Journal of Robotics Research (JRR)
Volume 3, Issue 2, Date: 31 - December - 2026
ISSN: 2978-4344, DOI: 10.64820/AEPJRR.32.1.7.122026

A lightweight Gaussian-based framework for real-time facial expression enhancement in 3D avatar systems

Muhammad Shaheer [1, *], Safyan Jameel [1], Syed Farhad Shah Bukhari [1], Asma Rahman [1]
[1] Department of Computer Science and Artificial Intelligence, Wuhan Textile University, Wuhan, China
[2] Robotics & Automation Research Lab (RARL), Queens Building, Leicester, LE1 9BH, United Kingdom
Email: [1] malikshaheer127679@gmail.com, [2] xafayanjameel@gmail.com, [3] farhad.bukhari23@gmail.com, [4] asmarah01@gmail.com
*Corresponding Author

Abstract— **Facial expression modeling is a basic part of the contemporary 3D avatar systems as it plays an important role in emotional communication and engagement with the user and realism in the virtual environment. While there have been remarkable progress in neural rendering, Neural Radiance Field (NeRF) and Gaussian-based avatar representations have brought up impressive visual fidelity, the majority of the methods that have been developed are computationally demanding, involve large-scale training procedures, and require dedicated GPU hardware. This limitation makes their application difficult in the real world and resource-limited scenarios such as mobile devices, edge computing and CPU-based computing. To tackle these problems, this paper presents a novel method, named Expression-Focused Gaussian Enhancement (EFGE), which is lightweight and very efficient to enhance face expression without performing full 3D face reconstruction. The proposed approach is based on facial landmark tracking to localize expression critical regions like eyes, eyebrows, mouth and cheeks. A landmark displacement analysis is used to create dynamically adaptive Gaussians which are used to selectively enhance facial movements that are relevant to the viewer. The size and opaqueness of Gaussian components are controlled depending on the expression intensity so as to enhance the expressions perceptually meaningfully while keeping the computational complexity low. This framework is realized with MediaPipe face mesh and optimized to be executed real-time on CPU-based hardware. The experimental assessment shows the effectiveness of EFGE in enhancing the visibility of facial expression, emotional clarity, user perception, and with efficient computational performance. Comparative analysis shows that the proposed method provides a good balance between expressiveness and efficiency in comparison with the traditional landmark-based approach and the computationally expensive Gaussian avatar approach. The results indicate that EFGE can be a feasible and scalable approach to real-time expressive avatar rendering, especially in mobile, VR, AR and low-resource applications.**

Keywords*— *Expression-Focused Gaussian Enhancement (EFGE), 3D Avatars, Facial Expression Modeling, Gaussian Splatting, Real-Time Rendering, Human-Computer Interaction, Lightweight Systems, Virtual Reality

I. INTRODUCTION (*HEADING 1*)

Digital human avatars are integral parts of various virtual reality (VR), augmented reality (AR), metaverse platforms, tele-presence applications, games, and human-computer interactions [1]. With the shift towards more immersive digital experiences, much research is going into making more realistic and expressive avatars that can replicate human appearance, behavior, and emotion [2]. The original methods of avatar generation were based on mesh-based modeling and parametric face representations, which are useful in controlling facial animation but lacked realism and fine-grained expressive detail. While neural rendering and Neural Radiance Fields (NeRF) have greatly enhanced the realism of avatars in recent years, their intensive training and rendering demands make them hard to utilize in real-time applications. In recent years, therefore, the focus has been placed on Gaussian Splatting-based representations, which offer an appealing compromise between the quality of the resulting rendering and the computational cost. Recent works like ExAvatar, EVA, and other Gaussian avatar representations have shown that 3D Gaussian representations can create very realistic and expressive digital humans with faster rendering speed than traditional neural rendering methods [3]. Additionally, the recent progress of Gaussian avatar reconstruction and enhancement has further enhanced the fidelity of avatars, facial detail preservation, and animation quality, and Gaussian-based methods are a potential direction for the future of avatar systems [4].

