The Indonesian Journal of Computer Science
www.ijcs.net
Volume 13, Issue 6, December 2024
https://doi.org/10.33022/ijcs.v13i6.4505

3D Image Generation Using Generative Adversarial Network for Virtual Art Gallery

Dorcas Oladayo Esan[1], Pius Adewale Owolawi[2], Chunling Tu[3]
oladayojadesola10@gmail.com[1], owolawi@tut.ac.za[2], du@tut.ac.za[3]
[1,2,3] Tshwane University of Technology, Pretoria, South Africa

Article Information

Received: 16 Nov 2024
Revised : 23 Dec 2024
Accepted : 30 Dec 2024

Keywords

3D Artistic Images, virtual art gallery, VAE, GANs, SDF

Abstract

Gallery art websites are often used to display artistic work for user accessibility. The conventional web-based gallery is built on the hyper-text Markup Language (HTML) which is limited in terms of content dynamism, interactivity, and scalability. GANs provide advanced capabilities with continuously evolving art experiences, making them a powerful tool for modern digital art galleries. Artistic websites have significantly contributed to the exhibition of artistic images since they can be accessed anywhere. However, the artistic web suffers from being too passive and lacks in-depth interactivity to keep people meaningfully engaged with an exhibition virtually. This paper explores the exhibition of 3D images within a virtual art gallery using an intelligent artistic web-based applications framework that integrates Variational Autoencoder and 3-D Signed Distance Function Cycle GAN (VAE-3DSDFCycleGAN) and quantitative questionnaire methods. The virtual gallery utilises GAN architectures to produce diverse and original 3D artworks, addressing traditional art galleries' spatial, viewing dimensions, image quality, and accessibility limitations. The questionnaire was used to evaluate the user's satisfaction. The experiment was done on the Coco African Mask dataset to generate 3-D images, yielding a high result and satisfaction in terms of the ease of use and viewing of the artistic image contents.

A. Introduction

Art exhibitions involve artists showing case their creative artwork for people to see and buy [1]. Protecting cultural history and heritage fosters a greater sense of community and collaboration, ultimately promoting a strong and thriving economy[2]. Museums play a crucial role in preserving historically significant artworks and artefacts, but recent data indicates a decline in attendance and overall impact, posing a threat to cultural preservation efforts.(Soumya et al., 2023). The advancements in technology have made this historical heritage in particular images to be kept using virtual applications[3]. These applications have assisted artists in the management and selling of their art content.

Another benefit of this application is that users are not required to be physically present at the museum to view art images. These virtual art applications are done using HTML pages containing photographs, texts, links, and virtual reality [4]. An immersive 360-degree view of the area surrounding a fixed location can be obtained by using panoramic photos, such as those based on QuickTime VR.

In 2021, David Zwirner an American artist presented the collections of his artworks in 50 online viewing rooms to potential buyers [5]. According to Zwirner, in 2019 his gallery sales increased by 400%. This sales increment attests to the fact that many artists are selling their artwork collections online without people physically present at the museum or exhibition building.

Artistic web-based galleries have made a significant contribution to the exhibition of artistic images since they can be accessed anywhere [6]. Traditional web-based art galleries are often built with HTML which primarily displays static content, requiring manual updates to add new artwork, which is a time-consuming process and limits the variety and novelty of the content. While HTML can support basic interactivity, it is limited in creating deeply immersive experiences. The static nature of HTML-based galleries may not fully engage viewers [4]. Furthermore, the conventional artistic web lacks a photorealistic image quality in real-time interactivity consequently leading to a user's lack of satisfaction [7]

Numerous works have attempted to address this problem focusing on the user's satisfaction. These methods include Dimensional Augmenter GAN (DiAGAN) [8], X-dimensional GAN(XDGAN) [9], geometry-aware 3D GAN [10], 3D Deep Convolution Generative Adversarial Networks (DCGAN) [11], etc., which have performed to varying degrees of success, but none have given conclusive solutions to address the challenging gaps. It is crucial to address the issue of imperfect 3D image generation, lack of in-depth interactivity and lack of user satisfaction in art web-based galleries. To generate three-dimensional images, this maintains real-time interactivity.

