; revised .

Affective Computing and the Road to an Emotionally Intelligent Metaverse

Farrukh Pervez[1], Moazzam Shoukat[2], Muhammad Usama[3], Moid Sandhu[4], Siddique Latif[5], Junaid Qadir[6]

[1]College of Aeronautical Engineering, National University of Sciences & Technology (NUST), Pakistan
[2]EmulationAI
[3]National University of Computer and Emerging Sciences (NUCES), Pakistan
[4]Australian e-Health Research Centre (AEHRC), CSIRO, Australia
[5]Queensland University of Technology (QUT), Australia
[6]Qatar University, Qatar

Corresponding Author: Junaid Qadir (e-mail: jqadir@qu.edu.qa)

ABSTRACT The metaverse is currently undergoing a profound transformation, fundamentally reshaping our perception of reality. It has transcended its origins to become an expansion of human consciousness, seamlessly blending the physical and virtual worlds. Amidst this transformative evolution, numerous applications are striving to mould the metaverse into a digital counterpart capable of delivering immersive human-like experiences. These applications envisage a future where users effortlessly traverse between physical and digital dimensions. Taking a step forward, affective computing technologies can be utilised to identify users' emotional cues and convey authentic emotions, enhancing genuine, meaningful, and context-aware interactions in the digital world. In this paper, we explore how integrating emotional intelligence can enhance the traditional metaverse, birthing an emotionally intelligent metaverse (EIM). Our work illuminates the multifaceted potential of EIM across diverse sectors, including healthcare, education, gaming, automotive, customer service, human resources, marketing, and urban metaverse cyberspace. Through our examination of these sectors, we uncover how infusing emotional intelligence enriches user interactions and experiences within the metaverse. Nonetheless, this transformative journey is riddled with challenges, and we address the obstacles hindering the realisation of EIM's full potential. By doing so, we lay the groundwork for future research endeavours aimed at further enhancing and refining the captivating journey into the world of EIM.

INDEX TERMS Metaverse, Affective Computing, AR/VR Technologies, Artificial Intelligence, Speech Emotion Recognition, Emotionally Intelligent Metaverse

I. Introduction

The concept of the metaverse is a captivating vision that combines a fully immersive and interconnected digital realm [1] through the convergence of virtual reality (VR), augmented reality (AR), and other digital technologies such as spatial computing, in an expansive ecosystem that allows seamless transition between physical and digital domains [2]. The metaverse concept goes beyond the traditional boundaries of VR, forging a comprehensive virtual universe where individuals can interact, collaborate, and engage with digital environments and entities in ways that mirror aspects of the physical world [3], [4]. As technology advances, the metaverse holds the potential to redefine social interactions, entertainment, education, commerce, and more, offering a new dimension of interconnectedness and experience [5] [6].

The success of metaverse technology relies heavily on its human-centric and social dimensions. Considering the fundamental role of affect and emotions in human interactions, it becomes imperative to incorporate insights and technologies from 'affective computing' [7] into the metaverse [8] to enable a virtual world possessing the capability to simulate, interpret, and respond to human emotions in real-time. In this paper, we posit the idea of an 'emotionally intelligent metaverse' (EIM), the envisaged future of traditional metaverse that leverages affective computing technologies to recognise users' emotional cues and express realistic emotions thus fostering genuine, meaningful and context-aware interactions within the digital realm. EIM offers multi-faceted benefits including but not limited to human-like social virtual interactions, enhanced users' emotional immersion, personalised therapeutic interventions, adaptive emotional designs and creation-based learning, elevating the overall quality of user experience. Metaverse designers use various elements like materials, lighting, layouts, and cultural nuances to craft virtual spaces

This article has been accepted for publication in IEEE Open Journal of the Computer Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCS.2024.3389462

FIGURE 1: Affective computing can transform the traditional metaverse into an emotionally intelligent metaverse.

that boost social interaction, learning, healing, and, most importantly, emotional connections within such EIM [9].

The concept of an EIM holds immense potential but introduces intricate challenges demanding thoughtful consideration. EIM can be used to support practical activities like meditation, personalised healthcare, gaming, content creation, emotion-driven education, and refining products with feedback (as depicted in Figure 1).

Translating human emotions into the digital realm however requires meticulous precision due to the complexity of emotional expressions, spanning facial features, vocal intonations, gestures, body language, and physiological responses.

To capture and transmit emotional nuances accurately, avatars and digital entities require advanced algorithms that decode real-time, context-dependent cues. Adapting affective computing models to accommodate this variability necessitates technological sophistication and deep comprehension of the interplay between emotions and human behaviour within the EIM.

This paper embarks on a thorough exploration, envisioning the creation of an EIM empowered by affective computing. Our primary focus centres on delineating potential features and applications of EIM, illustrating how affective computing can enhance the traditional metaverse in various ways, as demonstrated in Figure 1. However, amidst these opportunities, we engage in an open discussion regarding the challenges that might hinder seamless integration. These challenges encompass intricacies in accurately translating human emotions, ethical considerations regarding privacy and data security, and the need to adapt to the ever-evolving emotional landscape of the metaverse. Hence, this paper orchestrates a scholarly discourse navigating both the promises and complexities involved in harnessing affective computing to propel the metaverse towards heightened engagement and authenticity. The key contributions of this paper are appended below:

1) **Metaverse overview:** The paper introduces the metaverse, covering its history, enabling technologies and limitations.
2) **Exploration of EIM:** It explores how affective computing transforms the metaverse, with a focus on its impact in healthcare, education, gaming, and business.
3) **Challenges outlined:** The paper highlights challenges in designing EIM systems, including ethical considerations, responsible AI practices, and security.
4) **Future research directions:** This paper also identifies promising research avenues for advancing emotional intelligence integration in the metaverse, fostering innovation and ethical use.

Numerous surveys [3], [10]–[18] elucidating latest advancements in the metaverse have been published in the recent years. Various others [19], [20] focus on metaverse potential with regards to a particular application/area such as healthcare. Similarly, existing affective computing surveys [21], [22] focus on specific aspects of emotion recognition. We present the contributions of our paper in contrast to the existing metaverse surveys in Table 1.

A. Methodology

For the literature search in our survey, we curated references from well-established databases, including IEEE Xplore, Google Scholar, Scopus, and Web of Science, using search terms related to healthcare, metaverse, and artificial intelligence. To ensure comprehensive coverage, we also examined the bibliographies of relevant papers, identifying additional sources that extend the breadth and depth of our discussion on affective computing within the metaverse. This layered search strategy allowed us to incorporate seminal works and contemporary studies, fostering a rich dialogue on the evolution and future of emotionally intelligent interactions in EIM. We selectively gathered papers pivotal to advancing the conversation at the nexus of these fields, ensuring that each selected work significantly enriched our understanding of EIM. After thoroughly reading each paper, we made informed decisions on whether to include or exclude it from our survey based on its relevance, alignment with our paper's scope, and publication date, selecting papers not earlier than 2015. Our narrative synthesis, informed by an extensive array of academic contributions captures the historical evolution and the anticipated forward momentum of EIM, presenting a nuanced and comprehensive perspective shaped by a wealth of academic discourse.

B. Organisation

The paper is structured as follows: Section II explores the metaverse's history and core technologies, noting limitations. Section III covers affective computing in EIM, its impact on user immersion, and uses in healthcare, education, and gaming. Section IV addresses EIM development challenges like performance, adversarial robustness, real-time processing, and privacy. Section V looks at future EIM research directions, including methodological and ethical aspects. The conclusion

TABLE 1: Comparison of our paper with existing papers

Research Paper	Year	Metaverse			Emotionally Intelligent Metaverse			
		History	Enabling Technologies	Limitations	Literature Review	Features & Applications	Challenges	Future Directions
Lee et al. [10]	2021	✓	✓	✗	✗	✗	✗	✗
Sun et al. [11]	2022	✓	✗	✓	✗	✗	✗	✗
Wang et al. [12]	2022	✗	✓	✗	✗	✗	✗	✗
Ning et al. [13]	2023	✓	✓	✓	✗	✗	✗	✗
Coutinho et al. [8]	2023	✗	✓	✓	✗	very limited	very few	✓
This work	2023	✓	✓	✓	✓	detailed	comprehensive	✓

TABLE 2: Organisation of this paper

Sections	Sub-sections
I. Introduction	I-A. Methodology
	I-B. Organisation
II. Background of Metaverse: History, Enabling Technologies, and Limitations	II-A. The Timeline of Metaverse
	II-B. Metaverse's Enabling Technologies
	II-C. Limitations
III. Affective Computing for EIM	III-A. Authentic Human–AI Interaction
	III-B. Enhanced User Immersion
	III-C. Emotion–Aware Resource Dimensioning
	III-D. Meditation and Adaptive Health Surgeries
	III-E. Personalised Gameplay and Content
	III-F. Emotion-driven Virtual Education
	III-G. Product Refinement and Feedback
	III-H. Impact and Practicality of EIM in Diverse Industries
IV. Challenges in Developing EIM	IV-A. Performance considerations of emotion recognition
	IV-B. Securing EIM Against Adversarial Attacks
	IV-C. Privacy and Security Considerations in EIM
	IV-D. Addressing the Uncanny Valley Phenomenon in EIM
	IV-E. Modelling Dynamic and Evolving Emotional States in EIM
	IV-F. Challenges of Real-time Emotion Processing in EIM
V. Future Prospects for EIM Research	V-A. Foundational Models and EIM
	V-B. Cultural and Contextual Sensitivity in EIM using Deep Reinforcement Learning
	V-C. Improving Transparency and Explainability in EIM
	V-D. Ethical and Social Considerations in Developing Applications for EIM
VI. Conclusions	

summarises key findings and their significance for EIM's future. The paper's structure is summarised in Table 2 for reference.

II. Background of Metaverse: History, Enabling Technologies, and Limitations

The 'metaverse' combines 'meta' (beyond) and 'verse' (universe) to describe a three-dimensional space-time internet. It aims to create a virtual world parallel to the real one, with its own societal and economic systems. The metaverse's evolution, technological underpinnings, and challenges have been extensively discussed in research [23], [24]. Next, we outline its historical progress, key technologies, and inherent constraints.

A. A Timeline of Metaverse

The metaverse has evolved notably since the introduction of the first head-mounted VR and AR displays in the 1960s. NASA's utilisation of VR for astronaut training, followed by AR's application in theatre by 1996, demonstrates these technologies' versatile use cases. The term 'metaverse' was popularised by Neil Stephenson's 1992 novel "Snow Crash" [28]. Gaming applications of VR began in 1994, and by 2006, Roblox introduced an early version of the metaverse, integrating user-generated content. Google's VR-supported street view in 2007 and the emergence of NFTs in 2012 marked significant milestones in virtual technologies. The COVID-19 pandemic's influence expedited the transition towards virtual social interaction, boosting the metaverse's development, as outlined in Table 3.

TABLE 3: The timeline of the emergence of AR/VR technologies and the metaverse [11], [25]–[27]

Year	Event
1960	First VR head-mounted display
1968	First head-mounted AR display
1989	NASA used VR simulator to train astronauts
1991	A NASA engineer developed a VR system that lets you pilot a MARS rover
1992	Snow Crash: The first novel to introduce the metaverse and avatar
1994	VR in the gaming sector
1996	AR used in theatre and entertainment
2000	ARToolKit, an open-source computer tracking library for creating strong AR applications
2003	Second Life: 3D personalised virtual space
2006	Roblox: a rudimentary form of the metaverse
2007	Google brings its maps service with street-level 360-degree images
2009	AR emerged in print media
2012	Emergence of unique cryptographic tokens (NFT)
2014	Wearable AR technology (Google Glass)
2016	AR headset (Microsoft HoloLens)
2017	AR applications in retail
2018	Two standalone VR systems, Oculus Go and Oculus Quest, that need no computer or phone to work
2020	The critical point of virtual social
2021	The birth of metaverse

In 2021, Roblox's public offering and Google's Starline project underscored the growing focus on metaverse technologies. Facebook's rebranding to Meta signified its commitment to developing a metaverse ecosystem. Microsoft's Mesh for Teams and NVIDIA's Omniverse platform reflected the corporate world's increasing investment in the metaverse. Disney's venture into the metaverse with its IP imagery and Microsoft's acquisition of Activision Blizzard in January 2022 for $68.7 billion have both been pivotal to the metaverse's rapid progression [11]. These developments illustrate a sustained trend towards immersive digital environments supported by advanced technology.

B. Metaverse's Enabling Technologies

The metaverse, characterised by spatiotemporal extensibility, virtual-real interaction, and human-computer symbiosis, leverages advanced technologies for its existence and growth

This article has been accepted for publication in IEEE Open Journal of the Computer Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCS.2024.3389462

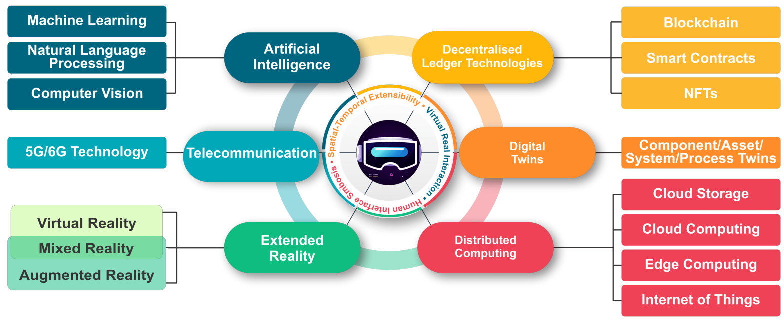

FIGURE 2: Metaverse enabling technologies

as shown in Figure 2. Spatiotemporal extensibility allows it to extend beyond physical limitations, offering endless virtual dimensions, while the virtual-real interaction integrates digital and physical realities. Human-computer symbiosis in the metaverse indicates a cooperative relationship, transcending traditional internet capabilities [11], [29], [30].

