IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 14 June 2023, accepted 8 August 2023, date of publication 22 August 2023, date of current version 31 August 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3307639*

RESEARCH ARTICLE

# Exploring Emotion Analysis Using Artificial Intelligence, Geospatial Information Systems, and Extended Reality for Urban Services

**SOMAIIEH ROKHSARITALEMI, (Member, IEEE), ABOLGHASEM SADEGHI-NIARAKI, (Member, IEEE), AND SOO-MI CHOI, (Member, IEEE)**

Department of Computer Science and Engineering, XR Research Center, Sejong University, Seoul 05006, Republic of Korea
Department of Convergence Engineering for Intelligent Drone, XR Research Center, Sejong University, Seoul 05006, Republic of Korea

Corresponding author: Soo-Mi Choi (smchoi@sejong.ac.kr)

This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation under ITRC Support Program under Grant IITP-2023-RS-2022-00156354; in part by the Metaverse Support Program to Nurture the Best Talents funded by the Korea Government [Ministry of Science and ICT (MSIT)] under Grant IITP-2023-RS-2023-00254529; and in part by the Ministry of Trade, Industry and Energy and Korea Institute for Advancement of Technology under Grant P0016038.

**ABSTRACT** The advent of human emotion data availability signifies a pivotal advancement, propelling intelligent systems to elevate human-technology interaction. The fusion of emotions with pioneering technologies such as artificial intelligence (AI), geospatial information systems (GIS), and extended reality (XR) heralds a groundbreaking era of emotion-driven intelligent systems, poised to revolutionize urban applications. In this context, our research carves a distinctive niche by engaging in a comprehensive review of pertinent literature, culminating in the formulation of an incisive 3-layer framework. Encompassing the bedrock of emotion analysis, seamless technology integration, intricate privacy considerations, and pragmatic implementation, this framework serves as a beacon for future endeavors. Most notably, we introduce the trailblazing concept of emotion-intelligent systems, an outcome of the harmonious blend of emotion data and pioneering technologies. Our research navigates the intricate terrain of cultivating emotionally sentient intelligent systems, deftly addressing complexities and surmounting challenges to forge a path towards a more emotionally attuned generation. With transformative implications, these avant-garde systems stand poised to reshape human-technology interfaces, thereby beckoning novel avenues for future exploration.

**INDEX TERMS** Emotion analysis, artificial intelligence (AI), geospatial information system (GIS), extended reality (XR).

## I. INTRODUCTION

Emotions are intricate, individual reactions to different stimuli that encompass changes in physiology, behavior, and cognition [1]. The increase in emotion data analysis with the advancement of intelligent systems has spurred the development of frameworks that foster improved human interactions [2]. Detecting a person's emotional state is pivotal in affective computing and can open new avenues for creating applications in urban services [3]. The surge in intelligent systems that make use of user behavior data—including emotions—has led to enhanced interactivity between humans and technology [3].

These applications demonstrate the versatility and potential impact of emotion detection and analysis technologies across different fields, from healthcare and transportation to entertainment and crime detection. For example, in healthcare, real-time emotion detection can be used to monitor health and provide distance learning support during pandemics—such as COVID-19 [4]. Understanding user

The associate editor coordinating the review of this manuscript and approving it for publication was Shadi Alawneh.

emotions can also be valuable for evaluating the quality of urban services—such as airports—by providing airport officials with insight into passengers' overall satisfaction. This can lead to more informed decisions regarding the allocation of resources and the development of more appropriate equipment and services to meet passenger needs [5]. Emotion analysis can also be used to improve driving safety in transportation systems. For instance, an application that analyzes user emotions using electroencephalography (EEG) can adjust the music played in a car to create a calmer and more focused driving experience. This could reduce distractions and potentially hazardous situations caused by emotional stress, ultimately leading to safer driving conditions [6]. Moreover, these emotion detection and analysis services can be applied to design game features for entertainment or therapy purposes—such as providing joyful game activities to users experiencing unpleasant emotions [7].

The potential impact of emotion detection and analysis technologies can be further amplified by the emergence of new technologies—such as artificial intelligence (AI), geospatial information systems (GIS), and extended reality (XR)—which enable the development of emotion-intelligent services that can enhance user experiences. For instance, GIS and emotion mapping can be used to analyze people's emotional responses to their environments, aiding in urban planning and design [8]. Moreover, AI combined with emotion detection can aid crime detection, as demonstrated by the use of multimodal sentiment analysis to classify crime through text and speech [9]. In emergency response training, a VR game that uses voice input to assess user reactions to incidents can detect emotional and psychological states [10]. Despite considerable advances in emotion detection and recognition technology, accurate analysis of emotions remains challenging owing to their subjective and individual nature. Additionally, the fuzzy borders between different emotional states and the lack of suitable data distributed across different moods pose further problems for emotion analysis. Consequently, it is imperative to develop methods for improving the accuracy of emotion analysis. Furthermore, ethical and privacy issues related to the collection and application of emotion data must be considered to ensure the responsible use of these technologies [11], [12], [13]. In pursuit of our goal, we have conducted extensive research on the integration of emotion analysis using technologies such as AI, GIS, and XR. The three main objectives of our research are as follows:

1) First, we aim to introduce the fundamental concepts of emotion analysis, including the definitions, techniques, modeling, and implementation environments in relevant studies. This section serves as an informative starting point for those unfamiliar with the concept.
2) The second part of our research is dedicated to exploring the combining of emotion analysis with new technologies—such as AI, GIS, and XR. We offer examples of urban applications to illustrate the potential of this integration.
3) In the third section, we introduce the concept of emotion-intelligent systems based on the integration of emotion data and new technologies. We also discuss the challenges involved in developing emotionally aware intelligent systems and the obstacles that must be overcome to achieve a more emotionally intelligent generation.

In general, the core contribution of this paper lies in its comprehensive exploration of emotion analysis integration with cutting-edge technologies, namely AI, GIS, and XR. The paper delves into the fundamental concepts of emotion analysis, bridging the gap for those unfamiliar with the field. It then advances to the heart of its contribution by detailing the merging of emotion analysis with new technologies. Through insightful urban application examples, the paper showcases the potential of this integration in enhancing urban services and user experiences. Additionally, the paper introduces the novel concept of emotion-intelligent systems that fuse emotion data with emerging technologies. This forward-looking approach underscores the importance of emotionally aware intelligent systems and addresses the challenges involved in their development. By identifying obstacles and offering strategies to overcome them, the paper presents a roadmap towards achieving a more emotionally intelligent generation of applications. Overall, the paper's contribution lies in its multifaceted exploration, ranging from foundational concepts to practical integration scenarios and forward-thinking insights.

Our research is organized into several sections to achieve our goal—that is, to explore the combining of emotion analysis with new technologies to develop a new generation of emotion-aware intelligent systems. **Section I** includes the introduction and provides an overview of the research and its goals. **Section II** discusses the research framework. **Section III** covers the basic concepts of emotion analysis, including modeling techniques and implementation environments. **Section IV** presents a literature review that discusses relevant studies on the combining of emotion analysis and new technologies comprising AI, GIS, and XR elements. **Section V** covers the next generation of emotion-aware intelligent systems, which introduces a new generation of intelligent systems that are emotion-aware and discusses the technologies involved in building such systems and the challenges in their development. **Section VI** presents the results of this study and discusses them. **Section VII** presents the conclusions, summarizing the research and its implications.

## II. RESEARCH FRAMEWORK

This section presents the general framework of the research, which can be divided into two parts. The first part discusses the search methodology used to gather relevant articles, the second outlining the components of the research framework formed based on the articles collected.

### A. SEARCH METHODOLOGY

We conducted a search using academic databases—including IEEE Xplore, Science Direct, Springer, and Google Scholar.

The search was limited to articles published between 2016 and 2023. Our aim is to achieve a complete understanding of the fundamental principles of emotion analysis and explore their integration with emerging technologies—such as AI, GIS, and XR. We divided our search into three categories.

The first category focused on articles related to the basic concepts of emotion analysis, including the definitions, technologies, models, and implementation frameworks.

The second category included articles that combined emotion detection and recognition with keywords—such as GIS, XR, and AI.

Finally, we examined articles on building emotional intelligence into next-generation technologies—such as emotionally intelligent XR, empowering AI with emotional intelligence, and spatial analysis of emotions. We also addressed the challenges involved in developing emotionally aware intelligent systems and the obstacles that need to be overcome to achieve a more emotionally intelligent generation. Our search strategy collected approximately 107 relevant articles, which are discussed in different sections of this paper.

### B. THE GENERAL FRAMEWORK OF OUR RESEARCH

As shown in **Fig. 1**, the proposed framework comprises three layers. The first layer contains the definition, technologies, methods, and implementation framework, focusing on the foundational concepts of emotion analysis. This includes defining key terms, outlining general steps, and discussing various technologies for data collection—such as visual, textual, audio, physiological, and multimodal methods. It also explores different categorizations of emotion analysis methods—such as rule-based, machine learning, and deep-learning techniques—while addressing privacy concerns. The implementation environment—including the relevant libraries, software tools, and programming languages—is also discussed.

The second layer demonstrates the potential of emotion analysis using new technologies for urban applications. This section categorizes the literature into three categories—that is, AI-based, GIS-based, and XR-based applications—and provides examples of developed urban applications. It also explores the benefits of combining AI, GIS, and XR with emotion analysis for different urban applications.

Finally, the third layer discusses hot topics for future research on building emotional intelligence into next-generation technologies, including emotionally intelligent XR, empowering AI with emotional intelligence, and spatial analysis of emotions: leveraging GIS for emotion detection and mapping. These areas of research have great potential to enhance the emotional intelligence of future technologies and advance the understanding of human emotions in various contexts. Additionally, we discuss the challenges involved in developing emotionally aware intelligent systems and the obstacles that need to be overcome to achieve a more emotionally intelligent generation.

**TABLE 1. Emotion model [15].**

| Model Name | Emotional State |
|---|---|
| Paul Ekman | Happiness, anger, fear, surprise, sadness, and disgust |
| Pleasure, arousal, and dominance (PAD) model | The 3D model based on pleasure, arousal, dominance |
| Russell's circumplex model | 2D, many states are categorized by HVHA, HVLA, LVHA, LVLA |
| Plutchick's model | Anger, fear, sadness, disgust, surprise, anticipation, trust, and joy |
| Positive activation negative activation (PANA) model | A 2D model where the vertical axis represents low-to-high positive affect and the horizontal axis represents low-to-high negative affect |
| Vector model | Arousal data with a vector for valence |
| Parrott, W.G | Joy, anger, sadness, fear, love, and surprise |

The details of the framework layers are discussed in various sections of the research. In Section III, which covers the fundamental concepts of emotion analysis, the first layer's specifics are addressed. Section IV presents a literature review related to emotion-aware AI-GIS-XR with a combination, which is linked to the second layer. Moving on to Section V, which delves into emotion intelligent systems and the challenges they pose for development, it pertains to the third layer's intricacies.

## III. BASIC CONCEPTS OF EMOTION ANALYSIS—DEFINITION, TECHNOLOGIES, METHODS, AND IMPLEMENTATION FRAMEWORK

Emotion recognition focuses on determining a person's emotional state at a specific time. It is possible to employ audio, visual, textual, or physiological data as inputs and anticipate emotions to achieve this goal. Basic characteristics—such as anger, disgust, fear, happiness, sadness, or surprise—or a continuous scale (arousal, valence, dominance, etc.), are applicable for this purpose [14]. The emotion level can be measured using the diverse models listed in **Table 1**. The models can be two-dimensional or three-dimensional (comprising arousal and valence, for example), or three-dimensional model (comprising valence, arousal, and dominance). In such models, valence refers to the degree of pleasantness or unpleasantness of an emotion, whereas arousal and dominance refer to the level of activation or excitement and the degree of control or dominance of the emotion, respectively [15].

As mentioned above, there are several methodologies and techniques for emotion detection and recognition, which can be categorized as follows:

### A. VISUAL EMOTION RECOGNITION

Research on visual emotion recognition (VER) is crucial in computer vision and deep learning applications. Numerous researchers in this field have studied facial gestures, and several have discussed the importance of observing additional visual factors—such as body posture, the movement of limbs, clothes, or breath amongst others. Analyzing multiple visual factors can enhance emotion recognition accuracy [16].

