Proceedings of 7th International Conference on Natural Sciences and Technology (ICNST'23)
March 3 - 4, 2023, Asian University for Women, Chittagong, Bangladesh

Facial Expression Driven Medical Rehabilitation Robot

Chowdhury Mohammad Masum Refat [a], Norsinnira Zainul Azlan [b], Anis Nurashikin Nordin [c].

[a] *Ph.D. Researcher, Dept. Of Mechanical Engineering, Osaka University, Japan.*
[b] *Assistant Professor, Dept. Of Mechatronics Engineering International Islamic University Malaysia.*
[c] *Professor, Department of Electronics and Computer Engineering, International Islamic University Malaysia.*
Email: masumrefat3@gmail.com [a], sinnira@iium.edu.my[b], anisnn@iium.edu.my[c]

Abstract

After a stroke, most patients suffer from various physical movement disorders. Recovery of paralyzed parts can be facilitated through proper rehabilitation exercises. However, many paralyzed patients rely on others for assistance with daily activities and rehabilitation therapy. Consequently, they often do not perform rehabilitation exercises regularly, leading to prolonged recovery periods. In recent years, various rehabilitation robots have become commercially available for home-based exercise environments. However, most of these robots are unable to interpret patients' feedback during rehabilitation exercises. In this paper, a deep learning-based Facial Expression Recognition (FER) system is developed to control a rehabilitation robot. The model is trained using the JAFFE and FER2013 facial expression datasets, and a prototype rehabilitation robot is used to evaluate the system's performance. Experimental hardware tests show that the model achieves an accuracy of 97.01% on the JAFFE dataset and 87% on the FER2013 dataset. In future work, the proposed facial expression model will be tested with an upper limb rehabilitation system to further validate its effectiveness.

Keywords: Facial expression recognition (FER); Convolution Neural Networks (CNN); Facial Expression Classification; Medical Rehabilitation Robot.

1.0 Introduction

Computers and Robots have become a leading aspect of our lives, and their behavior will only continue to grow, giving way to modern technologies. However, numerous barriers exist before robots can communicate fluidly with humans daily. Assume a robot can perform as a psychoanalyst. This robot can identify the patient's facial sensations and express a proper response. Reading facial emotions is a complicated process, but one that humans are perfect. (Subiksha, 2018) There are some technical difficulties in improving facial expression recognition methods. Various techniques have been recommended for facial expression identification for several years. However, much progress has been made toward recognizing facial expressions based on static face images without regard to partial occlusions. Facial appearance is a human response, and it often shows the properties of facial characteristics' deformation, the relative timing of facial movements, and their temporal evolution.

Lyons et al. developed a facial expression analysis method based on conventional neural networks, but the usual neural network accuracy is not good. The deep learning model can make more accurate predictions in this situation than the traditional neural network model. More importantly, transient fluctuations often provide essential data about what we try to infer and understand human emotions that link to facial expressions. Though efforts for facial expression recognition have renewed most newly, the existing methods have typically concentrated on one part of the issues or the other. In the modern world, facial expression is used in different applications such as taking pictures, face detectors device, robotics, and gaming which use an avatar to measure customer interest. However, the Facial action coding system, Feature method, Biorthogonal Wavelet Entropy, Local Binary Pattern techniques, and recently traditional neural networks have been used for recognizing the expression. However, previous approaches have some difficulty solving some classification problems why deep learning comes out to solve these kinds of issues (Yazhini and Loganathan, 2019).

The rehabilitation robot is an automatically operated machine designed to improve physical movements in persons with impaired physical functioning. There are two principal characteristics of rehabilitation robots: an assistive robot that substitutes for lost limb movements. Another rehabilitation robot is a therapy robot, sometimes called a rehabilitator—the assistive robot used as a substitute for lost limb movements. An assistive robot, for example, is the Manus ARM (assistive robotics manipulator), which is a wheelchair mounter robotics arm that is controlled using a chine switch or other input device. (Subiksha, 2018) The process is called telemanipulation and is similar to an astronaut controlling the spacecraft's robot arm from inside the cockpit. Teleoperated assistive robots, for example, are powered wheelchairs. Therapy robots are devices or machines for rehabilitation therapists that enable patients to complete practice movements aided by the robot. A rehabilitation therapy robot is also called a rehabilitator.

