

# Multi-Modal Emotion Recognition for Enhanced Requirements Engineering: A Novel Approach

Ben Cheng\*, Chetan Arora†, Xiao Liu\*, Thuong Hoang\*, Yi Wang\*, John Grundy†

\*School of Information Technology, Geelong, Victoria, Australia

{chengye, xiao.liu, thuong.hoang, xve}@deakin.edu.au

†Faculty of Information Technology, Monash University, Clayton, Victoria, Australia

{chetan.arora, john.grundy}@monash.edu

**Abstract**—Requirements engineering (RE) plays a crucial role in developing software systems by bridging the gap between stakeholders’ needs and system specifications. However, effective communication and elicitation of stakeholder requirements can be challenging, as traditional RE methods often overlook emotional cues. This paper introduces a multi-modal emotion recognition platform (MEMoRE) to enhance the requirements engineering process by capturing and analyzing the emotional cues of stakeholders in real-time. MEMoRE leverages state-of-the-art emotion recognition techniques, integrating facial expression, vocal intonation, and textual sentiment analysis to comprehensively understand stakeholder emotions. This multi-modal approach ensures the accurate and timely detection of emotional cues, enabling requirements engineers to tailor their elicitation strategies and improve overall communication with stakeholders. We further intend to employ our platform for later RE stages, such as requirements reviews and usability testing. By integrating multi-modal emotion recognition into requirements engineering, we aim to pave the way for more empathetic, effective, and successful software development processes. We performed a preliminary evaluation of our platform. This paper reports on the platform design, preliminary evaluation, and future development plan as an ongoing project.

**Index Terms**—Requirements Engineering, Requirements Elicitation, Emotions, Multi-modal Data Analysis.

## I. INTRODUCTION

Requirements Engineering (RE) is a critical phase in the software development lifecycle, responsible for capturing, documenting, and managing stakeholders’ needs to create a set of well-defined and comprehensive system requirements [1]. The effectiveness of RE directly impacts the quality and success of the resulting software system. RE involves complex human interactions, which in turn involve the emotions of all stakeholders. The emotional cues, if overlooked in RE, can often lead to misunderstandings and misinterpretations [2], [3].

Emotions are significant in human communication, influencing decision-making, problem-solving, and collaboration [3], [4]. In the context of RE, understanding stakeholders’ emotions can provide valuable insights into their needs, priorities, and expectations, ultimately leading to more effective communication and elicitation of requirements. Despite its importance, emotional awareness has not been extensively explored in the field of RE or software engineering in general [5].

Recent advances in Artificial Intelligence (AI) and machine learning have led to significant improvements in emotion recognition techniques in videos, audios, and texts [6]–[8]. These advancements have opened new opportunities to integrate emotion awareness into RE processes. However, there are also many practical challenges in using emotion recognition techniques in RE processes. For example, how to achieve real-time emotion recognition in various RE environments; how to generate useful emotion analysis result when stakeholders are experiencing intensive emotional changes; and how to protect user privacy since many personal videos and audios are being used by emotion recognition models. Therefore, the application of emotion recognition in RE processes is still in its infancy and there are many open issues to be addressed. This paper presents a novel multi-modal emotion recognition platform (MEMoRE) to enhance the RE process by capturing and analyzing emotional cues in stakeholder interactions. We plan to leverage MEMoRE in different RE activities, including stakeholder interviews (elicitation), change management (specification and validation), requirements review meetings (validation), and usability testing (validation).

The primary objectives of this paper are:

- 1) Introduce MEMoRE - a multi-modal emotion recognition platform that integrates facial expression, vocal intonation, and textual sentiment analysis to comprehensively understand stakeholder emotions during RE.
- 2) Discuss the potential benefits and challenges associated with incorporating emotion awareness into requirements engineering, emphasizing its impact on communication, elicitation strategies, and overall project success.
- 3) Present a preliminary evaluation of the proposed platform, highlighting its effectiveness in a real-world scenario and outlining a future evaluation plan.

