Ergonomics in Design, Vol. 129, 2024, 192–202
https://doi.org/10.54941/ahfe1004824

AI-Powered Emotion Recognition for Maritime Safety Enhancement

Mario Ivan Zignego[1], Paolo Andrea Gemelli[1], Alessandro Bertirotti[1], and Laura Pagani[2]

[1]DAD, Università Studi di Genova, 16100 Genoa, Italy
[2]PhD Research Group in Marine Science and Technology at the University of Genoa, 16100 Genoa, Italy

ABSTRACT

This paper delineates an initiative to devise an emotion recognition system bespoke for nautical settings, employing artificial intelligence (AI) and an array of biosensors. The system will integrate speech emotion recognition (SER) technology to discern emotions such as anger, fear, joy, and sorrow through voice patterns. Additionally, biosensors like skin conductance (EDA/GSR), electrocardiography (ECG), and eye-tracking technology will collect further physiological data. The aggregated data will be analysed to assess the emotional state of the individuals involved. A virtual mock-up, simulating a maritime environment complete with weather dynamics and potential equipment malfunctions that could induce stress or anxiety, will be created. This virtual environment will endeavour to achieve a high degree of realism to heighten the immersion of participants in the simulated scenario. AI algorithms will adjust this environment in real time in response to the user's emotional state to enhance the user experience (UX). For example, should the system detect stress or anxiety, the virtual scenario may transition to a calming seascape to encourage relaxation and alleviate stress. Future developments will see the incorporation of a real-time alarm mechanism capable of notifying essential personnel, such as the ship's captain or supervisor, upon detection of emotional states that warrant attention. This alert system is designed as a preventive measure, aiming to avert more serious consequences like accidents, errors, or more grave incidents. The system will also proffer recommendations for intervention or actions based on the detected emotional state. The project's goal is to introduce an avant-garde approach to monitoring and supporting the mental health and safety of maritime professionals. It aspires to detect emotional fluctuations instantaneously, provide interventions, and adjust the environment to promote relaxation and minimise stress. The project is expected to yield practical implications for the maritime industry and other high-risk work environments where the mental health and safety of workers are paramount. In conclusion, the project's outcomes are projected to offer practical benefits for the maritime sector and other high-risk occupations, ensuring the monitoring and support of workers' mental health and safety.

Keywords: Multimodal emotion recognition, Maritime safety and ergonomics, AI in high-risk environments

© 2024. Published by AHFE Open Access. All rights reserved.

INTRODUCTION

Human factors are a critical component in the maritime sector, serving as a keystone for ensuring safety and operational efficiency in shipping operations. The field has long been challenged by issues emanating from human error, with fatigue, stress, and emotional upheaval significantly contributing to maritime mishaps. Such incidents, which often have serious repercussions, underline the pressing demand for innovative solutions to manage and mitigate risks associated with human elements.

This paper elucidates the results of a seminal research project carried out by the doctoral research team in Marine Science and Technology at the University of Genoa, an effort greatly enhanced by the partnership with the Genoa Merchant Marine Academy. At the heart of our venture is the creation of an advanced multimodal system designed to detect and decode the emotional states of individuals on maritime vessels. Our aim is to pioneer advancements not only in the sphere of shipbuilding ergonomics but also in improving the user experience (UX) during sea journeys. Our holistic strategy takes into account a variety of environmental factors, from the colour palette to auditory and olfactory stimuli, which are pivotal in influencing the crew's emotions and overall well-being at sea.

We have successfully developed a multimodal system capable of effectively applying facial expression analysis and speech recognition to ascertain emotional states. The system incorporates sophisticated algorithms that process these human cues, providing essential insights into the emotional welfare of personnel aboard. Our ambitions, however, are broader; we foresee the integration of an extensive range of communicative modalities, including body language, postures, and a collection of biometric sensors for gathering data such as electroencephalograms (EEG), electrocardiograms (ECG), body temperature, and eye movement metrics. These additional modes of communication are expected to considerably enhance our understanding of human factors within maritime contexts.