1) AI plays a critical role, with ML, NLP, and computer vision driving content diversity and enhancing user interaction within the metaverse.
2) Telecommunications enable real-time virtual-physical synchronicity. The advent of 5G, and the anticipated transition to 6G, is key to minimizing latency in the metaverse [31].
3) XR, comprising AR, VR, and MR, serves as a pillar for the metaverse, merging the physical and virtual through varied immersive experiences [12].
4) Blockchain, a part of decentralised ledger technologies, ensures secure, decentralised metaverse transactions, fundamental to its widespread adoption and functionality [12].
5) Digital twins provide digital replicas of physical entities, underpinned by distributed computing and ledger technologies, enabling real-time, authentic metaverse interactions [32], [33].
6) Distributed computing, along with cloud and edge computing, optimises data processing for the metaverse, essential for delivering a real-time experience [34]–[36].

C. Limitations

Despite its potential, the metaverse encounters challenges in emulating the complex dynamics of human communication that are intrinsic to the physical world [37]. Avatars and virtual characters, the metaverse's inhabitants, are often limited in their capacity to interpret and react to human emotions, social cues, and the subtleties of context [38]. As a result, metaverse interactions can lack the depth, authenticity, and richness that characterise face-to-face encounters, leading to experiences that may feel superficial and disconnected. When users engage in the metaverse for specific activities, like business negotiations or personal conversations, the virtual entities' insufficient emotional intelligence can lead to inaccurate perceptions of the users' emotional states. This could prevent avatars from providing empathetic responses to emotions like joy or sadness, frustration or anger, making interactions within the metaverse less compelling, authentic, and satisfying.

The traditional metaverse is constrained by several limitations, particularly in terms of personalisation options, as identified in [39]. These constraints inhibit the system's ability to tailor experiences to individual user needs, leading to possible frustration and diminished engagement. This is of significant concern in domains such as healthcare and mental health therapy where customisation is vital. Table 4 delineates these limitations in contrast to the capabilities of EIM. In virtual environments, fostering meaningful human interactions, which are fundamental, depends on the ability to establish deep emotional connections, which is a challenging task without emotional intelligence. Traditional metaverses, limited by poor content moderation and marketing from lacking emotional intelligence, face negative impacts on user experience and revenue. This issue also perpetuates problems like hate speech and cyberbullying, further damaging user experience and platform standing. To overcome these hurdles, incorporating emotional intelligence into the metaverse, and creating an EIM is imperative. The following section delves into how affective computing can revolutionise the metaverse, addressing the aforementioned limitations to realise the full potential of the metaverse.

III. Affective Computing for EIM

Affective computing, an interdisciplinary field, focuses on understanding and utilizing human emotions. It integrates engineering, psychology, education, cognitive science, and sociology to explore how technology can enhance our comprehension of emotions. This field examines the complex relationship between emotions, human-technology interactions, and system design for leveraging affective states. It provides numerous solutions to challenges in the traditional metaverse, aiding the shift to an 'emotionally intelligent metaverse' (EIM). The key emotional indicators that EIM systems aim to interpret include facial expressions [73], body language [74], and voice tones [75] etc. that are captured via advanced metaverse technologies and further processed by sophisticated AI algorithms.

EIM is defined as a novel iteration of the metaverse that integrates affective computing, distinguishing itself from

This article has been accepted for publication in IEEE Open Journal of the Computer Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCS.2024.3389462

TABLE 4: Comparison of feature limitations in the traditional metaverse versus enhancements in EIM

Features	Traditional Metaverse	EIM
Adaptive Interaction to User Emotions	X	✓
Dynamic Emotion-Aware Content Personalization	X	✓
Advanced Emotional Communication Tools	X	✓
Personalized User Experience Based on Emotional Data	X	✓
Emotionally Responsive Healthcare Solutions	X	✓
Learning Modules Tailored to Emotional and Psychological State	X	✓
Deep Emotional Data Analysis and Insights	X	✓
Enhanced Privacy Measures for Sensitive Emotional Data	X	✓
Affective Human-AI Interaction	X	✓

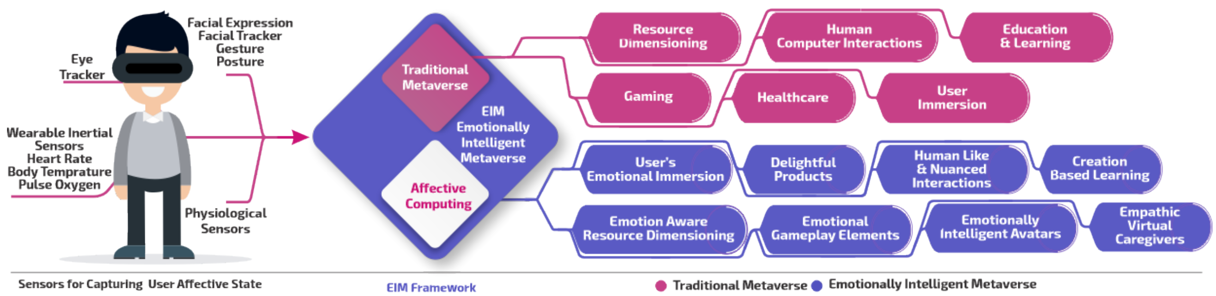

FIGURE 3: Presenting the EIM framework, showcasing the incorporation of emotional intelligence into the traditional metaverse. The conventional metaverse applications are depicted in shades of pink, while the EIM applications are represented in blue. Users' affective states are incorporated in traditional metaverse through sophisticated state-of-the-art metaverse technologies, as presented in Table 6. This integration of the traditional metaverse and affective computing unlocks a broad range of possibilities, enriching interactions and immersing users in the EIM experience.

TABLE 5: Features of emotionally intelligent metaverse

Feature	Description	References
Enhanced User Immersion	(a) Emotionally responsive avatars enhance users' immersive experience (b) Avatars exhibiting realistic emotions make virtual gatherings genuine and emotionally resonant	[40], [41], [38], [42], [43]
Emotion-Aware Resource Dimensioning	Optimal allocation of computing resources based on users' emotions	[8]
Meditation and Adaptive Health Surgeries	(a) Elevation of doctor-patient interactions would add to patients' convenience (b) Mindful exercises and stress reduction (c) Immersive adaptive virtual therapy sessions	[44], [45], [46], [47], [48], [49]
Personalised Gameplay and Content	Emotionally tailored gameplay elements and video content	[50], [51], [52]
Emotion-driven Virtual Education	Adaptive teaching techniques and learning preferences	[53], [54], [55], [56]
Product Refinement and Feedback	(a) Product refinement based on users' emotional feedback (b) Personalised advertisements and marketing promos	[57], [58], [59]
Authentic Human-AI Interaction	(a) More nuanced and authentic virtual interactions (b) Emotionally responsive narratives and experiences	[60], [61], [40]

TABLE 6: Suite of metaverse technologies and AI mechanisms for affective signal analysis in EIM

Signal	Metaverse Technologies	AI Technologies
Facial Expressions	VIVE Focus 3 Facial Tracker [62] & SALSA Lip Sync Suite v2 [63]	AI recognition algorithms
Voice	HMD Microphones [64], [65]	Deep learning voice analysis
Body Language	VIVE Tracker 3.0 [66], Ultimate Tracker [67], Wrist Tracker [68] and VRIK module [69]	AI motion analysis systems
Physiological and Respiratory	Biosignalsplux Toolkit [70] with NMFSs [71], EEG, GSR	NNs for signal processing
Eye Movements	VIVE Focus 3 Eye Tracker [72] and Microsoft Hololens 2 [64]	ML algorithms for eye data
Contextual	Environment sensors	DL for context analysis

existing models by its ability to interpret and adapt to users' emotional states in real time. The overarching objective of EIM is to forge a virtual realm that maximises users' emotional immersion by imbuing avatars with emotional intelligence as well as crafting emotionally resonant virtual environments, thereby elevating the overall quality and genuineness of user interactions and experiences. Such deeper emotional connections within EIM enable design of users' emotion-aware content and applications e.g. empathetic virtual caregivers complementing physicians in various healthcare-related interventions, creative learning environment tailored according to learners' emotional characteristics and needs [56], etc.

We present the framework of EIM as illustrated in Figure 3. The proposed framework underscores affective computing as a fundamental and integral component of EIM, signifying its pivotal role in enabling a digital realm capable of understanding, interpreting and adapting to users' emotions. Affective computing represented by the white square in the framework employs advanced metaverse technologies to capture a spectrum of users' affective states including facial expressions, body language, voice tones, and physiological signals and processes this data through a suite of AI technologies, to enable a transformative EIM, as detailed in Table 6. State-of-the-art technologies for integrating affective states include VIVE Focus 3 Facial Tracker [62] and SALSA Lip Sync Suite [63] for facial expressions, Meta Quest 3 [65] and Microsfot Hololens 2's [64] built-in microphones for voice, VIVE Ultimate Tracker [67], VIVE Tracker 3.0 [66], VIVE Wrist Tracker [68] and VRIK module of Final IK [69] for body language and gestures. Physiological and respiratory signals are acquired through Biosignalsplux wireless toolkit [70] with Nanomaterials-based flexible sensors (NMFSs), EEG and GSR, while eye movements are tracked using VIVE Focus 3 Eye Tracker [72] and Microsoft Hololens 2 [64]. Figure 4 emphasises the significance of affective computing within the EIM framework that illustrates how affective computing in EIM can elevate the quality of user interactions, aligning responses with their emotional states. This section delineates an array of potential strategies for embedding affective computing into the metaverse. For a concise overview, Table 5 encapsulates the essence of this section.

A. Authentic Human-AI Interaction

The metaverse has transformed human-AI interactions with its virtual space, offering modalities from language to multimodal engagement. AI avatars overcome language barriers, enriching communication. Within this dynamic landscape, AI avatars serve as linguistic bridges, seamlessly connecting users who would otherwise face language barriers. The pioneering work of Miller et al. [76] with ParlAI, a versatile framework for training and testing diverse dialogue models, underscores the ever-evolving nature of dialogue systems through multi-task learning, human evaluation, and reinforcement learning. The pursuit of goal-driven training and multi-domain dialogue policy learning [77] further enhances the versatility of AI-powered conversations. In the realm of decentralised digital assets, Wang et al. introduce a framework for decentralised multimodal interactive NFTs (Non-fungible tokens) [78], leveraging multimodal classifiers to analyse social media posts encompassing text and images. A unified vision-language pre-training (VLP) model, as proposed by Wang et al., demonstrates the metaverse's capacity to excel in image captioning and visual question-answering tasks [79].

Meanwhile, the groundbreaking work of [80] showcases state-of-the-art performance across multiple tasks through deep reinforcement learning-based multi-task learning. The metaverse extends its capabilities even further into embodied interaction, as seen in embodied question answering (EQA) [81], where AI agents navigate 3D environments and address questions, highlighting the fusion of language understanding and goal-oriented navigation. As we venture into the metaverse's future, the integration of emotional intelligence emerges as a compelling catalyst, promising to imbue these interactions with nuanced emotional understanding and responsiveness, thereby enriching the experience in unprecedented ways in an EIM.

EIM can revolutionise interactions with AI-driven characters within the metaverse by infusing them with empathy and heightened responsiveness to users' emotions and needs. Emotionally intelligent avatars possess the ability to discern emotional cues like facial expressions, tone of voice, body language, and gestures, resulting in virtual interactions that are more nuanced and authentic [60]. This, in turn, elevates the quality of social engagements within the metaverse. In the context of EIM, Figure 4 vividly portrays how digital entities respond empathetically to users' emotions, making human-AI interactions remarkably genuine and authentic. Recent research, such as that by [61], delves into the enhancement of real-time multi-sensory interactions (RMSIs), especially when accessed through VR headsets, emphasising the profound impact of users' emotions. The concept of human-like interaction metaverse (HIM) [40] promotes an adaptive virtual world, guided by users' emotional states, where interactions closely mimic human-like responses. Beyond this, affective computing has the potential to craft emotionally responsive narratives and experiences within the metaverse. Virtual stories and environments can dynamically adapt based on users' emotional engagement, ushering in a new era of more impactful and personalised storytelling.

B. Enhanced User Immersion

Metaverse emphasises social interactions enabling users to engage and interact in the digital world through the embodied representation of themselves: avatar [82]. Conventional metaverse environments focus on high-quality visuals, interactivity, and engaging stories to create a sense of presence but often miss delivering personalised, emotionally rich experiences. However, as highlighted by Zhang et al. [83], more fluent and realistic interactions can be achieved by tailoring digital persona dialogues to the profiles of the participants involved Motivated by the dual desires of self-verification and self-enhancement, users can personalise their avatars in appearance and behaviour. Figure 5 illustrates Cairns' 3-level immersion model [84], which outlines the progression from initial engagement to complete immersion in various activities. This model underscores the importance of understanding emotional dimensions in crafting and evaluating user engagement, emphasising the need to consider emotional immersion in metaverse technologies. Avatars serve as digital representations, granting users control and flexibility over their virtual identity [85]. The experiences of virtual counterparts in the metaverse can directly mirror users' real-world learning behaviour.

Various studies [86] have highlighted the direct relationship between users' immersive experiences and affective environments in VR. Expanding on this notion, the concept of HIM is introduced in [40], where avatars tailor their responses based on users' human behaviour vectors, including emotional states. Daneshfar et al. [41] have explored a novel echo state network (ESN) structure and its integration into the metaverse

This article has been accepted for publication in IEEE Open Journal of the Computer Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCS.2024.3389462

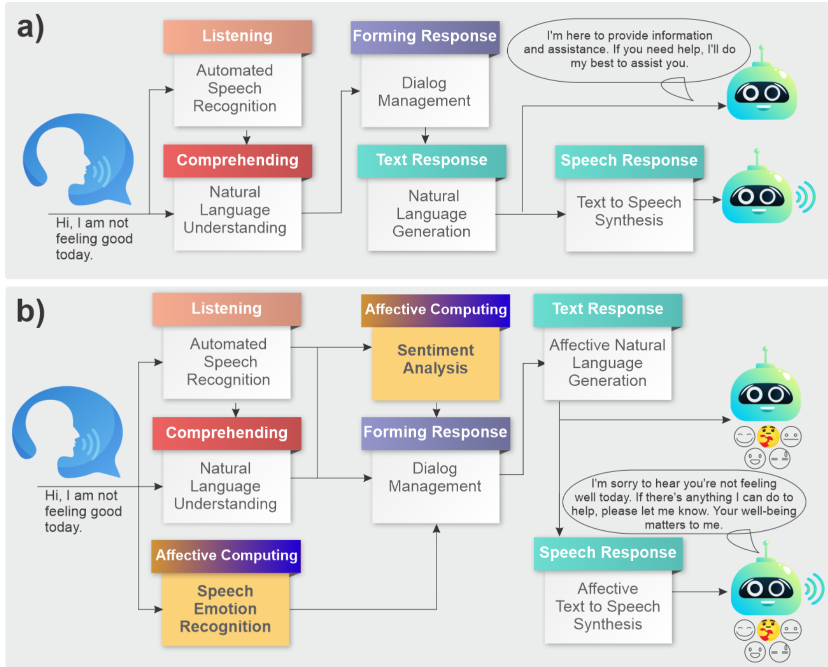

FIGURE 4: Comparing human-AI interactions in the traditional metaverse with the EIM. Here a) represents traditional metaverse interaction, lacking emotional comprehension, and provides basic responses. Whereas b) shows EIM interaction, enriched with affective computing, offers empathetic understanding and responses.