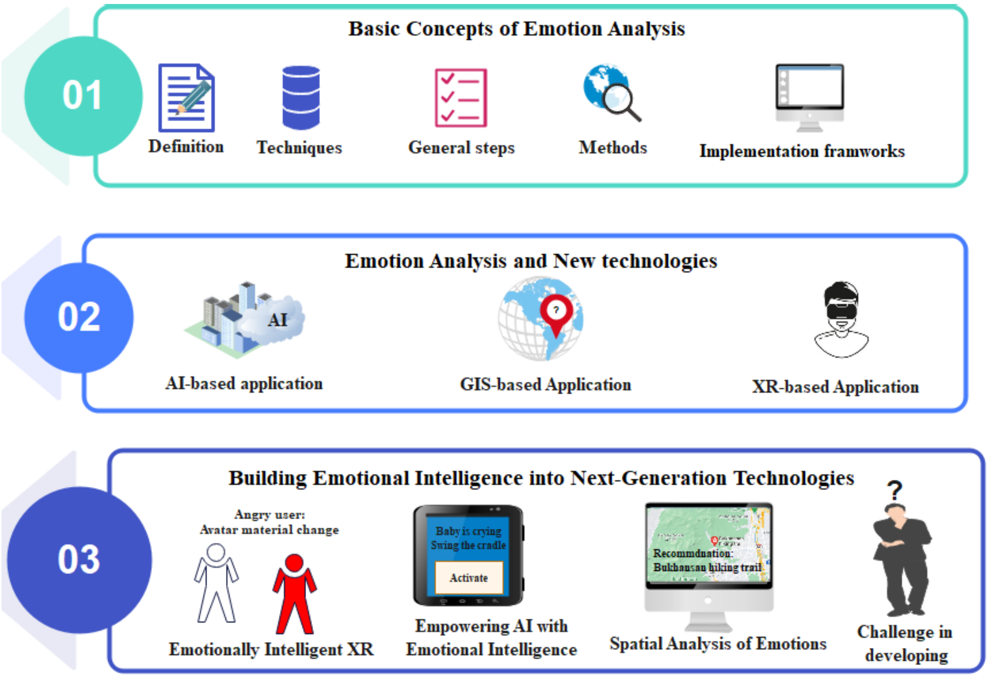


**FIGURE 1.** The research framework.

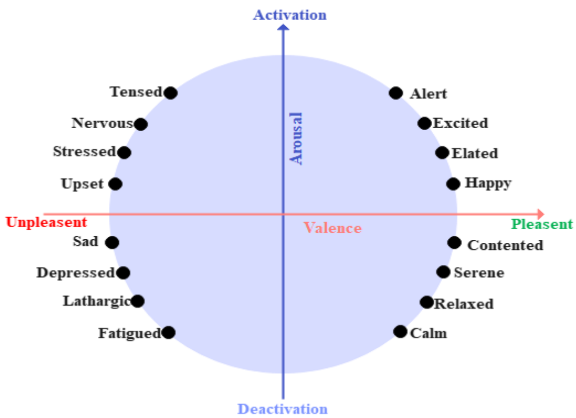


**FIGURE 2.** The 2D emotion model comprising valence and arousal [15].

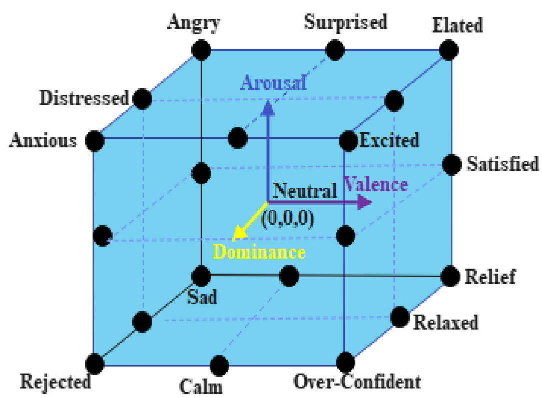


**FIGURE 3.** 3D emotion model comprising valence, arousal, and dominance [15].

### B. AUDIO EMOTION RECOGNITION

Speech-recognition systems use training data created from active, elicited, or natural sources. Further processing can be performed to render these signals suitable for feature extraction, supporting various emotions using prosodic and spectral features. After feature extraction, this system can employ classification algorithms to identify emotions using classic or new techniques—such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs). The addition of linguistic or visual features can enhance this system [17].

### C. TEXTUAL EMOTION RECOGNITION

With the rise of social networks, big data have been generated through user opinions on various platforms—such as Facebook and Twitter—the feelings of users being entered into these sources via texts, making these sources and their analyses crucial for learning about user emotions and behavior. Different models have been used to solve this problem, including deterministic models with continuous fine-grained alternatives for affective text analysis, and dimensional models with supervised or unsupervised classification approaches—such as deep-learning techniques and machine-learning techniques [18]. In other words, we can group textual emotion recognition methods into four categories—that is, keyword-based, lexicon-based, machine-learning-based, and hybrid methods. The first category uses pattern-based methods to identify emotion keywords for emotion recognition.

The second category applies a knowledge base with the text labeled according to the classified text. The third model uses supervised and unsupervised models—such as the Naïve Bayes (NB), support vector machine (SVM), and decision tree (DT) models—for emotion recognition, and the final model applies a combination of the others [19].

### D. PHYSIOLOGICAL EMOTION RECOGNITION

Physiological signals can be used to analyze emotions. Cardiac activity, electrical activity in the brain, temperature, respiration, and skin conductance can be used to analyze emotions and emotional states. In this regard, many devices and sensors exist. Blood volume, pulse, and photoplethysmography (PPG) can reveal blood flow using fingerprints; an electrocardiogram (ECG) can be used to assess cardiac electrical activity; temperature (ST) sensors can quickly measure temperature; electrodermal activity (EDA) can be used to measure skin activity; respiration sensors (RSP) can be used to measure respiration; and EEG can be used to measure the brain signals. See [20] for more information.

For this purpose, emotion data analysis involves several steps, including preprocessing to eliminate problems—such as noise, filtering, and normalization—to make the data more suitable for analysis, feature extraction, and feature selection—which entails removing and choosing features that are not useful because of the high dimensionality of the data space. Classification methods can be used to interpret data and obtain results using models such as the k-nearest neighbor (KNN), linear discriminant analysis (LDA), SVM, and DT models [20].

### E. MULTIMODAL EMOTION RECOGNITION

The use of a method for emotion recognition is an important problem to solve because of the specificities of each method and its external and environmental effects. For example, noise in the environment can affect speech recognition. Additionally, dark lighting and various angles should be considered in facial expressions. Because using only one method has limitations, combining different techniques—such as multimodal emotion recognition—can improve the performance and learning of each method [21]. Sharafi et al. [22] used a bidirectional long short-term memory (BiLSTM) network and two CNNs to combine audio and visual data for emotion recognition. Mel-frequency cepstral coefficients (MFCCs), energy features taken from audio signals, and BiLSTM network outputs were combined with the spatial and temporal features obtained by analyzing videos using this method. To classify emotions into various categories, this method used the SoftMax function [22]. Huang et al. [23] combined EEG with facial expressions for emotion detection. In this work, a neural network classifier was used in the facial expression method to classify emotions, and an SVM was used in the electroencephalogram (EGG) method. These two techniques were then combined by employing a sum or production rule [23]. Poria et al. [24] employed a method based on visual and textual models using a CNN. For emotion detection and recognition, this method could identify features from both approaches, applying multiple kernel learning classifiers to combine the results [24]. Sun et al. [25] proposed a novel approach for multimodal sentiment analysis using a multi-granularity feature attention fusion network that combined textual and image data from social media to perform emotion analysis [25].

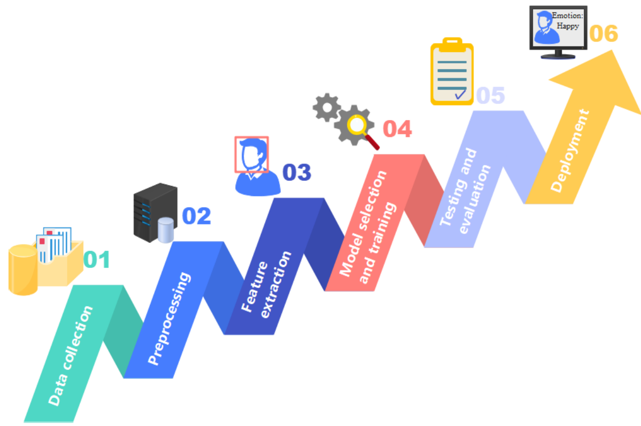

**FIGURE 4.** Emotion detection and recognition general steps.

#### 1) STRUCTURE OF EMOTION DETECTION AND RECOGNITION

The creation of frameworks based on emotion recognition in much research generally involves the steps shown in **Fig. 4**. Data collection is the first step and involves preparing the necessary data for emotion analysis—such as facial expressions, speech recordings, physiological signals, and textual data. Preprocessing aims to process data to prepare it for emotion analysis and includes techniques such as filtering, normalization, augmentation, and feature extraction. Feature extraction attempts to obtain relevant features related to emotions—such as facial expressions, body language, and gaze direction for visual emotion recognition; pitch, intensity, and spectral characteristics for audio emotion recognition; sentiment analysis; emotion lexicons; topic modeling for textual emotion recognition; and frequency bands, heart rate variability (HRV), and skin conductance for physiological emotion recognition.

Model selection and training is the next step, which involves selecting appropriate models—such as the SVM, CNN, or RNN—and training them on preprocessed data. Testing and evaluation follow, which evaluate the performance of the model on a separate dataset using metrics such as accuracy, precision, recall, F1-score, and area under the curve (AUC). The final step is deployment, which aims to deploy the model in real-world applications—such as tourism, healthcare, and other related fields [26], [27], [28], [29].

Emotion analysis encompasses various techniques—such as emotion detection, recognition, and classification. The goal of emotion detection is to determine the presence of

