Advancements in Emotion and Gesture Recognition Using Support Vector Machines (SVM): A Review Study

Shewale Sunita Bhiwsan
Research Scholar

Suhas Rajaram Mache
Computer Science, Department of Research, University of Technology, Jaipur, Rajasthan.

Corresponding Author Email: SUNITA.SHEWALE@Gmail.Com

ABSTRACT

The recognition of emotions and gestures plays a critical role in human-computer interaction (HCI), enabling more intuitive, empathetic, and natural communication between humans and machines. As technology advances, systems capable of understanding human emotional states and non-verbal cues, such as facial expressions and hand gestures, have numerous applications. These range from virtual assistants and customer service interfaces to mental health monitoring and assistive technologies for individuals with disabilities. One of the most powerful tools for emotion and gesture recognition is Support Vector Machines (SVM), which excel in handling high-dimensional data and can operate effectively with limited datasets. Through analyzing multimodal inputs like facial expressions, speech signals, and physiological measurements, SVM-based systems can achieve accurate classification of emotions and gestures, which enhances the user experience across various domains. This paper explores the importance of emotion and gesture recognition in HCI, the application of SVM in these systems, and the integration of advanced techniques to improve recognition accuracy and scalability.

Keywords*: *Emotion Recognition, Gesture Recognition, Support Vector Machines (SVM)

1. Introduction

The recognition of human emotions and gestures has become a crucial aspect of human-computer interaction (HCI), enabling machines to interpret and respond to human behaviours in a more intuitive and empathetic manner. This has numerous applications, from enhancing virtual assistants and customer service interfaces to improving mental health monitoring systems and developing assistive technologies for individuals with disabilities. The challenge, however, lies in the complexity and variability of human emotions and gestures, which are influenced by a multitude of factors such as cultural differences, personality traits, and contextual variables. Therefore, effective emotion and gesture recognition systems require sophisticated models that can accurately identify these expressions from raw data. In recent years, advancements in machine learning, particularly in the field of Support Vector Machines (SVM), have significantly improved the accuracy and reliability of emotion and gesture recognition systems. SVM, a supervised learning algorithm, has emerged as a powerful tool due to its ability to handle high-dimensional data and its robustness in classification tasks, particularly in scenarios with limited data. SVM works by finding an optimal hyperplane that separates data points into distinct classes, making it particularly effective for classification tasks such as recognizing different emotions or gestures based on physiological, facial, and behavioural data. The application of SVM to emotion recognition typically involves analysing data from multiple modalities, such as facial expressions, speech signals, and physiological measurements (e.g., heart rate, EEG signals). Facial expressions, for instance, are one of the most common indicators of emotional states. However, facial recognition alone may not provide sufficient accuracy due to variations in individual expressions and the subtleties of different emotional states. To address this, multimodal approaches have been developed, where data from multiple sources are combined to provide a more comprehensive understanding of the emotional state of an individual. SVM has proven particularly effective in these multimodal systems, as it can handle the complexities of integrating and classifying data from diverse sources. Gesture recognition, on the other hand, focuses on the identification of hand movements, body posture, and other non-verbal cues that convey meaning. In gesture recognition systems, SVM is used to classify different gestures based on features extracted from video frames or sensor data (such as accelerometer or gyroscope readings). These systems have applications in areas such as sign language interpretation, virtual reality (VR), and human-robot interaction, where accurate gesture recognition is essential for seamless communication. The integration of SVM with other advanced techniques, such as feature extraction methods, neural networks, and deep learning, has further enhanced the performance of emotion and gesture recognition systems. Feature extraction methods, like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), help reduce the dimensionality of the data while preserving the essential characteristics needed for classification. These techniques, combined with SVM's powerful classification capabilities, have led to significant improvements in recognition accuracy. Moreover, with the increasing availability of large datasets and advances in computational power, SVM-based emotion and gesture recognition systems have become more efficient and scalable. The development of benchmark datasets, such as the EmoReact and Facial Expression Recognition (FER) datasets, has provided a standard for evaluating the performance of these systems, allowing researchers to fine-tune their models for better generalization across different populations and contexts.