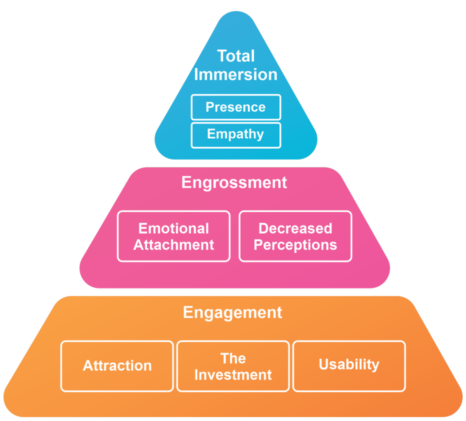

FIGURE 5: Enhanced immersion can be achieved by transforming the traditional metaverse into EIM platform.

for speech emotion recognition. Furthermore, a framework for an emotionally enhanced metaverse, driven by users' physiological signals, is presented in [38]. This framework has been validated through a proof of concept, demonstrating the synchronisation of users' emotions in virtual reality. These emotionally meaningful user-digital interactions promise to greatly enhance metaverse immersion.

The integration of wearable devices like wristbands into the metaverse, as demonstrated in the work by Rojas et al. [42], offers an avenue to enhance empathy in avatars and non-player characters (NPCs). These technologies enable avatars and NPCs to display realistic emotions, including empathy, thereby contributing to more authentic and emotionally resonant virtual gatherings. By leveraging empathy-enabled avatars, smart cities digital twins [87] can enhance virtual services and manifest immersive virtual tours of urban spaces. The empathetic interactions also Research shows that VR enhances empathy in users who interact with avatars representing different races, genders, or age groups [88]. Herrera et al. [89] also demonstrate that people experiencing embodied interactions with underprivileged individuals are more likely to exhibit long-term empathy compared to those who do not have such experiences. These studies collectively emphasise that infusing emotional intelligence into traditional metaverse experiences enables avatars to adeptly respond to users' emotions, enhancing the overall sense of immersion [90] and enabling EIM.

C. Emotion–Aware Resource Dimensioning

Within the metaverse, a prominent challenge is the efficient allocation of resources among users, all while meeting demanding requirements for high rendering quality and ultra-low latency across various applications. Studies present different techniques for resource allocation. For instance, Du et al. propose an attention-aware resource allocation scheme that customises metaverse services, involving a two-step process that predicts users' interests and enhances the quality of experience (QoE) [91]. Another framework, Metaslicing, groups applications with common functions into MetaInstances, addressing high resource requirements [92]. Zhou et al. introduce a federated learning-based resource allocation strategy [93], which optimises the weighted combination of energy consumption, model accuracy, and execution latency, outperforming incumbent benchmarks. Chua et al. [94] minimise transmission latency during the download of 3D world graphics and execution latency through a deep reinforcement learning-based approach and MDTRAP resource allocation mechanism, respectively.

As highlighted, resource dimensioning in the traditional metaverse optimises computational resources for rendering tasks like 3D graphics, physics simulations, and real-time interactions, ensuring peak performance, scalability, security, cost-efficiency, and future growth. However, resource dimensioning alone falls short in the evolving metaverse landscape. It lacks emotional intelligence, preventing avatars and virtual environments from effectively responding to user emotions, thus limiting immersion and authenticity. The integration of emotional intelligence enhances user engagement, empathy, and the overall metaverse experience, ultimately contributing to the creation of an EIM. For instance, one effective approach involves leveraging users' emotional states, inferred from various audiovisual cues, during sessions to optimise resource allocation. For example, resources for users expressing negative emotions like anger, sadness, and annoyance can be augmented to enhance the overall user experience [8].

D. Meditation and Adaptive Health Surgeries

Since the outbreak of COVID-19 in 2019, the healthcare sector has undergone a remarkable transformation, embracing innovations such as telehealth and robotic surgeries [17], [95], [96]. Post-2020, the adoption of telemedicine skyrocketed to an impressive 95%, up from a mere 43% pre-pandemic [97]. The metaverse now emerges as a promising frontier in healthcare, offering telemedicine, remote consultations, and immersive health experiences in a dynamic 3D environment, greatly enhancing accessibility and convenience in healthcare services [98], [99]. This integration within the metaverse facilitates uninterrupted patient monitoring, enables global expert consultations, and bolsters the security of electronic healthcare data through blockchain technology [45], [100]. Moreover, it empowers precise medical diagnosis through cutting-edge technologies [101] and enhances clinical interventions, as exemplified by the positive outcomes observed with 360-degree immersive videos in mental health care [102] and distraction therapy in high-intensity pain cases using immersive VR environments [103]. The incorporation of the metaverse into healthcare signifies a groundbreaking step toward more accessible, personalised, and effective medical services.

Metaverse-based healthcare enhances human-computer interaction and merges physical and virtual realities. It supports medical education through customised virtual scenarios for learning without real patient interaction and allows practice on simulated patients to understand treatments before actual application [104]. For instance, authors in [105] advocate the idea of creating ML-enabled digital twins of breast cancer for diagnostic and therapeutic purposes. Moreover, research is also focused on the digital twinning of microorganisms in metaverse that helps reduce diagnostic errors [106]. However, it's important to note that the metaverse's current limitations, especially in healthcare, include the absence of emotional intelligence and the human touch element. While it offers impressive technical capabilities, it may struggle to fully replicate the emotional nuances and empathy that human healthcare providers can offer. Patients often require not just medical treatment but also emotional support, which can be challenging for a purely digital environment to provide. Affective computing can help transform the traditional metaverse into an EIM, bridging the gap between technical capabilities and the vital human elements required in healthcare. EIM offers a wide range of benefits in healthcare, encompassing realistic and enhanced virtual doctor-patient interactions and various treatments for phobias. By assessing users' affective states, personalised therapeutic interventions can be delivered to enhance patients' comfort levels and sense of presence, resulting in significantly positive outcomes. Furthermore, affective digital entities, trained to convey appropriate emotions, can complement clinicians in patient triage and healthcare interventions. Loveys et al. advocate for the use of affective virtual humans in psychological and behavioural analysis of patients [44]. Virtual humans tailored to match a patient's community can yield favourable results. EIM can also encourage users to participate in mindfulness exercises, stress reduction activities, and emotional well-being practices [45].

Additionally, EIM can serve as a platform for immersive virtual therapy sessions, where affective computing detects users' emotions and adjusts therapy content in real-time [107]. This approach provides personalised emotional support, making therapy more effective and accessible. For example, cyber-therapy adapts itself based on psycho-physiological signals to reduce stress and anxiety in patients [47]–[49]. When integrated with patients' affective behaviour obtained through physiological sensors, VR can effectively address a range of phobias. In another work, Gromala et al. utilised VR-enabled mindfulness-based stress reduction (MBSR) meditation techniques for the therapy of chronic pain patients and showcased the effectiveness of such techniques in relieving pain when compared to conventional methods [46].

E. Personalised Gameplay and Content

The latest advancements in key enabling technologies of the metaverse, including XR, have profoundly shaped the landscape of the gaming and entertainment industry. XR, which encompasses VR, AR, and MR, offers a persistent and interconnected digital world that significantly enhances users' immersive gaming experiences. Researchers have delved into digital embodiment and avatar customisation in various virtual worlds. In addition to gaming, blockchain-based decentralised virtual spaces have been proposed for collaborative purposes [108], and even seemingly simple games like Hide and Seek have been used to study multi-agent dynamics [109]–[111]. Furthermore, Stanica et al. have proposed the innovative Immersive Neurorehabilitation Exercises Using VR (INREX-VR) system within gamified settings, aiming to rehabilitate physical functions inhibited by neurological disorders [112].

This article has been accepted for publication in IEEE Open Journal of the Computer Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCS.2024.3389462

FIGURE 6: EIM can host impactful games, characterised by enduring and delightful experiences, that combine all three levels of emotional designs.

This pioneering system has shown remarkable results in improving physical movements, as evidenced by enhanced accuracy and positive user feedback. However, it is crucial to acknowledge that without emotional intelligence, the gaming metaverse is inherently limited.

Emotional intelligence introduces a novel dimension to gaming by tailoring gameplay elements based on users' emotional responses, creating a more engaging and personalised experience. For instance, if negative emotions like frustration or anger are detected, the game's difficulty level can be adjusted to maintain user engagement and motivation. Characters and storylines can also be dynamically modified based on emotional data captured by affective computing technologies. Dozio et al. have developed emotionally adaptive virtual environments designed to elicit various emotions among users [50]. Similarly, Kerdvibulvech et al. have modelled human emotions to personalise metaverse content, enhancing overall immersion [51]. Video content can be suitably tailored through valence monitoring using innovative facial masks as proposed in [52]. Following Norman's principles of emotional design [113], gameplay elements can be crafted to evoke three levels of emotions, including reflective emotions that leave lasting impressions on users. EIM serves as a versatile platform for hosting impactful games that offer a delightful user experience, combining all three levels of emotional design as depicted in Figure 6. Interested readers can explore [114], [115] for a comprehensive overview of various emotional designs featuring exclusivity and perceived rarity.

F. Emotion-driven Virtual Education

Online education has rapidly progressed since the outbreak of the COVID-19 pandemic in 2020, which hindered the physical movement of people. However, traditional online teaching methods through mobiles and digital screens are not at par with face-to-face / offline education in terms of convenience and effectiveness. Metaverse couples trainers and trainees in an immersive virtual environment, while providing an opportunity for hands-on training with minimal cost [116]–[118]. Additionally, students are able to acquire complex educational concepts through experiential education in metaverse thus retaining information for long duration [18]. For instance, a virtual STEM class followed by hands-on training is organised for students to grasp the concept of radioactivity and associated safety precautions. Hearing-impaired people can get the opportunity to learn sign language from an avatar mimicking human gestures [119]. A hybrid learning model comprising a multi-user virtual environment and traditional teaching methods is explored in [120]. Collins [121] presents a comprehensive review regarding emergent behaviours of the metaverse and the envisaged role of higher education in the expansion of the digital world. Khan et al. [122] devise a 3D immersive training for enhancing users' awareness about road safety. However, without emotional intelligence, education in the metaverse may lack the crucial element of empathetic and nuanced interactions, potentially limiting the depth of understanding and human connection that traditional education can provide.

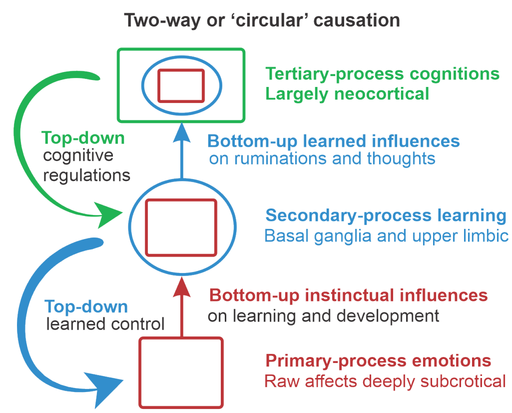

FIGURE 7: A conceptual framework depicting the significance of primary affective states (lower order brain functions) in learning and memory processes [123]. EIM-based virtual education applications can employ this framework to enhance the learning experience.

Affective computing empowers educators to tailor their teaching methods and learning approaches to align with students' emotional needs, ultimately fostering a more engaging and effective learning experience [53]. In the context of EIM, this adjustment, guided by emotional cues, can result in a more efficient and captivating learning journey, effectively closing the divide between conventional in-person education and the immersive digital universe of the metaverse. For instance, a virtual PBL system helps teachers gauge the efficacy of PBL through analysis of students' emotions to questions asked by keeping track of their eye blinks [54]. Hwang suggests an education model named Thinkering, Making, Improving, Ownering, Sharing (TMIOS), which allows students for digital content creation and gauges the effectiveness of the proposed model by observing students' emotions [55]. Technology-enhanced Edu-metaverse framework is proposed in [56] that enhances engagement in human-machine interactions and enables learners to pursue collaboration and creation-based learning. Creative engagement observes learners' cognitive and affective states and allows them to explore more flexible teaching methods. Moreover, the role of affective states in associative learning and memory phenomena is emphasised in [123]. The conceptual framework comprising two-way emotional control presented in Figure 7 highlights that lower-order primary affective processes influence and augment higher-order brain functions including learning and memory processes. The same may be utilised as a guiding framework by future EIM-based virtual education applications.

G. Product Refinement and Feedback

With rapidly evolving digital marketing, the metaverse has emerged as a game changer for business companies in the world. It enables brands and businesses to extend their reach and engage their target audience in a more interactive and entertaining way [124]. It is expected that the metaverse economy would overshoot the real-world economy with overall growth peaking at $12 trillion [125]. Brands are exploring the digital realm to capitalise on the business and marketing opportunities being offered by metaverse [126]. Companies like Gucci and Walmart have forayed into the metaverse to provide an immersive experience along with incentive-based interactive games to customers [127], [128]. Jeong et al. overcome various disadvantages of online shopping including lack of experience with products/items through an innovative e-commerce platform integrating metaverse with live-commerce [129]. The proposed platform employs the concept of "experience-to-buy" instead of "click-to-buy" and results in a more satisfying experience for shoppers during online purchasing. However, it is important to note that without emotional intelligence, the metaverse may struggle to build the emotional connections necessary for fostering meaningful customer engagement and refining products and services, potentially limiting its full potential of digital business and marketing.