Hence, it is crucial to address the issue of imperfect 3D image generation and lack of in-depth interactivity to art web-based galleries to keep people meaningfully engaged with an exhibition to show the creativity of the art irrespective of users' locations. The idea adopted is to allow the user to navigate the shared features and dimensions underlying the artist's web-based gallery art collection visually and creatively.

This study presents a Variational Autoencoder (VAE) and 3-D Signed Distance Function Cycle GAN (VAE-3D-SDFCycleGAN). An encoder network and a decoder network make up the VAE network. With the help of the prior distribution

p(z), the encoder network compresses the 2D input data into a latent representation, which the decoder network then uses to create new input data. Reduced during training are the reconstruction loss and the Kullback-Leibler divergence between the prior and the learned latent distribution optimize the weights of the VAE.

The Web-based gallery applications can benefit artists in content management and easy accessibility of the artist's work to the users. The contributions of this paper are as follows:

- Application Development: designing and developing a newly proposed 3D artistic web-based gallery application that could intelligently assist artists in their art image contents management for users' accessibility, thereby addressing a critical aspect of being physically present at an art gallery building.
- Performance Evaluation: Detailed experimental evaluations of the proposed framework were conducted on African art image datasets to test the performance of the model and assign questionnaires to users to evaluate their level of satisfaction in using the newly designed applications in real-life environments.
- New GANs Framework: The development of a new GANs framework that integrates Variational Auto-Encoder (VAE) with Variational Autoencoder (VAE) and 3-D Signed Distance Function Cycle GAN (VAE-3DSDFCycleGAN) models. This offers an innovative approach to generating 3D artistic images, thereby improving image viewing.

To the best of our knowledge, not enough research has presented the user-friendly artistic web-based gallery applications on inclined on Variational Autoencoder (VAE) and 3-D Signed Distance Function Cycle GAN (VAE-3DSDFCycleGAN). This paper is arranged in the following order: Section II provides a preview of the work done in the area of virtual art and the selected theoretical background of the proposed model, Section III presents a detailed explanation of a VAE-3D-SDFCycleGAN model with a detailed design of the gallery applications, Section IV focuses on the various experiments, evaluations, and questionnaires used to test the model. The concluding remarks are shared in Section V.

B. Related Works

Analysis of user experience during the pandemic was presented [1]to give a clear perspective of experiences in a virtual art exhibition as an alternative art space by using interview methods along with the artists and ecosystem. The approach shows that virtual exhibitions can be reached by all groups in various regions at any time.

The 3D cultural heritage was introduced by [12] to address the significance of restoration, conservation, engagement, education, research, and ethics that hardly exists in virtual 3D cultural reproduction. The authors combine a literature review with quantitative and qualitative analyses of data through questionnaires and workshops with groups of specialists and non-specialists. The results demonstrate the importance of carefully designing 3D interactions in the personal

and cultural contexts of end-users and cultural institutions to create authentic cultural experiences.

Research [13] examined how digital interpretation and presentation technologies affect visitor experience and developed a measurement framework for the digital heritage experience to address the influence and effects of these technologies on visitor experience. Digital display technologies were found to be well-received by visitors and to have had a positive impact on certain aspects, such as promoting exploration and advancing knowledge about the site. The study employed a mixed method, combining semi-structured interviews with a questionnaire based on a digital experience. The authors were unable to develop an interactive web-based artistic museum for visitors.

Furthermore, the study by [14] focused on assessing the user experience associated with mobile museums, tackling the problem of creating and assessing an interactive digital storytelling platform intended for museum visitors. The User Experience (UX) and data logging were assessed using methods like ethnography, pre- and post-experience in-depth interviews, and questionnaires. The study demonstrated how difficult it is to comprehend user experience (UX) and what factors contribute to its effectiveness in diverse cultural contexts. The mobile museum was able to show the artworks in a static position as the visitors can only view the artworks in 2D.

A thorough evaluation of the virtual online tour in a museum was conducted by work done in [15]. The authors interviewed users to learn about and analyze their satisfaction with the tour. According to the study's findings, the exhibition was interactive and had visual authenticity, but not enough behavioral authenticity.