In addition, we are in the process of developing a cutting-edge augmented and virtual reality platform. This element of our project promises to revolutionise the way we simulate and appraise our ergonomic and UX designs, affording us the opportunity to trial and refine them within virtual realms. This innovative method is set to improve product design in a cost-effective and risk-free manner, ensuring that the end products not only meet but exceed the expectations of end-users.

From a commercial vantage point, this system has the potential to redefine the maritime industry. By proactively identifying and addressing conditions such as stress, fatigue, and discomfort, our system is capable of significantly diminishing risks and enhancing safety and efficiency on board. In a domain where human error can lead to severe consequences, the facility to manage the well-being of crew members is of paramount importance. Furthermore, the system offers substantial benefits for staff recruitment and training, enabling the identification of traits conducive to maritime roles, as well as those that may be detrimental. This discernment is vital to ensure the selection and preparation of the most apt candidates for these demanding roles.

To summarise, this research signifies a notable leap forward in the amalgamation of human factors engineering with the latest technology in the maritime industry. Our interdisciplinary approach, merging expertise from marine science, technology, psychological science, and design, seeks to set new standards in the enhancement of maritime occupational and living standards. We hope that this paper will contribute meaningfully to the discourse on applied human factors and ergonomics, particularly in relation to high-risk sectors such as maritime operations.

LITERATURE REVIEW

The growing field of multimodal emotion recognition, particularly through the integration of systems like the Facial Action Coding System (FACS) and the Geneva Minimalistic Acoustic Parameter Set (GeMAPS), is revolutionizing the way we understand human emotions in various settings, including maritime environments. These systems offer comprehensive methods to analyse complex emotional states through facial and vocal cues, which are vital in environments where understanding human emotions is crucial for safety and efficiency (Tang et al., 2014).

The Facial Action Coding System (FACS) is pivotal in interpreting subtle changes in facial expressions related to various emotional states. Developed by Ekman and Friesen, FACS provides a detailed taxonomy of all human facial movements and has been adapted for use in automated systems for emotion recognition (Ekman and Friesen, 1978). Its application is critical in maritime settings where accurate interpretation of crew emotions can inform user-centric design and safety measures.

In parallel, the Geneva Minimalistic Acoustic Parameter Set (GeMAPS), proposed by Eyben et al. (2016), offers a minimalistic yet effective set of vocal parameters for emotion recognition. GeMAPS is essential for analysing vocal patterns indicative of stress, fatigue, or other emotional states, contributing significantly to operational safety in maritime environments.

Integrating these powerful tools in a multimodal approach enhances the reliability and accuracy of emotion detection systems. The use of such systems in maritime contexts, where understanding and responding to crew emotions is key, can lead to improved user experience and proactive incident prevention. These advanced emotion recognition systems can inform ergonomic ship designs and development of safety protocols, ensuring a safer and more efficient maritime industry (Eyben et al., 2016).

In summary, the application of multimodal emotion recognition systems using FACS and GeMAPS is showing promising results in the maritime sector. Their use in real-time monitoring of crew emotions can lead to significant advancements in safety standards and operational procedures, aligning with the evolving needs of the industry.

METHODOLOGY

Our methodology (Fig. 1) centred on an empirical study conducted with a cohort of 31 participants (26 males and 5 females) during their selection interviews for the officer cadet course at the Genoa Merchant Marine Academy. Prior to the study, we ensured ethical compliance by securing informed consent from all individuals, explicitly stating the research objectives and the use of the recordings. The data collection involved filming the interviews, which were then meticulously edited into two-minute clips. This editing process was crucial to isolate the audio from the visual data, allowing for independent processing. The video clips retained facial expressions, while the audio clips captured vocal nuances and speech patterns. To process these data streams, we employed four emotion detection systems (Morphcast, HumeAI, OpenFace, and an internally developed code) based on advanced machine learning and deep learning frameworks.

