4/25/2020

ADVANCING THE COURSE OF TELEMEDICINE WITH ARTIFICIAL INTELLIGENCE(A.I)

A CAPSTONE PROJECT

BY

ELIE TETTEH-WAYOE

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Background of Study

The emerging trends in health care delivery has made tele-health or tele-medicine an area that requires innovation. The use of machine learning techniques in health care sector has been in existence for quite some time now. Ironically, there has not been much work done in an attempt to apply these concepts to the betterment of tele-medicine.

This research paper provides the techniques which are very crucial to the overall enhancement of tele-health. Emphasis would be placed on the analysis of the sessions performed by these health professionals to ascertain the shortcomings and areas that require improvement.

1.2 Problem Statement

Tele-health or tele-medicine is an emerging area in health care delivery, yet there is no form of assessment to provide feedback to health care professionals on the quality of service rendered during their sessions. This affects the performance of these professionals and could have damaging effects on patient.

A system to assess the performance in a form of facial emotion recognition and speech evaluation to provide feedback in a timely fashion would help a great deal in the improvement of the quality of health care delivery.

1.3 Problem Statement Elaboration

The mode of administering health care in the form of video conferencing has made significant strides in salvaging many life threatening situations. As such, it is very essential to ensure that the quality and efficiency of these methods are not overlooked nor underestimated. A major area of concern is when health professionals have no form of performance metric to afford them the opportunity to strengthen or improve upon key areas which might be detrimental. The application of machine learning tools in a speech or content metric evaluation and facial emotion assessment would be of great benefit in addressing such lingering issues. With this in place, trainees would also benefit from such technological advancements to in their course work.

CHAPTER 2

LITERATURE REVIEW

1.1 Introduction

There is not much literature on the application of AI to the area of telehealth or telemedicine since this is an emerging area in the health care sector. The publications out there cut across a broad spectrum of telemedicine and AI, as such this paper would incorporate some aspects of both and propose the impact such advancements would bring on board.

1.2 Artificial Intelligence In Telemedicine

Artificial Intelligence (A.I) has over the past few years brought a lot of innovation and enhancement to the health fraternity and specifically telehealth. The introduction of such applications in a form disease detection and prediction has made giant strides in enhancing the access to telehealth. Many health firms have introduced these mechanisms to provide health care delivery to the under-deprived areas in several countries like India. Again, training of health care professionals have now become a bit easier through such mediums.

The use of machine learning techniques like natural language processing and facial emotion recognition are some of the steps to further advance the course of telehealth care. Being able to recognize the facial expressions is key to non-verbal communication between humans and the production perception. (Pramerdorfer & Kampel, 2016). More often than not, the facial emotion of a health professional administering health care services can have a negative impact by aggravating the situation of the patient. This occurs when the diagnosis being made tend to portray a worsening condition than expected and uncontrolled facial expressions of fear and panic could be sent across unknowingly to the patient. As such, machine learning in a form of facial emotion recognition can be made to signal such health professionals who may be found in such circumstances to adjust their facial expressions.

Again, with respect to verbal communication, natural language processing to tools can be used to transcribe the speech transcripts into text ascertain whether the sessions meet the basic requirements. This can be attained through the frequency of the essential words that need to be queried during such information sessions.

CHAPTER 3

METHODOLOGY

1.1 Data Description

The data involves two kinds of data sets. A text file of the transcribed voices from the telemedicine sessions and another which contains facial emotion images. The images were about 35000 in total divided across the various emotion categories of Angry, Sad, Disgust, Fear, Surprise, Happy and Neutral. The text files were six in total and were mostly sessions on stroke patients.

1.2 Data Collection

Two major steps were taken in order to get the required data for this project. The image data set was obtained from Kaggle and the text data was retrieved from the voice of the telemedicine sessions on stroke patients. The image data set is called FER 2013 and was used in Kaggle competition some years ago.

1.3 Data Preprocessing and Feature Engineering

Both data sets needed some data preprocessing to get the data into the right format for model building. The image data set was imbalanced as can be seen in Fig 3.1 with a one class being underrepresented. Also, the text files were very clean as transcribed using google speech to text transcriber yet the time stamp had to be stripped off each line to get in ready for further analyses.

