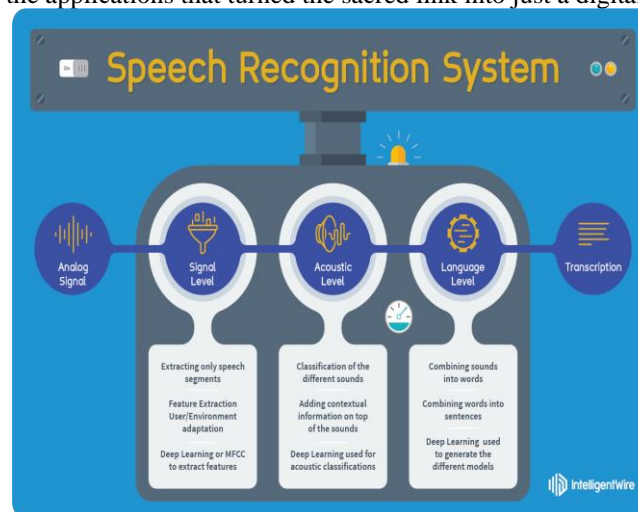

Figure.2: Speech Recognition

VI. IMPLEMENTATION:

This study aims to classify the emotion on a person's face into one of **seven categories**, using deep convolutional neural networks. This repository is an implementation of this research paper. The model is trained on the **FER-2013** dataset[8]. This dataset consists of 35887 grayscale, 48x48 sized face images with **seven emotions** - angry, disgusted, fearful, happy, neutral, sad and surprised[7].

- This implementation by default detects emotions on all faces in the webcam feed. With a simple 4-layer CNN, the test accuracy reached 63.2% in 50 epochs [1].
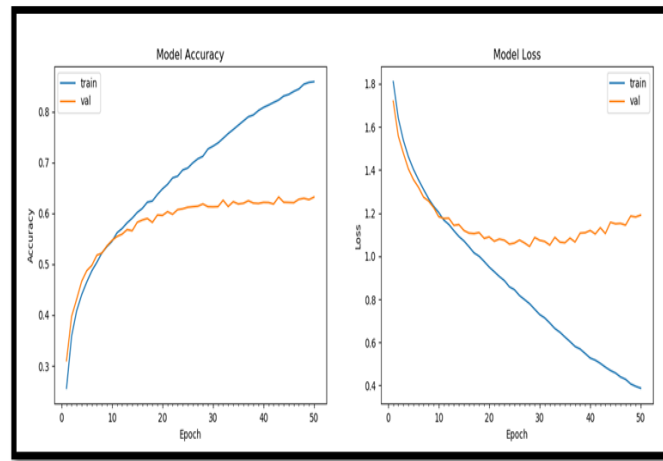
Figure.3: Implementation of Model

a) DATA PREPARATION

- The original FER2013 dataset in Kaggle is available as a single csv file. I had converted into a dataset of images in the PNG format for training/testing and provided this as the dataset in the previous section[10].
- In case you are looking to experiment with new datasets, you may have to deal with data in the csv format. I have provided the code I wrote for data preprocessing in the dataset_prepare.py file which can be used for reference[10].

Fig.4: Data Preparation

b) ALGORITHM

- First, the haar cascade method is used to detect faces in each frame of the webcam feed.
- The region of image containing the face is resized to 48x48 and is passed as input to the CNN.
- The network outputs a list of softmax scores for the seven classes of emotions.
- The emotion with maximum score is displayed on the screen [1].

c) APPLICATIONS

- Criminal investigations
- Intelligence gathering
- Examination of employment
- Criminal investigations
- Mental Health Assessment
- Identify social network users
- Supplement for polygraph testing

VII. CONCLUSION:

In this study, explained that the field of emotion recognition through advanced smart technology, neural network technology and deep learning will help a human being on another dimension of studies through applications and devices that recognize or deal with a person through emotions or recognize a person's emotion through an image or sound or even through a live broadcast. This will assist the medical, military or academic aspect of a technological revolution in the coming years.

REFERENCES

[1] Asaad, R., & Ali, R. (2019). Back Propagation Neural Network(BPNN) and Sigmoid Activation Function in Multi-Layer Networks. Academic Journal Of Nawroz University, 8(4), 216. doi: 10.25007/ajnu.v8n4a464.
[2] Zhang C., Zhang Z. A, (2010). Survey of Recent Advances in Face Detection. Microsoft Corporation; Albuquerque, NM, USA. TechReport, No. MSR-TR-2010-66.

[3] 3. Ekman P., Friesen W., Hager J. (2002). Facial Action Coding System: The Manual on CD ROM. A Human Face; Salt Lake City, UT, USA.

[4] Li, M., Zang, S., Zhang, B., Li, S., & Wu, C. (2014). A review of remote sensing image classification techniques: The role of spatial-contextual information. European Journal of Remote Sensing, 47(1), 389-411.

[5] Kwon, O. W., Chan, K., Hao, J., & Lee, T. W. (2003). Emotion recognition by speech signals. In Eighth European Conference on Speech Communication and Technology.

[6] Schuller, B., Rigoll, G., & Lang, M. (2003, April). Hidden Markov model-based speech emotion recognition. In 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). (Vol. 2, pp. II-1). IEEE.

[7] El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition, 44(3), 572-587.

[8] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001). Emotion recognition in human-computer interaction. IEEE Signal processing magazine, 18(1), 32-80.

[9] Nwe, T. L., Foo, S. W., & De Silva, L. C. (2003). Speech emotion recognition using hidden Markov models. Speech communication, 41(4), 603-623.

[10] Busso, C., Lee, S., & Narayanan, S. (2009). Analysis of emotionally salient aspects of fundamental frequency for emotion detection. IEEE transactions on audio, speech, and language processing, 17(4), 582-596.

[11] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. Neurocomputing, 187, 27-48.