2. Related Reviews

Singh and Patterson (2010) aimed to develop a system using accelerometers to identify involuntary motions in premature infants, which are linked to cerebral palsy, specifically cramped-synchronized general movements. They collected data from ten infants hospitalized at the University of California, Irvine Medical Center's neonatal critical care unit to test the system. The methodology involved applying machine learning algorithms to features extracted from the accelerometer data, achieving accuracies between 70% and 90%. The accuracy varied depending on the cost of false positives and false negatives. To establish the ground truth, the researchers used validated video observation annotations. The study also involved investigating the foundation of the algorithm's predictions to understand its decision-making process. The findings demonstrated the potential of using machine learning for early detection of cerebral palsy in premature infants, highlighting the relevance of this approach for improving neonatal care and diagnosis accuracy.

Jerritta, S., Murugappan, M., Nagarajan, R., and Wan, K. (2011) conducted a study in the field of Human-Computer Interaction, focusing on emotion detection systems to enhance user-computer interaction. The objective was to identify the emotional state of users to create more seamless interactions in various sectors, including education and health. The methodology included a review of emotion identification through physiological signals, such as facial expressions, gestures, and brain imaging. While many researchers developed user-dependent emotion systems with high classification rates, few tackled user-independent systems, which yielded lower accuracy. To improve the classification rate of these systems, the study highlighted the need for better emotional stimulation techniques, larger data samples, and advanced signal processing algorithms. The findings suggested that while significant progress had been made in user-dependent systems, challenges remained in user-independent emotion detection. The study underscored the importance of physiological data in emotion identification and pointed out the existing difficulties in achieving accurate emotion recognition. The relevance of this study lies in its potential to advance emotion detection research by addressing the current limitations and suggesting avenues for future improvements.

Vasuki and Aravindan (2012) aimed to enhance emotion recognition from speech by introducing a hierarchical ensemble of classifiers with two levels. The first level involved the classification of Mel Frequency Cepstral Coefficients (MFCC) of input speech using Support Vector Machine (SVM) and Gaussian Mixture Model (GMM) classifiers. The outputs from these classifiers, which included posterior probabilities from the GMM and discriminate function values from the SVM, were then fed into a second-level SVM classifier for final emotion classification. The researchers utilized the Berlin Emo-DB database, which contains speech samples for seven emotions: anger, fear, boredom, happiness, neutrality, disgust, and sadness. The findings revealed that the proposed fusion method achieved a recognition accuracy of 75%, outperforming the individual SVM and GMM classifiers, which achieved accuracies of 67% and 66%, respectively. This study was relevant as it demonstrated the effectiveness of combining multiple classifiers in a hierarchical structure to improve emotion recognition accuracy, offering potential applications in fields such as human-computer interaction and affective computing.

Saha et al. (2013) proposed a gesture detection method for Indian classical dance using the Kinect sensor to capture the human body’s skeleton. The study's objective was to identify emotions such as "Anger," "Fear," "Happiness," "Sadness," and "Relaxation" through gesture analysis. The methodology involved deriving twenty-three features, including six joint coordinates from the right and left hands and five from the upper body, focusing on the distance between body parts, velocity, acceleration, and joint angles. These features were used to extract emotional intensity information, determining whether the emotion was positive or negative. Support vector machines (SVM) were employed to classify the emotions, achieving a recognition rate of 86.8%. The study’s findings demonstrated the effectiveness of SVM in emotion detection using skeletal gesture data from the Kinect sensor. The relevance of this study lies in its application to gesture recognition for emotion detection, particularly in the context of Indian classical dance, providing insights into emotion recognition using minimal and precise skeletal data.