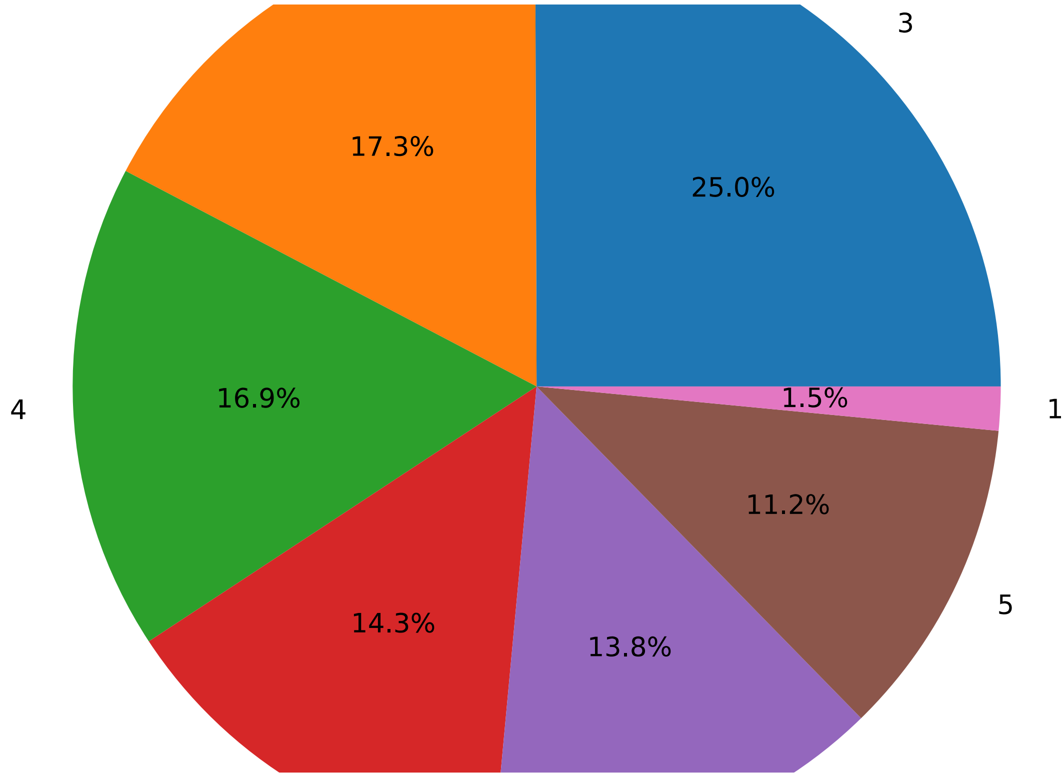


FIG 3.1

1.4 Data Modeling and Visualizations

The models that were used to achieve the desired output included convolution neural networks and natural language processing methods. Python programming language was used in the model building. With respect to the facial emotion recognition model, Keras package was used to build the convolution neural network. The data set was split into 80% training, 10% validation and 10% testing set. The model had three hidden layers and two fully connected layers as can be seen in Fig 3.2

```
Model: "sequential_11"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_81 (Conv2D)           (None, 48, 48, 64)        640
_________________________________________________________________
batch_normalization_101 (Bat (None, 48, 48, 64)        256
_________________________________________________________________
conv2d_82 (Conv2D)           (None, 48, 48, 64)        36928
_________________________________________________________________
batch_normalization_102 (Bat (None, 48, 48, 64)        256
_________________________________________________________________
max_pooling2d_41 (MaxPooling (None, 24, 24, 64)        0
_________________________________________________________________
conv2d_83 (Conv2D)           (None, 24, 24, 128)       73856
_________________________________________________________________
batch_normalization_103 (Bat (None, 24, 24, 128)       512
_________________________________________________________________
conv2d_84 (Conv2D)           (None, 24, 24, 128)       147584
_________________________________________________________________
batch_normalization_104 (Bat (None, 24, 24, 128)       512
_________________________________________________________________
max_pooling2d_42 (MaxPooling (None, 12, 12, 128)       0
_________________________________________________________________
dropout_41 (Dropout)         (None, 12, 12, 128)       0
_________________________________________________________________
conv2d_85 (Conv2D)           (None, 12, 12, 256)       295168
_________________________________________________________________
batch_normalization_105 (Bat (None, 12, 12, 256)       1024
_________________________________________________________________
conv2d_86 (Conv2D)           (None, 12, 12, 256)       590080
_________________________________________________________________
batch_normalization_106 (Bat (None, 12, 12, 256)       1024
_________________________________________________________________
max_pooling2d_43 (MaxPooling (None, 6, 6, 256)         0
_________________________________________________________________
dropout_42 (Dropout)         (None, 6, 6, 256)         0
_________________________________________________________________
conv2d_87 (Conv2D)           (None, 6, 6, 512)         1180160
_________________________________________________________________
batch_normalization_107 (Bat (None, 6, 6, 512)         2048
_________________________________________________________________
conv2d_88 (Conv2D)           (None, 6, 6, 512)         2359808
_________________________________________________________________
batch_normalization_108 (Bat (None, 6, 6, 512)         2048
_________________________________________________________________
max_pooling2d_44 (MaxPooling (None, 3, 3, 512)         0
_________________________________________________________________
dropout_43 (Dropout)         (None, 3, 3, 512)         0
_________________________________________________________________
flatten_11 (Flatten)         (None, 4608)              0
_________________________________________________________________
dense_31 (Dense)             (None, 1024)              4719616
_________________________________________________________________
```

