Computerized Identification of Mental State for Mental Health Care

Frederick Watkins
Angel Cortez
Veronica Ford

Austin Community College: Northridge Campus

2022

Introduction

Sebastian, a 30-year-old man, was just diagnosed with depression and is undergoing cognitive-behavioral therapy (CBT) sessions on a weekly basis. As he goes about his daily routine, he wears a gadget that measures his physiological activity. His therapist enters his physiological data for the week into a computer software at the start of each CBT session. The computer compiles Sebastian's levels of happy and negative affect and pinpoints the times when these reactions peaked. The therapist utilizes this information to track Sebastian's development and dynamically customize the therapy, for as by asking Sebastian to recollect experiences that correlate to some of the emotion peaks.

Camila is a 3-year-old child who recently had her right shoulder dislocated. She is in pain, but she is unable to explain the extent of her discomfort to her doctor. The doctor instructs her to keep motionless for 5 seconds before rotating her arm using routine range of motion testing (e.g., abduction, flexion). This causes her to wince and she lets out a little cry. The doctor films her face both while her arm is motionless and during the examinations. The footage is accompanied by an audio recording of her sob. Camila's pain intensity is estimated using a computer software that analyzes changes in her facial expressions and voice patterns during range of motion tests versus when her arm is held motionless. In addition to the X-ray and physical exam, he chooses a course of therapy based on the computer-generated assessment of Camila's pain severity.

Lewis just received a diagnosis of attention deficit hyperactivity disorder (ADHD) [1], [2]. He just cannot focus on his studies, and his grades are beginning to suffer as a result. This makes him unhappy since he believes he is not "clever enough" to succeed in college. Lewis' psychiatrist prescribes a typical amphetamine stimulant. He also hands Lewis a little gadget that he may attach to his laptop. When Lewis uses his laptop to do school assignments, assigned readings, or study for an exam, the gadget analyzes his eye stare. A computer software analyzes his eye gazing and produces estimations of Lewis's levels of attentiveness during the study session. It can tell when Lewis is focused, when he is distracted, and even when he is zoning out despite his best efforts. The program gives Lewis feedback on his levels of attention as they relate to the various study tasks. Lewis utilizes this knowledge to restructure his study technique and re-study specific areas. Furthermore, a new software tool he downloaded may use this information in real-time to recommend areas for further investigation [3]. Lewis's grades begin to improve, and he begins to feel more powerful.

These hypothetical examples show how technology (devices and computer programs) that can identify a person's mental state automatically can give actionable information to improve mental health care. As in the situations of Sebastian and Camila, these machine-provided mental state assessments can supplement self- or observer-reports of the same constructs. They can also allow for introspection and dynamic action, as demonstrated by Lewis. These machines may detect physiological arousal, emotions (for example, pain), affective dimensions (for example, valence and arousal), cognitive states (for example, attention and mental effort), emotional states (for example, sad, pleased, furious), and even complex cognitive-affective blends (e.g., confusion, frustration). They can target both brief occurrences (such as particular attentional lapses) and intermediate mood states (such as a poor day) (e.g., stress and depression). Some are ideally suited for usage in controlled settings (e.g., a physician's office), others for home and office use, and yet others for long-term ambulatory monitoring of mental states.

The term "AI" refers to the broad objective of producing computers capable of perceiving complex mental events, which was previously a uniquely human talent, as well as the several subfields of AI engaged in the creation of such machines (e.g., computer vision, machine learning). The goal of this research is to debunk the myths surrounding these supposedly miraculous technologies that can "read" a person's mental state. This is accomplished by first laying the theoretical and technological groundwork for the extremely multidisciplinary topic of "automated mental state detection" in 118 Artificial Intelligence in Behavioral and Mental Health Care. Following that, a few exemplary instances of modern mental state detecting systems are shown. The research finishes with a review of the field's outstanding concerns and some speculative thoughts on its future.

