Universe International Journal of Interdisciplinary Research
(Peer Reviewed Refereed Journal)

© UIJIR | ISSN (O) – 2582-6417
APRIL 2023 | Vol. 3 Issue 11
www.uijir.com

AFFECTIVE COMPUTING: A COMPREHENSIVE OVERVIEW OF APPROACHES AND ETHICAL IMPLICATIONS FOR EMOTION RECOGNITION

Author's Name: [1]Mafas Raheem, [2]Satwik Gawand

Affiliation: School of Computing, Asia Pacific University of Technology and Innovation, Kuala Lumpur, Malaysia

DOI No. – 08.2020-25662434

Abstract

It is not an overstatement to say that we are living in the age of Artificial Intelligence with new developments and advancements happening at a precipitous speed. While Artificial Intelligence accomplishes amazing feats and becomes more 'human-like', emotions are one of the aspects maintaining the distinction even now. Tremendous interest, understanding and growth have been acquired in emotion recognition whereas affective computing is still considered an old concept. This paper explores the concepts and studies around Affective Computing such as the understanding of emotions and different models that view how emotions can be defined and recognized as well as various ways of implementation like sentiment analysis, facial emotion recognition, and speech analysis. To conclude, the paper goes over the most critical considerations in affective computing and the ethical implications of emotion recognition systems that can have a significant impact on the lives of people.

***Keywords**: Affective Computing, Emotion AI & Models, Facial Emotion Recognition, Sentiment Analysis, Speech Analysis*

INTRODUCTION

Artificial Intelligence has been advancing at a breakneck pace in recent years. Developing Machine Learning and Deep Learning models that could handle and analyse various data for more complex applications like synthesizing speech, understanding images, and generating various forms of media has become vibrant in the contemporary analytics arena. In the pursuit of making Artificial Intelligence closer to humans, vast developments in multiple aspects such as knowledge & learning with Machine Learning algorithms, understanding images with Computer Vision, understanding speech with Speech Synthesis and physical touch with Robotics have been noticed. Furthermore, significant developments to recreate the senses of scent and taste in the field of robotics have been done, but one domain which recently started getting the spotlight is Affective Computing.

Affective Computing alias Emotion AI, is a significant wing of Artificial Intelligence which recognises, understands, and simulates human emotions. It comprises the application of Natural Language Processing and Deep Learning to examine and decode emotions through facial expressions, voice, and textual data. Affective Computing is a barely recognized field and deals with the concepts of motive and intent. Many advanced applications, although mathematical, like deep learning techniques, are black boxes where a complete understanding of how these algorithms make certain decisions are hidden. However, applying modern data science and AI techniques on the sensitive data needs more attention on the ethical aspects along with achieving remarkable and useful results. Countless ethical issues have been raised around the development and implementation of Artificial Intelligence for various use cases such as impersonating a human through text or voice or even manipulating the recording of voice or videos.

Emotion AI can be used in various ways in the commercial market. It can improve the personalization of virtual assistants and provide more natural and empathetic responses. Emotion recognition systems were already deployed that analyse the emotions of customers towards a promotional activity to increase the adoption of certain goods and services as well as help providing good customer service to improve customer satisfaction. Healthcare domain benefits from the integration of Emotion AI where empathetic chatbots help in detecting mental health disorders by targeting certain patterns, emotions, behaviours, or keywords as well as tracking the patient's emotional well-being. Strange use of Emotion AI is also possible in the entertainment industry where virtual characters can become more realistic and engaging in games and other media.