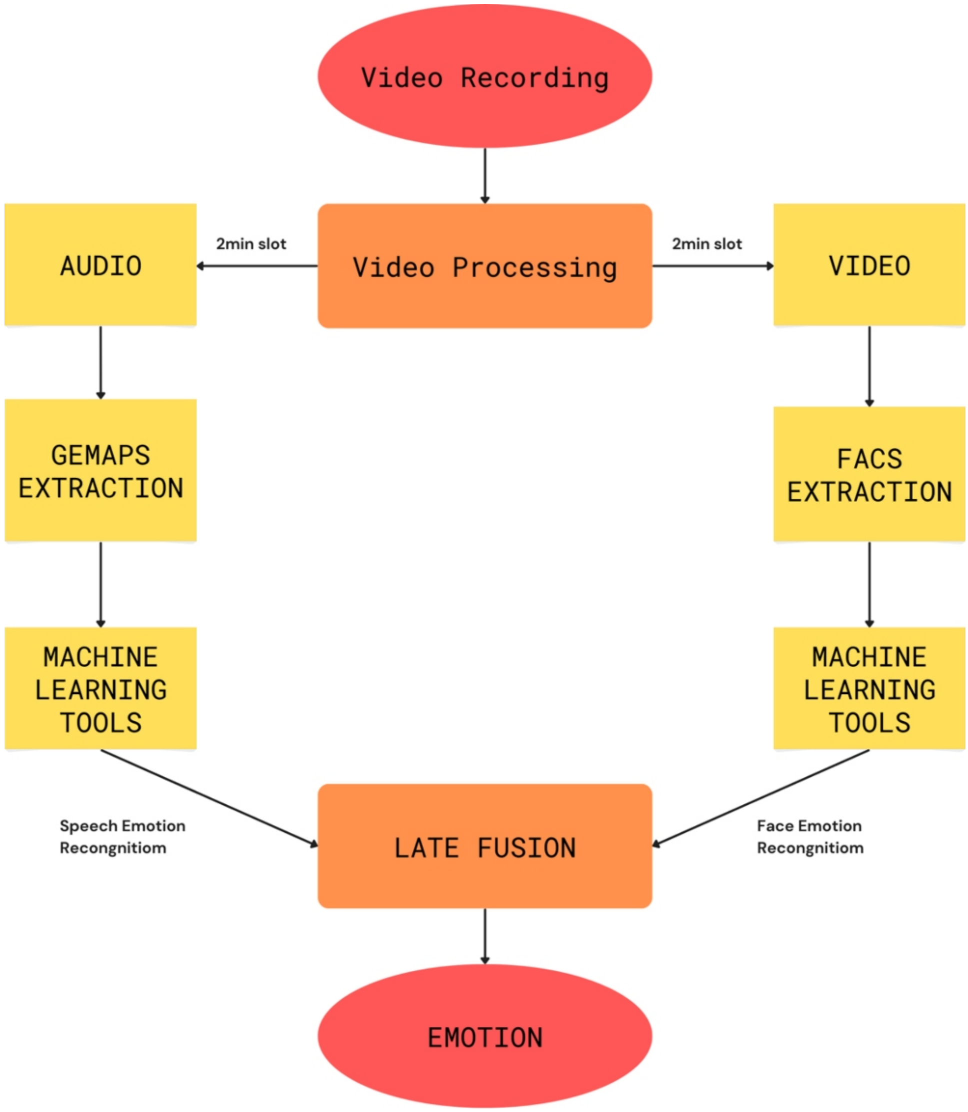

Figure 1: Our methodology.

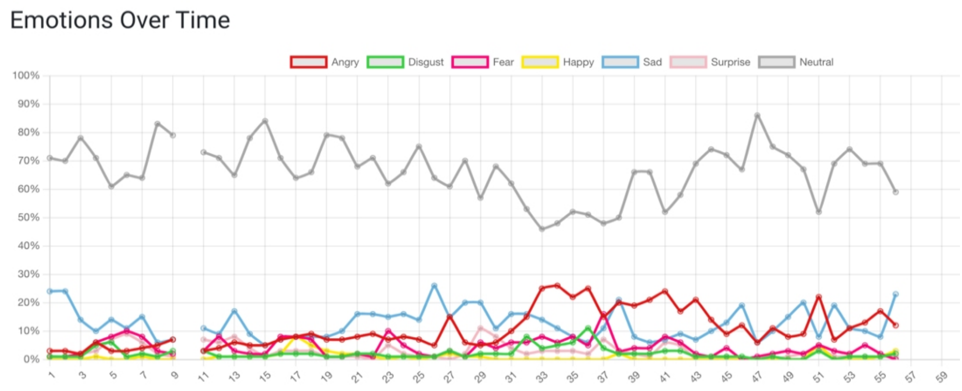

Figure 2: Morphcast FER analysis.