In an EIM, brands can leverage affective computing to collect real-time emotional feedback from users, gaining valuable insights into how their products and experiences resonate with their audiences. This capability empowers businesses to refine their offerings in alignment with users' emotional responses. This approach aligns with Norman's concept of crafting emotional designs, encompassing visceral, behavioural, and reflective elements, which leads to enduring, engaging, and delightful product experiences [113]. For instance, delightful experiences can be curated through expressive imagery and personalised marketing promotions, fostering an enhanced sense of ownership and relatability. Moreover, research by authors such as [57] emphasises the optimisation of the shopping experience based on users' physiological signals, while studies conducted by [58], [59] underscore the pivotal role of affective brand attributes in shaping positive perceptions among students and users regarding corporate brand image. This integration of emotional intelligence in the metaverse provides a powerful tool for businesses to create more meaningful and engaging interactions with their audience.

H. Impact and Practicality of EIM in Diverse Industries

EIM can transform service offerings across multiple industries, from healthcare to urban metaverse cyberspace, by improving user experience and engagement. Its impact promises a new era of innovation and efficiency. Below, we list down EIM's practical applications in various sectors, with Table 7 summarising its effects.

1) **Healthcare:** EIM enhances mental health care by tracking emotional well-being, enabling tailored treatments for better patient outcomes and personalised care. Its ability to detect mental health conditions early improves healthcare quality, allowing timely interventions for effective mental well-being management.
2) **Learning and Education:** EIM facilitates personalised education by analysing students' emotions in real-time, enhancing engagement and enabling tailored instruction. This adaptability improves learning outcomes by meeting each student's unique needs effectively and promptly.
3) **Retail:** EIM leverages emotion analysis to revolutionise personalised shopping, offering tailored product suggestions and improving overall service for a more engaging customer experience. This integration significantly enhances customer satisfaction.
4) **Gaming and Entertainment:** EIM uses emotional recognition technology to customise in-game elements and challenges in real-time, creating a personalised and engaging gaming environment. This adaptability enhances player engagement and boosts user retention by responding to individual emotional reactions.
5) **Automotive:** EIM can effectively monitor drivers' alertness and emotional states, enhancing safety and enabling the creation of responsive vehicle environments tailored to individual needs and conditions.
6) **Customer Service:** By leveraging advanced emotion analysis capabilities, EIM can discern and interpret customer sentiments, allowing for the delivery of empathetic service interactions. This not only enhances the effectiveness of issue resolution but also cultivates more robust and enduring customer relationships.
7) **Human Resources:** EIM enhances employee emotional well-being, boosting workplace productivity and morale. This leads to a more dynamic, collaborative environment and weaves a tapestry of satisfaction and engagement, elevating organisational success.
8) **Media and Marketing:** Leveraging EIM in marketing allows businesses to analyse emotional responses to campaigns, refining messaging and targeting. This optimises promotions and deepens understanding of consumer sentiment, leading to more effective, resonant campaigns.
9) **Urban Metaverse Cyberspace:** EIM can significantly impact and transform urban metaverse cyberspace [130], [131] in several ways. For instance, EIM not only leads

TABLE 7: Impact and practicality of EIM across various industries

Industry	Impact and Practicality of EIM
Healthcare	*Personalised Mental Health Treatments*: EIM monitors individuals' emotional well-being to tailor mental health treatments, enhancing patient outcomes and delivering personalised care.
Learning and Education	*Adaptive Emotional Learning Environments*: EIM provides learning settings that dynamically adjust to students' emotional states, significantly boosting engagement.
Retail	*Emotion-Driven Personalised Shopping*: EIM enhances the shopping experience by suggesting products and services tailored to customer emotions, increasing satisfaction.
Gaming & Entertainment	*Emotionally Responsive Gaming*: EIM offers gaming experiences that adapt to players' emotional states, leading to heightened engagement and improved user retention.
Automotive	*Emotion-Aware Safe Driving*: By monitoring drivers' alertness and emotional states, EIM contributes to driving safety and fosters a responsive vehicle environment.
Customer Service	*Empathetic Service Interactions*: EIM's emotion analysis capabilities enable more empathetic customer service, improving resolution effectiveness and strengthening customer relationships.
Human Resources	*Employee Emotional Well-Being*: EIM enhances workplace productivity and morale by focusing on the emotional well-being of employees.
Media and Marketing	*Targeted Emotional Campaign Analysis:* EIM evaluates emotional responses to marketing campaigns, enabling more precise message targeting and campaign refinement.
Urban Metaverse Cyberspace	*Improved Emotion-Aware Virtual City Services:* EIM not only leads to emotion-tailored education and health treatments but also fosters more personalised and empathetic interactions in smart city digital twin, manifesting improved virtual city services and immersive virtual tours of urban spaces.

to emotion-tailored education and health treatments but also fosters more personalised and empathetic interactions in smart city digital twins, manifesting improved virtual city services and immersive virtual tours of urban spaces. Moreover, taking into account the emotional states of the participants, virtual meetings and collaborations become more engaging and productive, leading to better decision-making and community engagement in urban development projects. The emotional cues of urban metaverse residents also facilitate the creation of more emotionally responsive urban spaces that promote well-being and social cohesion.

IV. Challenges in Developing EIM

Affective computing, crucial for recognising and interpreting human emotions, is key to evolving the traditional metaverse into an EIM. While its potential is vast, realising emotional intelligence in the metaverse faces notable challenges that require careful attention and resolution for effective integration

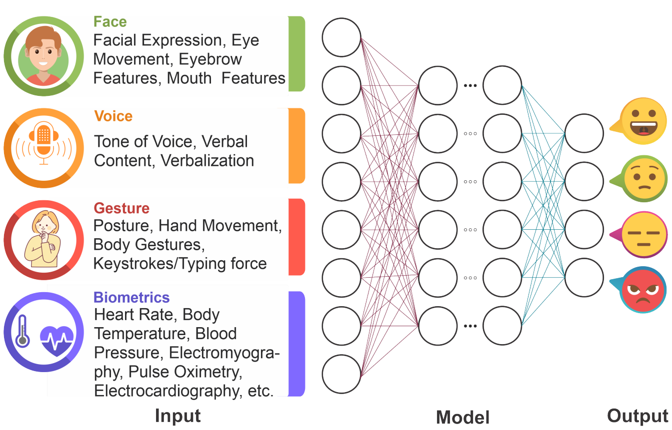

FIGURE 8: Multimodal affective computing, where deep learning models can be trained using the fusion of multimodal data.

A. Performance considerations of emotion recognition

Accurately recognising emotions in EIM presents a notable challenge. Variations in facial expressions, speech patterns, body language, and physiological signals within the EIM can impact the reliability of emotion detection systems. Facial emotion recognition, deemed a prominent facet of emotion recognition, is widely recognised as a crucial component of EIM [132]. Regrettably, computer vision and machine/deep learning techniques for facial emotion recognition are not without their inherent limitations. The presence of variables such as lighting conditions, facial obstructions, head poses, and subtle nuances in expressions impedes the accuracy of emotion detection in EIM [73], [133]. Moreover, cultural variations in the display of emotions introduce additional complexities. The performance of these visual systems experiences a significant decline when deployed in real-world environments, as opposed to controlled laboratory settings [132], [133]. Consequently, relying solely on facial expressions may prove inadequate in understanding the experienced emotion.

Moreover, identifying emotions conveyed through body language/posture presents a substantial hurdle. This necessitates the use of costly wearable devices and advanced algorithms capable of discerning movements and subsequently translating them into the user's emotions. The need for multiple wearable devices to accurately capture the movements of various body parts may pose an inconvenience to users [134]. Additional constraints encompass device portability and resource limitations, such as restricted battery capacity, storage, and processing speed [135]. Furthermore, the synchronisation of collected data proves to be a significant challenge, especially in the presence of unreliable network connectivity, impacting real-time emotion recognition [136].

Speech signals offer a readily available source that holds potential for utilisation in EIM for understanding emotions and adapting accordingly. Research communities have made commendable progress in achieving competitive performance by developing speaker-independent emotion recognition systems. Nonetheless, achieving language-invariant emotion detection remains a significant hurdle [137]. Despite emotions being deemed language invariant, the efficacy of emotion recognition wanes when assessed across diverse language-emotional corpora. To address this concern, the concept of leveraging representations derived from few-shot learning emerges as a potential strategy for adapting emotion recognition systems [138]. This approach necessitates a limited number of samples from the target language dataset. Nevertheless, the existing corpora cover only a small fraction of the world's languages, in stark contrast to the multitude of languages spoken globally [75].

To rectify the limitations observed in vision, body language, and speech-based emotion recognition systems, scholarly literature proposes the adoption of multimodal approaches [139]. These approaches advocate for the simultaneous analysis of video and audio inputs, harnessing the powerful synergy between facial displays and paralinguistic speech features. By incorporating both modalities (as shown in Figure 8), a more resilient and accurate emotion recognition system can be achieved, particularly when confronted with intricate or amalgamated emotions in real-world scenarios. This multimodal strategy enriches the emotional context by surpassing the confines of a solitary input type, thereby enhancing the overall efficacy of the recognition process [139].

B. Securing EIM Against Adversarial Attacks

Research has shown challenges in designing robust affective computing systems due to adversarial examples, which demonstrate deep model vulnerability to imperceptible perturbations [75], [140]. Prominent attacks craft perturbation noise guided by targeted output gradients, including fast gradient sign method [141], Jacobian-based saliency map attack [142], and DeepFool [143]. Affective computing systems are also susceptible to such attacks [140], [144]. Attack success underscores lack of robustness in DL model representations. As described by Li et al. [145], they proposed a robust detection-deactivation method for adversaries in EIM. Their method can restrict and separate access of potential malicious participants, meaning it can block vulnerable backpropagation and GAN attacks. The motivation was drawn from anomaly detection systems. Though presenting a potential solution, an adaptive adversary can learn this behaviour and modify attack vectors to inject EIM model vulnerabilities. This shows adversarial attacks significantly threaten affective computing in EIM. Malicious actors can manipulate input to deceive emotion models, enabling incorrect responses or unauthorised access to sensitive user emotions. Attacks can undermine trust, compromise privacy, and enable harm. To counter susceptibility, training DL models to generate robust representations against transformations has been explored [146], [147]. Notably, very deep architectures demonstrated SER robustness. Yet, further exploration is needed into what DL models capture from speech and how to define adversarial examples without adversary knowledge. This will help develop robust defences tailored for EIM using techniques like adversarial training, anomaly detection, and ensembling [148], [149].

C. Privacy and Security Considerations in EIM

EIM draws upon various data sources from diverse sensors, cameras, microphones, etc. Assuring the privacy and security of users' emotional data is paramount, requiring robust encryption, secure data handling, and user consent mechanisms [150]. Emotion tracking, inference of private state, surveillance, data leakage, data manipulation, lack of control over automatic inference, identifying despite anonymity, etc., are a few data privacy challenges that require immediate attention [151]. Research should focus on developing robust encryption techniques, user-controlled data-sharing mechanisms, and secure storage practices to address privacy and security concerns of affective computing in EIM [152].

While the literature has indeed touched upon privacy challenges in the metaverse, as evidenced by notable works such as [153]–[156], it is important to acknowledge that the introduction of affective computing exacerbates the privacy predicament. Therefore, it is imperative that we closely examine the unique issues that arise when affective computing is integrated into EIM, as this intersection presents heightened privacy challenges that warrant diligent attention and understanding. Indeed, while there have been some endeavours to comprehend the challenges surrounding the rise of privacy issues of utilising affective computing component techniques in the metaverse [8], [157], [158], it is worth noting that the empirical testing and validation of these issues are relatively scarce in the existing literature.

D. Addressing the Uncanny Valley Phenomenon in EIM

Creating an authentic and natural interaction involving affective computing within virtual environments, while avoiding the uncanny valley phenomenon, necessitates a delicate balance between preserving authenticity and ensuring user comfort [159]. The uncanny valley effect occurs when avatars, possessing human-like attributes, exhibit subtle deviations that evoke a sense of unease among users [160]. To address this challenge, a comprehensive exploration of strategies is imperative, encompassing not only the enhancement of avatar design and animation but also the optimisation of emotional representation. Transcending the uncanny valley requires researchers to delve into refining avatar design, aiming to achieve a closer resemblance to humans while still being well-received by users [161]. This may involve meticulous adjustments to facial features, body proportions, attire, and other visual elements [162]. Furthermore, investigating innovative animation techniques that emulate human movement patterns can contribute to a more genuine and lifelike user experience [163].

Additionally, the concept of portraying emotions plays a pivotal role in tackling the uncanny valley issue. Techniques that accurately capture and convey emotional states through avatars' facial expressions, gestures, and vocal nuances can significantly heighten emotional authenticity [164]. The integration of algorithms for emotion recognition and real-time adaptation empowers avatars to respond appropriately to users' emotional cues, establishing a deeper connection and mitigating any potential discomfort. Achieving a harmonious synthesis among the intricate facets of avatar design, animation, and emotional representation is a multifaceted endeavour. Research endeavours should be focused on unravelling the intricacies of the uncanny valley while embracing novel technologies to cultivate emotionally resonant interactions that are both immersive and comfortable for users. Through a systematic exploration of these methodologies, the foundation can be laid for enriched and emotionally fulfilling interactions in EIM.

E. Modelling Dynamic and Evolving Emotional States in EIM

The EIM offers long-term interactions and experiences, leading to dynamic shifts in users' emotional states over time. Affective computing models in EIM need to adapt to these changing emotional dynamics to maintain accurate recognition and response. Ensuring that the system adapts without causing user discomfort due to sudden emotional changes or failing to recognise gradual shifts presents a significant challenge [165].

In order for EIM to thrive, it is paramount to design stochastic models for continuously evolving emotional states and dynamics. The non-stationarity and long-context dependence inherent in emotional expressions pose significant challenges when it comes to modelling and interpretation. Traditional approaches may struggle to capture the intricate nuances and the evolving relationship between different modalities used to convey various emotions over time in EIM [166]. To address these complexities, a multi-model integrated solution is necessary. Such a solution should be capable of understanding and incorporating the nuances and dynamics of emotional expressions within EIM [166], [167]. By integrating multiple models that account for various modalities, including facial expressions, body language, voice tone, and other contextual cues, we can strive for a more comprehensive and accurate understanding of users' emotions [168]. Furthermore, this integrated solution should be designed to capture the temporal aspect of emotional expressions, recognising how emotions evolve and change over time. By considering the historical context and the progression of emotional states, we can develop a more nuanced understanding of the emotional landscape within the metaverse [51].