A user-centred evaluation model for audience experience and display was researched and presented by [16] to address the problem of user experience on websites in the context of the ongoing development of digital museums. To analyze using three models, the authors employed twenty digital museums. An evaluation of the three models' user experiences was conducted through the implementation of prototype websites. According to the findings, user identity variations and website content have an impact on user intentions, attitudes, and quality.

Effective visual communication in art design is a problem that was addressed by the work presented in [17] a quantitative analysis of the impact of virtual reality technology on visual communication and art design. The author was able to prove, through the use of virtual reality technology, that a system of visual communication design based on virtual reality can successfully increase the impact of visual communication art design, with good design performance.

Furthermore, a user-focused platform was introduced by [18] for developing virtual 3D exhibitions with VR capabilities. The goal was to make a wide array of exhibits easily accessible and explorable beyond the limitations of traditional museums. This platform allows users to collaborate with museum curators to create interactive and immersive virtual 3D/VR exhibitions.

Theoretical Background

The following section will cover a brief overview of the theoretical framework that was used to implement the suggested technique.

1. Variational Auto-Encoder (VAE)

A generative model called a variational autoencoder (VAE) is specifically made to capture the underlying probability distribution of a given data set and produce new techniques. The autoencoder neural network is composed of two deep learning-based modules, the encoder and the decoder [19]. Data is compressed by the encoder and stored in a high-dimensional space called hidden space. Equation (1) displays the encoder loss function.

$$loss = \| \hat{x} - x \|_2 + KL[N(\mu, \Sigma) \| \mathrm{N}(\mathrm{o}, \mathrm{I})] \quad (1)$$

Where x and y are the target sample and generated sample respectively, KL is the Kullback-Leibler divergence. The encoder's output is represented by the means and variance, μ and Σ. The opposite is done by the decoder. In other words, it creates a high-dimensional space out of the hidden space. The decoder's goal is to generate an output equal to its input to guarantee that the hidden space extracts the greatest amount of data from the dataset space. The decoder's loss function is given as in equation (2).

$$loss = D(\hat{x}) - \delta \| \hat{x} - x \|_2 \quad (2)$$

Where $\hat{x}$ is the target sample, y is the generated sample, D is the discriminator, and μ and Σ are the means and variance that the encoder produced.

2. Signed Distance Function (SDF)

The SDF is a powerful data format for storing distance values in 3D textures. These distance values are a measure of how far the selected location is from the surface. The specified distance function (SDF) represents the shortest distance function from a point in space to a surface [20]. Typically, distances are given as positive for points outside the surface and negative for points inside the surface. In theory, a function can act as an SDF for an object if it has a gradient norm at every point in space. Also, the gradient norm of the network output is constrained to 1 for the input points, these networks turn out to be hidden SDFs. The 3d-SDF is defined as in equation (3)

$$F\_S(x) = (-d(x, \partial S)\, d(x, \partial S) = (x \in S, x \notin S) \quad (3)$$

$$X = \{(P_i, s_i, V_i)\}_{i \in [n]} \quad (4)$$

Where the unsigned distance function is defined by $d(x, \partial S) = min_{y \in \partial S} \|x - y\|_2$, Shape $S \subset \mathbb{R}^3$ where $x \in S$ are points inside the shape, $x \notin S$ are outside, $x \in \partial S$ are shape boundaries, $[n] = \{1 \dots \dots n\}, P_i \in \mathbb{R}^3$ are 3D point locations. The 3DSDF approximation is defined as in equation (5)

$$F_X(x) = \Sigma_{i \in [n]} w_i(x)\, I_{V_i}\left(\frac{x - P_i}{s_i}\right) \quad (5)$$

Given an [x, y, z] spatial point as input, SDF derives the distance from that point to the nearest face of the underlying displayed object. When a spatial point is

inside (negative) or outside (positive) of an object's surface, it is indicated by the sign of the SDF output as in (6).

$$SDF(X) = s : X \epsilon \mathbb{R}^3, s \epsilon \mathbb{R} \quad (6)$$

Where x, y, and z are groups of points that are used to represent an underlying geometry.