EMOTION AND RELEVANT MODELS

Emotions are defined as a composite phenomenon of evaluative, expressive, behavioural, psychological, phenomenological, and mental concepts.1 Further, emotions play a crucial role in human experience and psychiatric illness.2 At the same time, the term "emotion" has been defined 92 different ways and 9 opposing statements by different researchers over 75 years, hence which were then classified into 11 distinct categories.3 Categorical and Dimensional are the main categories of emotions, both of which are used as the approach for numerous studies and experiments. Emotions are classified into distinct categories which are usually exclusive in the categorical model wheras emotions are represented in a two-dimensional space with valence and arousal as the two axes or dimensions in the dimensional model. Emotions are usually termed as a continuous value based on the intensity of the axes. 4

Basic Emotion Theory (BET) is a popular psychological theory which suggests that emotions are innate and instinctive responses to certain stimuli. This theory is built on the idea that there are a small number of basic emotions (happiness, anger, fear, sadness, and surprise) that are universal, and consistent across regions and cultures and are expressed in similar ways as there are specific psychological, and expressional cues triggered. BET defines emotions as a mental state that emerges in presence of interpersonal or intrapersonal events and involves signals & actions that display some coherence which are like primate relatives.5 Basic Emotion Theory has been the foundation of many studies while acknowledging the latest breakthroughs ranging from unexplored specific emotions to a better understanding of basic mechanisms and the study of emotional expression and understanding. One of the earliest and long-lasting studies by Paul Ekman still displayed universality in emotional understanding and served as the base for many theories and implementations over the years.6

Numerous attempts were witnessed in the past in criticising and improving BET due to the consideration of it being too basic and primitive. One of the recent claims asserted that classifications and inferences were oversimplified and that the developments were inconsistent with the base idea leading to staggering difficulty in testing while emphasizing the parsing of facial expressions.7 Additionally, it investigated the behavioural and ecological view of facial expressions that clears up the dilemmas and divergences in BET's base idea and subsequent developments. Numerous empirical pieces of literature exhibited that BET does not constrain the science of emotion as asserted by Crivelli and Fridlund.5

Numerous philosophers have criticized BET in terms of both theory and methods. Some arguments were made against BET, among that BET has a narrow approach towards emotions where there is no clear distinction between mutually exclusive emotion classes.8 Many researchers in the affective computing domain were exploring and widening the approach and found that there is no one "right" way to define this distinction. Moreover, another approach was suggested which includes programs that BET relies on, through radical enactivism; a way of crude behaviourism that rejects instrumentality assumption, and the foundation upon which other positive approaches about basic and non-basic minds rest is that the "embodied activity of living beings provides the right model for understanding minds."

Numerous other studies opposed BET and asserted that not just machines, but even humans cannot reliably perceive emotions through facial expressions. However, many technology-based companies have claimed that their solutions can recognize emotions with an accuracy of the 90th percentile. With a lack of scientific backing and secrecy maintained around the algorithms, many researchers are sceptical and consider it to be an ethical pitfall. However, realizing emotional responses for a product advertisement got lower stake while emotional responses towards job interviews, loans, international borders, etc. have higher stakes.9 Besides BET, many pieces of research works also used the variations of Russel's circumplex model of emotions where the basic emotions were explained in a 2D space with valence and arousal as the two dimensions.10 An emotion model named "Three Primary Color Model of Basic Emotions" was proposed with four basic emotions corresponding to three core affects that correlate to the three primary colours and mixing these colours in varying proportions which can result in a slew of complex emotions.2

EMOTION AI APPROACHES

There are various ways for emotion recognition. First comes the methods for data collection, which is either gathering responses from participants that obtain a general perspective or using medical-grade equipment in a controlled scenario to get precise information; to understand the relationship between emotions and conscious or unconscious expressions and actions. Although these approaches might seem impractical to implement, the goal of these is to extract insights and understandings about different features and their relationship with emotions. Furthermore, it was understood that use of machine learning techniques for data analysis would pave the way towards breakthroughs in practical implementations.10 The study of emotion recognition through expression has forged beyond face recognition for six emotions where it was understood that facial movement is just one of many forms of expressions.5 In addition to that, emotion recognition and sentiment analysis has been applied to multiple forms of data like text and speech.