F. Challenges of Real-time Emotion Processing in EIM

Affective computing in the EIM requires real-time processing to provide seamless and responsive emotional interactions. However, achieving low-latency emotion recognition and generation poses technical challenges, especially in resource-constrained environments. Federated computing architecture [169] holds great potential in mitigating latency issues in EIM by distributing certain aspects of data analysis to the edge, such as the headset processor, thereby circumventing transmission losses and delays. However, it is important to note that a federated architecture necessitates a delicate balance between performance and accuracy across various devices [170]. By effectively managing this trade-off, we can optimise the overall efficiency and effectiveness of data analysis in EIM. When it comes to EIM, addressing communication issues and network latency becomes even more critical. The nature of short-lived and rapidly changing emotional states necessitates an extremely low latency for accurate and real-time processing. To date, providing real-time processing and low latency for affective computing applications in the EIM remains an open question that requires further exploration. Tang et al. [171] presented a forward-thinking roadmap for the design of 6G network which can be opted as a potential solution for mitigating communication challenges and latency concerns in EIM. Researchers must grapple with this challenge and seek innovative solutions to enable efficient real-time emotion processing. This can be achieved by investigating novel architectures, compression techniques, and hardware optimisations explicitly tailored for the metaverse environment.

V. Future Prospects for EIM Research

This section highlights key areas requiring attention for EIM's successful implementation and its related technologies. These future directions are crucial in developing a robust and ethically guided framework for EIM.

A. Foundational Models and EIM

The recent convergence of large language models and large audio models with the capabilities of 6G technology is ushering in a new era of real-time interactions. These models such as GPT-4 [174] and SeamlessM4t [175] can enable seamless and instantaneous communication between users, transcending language barriers through natural language and speech processing and voice translation and recognition [176]. With the support of 6G's low latency, and high bandwidth, these models can provide lightning-fast responsiveness in applications ranging from gaming to healthcare. In summary, future work in this domain focuses on improving model efficiency, addressing ethical considerations, enhancing emotional intelligence, and fully utilising 6G technology to redefine real-time interactions. This includes developing affect-aware large audio models for real-time emotion recognition [176], benefiting applications like mental health support, virtual therapy, and personalised entertainment in EIM.

B. Cultural and Contextual Sensitivity in EIM using Deep Reinforcement Learning

EIM research should explore continuous adaptation techniques capable of modelling users' emotional trajectories and dynamically adjusting affective responses. These approaches enable more natural and empathetic interactions to unfold seamlessly over extended periods within the EIM. Additionally, it's imperative to recognise the influence of cultural and contextual variations on emotional expressions. Emotions are expressed differently across cultures and within various contexts, potentially leading to misinterpretations by models that lack exposure to diverse data. To mitigate this challenge, researchers can explore developing methods that train models using diverse and representative datasets to ensure that EIM

This article has been accepted for publication in IEEE Open Journal of the Computer Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCS.2024.3389462

TABLE 8: Brief description of the EIM-related challenges, existing solutions, existing gaps, and future direction

Challenges	Solution (Explored in Literature)	Existing Gaps	Future Directions
Performance considerations of emotion recognition in EIM	**(1):** Independent emotion recognition **(2):** Multi-modal emotion recognition [139]	**(1):** Limited annotated data **(2):** Language detection	**(1):** Cross-modal integration **(2):** leveraging Foundational models (GPT-4, SeamlessM4t, Whisper, etc.)
Securing EIM against adversarial attacks	**(1):** Using Deep architectures for SER in EIM [172] **(2):** Using preemptive defences [140], [173]	**(1):** Poor Performance **(2):** Lack of defences against adaptive adversary	**(1):** Adaptive defence against adversarial attacks **(2):** Multi-modal adversarial defences for EIM
Privacy and security considerations in EIM	**(1):** Proactive security approaches **(2):** Audits and Standards **(3):** Ethical Data Use Practices **(4):** Regulatory Compliance	**(1):** Lack of empirical testing **(2):** Ethical EIM **(3):** Universal data collection and processing guidelines	**(1):** Empirical privacy testing and validation in EIM **(2):** Ethics by design
Addressing the Uncanny Valley phenomenon in EIM	**(1):** Improving avatar design **(2):** Improved animation techniques **(3):** Optimising emotional representation.	**(1):** Disconnection between avatar realism and user comfort **(2):** Human-like emotion portrayal	**(1):** Integrating refined avatar design with lifelike animation **(2):** Accurate emotional expressiveness.
Modelling dynamic and evolving emotional states in EIM	**(1):** Understanding the dynamics of emotions in EIM **(2):** Designing responses for the emotional cues of avatars in EIM	**(1):** Limited capability to dynamically model and adapt to the changing emotional states of users. **(2):** Lack of annotated data for dynamic and long-context emotions	**(1):** Multi-modal solutions for understanding and incorporating emotional nuances in EIM.
Difficulties in the real-time processing of emotions in EIM	**(1):** Federated computing architecture for EIM **(2):** 6G technology combined with large language and audio models for EIM	**(1):** A delicate balance between performance and accuracy is required **(2):** Hardware and computational resources issues in federated environment.	**(1):** Improved federated networks for EIM **(2):** 6G systems with low latency and high bandwidth for swift responsiveness for EIM applications .

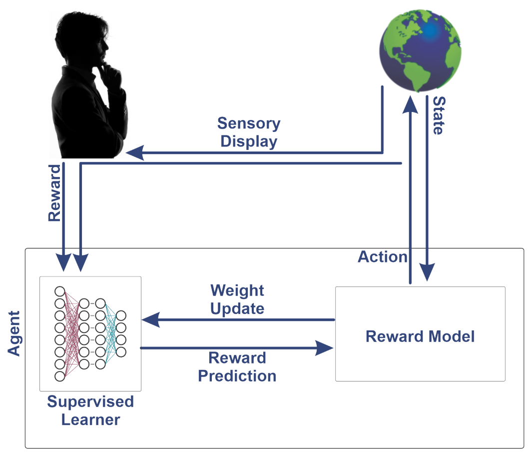

FIGURE 9: Reinforcement learning from human feedback (RLHF), which optimises learners through human feedback and guides model training iteratively until the desired performance is achieved.

is culturally sensitive and contextually accurate. Furthermore, within the evolving metaverse landscape, deep reinforcement learning (DRL) can be considered a crucial component in comprehending and adapting to long-term emotional dynamics [177]. DRL's distinctive framework can prioritise sustained emotional satisfaction over immediate gains and can expertly balance the exploration-exploitation equation [178]. This helps empower agents to continuously experiment with emotional responses while harnessing effective strategies over time according to cultural and contextual factors variations. The incorporation of memory mechanisms in DRL models aids in preserving past emotional states and interactions, ensuring that agents maintain context during extended interactions. Interestingly, a primary factor contributing to ChatGPT's outstanding performance is its training method, known as reinforcement learning from human feedback (RLHF) as shown in Figure 9. The principles of RLHF could be explored within EIM to enhance emotional modelling, fostering more empathetic and context-aware virtual interactions.

C. Improving Transparency and Explainability in EIM

In EIM, it is a necessity to grasp and understand learned emotional predictions. Current affective computing systems often lack clarity and user-friendliness, resulting in outputs that are difficult for people to make sense of. Improving transparency in AI models emerges as a potential solution. Model transparency involves clarifying "how the system reached a decision" [179]. While decision tree models are naturally transparent, "black box" models require extra tools to explain them [180]. Elaborate models may capture emotions in intricate ways, hindering straightforward human comprehension and complicating the validation of accuracy and fairness. Recently, explanation tools have emerged, providing users with insights into how systems make decisions [181], [182]. Upcoming research should focus on techniques to create understandable emotional representations in affective computing models. This might encompass visualisations, attribution methods, or embeddings that shed light on how models capture and use emotions.

D. Ethical and social considerations in developing applications for EIM

EIM holds the potential to reshape industries, human interactions, and address global challenges while promoting equal opportunities. However, it also presents complex ethical and social considerations that necessitate careful navigation [105], [183]. As affective computing applications in EIM gather and analyse user emotional data, it is crucial to obtain informed consent and address ethical concerns [184]. Users must be empowered with control over the collection and use of their emotional data, while simultaneously ensuring transparent communication of potential risks associated with emotional insights [109], [185]. Since affective computing techniques for EIM are still in their early stages of development, it is crucial to address the ethical challenges before rolling out applications that rely on these technologies [185].

The integration of real and virtual economies in EIM brings forth risks encompassing security, fraud, scams, and instability [186]. It is equally important to contemplate the impact of EIM on health and well-being, as prolonged immersion in virtual realms can give rise to addiction or isolation [187]. Moreover, the absence of moderation and the existence of illegal content only serves to amplify doubts regarding trustworthiness [188]. The decentralised and occasionally unregulated nature of metaverse activities creates opportunities for misconduct and exploitation [186], [188]. The fundamental challenges hindering the realisation of EIM's full potential stem from the absence of ethical and social considerations within the broader technology landscape [189]. As EIM involves capturing human emotional states through various sensors, the entrenched lack of ethical and socially conscious algorithm and product design presents a daunting challenge that demands immediate attention [189], [190].

Rosenberg's research [191] delves deep into the metaverse's potential risks, shedding light on its potential as a potent tool for manipulation and persuasion. The author uses the concept of feedback control and draws an intriguing analogy between the metaverse and engineering control systems as depicted in Figure 10. In this intriguing context, users effectively become subjects of control, exposing themselves to an environment where their behaviours and emotions can be actively monitored and influenced in real time. This scenario accentuates the urgent necessity to address the ethical implications of emotional manipulation within the EIM. The notable absence of explicit consent and transparency regarding data collection, coupled with the potential for unethical practices by malicious actors, raises valid concerns. It is imperative to handle these concerns with utmost care and sensitivity to ensure that users maintain full agency and control over their emotional data [157], [192].

Social norms could act to limit the development and implementation of affective computing applications in EIM [193]. For instance, most societies have strong norms around privacy, especially concerning sensitive personal data such as emotions [194]. Constantly monitoring users' emotions could also be perceived as odd, invasive, or 'creepy' by some, going against social expectations of normal human interaction and potentially limiting adoption [195], [196]. Certain contexts may also have clearer guidelines around the appropriate use of these technologies compared to others – for example, work, education, or healthcare versus leisure contexts [197]. Displays and interpretations of emotion also vary significantly between cultures, so EIM applications would need to be sensitive to cultural norms to avoid offence or misunderstanding [198].

While ethical challenges in affective computing have received attention, the specific ethical implications of applying affective computing in EIM remain an under-explored area [75], [199], [200]. A brief summary of challenges in developing EIM are also provided in Table 8. This entails respecting privacy, mitigating perceived creepiness, handling cultural differences sensitively, and avoiding direct emotional manipulation. As technology advances in this domain [201], [202], it is crucial to establish frameworks and guidelines for obtaining user consent, safeguarding data privacy, and addressing potential biases and discrimination arising from emotional insights in the metaverse.

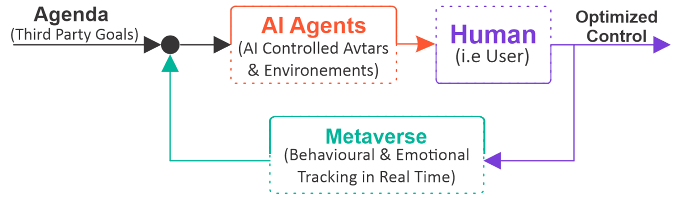

FIGURE 10: Rosenberg [191] scenario for metaverse mind control.

VI. Conclusions

In this paper, we explore how affective computing enhances the traditional metaverse, creating an emotionally intelligent metaverse (EIM) which redefines user interactions with virtual environments and AI entities for richer experiences, impacting domains like healthcare and entertainment. A brief summary of EIM features and potentials is presented as follows:

- EIM enables avatars and digital counterparts to adapt responses based on users' affective states for immersive experiences.
- EIM heightens empathy through interactions with avatars representing diverse backgrounds.
- Exploiting users' emotional states during sessions, EIM can amicably address the challenging issue of computing resource dimensioning among users.
- Multi-faceted benefits in healthcare include enhanced virtual doctor-patient interactions and personalised interventions.
- Following Norman's emotional design, game developers evoke user emotions with expressive illustrations in EIM, creating enduring, delightful experiences, and tailoring gameplay elements to emotional states.
- Brands refine products using emotional feedback for an enhanced shopping experience in EIM.
- Considering students' affective behaviour, EIM can aid educators in adapting techniques to improve the learning process.

Integrating emotional intelligence into the metaverse, embodied by EIM, is both compelling and intricate. As we explore affective computing within EIM, ethical considerations and responsible AI practices come to the forefront. Discerning emotions from digital expressions carries responsibility, necessitating transparency, fairness, and accountability in system design. The unique challenge of interpreting emotions in digital spaces underscores the importance of security and privacy, urging us to safeguard users' data, ensure consent, and protect against unauthorised access in EIM. The intricate algorithms underpinning affective computing require enhanced interpretability for user trust. Moving forward, the pursuit of robust, interpretable, and ethically responsible affec-

This article has been accepted for publication in IEEE Open Journal of the Computer Society. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/OJCS.2024.3389462

tive computing remains paramount within EIM. Techniques integrating data from various modalities promise improved precision. Our paper navigates this evolving landscape of affective interactions in virtual environments, emphasising the need of accurate and reliable algorithms, responsible AI practices, and an enhanced understanding of these dynamics within EIM.

REFERENCES

[1] S. Mystakidis, "Metaverse," *Encyclopedia*, vol. 2, no. 1, pp. 486–497, 2022.

[2] M. Jamshidi, A. Dehghaniyan Serej, A. Jamshidi, and O. Moztarzadeh, "The meta-metaverse: Ideation and future directions," *Future Internet*, vol. 15, no. 8, p. 252, 2023.

[3] J. P. Venugopal, A. A. V. Subramanian, and J. Peatchimuthu, "The realm of metaverse: A survey," *Computer Animation and Virtual Worlds*, p. e2150, 2023.