Mohseni, Zarei, and Ramazani’s (2014) develop a system for identifying facial movements, focusing on detecting human facial expressions and emotions. The methodology involved creating a face model graph using facial expression muscles in each frame, followed by an algorithm that extracted features by evaluating the size and angle fluctuations of facial graph edges. Seven facial expressions, including a neutral posture, were categorized using Support Vector Machine (SVM) and other classifiers, utilizing the MMI database. The researchers emphasized that their approach, which did not rely on the action unit system, avoided errors typically associated with the faulty identification of action units. The findings demonstrated that analyzing facial movements led to accurate and efficient recognition of facial expressions. The study's relevance lies in its contribution to improving human-computer interactions, social robots, and behaviour monitoring systems by enhancing the automatic understanding of human facial emotions without the typical errors caused by initial identification failures.

Khowaja et al. (2015) aimed to develop a system for identifying human emotions through facial analysis as part of a smart home automation system. The methodology involved capturing the user’s image, detecting the face, and segmenting regions of interest (eyes, nose, and mouth) to analyze emotions such as sadness, anger, happiness, and neutrality. This process utilized feature extraction methods and Principal

Component Analysis (PCA) along with Support Vector Machines (SVMs) for emotion classification. Java-based policies were implemented to simulate a home automation environment for testing and validation purposes. The system was initially tested on a single user and was able to detect four basic emotions. The findings highlighted the feasibility of using facial features for reliable emotion detection, laying the groundwork for future development of more advanced systems capable of recognizing a wider range of emotions for multiple users. This research is relevant as it contributes to the development of user-friendly smart home technologies, enhancing user interaction by adapting to emotional states without the need for manual input.

Jaiswal, M., Tabibu, S., & Bajpai, R. (2016) conducted a study aimed at developing a data-driven strategy for detecting signs of dishonesty in trial data by analyzing both visual and linguistic signals. The methodology involved using Open Face for facial action unit recognition to capture facial feature movements of witnesses during questioning. Simultaneously, Open Smile was employed to analyze audio patterns, particularly focusing on smiles. Lexical analysis was also performed on spoken words, concentrating on pauses and utterance breaks. The gathered data were then processed through a Support Vector Machine (SVM) to predict the likelihood of dishonesty. To enhance prediction accuracy, the researchers further explored a fusion technique that integrated visual and lexical analyses through string-based matching. The findings showed that the combined approach of using facial features, audio patterns, and lexical cues significantly improved the prediction of dishonesty. This study is relevant in the field of automated deception detection, particularly in legal and forensic contexts, offering a promising method for more accurate assessments of witness credibility based on multimodal signals.

Kundu and Saravanan (2017) aimed to explore advancements in automated systems for recognizing human emotions and gestures, specifically through facial imagery. The study focused on evaluating the five basic emotions—surprise, happiness, disgust, normalcy, and sleepiness—commonly detected in facial expressions. The methodology included two key techniques: the first used a merged image of facial regions, particularly the eyes and mouth, which was analyzed through a feed-forward neural network trained via backpropagation. The second technique involved the use of Orientated Fast and Rotated (ORB) descriptors on single image frames to extract texture information, with classification performed using Support Vector Machines (SVM). The study highlighted the effectiveness of combining SVM with neural networks for emotion categorization, particularly in enhancing facial expression detection. It also explored the application of drowsiness detection systems using facial images, showcasing their potential in reducing road fatalities caused by driver fatigue. The findings demonstrated the promise of these automated systems in human-computer interaction, offering new insights into emotion recognition technology. This study is highly relevant to fields such as automotive safety, AI-driven communication, and mental health monitoring, where emotion and gesture recognition plays a crucial role.