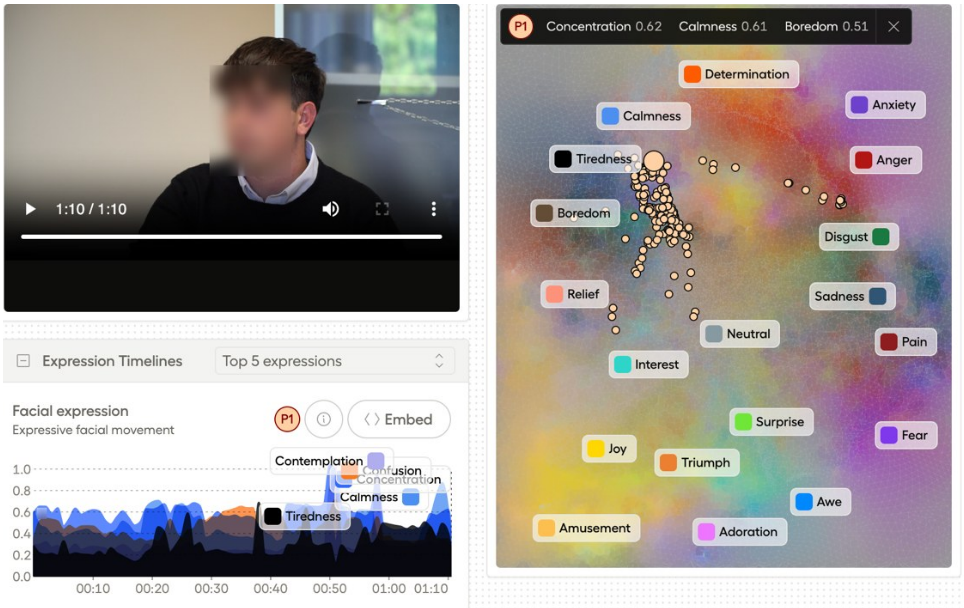

Figure 3: HumeAI FER analysis.

These systems were thoroughly compared to pinpoint and analyse a range of emotional states extracted from the gathered data. The following stage of our evaluation entailed a side-by-side review of the outcomes obtained from the video and audio feeds. The intent of this juxtaposition was to determine the consistency with which both channels recognized identical emotions (Fig. 2, 3, and 4). We fortified our methodological rigour by incorporating a post-interview element, where participants were requested to complete a structured questionnaire (Fig. 5, 6).

This self-report measure captured their perceived emotional experiences during the interview process, providing a subjective data point for later correlation with our objective findings.

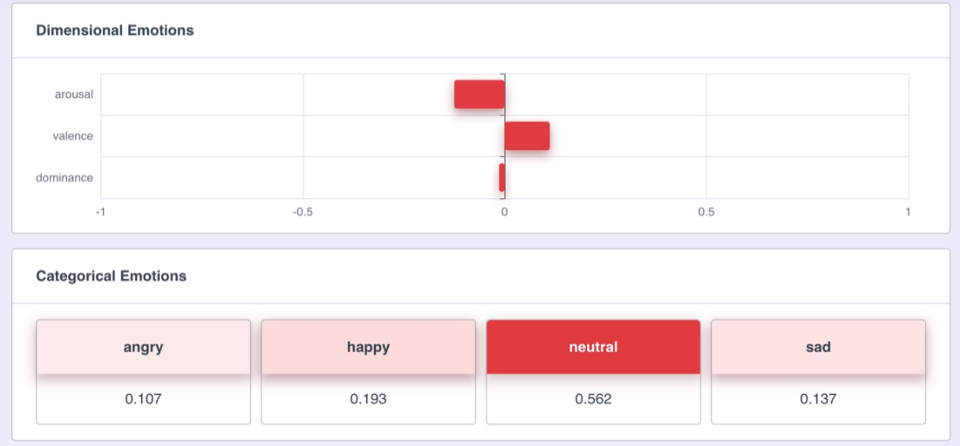

Figure 4: OpenSMILE SER analysis.