[4] M. Xu, W. C. Ng, W. Y. B. Lim, J. Kang, Z. Xiong, D. Niyato, Q. Yang, X. S. Shen, and C. Miao, "A full dive into realizing the edge-enabled metaverse: Visions, enabling technologies, and challenges," *IEEE Communications Surveys & Tutorials*, 2022.

[5] A. Kaushik, "Xr for social impact: A landscape review," *Games for Change*, 2020.

[6] C. Polona, M. T. André, and N. Maria, "Metaverse: Opportunities, risks and policy implications," 2022.

[7] R. W. Picard, *Affective computing*. MIT press, 2000.

[8] R. W. Coutinho and A. Boukerche, "When smart metaverse meets affective computing: Opportunities and design guidelines," *IEEE Communications Magazine*, 2023.

[9] K. Lee, "The interior experience of architecture: An emotional connection between space and the body," *Buildings*, vol. 12, no. 3, p. 326, 2022.

[10] L.-H. Lee, T. Braud, P. Zhou, L. Wang, D. Xu, Z. Lin, A. Kumar, C. Bermejo, and P. Hui, "All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda," *arXiv preprint arXiv:2110.05352*, 2021.

[11] J. Sun, W. Gan, H.-C. Chao, and P. S. Yu, "Metaverse: Survey, applications, security, and opportunities," *arXiv preprint arXiv:2210.07990*, 2022.

[12] Y. Wang, Z. Su, N. Zhang, R. Xing, D. Liu, T. H. Luan, and X. Shen, "A survey on metaverse: Fundamentals, security, and privacy," *IEEE Communications Surveys & Tutorials*, 2022.

[13] H. Ning, H. Wang, Y. Lin, W. Wang, S. Dhelim, F. Farha, J. Ding, and M. Daneshmand, "A survey on the metaverse: The state-of-the-art, technologies, applications, and challenges," *IEEE Internet of Things Journal*, 2023.

[14] W. Wen-xi, Z. Fang, W. Yue-liang, and N. Huan-sheng, "A survey of metaverse technology," *Chinese Journal of Engineering*, vol. 44, no. 4, pp. 744–756, 2022.

[15] H. Sami, A. Hammoud, M. Arafeh, M. Wazzeh, S. Arisdakessian, M. Chahoud, O. Wehbi, M. Ajaj, A. Mourad, H. Otrok *et al.*, "The metaverse: Survey, trends, novel pipeline ecosystem & future directions," *arXiv preprint arXiv:2304.09240*, 2023.

[16] G. K. Sriram, "A comprehensive survey on metaverse," *International Research Journal of Modernization in Engineering Technology*, vol. 4, no. 2, pp. 772–775, 2022.

[17] H. Ullah, S. Manickam, M. Obaidat, S. U. A. Laghari, and M. Uddin, "Exploring the potential of metaverse technology in healthcare: Applications, challenges, and future directions," *IEEE Access*, 2023.

[18] M. Uddin, S. Manickam, H. Ullah, M. Obaidat, and A. Dandoush, "Unveiling the metaverse: Exploring emerging trends, multifaceted perspectives, and future challenges," *IEEE Access*, 2023.

[19] R. Chengoden, N. Victor, T. Huynh-The, G. Yenduri, R. H. Jhaveri, M. Alazab, S. Bhattacharya, P. Hegde, P. K. R. Maddikunta, and T. R. Gadekallu, "Metaverse for healthcare: A survey on potential applications, challenges and future directions," *IEEE Access*, 2023.

[20] L. Petrigna and G. Musumeci, "The metaverse: A new challenge for the healthcare system: A scoping review," *Journal of functional morphology and kinesiology*, vol. 7, no. 3, p. 63, 2022.

[21] Y. Wang, W. Song, W. Tao, A. Liotta, D. Yang, X. Li, S. Gao, Y. Sun, W. Ge, W. Zhang *et al.*, "A systematic review on affective computing: Emotion models, databases, and recent advances," *Information Fusion*, vol. 83, pp. 19–52, 2022.

[22] S. Afzal, H. A. Khan, I. U. Khan, M. J. Piran, and J. W. Lee, "A comprehensive survey on affective computing; challenges, trends, applications, and future directions," *arXiv preprint arXiv:2305.07665*, 2023.

[23] J. Huggett, "Virtually real or really virtual: Towards a heritage metaverse," *Studies in digital heritage*, vol. 4, no. 1, pp. 1–15, 2020.

[24] H. Duan, J. Li, S. Fan, Z. Lin, X. Wu, and W. Cai, "Metaverse for social good: A university campus prototype," in *Proceedings of the 29th ACM international conference on multimedia*, 2021, pp. 153–161.

[25] S. B. Daily, M. T. James, D. Cherry, J. J. Porter III, S. S. Darnell, J. Isaac, and T. Roy, "Affective computing: historical foundations, current applications, and future trends," *Emotions and affect in human factors and human-computer interaction*, pp. 213–231, 2017.

[26] B. Poetker, "A brief history of augmented reality (+ future trends & impact)," https://www.g2.com/articles/history-of-augmented-reality, 2023 (accessed September 03, 2023).

[27] V. R. Society, "History of virtual reality," https://www.vrs.org.uk/virtual-reality/history.html, 2023 (accessed September 03, 2023).

[28] N. Stephenson, *Snow crash: A novel*. Spectra, 2003.

[29] S. Kaddoura and F. Al Husseiny, "The rising trend of metaverse in education: challenges, opportunities, and ethical considerations," *PeerJ Computer Science*, vol. 9, p. e1252, 2023.

[30] D. C. Smith, A. E. Douglas *et al.*, *The biology of symbiosis.* Edward Arnold (Publishers) Ltd., 1987.

[31] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6g wireless communication systems: Applications, requirements, technologies, challenges, and research directions," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 957–975, 2020.

[32] S. Mihai, M. Yaqoob, D. V. Hung, W. Davis, P. Towakel, M. Raza, M. Karamanoglu, B. Barn, D. Shetve, R. V. Prasad *et al.*, "Digital twins: A survey on enabling technologies, challenges, trends and future prospects," *IEEE Communications Surveys & Tutorials*, 2022.

[33] A. Masaracchia, V. Sharma, B. Canberk, O. A. Dobre, and T. Q. Duong, "Digital twin for 6g: Taxonomy, research challenges, and the road ahead," *IEEE Open Journal of the Communications Society*, 2022.

[34] P. Bellavista, J. Berrocal, A. Corradi, S. K. Das, L. Foschini, and A. Zanni, "A survey on fog computing for the internet of things," *Pervasive and mobile computing*, vol. 52, pp. 71–99, 2019.

[35] B. Hu and H. Gharavi, "Artificial intelligence-assisted edge computing for wide area monitoring," *IEEE Open Journal of the Communications Society*, 2023.

[36] H. Zhang, S. Mao, D. Niyato, and Z. Han, "Location-dependent augmented reality services in wireless edge-enabled metaverse systems," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 171–183, 2023.

[37] D. Mourtzis, N. Panopoulos, J. Angelopoulos, B. Wang, and L. Wang, "Human centric platforms for personalized value creation in metaverse," *Journal of Manufacturing Systems*, vol. 65, pp. 653–659, 2022.

[38] L. Angelini, M. Mecella, H.-N. Liang, M. Caon, E. Mugellini, O. Abou Khaled, and D. Bernardini, "Towards an emotionally augmented metaverse: a framework for recording and analysing physiological data and user behaviour," in *13th augmented human international conference*, 2022, pp. 1–5.

[39] M. Zallio and P. J. Clarkson, "Designing the metaverse: A study on inclusion, diversity, equity, accessibility and safety for digital immersive environments," *Telematics and Informatics*, vol. 75, p. 101909, 2022.

[40] M. Zyda, "Building a human-intelligent metaverse," *Computer*, vol. 55, no. 9, pp. 120–128, 2022.

[41] F. Daneshfar and M. B. Jamshidi, "An octonion-based nonlinear echo state network for speech emotion recognition in metaverse," *Neural Networks*, vol. 163, pp. 108–121, 2023.

[42] C. Rojas, M. Corral, N. Poulsen, and P. Maes, "Project us: A wearable for enhancing empathy," in *Companion Publication of the 2020 ACM Designing Interactive Systems Conference*, 2020, pp. 139–144.

[43] A. Streep, "Crime and opportunity : How one VR startup is capturing the 360 degree reality of the world's most vulnerable people," July 2016. [Online]. Available: Accessedon28Aug2023https://www.wired.com/2016/07/ryot-darg-mooser-disaster-vr

[44] K. Loveys, M. Sagar, M. Antoni, and E. Broadbent, "The impact of virtual humans on psychosomatic medicine," *Psychosomatic Medicine*, pp. 10–1097, 2023.

[45] B. K. Wiederhold, "Metaverse games: Game changer for healthcare?" pp. 267–269, 2022.

[46] D. Gromala, X. Tong, A. Choo, M. Karamnejad, and C. D. Shaw, "The virtual meditative walk: virtual reality therapy for chronic pain management," in *Proceedings of the 33rd Annual ACM conference on human factors in computing systems*, 2015, pp. 521–524.

[47] C. Lucifora, L. Angelini, Q. Meteier, C. M. Vicario, O. A. Khaled, E. Mugellini, and G. M. Grasso, "Cyber-therapy: the use of artificial intelligence in psychological practice," in *Intelligent Human Systems Integration 2021: Proceedings of the 4th International Conference on Intelligent Human Systems Integration (IHSI 2021): Integrating People and Intelligent Systems, February 22-24, 2021, Palermo, Italy*. Springer, 2021, pp. 127–132.

[48] M. H. Browning, K. J. Mimnaugh, C. J. Van Riper, H. K. Laurent, and S. M. LaValle, "Can simulated nature support mental health? comparing short, single-doses of 360-degree nature videos in virtual reality with the outdoors," *Frontiers in psychology*, vol. 10, p. 2667, 2020.

[49] T. Nukarinen, H. O. Istance, J. Rantala, J. Mäkelä, K. Korpela, K. Ronkainen, V. Surakka, and R. Raisamo, "Physiological and psychological restoration in matched real and virtual natural environments," in *Extended abstracts of the 2020 CHI conference on human factors in computing systems*, 2020, pp. 1–8.

[50] N. Dozio, F. Marcolin, G. W. Scurati, L. Ulrich, F. Nonis, E. Vezzetti, G. Marsocci, A. La Rosa, and F. Ferrise, "A design methodology for affective virtual reality," *International Journal of Human-Computer Studies*, vol. 162, p. 102791, 2022.

[51] C. Kerdvibulvech, "A digital human emotion modeling application using metaverse technology in the post-covid-19 era," in *International Conference on Human-Computer Interaction*. Springer, 2023, pp. 480–489.

[52] I. Kiprijanovska, B. Jakimovski, P. Krstevski, S. Stankoski, I. Mavridou, C. Nduka, H. Gjoreski, and M. Gjoreski, "Monitoring valence with facial emg sensors using machine learning," in *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*, 2022, pp. 178–182.

[53] G. Lăzăroiu, G. H. Popescu, I. Mohîrţă, R. Balica, and G. Cojocariu, "Does the capacity of identifying the student emotion expand the potentialities for e-learning?" *eLearning & Software for Education*, vol. 1, 2018.

[54] D. M. Barry, N. Ogawa, A. Dharmawansa, H. Kanematsu, Y. Fukumura, T. Shirai, K. Yajima, and T. Kobayashi, "Evaluation for students' learning manner using eye blinking system in metaverse," *Procedia computer science*, vol. 60, pp. 1195–1204, 2015.

[55] Y. Hwang, "When makers meet the metaverse: Effects of creating nft metaverse exhibition in maker education," *Computers & Education*, vol. 194, p. 104693, 2023.

[56] Z. Han, Y. Tu, and C. Huang, "A framework for constructing a technology-enhanced education metaverse: Learner engagement with human–machine collaboration," *IEEE Transactions on Learning Technologies*, 2023.

[57] P. Kral, K. Janoskova, and A.-M. Potcovaru, "Digital consumer engagement on blockchain-based metaverse platforms: extended reality technologies, spatial analytics, and immersive multisensory virtual spaces," *Linguistic and Philosophical Investigations*, vol. 21, pp. 252–267, 2022.

[58] S. F. S. Alwi and P. J. Kitchen, "Projecting corporate brand image and behavioral response in business schools: Cognitive or affective brand attributes?" *Journal of Business research*, vol. 67, no. 11, pp. 2324–2336, 2014.

[59] G. Davies, "In support of personality as a measure of reputation: A rejoinder to clardy's 'organizational reputation: Issues in conceptualization and measurement'(2012)," *Corporate Reputation Review*, vol. 16, pp. 168–173, 2013.

[60] Y. Sun, Y. Xu, C. Cheng, Y. Li, C. H. Lee, and A. Asadipour, "Travel with wander in the metaverse: An ai chatbot to visit the future earth," in *2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2022, pp. 1–6.

[61] T. Hennig-Thurau, D. N. Aliman, A. M. Herting, G. P. Cziehso, M. Linder, and R. V. Kübler, "Social interactions in the metaverse: Framework, initial evidence, and research roadmap," *Journal of the Academy of Marketing Science*, vol. 51, no. 4, pp. 889–913, 2023.

[62] "Communicate through expression - VIVE focus 3 facial tracker." [Online]. Available: Accessedon23Mar2024https://business.vive.com/us/product/vive-focus-3-facial-tracker

[63] "Unity asset store - SALSA lip sync suite." [Online]. Available: Accessedon23Mar2024https://assetstore.unity.com/packages/tools/animation/salsa-lipsync-suite-148442

[64] P. Lewicki, "Microsoft hololens 2; enterprise mixed reality glasses," 13 Jan 2020. [Online]. Available: Accessedon24Mar2024https://www.afternow.io/microsoft-hololens-2-enterprise-mixed-reality-glasses/

[65] "Meta Quest 3." [Online]. Available: Accessedon24Mar2024https://www.meta.com/quest/quest-3

[66] "Introducing VIVE tracker 3.0 - go beyond controllers." [Online]. Available: Accessedon23Mar2024https://www.vive.com/us/accessory/tracker3/

[67] "Unlock a new dimension of movement and freedom with the first AI-powered self-tracker from HTC VIVE." [Online]. Available: Accessedon23Mar2024https://www.vive.com/us/accessory/vive-ultimate-tracker/

[68] "VIVE wrist tracker." [Online]. Available: Accessedon23Mar2024https://www.vive.com/us/accessory/vive-wrist-tracker/

[69] "Unity asset store - VRIK module." [Online]. Available: Accessedon23Mar2024https://assetstore.unity.com/packages/tools/animation/final-ik-14290

[70] "Professional kit." [Online]. Available: Accessedon23Mar2024https://www.pluxbiosignals.com/products/professional-kit

[71] J. Wang, J. Suo, Z. Song, W. J. Li, and Z. Wang, "Nanomaterial-based flexible sensors for metaverse and virtual reality applications," *International Journal of Extreme Manufacturing*, vol. 5, no. 3, p. 032013, 2023.