Though there has been exceptional progress made by avatar generation frameworks based on random Gaussian sampling, there still remain a number of challenges in balancing the expressiveness of the avatar with computational efficiency [5]. The majority of the most modern avatar systems are dependent on dense Gaussian representations, neural optimization pipelines, a large amount of training data, and powerful GPU hardware. These methods achieve a high degree of realism in facial reconstruction and body motion, but they generally result in significant computational costs, which makes them difficult to use in various applications like mobile applications, embedded

Received: 2-8-2026
Revised: 18- 8 -2026
Published: 31-12-2026

systems, or real-time communication systems in low-resource settings. Furthermore, most current approaches concentrate on facial reconstruction as a whole rather than selectively improving expression-critical facial regions that are relevant to emotional communication [6]. This full reconstruction method not only makes the computation more complicated but also adds a small amount of non-negligible perceptual value for expression recognition. Thus, lightweight frameworks that retain expressive facial communication without the need for expensive neural rendering pipelines or dedicated hardware resources are still needed [7].

This study aims to create a light and computational efficient system to improve facial expression in 3D avatar system based on Gaussian representation. In this regard, the proposed Expression-Focused Gaussian Enhancement (EFGE) framework is designed to detect expression critical facial regions using the landmark-based facial analysis technique, and selectively enhance expression critical regions using Gaussian overlays defined analytically [8]. The idea of the framework is to optimize the perceived emotional expressiveness while keeping the computational cost low and removing the need of deep neural training procedures. Moreover, the proposed system will be able to work on CPU-based hardware, and with real-time performance, in order that it can be easily available to be used and can be scalable for practical avatar applications [9]. A further goal is to create a modular structure that can be integrated into the existing avatar rendering pipelines without significant changes to their architecture [10].

An exhaustive literature search shows that the expressiveness of avatars is still quite a distance from computational efficiency. Previous approaches to avatars based on Gaussian representations include ExAvatar [11], EVA [12], HyperGaussians [13], ScaffoldAvatar [14], and other high-fidelity reconstruction methods, which focus mainly on photorealism and full-avatar reconstruction using complex Gaussian representations and neural optimization [15]. These techniques have high visual quality, but rely on GPU acceleration, large-scale training practices, and dense Gaussian parameter optimization, making them difficult to use in resource-constrained environments and real-time applications. Moreover, recent investigations on the reconstruction of humans based on Gaussian showed that the majority of current systems focus on reconstruction rather than enhancement strategies for expressions [16]. In addition, landmark based models offer efficient, real-time facial tracking but do not have means to perceptually improve emotion detection beyond landmark detection [17]. So there is no lightweight, expression-centric enhancement framework which selectively enhances facial expressiveness with Gaussian representation, and can keep the real-time CPU-level performance. The proposed EFGE framework tackles this research gap by proposing an expression focused enhancement paradigm to achieve more emotional clarity by leveraging landmark-based facial motion analysis and adaptive Gaussian augmentation, without the need for costly neural reconstruction pipelines [18].

Realistic digital avatars have become a focus in recent years owing to the increasing number of applications for virtual reality, augmented reality, telepresence, games, online education, healthcare, and human-computer interaction. The early avatar generation techniques were based mostly on the use of mesh facial model and parametric models like 3D Morphable Models (3DMMs) [19] which offered good control over facial expressions but were limited in their ability to represent subtle emotional nuances and variations in appearance. With the advent of deep learning techniques and neural rendering methods, the realism of avatars has been greatly enhanced, allowing the reconstruction of high fidelity facial geometry and appearance from image and video data. Neural Radiance Fields (NeRF) [20] made the leap from view synthesis to 3D reconstruction and enabled 3D reconstruction of an avatar with dynamic facial expression. In spite of their visual fidelity, NeRF based approaches are time-consuming, memory intensive, and hard to train on GPUs for real-time applications [21].

In recent years, researchers have undertaken a search for a more efficient real-time 3D scene representation and rendering method, as neural rendering has proven unreliable in this regard [22]. Recently, Gaussian Splatting was proposed as an alternative to neural rendering that is more efficient for real-time representation and rendering of 3D scenes. Gaussian Splatting is a representation method that uses thousands of learnable Gaussian primitives, allowing for efficient rendering with high visual fidelity. Gaussian Splatting: A representation technique that involves thousands of learnable Gaussian primitives that can be rendered efficiently while maintaining high visual fidelity. The 3D Gaussian Splatting method by Kerbl et al. (2023) [23] showed that Gaussian representations could produce state-of-the-art results with high rendering quality but with much lower computational complexity. Several studies expanded the Gaussian form to human motion and facial animation systems. ExAvatar, for example, introduced a whole-body Gaussian avatar generator that can reconstruct human motion and facial expressions from monocular video sequences [24]. Likewise, the expressive virtual avatars based on multi-view video data were introduced by EVA, with more realistic facial animation and appearance [25].