The remainder of this paper is structured as follows: Section II reviews the related work. Section III presents the platform overview and its key components. Section IV presents the details of the preliminary evaluation results, discussions, and current limitations. Section V discusses the research roadmap, and Section VI concludes the paper.

## II. RELATED WORK

**Emotions in RE.** Requirements engineering (RE) focuses on capturing and analyzing user requirements, such as user feedback data. While user data often reflect user behavior and experience, there is a notable lack of research exploring the relationship between user behavior, experience, and emotions [9]. In recent years, some studies have explored human emotions' role in the RE process, such as requirements changes [10], [11] and sentiment analysis of user reviews [12], [13]. While these studies have presented the connection between human emotions and RE, emotions have not been extensively considered in the software development lifecycle. If emotions are not adequately considered, this can lead to misunderstandings of user needs, user dissatisfaction with the software product, and even uninstallation of the application [9]. Emotion-oriented requirements engineering aims to explore the impact of human emotions in various application scenarios of RE activities [14]. This paper presents the application of a multi-modal emotion recognition approach to enhance the accuracy of analyzing user emotional results and understanding the relationship between emotions and user requirements in RE activities.

**Detecting emotions.** Measuring emotions primarily includes physiological data, behavioral data, facial data, and self-report data [15]. For example, to address deliberately fabricated smiles, Perusque-Hernandez et al. [16] employed computer vision (CV) and facial distal electrocardiography (EMG) to identify different types of user smiles. The study ultimately demonstrated that CV could recognize user behavior that is not discernible by humans, while EMG can detect concealed behaviors that are not visible. However, most studies currently focus on the influence of a single data source on user requirements. Single-source emotion data frequently produce imprecise results. Nevertheless, a recent study focused on applying multi-source emotion recognition in RE activities [17]. Cheng et al. [17] reported that single-source emotion recognition is insufficient for providing comprehensive emotional analysis, which may hinder engineers' understanding of user emotions. They presented an edge computing-based multi-source emotion recognition platform (Edge4Emotion) that can analyze user emotions from various angles, such as facial expressions and audio data. Meanwhile, some research focuses solely on multi-modal emotion frameworks, models, and algorithms [6], [18]. Our approach is novel and can substantially benefit the RE community.

**Applications.** Currently, several studies mainly focus on evaluating the relationship between emotions and specific user groups (e.g., elderly) [19], [20], teams (e.g., agile teams) [11], and application scenarios (e.g., smart home systems, video games, and robot assistance) [14], [19], [21]. Human emotions are key in accepting software applications, especially in social systems. For example, Miller et al. [22] reported that adding the concept of emotional goals improved emergency systems, resulting in higher user satisfaction compared to earlier systems. Curumsing et al. [14] also reported an evaluation exam-

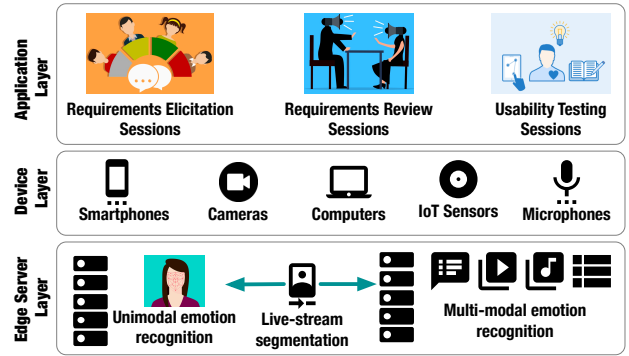


Fig. 1. MEmoRE Platform Overview

ple of employing an emotion-oriented approach to evaluate the SofiHub application. They found that the SofiHub application effectively mitigated the loneliness experienced by older users and made them feel safer and more cared for. Furthermore, emotions have a significant impact on software developers. For example, agile teams have significant emotional effects on developers during requirement changes. However, currently, no effective tools are available for detecting and analyzing developers' emotions [11], [23]. Therefore, there is a need for tools and techniques that can accurately detect and analyze emotions in the context of software development, which can improve the effectiveness and efficiency of software development and enhance user satisfaction.