Foundations, Theoretical And Technical

Automated detection of mental states is a growing topic of study within the wider discipline of human computer interaction and its sister fields of human factors and cognitive ergonomics. It is made up of several subfields, including social signal processing, emotional computing, attentionaware computing, and augmented cognition, each of which focuses on distinct mental states in various circumstances. The topic of automated mental state detection is really multidisciplinary. Its psychological foundations are found in cognitive psychology, affective sciences, social psychology, nonverbal behavior research, and psychophysiology. Its technological origins may be traced back to engineering and computer science, notably sensors and wearable devices, digital signal processing, and machine learning.

The psychological component of automated mental state detection is based on ideas that emphasize the embodied aspect of mental processes. According to embodied theories of cognition and emotion, mental states are expressed in the body rather than the mind. The heightened activation of the sympathetic nervous system during fight-or-flight reactions is one of the most direct examples of a mind-body interaction. There are also well-known associations between facial expressions and affective states, such as the wrinkled brow during perplexity. There is also a long history of investigating cognitive processes such as attention and cognitive load utilizing bodily/ physiological responses. For example, the study of eye movements (oculesics) has emerged as a significant technique for investigating visual attention, but electroencephalography has long been employed as a mental strain metric. This intimate mind-body interaction is understandable when one considers that cognition and affect are used to facilitate action. Simply put, we act on what we believe and feel. Because bodies are the actors of action, tracking observable physiological changes can reveal vital insights into unobservable mental states. This central concept underpins the automated detection technique, which aims to deduce mental states from body reactions.[4]–[8]

Figure 1 summarizes the fundamental concepts of automated mental state detection. The main assumption is that a person's interactions with the world (situational context) produce latent (or hidden) mental states that cannot be assessed directly. Circular causation describes how mental states are related with changes at several levels (neurobiological, physiological, physical expressions, overt behaviors, and subjective/metacognitive feelings/reflections), which in turn impact the mental states themselves. Some of these changes are implicit (for example, neurological and physiological changes), occurring outside of conscious awareness, whilst others are more clear (e.g., overt actions, metacognitive reflections). Some of these implicit and explicit changes are visible to machine sensors and human observers, while others are solely visible to the self (dotted lines in Figure 1). The computational challenge is to infer (or estimate) hidden mental states from machine-readable data captured by sensors.

To solve the aforementioned inference problem, two interconnected computational problems must be overcome. The initial difficulty is to extract diagnostic information (called features) from sensor inputs. Because the method utilized changes depending on the sensor and related signal, this is a sensor-specific component of mental state detection. For instance, if the sensor is a camera, the signal is video (presumably of the face). The characteristics in this situation might be the activation of certain facial muscles or Action Units (AUs) (Ekman & Friesen, 1978), such as the inner brow lift (AU 1) or the lip pucker (AU 18). To automatically calculate these face traits from video, computer vision-based algorithms are required. Similarly, pitch and amplitude are frequent paralinguistic (acoustic-prosodic) variables retrieved from a microphone-recorded audio signal (the sensor). Digital signal processing technologies applied to the speech domain are required in this case. If the content of the spoken signal is to be studied, automated speech recognition must be used, followed by natural language processing techniques to find relevant aspects.

The second problem entails inferring a person's mental state based on the properties collected from the signals. This is classified as machine learning and is partially, but not fully, signal-independent. To overcome this challenge, most (but not all) researchers use supervised learning approaches. In its most basic form, supervised learning attempts to learn a program automatically from training data. In general, supervised learning proceeds as follows. Annotated data in the form of features (extracted from signals recorded by sensors, as previously mentioned) with temporally synchronized annotations of mental states (usually provided by humans) is collected as a large number of people are experiencing the mental states. The relationship between the features and their corresponding annotations is then automatically modelled (learned) using supervised learning methods (dashed lines in Figure 1). Aspects of the situational context are occasionally used as extra input to provide context to the learning process. When presented with new Automated Mental State Detection for Mental Health Care 121 data without correlating annotations (e.g., collected at a later time and/or from a person not in the training data solid lines in Figure 1), the resulting model created during supervised learning is then used to produce computer-generated estimates of mental states. The two most immediate performance metrics are accuracy and generalizability. Accuracy is defined as the degree to which automated mental state estimations match some objective benchmark, often self- or observer-reported mental states, while, generalizability is concerned with the detectors' resilience when used to data other than that used to train the supervised classifiers [9], [10].