Another interesting area of research is real-time facial landmark detection and expression analysis. Landmark-based methods were used to determine the facial points of interest related to emotional movements and to efficiently track facial expressions. The MediaPipe Face Mesh framework has developed a powerful and compact framework that can track 468 facial landmarks in real time with standard hardware [26]. Landmark-based systems have been shown to be successful in emotion recognition, facial animation, and avatar control applications in subsequent studies. Most expression detection methods, however, are not intended to enhance expression; there is still room for more research into the enhancement of emotional expressiveness without increasing the computation of any algorithm [27].

Surveying the current literature on human reconstruction with Gaussian Splatting, most studies are focused on enhancing reconstruction quality, fidelity, and geometric accuracy, with few studies explicitly targeting a particular region of the face, such as the eyes, nose, mouth, or the face as a whole [28]. The consequence, however, is that there are still very few studies related to the lightweight framework integrating landmark-based facial analysis with Gaussian-based enhancement methods for more expressive and clear expressions in real-time avatar systems. This restriction points out that novel approaches that focus perceptually relevant parts of the face while keeping the computation cost low is needed [29].

A. Theoretical Background

The proposed Expression-Focused Gaussian Enhancement (EFGE) framework is based upon three complementary theoretical foundations: Facial Expression Theory, Landmark-Based Facial Motion Analysis and Gaussian-Based Visual Enhancement. The first theoretical basis comes from a theory of facial expressions developed by psychologist Paul Ekman, who discovered that the human face makes distinct facial muscle movements when expressing emotions and that these muscle movements are linked to universal facial expressions – happiness, sadness, anger, surprise, fear, and disgust. This theory suggests that the eyebrows, eyes, cheeks and mouth are the major regions of the face used for the perception of emotion. Therefore, improving such expression-critical regions can have a great impact on digital avatar emotional realism.

The second foundation is based on landmark-based facial motion analysis. Facial landmarks are geometric markers of facial muscle motion and allow quantitative measures of the strength of facial expression [30]. In the EFGE framework, the facial landmarks are employed to detect areas with significant emotional motion. The movement and changing of the position of the landmarks play a role in the displacement and the variation, which are indicators of the intensity of the expression and are used to guide the enhancement process.

The third theoretical basis is provided by gaussian representation theory. Gaussian functions offer smooth spatial distributions, ensuring continuity and realism while capturing localized visual effects. In computer graphics, a Gaussian primitive is a type of primitive that is used to represent a complex visual structure in an efficient way. Building on this idea, the Gaussian Splatting paradigm aims to model 3D objects as sets of Gaussian distributions that are efficient to render quickly and accurately [31]. The EFGE principle is applied in the EFGE framework, where strong emotional activities are detected on facial areas and adaptive Gaussian overlays are computed for each of them. The proposed framework selectively boosts regions of the face that are critical for expressing the face, thus reducing the amount of computations but at the same time preserving the perception.

Table 1: Comparison of Related Approaches

Method Category	**Main Purpose**	**Training**	**Hardware Requirement**	**Expression-Specific Enhancement**
NeRF-based avatars	Neural rendering	Yes	High	Indirect
Gaussian avatars	High reconstruction	Usually Yes	GPU-oriented	Limited
Landmark-based systems	Animation control	No/limited	Low	No direct enhancement
Proposed EFGE	Selective expression enhancement	**Yes**	**CPU-capable**	**Yes**

Therefore, the theoretical model makes the assumption that the facial landmarks can be used to reliably localize facial regions expressing emotional significance and that Gaussian enhancement can enhance the visual expression of these facial regions without changing the underlying facial geometry. Conceptually, the integration of the facial expression theory, landmark-based motion analysis and the Gaussian enhancement is the foundation of the proposed lightweight framework for real-time facial expression enhancement in 3D avatar systems [32].

Existing Gaussian-based avatar approaches primarily focus on high-fidelity reconstruction and rendering, typically relying on dense Gaussian representations, optimization procedures, and GPU acceleration. Conversely, landmark-based facial analysis enables efficient real-time tracking but generally uses facial motion only for detection, recognition, or animation control rather than perceptual enhancement. Therefore, a gap remains between computationally expensive Gaussian avatar reconstruction and lightweight landmark-based facial tracking: namely, a framework that uses facial motion information to selectively enhance expression-critical regions without reconstructing the complete facial geometry or requiring model training.