In summary, the application of multi-modal emotion recognition has been unexplored in RE activities. Therefore, our study will be an important step in the direction of applying multi-modal emotion recognition in RE activities.

## III. MULTI-MODAL EMOTION RECOGNITION PLATFORM

In this section, we describe our multi-modal emotion recognition platform for RE (MEmoRE), the main components developed and planned, and the planned applications in RE.

### A. System Overview

The architecture of the system is illustrated in Figure 1. It can be categorized into three distinct layers: the application, the device and the edge server layers. The **device layer** encompasses a variety of devices such as cameras, smartphones, microphones and wearable devices, which can gather motion data and physiological data from multiple sources. These devices can connect to the edge servers through wired networks and wireless connections. The **edge server layer** comprises multiple edge servers capable of processing various types of data and supporting different emotion recognition models for uni-modal and multi-modal. For instance, the platform includes dedicated edge servers for recognizing emotions based on video and audio data separately. An important edge server for multi-modal emotion recognition processes both visual and audio data simultaneously. The **application layer** consists of various tools and applications catering to different RE activities. For example, the emotion recognition results for interviewees can be utilized in requirements gathering to

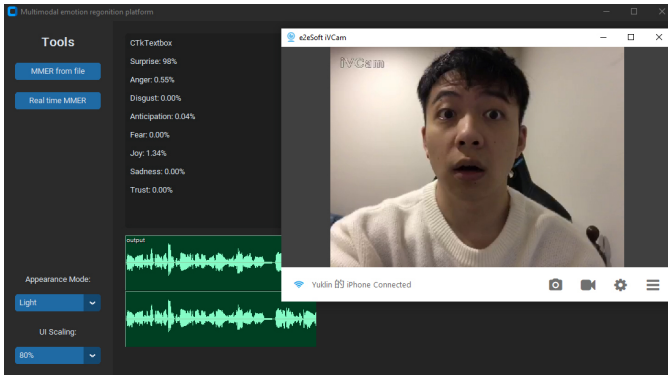


Fig. 2. Example UI and emotion recognition accuracy for video segments of various lengths.

understand their genuine opinions about specific requirements, while the results for testers in usability testing can help identify their actual experiences with certain functionalities. Further details about the various components of the MEMoRE platform are provided below. An example UI of the platform is shown in Fig. 2, also demonstrating the accuracy of emotion recognition for video segments. The process is as follows:

- 1) The end devices will capture the user's physiological signals which constantly change during a conversation. At the moment, for our multi-modal emotion analysis, we focus on facial expressions and vocal cues, so we will use cameras and microphones to capture the user's facial expressions and vocal changes during the conversation.
- 2) The data captured by the camera and microphone will be pre-processed by the video processing edge server. The video processing edge server will segment the data stream (e.g., every 10-seconds/30-seconds) and save these video segments in the MP4 video format. This facilitates data recording and manual review.
- 3) The edge server for multi-modal emotion recognition will extract features from the video and audio data separately, which will then be fed into a pre-trained deep learning model. Ultimately, the predicted emotional state will be generated as the output for further usage by various RE activities.
- 4) Our approach involves incorporating the output results into our requirements analysis. This enables us to identify the emotional state of the target user during the requirements gathering or testing phase, helping software engineers to focus on the users' specific requirements and improve the design accordingly.