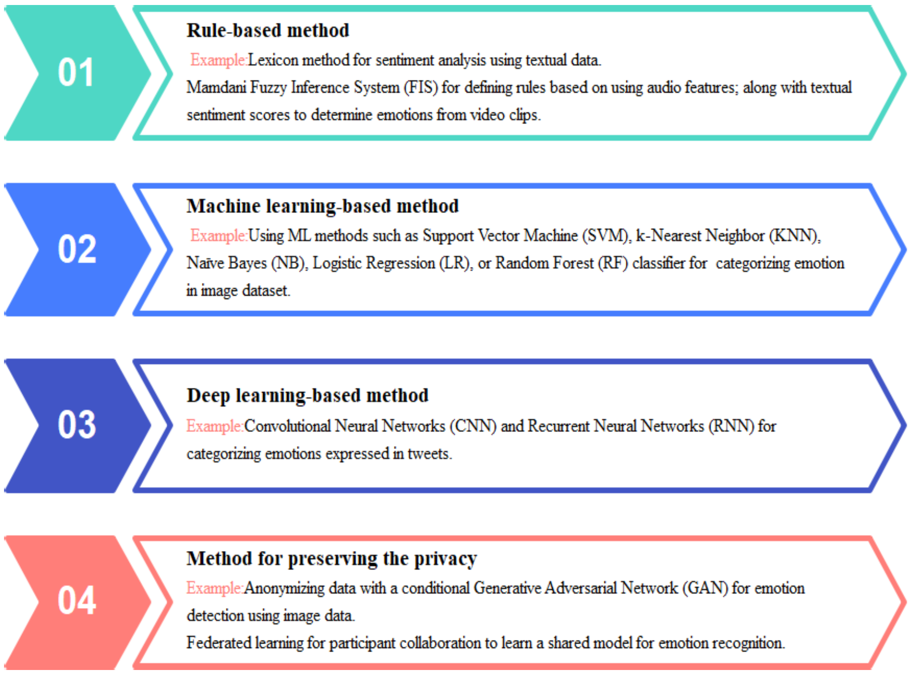


**FIGURE 5.** Emotion analysis methods general category.

**TABLE 2.** Emotion analysis methods categories based on different techniques [37] and [38].

| Category | Description | Models |
|---|---|---|
| 1 | Emotion analysis using facial data | Artificial neural networks, support vector machines, AdaBoost, constructive feedforward neural networks, face profile sequences, neural networks, wavelet transform, and neural network combination send tabular, 3D face recognition using swarm intelligence, 3D deformable model, MANFIS model, etc. |
| 2 | Emotion analysis using speech | Artificial neural networks tested on gender-dependent databases, anchor models, ASR, CNN, semi-CNN, hybrid deep neural network hidden Markov model (DNN-HMM), GMM-CHMM, fuzzy C-means, Bayesian hypothesis testing, HMM N-D HMM, etc. |
| 3 | Emotion analysis with physiological signals using ECG and EGG | ECG methods: Human emotion recognition using electrocardiogram signals, empirical mode decomposition (EMD) and discrete Fourier transform, Emotion recognition using wireless signals, supervised dimensionality reduction, etc.<br>EGG methods: Frequency domain features and SVMs, SVM and linear dynamic system, kernel Eigen-emotion pattern and adaptive SVM, combination of spatial filtering and wavelet, higher order crossings method, etc. |
| 4 | Emotion analysis using textual data | Knowledge-Based ANN, rough set theory and SVM, semantic labels and ATTs, separable mixture model, deep-learning-based personality detection, etc. |
| 5 | Multimodal techniques and models | CNN and RNN for facial expression and speech, CNN+LSTM for facial expressions and EEG, multi-level multi-head fusion attention + RNN for audio and text, CNN attention model for speech and text, 1D and 2D CNN for speech and visual data, convolution bidirectional long short-term memory neural network (CBLNN) for facial and heart rate, multimodal deep regression Bayesian network (MMDRBN) for audio and visual, 3D convolutional-long short-term memory for visual, textual, and audio, multi-directional regression for speech, face, deep multi-task learning framework for text, acoustic and visual frames of a video, etc. |

emotions in different forms of data (including images and text). Emotion recognition goes a step further by detecting the presence of emotions and assessing their intensity in specific contexts. Emotion classification involves categorizing emotions into different groups—such as happiness, sadness, anger, and fear, among others [30], [31], [32].

### 2) EMOTION DETECTION AND RECOGNITION METHODS

Various studies have considered the use of suitable analytical models for emotion extraction. Accordingly, emotion analysis methods can be categorized as follows:

1) **Rule-based methods**: To identify emotions in text or speech, rule-based methods use defined rules or heuristics, such as looking at words or phrases connected to emotions. One example of a rule-based method is the Lexicon method, which uses a dictionary to find positive words (such as "happy,") and negative words (such as "hate,") and assesses their frequency to assign a score that reflects the text's positive or negative emotional connotations [33]. Another example is a rule-based method based on the Mamdani fuzzy inference system (FIS) which defines rules based on audio

features (such as pitch and energy), along with textual sentiment scores to identify emotions (such as happiness, sadness, anger, and neutrality) in video clips [34].

2) **Machine learning-based methods**: Machine learning methods use statistical models to learn patterns and relationships in data, which can be used to detect emotions. These models can be trained on labeled data to recognize patterns associated with emotions. However, extensive data are usually required to train these models to ensure their accuracy and generalizability. Once trained, the model can be applied to identify emotions in newly unlabeled data. A good example of this is the use of machine learning (ML) methods—such as SVM, KNN, NB, logistic regression (LR), or random forest (RF) classifiers—for categorizing emotions in an image dataset [35].

3) **Deep-learning-based methods**: Deep learning-based methods use neural networks to process and analyze data for emotion analyses. These methods are particularly practical for complex tasks—such as speech recognition and sentiment analysis—as they can learn from unstructured data and extract the necessary features for accurate emotion detection and recognition. However, it is important to note that the training and execution of these methods require considerable computing resources and large volumes of data. An effective application of neural network models for emotion classification is the analysis of tweets. In the context of emotion analysis, CNNs and RNNs have been shown to perform well at accurately categorizing emotions expressed in tweets. Consequently, this approach has been used in several studies and has demonstrated promising results [36].

Several studies have examined emotion detection and recognition methods. Saxena et al. [37] categorized different methods of recognizing emotions based on facial expressions, speech signals, physiological signals, and textual analyses [37]. Maithri et al. [38] conducted a study on various emotion detection and recognition models using EEG, facial, and speech signals, and employing multi-modal emotion recognition techniques [38]. Marechal et al. [39] explored emotion analysis methods for different data types including text, sound, images, videos, and physiological signals, as well as multimodal emotion detection techniques [39].

Diverse citation approaches are accompanied by distinctive combinations of strengths and limitations, which are elaborated upon in the following discussion.

*Strengths:* The diversity of methods presented here, which includes rule-based, machine learning, and deep learning approaches, underscores their adaptability to a wide array of data sources such as facial expressions, speech signals, physiological signals, and textual data. Rule-based techniques, exemplified by the Lexicon method and Mamdani fuzzy inference system, enable the identification of emotions through predefined linguistic rules. Machine learning methods, employing SVM, KNN, and neural network classifiers, excel at recognizing intricate emotional patterns by learning from labeled data. Deep learning methods, including CNNs and RNNs, are adept at processing unstructured data like tweets, allowing for accurate emotion classification.

*Limitations:* However, these methods are not without their limitations. Rule-based approaches rely on predefined rules or lexicons, which may struggle to precisely capture the nuances of emotions and cultural contexts. Machine learning methods necessitate a significant amount of labeled data, which can be challenging to collect and limit their applicability. Additionally, these methods may encounter difficulties in generalizing to unseen data, leading to potential overfitting or compromised performance. Deep learning techniques excel at extracting intricate features, but their usage requires substantial computational resources and extensive datasets. Furthermore, these methods can suffer from interpretability issues, making it challenging to comprehend their decision-making process, particularly in critical applications.

#### 3) PRIVACY AND DATA SECURITY FOR EMOTION DETECTION AND RECOGNITION

In emotion detection and recognition that uses different data—such as faces, voices, and texts related to users—saving information related to user identification on third-party servers is always a major concern. Therefore, attention to security and the application of appropriate privacy-protection methods are increasingly important. Ramesh et al. [40] focused on anonymizing data to maintain the privacy of users whose data were used for emotion detection and recognition. The author proposed a technique based on a conditional generative adversarial network (GAN) that used a chosen landmark as the condition for the generator and discriminator to generate realistic faces with the least amount of identity-based information and the highest amount of emotion-based information in comparison with the original face for emotion detection analysis [40]. Khalifa et al. [41] proposed a method based on federated learning to address privacy issues for speech emotion recognition. This approach could be used when participant collaboration was required for learning a shared model without disclosing the participant's raw data [41]. Dias et al. [42] introduced a method based on two approaches, the first of which used homomorphic encryption and distance-preserving hashing techniques to maintain privacy in speech emotion recognition [42]. Feng et al. [43] presented a model for maintaining privacy in the federated learning of speech-emotion recognition. The method used user-level differential privacy (UDP) which was controlled by privacy parameters $\epsilon$ and $\delta$ to protect user privacy in federated learning [43]. Anwar et al. [44] proposed a novel technique for privacy-preserving emotion detection using EEG data. Their method involved the use of an artificial neural network (ANN) to classify emotional states based on arousal, valence, and dominance. This framework employed a federated learning method to ensure user data privacy [44].

#### 4) EMOTION DETECTION IMPLEMENTATION FRAMEWORK

It is important to use the appropriate libraries, tools, and programming languages for emotion detection and recognition. This section presents an example of the resources used to perform emotion analysis.

Libraries such as TensorFlow and PyTorch are commonly used for machine learning tasks, particularly natural language processing, sentiment analysis, and emotion analysis. NLTK is a powerful Python library designed specifically for emotion analysis, while TextBlob offers functionality for both sentiment and emotion analysis. Hugging Face is a popular library that provides pre-trained models for various natural language processing tasks, including sentiment and emotion analysis. Overall, these libraries can assist in the design of machine learning models that can effectively identify and analyze emotions and sentiments in textual data [45]. Additionally, the R language has several packages—such as Syuzhet, RSentiment, sentimentr, SentimentAnalysis, and meanr—that are commonly used for sentiment analysis. These packages offer a combination of machine learning and lexicon-based methods, as well as the necessary statistics for sentiment analysis [46].

Several excellent libraries are available for emotion analysis using image processing. OpenCV is a powerful open-source computer vision library that can be used for emotion analysis through image processing. TensorFlow provides tools for emotion analysis in images through facial recognition and detection, whereas PyTorch offers capabilities for emotion analysis in images, specifically for facial recognition and detection. Keras is a high-level deep-learning library that can be used for emotion analysis of images, including image classification and recognition. In addition, the scikit-image Python library provides tools for image processing and emotion analyses. These libraries can be used to build machine learning methods that accurately analyze emotions in images [47].

Several software tools are available for real-time emotion detection through facial analysis, such as DeepFace [48] which identifies six basic emotions—that is, happiness, sadness, anger, surprise, fear, and disgust. Affectiva [49] is a cloud-based API that provides real-time emotion detection from facial expressions, whereas Microsoft Cognitive Services [50] offers emotion recognition APIs that can detect emotions from facial expressions. Amazon Recognition [51] also provides emotion detection using deep-learning algorithms. Additionally, FaceReader [52] is a software tool that can recognize several basic emotions—including happiness, sadness, anger, fear, surprise, disgust, and contempt—from facial expressions.

Based on the work of Garcia-Garcia et al. [53], several SDKs and APIs are available for speech-based emotion detection and recognition—including Beyond Verbal API, Votakuri, Emotional Voice, and Good Vibrations. These tools enable emotion detection using speech signals. Additionally, SDKs and APIs—such as Emotion API, Affectiva, nViso, and Kairos—specialize in detecting emotions from facial expressions. Moreover, the tone analyzer, Receptiviti, BiText, and Synesketch are commonly used for emotion detection technologies from text [53]. **Tables 3** and **4** present the resource categories mentioned in this section for emotion analysis and compare their speeds.

**TABLE 3.** Tools, libraries, SDKs, and APIs for emotion analysis [45], [46], [47], [48], [49], [50], [51], [52], [53].

| Category | Resource Name | Description |
|---|---|---|
| 1 | TensorFlow, PyTorch, NLTK, TextBlob, Hugging Face, syuzhet, RSentiment, sentimentr, SentimentAnalysis, and meanr. | Useful for natural language processing, sentiment analysis, and emotion analysis. |
| 2 | OpenCV, TensorFlow, PyTorch, Keras, and scikit-image. | Practical for Emotion analysis through image processing. |
| 3 | DeepFace, Affectiva, Microsoft Cognitive Services, Amazon Recognition, FaceReader. | Software tools are available for real-time emotion detection through facial analysis. |
| 4 | 4.1: Beyond Verbal API, Votakuri, EmoVoice, and Good Vibrations<br>4.2: Emotion API, Affectiva, nViso, and Kairos.<br>4.3: Tone Analyzer, Receptiviti, BiText, and Synesketch. | Garcia-Garcia et al. [2017] categorize SDK and API as follows:<br>4.1: Speech-based emotion detection.<br>4.2: Detecting emotions from facial expressions.<br>4.3: Emotion detection from text. |

## IV. LITERATURE REVIEW

The combination of emotion analysis and new technologies offer tremendous potential for enhancing urban services. This field of study can be broadly categorized into three groups—that is, AI-based, GIS-based, and XR-based applications.