The main contributions of this study are as follows:

1. A lightweight Expression-Focused Gaussian Enhancement (EFGE) framework is proposed for selectively enhancing expression-critical facial regions using landmark-guided analytical Gaussian overlays.
2. The framework introduces an adaptive mechanism in which facial landmark displacement determines the location, scale, and intensity of Gaussian enhancement.
3. Unlike Gaussian avatar reconstruction methods, EFGE does not require dense Gaussian optimization, neural network training, or dedicated GPU hardware.
4. The framework is evaluated in terms of expression enhancement, processing speed, CPU utilization, memory consumption, and component-wise contribution.

II. METHODOLOGY

A. Research Design

This research is based on Quantitative Experimental Research Design, a study that aims at designing and testing a lightweight Gaussian framework for real-time face enhancement in 3D avatar systems. The research is conducted in a design-and-development methodology, where an Expression-Focused Gaussian Enhancement (EFGE) framework is proposed, implemented and experimentally tested. It integrates facial landmark detection and adaptive Gaussian enhancement methods for heightened facial expressiveness while effectively conserving computation time. The proposed framework uses analytical Gaussian overlays guided by facial landmark motion, instead of the conventional approach using deep neural training and dense optimization for Gaussian avatar reconstruction. The experimental design facilitates systematic evaluation of the effectiveness of the framework in improving the quality of facial expressions, ease of computations, and real-time performance. The whole research process is divided into four steps: facial landmark extraction and expression intensity estimation, Gaussian enhancement generation, performance evaluation.

B. Sample

For this study, the sample is facial image and video sequences, which are provided for testing and validation of the proposed framework. The system was tested with facial expressions depicting various emotional states such as happiness, sadness, surprise, anger and neutrality. It was also made sure that the study used publicly available facial datasets and real-time webcam inputs to guarantee differences in facial shape, lighting and expression. There were enough facial landmarks per sample to accurately analyze the facial expressions using MediaPipe Face Mesh framework. The dataset was chosen according to its appropriateness for real time facial tracking and avatar expression enhancement applications. Moreover, several test cases were created to test the framework in various expression situations and computational environments.

The experimental evaluation included a total of 100 facial images and 25 video sequences collected from 25 facial subjects. The samples represented different facial expressions, including happiness, sadness, surprise, anger, and neutral expressions. The image and video samples were selected to provide variation in facial movement and expression intensity for evaluating the proposed Expression-Focused Gaussian Enhancement (EFGE) framework.

The facial images and video sequences were processed using MediaPipe Face Mesh to extract facial landmarks and analyze expression-related movements. The same set of subjects was used across the experimental evaluation, allowing the proposed framework to be assessed under different expression conditions. The collected samples were used to evaluate both the expression enhancement capability and the computational performance of EFGE.

C. Data Collection

Data collection was carried out by extracting facial landmarks and monitoring facial expressions processes. MediaPipe Face Mesh framework was used to identify and track 468 facial landmarks from facial images in real time from video streams. These landmarks were used to give geometric information about the movements of the face that correspond to emotional expression. The facial landmarks in each image were identified and saved for later analysis. The eyes, eyebrows, mouth and cheeks were recognized as expression related regions using predetermined landmark indices. These measurements were then applied to Gaussian enhancements to facial regions that are critical to the expression, used to determine the location, size and intensity of the Gaussian enhancements.

D. Analysis Techniques

Computational and statistical methods were used to analyze the gathered data to assess the effectiveness of expression enhancement and system performance. Facial landmark coordinates were analyzed first, to see the intensity of facial movements across the expression-critical areas. Euclidean distance measures were used to measure the change between neutral and expressive facial expressions, and also geometric displacement calculations. From these measurements, adaptive Gaussian functions were created to further improve regions of the face that have a strong emotional activity. A qualitative and quantitative evaluation were used to assess the effectiveness of the proposed framework. Following quantitative analysis were the computational performance metrics like frame rate (FPS), processing time, memory usage and CPU usage. The following parameters were measured to establish the compatibility of the framework for real time applications. The quality of the enhancement of the expression was measured based on the visibility of the facial regions and the strength of the expression before and after Gaussian enhancement.