FIG 3.2

1.5 Graphic User Interface (G.U.I)

A graphic user interface was created to portray how the text analyzer works on a transcribed telemedicine session as can be seen and a given image for facial emotion recognition in Fig 3.3. Also, in Fig 3.4, the output from the Facial Emotion recognizer can be seen with the category of emotion and the probability of prediction.

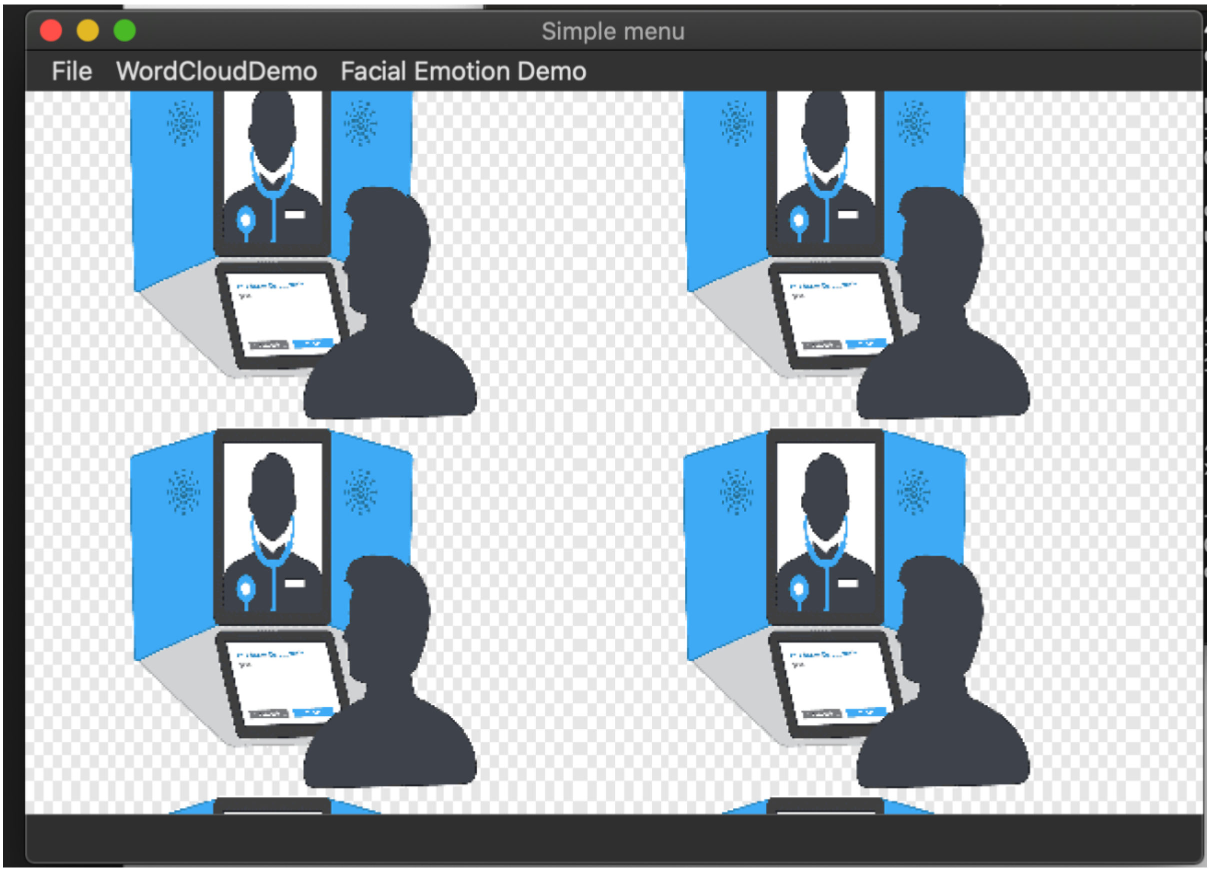


FIG 3.3

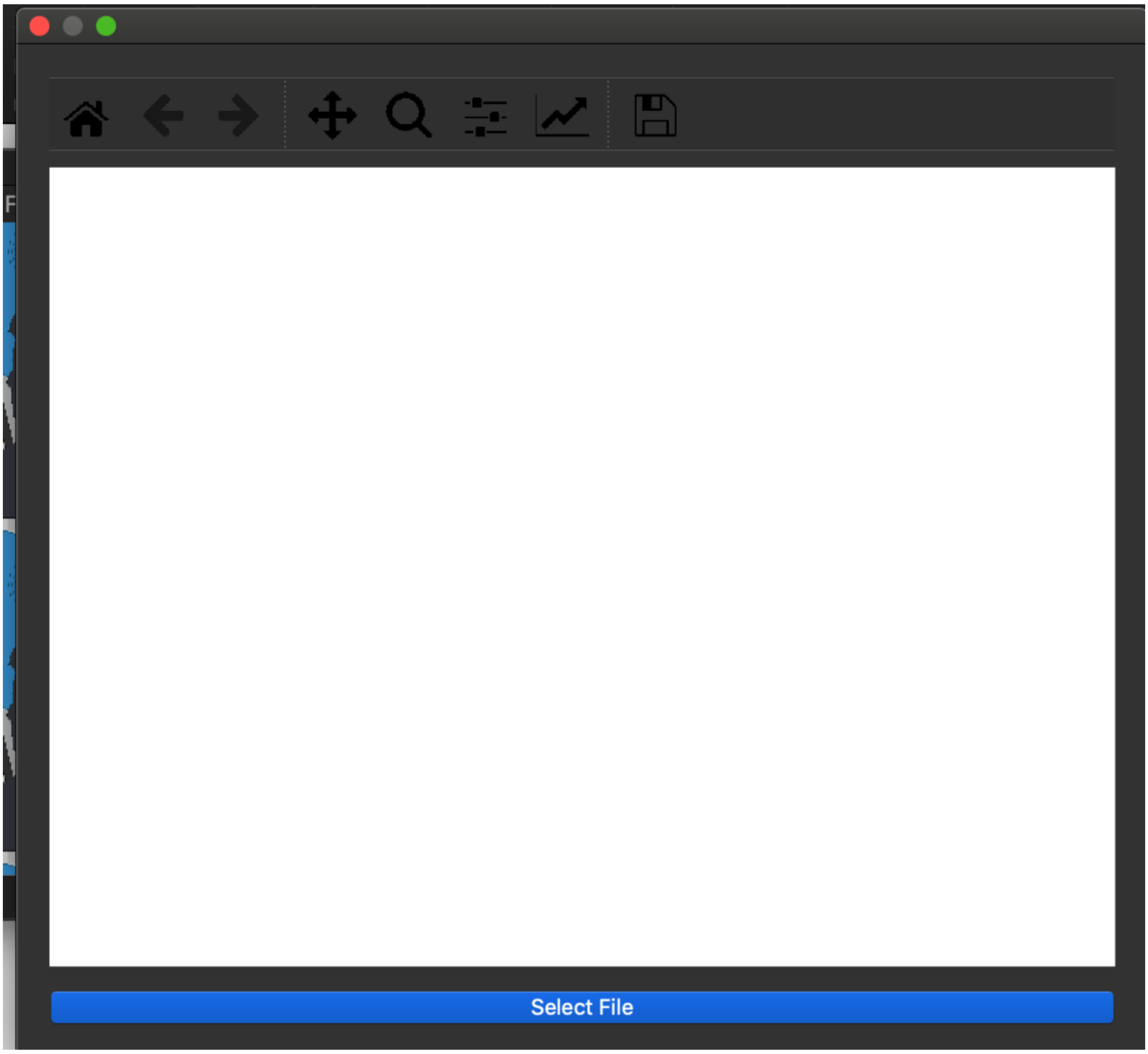


FIG 3.4

CHAPTER 4

RESULTS & ANALYSIS

1.1 Facial Emotion Recognition

The training accuracy was about 74% and a validation accuracy of about 64%. Fig 4.1 and Fig 4.2 shows the model training accuracy and loss respectively. Also, the confusion matrix in Fig 4.2 give the accuracy of the model in predicting each class in the validation data set. It can be seen that the model has generalized though it performs better on certain categories as compared to others.