### A. AI-BASED APPLICATION

Oh et al. [54] suggested a deep-learning-based method to identify driver emotions such as fatigue and rage. This model combined a facial expression recognition algorithm using a CNN architecture with a sensor-fusion emotion recognition algorithm that integrated the state of the facial expression with EDA, a biophysiological signal based on Russell's arousal-valence model. This combination increased the detection accuracy to address circumstances in which emotions could not be determined based on facial expressions [54]. Kanaparthi et al. [55] presented a methodology based on machine learning using a KNN classifier to identify stress among IT employees. In this study, stress was identi-

**TABLE 4. Tools, libraries, SDKs, and APIs for emotion analysis based on speed [45], [46], [47], [48], [49], [50], [51], [52], [53].**

| Resource Name | Speed | Description |
| --- | --- | --- |
| TensorFlow, PyTorch, Keras | Fast | Deep learning libraries are optimized for achieving high performance and can process large amounts of data with speed. They are commonly used for real-time emotion detection applications. However, to achieve optimal performance, powerful hardware may be necessary. |
| OpenCV, scikit-image | Fast | These Image-processing libraries are designed to achieve high performance and can efficiently process image and video data. They are suitable for real-time applications that involve emotion detection, but optimal performance may necessitate the use of powerful hardware. |
| Hugging Face, TextBlob, NLTK | Medium | These natural language processing libraries are optimized for accuracy and ease of use but may not be as fast as deep learning libraries when processing diverse text data. They can still be employed for real-time emotion detection applications with proper optimization. |
| Affectiva, DeepFace, Microsoft Cognitive Services, Amazon Recognition, FaceReader | Medium | These facial analysis software tools use complex algorithms to analyze facial expressions and emotions, which may require more processing time. They can still be used for real-time emotion detection applications with proper optimization. |
| Beyond Verbal API, Votakuri, EmoVoice, Good Vibrations, Emotion API, nViso, Kairos, Tone Analyzer, Receptiviti, BiText, Synesketch | Slow | These emotion analysis tools typically rely on language models and complex algorithms, which can require significant processing time to analyze text, speech, or facial expressions. They may not be suitable for real-time emotion detection applications but can still be useful for offline analysis of emotion data. |

fied based on indicators—such as angry, disgusted, fearful, sad, and neutral—using image data analysis. The results of this research could be used to manage stress by creating a health-based work environment [55]. Feng et al. [56] proposed a user-group related topic-emotion detection (UGTE) model. In this model, a user's character was used to check the similarity of individuals' emotions to aggregate information in the group [56]. Zhu et al. [57] introduced a technique based on the use of an EGG and eye movement (EMO) data using the content-based ensemble method (CBEM) for depression detection. This approach divided the data into context-based subsets, and then used a majority vote strategy to assign the corresponding label to each subject. The use of these sources was extremely beneficial because of their ease of recording and noninvasive nature [57]. Bandyopadhyay et al. [58] developed an emotion detection system for online shopping recommendations. In this study, a webcam was used to detect customer emotions, and a recommendation was made in response to this detection using an ANN [58].

Kim et al. [59] presented a framework for detecting emotions in smart cities based on the Internet of Things (IoT). To improve the cumulative accuracy, the key issues surrounding the problem of emotion detection through wireless signals and their reflections were discussed in this study, the virtual emotional barrier being the name given to this problem. The max-k-cumulative-accuracy selection method has also been suggested as a solution to this problem. The results of this study could be applied to several smart city applications, including subways, parks, and other services [59]. Lu et al. [60] proposed a method for interactive greetings based on emotion detection. This method employed Viola-Jones for segmentation of the area around the face, eye, and mouth, face-CNN for identification of a person, eye-CNN, and mouth-CNN for emotion detection. 3D animation was then displayed based on the identified emotions [60]. Guo et al. [61] suggested a novel method based on deep learning—called deep-learning-assisted semantic text analysis (DLSTA)—to extract sentiments from textual data. This model handled big data, and its results could be used to analyze information from large sources on social networks such as Twitter [61]. Torcate et al. [62] presented a method that applied a pre-trained deep network and classical machine learning to recognize emotions in elderly people. In this method, a hybrid architecture based on the LeNet network was employed to extract features, and the random forest algorithm was used for the classification of emotions. The results could be used for music therapy and to support protection against dementia [62].

Seo et al. [63] applied a method for classifying emotions in patients based on EGG signals. To classify emotions like "happy, peaceful, or bored," this technique used a multilayer perceptron (MLP) and SVM along with deep-learning RNN-based models. The output of this system could be used by the healthcare systems [63]. Hwang et al. [64] analyzed the emotional states of construction workers, evaluating their activity, safety, and health using a bipolar dimensional emotional valence model. The research output could be used to identify negative emotions such as stress in workers who were in stressful situations, such as working on a ladder [64]. Meng et al. [65] proposed a technique using brain wave signals obtained via an EGG on data to analyze emotions for health monitoring applications. This study used a classifier—such as a feedforward neural network classifier model with a first-order stochastic optimization algorithm (ADAM) for emotion analysis—along with a layer incorporating touch behavior analysis of the user to improve emotion recognition results [65]. Menychtas et al. [66] proposed an interactive method that used the emotional states of patients during video conferences. This study analyzed user emotions by applying a

CNN using TensorFlow and an OpenCV library [66]. Chen et al. [67] applied emotion recognition through sentiment analysis and topic modeling using latent Dirichlet allocation (LDA) to analyze people's views on reopening measures for COVID-19. The main discussion topics—such as school, hospitality, and shop reopenings—were evaluated in this study. Emotions were measured using negative, positive, and neutral sentiments [67].

Hossain [68] explored the integration of 5G technology and emotion analysis, leveraging 5G's high-speed connectivity and low-latency capabilities for healthcare applications. This convergence facilitates the seamless incorporation of emotion analysis into different contexts. Real-time detection and response to emotional states are empowered, enhancing service quality and personalized experiences. This study introduces an emotion-aware connected healthcare system that employs IoT devices to gather and process patients' speech and image signals. Classification techniques are utilized to determine emotional states, aiding caregivers' effective response, particularly in cases involving pain detection.

Malik et al. [69] emphasized the significance of integrating Edge computing into emotion-aware applications to alleviate the data transmission strain on centralized cloud systems. By localizing computation closer to data sources, this approach effectively reduces the workload on remote servers. This study introduces a novel non-speech-based emotion recognition system, leveraging the potential of edge computing. The system's objective is to decipher emotions conveyed through non-speech cues like screams and cries. Additionally, the research investigates the application of knowledge distillation as a means to augment the system's efficiency and overall performance.

Rathee et al. [70] emphasized the significance of blockchain technology for emotion-aware applications. Their study introduced a blockchain-based intelligent mechanism that offers reliable trust-based schemes for data communication, recording, and transmission across heterogeneous networks. The research proposed a computing system leveraging blockchain for secure data transmission in emotion-aware applications. It employs the Analytical Hierarchical Process (AHP) for information exchange between patients and doctors, while identity-based trust models and blockchain mechanisms are incorporated to strengthen privacy among communicating entities.

### B. GIS-BASED APPLICATION

Cardone et al. [71] suggested a spatial framework based on fuzzy sentiment analysis using extended fuzzy C-means (EFCM) to analyze user emotions based on Twitter. The outcomes of this study assisted in identifying the modes of users' pleasant emotional categories—such as optimism—and unpleasant emotional categories related to unpleasant feelings—such as anguish— in areas that could be considered in tourism studies to determine the quality and service level based on citizen and tourist comments [71]. Cardone et al. [72] proposed a GIS-based framework for emotional analysis to find event hot and cold spots, such as a heat wave. This research employed user comments on social media—such as Twitter—and applied a method based on fuzzy-emotion-based hot- and cold-spot classification to fuzzify unpleasant and pleasant emotions and determine hot-spot and cold-spot areas based on these emotions [72].

Kariryaa et al. [73] explained the significance of modeling suitable routing criteria for autonomous vehicles based on the user mode and activity. An interesting example in this study was a scenario in which when the user was in an unpleasant mood, the system needed to change the route from the desired route—such as preferred restaurants or a place of natural beauty—altering the current user situation [73]. Park et al. [74] proposed a methodology based on exploratory spatial data analysis using a GIS and emotion analysis of Twitter data to analyze the emotions of theme park visitors, the Circumplex model being used to categorize emotions. This study analyzed the most frequently used words in each quadrant associated with different topics. A good example was the discovery of topics such as *mad* and *ride*, which were connected to emotions such as *annoyance*. The results of this study were applicable for identifying riding attractions in a theme park zone by analyzing emotion patterns and finding the regions of clustered emotion in each quadrant [74].

To use emotion analysis for landscape planning, Zhang et al. [75] analyzed images using FireFACE v1.0, obtained from microtwitters on Sino Weibo and a GIS. Multivariate linear regression was used to link emotions with urban spaces. For instance, it was determined that green space made an enormous contribution to happiness scores and blue areas made a negative contribution to sadness scores [75]. Kang et al. [76] proposed a method for GIS analysis to extract emotions from large-scale georeferenced social media photos. In this study, facial extraction and spatial clustering techniques were used to analyze the image data using the Face++ emotion API to identify happiness lists for tourist destinations. The effects of geographic factors and human emotions on these locations were evaluated using correlation analysis and multilinear regression [76].

Neppalli et al. [77] proposed a method for analyzing user comments made during Hurricane Sandy, using GIS and textual data from Twitter. The SentiStrength algorithm and SVM classifier sentiment analysis were applied to this method for sentiment analysis, and a GIS was used for presentation and cluster measures to find variations in tweets while considering Hurricane Sandy, the results being beneficial to emergency responders [77]. Yao et al. [78] applied a method based on topic modeling and domain-adaptive sentiment analysis to create a real-time analytical and geospatial visual system for urban events, the results of which could be useful for providing prompt responses during disaster management [78]. Xiao et al. [79] presented a methodology using emotion classification based on bidirectional encoder representations from transformers (BERT) and CNNs. Additionally, hot spot analysis and word-shift graphs were applied to understand similar topics, the results of which could help understand the happiness of residents, a good source for urban planning [79]. Rybarczyk et al. [80] investigated user comments on travel modes and determined the relationship between influencing factors and user opinions. This study was conducted based on an estimation of the sentiment index (pleasure or valence), followed by spatial regression to determine the relationship between opinions and elements related to travel modes. Understanding positive opinions about bicycling and the impact of factors—such as humidity—on negative opinions in this regard, was one of the positive outcomes of this research [80]. Resch et al. [81] used physiological data from sensors and GIS in urban planning to identify stress and relaxation areas. The research output was shown by applying GIS hot spot analysis to identify hot and cold spots in these areas [81].

Zhang et al. [82] employed emotion detection and a GIS to determine the relationship between the spatial characteristics of campus spaces and student emotions. In this study, facial emotion recognition with images, ECG analysis, and a wearable sensor for students' psychological parameters were used, regression being employed to determine the correlation between these emotions and spatial characteristics [82].

### C. XR-BASED APPLICATION

Wang et al. [83] discussed the importance of combining XR technology with user emotions to develop intelligent driving systems, a good example being one that uses augmented reality to display details of nearby drivers, including their emotions—for example, anger. This system could be used by drivers to assist with driving choices, such as lane changes and braking [83]. Guan et al. [84] described a platform that set the display features in the virtual reality space based on the detection of user emotions. In this study, deep learning was used to categorize user moods as neutral, happy, surprised, sad, angry, disgusted, fearful, or contemptuous. Using this model, the color of a virtual tree was changed to reflect these states—for example, the virtual trees could display green as a happy emotion [84].

Shaltout et al. [85] proposed a method for user emotion detection that involved interactions with an embodied XR agent. This methodology used analysis heat maps created by the user's eye gaze, in addition to tracking anxious users gaze patterns. This theory presumed that a worried user avoided looking at an avatar's face [85]. Joshi et al. [86] developed a tool for an online class in which the teacher could check student feedback based on a determination of their emotions during the presentation, and this analysis could then be simulated by a virtual agent. This research applied an open face for facial feature extraction and then used the Gaussian Naïve Bayes (GaussianNB) classifier model to recognize emotional states, such as happy, surprised, angry, sad, disgusted, and afraid. They also mentioned emotional states related to engagement, boredom, confusion, and frustration [86].