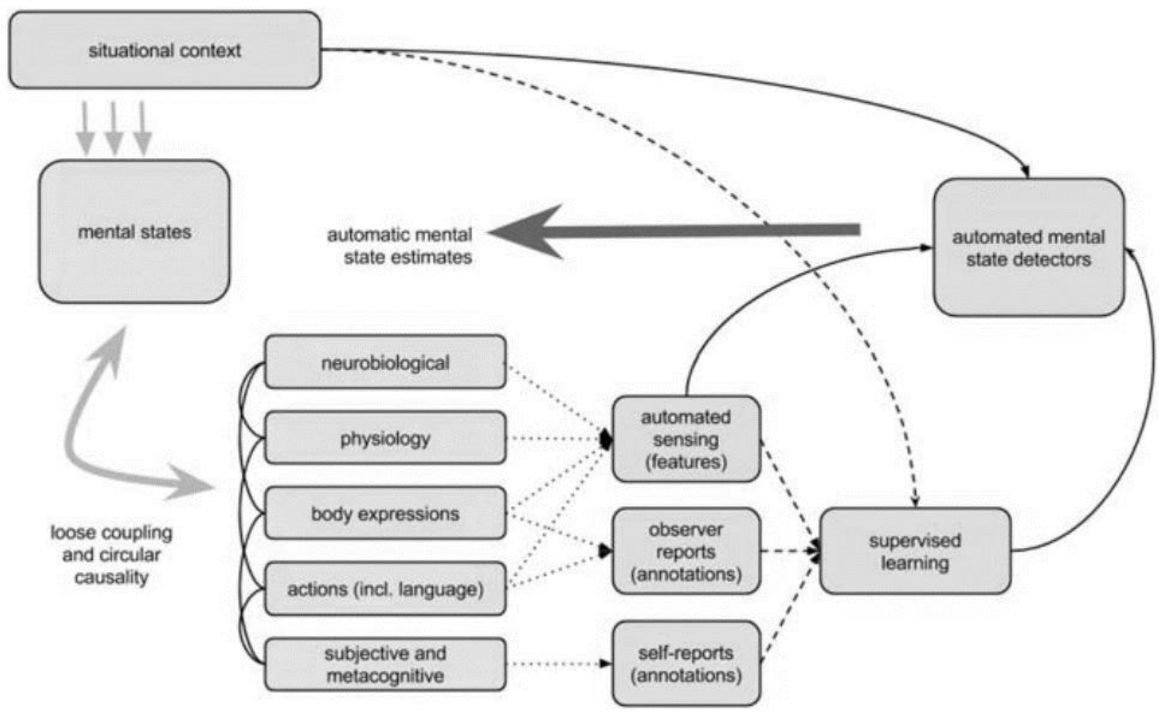

Figure 1.

It should be emphasized that the preceding article purposefully glosses over many of the complexity required in the various phases of developing an automated mental state detector. There are numerous intricacies to data gathering and annotation that must be learned. Computing diagnostic characteristics necessitates the resolution of several outstanding problems in the relevant disciplines. Then there's the problem of picking a subset of diagnostic features, modeling feature connections, and lowering the dimensionality of the feature subspace. Before supervised learning can begin, the training data must often be collected and altered in a variety of ways. The choice of supervised learning technique is next made, followed by ways to parameterize the model. When various modalities (e.g., auditory and visual) are employed, the extra difficulty of selecting how to blend modalities arises. Finally, adequate validation techniques and measurements must be chosen, which is not an easy task. These difficulties, when taken together, have encouraged rich and productive multidisciplinary research agendas and will continue to do so for many years to come.[6], [11]