Sentiment Analysis

Sentiment analysis helps computers to predict the emotions underlying a particular context. In its base form, sentiment analysis classifies whether a particular piece of content is positive, neutral, or negative. This is further developed by categorizing it into six basic emotions listed in BET and researchers are trying to detect and recognize more complex emotions.11 Numerous approaches were deployed in Sentiment Analysis in the field of text classification. At the same time, an overview of different approaches in Emotion AI-driven Sentiment Analysis as well as the steps involved in the process were also discussed.11 Furthermore, different approaches and techniques were compared and determined that Support Vector Machines (SVM) and Term-Frequency demonstrated the highest accuracy in the ontology-based approach.11 One of the most frequent applications of sentiment analysis is to get an overview of general sentiment towards a product or any tangible or intangible target. Tweets related to government policies regarding online learning was used to build a K-Nearest Neighbour model which obtained an accuracy of 84.93% by forming 10 clusters.12

Data Analysis is a significant discipline in the data era which provide lot of useful insights to grow and improve the businesses.13 Robust techniques and algorithms are generally used to perform sentiment analysis that provide a comprehensive overview of the results with a detailed analysis. Many optimization and regularization techniques are also used to fine tune the models and its results in the contemporary data science platform. Particle Swarm Optimization (PSO) was revealed to be one of the most frequently used optimization algorithms along with Support Vector Machine (SVM) for sentiment classification. However, literature evidances of a variety of optimization techniques and algorithms that can be applied in this area to more accurate and useful results.

Text is one of the most widely used communication tool for sharing opinions and thoughts especially on the internet. So, emotion recognition from text is a recent essential research area in Natural Language Processing (NLP) where many people live on the digital platform.14 The authors argue that the previous models adopted word embedding vectors to represent semantic and syntactic information which cannot infer the emotional relationship between words. However, a model was proposed named Semantic-Emotional Neural Network (SENN) which is an amalgamation of two sub-networks such as bidirectional Long-Short Term Memory (biLSTM) which grasps the context and semantic relationship, and a Convolutional Neural Network (CNN) which extracts the emotional features and understand the emotional relationship between words.14 SENN demonstrated improved performance compared to other state-of-the-art methods across a slew of text datasets with various contexts.

Facial Emotion Recognition

Facial Emotion Recognition (FER) is one of the most explored domains in Emotion AI. Basic Emotion Theory is the cornerstone for Facial Emotion Recognition where emotions are identified based on the facial expressions of a person. A comparison has been provided between exploring various conventional methods and recent approaches based on deep learning. There are various approaches in conventional facial emotion recognition, however, the fundamental idea behind these is detecting the face and extracting the geometric features, or a composite of geometric and appearance features.15 With the recent developments in machine learning and deep learning, these techniques have trickled down into the domains of Computer Vision (CV) and facial emotion recognition. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are the most used deep learning techniques used for computer vision as well as feature extraction, classification, and recognition. It is also mentioned that although the research behind facial emotion recognition has been conducted over the last decade, the performance of facial emotion recognition has improved drastically with the implementation of a combination of deep learning techniques.

Basic Emotion Theory asserted that facial expressions are universal. However, several researchers conducted studies to contradict this belief and many studies niched down to just one demographic

© UIJIR | ISSN (O) – 2582-6417
APRIL 2023 | Vol. 3 Issue 11
www.uijir.com

factor like age, gender, or race. The factors such as gender, age and race were used for one emotion (happiness) detection by observing the intensity of two peculiar Facial Action Units (FAUs or AUs), Cheek Raiser (AU6) and Lip Corner Puller (AU12) using a deep learning algorithm developed on a massive training set consisting of 75,000 images.16 It was concluded that females have a higher AU12 intensity than men, African Americans have a higher AU6 and AU12 intensity, and people above 40 display a stronger AU12 intensity.16

Covering faces has been considered a main obstacle for any facial recognition software, hence accessories like glasses, hats, etc. have been identified as obstecles too. The recent pandemic brought about many challenges for facial recognition systems as people had to cover their faces with masks which affected the facial emotion recognition. A study was conducted to quantify the effect of masks on emotion detection using data from 41 participants for 6 basic emotions across a population of varying ages and examined the accuracy of FER using apriori power test.17 Although the comparison demonstrated a decrease in accuracy, the magnitude of performance change differed from emotion to emotion as different emotions make use of different parts of the face and action units across the face. Happiness and Disgust displayed a significant drop in accuracy as mouth as a primary means to express these feelings. While on the other hand, fear and anger didn't see much difference in performance as these emotions mostly expressed through eyes. Moreover, different age groups exhibited differences in the degree of performance shift with the middle-ages being the least affected as compared to younger and senior participants.