HOW DID YOU FEEL?
COME TI SEI SENTITO DURANTE L'ESERCITAZIONE?

English	Italiano	Per niente: 0	Un po': 1	Abbastanza: 2	Molto: 3	Moltissimo: 4	Non so
Afraid	Paura						
Amused	Divertito						
Angry	Arrabbiato						
Annoyed	Infastidito						
Anxious	Ansioso						
Apathetic	Apatico						
Aroused	Eccitato						
Astonished	Stupito						
Bored	Annoiato						
Calm	Calma						
Conceited	Presuntuoso						
Contemplative	Contemplativo						
Content	Contenuto						
Convinced	Convinto						
Delighted	Incantato						
Depressed	Depresso						
Determined	Determinato						
Disappointed	Deluso						
Discontented	Scontento						
Distressed	Angosciato						
Embarrassed	Imbarazzato						
Enraged	Infuriato						
Excited	Eccitato						
Feel Well	Sentirsi bene						
Frustrated	Frustrato						
Happy	Contento						
Hopeful	Speranzoso						
Impressed	Impressionato						
Melancholic	Malinconico						
Peaceful	Tranquillo						
Pensive	Pensieroso						
Pleased	Lieto						
Relaxed	Rilassato						
Sad	Triste						
Satisfied	Soddisfatto						
Tired	Stanco						
Uncomfortable	Scomodo						
Worried	Preoccupato						

Figure 5: Questionnaire.

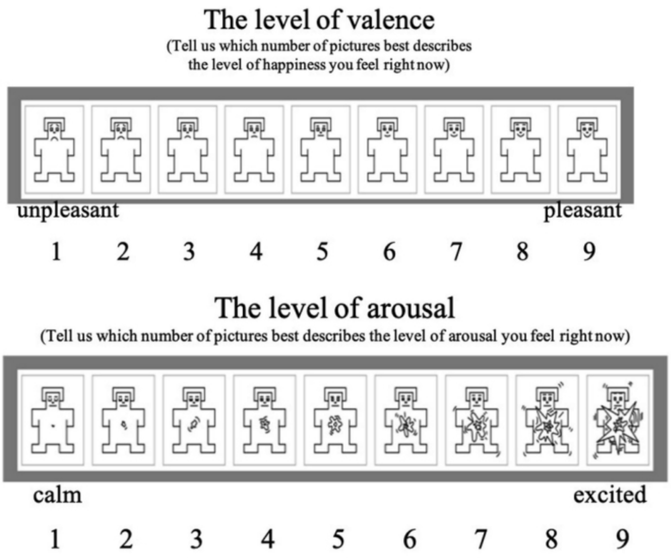

Figure 6: Arousal-Valence emotion questionnaire – self assessment manikin.

Additionally, we engaged an expert in cognitive anthropology to perform an independent analysis of the video recordings. The expert's role was to offer a professional interpretation of the emotional states displayed, which served as a benchmark for our AI-driven analysis.

In the final stage of our methodology, we triangulated the results from the AI algorithms, the candidates' self-assessments, and the expert anthropologist's evaluations. This comprehensive comparison was instrumental in validating the efficacy of our emotion recognition system. While the system demonstrated considerable promise as a novel tool in the domain of emotion recognition, it was acknowledged that it remained in a developmental stage, with potential for further refinement and optimisation.

PROGRESS AND FINDINGS

The research embarked upon an iterative process with the primary aim of validating a multimodal emotion recognition system in a maritime educational setting. The initial phase involved data collection from 31 prospective officer cadets at the Genoa Merchant Marine Academy, which yielded a rich dataset for preliminary analysis.