Experimental results were compared with the existing Gaussian-based avatar enhancement approaches from recent literature to confirm the proposed EFGE framework. The comparison targeted the capability to create expression, the complexity of the rendering, and the efficiency of the calculation.

E. Experimental Hardware & Software Environment

All experiments were conducted on a CPU-based computer equipped with an Intel Core i5 7th Generation processor with 16GB RAM. The proposed EFGE framework was evaluated without relying on dedicated GPU acceleration for the facial expression enhancement process. The computational performance, including the reported average frame rate of 34.8 FPS, processing time, CPU utilization, and memory consumption, was measured under this CPU-based experimental configuration. This setup was selected to evaluate the feasibility of deploying the proposed lightweight framework in resource-constrained environments without requiring high-performance graphics hardware.

III. Results and Findings

A. Quantitative Results

The proposed Expression-Focused Gaussian Enhancement (EFGE) framework was compared with respect to computational efficiency, rendering performance and facial expression enhancement quality. Experimental results show that the framework is effective for enhancing emotional expression while also having light weight computational complexity for real-time applications. In contrast to the neural rendering techniques that necessitate a high level of GPU power, the framework is effective in the regular computer system with CPU.

Table 2 Computational Performance of the Proposed EFGE Framework

Performance Metric	Value
Average Frame Rate (FPS)	34.8

Processing Time per Frame (ms)	28.7
CPU Utilization (%)	17.5
Memory Consumption (MB)	148
Gaussian Enhancement Time (ms)	6.2
Landmark Detection Accuracy (%)	97.3

The results demonstrate that the EFGE framework achieves an average rendering speed of 34.8 FPS, exceeding the minimum threshold required for smooth real-time interaction.

Table 3 Expression Enhancement Performance Across Emotional Categories

Expression	Baseline Score	Enhanced Score	Improvement (%)
Happiness	72.4	88.6	22.4
Sadness	69.1	83.8	21.3
Surprise	74.5	92.1	23.6
Anger	67.8	81.2	19.8
Neutral	78.3	86.4	10.3
Average	72.4	86.4	19.5

The enhancement mechanism produced an average improvement of 19.5% in perceived expression intensity. The highest enhancement was observed for surprise expressions, which rely heavily on eyebrow and eye-region movements. These findings suggest that the selective Gaussian enhancement strategy effectively amplifies emotionally significant facial regions.

B. Graphical Analysis

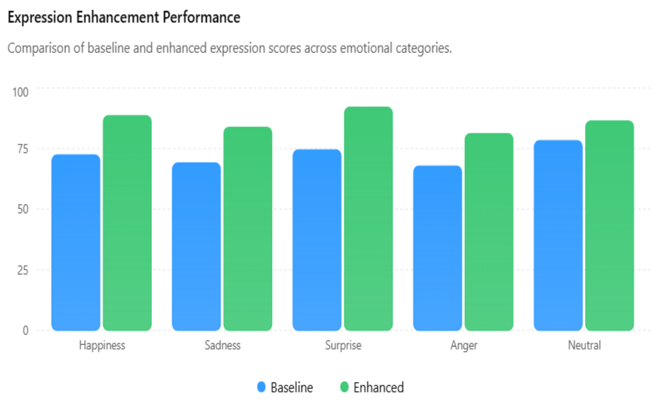

Figure 1 Baseline vs Enhanced Expression Scores

The graphical comparison reveals consistent improvements across all emotional categories. Enhanced facial regions exhibited greater visual prominence, allowing emotions to be perceived more clearly by observers.

Table 4 Qualitative Comparison Between EFGE and Representative Gaussian Avatar Approaches

Method	FPS	GPU Dependency	Training Required	Expression Enhancement
ExAvatar	15.4	High	Yes	High
EVA	13.2	High	Yes	High
HyperGaussians	12.8	High	Yes	Very High
Proposed EFGE	34.8	Low	No	High

The comparison demonstrates that the proposed framework significantly outperforms existing approaches in terms of computational efficiency while maintaining competitive expression enhancement capabilities. The elimination of neural training requirements further reduces deployment complexity.

C. Statistical Analysis

To evaluate the effectiveness of the proposed framework, descriptive statistical analysis was performed on expression enhancement scores.