[72] "An insightful add-on - VIVE focus 3 eye tracker." [Online]. Available: Accessedon23Mar2024https://business.vive.com/us/product/vive-focus-3-eye-tracker

[73] S. Li and W. Deng, "Deep facial expression recognition: A survey," *IEEE transactions on affective computing*, vol. 13, no. 3, pp. 1195–1215, 2020.

[74] F. Noroozi, C. A. Corneanu, D. Kamińska, T. Sapiński, S. Escalera, and G. Anbarjafari, "Survey on emotional body gesture recognition," *IEEE transactions on affective computing*, vol. 12, no. 2, pp. 505–523, 2018.

[75] S. Latif, R. Rana, S. Khalifa, R. Jurdak, J. Qadir, and B. W. Schuller, "Survey of deep representation learning for speech emotion recognition," *IEEE Transactions on Affective Computing*, 2021.

[76] A. H. Miller, W. Feng, A. Fisch, J. Lu, D. Batra, A. Bordes, D. Parikh, and J. Weston, "Parlai: A dialog research software platform," *arXiv preprint arXiv:1705.06476*, 2017.

[77] J. A. Mendez, A. Geramifard, M. Ghavamzadeh, and B. Liu, "Reinforcement learning of multi-domain dialog policies via action embeddings," *arXiv preprint arXiv:2207.00468*, 2022.

[78] A. Wang, Z. Gao, L. H. Lee, T. Braud, and P. Hui, "Decentralized, not dehumanized in the metaverse: Bringing utility to nfts through multimodal interaction," in *Proceedings of the 2022 International Conference on Multimodal Interaction*, 2022, pp. 662–667.

[79] L. Zhou, H. Palangi, L. Zhang, H. Hu, J. Corso, and J. Gao, "Unified vision-language pre-training for image captioning and vqa," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 07, 2020, pp. 13 041–13 049.

[80] M. Hessel, H. Soyer, L. Espeholt, W. Czarnecki, S. Schmitt, and H. Van Hasselt, "Multi-task deep reinforcement learning with popart," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 3796–3803.

[81] A. Das, S. Datta, G. Gkioxari, S. Lee, D. Parikh, and D. Batra, "Embodied question answering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1–10.

[82] S. B. Far, S. M. H. Bamakan, Q. Qu, and Q. Jiang, "A review of non-fungible tokens applications in the real-world and metaverse," *Procedia Computer Science*, vol. 214, pp. 755–762, 2022.

[83] S. Zhang, E. Dinan, J. Urbanek, A. Szlam, D. Kiela, and J. Weston, "Personalizing dialogue agents: I have a dog, do you have pets too?" *arXiv preprint arXiv:1801.07243*, 2018.

[84] E. Brown and P. Cairns, "A grounded investigation of game immersion," in *CHI'04 extended abstracts on Human factors in computing systems*, 2004, pp. 1297–1300.

[85] E. Han, M. R. Miller, N. Ram, K. L. Nowak, and J. N. Bailenson, "Understanding group behavior in virtual reality: A large-scale, longitudinal study in the metaverse," in *72nd Annual International Communication Association Conference, Paris, France*, 2022.

[86] A. Felnhofer, O. D. Kothgassner, M. Schmidt, A.-K. Heinzle, L. Beutl, H. Hlavacs, and I. Kryspin-Exner, "Is virtual reality emotionally arousing? investigating five emotion inducing virtual park scenarios," *International journal of human-computer studies*, vol. 82, pp. 48–56, 2015.

[87] K. Kuru, "Metaomnicity: Towards immersive urban metaverse cyberspaces using smart city digital twins," *IEEE Access*, 2023.

[88] C. D. Cogburn, "Experiencing racism in vr," 2017. [Online]. Available: Accessedon01Sep2023https://www.ted.com/talks/courtney_cogburn_experiencing_racism_in_vr_courtney_d_cogburn_phd_tedxrva

[89] F. Herrera, J. Bailenson, E. Weisz, E. Ogle, and J. Zaki, "Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking," *PloS one*, vol. 13, no. 10, p. e0204494, 2018.

[90] J. Collange and J. Guegan, "Using virtual reality to induce gratitude through virtual social interaction," *Computers in Human Behavior*, vol. 113, p. 106473, 2020.

[91] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, X. S. Shen, and D. I. Kim, "Exploring attention-aware network resource allocation for customized metaverse services," *IEEE Network*, 2022.

[92] N. H. Chu, D. T. Hoang, D. N. Nguyen, K. T. Phan, E. Dutkiewicz, D. Niyato, and T. Shu, "Metaslicing: A novel resource allocation framework for metaverse," *IEEE Transactions on Mobile Computing*, 2023.

[93] X. Zhou, C. Liu, and J. Zhao, "Resource allocation of federated learning for the metaverse with mobile augmented reality," *arXiv preprint arXiv:2211.08705*, 2022.

[94] T. J. Chua, W. Yu, and J. Zhao, "Resource allocation for mobile metaverse with the internet of vehicles over 6g wireless communications: A deep reinforcement learning approach," in *2022 IEEE 8th World Forum on Internet of Things (WF-IoT)*, 2022, pp. 1–7.

[95] C. Pensieri and M. Pennacchini, "Virtual reality in medicine," in *Handbook on 3D3C Platforms: Applications and Tools for Three Dimensional Systems for Community, Creation and Commerce*. Springer, 2016, pp. 353–401.

[96] S. G. Alonso, G. Marques, I. Barrachina, B. Garcia-Zapirain, J. Arambarri, J. C. Salvador, and I. de la Torre Díez, "Telemedicine and e-health research solutions in literature for combatting covid-19: a systematic review," *Health and technology*, vol. 11, pp. 257–266, 2021.

[97] B. Marr, "The amazing possibilities of healthcare in the metaverse," *Forbes*, 2022. [Online]. Available: Accessedon29Aug2023https://www.forbes.com/sites/bernardmarr/2022/02/23/the-amazing-possibilities-of-healthcare-in-the-metaverse/?sh=492c99c19e5c

[98] J. Thomason, "Metahealth-how will the metaverse change health care?" *Journal of Metaverse*, vol. 1, no. 1, pp. 13–16, 2021.

[99] A. Musamih, I. Yaqoob, K. Salah, R. Jayaraman, Y. Al-Hammadi, M. Omar, and S. Ellahham, "Metaverse in healthcare: Applications, challenges, and future directions," *IEEE Consumer Electronics Magazine*, 2022.

[100] D. Yang, J. Zhou, R. Chen, Y. Song, Z. Song, X. Zhang, Q. Wang, K. Wang, C. Zhou, J. Sun *et al.*, "Expert consensus on the metaverse in medicine," *Clinical eHealth*, vol. 5, pp. 1–9, 2022.

[101] H. Werner, G. Ribeiro, V. Arcoverde, J. Lopes, and L. Velho, "The use of metaverse in fetal medicine and gynecology," *European Journal of Radiology*, vol. 150, 2022.

[102] A. Ionescu, T. Van Daele, A. Rizzo, C. Blair, and P. Best, "360 videos for immersive mental health interventions: A systematic review," *Journal of Technology in Behavioral Science*, vol. 6, no. 4, pp. 631–651, 2021.

[103] G. Chen, "Designing and evaluating a social vr clinic for knee replacement surgery," 2019.

[104] "Metaverse: The next frontier for health 4.0," Apr 2022. [Online]. Available: Accessedon29Aug2023https://www.netscribes.com/metaverse-in-healthcare/

[105] O. Moztarzadeh, M. Jamshidi, S. Sargolzaei, F. Keikhaee, A. Jamshidi, S. Shadroo, and L. Hauer, "Metaverse and medical diagnosis: A blockchain-based digital twinning approach based on mobilenetv2 algorithm for cervical vertebral maturation," *Diagnostics*, vol. 13, no. 8, p. 1485, 2023.

[106] M. B. Jamshidi, S. Sargolzaei, S. Foorginezhad, and O. Moztarzadeh, "Metaverse and microorganism digital twins: A deep transfer learning approach," *Applied Soft Computing*, vol. 147, p. 110798, 2023.

[107] "Excitement is real, it's not digital." [Online]. Available: Accessedon06Sep2023https://www.metaverseme.io/blog/booklet-excitement-is-real

[108] B. Ryskeldiev, Y. Ochiai, M. Cohen, and J. Herder, "Distributed metaverse: creating decentralized blockchain-based model for peer-to-peer sharing of virtual spaces for mixed reality applications," in *Proceedings of the 9th augmented human international conference*, 2018, pp. 1–3.

[109] J. Baker-Brunnbauer, "Ethical challenges for the metaverse development," 2022.

[110] N. Ilinykh, S. Zarrieß, and D. Schlangen, "Meetup! a corpus of joint activity dialogues in a visual environment," *arXiv preprint arXiv:1907.05084*, 2019.

[111] B. Chen, S. Song, H. Lipson, and C. Vondrick, "Visual hide and seek," in *Artificial Life Conference Proceedings 32*. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ..., 2020, pp. 645–655.

[112] I.-C. Stanica, F. Moldoveanu, G.-P. Portelli, M.-I. Dascalu, A. Moldoveanu, and M. G. Ristea, "Flexible virtual reality system for neurorehabilitation and quality of life improvement," *Sensors*, vol. 20, no. 21, p. 6045, 2020.

[113] D. A. Norman, *Emotional design: Why we love (or hate) everyday things*. Civitas Books, 2004.

[114] A. Altonen, "Emotional ux designs for creating desirable products," 2020. [Online]. Available: Accessedon15Sep2023https://www.qt.io/blog/emotional-design-for-creating-desirable-products#:

[115] B. Dalvi, "Emotions in video game ux," 2021. [Online]. Available: Accessedon15Sep2023https://uxdesign.cc/emotions-in-video-game-ux-d0aa7027be3

[116] J. Díaz, C. Saldaña, and C. Avila, "Virtual world as a resource for hybrid education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 15, pp. 94–109, 2020.

[117] K. Hirsh-Pasek, J. Zosh, H. Hadani, R. Golinkoff, K. Clark, C. Donohue, and E. Wartella, "A whole new world: Education meets the metaverse. the brookings institution," 2022.

[118] A. Almarzouqi, A. Aburayya, and S. A. Salloum, "Prediction of user's intention to use metaverse system in medical education: A hybrid sem-ml learning approach," *IEEE access*, vol. 10, pp. 43 421–43 434, 2022.

[119] G. Batnasan, M. Gochoo, M.-E. Otgonbold, F. Alnajjar, and T. K. Shih, "Arsl21l: Arabic sign language letter dataset benchmarking and an educational avatar for metaverse applications," in *2022 ieee global engineering education conference (educon)*. IEEE, 2022, pp. 1814–1821.

[120] J. Kemp and D. Livingstone, "Putting a second life®"metaverse" skin on learning management systems (2007)."

[121] C. Collins, "Looking to the future: Higher education in the metaverse," *Educause Review*, vol. 43, no. 5, pp. 50–52, 2008.

[122] N. Khan, K. Muhammad, T. Hussain, M. Nasir, M. Munsif, A. S. Imran, and M. Sajjad, "An adaptive game-based learning strategy for children road safety education and practice in virtual space," *Sensors*, vol. 21, no. 11, p. 3661, 2021.

[123] J. Panksepp and M. Solms, "What is neuropsychoanalysis? clinically relevant studies of the minded brain," *Trends in cognitive sciences*, vol. 16, no. 1, pp. 6–8, 2012.

[124] J. Kevins, "Metaverse as a new emerging technology: An interrogation of opportunities and legal issues: Some introspection," *Available at SSRN 4050898*, 2022.

[125] G. Sachs, "Understanding the metaverse and web 3.0," 2022. [Online]. Available: Accessedon01Sep2023https://www.goldmansachs.com/intelligence/podcasts/episodes/01-11-2022-eric-sheridan.html

[126] A. Joy, Y. Zhu, C. Peña, and M. Brouard, "Digital future of luxury brands: Metaverse, digital fashion, and non-fungible tokens," *Strategic change*, vol. 31, no. 3, pp. 337–343, 2022.