## B. Major Components

**Application Layer.** MEMoRE platform can support different RE activities, including requirements elicitation, validation, and so on. Requirements elicitation is an important step for stakeholders in the software development process. However, traditional requirements elicitation techniques can limit their ability to dig into stakeholders' true needs and desires. Using the emotion recognition approach enhances the requirements elicitation process, which better feeds the expectation of stakeholders. With our multi-modal emotion recognition platform,

we can easily analyze stakeholders' emotions in real-time discussion environments to identify important requirements. When stakeholders express high positive emotions during a discussion, it might indicate that particular requirements are essential to them. In contrast, if stakeholders exhibit confusion or frustration, it often suggests that our engineers may not have understood their requirements correctly. We need to reconsider the proposed solution or clarify the requirements further. In this way, our platform can help facilitate more effective communication and collaboration between stakeholders and engineers, leading to the identification of requirements that better address the needs and desires of stakeholders. Furthermore, emotion recognition can prioritize requirements based on their emotional impact. For instance, requirements related to user satisfaction or emotions such as joy, trust, or confidence can be given higher priority than others. This prioritization can ensure that requirements with top importance are given priority during the development process.

Requirements validation is essential to ensure that they accurately reflect the stakeholders' needs and expectations. This can be done by observing stakeholders' emotional responses to prototypes or demos. It is a common practice to provide prototypes to stakeholders which allows engineers to gather feedback on the design, functionality, and overall user experience. During the demonstration or the usage of the prototype, stakeholders may reveal various emotional responses. For example, positive emotions such as surprise and joy indicate that the software's functionality has met or exceeded the expectation of users. Conversely, if a user expresses disgust or frustration, it may indicate that the software does not meet the user's expectations or is difficult to use. Our platform can automate this process and generate reports that provide engineers with detailed records of emotional responses to specific requirements so that engineers can adjust the development direction and improve the user experience. Furthermore, emotion-based requirement validation can be integrated into the requirement engineering process to enhance the quality of requirements and ensure their alignment with stakeholders' emotional needs and preferences.

**Device Layer.** The end device layer in our platform supports the connection of various types of devices and sensors for emotion data collection to collect real-time emotion data. The data is then transmitted to different edge servers for emotion recognition. Cameras provide live-stream video data and microphones provide audio data which supports different edge servers to perform emotion recognition. Moreover, smartphones can serve as integrated devices that facilitate the collection of multi-modal data, such as video and audio, through embedded cameras and microphones. Overall, the end device layer is an essential component of our platform as it allows for the integration of various devices for emotion data collection. By utilizing a combination of different sensors and devices, our platform can collect multi-modal data to provide a more accurate, reliable and comprehensive emotion analysis.

**Edge Server Layer.** This layer consists of various computing

nodes near the network edge (namely edge servers) for processing data transmitted directly from the end devices. Our novel approach focuses on the MEmoRE server to provide emotion fusion results from human facial expression, human voice and text. Specifically, it first segments the real-time video and audio streams into small clips with fixed lengths as input for a pre-trained multi-modal emotion recognition model, and then outputs the probabilities for eight different emotions as classified in Joy, Sadness, Anger, Anticipation, Disgust, Fear and Trust. For multi-modal emotion recognition in our platform, we utilized the model proposed in [24] which employ a deep Visual Audio Attention Network (VAANet). VAANet is a novel architecture that integrates spatial, channel-wise, and temporal attentions into a visual 3D convolutional neural network (CNN) and temporal attention into an audio 2D CNN. In addition, our platform also supports uni-modal emotion analysis, specifically for image and audio modalities. Each of these modalities can be handled by a dedicated edge server that utilizes deep learning models for emotion recognition. In fact, applying different uni-modal emotion recognition methods under specific requirements engineering environments can be beneficial. For example, if the environment involves a lot of phone conversations, using an audio-based emotion recognition method would be more appropriate. In contrast, if the environment involves video conferencing or surveillance footage, a video-based emotion recognition method would be more useful. Furthermore, incorporating text-based emotion recognition methods could be more useful for requirements engineering environments that involve a lot of written communication, such as email or chat-based communications.

#### IV. PRELIMINARY EVALUATION

##### A. Preliminary Results

Here, we demonstrate the preliminary results of the experimentation on MEmoRE. We divided the experiment into two steps. First, we aimed to determine the optimal length of segments, in seconds, for achieving the highest accuracy when segmenting the interview data. Second, we aimed to evaluate the emotion results of the interview data, using MEmoRE. In the following, we describe the details of the experiments.