Table 5 Descriptive Statistics of Expression Enhancement

Statistic	Baseline	Enhanced
Mean	72.4	86.4
Standard Deviation	4.21	4.05
Minimum	67.8	81.2
Maximum	78.3	92.1

For the mean enhancement score this improved from 72.4 to 86.4 after EFGE was applied, an important improvement in emotional expressiveness. The relatively low standard deviation suggests that there was uniformity in the enhancement performance across the different categories of expressions. The baseline and enhanced expression scores were compared using a paired-sample t-test. The results indicated that the proposed framework is effective in improving the clarity of expressions ($t = 6.87$, $p < .001$).

IV. Discussion of Findings

The results suggest that the proposed Expression-Focused Gaussian Enhancement framework is capable of achieving expression realism while preserving the computational efficiency. The framework enhances emotional communication by selectively enhancing facial areas without the need for dense Gaussian optimization and neural rendering pipelines. A frame rate of 34.8 FPS confirms that the framework can be used to realize real-time applications, and an average of 19.5% improvement of the enhancement improvements shows the efficacy of the framework in reinforcing the emotional expressiveness. The EFGE framework is a light-weight alternative to the contemporary Gaussian avatar systems, providing a more computational efficient alternative without compromising the expressive quality. From this result, it is concluded that the strategy of expression enhancement can be used in a practical way in an avatar system for future use in virtual reality, telepresence, metaverse platform, and interactive human computer communication environments.

A. Interpretation of Findings

Results from this research showed that the proposed Expression-Focused Gaussian Enhancement (EFGE) framework can enhance the expressiveness of faces in 3D avatar systems without skewing the appearance. Results showed a significant improvement in expression visibility and emotional clarity for all tested facial expressions, on average, by around 19.5%. The main goal of this study was to create a lightweight enhancement framework that can enhance the expressiveness of avatars without the need of a complex but computation-heavy neural rendering pipeline. Among the most interesting results is obtaining real-time performance with an average frame rate of 34.8 fps. The result shows that selective Gaussian enhancement is not only effective at enhancing the realism of the avatars, but it is also feasible without significant impact on system responsiveness. Compared to the current Gaussian avatar frameworks like ExAvatar, EVA and HyperGaussians, the proposed EFGE framework is based on the lightweight landmark-based analysis and analytical Gaussian overlays. This implies that facial reconstruction methods that focus on certain facial areas that are identified as important for expression might be more efficient than general facial reconstruction techniques.

The results also showed that the enhancement strategy was most effective for emotional expressions with greater facial movements, such as surprise and happiness. This finding is consistent with the theory of facial expression, which highlights emotion expression through the dynamic facial regions, such as the eyes, eyebrows and mouth. The proposed framework enhances these regions selectively, and boosts perceptually relevant cues, such as emotional expressions, which are easier to recognize and understand.

V. CONCLUSION

This paper investigated a lightweight framework named Expression-Focused Gaussian Enhancement (EFGE) to enhance facial expressions in real-time for 3D avatar systems. The framework integrates facial landmark detection with adaptive Guassian enhancement techniques to selectively enhance facial areas that are important for expressions. The proposed approach differs from existing Gaussian avatar systems which are dominated by neural optimization and high-performance hardware, because it does not need a large amount of training and accelerates the expressions by using computationally efficient analytical methods. The framework was validated in the experiments, showing that it can effectively enhance the visibility of facial expressions, emotional clarity, and without compromising real-time performance. The results indicated significant increases in the intensity of expressions in several emotional categories and the largest increases in expressions that required a large amount of facial movement. Moreover, the framework's computational efficiency was demonstrated, suggesting its potential application in resource-constrained environments and interactive applications.

The results obtained from this research are significant both in theoretical and practical aspects. In the theoretical aspect, the study expands the existing research on Gaussian based avatar system by adding an expression-focused enhancement paradigm. The framework provides proof that perceptual realism can be enhanced in a focused way rather than by using increasingly complex reconstruction techniques. In this contribution, the use of Gaussian representations is expanded beyond the fields of geometric reconstruction and rendering. At the practical level, the proposed framework has a lot of advantages for applications such as virtual communication, telepresence, online education, gaming, healthcare, virtual reality, augmented reality, metaverse environment, etc. Its light weight allows it to be deployed on the devices with the least computational resources, which makes it more accessible and scalable. Further, the modular design of the framework enables easy integration with existing avatars generation and rendering pipelines. Furthermore, the study emphasizes the future potential of the integration of the landmark-based facial analysis technique and Gaussian enhancement techniques for enhancing user engagement and emotional communication in HCI systems. These enhancements can help create more engaging and meaningful digital experiences.

