Quantifying human factor in ships operation via an AI based emotions recognition

Mario Ivan Zignego*, Paolo Gemelli*, Alessandro Bertirotti*, Laura Pagani*

***University of Genoa, Genoa, Italy**, mario.ivan.zignego@unige.it, paolo.gemelli@edu.unige.it, laura.pagani@edu.unige.it, alessandro.bertirotti@edu.unige.it

Abstract

Human factor evaluation via emotion recognition from face micro-expressions and speech analysis holds significant value in the field of shipping and logistics. Understanding the emotional state and engagement levels of individuals involved in shipping operations can greatly impact efficiency, customer satisfaction, and overall success in the industry.

Emotions recognition from face micro-expressions plays a crucial role in assessing the well-being and mental state of shipping personnel. The ability to detect micro-expressions can help identify signs of stress, fatigue, frustration, or other emotional states that may affect their performance. By recognizing these emotions, shipping companies can take proactive measures to address issues, provide the necessary support, and ensure the well-being of their employees. This can lead to improved job satisfaction, reduced turnover rates, and increased productivity.

Speech analysis also plays a significant role in human factor evaluation within the shipping industry. Effective communication is essential for successful shipping operations, and analyzing speech patterns can provide insights into the emotional state, engagement, and clarity of individuals involved. By analyzing the tone, pitch, and speed of speech, shipping companies can identify signs of confusion, frustration, or lack of engagement. This information can be used to optimize training programs, improve communication protocols, and enhance overall operational efficiency.

Moreover, emotion recognition and speech analysis can be used in customer service within the shipping industry. Understanding the emotions and satisfaction levels of customers through their facial expressions and speech can enable shipping companies to tailor their services and address any concerns promptly. This personalized approach enhances customer experience, strengthens relationships, and ultimately leads to customer loyalty and repeat business.

While utilizing emotion recognition and speech analysis in the shipping industry offers numerous benefits, ethical considerations must be taken into account. Ensuring the privacy and consent of individuals involved is of utmost importance. Companies must handle personal information responsibly and transparently, adhering to legal and ethical standards. Additionally, biases and discrimination risks associated with emotion recognition technology should be acknowledged and addressed to ensure fair and equitable use.

In conclusion, human factor evaluation via emotion recognition from face micro-expressions and speech analysis holds substantial value in the shipping industry. By understanding and interpreting emotions, shipping companies can enhance employee well-being, improve communication, optimize operations, and provide exceptional customer service. However, ethical guidelines must be followed to protect privacy, avoid biases, and ensure the responsible use of these technologies in the shipping sector.

Keywords: Human factor, performance assesment

1. Introduction

Human factors are a major contributor to accidents in the shipping industry. Fatigue, stress, and other emotional factors can all lead to errors in judgment and decision-making, which can have serious consequences.

In recent years, there has been a growing interest in using artificial intelligence (AI) to quantify human factors in ships operation. AI-based emotions recognition technology can be used to track the emotional state of seafarers, and to identify patterns that may be indicative of fatigue, stress, or other potential problems.

This information can then be used to take preventive measures, such as rotating crew members, providing breaks, or offering counseling. By quantifying human factors in this way, AI can help to improve safety and efficiency in the shipping industry.

Here are some of the benefits of using AI-based emotions recognition in ships operation:

- Early identification of potential problems: AI can be used to track the emotional state of seafarers in real time, which can help to identify potential problems early on. This can give shipowners and operators a chance to take preventive measures before an accident occurs.
- Improved decision-making: By understanding the emotional state of seafarers, AI can help to improve decision-making in critical situations. This can help to prevent accidents and improve safety.
- Increased efficiency: AI can be used to optimize crew schedules and workloads, which can help to reduce fatigue and improve efficiency.

Overall, AI-based emotions recognition has the potential to significantly improve safety and efficiency in the shipping industry. By quantifying human factors in this way, AI can help to prevent accidents and make the shipping industry safer for everyone involved.