Speech Analysis

In the last decade, with the invention and advancements of smart devices, a gradual change has been noticed on the way of interaction with these devices; from physical buttons to capacitive touch and finally touch screens. Accessibility and convenience have been significantly improved via a choice to interact with these devices using the voice too. Voice-controlled systems focus only on the keywords and understanding the content, Speech Analysis for Emotion AI is difficult as there are several vocal features that must be taken into consideration like the tone, pitch, loudness, and pauses between words and sentences.18

Voice is a rich medium of communication and expression as it conveys not only the content but a great deal of underlying context. Several studies have demonstrated and provided a meta-analysis about these studies via a quantitative review of several algorithms applied on culturally specific and cross-cultural datasets.19 The analysis obtained a mean accuracy of 89.7% for culturally specific dataset and a mean accuracy of 84.3% for cross-cultural datasets.19 Another significant approach proposed a new framework that could extract a host of features from audio files to utilize as the input for a 1-D Convolutional Neural Network for emotion recognition using numerous datasets like the TyersonAudio-Visual Database of Emotion Speech and Song (RAVDESS), Berlin (EMO-DB), and Interactive Emotional Dyadic Motion Capture (IEMOCAP).20 The proposed model obtained an accuracy of 71.61% with 8 classes on RAVDESS dataset and an accuracy of 64.3% with 4 classes on IEMOCAP dataset proving an outperformance than the other state-of-the-art models.20 Further, an accuracy of 95.71% with 7 classes was obtained on EMO-DB while proving the model can be used for a generalized scenario.20

Another emotion recognition approach was proposed from the speech where age and gender were used to support the emotion recognition task using hierarchical models where an accuracy of 64.2% was obtained by the base model, an accuracy of 65.1% was obtained by the age-based model and an accuracy of 70.59% was obtained by the age-gender model.21 Human speech depends on numerous factors; primarily it involves the combination of individual traits and the emotional states that determine how human expresses. A study proposed a 4-D conceptual framework that analyses the dimensions such as time, amplitude, frequency and spectral.18

A number of approaches were explored for Speech Emotion Recognition using deep learning algorithms such as the Deep Boltzmann Machine (DBM), Recurrent Neural Network (RNN), Recursive Neural Network (RvNN), Deep Belief Network (DBN), Convolutional Neural Network (CNN), and Auto Encoder (AE) to recognize emotions on various datasets, and provides a conceptual understanding of these techniques as well as summarizing numerous studies that implement variations of techniques and datasets.22 Many studies conducted around the demographic effects on expressions and emotions usually dealt with the observation of facial expressions.

ETHICAL CONSIDERATIONS IN AFFECTIVE COMPUTING

Affective Computing deals with recognizing and understanding emotions using information like facial expressions, voice patterns, etc. which poses a slew of ethical challenges to deal with. Emotions guide human behaviour and can be influenced by external factors. Hence, it is extremely critical to consider the ethical implications of any research, experiment, or implementation of affective computing.

In this digital era, the amount of new data and information generation is increasing exponentially year over year where the analysis of these data significantly supports digital economy. The personal information, thoughts, and opinions as well as medical data collected through health tracking apps and wearable devices are significant among these data. Healthcare and the retail industries are the primary industries that generates more data and perform rigorous analysis. It's quite crucial to understand the data of errors and biases, which could lead to the possibility of harming the people either medically or non-medically. Moreover, due to the lack of robustness in the security of this data, frequent data breaches can occur even within fully fledged analytical environments.