Neuroscience research has shown that the brain and spinal code can adapt, even after injury, using practiced movements. Therapy rehabilitation robots are machines or tools for rehabilitation therapists, allowing the patients to practice activities aided by the robot (Masum Refat and Zainul Azlan, 2021). The rehabilitation robot's primary limitation is not understanding the patient's feedback when running the rehabilitation process. Deep Neural Networks (DNN) are models animated by the human brain; mostly, it is the capacity to separate structures(patterns) from raw information. (Irfan and Hameed, 2017) Deep-learning models work enormous changes from essential information input to find portrayals progressively unique to such information. The performed changes are mixes of linear and nonlinear tasks. These changes are utilized to speak to the statement at various levels of deliberation. In this paper, we have developed a deep-learning facial expression model for a medical rehabilitation robot.

This paper is organized as follows: Section 2 idea about the convolution neuron network (CNN) for facial expression recognition, Section- 3 describes the architecture & methodology shortly, and Section-4 reports the analysis and discussion of the results. Lastly, Section-5 concludes this paper.

2.0 Architecture and Methods

Different architecture and methods have been developed for facial expression recognition and classification. However, the most effective FER methods are convolution neural networks. This work uses a convolution neural network for facial expression recognition and classification.

2.1 Convolution Neural Network:

The convolution neural network (CNN) is the most popular for facial expression recognition (FER) and classification. Various types of FER are developed by using convolution neural network algorithms. [] CNN is capable of predicting the local facial expression object detection, and it can do the multitask deep neuron network (DNN) [].CNN is the deep learning method with local weight sharing, state-an-art translation, scaling distortion, and deformation. Convolution neural network (CNN) has been applied to many fields, such as speech analysis and image recognition. CNN is constructed by three main processing layers convolution layers, a pooling layer, and a fully connected layer. Figure 1 shows the architecture of the convolution neural networks. (Refat and Azlan, 2019) Convolution neural networks (CNN) initially focus on input comprising images. CNN model architecture is developed for maintaining the specific type of data. The critical difference between CNN neurons and other neurons is that the layer of CNN comprises the neuron organized into three dimensions: input spatial dimensionality (width and height) and depth. In the tradition, an Artificial neural network (ANN) does not mention the depth of the total number of layers, but the third dimension is the activation volumes. Unlike the standard ANN, the neuron in any given layer will be connected to the small region of the preceding layers. (Deopa et al., 2019)

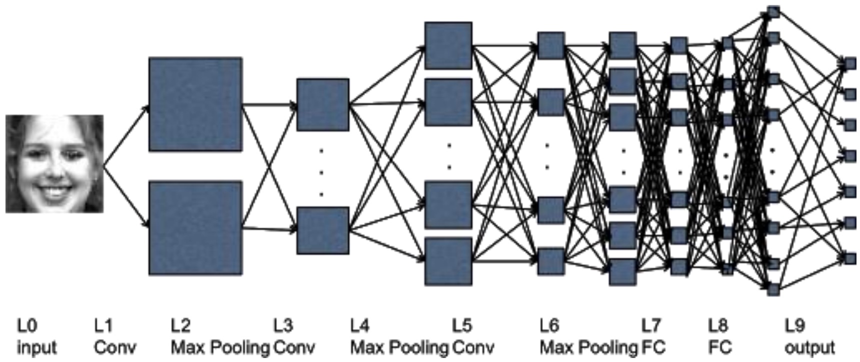


Fig. 1: Architecture of the convolution neural network (Deopa et al., 2019)

2.2 Database

To train our model, we used two datasets: Japanese Female facial expressions (JAFFE) and FER2013. The JAFFE dataset has 213 facial expression images with seven different emotions: simple, sad, degusting, surprise, angry, and neutral also, all the image is grayscale image fig.2. JAFFE all image size is 256 x 256. We divided the dataset into three sections with a 70:15:15 ratio. 70% dataset is used to train the model, 15% for cross-validation, and 15% for testing the model for how well the model is predicted.

The Fer2013 dataset is among the most popular for facial expression recognition and classification because it has around 35,685 examples. Image size 48x48 pixel grayscale images of faces shown in fig.3. Images are labeled based on emotions such as (neutral, happiness, sadness, disgust, anger, surprise, and fear.