Here are some of the challenges of using AI-based emotions recognition in ships operation:

- Data collection: In order to train an AI model to recognize emotions, a large dataset of labeled data is needed. This data can be difficult and expensive to collect, especially in the shipping industry.
- Accuracy: The accuracy of AI-based emotions recognition models is still relatively low. This means that there is a risk of false positives and false negatives, which can lead to incorrect decisions being made.
- Acceptance: There is some resistance to the use of AI in the shipping industry, due to concerns about privacy and safety. This will need to be addressed in order for AI-based emotions recognition to be widely adopted.

An additional challenge is the need to train a neural network. This means having a large amount of data that is necessary to train the network itself, making it suitable for the purpose for which it is designed.

Despite these challenges, the potential benefits of using AI-based emotions recognition in ships operation are significant. As the technology continues to develop, it is likely that these challenges will be overcome. In the future, AI-based emotions recognition could become a standard tool for improving safety and efficiency in the shipping industry.

2. Defining the user stress level via a tensor

A user's stress level can be characterized by a tensor of measurable variables. A tensor is a mathematical object that generalizes scalars, vectors, and matrices to higher dimensions. Scalars are numbers, vectors are lists of numbers, and matrices are two-dimensional arrays of numbers. Tensors can be thought of as multidimensional arrays of numbers. In machine learning, tensors are used to represent data, such as images and videos. The values of these variables can be used to create a tensor to gauge worker stress level.

The accuracy of the classification depends on the quality of the data that is used to create the vector. The more data that is available, the more accurate the classification will be. However, even with a small amount of data, it is often possible to get a good estimate of the user's emotional state.

The tensor of measurable variables can also be used to track the user's emotional state over time.

These variables can include:

- Acoustic features: These features measure the physical properties of speech, such as pitch, loudness, and duration. Pitch is the frequency of a sound wave, and it is often associated with emotions such as happiness and excitement. Loudness is the amplitude of a sound wave, and it is often associated with emotions such as anger and fear. Duration is the length of a sound, and it is often associated with emotions such as sadness and boredom.

- Lexical features: These features measure the words that are used in speech, as well as the way that they are used. Some of the most commonly used lexical features for emotion detection include the words that are used, the order in which they are used, and the way that they are used. For example, the word "happy" is often associated with positive emotions, while the word "sad" is often associated with negative emotions. The order in which words are used can also be important. For example, the phrase "I am happy" is more likely to be associated with positive emotions than the phrase "I am sad."
- Physiological features: These features measure the user's physiological state, such as heart rate, blood pressure, and skin conductance. Some of the most commonly used physiological features for emotion detection include heart rate, blood pressure, and skin conductance. Heart rate is the number of times that the heart beats per minute, and it is often associated with emotions such as anxiety and fear. Blood pressure is the force of blood pushing against the walls of the arteries, and it is often associated with emotions such as anger and stress. Skin conductance is the electrical conductivity of the skin, and it is often associated with emotions such as excitement and fear.
- Behavioral features: These features measure the user's behavior, such as facial expressions, body language, and eye movements. Facial expressions are often used to communicate emotions, and they can be very effective at conveying emotions such as happiness, sadness, anger, and fear. Body language can also be used to communicate emotions, and it can be very effective at conveying emotions such as confidence, nervousness, and boredom. Eye movements can also be used to communicate emotions, and they can be very effective at conveying emotions such as interest, boredom, and surprise.

The combination of acoustic, lexical, physiological, and behavioral features can be used to create a very accurate and reliable representation of a user's emotional state. This information can then be used to improve a wide range of applications, such as healthcare, customer service, and education.

3. Technologies

Technology has the potential to revolutionise the field of human factor assesment by providing new tools for measuring and analysing the impact of stressful situations on workers' performances.
The principal technologies used can be summarized as follows.

3.1. Speech Emotion Recognition (SER)

Speech emotion recognition (SER) is the process of identifying the emotional state of a person based on their speech.
There are two main approaches to SER: acoustic-based and lexical-based.