Emotional States

Because affect is involved in a variety of mental diseases, automated affect detection might be a potential technique to acquire an indirect evaluation of a person's underlying mental health. Affect is a broad term that encompasses both moods and emotions, which may be classified along a variety of aspects. Emotions are short, intense experiences that dominate the forefront of awareness, have major physiological and behavioral expressions, and swiftly prepare the physical systems for action, whereas moods are more ephemeral and have a background impact on consciousness. The vast bulk of studies on affect detection has concentrated on identifying emotions rather than moods, with an emphasis on the so-called "basic emotions," which commonly include anger, surprise, pleasure, disgust, sorrow, and fear. Nonbasic emotions, such as boredom, perplexity, annoyance, engagement, and curiosity, have some of the characteristics associated with basic emotions but have gotten significantly less attention. Some studies prefer to define degrees of intensity on one or more fundamental affect dimensions, with a particular emphasis on valence and arousal, rather than discrete affect representations (e.g., sad versus furious) (sleepy to active). [12]–[14]

Numerous recent studies show that affect detection is one of the most frequently investigated mental state detection challenges. Affect detection systems range substantially in terms of sensors/signals employed, affect representation, particular affective states observed, whether the states occur naturally or are produced experimentally, and the circumstances in which affect detection occurs. Three affect detection projects are examined to provide background for the work in this field, each stressing a distinct mix of sensor/signal, affecting representation, affecting state, and situational context. [15], [16]

Fundamental Emotions

The first study looked at was a lab study that focused on detecting fundamental emotions evoked by an affect elicitation process. Janssen et al. (2013) compared automatic detection to human perception of three primary emotions (happy, sad, furious), calm, and neutral caused by an autobiographical recall process in Automated Mental State Detection for Mental Health Care 123. This approach required 17 participants to write about two occurrences in their lives that were related with feelings of happiness, anger, sadness, or neutrality. Participants were then asked to recollect a subset of those incidents in such a way that they relived the feelings they had felt, as well as to vocally describe each occurrence (in Dutch). While participants recalled and reported the events, audio, video, and physiological data (electrodermal activity, skin temperature, breathing, and electrocardiography) were captured. Each tape was labeled with the name of the feeling that the subject was asked to recollect. [17]–[19]

A number of standard characteristics were automatically calculated, including specified face landmarks, head position, basic frequency of speech, and overall level and variation in each physiological signal. When just facial and physiological features were included, a support vector machine classifier (supervised learning approach) generated the best results. It properly identified the emotion label of each recording 82% of the time. Furthermore, the authors directly compared machine identification of emotion to human detection of emotion. This was accomplished by asking a group of human judges (both American and Dutch) to determine the moods of the participants based on various stimulation combinations (audio-only, video-only, audio-video). The Dutch judges were the most correct (63%) when only given the audio (which was also in Dutch), whereas the US judges were the most accurate (31%), when both audio and video were offered. However, human accuracy (63% and 31%) was significantly lower than automated detection accuracy (82%), a result with far-reaching ramifications.[20], [21]

Nonstandard Emotions

The second research used a very different strategy from Janssen's, focusing on multimodal detection of (mainly) fundamental emotions induced experimentally in controlled laboratory conditions. Bosch investigated unimodal detection of nonbasic emotions in a noisy real-world scenario of a computer-enabled classroom. As part of their usual physics/physical science lessons, 137 middle and high school students participated in this study by playing a conceptual physics instructional game in small groups for 2.5 hours over three days. Trained observers performed live affect annotations by monitoring students one at a time using a round-robin approach (observing one student until observable effect was identified or 20 seconds had passed and then moved on to the next student in a preplanned order). Boredom, perplexity, joy, engagement, and frustration were the feelings of interest. During game play, videos of students' faces and upper bodies were captured and synced with the impact annotations. The FACET computer-vision tool was used to analyse the movies, which offers estimations of the probability of 19 facial AUs, head attitude (orientation), and location. Using motion-filtering techniques, body movement was also inferred from the videos. To automatically distinguish each emotional state from the others, a machine learning technique was used, which was verified in a way that generalizes to new pupils. Individual automatic detection accuracy rates ranged from 62% (frustrated vs. other states) to 83% (delighted vs. other states), which is noteworthy given the noisy nature of the environment, with students constantly fidgeting, talking to one another, asking questions, leaving to use the restroom, and even occasionally using cellphones (against classroom policy).[22]–[24]