Noroozi et al. (2018) explore the automated identification of emotional body gestures, an area that had received less attention compared to facial expressions and voice. The authors introduced emotional body gestures as part of "body language" and examined gender differences and cultural dependence. The methodology involved creating a comprehensive framework for detecting emotional body gestures in both RGB and 3D spaces, with techniques for human detection, posture estimation, and static and dynamic body posture assessment. The study also reviewed the latest research on emotion detection using expressive motion images and discussed multi-modal techniques combining facial expressions or speech with body motions. The findings highlighted that despite advances in pre-processing techniques like human detection and posture estimation, the availability of labeled data for emotion identification remained limited. The study also pointed out the lack of consensus on output spaces and the superficial nature of current representations, which were mostly based on basic geometrical representations. The study's relevance lies in its effort to expand the field of emotion recognition by focusing on body movements and providing a framework for future research in this underexplored area.

Rao, Rao, and Chowdary (2019) explore the integration of speech and facial expression features for more accurate emotion recognition. The methodology involved combining Mel Frequency Cepstral Coefficients (MFCC) from speech and Maximally Stable Extremal Regions (MSER) from facial expressions to enhance emotion prediction. The researchers utilized the Indian Face Database and Berlin Speech Database to assess recognition accuracy. Their findings indicated that combining MSER and MFCC improved recognition rates by an additional two to three percent compared to using either feature alone. The study highlighted that emotion recognition systems, which previously struggled with issues such as subject variability and environmental factors like illumination change and noise, could benefit from the combination of these two modalities. This approach was particularly relevant as it addressed the challenges of dynamic, subject-independent emotion recognition, contributing to more robust systems for real-world applications. The study was significant in advancing emotion recognition technology, offering a method to improve its reliability and effectiveness across diverse conditions.

Amin et al. (2020) developed *E-Voice*, a smart glove designed to bridge communication gaps between deaf-mute individuals and the general population by translating gestures into voice and text. The objective was to create a lightweight and portable prototype that could facilitate gesture translation using contemporary sensors, making it accessible and practical for everyday use. The methodology involved integrating the American Sign Language template for gesture recognition, allowing the glove to convert gestures into speech and text in real time. The findings highlighted the effectiveness of the E-Voice glove in reducing the time required for the general public to learn sign language, thus promoting inclusivity. The prototype was noted for its ease of use and potential to reduce social prejudices between deaf-mute groups and hearing communities. The study's relevance lies in its contribution to assistive technology, offering a practical solution for enhancing communication and fostering social integration for individuals with speech impairments.

Ilyas et al. (2021) aimed to improve human-robot interaction by addressing challenges in emotion detection systems, particularly in correlating upper body motions with facial expressions. The researchers developed an emotion-learning model that utilized a deep convolutional neural network (CNN) trained on benchmark datasets featuring both facial expressions and body movements. The methodology involved combining features from facial and body motions using various fusion techniques to enhance emotion detection. The results demonstrated that the system achieved an emotion identification accuracy of 76.8% using only upper body motions, surpassing the 73.1% accuracy on the FABO dataset. Furthermore, by incorporating multimodal compact bilinear pooling with temporal information, the approach outperformed existing state-of-the-art methods, achieving an impressive 94.41% accuracy on the same dataset. The findings underscore the relevance of integrating multimodal data for improving emotion and gesture recognition, particularly in human-robot interactions, ultimately contributing to better user experiences through more responsive robotic systems.

Kumar GS, S., Arun, A., Sampathila, N., & Vinoth, R. (2022) aimed to improve the accuracy of emotion recognition by analyzing electroencephalogram (EEG) signals, which are not influenced by outward appearance or behavior. They investigated EEG signals from various scalp locations to accurately categorize emotional states for human-machine interaction. The methodology involved extracting power spectral densities (PSDs) from multivariate EEG data, followed by the selection and classification of features using Long Short-Term Memory (LSTM) and Bi-directional Long Short-Term Memory (Bi-LSTM) networks. The study utilized a two-dimensional emotion model, considering both positive and negative feelings, and applied a region-based categorization scheme for frontal, parietal, temporal, and occipital regions. They compared the performance of their model with other classifiers like Artificial Neural Networks (ANN), Support Vector Machine (SVM), and K-Nearest Neighbors (K-NN). The findings revealed that Bi-LSTM achieved an accuracy of 94.95% when using four prefrontal electrodes, outperforming the other models. This research is highly relevant as it addresses the limitations of traditional emotion detection methods and demonstrates the potential of EEG-based emotion recognition for precise, real-time human emotion analysis, which can be crucial in enhancing human-machine interactions.

