CALIFORNIA STATE UNIVERSITY
BAKERSFIELD

Assisting Telepsychology with Facial Emotion Recognition During the COVID-19 Pandemic Era

Yeana L. Bond and Albert C. Cruz

I. Abstract

- Computer algorithms can reduce the strain of online patient visits in U.S. healthcare, and this research will use facial emotion recognition algorithms to assist therapists and mental health professionals engaged in remote therapy sessions for better patient outcomes.
- An AI (Artificial Intelligence) program trained on real-world examples of music reaction videos from YouTube can use the subtle visual cues in captured images from the collected videos to identify emotional states since genuine expressions are induced by the music therein [1].

II. Background

- According to a 2021 survey, 96% of psychologists have continued providing telehealth services since the height of the COVID-19 pandemic [2].
- Music therapists have also shifted to providing services remotely through telehealth [3] to ensure the safety of their clients [4], and our work aims to help in their efforts to care for their patients remotely.

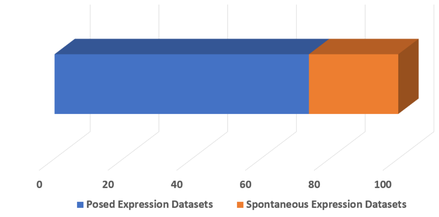

Figure 1: Types of datasets used in previous studies [5]

- While we are not the first to propose developing an AI system that detects facial expressions to determine a subject's emotional state, to the best of our knowledge we are the first to do so using YouTube reaction videos as our dataset whereas more than 74% of previous studies use databases consisting of posed, exaggerated expressions [5].

III. Dataset and Data Tagging

- Datasets containing "faked" facial expressions are not necessarily representative of actual emotional responses as they do not account for the full range of potential expression which would include more subtle, quick, and less apparent signs in a subject's face, also known as micro-expressions.

Figure 2: Comparison between some facial expressions. The top five images are from the posed dataset, CK+, and the bottom two images captured inside a reaction video from a YouTube channel, *The Adventures with TNT*

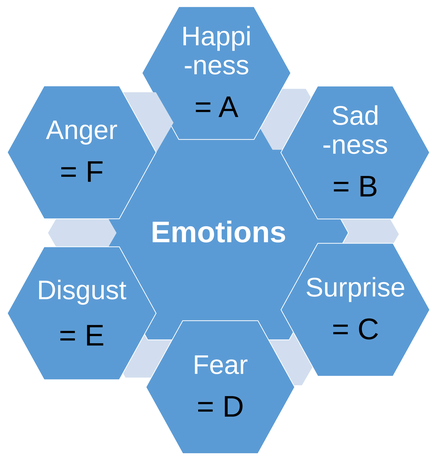

Figure 3: In this research, six basic emotions suggested by Dr. Ekman are used [7], and the alphabet letters from A to F represent an assigned emotion value.

- We will attempt to make up for the current lack of large-scale micro-expression datasets [6] by establishing our own dataset containing genuine micro-expressions obtained from YouTube reaction videos and training an AI program using such a real-world, spontaneous expression dataset.

Apex 1, emotion	Apex 2, emotion	Apex 3, emotion	Apex 4, emotion	···	Apex n, emotion	Name of the Artist	≈	Ethnicity/ Sex
1:36, E	1:49, A	2:39, A	3:19, C	···	6:06, B	Steely Dan	≈	Black/ Male

Meta data

At the above time stamps, we can open, grab, and retrieve images from a video using a Python's built-in class, cv2.VideoCapture().

Figure 4: A sample of data tagging and matching one of the emotions per an apex point. This is done manually on an Excel sheet saved as a .csv file. The letters in this table correspond to the emotions in Figure 3.

IV. Algorithms and Methods

- We will use Convolutional Neural Network (CNN), a type of Deep Neural Network (DNN) that focuses on extracting useful features from images and works as an image classifier, and we will link temporal dependencies between a set of images employing Long-Short Term Memory network (LSTM). This is both a DNN and a special type of Recurrent Neural Network (RNN) that can aid in establishing the temporal relationship between the frames in a video [8].

V. Reference

[1] X. Li, T. Pfister, X. Huang, G. Zhao and M. Pietikäinen, "A Spontaneous Micro-expression Database: Inducement, collection and baseline," in *2013 10th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG),* 2013, pp. 1-6, doi: 10.1109/FG.2013.6553717.

[2] American Psychological Association, "Worsening mental health crisis pressures psychologist workforce 2021 COVID-19 Practitioner Survey," 2021. [Online]. Available: https://www.apa.org/pubs/reports/practitioner/covid-19-2021

[3] N. Gelfand, "The Experience of Music Therapists in Delivering Telehealth," M.S. thesis, Dept. Music, Molloy College, Rockville Centre, NY, USA, 2021.

[4] S. B. David Knott, "Virtual Music Therapy: Developing New Approaches to Service Delivery," *Music Therapy Perspectives,* vol. 38, no. 2, pp. 151-156, 2020.

[5] A. Masson, G. Cazenave, J. Trombini, and M. Batt, "The current challenges of automatic recognition of facial expressions: A systematic review," in *AI Communications*, Dec. 2020. [Online]. Available: https://content.iospress.com/articles/ai-communications/aic200631

[6] Y. Jiao, M. Jing, Y. Hu, and K. Sun, "Research on a Micro-Expression Recognition Algorithm based on 3D-CNN," in *2021 3rd Int. Conf. on Intelligent Control, Measurement and Signal Processing and Intelligent Oil Field (ICMSP),* 2021, pp. 221-225, doi: 10.1109/ICMSP53480.2021.9513351.

[7] C. Mohn, H. Argstatter, and F.-W. Wilker, "Perception of six basic emotions in music," *Psychology of Music*, vol. 39, no. 4, pp. 503-517, 2011, doi: 10.1177/0305735610378183

[8] F. An and Z. Liu, "Facial expression recognition algorithm based on parameter adaptive initialization of CNN and LSTM," *The Vis. Comput.,* vol. 36, pp. 483-498, 2020. [Online]. Available: https://doi.org/10.1007/s00371-019-01635-4