Influence Dimensions

The third research examined is not actually a study, but rather a collection of many initiatives targeted at solving a specific affect detection problem. The concept is that when different researchers apply their own methodologies to their own data sets and use their own criteria to evaluate performance, it is impossible to determine progress in any particular study topic (in this example, effect detection). Direct comparisons of results from several research groups are muddled since any discernible difference can be attributed to the technique, the data, or the performance indicator. One solution to this problem is challenge contests, which are a recurrent subject in computer science and AI research. In this case, researchers are requested to apply their methods to a given dataset, and the results are assessed using a fixed metric(s), allowing for direct comparisons across approaches created by various research groups.[21], [25]

The Audio-Video Emotion Detection Challenge (AVEC) is an annual affect detection competition that began in 2011 as part of the Affective Computing and Intelligent Interaction (ACII) conference series. The 2012 AVEC challenge, which explored automated identification of affect dimensions during Automated Mental State Detection for Mental Health Care 125 human computer interactions, is the subject here. The SEMAINE corpus was utilized in the AVEC 2012 competition to capture naturalistic data of humans interacting with artificial agents. The artificial agents play distinct emotionally stereotypical roles (for example, Spike is furious and combative, whereas Prudence is even-tempered and rational), which biases the affective tone of the discourse. Researchers were given videos of individuals' expressions and audio of their words taken during these emotionally heated conversations. Each film was analyzed by two to six human raters along four affect dimensions: valence (negative to positive), arousal , power (low control to high control), and expectancy (unexpected to expected). Because the affect annotations were continually scaled from 21 to 1, the aim was to anticipate the strength of each affect dimension using automated audio-visual affect detection algorithms. This emphasis on dimensional representations of affect (e.g., valence, arousal, power, expectancy) is a more relevant distinguishing factor than the category or discrete representations used in the prior two research evaluated. [26], [27]

The researchers were given two subsets of the annotated data to create their models (training and development subsets), which were then applied to a second subset for which the annotations were not accessible (test subset). Each research group individually provided influence predictions for each dimension using their methodologies on the test subset. The findings of ten study groups were provided. When the winning team's correlation across the four dimensions was averaged, it was 0.45, a significant accomplishment given the task's complexity.

Pain

Pain has been linked to a variety of mental illnesses. After controlling for age and gender, a large research of 85,088 persons from 17 countries found that persistent back/neck pain over a 12-month period was a positive predictor of mood disorders, anxiety disorders, and alcohol abuse/dependence. Pain measurement, which mostly depends on self-report questionnaires, has well-known drawbacks in terms of subjectivity, interpretability, and administration practicality in specific groups.

Automatic pain detectors can help to reduce some of these issues by providing accurate pain monitoring. The number of automated pain detection systems is small, owing to the difficulties in gathering adequate datasets for detector construction. The recent publication of the UNBC-McMaster shoulder pain expression archive database, on the other hand, is intended to jumpstart research in this field. The database contains 129 people who self-identified as having shoulder discomfort. The data set includes films of individuals conducting eight range-of-motion tests (e.g., abduction, internal and external arm rotation) as well as self-reports of pain level after each test. A subset of the data (200 video sequences from 25 people) has been made accessible to the scientific community in order to aid in the development of automatic pain detection systems.[28], [29] UNBC-McMaster database to demonstrate one such automated pain detection method. Their method involved obtaining appearance-based characteristics from each frame of the video and filtering them using a set of log-normal filters. To create detectors of four degrees of pain, a support vector machine classifier was utilized (no pain, trace pain, weak pain, strong pain). When the validation approach assured generalizability to additional participants, they attained an average classification accuracy of 0.56 (F1 metric), a promising result given the complexity of the challenge and the early phases of research in this field. [30], [31]

Depression

Depression is one of the most frequent and significant mental health problems. Automatic depression detection systems have the ability to significantly reduce the harmful impacts of depression by giving early warning indications of depression and serving as an objective evaluation of the success of depression therapies. The inclusion of the Depression Detection Sub-challenge (DSC) as part of the 2013 and 2014 AVEC series appears to have advanced research in depression identification. The challenge challenges researchers to create and test their own depression detectors on the same dataset, allowing for valid comparisons of each technique because data and assessment criteria remain consistent.

