

Thrivee Use of Advanced Technologies

Supporting Evidence





USE OF ADVANCED TECHNOLOGIES

OVERVIEW

This document outlines supporting evidence of advanced technologies that supports how Thrivee delivers virtual care. Thrivee utilizes advanced technology to aid providers in the delivery of virtual care to augment the core telebehavioral health platform. Advanced technologies serve to augment the patient and provider experience to enhance provider insights for identification and diagnosis in addition to improving workflow. Thrivee intends to integrate real-time speech to text voice recognition in 2018 adding advanced voice recognition machine learning and artificial intelligence. Thrivee will add advanced facial recognition in 2019 and beyond as the technologies mature and where they demonstrate improvement in the delivery of care.

Speech to text voice recognition is used to create real-time transcripts of virtual care encounters and group counseling sessions and identification of depression and anxiety. Facial recognition analysis supports providers with enhanced real-time tools to improve the clinical accuracy and care delivery insights.

Thrivee speech-to-text uses commercial cloud computing technologies for voice recognition. Natural language processing leverages machine learning to find insights and relationships within the text. These insights and relationships are presented as tools within the virtual visit encounter and flags for post-visit review and follow-up.

CLINICAL IDENTIFICATION OF ANXIETY AND DEPRESSION

The clinical criteria for diagnosis for the continuum of anxiety and depression disorders is well established¹ along with clinical observations of anxiety from auditory and visual cues.² Yet the accuracy of diagnosis in the primary care setting is lacking. A 2011 study of Canadian primary care clinics assessing 840 patients by usual care versus a validated instrument for structured diagnostic psychiatric interview.³ Using this comparison, the researchers found detection rates by primary care providers to be considerably low. Misdiagnosis rates indicative of type II error (i.e. false negative) for general anxiety disorder was 71.0% and 65.9% for major depressive disorder.⁴ Therefore, primary care providers may benefit from any tool that aids in the identification of anxiety and depression.

AUDIO IDENTIFICATION OF DEPRESSION AND ANXIETY





USE OF ADVANCED TECHNOLOGIES

Audio cues can aid in the real-time identification and diagnosis of depression using verbal attributes and vocal acoustics. Verbal attributes such as word cadence, pauses, and vocal inflection are extractable features from audio files for analysis.⁵ The pyAudioAnalysis⁶ is an open-source library for audio file analysis able to estimate a patient's depression using real-time audio analysis with 70% accuracy.⁷ The researchers applied estimated a clinical depression score using the Beck Depression Inventory (BDI),⁸ a validated instrument for measuring depression. The accuracy of the analysis was evaluated using the 2013 Continuous Audio/Visual Emotion and Depression Recognition Challenge (AVEC) dataset.⁹ Vocal acoustics is used to identify depression where patients exhibiting a depressive state exhibit different vocal patterns, specifically in higher energy bands.¹⁰ Classification of depression is possible using Mel-Scale Frequency Cepstral Coefficients (MFCC).¹¹ A 2018 study of 36 patients with major depressive disorder and 36 healthy controls used the open-source OpenSMILE¹² speech feature extraction tool to analyze acoustic features. The researchers found the second dimension of MFCC was able to discriminate between patient groups with a sensitivity of 77.8% and specificity of 86.1%.¹³

VIDEO IDENTIFICATION OF DEPRESSION AND ANXIETY

Analysis of video enables identification of depression and anxiety in real-time involves the use of facial recognition and advanced analysis using algorithms. A generalized process involves mapping of facial features and landmarks, identifying transient changes, and correlating those changes to behavioral and emotional states. A key underlying assumption of many forms of video analysis is that "motion must be smooth in relation to the frame rate."¹⁴ Humans process emotional observations with a tendency to focus on and identify the eyes and the mouth.¹⁵ In contrast, Facial Emotion Recognition (FER) technology uses an input image to identify facial landmarks well beyond the eyes and mouth, extracting features, and subsequent classification of features.¹⁶ The Facial Action Coding System (FACS)¹⁷ is the benchmark for facial analysis that categorizes facial muscle activation of an expression into Action Units (AU).¹⁸ A 2016 experimental trial of inducing emotional states via experimental tasks and conditions yielded 80-90% accuracy for facial recognition identification of stress and anxiety.¹⁴ This readily exceeds the high variation noted via human observers for detection of emotional state and trait anxiety using visual and audio cues.² Photoplethysmography (PPG), a readily available optical method to detect microvascular blood flow changes, may aid in FER identification of emotional state¹⁹ along with other bio-signaling methods.²⁰

COMBINED AUDIO AND VIDEO ANALYSIS





USE OF ADVANCED TECHNOLOGIES

The benefit of multi-modal observation using both audio and visual cues is greatly improved accuracy and performance²¹ with deeper clinical insights. Various methods exist for fusion schemes and automation of algorithms used to merge and subsequently analyze audio and video feeds^{22,23,24,25,26} A 2015 review of automated analysis of combined audio and video for depression detection cited classification accuracy for depression between 70% and 80% and some studies reporting as high as 90% accuracy.²⁶

¹ American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed.). 2013; American Psychiatric Publishing, Arlington, VA.