3. The Importance of Emotion and Gesture Recognition in Human-Computer Interaction (HCI)

Enhancing Natural Human-Computer Interaction: Emotion and gesture recognition plays a critical role in making human-computer interaction (HCI) more natural and intuitive. Traditional HCI methods, such as command-based interfaces, often require users to interact with machines through rigid input devices like keyboards or touchscreens. However, these methods lack the flexibility to understand the user's emotional state or non-verbal cues, such as hand movements or facial expressions. By incorporating emotion and gesture recognition, systems can better understand the nuances of human behavior, enabling them to respond more empathetically and naturally. This can result in more engaging and fluid interactions, where the system can adapt its responses based on the user's emotional state or gestures. For example, a virtual assistant that recognizes a user's frustration through their voice or facial expression could change its tone or provide additional assistance, making the interaction more user-friendly and human-like.

Applications in Mental Health and Assistive Technologies: Emotion and gesture recognition systems have significant implications for fields such as mental health monitoring and assistive technologies. In mental health, these systems can help detect emotional distress or signs of mental health issues by analyzing a person's facial expressions, speech, and physiological signals. For instance, systems that monitor emotional changes over time can be used to track the emotional well-being of individuals with conditions like depression or anxiety. These insights can help caregivers, therapists, or even AI systems provide personalized support. In assistive technologies, such as for individuals with disabilities, gesture recognition can facilitate communication for people who may have difficulty using traditional input methods. For example, sign language recognition systems enable deaf or hard-of-hearing individuals to interact with machines, making technology more accessible to everyone. Emotion and gesture recognition, therefore, have the potential to improve the lives of individuals by offering tailored support and improving accessibility in various domains.

Enhancing User Experience in Consumer and Entertainment Applications: Emotion and gesture recognition are also transforming consumer and entertainment applications, where they can enhance user experience and interaction. In customer service, for example, systems that detect a customer's emotional state can adjust their responses accordingly, offering more personalized and empathetic support. Similarly, in virtual reality (VR) and gaming, gesture recognition allows users to control environments with natural body movements, creating a more immersive and interactive experience. Games and applications can adapt in real-time to a player's emotions, enhancing engagement and user satisfaction. Furthermore, in advertising and marketing, emotion recognition can be used to gauge a consumer's emotional response to products or campaigns, allowing businesses to tailor their marketing strategies based on real-time emotional feedback. As a result, emotion and gesture recognition can significantly improve user experiences, making interactions with technology more responsive, personalized, and enjoyable.

Support Vector Machines (SVM): A Powerful Tool for Classification Tasks

Support Vector Machines (SVM) have emerged as one of the most effective tools for emotion and gesture recognition due to their ability to handle high-dimensional data and their robustness in classification tasks. SVM is a supervised learning algorithm that is particularly powerful in scenarios with limited data, which is often the case in emotion and gesture recognition tasks. SVM works by finding an optimal hyperplane that separates data points into distinct classes. This hyperplane maximizes the margin between different classes, which helps achieve better generalization and reduces the risk of overfitting. The primary advantage of using SVM for emotion and gesture recognition lies in its ability to classify data points with high accuracy, even in cases where the data is noisy or complex. Since emotion and gesture recognition often involves analyzing multi-dimensional data—such as facial expressions, speech signals, and physiological measurements—SVM is particularly well-suited to handle these complexities. In contrast to other machine learning techniques, SVM does not require a large amount of labelled data to perform effectively, which is valuable when working with specialized datasets. Additionally, SVM can use different kernel functions (linear, polynomial, or radial basis function) to transform data into higher-dimensional spaces, enabling it to handle complex classification tasks more effectively.