The dataset utilized in the DSC challenge included 240 hours of video (with audio) of 84 individuals performing simple activities directed by a computer interface across numerous sessions. Participants differed in the number of tasks, sessions, and session length. The Beck Depression Inventory-II was used to assess participants' degrees of depression. The data were also tagged for fundamental affect dimensions, which are not included in this paper. For the 2013 and 2014 challenges, a portion of this data was used. This comprised 300 films of participants reading aloud in German extracts from a German fable (northwind task) and answering to basic questions in German (e.g., "what is your favorite meal" freeform task) in the most current 2014 challenge. During these exercises, videos of participants' faces and audio of their utterances were captured. Researchers were given access to a portion of this data, as well as the depression levels of each participant (training and development partitions). The test partition was utilized to analyze the results, and the depression levels of these people were suppressed.[27], [28], [32]

In answer to this difficulty, researchers took a wide range of tactics. The results were measured using the root mean square error (RMSE) between anticipated and actual depression levels for test partition individuals, which varied from 8 to 12. The winning system attained an RMSE of 8.12 by emphasizing simulating the timing and coordination of voice output and facial expression. This result was a little increase above the top result of the 2013 challenge (RMSE of 8.50), which was achieved by the same research team on a comparable but unrelated dataset. [23], [33]–[36]

Conclusion

Measurement is the first step toward genuine change. Fully automated systems that give fine-grained evaluations of a person's mental state over lengthy periods of time and in a variety of circumstances have much to offer the science and practice of mental health treatment. These mental state detection systems may be incorporated at numerous levels within the broader mental healthcare system, such as clinical decision-making, ambulatory monitoring, and technology-supported therapy. This research discussed some of the theoretical and technical issues underlying such systems, as well as grounding the key problems in the context of a few case studies focused on automatically detecting mental states relevant to mental health care pain, depression, and stress). Unfortunately, the highly selective nature of this review precluded a conversation of many other excellent system was developed by dedicated groups of international researchers in a variety of fields who are constantly making theoretical, technical, and practical innovations to tackle the challenge of automatic mental state detection.

Automatic mental state detection is a difficult problem to solve. Despite significant improvements over the years, modern technologies are not yet suitable for practical usage. Many early years in the discipline were spent exhibiting research prototypes as proof-of-concepts for the feasibility of automatic mental state detection. This was important to persuade the early skeptics and detractors who mocked the field's pioneers. These early (generation 1) systems concentrated on a restricted subset of mental states that were acted (or generated) by a small group of persons in the laboratory. Generation 1 was further distinguished by the employment of costly and intrusive sensing equipment that were intrinsically nonscalable, as well as the application of less technically proficient computing approaches and less strict validation processes. We are currently in generation 2, where the emphasis is on identifying realistic experiences of a wider range of mental states in more real-world scenarios, employing more scalable, wearable, and unobtrusive sensing systems, as well as more complex approaches and demanding validation processes. Although these generation 2 systems are expected to make significant progress, they may yet fall short in several areas. There must be a focus on improving detection accuracies, demonstrating applicability across a variety of real-world contexts, achieving generalizability across different populations, and adequately attempting to address thorny Automated Mental State Detection for Mental Health Care 131 ethical concerns. It is not a question of "will" but of "when" these challenges will be addressed, after which automatic mental state detection systems will make a meaningful and measurable difference in people's lives by improving their mental health.

References

[1] K. K. S. Voeller, “Attention-deficit hyperactivity disorder (ADHD),” *J. Child Neurol.*, vol. 19, no. 10, pp. 798–814, Oct. 2004.

[2] A. S. Rowland, C. A. Lesesne, and A. J. Abramowitz, “The epidemiology of attention-deficit/hyperactivity disorder (ADHD): a public health view,” *Ment. Retard. Dev. Disabil. Res. Rev.*, vol. 8, no. 3, pp. 162–170, 2002.