A study from 33 scientists unravelled the belief that discrete emotions may require different regulatory guidelines and processes, hence concluded that emotion and cognition have a brisk and instinctive correlation.20 There is still no consensus on the understanding of emotions and emotional models, that have considerable implications for systems that rely on these understandings and assertions. An idea of "Emotional Artificial Intelligence" was conceived which alludes to the growing number of AI systems that track, analyse and measure human expressions to determine their emotions.23 Furthermore, a set of guidelines around the ethical use of Emotional Artificial Intelligence were defined to enable companies to develop solutions that are acceptable in both legal and ethical considerations.1

Apart from the smart devices, the integration of technology and intelligent devices into toys and education methods meant for young children in the contemporary world. Although these devices aim to optimize the learning ability and assist their development, it is important to understand and address the possible issues and manipulations that might harm in some way. Ethical considerations and guidelines are usually strict when it comes to children, thus developers and companies must pay extreme attention to detail and ethics. A national survey conducted in UK to understand the opinions and perspectives about toys and child-oriented devices using emotion data from 1000 parents dealt with different perspectives and niches revealed 4 distinct themes: generational unfairness, parental vulnerability, guarded interest, and need for better governance.23 Additionally, some recommendations were shared reflecting the technology, parenting and governance landscape for regulatory bodies and toy companies.

CONCLUSION

Affective Computing is a domain of Artificial intelligence that deals with recognizing, understanding and simulating emotions. With the rapid advancement of Artificial Intelligence, it can replicate various aspects of humans like learning, vision and even voice. Emotion is an enormous obstacle that draws a clear distinction between humans and machines, but with the rise of Artificial Intelligence, emotion recognition is possible and viable. Defining emotion was a challenge although there is no clear definition. Two emotion models have been popular such as categorical, which divides emotions into discrete categories and dimensional where emotions are represented as a continuous value on a 2D place of arousal and valence in which emotion recognition approaches vary like understanding a general sentiment from a piece of text. New research have been successful in extracting or recognizing emotions through different ways of expression such as facial expression and vocal expression. The most critical aspect of Affective Computing is the ethical implication behind these systems. Emotions can reveal the intentions of a person and they guide our actions. Manipulating these can have a significant effect on the lives of people, hence there must be clear guidelines in place to avoid misuse and prevent any harm caused by these systems.

REFERENCES

1. Stark, L., & Hoey, J. (2021). The Ethics of Emotion in Artificial Intelligence Systems. 12. https://doi.org/10.1145/3442188.3445939
2. Gu, S., Wang, F., Patel, N. P., Bourgeois, J. A., & Huang, J. H. (2019). A model for basic emotions using observations of behavior in Drosophila. Frontiers in Psychology, 10 (APR), 781. https://doi.org/10.3389/FPSYG.2019.00781/BIBTEX
3. Izard, C. E. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. Emotion Review, 2(4), 363–370. https://doi.org/10.1177/1754073910374661
4. Poria, S., Majumder, N., Mihalcea, R., & Hovy, E. (2019). Emotion recognition in conversation: Research challenges, datasets, and recent advances. IEEE Access, 7, 100943–100953. https://doi.org/10.1109/ACCESS.2019.2929050
5. Keltner, D., Tracy, J. L., Sauter, D., & Cowen, A. (2019). What Basic Emotion Theory Really Says for the Twenty-First Century Study of Emotion. Journal of Nonverbal Behavior, 43(2), 195–201. https://doi.org/10.1007/S10919-019-00298-Y/TABLES/1
6. Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional Expression: Advances in Basic Emotion Theory. Journal of Nonverbal Behavior 2019 43:2, 43(2), 133–160. https://doi.org/10.1007/S10919-019-00293-3
7. Crivelli, C., & Fridlund, A. J. (2019). Inside-Out: From Basic Emotions Theory to the Behavioral Ecology View. Journal of Nonverbal Behavior 2019 43:2, 43(2), 161–194.