**(1) Experiment 1:** We assessed the accuracy of different video segment lengths using eight minute long interview videos. The video format was MP4. To ensure uniformity of the interview video data, we set the frame rate to 24 frames per second (24 fps), which is a common video standard. Next, we segmented the video into different intervals of 6-seconds, 10-seconds, 15-seconds, 30-seconds and 60-seconds. Fig. 3 shows the accuracy results for videos segmented into different lengths. We found that the lowest accuracy was achieved when then interview video data was segmented into 60 seconds. If the interview video data was segmented into 6-seconds, 15-seconds and 30-seconds, the accuracy was below 50%. After multiple validations, we discovered that the highest accuracy was achieved only when the video was segmented into 10 seconds. With longer videos, e.g., 60 seconds, the

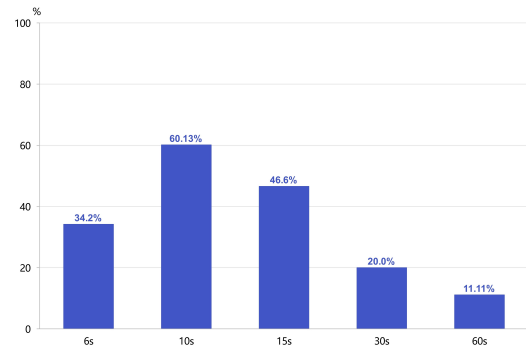


Fig. 3. Accuracy of emotion recognition at different segment lengths.

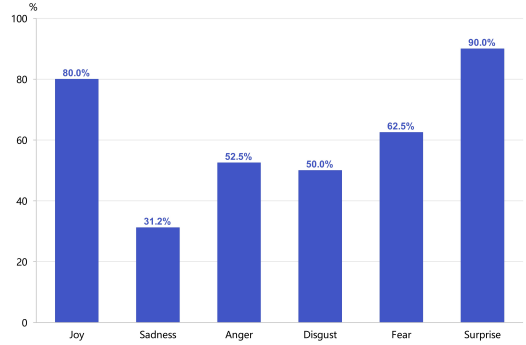


Fig. 4. Accuracy of emotion detection for different emotions

platform is inaccurate potentially due to multiple emotions in the video and since the platform is currently tuned to provide a single classification. With shorter videos, there is potentially insufficient content to accurately judge emotions.

**(2) Experiment 2:** We aimed to discover the model’s performance against different emotions. In this experiment, we selected 48 videos with six different emotions and eight videos for each type of emotion from the MELD dataset [25]. Fig. 4 shows the accuracy results of the model recognizing different emotions. The results showed that the accuracy of detecting positive emotions is high compared to negative emotions. For instance, the model performs well for detecting joy and surprise, whereas, not accurately at detecting sadness, anger and disgust. We plan to train further and improve our platform on more data and improve the accuracy for all emotions.

##### B. Discussion

**The Importance of Video Segmentation.** It is well known that human emotions play a significant role in requirements engineering activities. This also means that low accuracy in emotion recognition could result in incorrect decisions and analysis by engineers. In this work, it was observed that the highest accuracy is achieved using a 10-second interval for real-time video segmentation. However, this could be varied under different RE environments. Hence, we suggest to test and to determine the best video segmentation interval to improve the accuracy of real-time emotion analysis results.

This may assist requirement engineers in capturing complete conversations and better understanding end-users needs. Moreover, it was noted that using a 6-second interval might be too short for some conversations and reactions to complete. It is very challenging to accurately detect human emotions using a very short segment with limited information. For example, a person may cry out of extreme joy instead of sadness, but this emotion can be easily misunderstood if the video segment is very short. However, using much longer segments such as 30-second and 60-second intervals might also not yield accurate results as the conversation content and user's emotional state can change multiple times, making it difficult to determine the true emotion and hence less useful for RE activities.