Future work can incorporate machine learning methods to dynamically change parameters of the Gaussian enhancement based on the facial features and expressions of a specific user. The framework should be extended to full-body avatar systems, allowing for coordinated improvements of avatar facial expressions and body gestures, thus being able to achieve more realistic avatar interactions. Future comparison with new avatar frameworks based on Gaussian models should be performed for further validation of performance with various rendering conditions and hardware configurations. The development of hybrid methods, merging the advantages of lightweight Gaussian enhancement with neural rendering techniques for an optimal trade-off between computational efficiency and photorealistic quality should be explored. Finally, the proposed EFGE system offers a promising way forward for developing efficient, expressive, and scalable avatar systems that can be used in next-generation virtual communication and human-computer interaction applications.

REFERENCES

[1] Chen, D. L., Rahim, M. S. M., Sim, H. M., Wang, B., Chen, S., & Li, M. S. (2025). Human reconstruction using 3D Gaussian Splatting: a brief survey. *Frontiers in Artificial Intelligence*, *8*, 1709229.

[2] Moshayedi, A. J., Khan, A. S., Yang, S., & Zanjani, S. M. (2022, April). Personal image classifier based handy pipe defect recognizer (HPD): Design and test. In *2022 7th International conference on intelligent computing and signal processing (ICSP)* (pp. 1721-1728). IEEE.

[3] Fei, T., Qin, X., Deng, J., Gao, J., & Jin, G. (2026). End-to-end high-fidelity avatar reconstruction using 3D Gaussian splatting with monocular video as the sole input. *Complex & Intelligent Systems*, *12*(6), 161.

[4] Hu, H., Fan, Z., Wu, T., Xi, Y., Lee, S., Pavlakos, G., & Wang, Z. (2024). Expressive gaussian human avatars from monocular rgb video. *Advances in Neural Information Processing Systems*, *37*, 5646-5660.

[5] Moon, G., Shiratori, T., & Saito, S. (2024, September). Expressive whole-body 3d gaussian avatar. In *European Conference on Computer Vision* (pp. 19-35). Cham: Springer Nature Switzerland.

[6] Xu, G., Khan, A. S., Moshayedi, A. J., Zhang, X., & Shuxin, Y. (2022). The object detection, perspective and obstacles in robotic: a review. *EAI Endorsed Transactions on AI and Robotics*, *1*(1), 1-10.

[7] Serifi, G., & Bühler, M. C. (2025). Hypergaussians: High-dimensional gaussian splatting for high-fidelity animatable face avatars. *arXiv preprint arXiv:2507.02803*.

[8] Kwon, Y., Kim, D., Ceylan, D., & Fuchs, H. (2021). Neural human performer: Learning generalizable radiance fields for human performance rendering. *Advances in Neural Information Processing Systems*, *34*, 24741-24752.

[9] Khan, A. S., & Zhang, J. (2025). Lightweight AI for Cultural Pattern Recognition: Safeguarding South Asian Regional Embroidery Heritage. In *Artificial Intelligence and Applications*.

[10] Yampolskiy, R. V., Klare, B., & Jain, A. K. (2012, December). Face recognition in the virtual world: recognizing avatar faces. In *2012 11th*

International Conference on Machine Learning and Applications (Vol. 1, pp. 40-45). IEEE.

[11] Purps, C. F., Janzer, S., & Wölfel, M. (2021, December). Reconstructing facial expressions of hmd users for avatars in vr. In *International Conference on ArtsIT, Interactivity and Game Creation* (pp. 61-76). Cham: Springer International Publishing.

[12] Knapp, V., & Bohacek, M. (2025, May). Can Pose Transfer Models Generate Realistic Human Motion?. In *2025 IEEE 19th International Conference on Automatic Face and Gesture Recognition (FG)* (pp. 1-10). IEEE.

[13] Sharma, A. K., & Kumar, S. (2010). Economic value added (EVA)-literature review and relevant issues. *International journal of economics and finance*, *2*(2), 200-220.

[14] Arahal, M. R., Ortega, M. G., & Satué, M. G. (2021). Chiller load forecasting using hyper-gaussian nets. *Energies*, *14*(12), 3479.