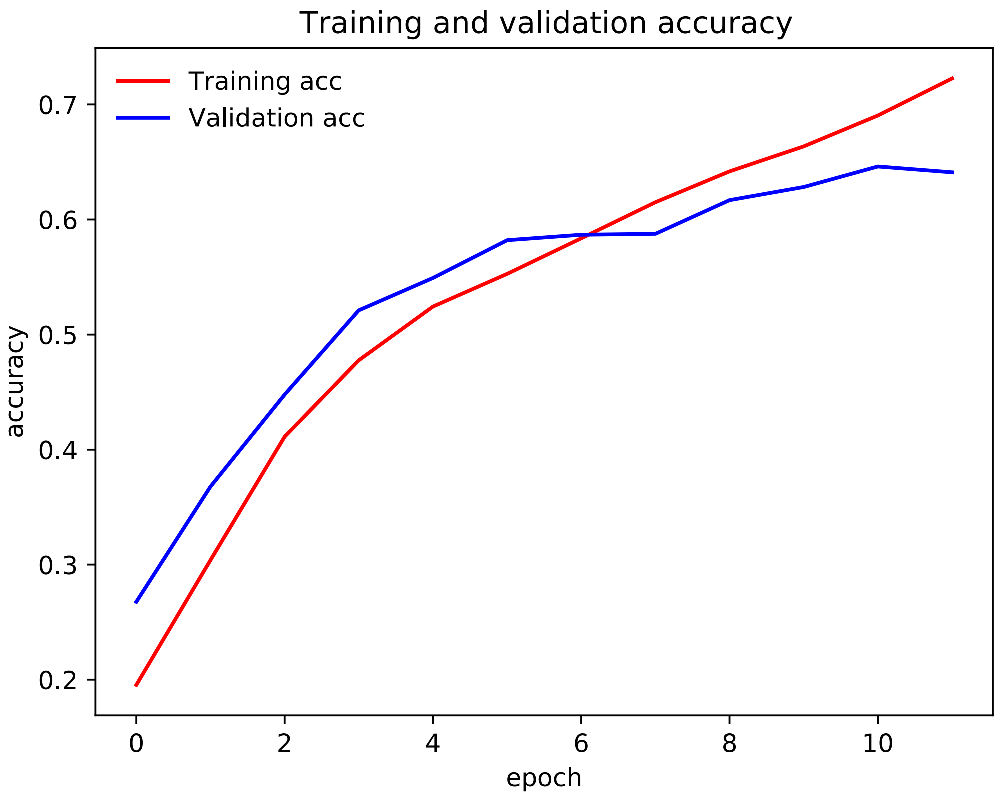


FIG 4.1

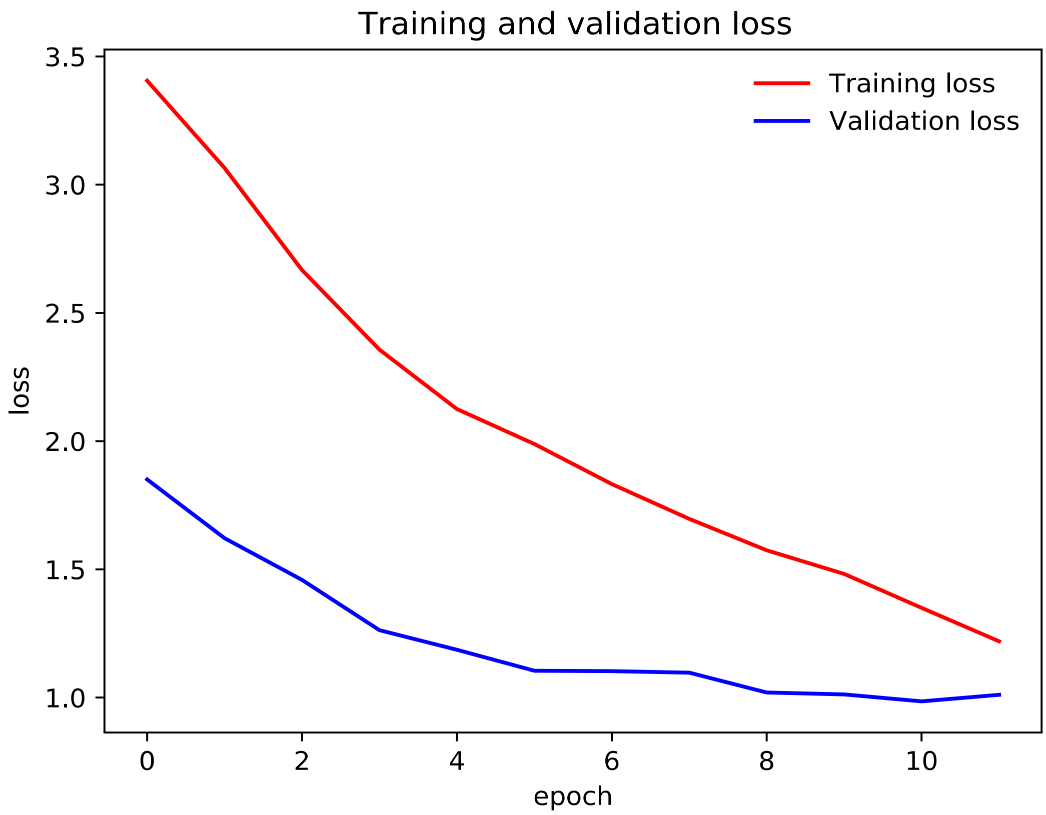


FIG 4.2

1.2 Text Analyzer

The output from the text analyzer is displayed in a world cloud that shows the most frequent occurring words found in the one of the sessions that was performed on a stroke patient as can be seen in fig 4.3

FIG 4.3

1.3 Confusion Matrix

This is the confusion matrix of the model’s prediction on the test set. The model does quite well in predicting almost on average 64% of the emotions correctly. The fig 4.4 shows the predictions on each emotion.


Angry
Disgust
Fear
Happy
Sad
Surprise
Neutral
0.60 0.04 0.08 0.06 0.12 0.02 0.09
0.11 0.75 0.00 0.02 0.09 0.00 0.04
0.12 0.01 0.42 0.05 0.23 0.07 0.10
0.04 0.00 0.02 0.83 0.02 0.03 0.06
0.15 0.03 0.11 0.06 0.48 0.01 0.16
0.04 0.00 0.08 0.04 0.02 0.80 0.01
0.10 0.01 0.08 0.11 0.16 0.01 0.53
Angry
Disgust
Fear
Happy
Sad
Surprise
Neutral
0.75
0.60
0.45
0.30
0.15
0.00

FIG 4.4

CHAPTER 5

CONCLUSION

1.1 Conclusion

The use of A.I in improving the course of Telemedicine is an area which has made tremendous strides and could be of great benefit in times like these where a global pandemic of COVID - 19 has made human contact very deadly. This research is a step in the right direction and requires a lot of contribution to make the needed impact required in the area of health delivery and improvement.