**The Importance of Emotion Recognition Accuracy in RE Activities.** Human emotions present a complex process in RE activities and the environments for various RE activities could also be very different. Therefore, relying on single-source emotional analysis may be insufficient to accurately capture users' emotions. In this work, we propose the MEMoRE platform that identifies users' emotions through multi-modal combination of facial expressions, vocal intonation, and textual sentiment. The results of the second experiment demonstrated (Figure 3) that the highest accuracy rates are observed for surprise and joy in the real-time interview video. In addition, the MEMoRE platform also detected some negative emotions, such as disgust, fear, and sadness. These negative emotions are also important in RE activities as they can help engineers identify whether stakeholders have negative emotions, and if so, engineers can further investigate the cause for them.

In summary, here we present some potential contributions of the MEMoRE platform to the RE community as follows:

- 1) Multi-modal emotion recognition can assist requirement engineers in validating requirements, gauging stakeholder engagement, and ensuring that the software product meets stakeholders' expectations [26].
- 2) By identifying users' positive and negative emotions, requirement engineers can gain a better understanding of how the software performs and where needs improvement.
- 3) Our MEMoRE platform can automate the process for emotion detection and analysis. It can assist requirement engineers in efficiently determining requirement priorities based on stakeholders' emotional responses.
- 4) Stakeholders' emotional responses can guide requirements engineers in selecting more effective elicitation methods, and help them better understanding stakeholders' needs and expectations.

### C. Current Limitations

There are still some limitations in our current work. Although our research indicates that segmenting every 10 seconds produces better results, in actual RE activities, the conversation interval is not always fixed. It is possible to switch between two or more emotions in seconds, increasing the probability of errors. Furthermore, the current design of MEMoRE includes text, video, and audio modalities. There is

still bias due to individual differences in facial expressions. To address this issue, we plan to incorporate more modalities. Moreover, we have not tested our platform for actual RE activities in the presented experiments. To address this, we plan to conduct user studies with university students, as detailed in the following section V.

## V. RESEARCH ROADMAP

### A. Evaluation Plan

**Recruit Participants.** We aim to recruit participants in information technology, computer science and design courses who have completed or are current students in tertiary education. To ensure valuable data from participants' feedback, we require that they are currently in the final year of their tertiary courses. We plan to recruit  $\approx 25$  students for this experiment. We further plan to conduct case studies with practitioners (from our industry partners) in healthcare, e-commerce and app development domains for different RE tasks, e.g., elicitation [1], negotiations [26], change management [27].

**Experimental Design.** We plan to deploy the experiment in software projects of student teams. The experiment time will be  $\approx 45$  minutes. We will ask the participant groups to discuss their team software project progress in  $\approx 15$  minutes. We plan to collect and analyze the participants' emotions during this time. After the discussion, we will give the participants their emotion recognition results during the interviews in  $\approx 10$  minutes. Participants can discuss these results together. For example, if some participants were confused during the interviews, they could focus on the negative emotion in the last 15 minutes. After completing the experiment, participants must fill out a five-point Likert scale questionnaire. The questionnaire aims to collect feedback on multi-modal emotion recognition's positive and negative aspects. Throughout the experiment process, we will record video and audio data for analysis and ensure the confidentiality of user data. Regarding practitioner studies, we aim to analyze recorded requirements elicitation and review sessions and perform a retrospective analysis of MEMoRE results with practitioners. This will focus on assessing the accuracy of MEMoRE and implications of real-time emotion recognition.

### B. Future Work

We plan to adopt a conversational segmentation method instead of the 10-second segmentation to address the limitations. This involves segmenting based on sentences, which allows for a more holistic recognition of an individual's emotions and addresses the limitation of the time-based segmentation method. As ongoing work, the platform will be used as an open-source project to host future studies. Here, we also provide research directions in the near future.

- 1) **Using MEMoRE to support requirements changes.** Requirements changes represent a significant challenge in software development, especially agile development. Based on multi-modal emotion recognition, we can explore users' real emotions, thus, determining their true needs and feelings.