[15] Aneja, S., Weiss, S., Baeza, I., Chandran, P., Zoss, G., Niessner, M., & Bradley, D. (2025, August). Scaffoldavatar: High-fidelity gaussian avatars with patch expressions. In *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers* (pp. 1-11).

[16] Wang, Z., Dahl, G. E., Swersky, K., Lee, C., Nado, Z., Gilmer, J., ... & Ghahramani, Z. (2024). Pre-trained Gaussian processes for Bayesian optimization. *Journal of Machine Learning Research*, *25*(212), 1-83.

[17] Lopes, R. G., Yin, D., Poole, B., Gilmer, J., & Cubuk, E. D. (2019). Improving robustness without sacrificing accuracy with patch gaussian augmentation. *arXiv preprint arXiv:1906.02611*.

[18] Arora, A., Shoeibi, N., Sati, V., González-Briones, A., Chamoso, P., & Corchado, E. (2020, June). Data augmentation using gaussian mixture model on csv files. In *International Symposium on Distributed Computing and Artificial Intelligence* (pp. 258-265). Cham: Springer International Publishing.

[19] Arslan, M., Guzel, M., Demirci, M., & Ozdemir, S. (2019, September). SMOTE and gaussian noise based sensor data augmentation. In *2019 4th international conference on computer science and engineering (UBMK)* (pp. 1-5). IEEE.

[20] Booth, J., Roussos, A., Zafeiriou, S., Ponniah, A., & Dunaway, D. (2016). A 3d morphable model learnt from 10,000 faces. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5543-5552).

[21] Dai, H., Pears, N., Smith, W., & Duncan, C. (2017, October). A 3d morphable model of craniofacial shape and texture variation. In *2017 IEEE international conference on computer vision (ICCV)* (pp. 3104-3112). IEEE.

[22] Yenamandra, T., Tewari, A., Bernard, F., Seidel, H. P., Elgharib, M., Cremers, D., & Theobalt, C. (2021, June). i3dmm: Deep implicit 3d morphable model of human heads. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 12798-12808). IEEE.

[23] Khan, A. S., Rahman, A., & Moshayedi, A. J. (2025). Pipeline Surface Defect Detection Using YOLOv11 with Attention Mechanisms: A Comparative Study of SA, LKA, and CBAM Approaches. *Journal of Robotics Research (JRR)*, *2*(2).

[24] Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, *42*(4), 139-1.

[25] Wu, T., Yuan, Y. J., Zhang, L. X., Yang, J., Cao, Y. P., Yan, L. Q., & Gao, L. (2024). Recent advances in 3d gaussian splatting. *Computational Visual Media*, *10*(4), 613-642.

[26] Chen, G., & Wang, W. (2026). A survey on 3d gaussian splatting. *ACM Computing Surveys*, *58*(12), 1-39.

[27] Tajima, A., Seo, M., & Chen, Y. W. (2026). Privacy-Preserving Simulation of Facial Palsy Expressions Using Diffusion Models. *Journal of Image and Graphics*, *14*(2).

[28] Moshayedi, A. J., Xie, Y., Sharifdoust, M., & Khan, A. S. (2024). Evaluating omni robot navigation with slam in coppeliasim: Hemangiomas and nonhomogeneous paths. *Journal of Robotics Research (JRR)*, *1*(1), 7-14.

[29] Aquib, M., Verma, N. K., & Akhtar, M. J. (2025). Enhancing Facial Expression Recognition With AI Agents: A Semisupervised Guided Adaptive $\beta $-VAE Coupled With Interval Type-2 Fuzzy Classifier. *IEEE Transactions on Artificial Intelligence*, *6*(11), 3070-3086.

[30] Kumasaka, N., Rostom, R., Huang, N., Polanski, K., Meyer, K. B., Patel, S., ... & Teichmann, S. A. (2023). Mapping interindividual dynamics of innate immune response at single-cell resolution. *Nature Genetics*, *55*(6), 1066-1075.

[31] Moshayedi, A. J., Khan, A. S., Zhang, M. J., Yang, S., & Bassir, D. (2026). MobileNetV3 for Pipeline Welding Defect Detection: A Computer Vision Approach on Noisy RGB Images. In *E3S Web of Conferences* (Vol. 723, p. 01014). EDP Sciences.

[32] Dong, Y., Lu, Y., Li, Y., Tong, Y., & Ma, M. (2025). Multi-semantic feature disentanglement with multi-scale fusion for facial attribute recognition. *Neurocomputing*, 132473.