Dillon et al. [87] suggested a stress detection model based on assessing voice feature correlation with physiological parameters—such as blood pressure, heart rate, and psychological factors—whose results could be incorporated into VR to improve the presentation process of a smart virtual audience [87]. Lampropoulos et al. [88] applied a method based on text mining, sentiment analysis, and topic modeling to determine trends in augmented reality and virtual reality in education, as well as people's sentiments related to these technologies. This study used tools such as Text Blob, a word-emotion association lexicon (EmoLex), a valence aware dictionary for sentiment reasoning (VADER), and LDA for this purpose [88]. Eom et al. [89] applied emotion modeling based on arousal and valence to display pictures as holograms. The user biosensor data were monitored for various levels of happiness (from happy to unhappy) and calmness (from calm to excited), and pictures were chosen as holograms to complete the task [89]. Hart et al. [90] applied mixed reality (MR) to share emotions between collaborators and its augmentation in the form of virtual avatars. This framework applied psychological sensors to obtain user emotions and changed an avatar's status based on these emotions [90]. Zhang et al. [91] created an AR application for testing using the Microsoft Azure Face API. It provided users with opinions on a new product by analyzing their emotions and displaying them using augmented reality technology [91]. Bellenger et al. [92] presented a methodology for predicting facial emotion recognition using a reduced feature set in metaverse-based avatars. This method involved the use of ML algorithms—such as SVM—based on features extracted from image datasets [92]. Reig et al. [93] used a bespoke CNN to detect the emotions of dogs—such as playful, alert/aroused, appeasing, and neutral—in an AR application designed for interaction between humans and pets [93]. Selfridge et al. [94] used emotion analysis and MR for a live music performance, the purpose of which was to create an MR visual using performers' emotions. In emotion sensing, emotion sensors are used to monitor participants' facial expressions, predict their expressed emotions, and execute predictions that are mapped onto the screen. [94]. Izountar et al. [95] developed a personalized exergame for motor rehabilitation comprising three parts—that is, computing and interpretation, emotion recognition, and adaptation. The method employed an interactive framework using deep face and CNN models to detect emotions with positive and negative results and then adapted game objects to provide an appropriate game for the patient [95].

## V. BUILDING EMOTIONAL INTELLIGENCE INTO NEXT-GENERATION TECHNOLOGIES: EXPLORING THE INTERSECTION OF XR, AI, AND GIS FOR EMOTION DETECTION AND ANALYSIS

The intersection of emotional data and various technologies has the potential to create a new generation of intelligent systems known as emotionally intelligent systems. These systems can perceive and respond to user emotions in real time, resulting in more personalized and engaging experiences. Emotionally Intelligent XR systems can adjust their content and environment based on user emotions, leading to more immersive experiences.

This section explores the concept of emotionally intelligent XR, empowering AI with emotional intelligence, and the spatial analysis of emotions, and leveraging GIS for emotion detection and mapping. The importance of these topics lies in their potential to create more emotionally intelligent systems that can provide valuable insights for businesses and organizations.

Overall, the combining of emotion data and various technologies has the potential to create more advanced and personalized systems that can enhance user experiences, while also providing insights for businesses and organizations to better understand their audiences.

### A. EMOTIONALLY INTELLIGENT XR

A good example of emotionally intelligent XR was proposed by Gupta et al. [96], who developed an emotion-adaptive VR game set in a virtual forest, where the XR elements changed based on the user's emotions. In this study, the authors developed an emotion-adaptive VR game and user data obtained from EEG, EDA, and HRV physiological signals were used for the learning and calibration stages of this system. The user could move through a virtual forest, where virtual elements such as audio and fog changed depending on the user's emotional state, with unpleasant emotions causing the fog and audio to be reduced [96]. Emotionally intelligent XR could act as a real-time emotion-based adaptation of XR content and interfaces, thereby providing a more personalized and immersive experience.

In another study, Zheng et al. [97] developed a VR-based system that used multi-model physiological signals, EMO, pupil, and an ECG to model the level of fear simulated by a VR game based on hitting a ball. This was an excellent example of multisensory emotion detection in an XR environment [97]. The use of emotional data provided an opportunity to increase the expressiveness of virtual elements. Bernal et al. [98] offered a compelling example of this, as they employed human motion and physiological signals to imbue virtual avatars with greater expressiveness in real time. Specifically, their particle-system-based avatar enabled the avatar's material to demonstrate a range of aesthetics, from dim to bright and blue to red. By using an arousal-valence model to capture a user's emotional state, the particle system changed color from blue to red when the user experienced a high level of anger, effectively conveying the user's emotional state to other avatars [98]. XR capabilities could also be used to improve data representation in various urban applications, including in the medical field. For instance, Monferrer et al. [99] developed a facial-emotion recognition system for patients diagnosed with major depressive disorder that used virtual faces to display data on six basic emotions for more effective health monitoring [99].

### B. EMPOWERING AI WITH EMOTIONAL INTELLIGENCE

The incorporation of emotional data into AI systems can unlock a wide range of potential applications. Alam et al. [100] developed an innovative AI-based application that leveraged the IoT and machine learning algorithms to take care of children. The application could detect children's emotions, including crying, and by monitoring this, parents could gain insights into their child's emotional state. With the mobile app, parents could remotely swing the baby's cradle, leading to better time management. This application was an example of how AI could be empowered with emotional intelligence to provide valuable support to parents and care givers [100].

Jang et al. [101] demonstrated the potential of multimodal emotion detection and recognition using physiological signals in medical applications. Their study analyzed the ECG, EDA, respiration (RESP), and peripheral temperature (PT) data during rest, stress, and recovery using machine learning methods such as logistic regression (LoR), KNN, SVM, RF, and MLP to detect patients with panic disorder (PD). This automated system had the potential to enhance the diagnostic accuracy and minimize human error, leading to improved patient outcomes and lower healthcare costs [101]. Additionally, by leveraging multiple modalities of emotional expression, AI systems could potentially achieve more accurate and context-sensitive emotion recognition across different applications and cultures—such as medical and customer service. A good example of this was the work of Mittal et al. [102], who proposed a method for context-aware, perceived human emotion recognition using videos and images. This method applied faces and gates (body movement), semantic context, and sociodynamics using a self-attention-based CNN and depth map analysis [102].

### C. SPATIAL ANALYSIS OF EMOTIONS: LEVERAGING GIS FOR EMOTION DETECTION AND MAPPING

The combining of geospatial data and emotion analysis can create a geospatial emotion-aware application that delivers personalized services based on the detection and recognition of users' emotional states and preferences. One example is the sentiment-aware delightful walking route recommendation system proposed by Li et al. [103], which considered both seasons and scenery. The system used the sentiment analysis of tweets to calculate sentiment scores and identify relevant keywords related to seasons and sentiments. Additionally, it employed cosine similarity analysis of related words and geotagged tweets to suggest walking routes based on the locations of the tweets. This approach enabled the system to provide tailored recommendations that were better suited to the user's current mood and location [103].

The use of geospatial analysis to map and analyze emotional data can provide valuable insights for urban decision making. For example, Edry et al. [104] implemented a web-based geo-visualization tool that used sentiment analysis from Twitter to display geographic clusters of negative emotions or stress based on time in relation to the spread of COVID-19. This type of analysis could help policymakers and researchers obtain a more comprehensive understanding of how people experience and respond to events in different geographic areas. By identifying patterns and trends in emotion data, they could make informed decisions on how to allocate resources, implement policies, and communicate with the public.

Geospatial analysis of emotional data can also be applied in other fields—such as transportation planning, emergency management, and public health [104]. Huang et al. [105] presented a novel framework for collecting and analyzing human emotions using AI-based emotion analysis methods—such as natural language processing (NLP) and computer vision (CV), along with GIS for spatiotemporal analysis. This framework allowed for the exploration of the connection between individuals and their surroundings through visualization and analysis of emotional patterns [105].

#### 1) CHALLENGES IN DEVELOPING EMOTIONALLY AWARE INTELLIGENT SYSTEMS: ADDRESSING THE OBSTACLES TO ACHIEVE A MORE EMOTIONALLY INTELLIGENT GENERATION

Implementing an emotion-aware system requires careful consideration of the various challenges that arise in emotion analysis, including issues with data sources and limitations of existing models. For instance, when working with textual data, one of the key issues is the absence of high-quality datasets that include a proper distribution of emotional categories. This shortage of sufficient data can result in incomplete emotional information, making it difficult to obtain a strong analysis of the text. Additionally, there can be fuzzy emotional boundaries between different emotions—such as love and happiness—which require multiple emotional labels to capture expressions accurately. Another challenge in emotion analysis is the need to consider context. Some models assume that a sentence is separate from its context and expresses static emotions, which can lead to limitations in the accuracy of the analysis. Consequently, context consideration is essential for conducting a thorough analysis of emotions in textual data.

Overall, these challenges highlight the need for continued research and development of more robust and accurate models for emotion analysis in emotion-aware applications [12]. Moreover, using nonverbal cues—such as facial expressions and body language—for emotion-aware applications can be challenging to interpret accurately without the proper context. To solve this problem, it can be helpful to use multimodal emotion analysis, owing to the lack of information in one module, the modules including vision, audio, text, EEG, and body gestures. As such, audiovisual emotion detection and sampling are widely used because of their wide application and simplicity.

Guo et al. [106] suggested a new method of a novel audio-visual fusion network based on a 3D-CNN and convolution-augmented transformer (conformer) to extract features from the image and then from the signal and combines them with the fusion module, using these features in the prediction for emotion detection and recognition [106]. Considering that user characteristics are an essential element in the creation of emotionally aware systems, Gogula et al. [107] demonstrated the importance of the appropriate extraction of emotional data based on user characteristics in developing emotionally aware systems. They developed an application in the cloud for recommending books, which involved semantic network grouping of comparable sentences, sentiment analysis, and reviewer clustering. Sentiment analysis was performed using a CNN-long short-term memory (CNN-LSTM) model, followed by a clustering model based on user characteristics—such as age and location. This approach helped to provide personalized book recommendations to users based on their emotional and contextual preferences. This study highlighted the significance of considering user characteristics and emotional context in the development of emotionally intelligent systems [107].

Memory management is a crucial consideration for mobile applications that rely on the IoT. An example of the proposed method for optimizing memory usage is the approach presented by Talaat et al. [108] in their real-time emotion identification system for children with autism. This system incorporated a cloud layer, fog layer, and IoT layer to enable a mobile application to capture a child's face, detect emotions—such as anger, fear, joy, neutrality, sadness, or surprise—and send an alert to parents when the child's emotions were not normal. This system used a deep convolutional neural network (DCNN) architecture to enhance the accuracy of facial expression recognition. Additionally, the application employed a novel cache-replacement strategy using fuzzy logic to manage resources more efficiently. This strategy freed up memory by removing old captured images based on various parameters, such as the location of the parents. Overall, the proposed method demonstrated effective memory management in the context of IoT-based mobile applications [108].

Another consideration is when using big and diverse data (different modalities)—such as different sensors—a method that can handle these large data and correlate them to create the final result is very important. For example, Kanjo et al. [106] proposed an algorithm based on hybrid models using a CNN and LSTM RNN (CNN-LSTM) for sensor data—such as on-body, environmental, and location data—which can be very useful for sentiment classification in the use of big sensor data [109]. Emotion-aware applications are frequently designed with a narrow focus on particular racial, cultural, ethnic, and gender groups, and often neglect diversity. This oversight can lead to unintentional bias or inaccurate identification, which can negatively affect certain individuals. Moreover, the surreptitious collection of emotional data raises ethical concerns as it may be misused for malicious purposes. For example, emotion-sensing technology in the workplace may result in discrimination against workers or pressure them to feel happy and conceal their authentic emotions, leading to heightened tension and stress [110].

Emphasizing security and ethical considerations in relation to data and technology, particularly when data is collected or used by external parties, is paramount. The implementation of robust encryption, the imposition of stringent data access controls, and the active monitoring of activities to detect potential intrusions all play a pivotal role in ensuring the protection of sensitive information. Additionally, incorporating blockchain technology can further enhance data security by providing a decentralized and tamper-proof ledger for transparent and secure data transmission.

Through the prioritization of these security measures, both organizations and individuals can effectively mitigate the risks linked to unauthorized access, data breaches, and the potential misuse of valuable information. Furthermore, this proactive approach, coupled with blockchain's capabilities, showcases a dedication to upholding ethical standards and safeguarding user privacy. This combined effort fosters a heightened sense of trust within the realm of data and technology, promoting secure and reliable data transactions and interactions.