[3] P. Stefanatou, E. Giannouli, Z. Antonopoulou, P. Tsellos, G. Vaslamatzis, and M. Typaldou, “The Concept of Time Perspective Within a Psychiatric Context,” *Eur. Psychiatry*, vol. 33, no. S1, pp. S507–S508, Mar. 2016.

[4] J. S. Obeid, E. R. Weeda, and A. J. Matuskowitz, “Automated detection of altered mental status in emergency department clinical notes: a deep learning approach,” *medical informatics and …*, 2019.

[5] S. K. D’Mello, “Automated mental state detection for mental health care,” *Intelligence in Behavioral and Mental Health Care*, 2016.

[6] E. Garcia-Ceja, M. Riegler, and T. Nordgreen, “Mental health monitoring with multimodal sensing and machine learning: A survey,” *Pervasive Mob. Comput.*, 2018.

[7] P. Stefanatou, G. Konstantakopoulos, E. Giannouli, N. Ioannidi, and V. Mavreas, “Patients’ needs as an outcome measure in schizophrenia,” *Eur. Psychiatry*, vol. 33, no. S1, pp. S453–S453, Mar. 2016.

[8] K. Kollias *et al.*, “The development of the Early Intervention in Psychosis (EIP) outpatient unit of Eginition University Hospital into an EIP Network,” *Psychiatrike*, vol. 31, no. 2, pp. 177–182, Apr. 2020.

[9] S. P. Stefanatou Pentagiotissa, G. K. George Konstantakopoulos, E. G. Eleni Giannouli, S. V. Silia Vitoratou, and V. M. Venetsanos Mavreas, “The relationship between patients’ needs and psychopathology in schizophrenia: Do patients and therapists agree?,” *Eur. Psychiatry*, vol. 30, p. 1376, Mar. 2015.

[10] P. Stefanatou *et al.*, “Fear of COVID-19 impact on professional quality of life among mental health workers,” *Int. J. Environ. Res. Public Health*, vol. 19, no. 16, p. 9949, Aug. 2022.

[11] M. A. U. Alam, N. Roy, and S. Holmes, “Automated functional and behavioral health assessment of older adults with dementia,” *on connected health …*, 2016.

[12] R. A. Calvo, D. N. Milne, M. S. Hussain, and H. Christensen, “Natural language processing in mental health applications using non-clinical texts,” *Natural Language Engineering*, vol. 23, no. 5, pp. 649–685, 2017.

[13] S. Chancellor, M. L. Birnbaum, E. D. Caine, V. M. B. Silenzio, and M. De Choudhury, “A Taxonomy of Ethical Tensions in Inferring Mental Health States from Social Media,” in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, Atlanta, GA, USA, 2019, pp. 79–88.

[14] P. Stefanatou *et al.*, “Patient-reported needs predict perceived psychosocial disability and quality of life beyond symptom severity in schizophrenia,” *Disabil. Rehabil.*, pp. 1–9, Feb. 2022.

[15] T. D. Pham, T. C. Thang, and M. Oyama-Higa, “Mental-disorder detection using chaos and nonlinear dynamical analysis of photoplethysmographic signals,” *Chaos Solitons Fractals*, 2013.