The editing and segmentation of the interview footage into two-minute intervals facilitated focused analysis on short, emotionally significant interactions. The isolated audio streams were processed using advanced acoustic analysis tools designed to pick up on subtle changes in tone, pitch, and speech tempo, which could indicate underlying emotional states.

In parallel, the visual data underwent a separate analysis using facial recognition algorithms capable of detecting micro-expressions and other non-verbal cues. The initial findings from the video data suggested a high incidence of stress-related expressions, a finding consistent with the high-pressure context of a selection interview.

Upon integrating the audio and video analysis, we observed a moderate to high correlation between the emotions detected by each modality, with both channels independently identifying stress, concentration, and moments of relief. These moments often aligned with the content and context of the interview questions, validating the relevance of our multimodal approach.

The self-reported emotional states from the participants' questionnaires provided a subjective backdrop to our objective measures. Interestingly, there was a notable variance in the self-assessment of stress levels, with some participants reporting lower stress levels than what was indicated by our AI system. This discrepancy highlights the complexity of self-awareness in high-stress situations and underscores the value of objective emotional recognition systems.

The expert anthropologist's input further corroborated our AI findings, particularly in recognising nuanced emotional responses that the candidates exhibited. This expert validation was instrumental in refining our algorithms to be more sensitive to the subtleties of human emotion in high-stakes environments.

In summary, our findings indicate that the multimodal emotion recognition system has significant potential for application in maritime settings. The system's ability to detect a range of emotions with a high degree of correlation between objective measures and expert analysis (up to 76% of accuracy) sets the stage for its use in enhancing maritime safety by monitoring and addressing the emotional well-being of seafarers.

DISCUSSION

The analysis of the multimodal emotion recognition system has yielded insights that are both affirming and enlightening. The moderate to high correlation between the emotions detected via audio and video channels provides a strong case for the system's efficacy. Notably, the AI algorithms were able to detect stress and concentration, emotions typically prevalent in the high-pressure environment of selection interviews. These findings align with existing literature that underscores the robustness of multimodal systems in emotion recognition tasks (Kuncheva et al., 2001; Sebe et al., 2005).

The variance observed between the AI-detected emotions and participants' self-reported emotional states opens an intriguing discourse on the perception of emotions. This divergence could suggest a potential lack of self-awareness or an unwillingness to report certain emotions, a finding that resonates with studies on affective forecasting and emotional intelligence (Gross, 2002; Hoerger et al., 2010). It highlights the necessity of objective measures in settings where accurate emotion recognition is crucial for safety and performance.

The anthropologist's expert analysis, largely corroborating the AI-detected emotions, also introduces an interesting dimension to the system's validation. The human expert's ability to discern subtleties in emotion aligns with the nuances our AI system is designed to detect, which suggests a complementary relationship between human expertise and AI capabilities. This duality is vital in designing systems that are not only technologically advanced but also deeply rooted in the understanding of human behaviour (Picard, 1997; Cowie et al., 2001).

Despite the promising results, the system's current limitations must be acknowledged. The research has thus far been conducted in a controlled environment, which, while beneficial for initial validation, does not fully replicate the complexities of a maritime setting. Future iterations of the system will require testing in more dynamic, real-world scenarios to ensure robustness and reliability.

Moreover, the ethical implications of emotion recognition technology, particularly regarding privacy and consent, necessitate careful consideration. As this technology progresses toward real-world applications, it is imperative to establish clear guidelines and practices that safeguard individual rights (McStay, 2018).

In conclusion, the progress made with the multimodal emotion recognition system sets a foundational stone for further research. The potential applications of such a system in maritime environments could revolutionise safety protocols and training processes, making it an invaluable asset in the industry. As we move towards operationalising this technology, a balanced approach that respects ethical boundaries while striving for technological innovation will be paramount.

CONCLUSION AND FUTURE WORK

The study has successfully demonstrated the potential of a multimodal emotion recognition system to enhance the understanding of emotional states in high-stress maritime settings. The correlation between the AI's emotion detection and the expert anthropologist's assessments provides a compelling argument for the system's viability. However, it is the discrepancy between these objective measures and the self-reported emotions that underlines the complexity and necessity of this technology in environments where self-assessment may be unreliable or insufficient.