Multimodal Approaches to Emotion Recognition: Emotion recognition typically involves analyzing data from various sources, such as facial expressions, speech signals, and physiological measurements. While facial expressions are one of the most common indicators of emotional states, relying solely on facial recognition may not always provide accurate results. Facial expressions can vary widely between individuals, and different emotional states can be expressed using similar facial movements. Furthermore, expressions may be subtle or influenced by factors like cultural background or personality, further complicating recognition. To address these challenges, multimodal approaches to emotion recognition have been developed. These systems combine data from multiple sources, such as video frames, speech, and physiological signals (e.g., heart rate, EEG signals), to obtain a more comprehensive understanding of an individual's emotional state. SVM plays a crucial role in multimodal systems because it can efficiently integrate and classify data from diverse sources. By combining facial recognition data with speech and physiological measurements, multimodal systems are able to provide a more accurate and reliable assessment of a person's emotional state. For instance, the tone and pitch of speech may provide context for the facial expressions observed, making it easier to distinguish between similar emotions. Similarly, physiological signals like EEG or heart rate can offer additional insights into the emotional response, especially when facial expressions are ambiguous or subtle. The integration of multimodal data has become a key development in emotion recognition systems, improving the robustness and accuracy of the models. Researchers are increasingly focusing on fusing these various data types and applying SVM to classify emotions with higher precision.

Gesture Recognition with SVM: Applications and Challenges

Gesture recognition is another crucial aspect of human-computer interaction, focusing on the identification of hand movements, body postures, and other non-verbal cues that convey meaning. This form of recognition is essential in applications such as sign language interpretation, virtual reality (VR), and human-robot interaction. In gesture recognition systems, SVM is used to classify different gestures based on features extracted from video frames or sensor data, such as accelerometer or gyroscope readings. Gesture recognition systems often face challenges related to variations in gestures between individuals, environmental factors (such as lighting conditions), and the need for real-time processing. SVM, however, can handle these issues effectively by focusing on the most relevant features for classification. By extracting features from sensor data or video frames, SVM can identify gestures and classify them into predefined categories. This ability to classify gestures based on a diverse range of features makes SVM particularly useful in scenarios where high accuracy and real-time performance are essential, such as in VR environments or when interacting with robots. Moreover, SVM-based gesture recognition systems have shown promise in facilitating communication for people with disabilities. For example, systems that recognize sign language gestures can enable deaf individuals to communicate more easily with machines, providing a more inclusive user experience. Similarly, in VR and gaming applications, gesture recognition enables more immersive interactions, allowing users to control virtual environments using natural hand movements.

Integration with Advanced Techniques and Scalability of SVM Systems: The integration of SVM with other advanced techniques, such as feature extraction methods, neural networks, and deep learning, has significantly enhanced the performance of emotion and gesture recognition systems. Feature extraction methods like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) help reduce the dimensionality of the data while preserving the essential characteristics needed for classification. These methods allow for the identification of the most important features of the data, making the SVM classification process more efficient and accurate. In addition to feature extraction, the combination of SVM with deep learning techniques, such as Convolutional Neural Networks (CNNs), has further improved recognition accuracy. While SVM alone is highly effective, its performance can be enhanced when integrated with deep learning models that automatically learn relevant features from raw data. This hybrid approach leverages the strengths of both SVM and deep learning, resulting in more robust and accurate emotion and gesture recognition systems. Moreover, with the increasing availability of large datasets and advances in computational power, SVM-based systems have become more scalable and efficient. Benchmark datasets such as EmoReact and Facial Expression Recognition (FER) have been developed, providing a standard for evaluating the performance of emotion recognition systems. These datasets allow researchers to fine-tune their models and ensure better generalization across different populations and contexts. The availability of large, diverse datasets also enables the development of more accurate and reliable systems that can be applied to real-world scenarios.