3. Cycle Generative Adversarial Networks (CycleGAN)

Cycle GAN loops are used to transfer features from one image to another or to compare the distribution of one image with another image. The Cycle GAN framework was introduced for the image-to-image conversion task without a paired training database [21]. Learning the mapping from the P domain (input) to the Q domain (output) and vice versa is done by losing loop consistency. The main connection between the two domains is the combination of the two losses and the adversary loss, which is the cyclic loss as in (7).

$$L_{GAN}(G, D_Q, P, Q) = E_{q \sim s_{data(q)}}[\log D_Q(q)] + E_{p \sim s_{data(p)}}[\log(1 - D_Q(G(P)))] \quad (7)$$

Where G attempts to produce an image (D) and D_Q seeks to distinguish between synthesized images $G(P)$ and images q (in the Q domain). Whereas opponent D seeks to maximize the objective cycle-consistency loss.

$$L_{cyc}(G, F) = E_{p \sim s_{data(p)}}\left[\|F(G(P)) - P\|_1\right] + E_{q \sim s_{data(q)}}\left[\|G(F(q)) - q\|_1\right] \quad (8)$$

The cycle GAN model shows the most extreme unaided learning environment possible by stretching its edges. The cycle GAN is fit for changing variety and surface, showing the model's generalizability in an assortment of more extensive territory applications led without matched information [21]. The CycleGAN configuration isn't exactly equivalent to other GANs to such an extent that it contains two arranging capacities (G and F) that go about as generators and their related Discriminators $D(X)$ and $D(Y)$: The generator arranging abilities are as in (9) and (10).

$$G : X \rightarrow Y \quad (9)$$

$$F : Y \rightarrow X \quad (10)$$

where X is the input image distribution and Y is the desired output distribution. These have discriminators that match them: $D(X)$ try to translate mapping F similarly, while $D(Y)$ Pushes G to convert X (generated Output) into an indistinguishable result of the target domain Y.

Baseline Models

The baseline theory that this research used is further explained in the following sections.

1. Deep Convolutional Generative Adversarial Network GAN (DCGAN)

A generative G network and a discriminative D network make up the DCGAN paradigm for image generation. The DCGAN image generation paradigm is

composed of a discriminative D network and a generative G network [22]. Using deconvolution layers, the G-network generates a neural deconvolution unit that transforms d-dimensional vectors into images. The D network employs the same structure as a traditional CNN to determine whether the data is an actual image from a predefined dataset or G. In equation (11), the model training is displayed.

$$\text{Min}\max(D,G) = E_{X\sim p_{data(x)}}[\log D(x)] + E_{z\sim p_z(z)}\left[\log\left(1 - D\left(G(z)\right)\right)\right] \quad (11)$$

where $p_{data(x)}$, and $p_{z(z)}$)) are the probability distributions of x and z, respectively, and z is a dimensional vector of random numbers. *(1-D(G(z))* is the result $p_{z(z)}$ and *D(x)* is the probability that an image $p_{data(x)}$ will generate the input. *D* trains to make more correct answers than G, and G trains to make D think less by lowering $\log(1 - D(G(z))$.

2. Conditional Generated Adversarial Networks (cGANs)

This model can learn the input data and control the generated data using the information obtained from generator G and discriminator D to establish the relationship between different classes [23]. This model utilizes the minimax objective function as in (12).

$$min_G max_D f(D,G) = E_x[\log(D(x|y))] + E_z[\log\left(1D\left(G(z|y)\right)\right) \quad (12)$$

Where x is the input, y is the output and z is the random noise vector. A thorough description of the proposed method for creating three-dimensional images from two-dimensional artistic images can be found in the following section.

C. Research Method

A secondary data approach was used for data collection, and the data were analyzed separately. The qualitative research method was used by the investigators. To better interpret the data, qualitative research aims to comprehend and explain the data. The use of quantitative research, which involves analyzing the data collected numerically to make decisions, was also adopted. Quantitative data were gathered for this study from simulations and empirical analysis.

1. Methodology Detail

The investigation in this study employs an experimental research design. This makes use of experimental data to back up positive claims about gaps and contributions in section 1. In this study, the artistic image data contains the historical African arts. The image is passed through data processing to remove unwanted artefacts. The preprocess image is passed to the feature extraction stage where the relevant features are extracted from the image. The extracted features are passed to the image generation for training purposes. Furthermore, a qualitative method was employed to validate user satisfaction with the newly developed 3D web-based art gallery.