As we look to the future, our research will transition from the controlled conditions of the academy to the unpredictable and dynamic environment of active maritime operations. This natural progression is essential to validate the system's efficacy in real-world conditions, where factors such as sea state, weather, and operational pressures come into play. The inclusion of additional biometric sensors will aim to capture a more comprehensive physiological profile, adding layers of data to enhance the accuracy of emotion recognition.

The development of an augmented and virtual reality platform represents the next frontier for our research. This will enable us to simulate maritime scenarios with high fidelity, providing a safe and controlled space for further testing and refinement of the system. Moreover, these virtual environments will facilitate the evaluation of ergonomic and UX designs, potentially revolutionising the way ships are built and operated.

From an ethical standpoint, the advancement of our research will be guided by a commitment to respect privacy and consent. As we refine the technology, we will engage with stakeholders, including seafarers, shipping companies, and regulatory bodies, to develop a framework that ensures the responsible use of emotion recognition systems.

Commercially, the implications of our work are vast. The system holds promise for enhancing the recruitment and training processes within the maritime industry, potentially leading to improved safety outcomes and operational efficiency. By identifying stress, fatigue, and other adverse emotional states, the system could inform targeted interventions, contributing to a safer and more productive maritime workforce.

In summary, the research thus far has laid a solid foundation for a transformative tool in maritime safety and ergonomics. Our continued efforts will focus on technological advancement, ethical considerations, and practical applications, with the ultimate goal of establishing a new benchmark in the industry for emotion recognition and response.

ACKNOWLEDGMENT

The authors would like to acknowledge the Merchant Marine Academy in Genoa for their kind collaboration.

REFERENCES

Bartlett, M. S., Littlewort, G., Frank, M. G. and Lee, K. (2014) 'Automatic Decoding of Facial Movements Reveals Deceptive Pain Expressions', Current Biology, 24(7), pp. 738–743.

Cohn, J. F. and De La Torre, F. (2014) 'Automated Face Analysis for Affective Computing', in Calvo, R. A., D'Mello, S. K., Gratch, J. and Kappas, A. (eds.) The Oxford Handbook of Affective Computing. Oxford University Press.

Ekman, P. and Friesen, W. V. (1978) Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press.

Eyben, F., Weninger, F., Gross, F., and Schuller, B. (2016) "The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing," IEEE Transactions on Affective Computing.

Gross, J. J. (2002). "Emotion regulation: Affective, cognitive, and social consequences." Psychophysiology, 39(3), 281–291.

Hoerger, M., Chapman, B. P., Epstein, R. M., & Duberstein, P. R. (2010). "Emotional intelligence: A theoretical framework for individual differences in affective forecasting." Emotion, 10(3), 412–421.

Kuncheva, L. I., Bezdek, J. C., Duin, R. P., & Whitaker, C. J. (2001). "Decision templates for multiple classifier fusion: An experimental comparison." Pattern Recognition, 34(2), 299–314.

McStay, A. (2018). "Emotional AI: The rise of empathic media." SAGE.

Picard, R. W. (1997). "Affective Computing." MIT Press.

Sebe, N., Cohen, I., Gevers, T., & Huang, T. S. (2005). "Emotion recognition based on joint visual and audio cues." Pattern Recognition, 18(10), 2051–2056.

Schuller, B., Steidl, S. and Batliner, A. (2011) 'The Interspeech 2011 Speaker State Challenge', in Proceedings of the 12th Annual Conference of the International Speech Communication Association.

Tang, K., Tie, Y., Yang, T., and Guan, L. (2014) "Multimodal Emotion Recognition (MER) System," CCECE 2014.

Zhang, Z., Girard, J. M., Wu, Y., Zhang, X., Liu, P., Ciftci, U.,... Morency, L.-P. (2016) 'Multimodal Spontaneous Emotion Corpus for Human Behavior Analysis', in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.