Fig. 2: Sample of JAFFE (Refat and Azlan, 2019)

Fig. 3: Sample of Fer2013 Dataset (Deopa et al., 2019)

3.0 Methodology

This work uses computer vision methods for facial expression recognition and classification. Two different datasets, Fer2013 and JAFFE, have been used for training. Moreover, the convolution neural network (CNN) deep learning algorithms are used for training the model. Before introducing the model, the dataset normalized both datasets (48 x 48) pixel grayscale image of the face. This research modifies the computer vision methods and adds rehabilitation robot algorithms, as shown in the figure. 4. These proposed methods used the camera to detect new face data classification and recognition. Initially, this research used four emotions (neutral, happy, sad, and disgust) for operating a rehabilitation robot. It proposed that when patient feedback was neutral and Happy, the rehabilitation robot worked as usual for these four emotions. Besides, when the patient started feeding terribly (sad) and uncomfortably (disgusted), the rehabilitation robot stopped immediately and slowed down the speed. In contrast, the proposed prototype (robotic hand) used for testing the proposed algorithm shows fig.6 when neutral, happy, sad, and disgusted; the systems show peace, thumbs up, and fast-moving all figures and hand rotate 180 degrees, respectively. Shortly, this work proposed that facial expression can improve rehabilitation robot performance.

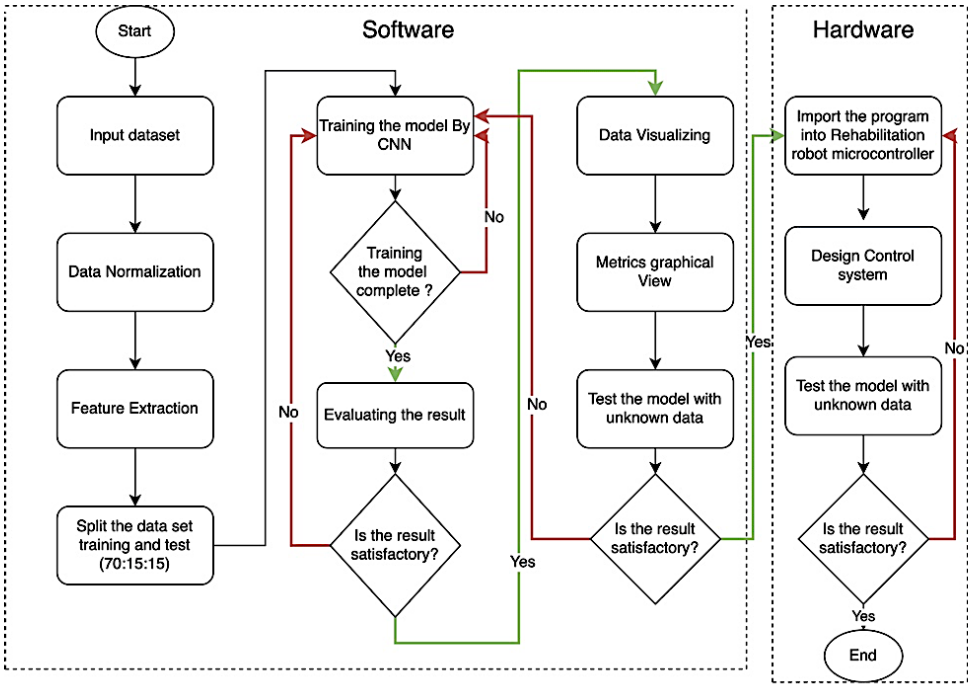
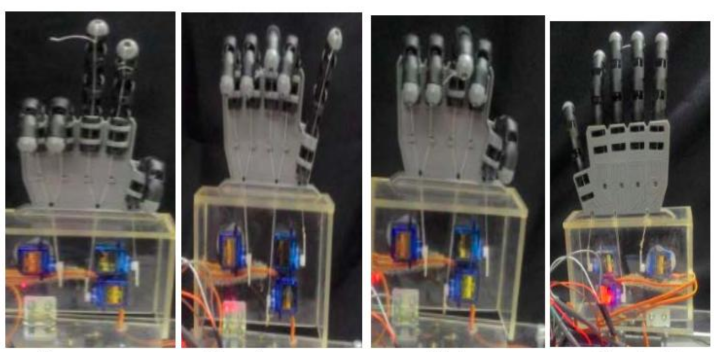


Fig. 5: Methodology flow chart



Peace | Thumbs up | Fist | Hand rotate

Fig. 6: The result of the robotics hand (prototype)

4.0 Results Analysis

This research uses two facial expression datasets (JAFFE & Fer 2013) to train the deep learning classification model. Before training the model, the dataset images are normalized JAFFE and Fer2013 dataset (48 x 48) pixel grayscale image and used data augmentation for both datasets. After data preparation, the convolution neural network (CNN) use for training the model for both datasets. After 100 epochs, the model achieved a higher accuracy of 79% for JAFFE and 96% for the fer2013 dataset. A prototype rehabilitation robot was used to test the facial expression model. The experimental hardware test achieved 87% and 69% accuracy for JAFFE and Fer2013 dataset accordingly for four emotions: neutral, happy, sad, and degust, shown in fig. 7 and 8. JAFFE shows low efficiency in the test performance. Besides, JAFFE and Fer2013 average sample time need for training is 22ms/step and 23ms/step, respectively.