1.2 Project Limitation

The project had some constraints that when looked at could help in the creation of a better product. The image size used in training the model was small in size 48 by 48 and in grayscale. This greatly affected the model building and accuracy generated. Also, the model required high computational resources to train which limited the explorations that could be done to ascertain model improvement.

1.3 Future Research

In an effort to improve upon this research, an integration of recommender system of potential alignments based on some knowledge can be incorporated into the model to ease the physician time taken to analyze possible infections related to the information obtained from the patient.

REFERENCES

Goodfeli-et-al-2013 Challenges in Representation Learning: A report on three machine learning contests.

E. Sariyanidi, H. Gunes, and A. Cavallaro, “Automatic analysis of facial affect: A survey of registration, representation, and recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 37, no. 6, pp. 1113–1133, 2015

http://arxiv.org/abs/1307.0414

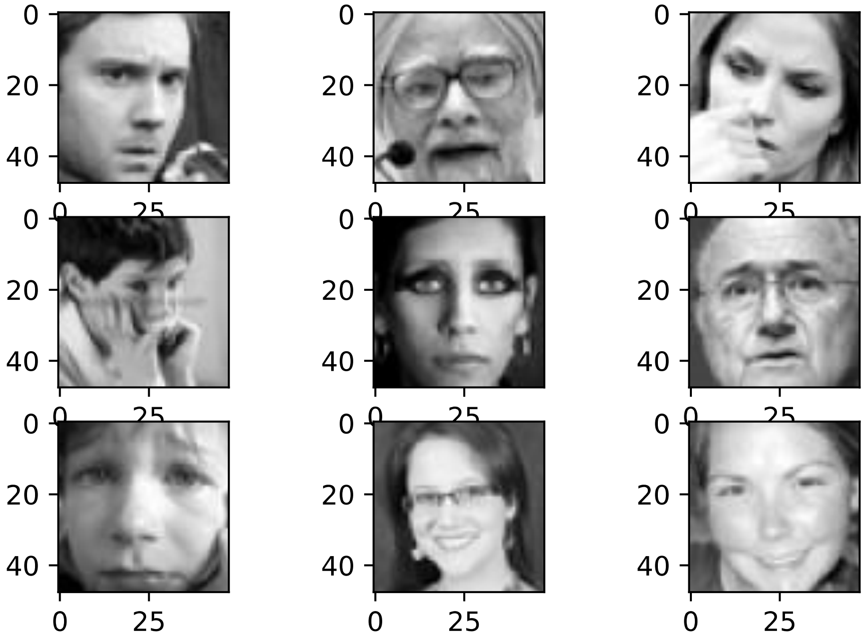
https://www.ncbi.nlm.nih.gov/pubmed/30844756

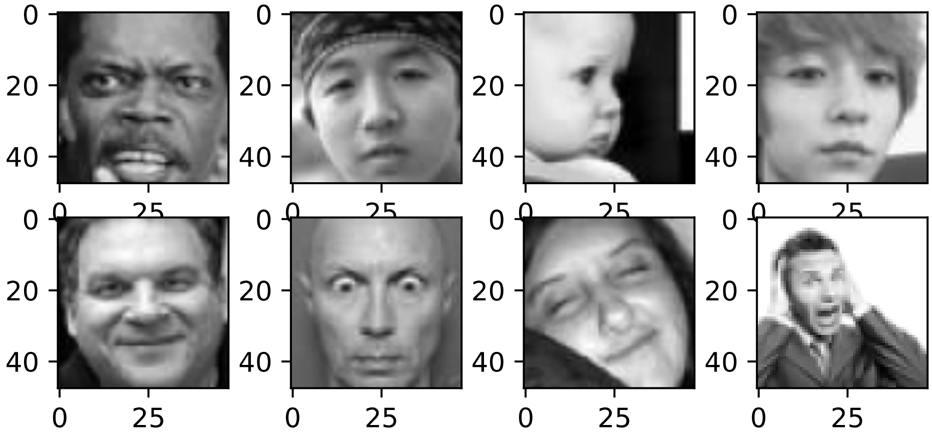
https://emerj.com/ai-sector-overviews/artificial-intelligence-in-telemedicine-and-telehealth/

https://arxiv.org/pdf/1612.02903.pdf

APPENDIX

SAMPLE IMAGES FROM DATA SET




0
20
40
0
25