- Acoustic-based SER systems analyze the acoustic features of speech, such as pitch, loudness, and duration. These features are then used to train a classifier that can identify the emotional state of the speaker.
- Lexical-based SER systems analyze the lexical content of speech, such as the words that are used and the way that they are used. These features are then used to train a classifier that can identify the emotional state of the speaker.

Acoustic-based SER systems are typically more accurate than lexical-based SER systems. This is because acoustic features are more closely linked to emotions than lexical features. However, acoustic-based systems can be more difficult to train and they are not as robust to changes in the speaker's voice. Lexical-based SER systems are easier to train and they are more robust to changes in the speaker's voice. However, lexical-based systems are typically less accurate than acoustic-based systems.

In recent years, there has been a growing interest in using deep learning for SER. Deep learning models have been shown to be very effective at learning complex patterns in data. This makes them well-suited for SER, which is a complex task that requires the ability to learn subtle patterns in speech.

There are a number of challenges that need to be addressed in order to improve the accuracy of SER systems. One challenge is the variability of human speech. The way that people speak can vary depending on a number of factors, such as their gender, age, and dialect. This variability can make it difficult to train SER systems that are accurate for all speakers.

Another challenge is the lack of large, well-labeled datasets. SER systems require large datasets of speech that have been labeled with the emotional state of the speaker. However, these datasets can be difficult to obtain.

Despite these challenges, SER is a promising field of research with a wide range of potential applications. As SER systems become more accurate, they will be able to play an increasingly important role in a variety of applications.

3.2. Biometric technologies

Eye-tracking

Eye-tracking technology has become increasingly popular as it provides valuable insights into how people perceive and interact with their environment, because it measures the movements of the eyes and it records how people perceive, react, and navigate through different sites.

There are different types of eye-tracking systems:

- remote systems which use cameras to track the movement of the eyes from a distance;
- head-mounted systems which use glasses or headsets that are equipped with cameras to track eye movement from a closer distance.

Both systems use infrared or corneal reflection methods to track the movement of the eyes.

Another application of eye-tracking technology is to measure cognitive load, or the amount of mental effort required to perform a task.
One limitation of eye-tracking technology is that it only measures visual attention and does not provide information on other sensory modalities, such as touch or smell. Additionally, eye-tracking data can be influenced by factors such as fatigue, cognitive load, and individual differences in eye movement patterns.

Electroencephalography (EEG)
EEG is a non-invasive technique that measures the electrical activity of the brain using electrodes placed on the scalp. It has been used for many years to diagnose neurological disorders and to study brain function, while more recently, it has been used to study emotional responses to the built environment, since it can measure different types of brain waves, including alpha, beta, delta, and theta waves. These waves are associated with different mental states, such as relaxation, arousal, and attention. By measuring changes in brain waves, researchers can identify emotional responses to different environmental stimuli.

Functional Magnetic Resonance Imaging (fMRI)
Functional Magnetic Resonance Imaging (fMRI) is a non-invasive technique that uses magnetic fields to measure changes in blood flow in the brain. Similarly to EEG; fMRI has been used extensively to study brain function and to diagnose neurological disorders, but it can be used to identify areas of the brain that are activated in response to different environmental stimuli, such as colours, sounds, taste, etc.

Heart Rate Variability (HRV)
Heart Rate Variability (HRV) is a measure of the variation in time between consecutive heartbeats. It is a non-invasive measure that has been used to study the autonomic nervous system, which controls functions such as heart rate, blood pressure, and digestion. More recently, it has been used to study emotional responses to the built environment.
HRV can be used to identify changes in the autonomic nervous system that are associated with different emotional states.

3.3. Behavioural technologies

We can quantify stress level using:

- head pose: The position and orientation of the head can be used to indicate stress. For example, a person who is feeling stressed may tend to hold their head down or to the side.