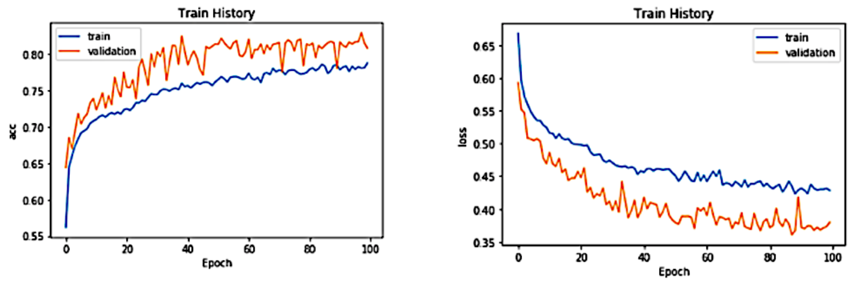


Fig. 7: Fer2013 dataset training and validation accuracy and loss

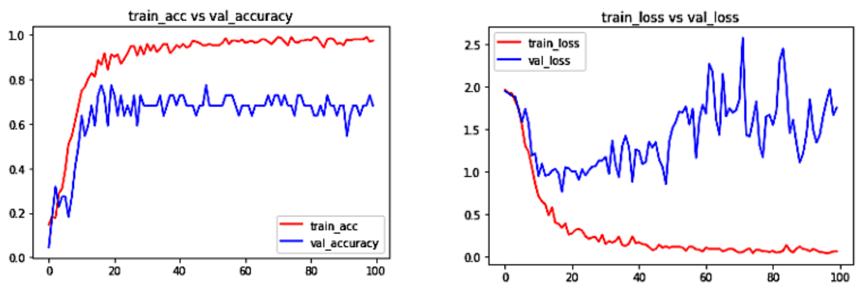


Fig. 8: JAFFE Dataset training and validation accuracy and loss

5.0 Conclusion

This paper proposed a deep-learning facial expression method to improve medical rehabilitation robot performance. Convolution neural network (CNN) algorithms are used in the proposed model. Two different facial expression datasets, JAFFE and Fer2013, are used to train the model. After training the model, this model integrates with a phototype rehabilitation robot to test the experimental hardware. The system simulation training accuracy for JAFFE and fer2013 dataset is 97.01% and 87%. Besides, the prototype test accuracy for four facial expression categories (neutral, happy, sad, and disgust) is 67% for JAFFE and 87% for fer2013. This research proposed that the rehabilitation robot worked as usual when patient feedback was neutral and Happy. Besides, when the patient started feeding terribly (sad) and uncomfortably (disgusted), the rehabilitation robot stopped immediately and slowed down the speed. This work implements the proposed methods into a robotics hand prototype with four emotions. In the future, the proposed methods will be implemented into a fully functional upper limb rehabilitation robot.

Reference

DEOPA, A., SINHA, A., PRAKASH, A. & SINHA, R. K. Facial Expression Recognition using Convolutional Neural Network and SoftMax function on Captured Images. 2019 International Conference on Communication and Electronics Systems (ICCES), 17-19 July 2019 2019. 273-279.

IRFAN, M. & HAMEED, I. A. Deep learning based classification for healthcare data analysis system. 2017 International Conference on Behavioral, Economic, Socio-cultural Computing (BESC), 16-18 Oct. 2017 2017. 1-6.

MASUM REFAT, C. M. & ZAINUL AZLAN, N. 2021. Stretch Sensor-Based Facial Expression Recognition and Classification Using Machine Learning. *International Journal of Computational Intelligence and Applications*, 20**,** 2150010.

REFAT, C. M. M. & AZLAN, N. Z. Deep Learning Methods for Facial Expression Recognition. 2019 7th International Conference on Mechatronics Engineering (ICOM), 30-31 Oct. 2019 2019. 1-6.

SUBIKSHA, K. P. Improvement in Analyzing Healthcare Systems using Deep learning Architecture. 2018 4th International Conference on Computing Communication and Automation (ICCCA), 14-15 Dec. 2018 2018. 1-4.

YAZHINI, K. & LOGANATHAN, D. A State of Art Approaches on Deep Learning Models in Healthcare: An Application Perspective. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 23-25 April 2019 2019. 195-200.