## VI. RESULTS AND DISCUSSION

User emotion analysis is a critical aspect of urban services and can offer useful insights for understanding the needs and preferences of individuals in various settings. Several technologies can be employed to achieve this goal, including visual, audio, textual, physiological, and multimodal emotion recognition applications, all of which combine several methods to address the limitations of each approach. By combining various methodologies, we can achieve a more profound and precise analysis of emotions. For instance, consider the fusion of facial expressions, body language, and voice analysis. This amalgamation allows us to capture intricate nuances in emotions; subtle shifts in facial expressions are complemented by the tone and pitch of the voice, resulting in heightened precision. Furthermore, the integration of users' textual input with physiological measurements provides another layer of insight. By examining the words someone uses and correlating them with physiological changes, we gain a holistic understanding of their emotional responses. The synergy between text analysis and physiological data offers a comprehensive view of emotions, showcasing both how someone expresses their emotions through words and the physiological alterations caused by those emotions. In essence, the integration of these diverse methodologies creates a synergistic.

The general steps in emotion detection and recognition include data collection, preprocessing, feature extraction, model selection, training, testing, evaluation, and deployment. Various libraries, software tools, and frameworks are available for emotion analysis—such as TensorFlow, PyTorch, NLTK, TextBlob, Hugging Face, OpenCV, Keras, scikit-image Python library, DeepFace, Affectiva, Microsoft Cognitive Services, Amazon Recognition, FaceReader, Beyond Verbal API, Votakuri, EmoVoice, Good Vibrations, Emotion API, Affectiva, nViso, Kairos, Tone Analyzer, Receptiviti, BiText, and Synesketch. These libraries, tools, and frameworks can be used for different types of emotion analyses, including image processing, speech-based emotion detection and recognition, and emotion detection technologies for text.

Choosing the appropriate technology depends on the specific use case and available data. Consequently, it is crucial to carefully evaluate the advantages and disadvantages of each approach before selecting a technology for emotion analysis, and to use appropriate analytical models to effectively analyze user emotions. Emotion analysis models can be categorized into various groups—that is, rule-based, machine-learning-based, and deep-learning-based models. Each of these models has its strengths and limitations, the choice of an appropriate model depending on the specific use case and available data.

Security is a major concern in emotion-based systems. Several studies have addressed this problem using different methods—such as data anonymization using GANs, federated learning, and homomorphic encryption. Anonymizing data ensures that user identities can be protected while allowing data to be used for analysis. GANs use two deep-learning models to generate synthetic data that are similar to real data, thereby helping to protect sensitive information. Federated learning enables multiple participants to collaborate in learning a shared model without disclosing the raw data. Homomorphic encryption allows data to be encrypted such that it can still be processed by an emotion analysis model—such as a ML algorithm—without the need to decrypt it.

Emotion analysis can be used to create intelligent emotion-aware systems that provide major benefits to urban services. For instance, in a virtual forest game, XR technology can be used to adjust the fog and audio levels by making use of a user's emotional state. This can improve user engagement and immersion in the virtual environment and provide opportunities for personalized and adaptive experiences. For example, emotionally intelligent XR can be used to change the virtual elements in a game—such as a virtual avatar that employs a user's emotional state. This approach can be applied to a particle-based system that uses a virtual avatar to change its appearance by making use of the user's emotions, thus making it easier for other avatars to understand the user's emotional state. Another example is empowering AI with emotional intelligence, which can be used to create child-care applications that detect a child's emotions and provide appropriate services—such as swinging the cradle when the child cries. The spatial analysis of emotions can also be made possible by providing location-emotion-aware services that can be used to recommend routes based on the mood of the user, to elevate their emotional state.

In summary, user emotion analysis is a critical aspect of urban services that can provide valuable insights into individuals' needs and preferences in different settings. Using appropriate analytical models and addressing security concerns, it is possible to create intelligent systems that can provide major benefits for urban services—such as emotionally intelligent XR, empowering AI with emotional intelligence, and spatial analysis of emotions.

Emotion-aware applications are crucial for better service delivery; however, their development faces certain challenges, including the lack of sufficient high-quality data and the need for appropriate distribution among different categories of emotion. Another challenge is the need for appropriate models that can handle the fuzzy nature of emotional states. Moreover, to address the weaknesses of current methods, further investigation is needed to explore new methods—such as multimodal emotion analysis that uses audiovisual features because of the limitations of one module. Additionally, mobile applications need to consider memory management strategies—such as cache-replacement strategies—to meet the specific needs of the platform, especially for the IoT. Additionally, it is important to use an appropriate strategy when using different sensors to handle the relationships between large data. Moreover, emotion-aware applications must consider user characteristics when delivering services. The development of AI algorithms that address racial, cultural, ethnic, and gender differences is also crucial and should be investigated. However, data tracking for emotion analysis could raise ethical concerns, as it could be misused for malicious purposes and lead to discrimination against certain groups. Consequently, the implementation of appropriate policies is important to address these concerns.

While emotion intelligent systems offer promising applications, their feasibility is contingent upon various factors. One primary requirement is the availability of diverse datasets, as these systems rely on comprehensive and representative data to accurately recognize emotions. Furthermore, the utilization of different emotion technologies, such as facial expression analysis, physiological and speech signal interpretation, and textual analysis, demands precise algorithms to prevent misinterpretations.

Privacy and ethical considerations are also pivotal aspects, raising concerns about data security and potential misuse. Striking a balance between delivering personalized experiences and ensuring user privacy presents a delicate challenge. Addressing this challenge entails implementing robust data protection measures and transparent user consent procedures to uphold ethical standards and foster user trust.

Amidst the promise and potential of integrating emotion analysis with advanced technologies, it is important to acknowledge the challenges that researchers encounter in this evolving field. The accurate detection and analysis of emotions remain complex due to their subjective and individualistic nature. The variability in human emotional responses, cultural influences, and contextual nuances pose significant hurdles in developing universally effective emotion recognition systems. Additionally, the inherent fuzziness of emotional boundaries and the scarcity of comprehensive and diverse datasets can hinder the robustness of emotion analysis algorithms. As this area involves the intersection of psychology, technology, and data science, ethical considerations related to privacy, data security, and the responsible application of collected emotion data further complicate the research landscape. Despite these difficulties, researchers and practitioners remain dedicated to refining and advancing emotion-aware technologies that respect human emotions while also contributing positively to various sectors.

## VII. CONCLUSION

After reviewing previous studies, we proposed a three-layer framework for analyzing emotions in combination with new technologies—such as AI, GIS, and XR. The first layer covers basic concepts—including definitions, technologies, models, and implementation environments—such as libraries, software, SDKs, and API—for emotion detection and recognition projects. This layer provides a suitable starting point for future studies. The second layer focuses on the combination of emotions and new technologies and provides examples of how this combination can be used to improve urban services. We also discuss various applications that can benefit from this approach. The third layer introduces a new generation of emotion-intelligence technologies with the potential to revolutionize the analysis and understanding of emotions. We discuss the capabilities of emotionally intelligent XR, empowering AI with emotional intelligence, and the spatial analysis of emotions, highlighting their benefits for urban services. We further discuss the potential of using these technologies to understand and improve the user experience in urban environments. We discuss the importance of incorporating emotional intelligence into urban planning and highlight the potential of these technologies in improving citizens' well-being. Overall, this framework provides a comprehensive approach for analyzing emotions and their interactions with new technologies in an urban environment. By integrating emotional intelligence into urban planning and services, we can create more user-centric and inclusive urban spaces that promote well-being and happiness. Additionally, we identify several challenges that must be addressed to successfully implement emotion-aware systems in real-world applications. These challenges include acquiring high-quality data, incorporating user context, utilizing multimodal models to improve model quality, implementing effective memory management strategies for mobile-based applications, developing strategies for processing vast amounts of sensor data, and addressing issues of diversity and ethical concerns in the development of models for a wide range of racial, cultural, ethnic, and gender groups.

## REFERENCES

[1] A. Moin, F. Aadil, Z. Ali, and D. Kang, "Emotion recognition framework using multiple modalities for an effective human–computer interaction," *J. Supercomput.*, vol. 79, no. 8, pp. 1–30, 2023.

[2] S. Alsubai, "Emotion detection using deep normalized attention-based neural network and modified-random forest," *Sensors*, vol. 23, no. 1, p. 225, Dec. 2022.

[3] S. Greene, H. Thapliyal, and A. Caban-Holt, "A survey of affective computing for stress detection: Evaluating technologies in stress detection for better health," *IEEE Consum. Electron. Mag.*, vol. 5, no. 4, pp. 44–56, Oct. 2016.

[4] M. Awais, M. Raza, N. Singh, K. Bashir, U. Manzoor, S. U. Islam, and J. J. P. C. Rodrigues, "LSTM-based emotion detection using physiological signals: IoT framework for healthcare and distance learning in COVID-19," *IEEE Internet Things J.*, vol. 8, no. 23, pp. 16863–16871, Dec. 2021.

[5] L. Martin-Domingo, J. C. Martín, and G. Mandsberg, "Social media as a resource for sentiment analysis of airport service quality (ASQ)," *J. Air Transp. Manage.*, vol. 78, pp. 106–115, Jul. 2019.

[6] C. Bankar, A. Bhide, A. Kulkarni, C. Ghube, and M. Bedekar, "Driving control using emotion analysis via EEG," in *Proc. IEEE Punecon*, Nov. 2018, pp. 1–7.

[7] H. Hadjar, P. McKevitt, and M. Hemmje, "Home-based immersive web rehabilitation gaming with audiovisual sensors," in *Proc. 33rd Eur. Conf. Cognit. Ergonom.*, Oct. 2022, pp. 1–7.

[8] T. H. M. de Oliveira, "The emotion-aware city: Using ambient geographic information (AGI) in order to understand emotion & stress within smart cities," in *Proc. AGILE PhD School*, 2015, pp. 1–5.

[9] M. Boukabous and M. Azizi, "Multimodal sentiment analysis using audio and text for crime detection," in *Proc. 2nd Int. Conf. Innov. Res. Appl. Sci., Eng. Technol. (IRASET)*, Mar. 2022, pp. 1–5.

[10] Y. Sermet and I. Demir, "Flood action VR: A virtual reality framework for disaster awareness and emergency response training," in *Proc. ACM SIGGRAPH Posters*, Jul. 2019, pp. 1–2.

[11] W. Yang, "Personalized physiological-based emotion recognition and implementation on hardware," Ph.D. thesis, Sorbonne Univ., France, 2018, pp. 1–164.

[12] J. Deng and F. Ren, "A survey of textual emotion recognition and its challenges," *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 49–67, Jan. 2023.

[13] A. Petrova, D. Vaufreydaz, and P. Dessus, "Group-level emotion recognition using a unimodal privacy-safe non-individual approach," in *Proc. Int. Conf. Multimodal Interact.*, Oct. 2020, pp. 813–820.

[14] D. Fedotov, Y. Matsuda, and W. Minker, "From smart to personal environment: Integrating emotion recognition into smart houses," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2019, pp. 943–948.

[15] M. R. Islam, M. A. Moni, M. M. Islam, D. Rashed-Al-Mahfuz, M. S. Islam, M. K. Hasan, M. S. Hossain, M. Ahmad, S. Uddin, A. Azad, S. A. Alyami, M. A. R. Ahad, and P. Lió, "Emotion recognition from EEG signal focusing on deep learning and shallow learning techniques," *IEEE Access*, vol. 9, pp. 94601–94624, 2021.

[16] J. Heredia, Y. Cardinale, I. Dongo, and J. Díaz-Amado, "A multi-modal visual emotion recognition method to instantiate an ontology," in *Proc. 16th Int. Conf. Softw. Technol.*, 2021, pp. 453–464.

[17] M. B. Akçay and K. Oğuz, "Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers," *Speech Commun.*, vol. 116, pp. 56–76, Jan. 2020.

[18] Y. Ghafoor, S. Jinping, F. H. Calderon, Y.-H. Huang, K.-T. Chen, and Y.-S. Chen, "TERMS: Textual emotion recognition in multidimensional space," *Int. J. Speech Technol.*, vol. 53, no. 3, pp. 2673–2693, Feb. 2023.