https://doi.org/10.1007/S10919-019-00294-2

8. Hutto, D. D., Robertson, I., & Kirchhoff, M. D. (2018). A new, better BET: Rescuing and revising Basic Emotion Theory. Frontiers in Psychology, 9(JUL), 1217. https://doi.org/10.3389/FPSYG.2018.01217/BIBTEX
9. Heaven, D. (2020). Why faces don't always tell the truth about feelings. Nature, 578(7796), 502–504. https://doi.org/10.1038/D41586-020-00507-5
10. Dzedzickis, A., Kaklauskas, A., & Bucinskas, V. (2020). Human emotion recognition: Review of sensors and methods. Sensors (Switzerland), 20(3). https://doi.org/10.3390/S20030592
11. Chakriswaran, P., Vincent, D. R., Srinivasan, K., Sharma, V., Chang, C. Y., & Reina, D. G. (2019). Emotion AI-driven sentiment analysis: A survey, future research directions, and open issues. Applied Sciences (Switzerland), 9(24). https://doi.org/10.3390/APP9245462
12. Isnain, A. R., Supriyanto, J., & Kharisma, M. P. (2021). Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 15(2), 121. https://doi.org/10.22146/ijccs.65176
13. Yadav, A., & Vishwakarma, D. K. (2020). A comparative study on bio-inspired algorithms for sentiment analysis. Cluster Computing, 23(4), 2969–2989. https://doi.org/10.1007/s10586-020-03062-w
14. Batbaatar, E., Li, M., & Ryu, K. H. (2019). Semantic-Emotion Neural Network for Emotion Recognition from Text. IEEE Access, 7, 111866–111878. HYPERLINK "https://doi.org/10.1109/ACCESS.2019.2934529"
15. Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. Sensors (Switzerland), 18(2). https://doi.org/10.3390/S18020401
16. Fan, Y., Lam, J. C. K., & Li, V. O. K. (2021). Demographic effects on facial emotion expression: an interdisciplinary investigation of the facial action units of happiness. Scientific Reports, 11(1). https://doi.org/10.1038/S41598-021-84632-9
17. Carbon, C. C. (2020). Wearing Face Masks Strongly Confuses Counterparts in Reading Emotions. Frontiers in Psychology, 11. https://doi.org/10.3389/FPSYG.2020.566886/FULL
18. Hildebrand, C., Efthymiou, F., Busquet, F., Hampton, W. H., Hoffman, D. L., & Novak, T. P. (2020). Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications. Journal of Business Research, 121, 364–374. https://doi.org/10.1016/J.JBUSRES.2020.09.020
19. Laukka, P., & Elfenbein, H. A. (2021). Cross-Cultural Emotion Recognition and In-Group Advantage in Vocal Expression: A Meta-Analysis. In Emotion Review (Vol. 13, Issue 1, pp. 3–11). SAGE Publications Ltd. https://doi.org/10.1177/1754073919897295
20. Issa, D., Fatih Demirci, M., & Yazici, A. (2020). Speech emotion recognition with deep convolutional neural networks. Biomedical Signal Processing and Control, 59, 101894. https://doi.org/10.1016/J.BSPC.2020.101894
21. Shaqra, F. A., Duwairi, R., & Al-Ayyoub, M. (2019). Recognizing Emotion from Speech Based on Age and Gender Using Hierarchical Models. Procedia Computer Science, 151, 37–44. https://doi.org/10.1016/J.PROCS.2019.04.009
22. Khalil, R. A., Jones, E., Babar, M. I., Jan, T., Zafar, M. H., & Alhussain, T. (2019). Speech Emotion Recognition Using Deep Learning Techniques: A Review. IEEE Access, 7, 117327–117345.

https://doi.org/10.1109/ACCESS.2019.2936124

23. McStay, A., & Rosner, G. (2021). Emotional artificial intelligence in children's toys and devices: Ethics, governance and practical remedies. Big Data and Society, 8(1). https://doi.org/10.1177/2053951721994877