- 2) **Expanding the platform to different methods.** There are other requirements elicitation methods besides interviews [28]. For example, using virtual reality (VR) to capture user needs; however, there is no prior work on stakeholder needs and emotions based on VR [29]. Therefore, we plan to deploy MEMoRE within a VR environment and explore user emotions and needs.
- 3) **Supporting special user groups and different types of application scenarios.** In the future, we plan to focus on the emotions of special user groups such as elderly, children and people with disabilities (e.g., speech impaired users). We also plan to support deployment in various scenarios, such as requirements elicitation workshops with various number of stakeholders. Overall, our platform will cover special user groups in different types of application scenarios.

## VI. CONCLUSION

Human emotions play a significant role in RE activities, as they are an important factor in effective communication and eliciting stakeholder needs. However, the application of emotion recognition in the RE process is still in its infancy despite the success of many deep learning based emotion recognition methods. In this paper, we presented the details of our MEMoRE platform which can assist requirements engineers in better understanding user responses towards software requirements through accurate detection of user emotions. Experimental results demonstrated that our MEMoRE platform can accurately analyze real-time interview video data and using a 10-second segmentation interval to achieve the highest accuracy. We also discussed the contributions and limitations of our MEMoRE platform for RE activities. Finally, as an ongoing research project, we have designed an evaluation plan and proposed a series of future work.

## ACKNOWLEDGEMENTS

Grundy is supported by ARC Laureate Fellowship FL190100035.