[19] K. Sailunaz, M. Dhaliwal, J. Rokne, and R. Alhajj, "Emotion detection from text and speech: A survey," *Social Netw. Anal. Mining*, vol. 8, no. 1, pp. 1–26, Dec. 2018.

[20] H. Z. Wijasena, R. Ferdiana, and S. Wibirama, "A survey of emotion recognition using physiological signal in wearable devices," in *Proc. Int. Conf. Artif. Intell. Mechatronics Syst. (AIMS)*, Apr. 2021, pp. 1–6.

[21] M. Li, X. Qiu, S. Peng, L. Tang, Q. Li, W. Yang, and Y. Ma, "Multimodal emotion recognition model based on a deep neural network with multiobjective optimization," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–10, Aug. 2021.

[22] M. Sharafi, M. Yazdchi, R. Rasti, and F. Nasimi, "A novel spatio-temporal convolutional neural framework for multimodal emotion recognition," *Biomed. Signal Process. Control*, vol. 78, Sep. 2022, Art. no. 103970.

[23] Y. Huang, J. Yang, P. Liao, and J. Pan, "Fusion of facial expressions and EEG for multimodal emotion recognition," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–8, Sep. 2017.

[24] S. Poria, I. Chaturvedi, E. Cambria, and A. Hussain, "Convolutional MKL based multimodal emotion recognition and sentiment analysis," in *Proc. IEEE 16th Int. Conf. Data Mining (ICDM)*, Dec. 2016, pp. 439–448.

[25] T. Sun, S. Wang, and S. Zhong, "Multi-granularity feature attention fusion network for image-text sentiment analysis," in *Proc. 39th Comput. Graph. Int. Conf. (CGI)*, 2023, pp. 3–14.

[26] A. Tripathi, U. Singh, G. Bansal, R. Gupta, and A. K. Singh, "A review on emotion detection and classification using speech," in *Proc. Int. Conf. Innov. Comput. Commun. (ICICC)*, 2020, pp. 1–5.

[27] H. Liu, Y. Zhang, Y. Li, and X. Kong, "Review on emotion recognition based on electroencephalography," *Frontiers Comput. Neurosci.*, vol. 15, p. 84, Oct. 2021.

[28] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial–temporal recurrent neural network for emotion recognition," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 839–847, Mar. 2019.

[29] A. N. Khan, A. A. Ihalage, Y. Ma, B. Liu, Y. Liu, and Y. Hao, "Deep learning framework for subject-independent emotion detection using wireless signals," *PLoS ONE*, vol. 16, no. 2, Feb. 2021, Art. no. e0242946.

[30] A. Jaiswal, A. Krishnama Raju, and S. Deb, "Facial emotion detection using deep learning," in *Proc. Int. Conf. Emerg. Technol. (INCET)*, Jun. 2020, pp. 1–5.

[31] P. Theerthagiri, "Stress emotion recognition with discrepancy reduction using transfer learning," *Multimedia Tools Appl.*, vol. 82, no. 4, pp. 5949–5963, Feb. 2023.

[32] S. A. Hussain and A. S. A. Al Balushi, "A real time face emotion classification and recognition using deep learning model," *J. Phys., Conf. Ser.*, vol. 1432, no. 1, Jan. 2020, Art. no. 012087.

[33] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, "Text-based emotion detection: Advances, challenges, and opportunities," *Eng. Rep.*, vol. 2, no. 7, p. e12189, Jul. 2020.

[34] S. Vashishtha and S. Susan, "Unsupervised fuzzy inference system for speech emotion recognition using audio and text cues (workshop paper)," in *Proc. IEEE 6th Int. Conf. Multimedia Big Data (BigMM)*, Sep. 2020, pp. 394–403.

[35] A. I. Siam, N. F. Soliman, A. D. Algarni, F. E. A. El-Samie, and A. Sedik, "Deploying machine learning techniques for human emotion detection," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–16, Feb. 2022.

[36] H. S. H. Kumar, Y. P. Gowramma, S. H. Manjula, D. Anil, and N. Smitha, "Comparison of various ML and DL models for emotion recognition using Twitter," in *Proc. 3rd Int. Conf. Intell. Commun. Technol. Virtual Mobile Netw. (ICICV)*, Feb. 2021, pp. 1332–1337.

[37] A. Saxena, A. Khanna, and D. Gupta, "Emotion recognition and detection methods: A comprehensive survey," *J. Artif. Intell. Syst.*, vol. 2, no. 1, pp. 53–79, 2020.

[38] M. Maithri, U. Raghavendra, A. Gudigar, J. Samanth, P. D. Barua, M. Murugappan, Y. Chakole, and U. R. Acharya, "Automated emotion recognition: Current trends and future perspectives," *Comput. Methods Programs Biomed.*, vol. 215, Mar. 2022, Art. no. 106646.

[39] C. Marechal, D. Mikolajewski, K. Tyburek, P. Prokopowicz, L. Bougueroua, C. Ancourt, and K. Wegrzyn-Wolska, "Survey on AI-based multimodal methods for emotion detection," in *High-Performance Modelling and Simulation for Big Data Applications*, vol. 11400. Cham, Switzerland: Springer, 2019 pp. 307–324.

[40] R. Ramesh, "Emotion detection with privacy preservation using adversarial learning," M.S. thesis, Texas A&M Univ., College Station, TX, USA, 2021, pp. 1–48.

[41] S. Latif, S. Khalifa, R. Rana, and R. Jurdak, "Federated learning for speech emotion recognition applications," in *Proc. 19th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2020, pp. 341–342.

[42] M. Dias, A. Abad, and I. Trancoso, "Exploring hashing and cryptonet based approaches for privacy-preserving speech emotion recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 2057–2061.

[43] T. Feng, R. Peri, and S. Narayanan, "User-level differential privacy against attribute inference attack of speech emotion recognition in federated learning," 2022, *arXiv:2204.02500*.

[44] M. A. Anwar, M. Agrawal, N. Gahlan, D. Sethia, G. K. Singh, and R. Chaurasia, "FedEmo: A privacy-preserving framework for emotion recognition using EEG physiological data," in *Proc. 15th Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2023, pp. 119–124.

[45] M. Singh, A. K. Jakhar, and S. Pandey, "Sentiment analysis on the impact of coronavirus in social life using the BERT model," *Social Netw. Anal. Mining*, vol. 11, no. 1, p. 33, Dec. 2021.

[46] M. Misuraca, A. Forciniti, G. Scepi, and M. Spano, "Sentiment analysis for education with R: Packages, methods and practical applications," 2020, *arXiv:2005.12840*.

[47] W. E. Villegas-Ch, J. García-Ortiz, and S. Sánchez-Viteri, "Identification of emotions from facial gestures in a teaching environment with the use of machine learning techniques," *IEEE Access*, vol. 11, pp. 38010–38022, 2023.
[48] Accessed: May 8, 2023. [Online]. Available: https://www.geeksforgeeks.org/facial-expression-detection-using-deepface-module-in-python/
[49] Accessed: May 8, 2023. [Online]. Available: https://www.affectiva.com/
[50] Accessed: May 8, 2023. [Online]. Available: https://learn.microsoft.com/ko-kr/xamarin/xamarin-forms/data-cloud/azure-cognitive-services/emotion-recognition
[51] Accessed: May 8, 2023. [Online]. Available: https://docs.aws.amazon.com/rekognition/latest/dg/what-is.html
[52] Accessed: May 8, 2023. [Online]. Available: https://www.noldus.com/facereader
[53] J. M. Garcia-Garcia, V. M. Penichet, and M. D. Lozano, "Emotion detection: A technology review," in *Proc. 18th Int. Conf. Hum. Comput. Interact.*, 2017, pp. 1–8.
[54] G. Oh, J. Ryu, E. Jeong, J. H. Yang, S. Hwang, S. Lee, and S. Lim, "DRER: Deep learning–based driver's real emotion recognizer," *Sensors*, vol. 21, no. 6, p. 2166, Mar. 2021.
[55] S. K. Kanaparthi, S. P, L. P. Bellamkonda, B. Kadiam, and B. Mungara, "Detection of stress in IT employees using machine learning technique," in *Proc. Int. Conf. Appl. Artif. Intell. Comput. (ICAAIC)*, May 2022, pp. 486–493.
[56] J. Feng, Y. Rao, H. Xie, F. L. Wang, and Q. Li, "User group based emotion detection and topic discovery over short text," *World Wide Web*, vol. 23, no. 3, pp. 1553–1587, May 2020.
[57] J. Zhu, Z. Wang, T. Gong, S. Zeng, X. Li, B. Hu, J. Li, S. Sun, and L. Zhang, "An improved classification model for depression detection using EEG and eye tracking data," *IEEE Trans. Nanobiosci.*, vol. 19, no. 3, pp. 527–537, Jul. 2020.
[58] S. Bandyopadhyay, S. Thakur, and J. Mandal, "Emotion detection for online recommender system using deep learning: A proposed method," *Innov. Syst. Softw. Eng.*, vol. 18, pp. 1–8, 2022.
[59] H. Kim, J. Ben-Othman, S. Cho, and L. Mokdad, "A framework for IoT-enabled virtual emotion detection in advanced smart cities," *IEEE Netw.*, vol. 33, no. 5, pp. 142–148, Sep. 2019.
[60] C.-T. Lu, C.-W. Su, H.-L. Jiang, and Y.-Y. Lu, "An interactive greeting system using convolutional neural networks for emotion recognition," *Entertainment Comput.*, vol. 40, Jan. 2022, Art. no. 100452.
[61] J. Guo, "Deep learning approach to text analysis for human emotion detection from big data," *J. Intell. Syst.*, vol. 31, no. 1, pp. 113–126, Jan. 2022.
[62] A. S. Torcate, F. S. Fonseca, M. A. de Santana, J. C. Gomes, and W. P. dos Santos, "Deep architectures based on convolutional neural networks and random forests for detection and recognition of emotions in facial expressions in the elderly," *Appl. Soft Comput.*, pp. 1–56, 2022.
[63] J. Seo, T. H. Laine, G. Oh, and K.-A. Sohn, "EEG-based emotion classification for Alzheimer's disease patients using conventional machine learning and recurrent neural network models," *Sensors*, vol. 20, no. 24, p. 7212, Dec. 2020.
[64] S. Hwang, H. Jebelli, B. Choi, M. Choi, and S. Lee, "Measuring workers' emotional state during construction tasks using wearable EEG," *J. Construct. Eng. Manage.*, vol. 144, no. 7, Jul. 2018, Art. no. 04018050.
[65] W. Meng, Y. Cai, L. T. Yang, and W.-Y. Chiu, "Hybrid emotion-aware monitoring system based on brainwaves for Internet of Medical Things," *IEEE Internet Things J.*, vol. 8, no. 21, pp. 16014–16022, Nov. 2021.
[66] A. Menychtas, M. Galliakis, P. Tsanakas, and I. Maglogiannis, "Real-time integration of emotion analysis into homecare platforms," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 3468–3471.
[67] Y. Chen, H. Niu, and E. A. Silva, "The road to recovery: Sensing public opinion towards reopening measures with social media data in post-lockdown cities," *Cities*, vol. 132, Jan. 2023, Art. no. 104054.
[68] M. S. Hossain and G. Muhammad, "Emotion-aware connected healthcare big data towards 5G," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2399–2406, Aug. 2018.
[69] I. Malik, S. Latif, S. Manzoor, M. Usama, J. Qadir, and R. Jurdak, "Emotions beyond words: Non-speech audio emotion recognition with edge computing," 2023, *arXiv:2305.00725*.
[70] G. Rathee, S. Garg, G. Kaddoum, and M. M. Hassan, "A secure emotion aware intelligent system for Internet of healthcare," *Alexandria Eng. J.*, vol. 75, pp. 605–614, Jul. 2023.
[71] B. Cardone, F. Di Martino, and S. Sessa, "GIS-based fuzzy sentiment analysis framework to classify urban elements according to the orientations of citizens and tourists expressed in social networks," *Evol. Intell.*, vol. 15, pp. 1–10, Apr. 2021.
[72] B. Cardone, F. Di Martino, and V. Miraglia, "A GIS-based hot and cold spots detection method by extracting emotions from social streams," *Future Internet*, vol. 15, no. 1, p. 23, Dec. 2022.
[73] A. Kariryaa, T. Veale, and J. Schöning, "Activity and mood-based routing for autonomous vehicles," 2020, *arXiv:2008.02016*.
[74] S. B. Park, J. Kim, Y. K. Lee, and C. M. Ok, "Visualizing theme park visitors' emotions using social media analytics and geospatial analytics," *Tourism Manage.*, vol. 80, Oct. 2020, Art. no. 104127.
[75] J. Zhang, Z. Yang, Z. Chen, M. Guo, and P. Guo, "Optimizing urban forest landscape for better perceptions of positive emotions," *Forests*, vol. 12, no. 12, p. 1691, Dec. 2021.
[76] Y. Kang, Q. Jia, S. Gao, X. Zeng, Y. Wang, S. Angsuesser, Y. Liu, X. Ye, and T. Fei, "Extracting human emotions at different places based on facial expressions and spatial clustering analysis," *Trans. GIS*, vol. 23, no. 3, pp. 450–480, Jun. 2019.
[77] V. K. Neppalli, C. Caragea, A. Squicciarini, A. Tapia, and S. Stehle, "Sentiment analysis during hurricane sandy in emergency response," *Int. J. Disaster Risk Reduction*, vol. 21, pp. 213–222, Mar. 2017.
[78] F. Yao and Y. Wang, "Towards resilient and smart cities: A real-time urban analytical and geo-visual system for social media streaming data," *Sustain. Cities Soc.*, vol. 63, Dec. 2020, Art. no. 102448.
[79] X. Xiao, C. Fang, H. Lin, L. Liu, Y. Tian, and Q. He, "Exploring spatiotemporal changes in the multi-granularity emotions of people in the city: A case study of Nanchang, China," *Comput. Urban Sci.*, vol. 2, no. 1, p. 1, Dec. 2022.
[80] G. Rybarczyk, S. Banerjee, M. D. Starking-Szymanski, and R. R. Shaker, "Travel and us: The impact of mode share on sentiment using geo-social media and GIS," *J. Location Based Services*, vol. 12, no. 1, pp. 40–62, Jan. 2018.
[81] B. Resch, I. Puetz, M. Bluemke, K. Kyriakou, and J. Miksch, "An interdisciplinary mixed-methods approach to analyzing urban spaces: The case of urban walkability and bikeability," *Int. J. Environ. Res. Public Health*, vol. 17, no. 19, p. 6994, Sep. 2020.
[82] J. Zhang and Y. Li, "The impact of campus outdoor space features on students' emotions based on the emotion map," *Int. J. Environ. Res. Public Health*, vol. 20, no. 5, p. 4277, Feb. 2023.
[83] H. Wang, Z. Wang, D. Chen, Q. Liu, H. Ke, and K. Han, "Metamobility: Connecting future mobility with metaverse," 2023, *arXiv:2301.06991*.
[84] J. Guan, "Extending the metaverse: Exploring generative objects with extended reality environments and adaptive context awareness," M.S. thesis, Toronto, ON, Canada, 2022, pp. 1–137.
[85] N. Shaltout, "[DC] emotion regulation via eye gaze," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2021, pp. 711–712.
[86] N. Joshi, N. Beecken, H. Bah, F. Steinicke, and J. Degner, "Advanced emotion analytics of virtual group meetings involving intelligent virtual agents," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2022, pp. 344–350.
[87] Arushi, R. Dillon, A. Ni Teoh, and D. Dillon, "Voice analysis for stress detection and application in virtual reality to improve public speaking in real-time: A review," 2022, *arXiv:2208.01041*.
[88] G. Lampropoulos, E. Keramopoulos, K. Diamantaras, and G. Evangelidis, "Augmented reality and virtual reality in education: Public perspectives, sentiments, attitudes, and discourses," *Educ. Sci.*, vol. 12, no. 11, p. 798, 2022.
[89] S. Eom, S. Kim, Y. Jiang, R. J. Chen, A. R. Roghanizad, M. Z. Rosenthal, J. Dunn, and M. Gorlatova, "Investigation of thermal perception and emotional response in augmented reality using digital biomarkers: A pilot study," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2023, pp. 170–173.
[90] J. D. Hart, T. Piumsomboon, G. A. Lee, and M. Billinghurst, "Sharing and augmenting emotion in collaborative mixed reality," in *Proc. ISMAR Adjunct*, 2018, pp. 212–213.
[91] J. Zhang, E. Kamioka, and P. X. Tan, "Emotions detection of user experience (UX) for mobile augmented reality (MAR) applications," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 1.4, pp. 63–67, Sep. 2019.
[92] D. Bellenger, M. Chen, and Z. Xu, "Facial emotion recognition with a reduced feature set for video game and metaverse avatars," Authorea Preprints, Univ. Huddersfield, U.K., 2023, pp. 1–21.