- facial expressions: Facial expressions can also be used to indicate stress. For example, a person who is feeling stressed may tend to frown, squint, or clench their jaw.
  Facial Action Coding System (FACS) is a tool that is widely used in neuroscience research to objectively measure and analyse facial expressions.
  Unlike subjective ratings of facial expressions, which can be influenced by factors such as observer bias or individual differences in perception, FACS provides a standardized and objective method for measuring facial movements, because it uses a series of codes to describe the movements of different facial muscles, such as raising the eyebrows, wrinkling the nose, or pulling the lips back. By coding facial movements in this way, researchers can objectively quantify and compare facial expressions across individuals and across different experimental conditions with greater accuracy and precision. FACS has been used in a wide range of neuroscience research, including studies on emotion, social interaction, and decision-making. It has been used to study the neural mechanisms underlying emotion regulation by measuring changes in facial expressions in response to emotional stimuli and it is also useful to study the neural basis of social interaction by measuring facial expressions during cooperative and competitive interactions.
  Additionally, FACS only measures facial expressions and does not provide information on other modalities, such as physiological responses or subjective ratings of emotional experience.
- Body language: Body language can also be used to indicate stress. For example, a person who is feeling stressed may tend to fidget, cross their arms, or avoid eye contact.
- Hand gestures: Hand gestures can also be used to indicate stress. For example, a person who is feeling stressed may tend to tap their fingers, clench their fists, or rub their temples.

4. Case Study

In the experiments conducted at the Merchant Marine Academy and the University of Genoa, two groups of 15 and 6 students, respectively, were monitored during the performance of activities that could induce significant levels of stress. In the first case, it was an exercise on electronic cartography that was a prerequisite for taking the final exam. In the case of the group at the University of Genoa, on the other hand, the students had to present the final project of a course followed during the year. In both cases, the participants were filmed during the entire activity and the videos were analyzed using two different software based on a convolutional neural network (CNN) and supervised machine learning algorithms. The data obtained were compared with those from a questionnaire (Perceived Stress Scale) previously completed by the participants. The comparison between the questionnaire and the results of the algorithms showed profiles of stress compatible with the activity performed.

7. Conclusions

The automatic analysis of the stress level during the operations of the crews of naval vessels represents a significant challenge in many ways. From a technological point of view, the development of artificial intelligence algorithms, combined with that of sensors, allows today to achieve extremely promising results. Nevertheless, significant efforts are still needed in the development of systems whose effectiveness is comparable, if not superior, to that of humans. Considering the great interest and the rapid technological development, it is legitimate to expect a great development of the sector in relatively short times.

References

Cohen, S., Kamarck, T., and Mermelstein, R. (1983). “A global measure of perceived stress.” Journal of Health and Social Behavior, 24, 386-396.

Cohen, S. and Williamson, G. (1988). “Perceived Stress in a Probability Sample of the United States”. Spacapan, S. and Oskamp, S. (Eds.) The Social Psychology of Health. Newbury Park, CA: Sage.

Gemelli Paolo, Zignego Mario Ivan, Bertirotti Alessandro. (2022). “A Smart Mockup for an Innovative Interior Yacht Design Approach”. In NAV 2022: 20th International Conference on Ship & Maritime Research

Giannakakis, Giorgos, Dimitris Grigoriadis, Katerina Giannakaki, Olympia Simantiraki, Alexandros Roniotis, and Manolis Tsiknakis. (2022). “Review on Psychological Stress Detection Using Biosignals.” IEEE Transactions on Affective Computing 13 (1): 440–60. https://doi.org/10.1109/TAFFC.2019.2927337.

Zhang, Kaihao, Yongzhen Huang, Yong Du, and Liang Wang (2017). “Facial Expression Recognition Based on Deep Evolutional Spatial-Temporal Networks.” IEEE Transactions on Image Processing 26 (9): 4193–4203. https://doi.org/10.1109/TIP.2017.2689999.

Zignego Mario Ivan, Bertirotti Alessandro, Gemelli Paolo, Pagani Laura. (2023). “Technology and neuroarchitecture”. In: IDEA investigatimg design in architecture.