## REFERENCES

- [1] A. Aurum and C. Wohlin, *Engineering and managing software requirements*. Springer, 2005, vol. 1.
- [2] I. Ramos and D. M. Berry, "Is emotion relevant to requirements engineering?" *Requirements Engineering*, vol. 10, 2005.
- [3] R. Colomo-Palacios, A. Hernández-López, Á. García-Crespo, and P. Soto-Acosta, "A study of emotions in requirements engineering," in *WSKS 2010*, 2010.
- [4] A. M. Isen, "An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications," *Journal of consumer psychology*, vol. 11, no. 2, 2001.
- [5] M. Sánchez-Gordón and R. Colomo-Palacios, "Taking the emotional pulse of software engineering—a systematic literature review of empirical studies," *Information and Software Technology*, vol. 115, 2019.
- [6] D. Hazarika, R. Zimmermann, and S. Poria, "Misa: Modality-invariant and-specific representations for multimodal sentiment analysis," in *28th ACM International Conference on Multimedia*, 2020.
- [7] K. Zamani, D. Zowghi, and C. Arora, "Machine learning in requirements engineering: A mapping study," in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*. IEEE, 2021.
- [8] R. Das and T. D. Singh, "Multimodal sentiment analysis: A survey of methods, trends and challenges," *ACM Comput. Surv.*, mar 2023.
- [9] M. Stade, S. A. Scherr, P. Mennig, F. Elberzhager, and N. Seyff, "Don't worry, be happy – exploring users' emotions during app usage for requirements engineering," in *2019 IEEE 27th International Requirements Engineering Conference (RE)*, 2019.
- [10] K. Madampe, R. Hoda, and J. Grundy, "A framework for emotion-oriented requirements change handling in agile software engineering," *IEEE Transactions on Software Engineering*, 2023.
- [11] M.-A. A. El-Migid, D. Cai, T. Niven, J. Vo, K. Madampe, J. Grundy, and R. Hoda, "Emotimonitor: A trello power-up to capture and monitor emotions of agile teams," *Journal of Systems and Software*, vol. 186, p. 111206, 2022.
- [12] E. Guzman and W. Maalej, "How do users like this feature? a fine grained sentiment analysis of app reviews," in *2014 IEEE 22nd international requirements engineering conference (RE)*. Ieee, 2014.
- [13] M. Bano, C. Arora, D. Zowghi, and A. Ferrari, "The rise and fall of covid-19 contact-tracing apps: when nfirs collide with pandemic," in *29th International Requirements Engineering Conference (RE)*. IEEE, 2021.
- [14] M. K. Curumsing, N. Fernando, M. Abdelrazek, R. Vasa, K. Mouzakis, and J. Grundy, "Emotion-oriented requirements engineering: A case study in developing a smart home system for the elderly," *Journal of Systems and Software*, vol. 147, pp. 215–229, 2019.
- [15] H. L. Meiselman, *Emotion measurement*. Woodhead publishing, 2016.
- [16] M. Perusquía-Hernández, S. Ayabe-Kanamura, K. Suzuki, and S. Kumano, "The invisible potential of facial electromyography: A comparison of emg and computer vision when distinguishing posed from spontaneous smiles," ser. CHI '19, 2019.
- [17] B. Cheng, Y. Wang, D. Shao, C. Arora, T. Hoang, and X. Liu, "Edge4emotion: An edge computing based multi-source emotion recognition platform for human-centric software engineering," in *21st International Symposium on Cluster, Cloud and Internet Computing*, 2021.
- [18] T. Mittal, U. Bhattacharya, R. Chandra, A. Bera, and D. Manocha, "M3er: Multiplicative multimodal emotion recognition using facial, textual, and speech cues," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 02, 2020.
- [19] A. Demaeght, C. Miclau, J. Hartmann, J. Markwardt, and O. Korn, "Multimodal emotion analysis of robotic assistance in elderly care," ser. PETRA '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 230–236.
- [20] G. Boatendal and T. Kowatsch, "Speech emotion recognition among elderly individuals using multimodal fusion and transfer learning," in *International Conference on Multimodal Interaction*, ser. ICMI '20 Companion, New York, NY, USA, 2021, p. 12–16.
- [21] D. Callele, E. Neufeld, and K. Schneider, "Balancing security requirements and emotional requirements in video games," in *2008 16th IEEE International Requirements Engineering Conference*, 2008.
- [22] T. Miller, S. Pedell, A. A. Lopez-Lorca, A. Mendoza, L. Sterling, and A. Keirnan, "Emotion-led modelling for people-oriented requirements engineering: The case study of emergency systems," *Journal of Systems and Software*, vol. 105, pp. 54–71, 2015.
- [23] A. Kołakowska, A. Landowska, M. Szwoch, W. Szwoch, and M. R. Wróbel, "Emotion recognition and its application in software engineering," in *2013 6th International Conference on Human System Interactions (HSI)*, 2013.
- [24] S. Zhao, Y. Ma, Y. Gu, J. Yang, T. Xing, P. Xu, R. Hu, H. Chai, and K. Keutzer, "An end-to-end visual-audio attention network for emotion recognition in user-generated videos," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 01, 2020.
- [25] S. Poria, D. Hazarika, N. Majumder, G. Naik, E. Cambria, and R. Mihalcea, "Meld: A multimodal multi-party dataset for emotion recognition in conversations," *arXiv preprint arXiv:1810.02508*, 2018.
- [26] A. Ferrari, T. Huichapa, P. Spoletini, N. Novielli, D. Fucci, and D. Girardi, "Using voice and biofeedback to predict user engagement during requirements interviews," *arXiv preprint arXiv:2104.02410*, 2021.
- [27] C. Arora, M. Sabetzadeh, A. Goknil, L. C. Briand, and F. Zimmer, "NARCIA: an automated tool for change impact analysis in natural language requirements," in *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering*, 2015.
- [28] O. Karras, S. Kiesling, and K. Schneider, "Supporting requirements elicitation by tool-supported video analysis," in *2016 IEEE 24th International Requirements Engineering Conference (RE)*. IEEE, 2016.
- [29] Y. Wang, B. Cheng, T. Hoang, C. Arora, and X. Liu, "Virtual reality enabled human-centric requirements engineering," in *36th International Conference on Automated Software Engineering Workshops (ASEW)*. IEEE, 2021.