[93] S. Reig, E. P. Cruz, M. M. Powers, J. He, T. Chong, Y. J. Tham, S. Kratz, A. Robinson, B. A. Smith, R. Vaish, and A. Monroy-Hernández, ''Supporting piggybacked co-located leisure activities via augmented reality,'' in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2023, pp. 1–15.

[94] R. Selfridge and M. Barthet, ''Augmented live music performance using mixed reality and emotion feedback,'' in *Proc. 14th Int. Symp. Comput. Music Multidisciplinary Res.*, 2019, p. 210.

[95] Y. Izountar, S. Benbelkacem, S. Otmane, A. Khababa, M. Masmoudi, and N. Zenati, ''VR-PEER: A personalized exer-game platform based on emotion recognition,'' *Electronics*, vol. 11, no. 3, p. 455, Feb. 2022.

[96] K. Gupta, Y. Zhang, Y. S. Pai, and M. Billinghurst, ''WizardOfVR: An emotion-adaptive virtual wizard experience,'' in *Proc. SIGGRAPH Asia XR*, Dec. 2021, pp. 1–2.

[97] R. Zheng, T. Wang, J. Cao, P.-P. Vidal, and D. Wang, ''Multi-modal physiological signals based fear of heights analysis in virtual reality scenes,'' *Biomed. Signal Process. Control*, vol. 70, Sep. 2021, Art. no. 102988.

[98] G. Bernal and P. Maes, ''Emotional beasts: Visually expressing emotions through avatars in VR,'' in *Proc. CHI Conf. Extended Abstr. Hum. Factors Comput. Syst.*, May 2017, pp. 2395–2402.

[99] M. Monferrer, A. S. García, J. J. Ricarte, M. J. Montes, P. Fernández-Sotos, and A. Fernández-Caballero, ''Facial affect recognition in depression using human avatars,'' *Appl. Sci.*, vol. 13, no. 3, p. 1609, Jan. 2023.

[100] H. Alam, M. Burhan, A. Gillani, I. U. Haq, M. A. Arshed, M. Shafi, and S. Ahmad, ''IoT based smart baby monitoring system with emotion recognition using machine learning,'' *Wireless Commun. Mobile Comput.*, vol. 2023, pp. 1–11, Apr. 2023.

[101] E. H. Jang, K. W. Choi, A. Y. Kim, H. Y. Yu, H. J. Jeon, and S. Byun, ''Automated detection of panic disorder based on multimodal physiological signals using machine learning,'' *ETRI J.*, vol. 45, no. 1, pp. 105–118, Feb. 2023.

[102] T. Mittal, P. Guhan, U. Bhattacharya, R. Chandra, A. Bera, and D. Manocha, ''EmotiCon: Context-aware multimodal emotion recognition using Frege's principle,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 14222–14231.

[103] D. Li, S. Ishitsubo, K. Yamauchi, P. Siriaraya, S. Nakajima, and Y. Kawai, ''A sentiment-aware delightful walking route recommendation system considering the scenery and season,'' in *Proc. Int. Conf. Data Mining Workshops (ICDMW)*, Dec. 2021, pp. 867–872.

[104] T. Edry, N. Maani, M. Sykora, S. Elayan, Y. Hswen, M. Wolf, F. Rinaldi, S. Galea, and O. Gruebner, ''Real-time geospatial surveillance of localized emotional stress responses to COVID-19: A proof of concept analysis,'' *Health Place*, vol. 70, Jul. 2021, Art. no. 102598.

[105] Y. Huang, T. Fei, M.-P. Kwan, Y. Kang, J. Li, Y. Li, X. Li, and M. Bian, ''GIS-based emotional computing: A review of quantitative approaches to measure the emotion layer of human–environment relationships,'' *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 9, p. 551, Sep. 2020.

[106] P. Guo, Z. Chen, Y. Li, and H. Liu, ''Audio-visual fusion network based on conformer for multimodal emotion recognition,'' in *Proc. 2nd CAAI Int. Conf. Artif. Intell. (CICAI)*, Beijing, China, 2023, pp. 315–326.

[107] S. D. Gogula, M. Rahouti, S. K. Gogula, A. Jalamuri, and S. K. Jagatheesaperumal, ''An emotion-based rating system for books using sentiment analysis and machine learning in the cloud,'' *Appl. Sci.*, vol. 13, no. 2, p. 773, Jan. 2023.

[108] F. M. Talaat, Z. H. Ali, R. R. Mostafa, and N. El-Rashidy, ''Real time facial emotion recognition model based on kernel autoencoder and convolutional neural network for autism childrens,'' Kafrelsheikh Univ., Egypt, 2023, pp. 1–17.

[109] E. Kanjo, E. M. G. Younis, and C. S. Ang, ''Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection,'' *Inf. Fusion*, vol. 49, pp. 46–56, Sep. 2019.

[110] M. Peter and M.-T. Ho, ''Why we need to be weary of emotional AI,'' *AI Soc.*, vol. 37, pp. 1–3, 2022.

**SOMAIIEH ROKHSARITALEMI** (Member, IEEE) received the B.Sc. degree in applied surveying engineering from Shahid Rajaee Teacher Training University, the M.Sc. degree in geospatial information system (GIS) from the University of Tehran, and the Ph.D. degree in computer science and engineering from Sejong University, South Korea, in 2023. She is currently a Postdoctoral Researcher in computer science and engineering with Sejong University. Her research interests include GIS, artificial intelligence (AI), XR(VR/AR/MR), and utilizing these technologies for smart city applications.

**ABOLGHASEM SADEGHI-NIARAKI** (Member, IEEE) received the B.Sc. degree in geomatics engineering and the M.Sc. degree in geospatial information system (GIS) engineering from the Faculty of Geomatics Engineering, K. N. Toosi University of Technology (KNTU), and the Ph.D. degree in geo-informatics engineering from Inha University, South Korea, in 2008. He is currently an Assistant Professor with the Computer Science and Engineering Department, Sejong University, South Korea. Previously, he was an Assistant Professor with the Geo-Informatics Engineering Department, Inha University. He is a member of the XR Research Center for Real-Virtual Interconnected Metaverse, approved by the Korean Ministry of Science and ICT. He has extensive teaching experience at various universities, including Sejong University, Inha University, and KNTU. His research interests include GeoAI, XR (MR, VR, and AR), metaverse, ubiquitous GIS, healthcare, HCI, and culture technology (CT).

**SOO-MI CHOI** (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in computer science and engineering from Ewha University, South Korea, in 1993, 1995, and 2001, respectively. She is currently a Professor in computer science and engineering with Sejong University, South Korea, and the Director of the eXtended Reality Research Center Designated by the Ministry of Science and ICT of Korea. She was a Visiting Researcher with the Fraunhofer Institute for Computer Graphics Research, Germany, in 1998. She joined the Faculty of Computer Science and Engineering, Sejong University, in 2002. From 2008 to 2009, she was a Visiting Professor with the CG Laboratory, ETH Zürich, Switzerland. She is currently the Chief Vice President of the Korea Computer Graphics Society and the Vice President of the HCI Society of Korea. Her research interests include XR(VR/AR/MR), metaverse, computer graphics, HCI, and biomedical imaging.