4. Conclusion

Emotion and gesture recognition systems have become essential components of human-computer interaction, contributing to more human-like and responsive technology. Through the application of Support Vector Machines (SVM), these systems can accurately classify emotions and gestures by analyzing complex and multidimensional data from multiple modalities. The integration of SVM with feature extraction methods, such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), alongside advances in deep learning techniques, has significantly enhanced the performance and accuracy of these systems. In real-world applications, such as virtual assistants, mental health monitoring, assistive technologies, and immersive gaming, the ability to understand and respond to human emotions and gestures has revolutionized user interaction with machines. Furthermore, the increasing availability of large datasets and computational power continues to improve the scalability and efficiency of emotion and gesture recognition systems, making them more applicable across diverse contexts and populations. This progress holds the potential to reshape how we interact with technology, creating more empathetic, accessible, and engaging experiences.

References

1. Singh, M., & Patterson, D. J. (2010, October). Involuntary gesture recognition for predicting cerebral palsy in high-risk infants. In *International Symposium on Wearable Computers (ISWC) 2010* (pp. 1-8). IEEE.
2. Jerritta, S., Murugappan, M., Nagarajan, R., & Wan, K. (2011, March). Physiological signals based human emotion recognition: a review. In *2011 IEEE 7th international colloquium on signal processing and its applications* (pp. 410-415). IEEE.
3. Vasuki, P., & Aravindan, C. (2012, November). Improving emotion recognition from speech using sensor fusion techniques. In *TENCON 2012 IEEE Region 10 Conference* (pp. 1-6). IEEE.
4. Saha, S., Ghosh, S., Konar, A., & Nagar, A. K. (2013, June). Gesture recognition from indian classical dance using kinect sensor. In *2013 fifth international conference on computational intelligence, communication systems and networks* (pp. 3-8). IEEE.
5. Mohseni, S., Zarei, N., & Ramazani, S. (2014, October). Facial expression recognition using anatomy based facial graph. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 3715-3719). IEEE.
6. Khowaja, S. A., Dahri, K., Kumbhar, M. A., & Soomro, A. M. (2015, December). Facial expression recognition using two-tier classification and its application to smart home automation system. In *2015 International Conference on Emerging Technologies (ICET)* (pp. 1-6). IEEE.
7. Jaiswal, M., Tabibu, S., & Bajpai, R. (2016, December). The truth and nothing but the truth: Multimodal analysis for deception detection. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)* (pp. 938-943). IEEE.
8. Kundu, T., & Saravanan, C. (2017, December). Advancements and recent trends in emotion recognition using facial image analysis and machine learning models. In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* (pp. 1-6). IEEE.
9. Noroozi, F., Corneanu, C. A., Kamińska, D., Sapiński, T., Escalera, S., & Anbarjafari, G. (2018). Survey on emotional body gesture recognition. *IEEE transactions on affective computing*, *12*(2), 505-523.
10. Rao, K. P., Rao, M. C. S., & Chowdary, N. H. (2019). An integrated approach to emotion recognition and gender classification. *Journal of Visual Communication and Image Representation*, *60*, 339-345.
11. Amin, M. S., Amin, M. T., Latif, M. Y., Jathol, A. A., Ahmed, N., & Tarar, M. I. N. (2020, November). Alphabetical gesture recognition of american sign language using e-voice smart glove. In *2020 IEEE 23rd International Multitopic Conference (INMIC)* (pp. 1-6). IEEE.
12. Ilyas, C. M. A., Nunes, R., Nasrollahi, K., Rehm, M., & Moeslund, T. B. (2021, February). Deep Emotion Recognition through Upper Body Movements and Facial Expression. In *VISIGRAPP (5: VISAPP)* (pp. 669-679).
13. Kumar GS, S., Arun, A., Sampathila, N., & Vinoth, R. (2022). Machine learning models for classification of human emotions using multivariate brain signals. *Computers*, *11*(10), 152.