² Harrigan JA, Wilson K, Rosenthal R. Detecting state and trait anxiety from auditory and visual cues: a meta-analysis. *Pers Soc Psychol Bull.* 2004;30(1):56-66.

³ Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry.* 1998;59 Suppl 20:22-33;quiz 34-57.

⁴ Vermani M, Marcus M, Katzman MA. Rates of Detection of Mood and Anxiety Disorders in Primary Care: A Descriptive, Cross-Sectional Study. *The Primary Care Companion to CNS Disorders.* 2011;13(2):PCC.10m01013.

⁵ Mundt JC, Snyder PJ, Cannizzaro MS, Chappie K, Geralt DS. Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *Journal of neurolinguistics.* 2007;20(1):50-64.

⁶ Giannakopoulos T. pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. *PLOS ONE.* 2015;10(12): e0144610.

⁷ Giannakopoulos T, Smailis C, Perantonis S, Spyropoulos C. Realtime depression estimation using mid-term audio features. *International Workshop on Artificial Intelligence and Assistive Medicine.* 2014;2(7):228-230.

⁸ Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry.* 1961;4:561-71.

⁹ Valstar M, Schuller B, Smith K, Eyben F, Jiang B, Bilakhia S, Schnieder S, Cowie R, Pantic M. AVEC 2013 - The continuous Audio/Visual Emotion and depression recognition challenge. *AVEC 2013 - Proceedings of the 3rd ACM International Workshop on Audio/Visual Emotion Challenge.* 2013;3-10.

¹⁰ Parikh A, Zafar S, Saini M. Audio/Speech Signal Analysis for Depression. *IJSRD.* 2014:228-230.

¹¹ Laosaphan T, Yingthawornsuk T. Classification of Depressed Speakers based on MFCC in Speech Samples. *ICAEEE 2012, Pattaya, Thailand.*

¹² audEERING GmbH. openSMILE: The Munich Versatile and Fast Open-Source Audio Feature Extractor. Available at: <https://audeering.com/technology/opensmile/> Accessed May 15, 2018.

¹³ Taguchi T, Tachikawa H, Nemoto K, Suzuki M, Nagano T, Tachibana R, Nishimura M, Arai T. Major depressive disorder discrimination using vocal acoustic features. *J Affect Disord.* 2018;1(225):214-220.

¹⁴ Giannakakis G, Pediaditis M, Manousos D, Kazantzaki E, Chiarugi F, Simos PG, Marias K, Tsiknakis M. Stress and anxiety detection using facial cues from videos. *Biomed Signal Process Control* 2017;31:89–101.

¹⁵ Wegryzn M, Vogt M, Kireclioglu B, Schneider J, Kissler J. Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS ONE* 2017;12(5): e0177239.

¹⁶ Ko BC. A Brief Review of Facial Emotion Recognition Based on Visual Information. *Sensors (Basel).* 2018; 30:18(2).

¹⁷ Ekman P, Friesen WV (1978) Facial Action Coding System: A Technique for the Measurement of Facial Movement. 1978; Consulting Psychologists Press, Palo Alto, CA.

¹⁸ Shichuan D, Tao Y, Martinez AM. Compound facial expressions of emotion. *PNAS.* 2014;111.15:E1454-E1462.





USE OF ADVANCED TECHNOLOGIES

- ¹⁹ Allen J. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas.* 2007;28(3):R1-39.
- ²⁰ Sim H, Lee WH, Kim JY. A Study on Emotion Classification utilizing Bio-Signal (PPG, GSR, RESP) [C]. *Art, Culture, Game, Graphics, Broadcasting and Digital Contents.* 2015: 73-77.
- ²¹ He L, Jiang D, Sahli H. Multimodal depression recognition with dynamic visual and audio cues. *Proc Int Conf Affective Comput Intell Interaction* 2015;260–266.
- ²² Datcu D, Rothkrantz LJM. Semantic Audio-Visual Data Fusion for Automatic Emotion Recognition. *Proceedings of the EuroMedia.* 2008.
- ²³ Dham S, Sharma A, Dhall A. Depression Scale Recognition from Audio, Visual and Text Analysis. *Proceedings of the Audio/Visual Emotion Challenge and Workshop.* 2017.
- ²⁴ Senoussaoui M, Sarria-Paja M, Santos JF, Falk TH. Model Fusion for Multimodal Depression Classification and Level Detection. *Proceedings of the Audio/Visual Emotion Challenge and Workshop.* 2014.
- ²⁵ Kächele M, Glodek M, Zharkov D, Meudt S, Schwenker F. Fusion of audio-visual features using hierarchical classifier systems for the recognition of affective states and the state of depression. *Proceedings of the International Conference on Pattern Recognition Applications and Methods.* 2014.
- ²⁶ Girard JM, Cohn JF. Automated Audiovisual Depression Analysis. *Curr Opin Psychol.* 2015;4:75-79.