[127] Roblox, "The gucci garden experience lands on roblox," 17 May 2021. [Online]. Available: Accessedon01Sep2023https://blog.roblox.com/2021/05/gucci-garden-experience/
[128] W. Seok, "Analysis of metaverse business model and ecosystem," *Electronics and Telecommunications Trends*, vol. 36, no. 4, pp. 81–91, 2021.
[129] H. Jeong, Y. Yi, and D. Kim, "An innovative e-commerce platform incorporating metaverse to live commerce," *International Journal of Innovative Computing, Information and Control*, vol. 18, no. 1, pp. 221–229, 2022.
[130] E. Faliagka, E. Christopoulou, D. Ringas, T. Politi, N. Kostis, D. Leonardos, C. Tranoris, C. P. Antonopoulos, S. Denazis, and N. Voros, "Trends in digital twin framework architectures for smart cities: A case study in smart mobility," *Sensors*, vol. 24, no. 5, p. 1665, 2024.
[131] S.-m. Wang and L. H. Vu, "The integration of digital twin and serious game framework for new normal virtual urban exploration and social interaction," *Journal of Urban Management*, vol. 12, no. 2, pp. 168–181, 2023.
[132] D. Mehta, M. F. H. Siddiqui, and A. Y. Javaid, "Facial emotion recognition: A survey and real-world user experiences in mixed reality," *Sensors*, vol. 18, no. 2, p. 416, 2018.
[133] Z.-Y. Huang, C.-C. Chiang, J.-H. Chen, Y.-C. Chen, H.-L. Chung, Y.-P. Cai, and H.-C. Hsu, "A study on computer vision for facial emotion recognition," *Scientific Reports*, vol. 13, no. 1, p. 8425, 2023.
[134] H. Park, J. Pei, M. Shi, Q. Xu, and J. Fan, "Designing wearable computing devices for improved comfort and user acceptance," *Ergonomics*, vol. 62, no. 11, pp. 1474–1484, 2019.
[135] S. Pal, S. Mukhopadhyay, and N. Suryadevara, "Development and progress in sensors and technologies for human emotion recognition," *Sensors*, vol. 21, no. 16, p. 5554, 2021.
[136] M. U. Hasan and I. Negulescu, "Wearable technology for baby monitoring: a review," *J. Text. Eng. Fash. Technol*, vol. 6, no. 112.10, p. 15406, 2020.
[137] S. Latif, J. Qadir, and M. Bilal, "Unsupervised adversarial domain adaptation for cross-lingual speech emotion recognition," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2019, pp. 732–737.
[138] S. Latif, R. Rana, S. Khalifa, R. Jurdak, and B. W. Schuller, "Self supervised adversarial domain adaptation for cross-corpus and cross-language speech emotion recognition," *IEEE Transactions on Affective Computing*, 2022.
[139] N. Ahmed, Z. Al Aghbari, and S. Girija, "A systematic survey on multimodal emotion recognition using learning algorithms," *Intelligent Systems with Applications*, vol. 17, p. 200171, 2023.
[140] S. Latif, R. Rana, and J. Qadir, "Adversarial machine learning and speech emotion recognition: Utilizing generative adversarial networks for robustness," *arXiv preprint arXiv:1811.11402*, 2018.
[141] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," *arXiv preprint arXiv:1412.6572*, 2014.
[142] N. Papernot, P. McDaniel, S. Jha, M. Fredrikson, Z. B. Celik, and A. Swami, "The limitations of deep learning in adversarial settings," in *2016 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, 2016, pp. 372–387.
[143] S.-M. Moosavi-Dezfooli, A. Fawzi, and P. Frossard, "Deepfool: a simple and accurate method to fool deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2574–2582.
[144] P. Liu, Y. Lin, Z. Meng, L. Lu, W. Deng, J. T. Zhou, and Y. Yang, "Point adversarial self-mining: A simple method for facial expression recognition," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12 649–12 660, 2021.
[145] P. Li, Z. Zhang, A. S. Al-Sumaiti, N. Werghi, and C. Y. Yeun, "A robust adversary detection-deactivation method for metaverse-oriented collaborative deep learning," *IEEE Sensors Journal*, 2023.
[146] X. Yuan, P. He, Q. Zhu, and X. Li, "Adversarial examples: Attacks and defenses for deep learning," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 9, pp. 2805–2824, 2019.
[147] N. Akhtar and A. Mian, "Threat of adversarial attacks on deep learning in computer vision: A survey," *IEEE Access*, vol. 6, pp. 14 410–14 430, 2018.
[148] J. Hu, J. Wen, and M. Fang, "A survey on adversarial attack and defense of deep learning models for medical image recognition," *Metaverse*, vol. 4, no. 1, p. 17, 2023.
[149] Z. Yi, Y. Qian, M. Chen, S. A. Alqahtani, and M. S. Hossain, "Defending edge computing based metaverse ai against adversarial attacks," *Ad Hoc Networks*, vol. 150, p. 103263, 2023.
[150] P. Porambage, G. Gür, D. P. M. Osorio, M. Liyanage, A. Gurtov, and M. Ylianttila, "The roadmap to 6g security and privacy," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1094–1122, 2021.
[151] C. B. Fernandez and P. Hui, "Life, the metaverse and everything: An overview of privacy, ethics, and governance in metaverse," in *2022 IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW)*. IEEE, 2022, pp. 272–277.
[152] R. Benjamins, Y. Rubio Viñuela, and C. Alonso, "Social and ethical challenges of the metaverse: Opening the debate," *AI and Ethics*, pp. 1–9, 2023.
[153] Y. Huang, Y. J. Li, and Z. Cai, "Security and privacy in metaverse: A comprehensive survey," *Big Data Mining and Analytics*, vol. 6, no. 2, pp. 234–247, 2023.
[154] R. Di Pietro and S. Cresci, "Metaverse: security and privacy issues," in *2021 Third IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)*. IEEE, 2021, pp. 281–288.
[155] B. Falchuk, S. Loeb, and R. Neff, "The social metaverse: Battle for privacy," *IEEE Technology and Society Magazine*, vol. 37, no. 2, pp. 52–61, 2018.
[156] R. Zhao, Y. Zhang, Y. Zhu, R. Lan, and Z. Hua, "Metaverse: Security and privacy concerns," *arXiv preprint arXiv:2203.03854*, 2022.
[157] J. M. Silvennoinen and P. Saariluoma, "Affective mimetics, emotional information space, and metaverse," *Journal of Design Thinking*, vol. 2, no. 2, pp. 293–304, 2021.
[158] F. Lu and B. Liu, "Affective digital twins for digital human: Bridging the gap in human-machine affective interaction," *arXiv preprint arXiv:2308.10207*, 2023.
[159] N. Mishra, M. Ramanathan, G. Tulsulkar, and N. M. Thalmann, "Uncanny valley for interactive social agents: An experimental study," *Virtual Reality & Intelligent Hardware*, vol. 4, no. 5, pp. 393–405, 2022.
[160] H. Brenton, M. Gillies, D. Ballin, and D. Chatting, "The uncanny valley: does it exist," in *Proceedings of conference of human computer interaction, workshop on human animated character interaction*. Citeseer, 2005.
[161] M. Sharma and K. Vemuri, "Accepting human-like avatars in social and professional roles," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 11, no. 3, pp. 1–19, 2022.
[162] Q. Cao, H. Yu, P. Charisse, S. Qiao, and B. Stevens, "Is high-fidelity important for human-like virtual avatars in human computer interactions?" *International Journal of Network Dynamics and Intelligence*, pp. 15–23, 2023.
[163] X. Cheng, S. Zhang, and J. Mou, "Hello, avatar: The impact of avatars on individuals' emotions and behaviors from three congruity perspectives," 2023.
[164] N. Gurung, J. B. Grant, and D. Hearth, "The uncanny effect of speech: The impact of appearance and speaking on impression formation in human–robot interactions," *International Journal of Social Robotics*, pp. 1–16, 2023.
[165] S.-M. Park and Y.-G. Kim, "A metaverse: Taxonomy, components, applications, and open challenges," *IEEE access*, vol. 10, pp. 4209–4251, 2022.
[166] L.-H. Lee, C. Bermejo Fernandez, A. Alhilal, T. Braud, S. Hosio, E. H. A. de Haas, and P. Hui, "Beyond the blue sky of multimodal interaction: A centennial vision of interplanetary virtual spaces in turn-based metaverse," in *Proceedings of the 2022 International Conference on Multimodal Interaction*, 2022, pp. 648–652.
[167] T. Huynh-The, Q.-V. Pham, X.-Q. Pham, T. T. Nguyen, Z. Han, and D.-S. Kim, "Artificial intelligence for the metaverse: A survey," *Engineering Applications of Artificial Intelligence*, vol. 117, p. 105581, 2023.
[168] S. A. M. Zaidi, S. Latif, and J. Qadi, "Cross-language speech emotion recognition using multimodal dual attention transformers," *arXiv preprint arXiv:2306.13804*, 2023.
[169] M. H. Hunt and S. Angelopoulos, "A federated network architecture perspective on the future of the metaverse," 2023.

[170] T. H. Falk, L. B. Le, and R. Morandotti, "The internet of senses: A position paper on the challenges and opportunities of multisensory immersive experiences for the metaverse," in *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*. IEEE, 2022, pp. 139–144.

[171] F. Tang, X. Chen, M. Zhao, and N. Kato, "The roadmap of communication and networking in 6g for the metaverse," *IEEE Wireless Communications*, 2022.

[172] S. Latif, R. Rana, S. Khalifa, R. Jurdak, and B. W. Schuller, "Deep architecture enhancing robustness to noise, adversarial attacks, and cross-corpus setting for speech emotion recognition," 2020.

[173] A. Qayyum, M. Usama, J. Qadir, and A. Al-Fuqaha, "Securing connected & autonomous vehicles: Challenges posed by adversarial machine learning and the way forward," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 998–1026, 2020.

[174] OpenAI, "Gpt-4 technical report," *https://arxiv.org/pdf/2303.08774.pdf*, 2023.

[175] L. Barrault, Y.-A. Chung, M. C. Meglioli, D. Dale, N. Dong, P.-A. Duquenne, H. Elsahar, H. Gong, K. Heffernan, J. Hoffman *et al.*, "Seamlessm4t-massively multilingual & multimodal machine translation," *arXiv preprint arXiv:2308.11596*, 2023.

[176] S. Latif, M. Shoukat, F. Shamshad, M. Usama, H. Cuayáhuitl, and B. W. Schuller, "Sparks of large audio models: A survey and outlook," *arXiv preprint arXiv:2308.12792*, 2023.

[177] S. Latif, H. Cuayáhuitl, F. Pervez, F. Shamshad, H. S. Ali, and E. Cambria, "A survey on deep reinforcement learning for audio-based applications," *Artificial Intelligence Review*, vol. 56, no. 3, pp. 2193–2240, 2023.

[178] E. Lakomkin, M. A. Zamani, C. Weber, S. Magg, and S. Wermter, "Emorl: continuous acoustic emotion classification using deep reinforcement learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–6.

[179] A. Rosenfeld and A. Richardson, "Explainability in human–agent systems," *Autonomous Agents and Multi-Agent Systems*, vol. 33, pp. 673–705, 2019.

[180] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM computing surveys (CSUR)*, vol. 51, no. 5, pp. 1–42, 2018.

[181] A. Rosenfeld, "Better metrics for evaluating explainable artificial intelligence," in *Proceedings of the 20th international conference on autonomous agents and multiagent systems*, 2021, pp. 45–50.

[182] J. V. Jeyakumar, J. Noor, Y.-H. Cheng, L. Garcia, and M. Srivastava, "How can i explain this to you? an empirical study of deep neural network explanation methods," *Advances in Neural Information Processing Systems*, vol. 33, pp. 4211–4222, 2020.

[183] J. R. Jim, M. T. Hosain, M. Mridha, M. M. Kabir, and J. Shin, "Towards trustworthy metaverse: Advancements and challenges," *IEEE Access*, 2023.

[184] K. Kostick-Quenet and V. Rahimzadeh, "Ethical hazards of health data governance in the metaverse," *Nature machine intelligence*, pp. 1–3, 2023.

[185] C. H. Smith, J. Molka-Danielsen, J. Rasool, and J.-B. Webb-Benjamin, "The world as an interface: exploring the ethical challenges of the emerging metaverse," 2023.

[186] A. N. Mitrushchenkova, "Personal identity in the metaverse: Challenges and risks," *Kutafin Law Review*, vol. 9, no. 4, pp. 793–817, 2023.

[187] H. J. Oh, J. Kim, J. J. Chang, N. Park, and S. Lee, "Social benefits of living in the metaverse: The relationships among social presence, supportive interaction, social self-efficacy, and feelings of loneliness," *Computers in Human Behavior*, vol. 139, p. 107498, 2023.

[188] D.-I. D. Han, Y. Bergs, and N. Moorhouse, "Virtual reality consumer experience escapes: preparing for the metaverse," *Virtual Reality*, vol. 26, no. 4, pp. 1443–1458, 2022.

[189] M. Okkerman, "Holding up a black mirror: exploring ethical issues of the metaverse from a user perspective," Master's thesis, University of Twente, 2023.

[190] K. Charamba, "Beyond the corporate responsibility to respect human rights in the dawn of a metaverse," *U. Miami Int'l & Comp. L. Rev.*, vol. 30, p. 110, 2022.

[191] L. Rosenberg, "Mind control: The metaverse may be the ultimate tool of persuasion," *VentureBeat, VentureBeat*, vol. 22, 2022.

[192] S. Latif, A. Qayyum, M. Usama, J. Qadir, A. Zwitter, and M. Shahzad, "Caveat emptor: the risks of using big data for human development," *Ieee technology and society magazine*, vol. 38, no. 3, pp. 82–90, 2019.

[193] R. Somarathna, T. Bednarz, and G. Mohammadi, "Virtual reality for emotion elicitation–a review," *IEEE Transactions on Affective Computing*, 2022.

[194] S. Lee, S. Lee, Y. Choi, J. Ben-Othman, L. Mokdad, K. Jun, and H. Kim, "Affective surveillance management in virtual emotion based smart complex infrastructure," *IEEE Communications Magazine*, 2023.

[195] M. Saker and J. Frith, "Contiguous identities: the virtual self in the supposed metaverse," *First Monday*, 2022.

[196] Y. K. Dwivedi, N. Kshetri, L. Hughes, N. P. Rana, A. M. Baabdullah, A. K. Kar, A. Koohang, S. Ribeiro-Navarrete, N. Belei, J. Balakrishnan *et al.*, "Exploring the darkverse: A multi-perspective analysis of the negative societal impacts of the metaverse," *Information Systems Frontiers*, pp. 1–44, 2023.

[197] S. Zuboff, "The age of surveillance capitalism," in *Social Theory Re-Wired*. Routledge, 2023, pp. 203–213.

[198] L. Evans, J. Frith, and M. Saker, *From microverse to metaverse: Modelling the future through today's virtual worlds*. Emerald Publishing Limited, 2022.

[199] S. Latif, H. S. Ali, M. Usama, R. Rana, B. Schuller, and J. Qadir, "Ai-based emotion recognition: Promise, peril, and prescriptions for prosocial path," *arXiv preprint arXiv:2211.07290*, 2022.

[200] S. Latif, "Deep representation learning for speech emotion recognition," Ph.D. dissertation, University of Southern Queensland, 2022.

[201] Z. Chen, W. Gan, J. Sun, J. Wu, and P. S. Yu, "Open metaverse: Issues, evolution, and future," *arXiv preprint arXiv:2304.13931*, 2023.

[202] Z. Chen, J. Wu, W. Gan, and Z. Qi, "Metaverse security and privacy: An overview," in *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022, pp. 2950–2959.