[16] E. J. Choi and D. K. Kim, “Arousal and valence classification model based on long short-term memory and DEAP data for mental healthcare management,” *Healthc. Inform. Res.*, 2018.
[17] A. E. Kazdin, “Annual Research Review: Expanding mental health services through novel models of intervention delivery,” *J. Child Psychol. Psychiatry*, vol. 60, no. 4, pp. 455–472, Apr. 2019.
[18] D. D. Luxton, “Artificial intelligence in psychological practice: Current and future applications and implications,” *Prof. Psychol. Res. Pr.*, vol. 45, no. 5, pp. 332–339, Oct. 2014.
[19] S. Dimitrakopoulos, A. Hatzimanolis, P. Stefanatou, L.-A. Xenaki, and N. Stefanis, “S125. The role of dup, DUI and polygenic score for schizophrenia on cognition in Athens fep study sample,” *Schizophr. Bull.*, vol. 46, no. Supplement_1, pp. S82–S83, May 2020.
[20] W. Moyle, M. Bramble, C. Jones, and J. Murfield, “Care staff perceptions of a social robot called Paro and a look-alike Plush Toy: a descriptive qualitative approach,” *Aging Ment. Health*, vol. 22, no. 3, pp. 330–335, Mar. 2018.
[21] J. Yang, J. Zhou, G. Tao, M. Alrashoud, K. N. A. Mutib, and M. Al-Hammadi, “Wearable 3.0: From Smart Clothing to Wearable Affective Robot,” *IEEE Netw.*, vol. 33, no. 6, pp. 8–14, Nov. 2019.
[22] D. D. Luxton, “An introduction to artificial intelligence in behavioral and mental health care,” *Artificial intelligence in behavioral and mental health*, 2016.
[23] T. Shibata and K. Wada, “Robot therapy: a new approach for mental healthcare of the elderly–a mini-review,” *Gerontology*, 2011.
[24] K.-J. Oh, D. Lee, B. Ko, and H.-J. Choi, “A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation,” in *2017 18th IEEE International Conference on Mobile Data Management (MDM)*, 2017, pp. 371–375.
[25] L. Bickman, “Improving Mental Health Services: A 50-Year Journey from Randomized Experiments to Artificial Intelligence and Precision Mental Health,” *Adm. Policy Ment. Health*, vol. 47, no. 5, pp. 795–843, Sep. 2020.
[26] L. Lu, R. Cai, and D. Gursoy, “Developing and validating a service robot integration willingness scale,” *Int. J. Hosp. Manage.*, 2019.
[27] S. M. Rabbitt, A. E. Kazdin, and B. Scassellati, “Integrating socially assistive robotics into mental healthcare interventions: applications and recommendations for expanded use,” *Clin. Psychol. Rev.*, vol. 35, pp. 35–46, Feb. 2015.
[28] E. Miller and D. Polson, “Apps, avatars, and robots: The future of mental healthcare,” *Issues Ment. Health Nurs.*, 2019.
[29] D. D. Luxton, S. L. Anderson, and M. Anderson, “Chapter 11 - Ethical Issues and Artificial Intelligence Technologies in Behavioral and Mental Health Care,” in *Artificial Intelligence in Behavioral and Mental Health Care*, D. D. Luxton, Ed. San Diego: Academic Press, 2016, pp. 255–276.
[30] C. C. Bennett and T. W. Doub, “Chapter 2 - Expert Systems in Mental Health Care: AI Applications in Decision-Making and Consultation,” in *Artificial Intelligence in Behavioral and Mental Health Care*, D. D. Luxton, Ed. San Diego: Academic Press, 2016, pp. 27–51.
[31] M. Moosaei, S. K. Das, D. O. Popa, and L. D. Riek, “Using Facially Expressive Robots to Calibrate Clinical Pain Perception,” in *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*, 2017, pp. 32–41.
[32] A. Fiske, P. Henningsen, and A. Buyx, “Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy,” *J. Med. Internet Res.*, vol. 21, no. 5, p. e13216, May 2019.
[33] H. Ashrafian, “Can Artificial Intelligences Suffer from Mental Illness? A Philosophical Matter to Consider,” *Sci. Eng. Ethics*, vol. 23, no. 2, pp. 403–412, Apr. 2017.
[34] A. A. Scoglio, E. D. Reilly, J. A. Gorman, and C. E. Drebing, “Use of Social Robots in Mental Health and Well-Being Research: Systematic Review,” *J. Med. Internet Res.*, vol. 21, no. 7, p. e13322, Jul. 2019.
[35] K. L. Fortuna *et al.*, “A Future Research Agenda for Digital Geriatric Mental Healthcare,” *Am. J. Geriatr. Psychiatry*, vol. 27, no. 11, pp. 1277–1285, Nov. 2019.

[36] F. Dino, R. Zandie, H. Abdollahi, S. Schoeder, and M. H. Mahoor, "Delivering Cognitive Behavioral Therapy Using A Conversational Social Robot," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 2089–2095.