2. Tools and Technologies

Since the application development is based on the client/server architecture. It is important to give a brief description of some of the basic tools and technologies that were used for the implementation of the system. Different software was used in the research for analysis. The implementation was done using PHYTON and Java programming languages. MongoDB Atlas database (NoSQL) is used as the database. Expressjs is used as the backend framework, react is used for the frontend framework and Threejs is utilised for the front end of 3D rendering was used for the website development environment, the reason is that this database has been used and tested in various application development environments. PHP was used as the main scripting language for the development of the 3D web-based gallery function.

3. Development of 3-D Image Generation with CycleGAN

As illustrated in the system architecture in Figure 1, this section explains the proposed approach. There are four stages in the proposed system framework: the image acquisition stage (a), the image pre-processing stage (b), the image feature extraction stage (c), and (e) the image generation stage. The following sections provide a detailed explanation of the stages used in the implementation.

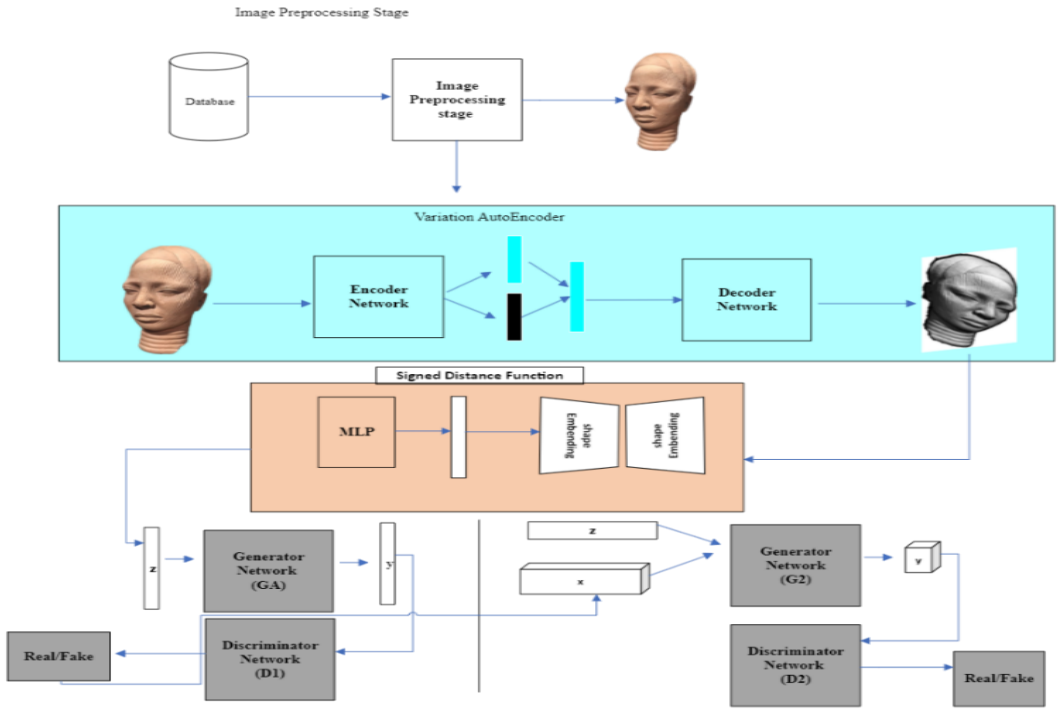

Figure 1. Artistic 3D Image Generation with VAE-3D-SDFCycleGAN Framework

i. Image Acquisition Stage

The experiments in this study used COCO Africa Mask art images [24]. The dataset is images of African masks that will help readers experience the pinnacle of African art. This dataset consists of 9300 images of African art.

ii. Image Pre-processing Stage

The image pre-processing step is an important step in computer vision to remove unnecessary noise and reduce the computational power of the proposed model. After the image pre-processing stage, the final image is sent for image normalization. To make it easier for machine learning algorithms to process the images, all of them are rescaled to a standard image range, which is usually between 0 and 1. This is done as shown in equation (13).

$$I^1 = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{13}$$

The data point within the features is denoted by x, and max(x) and min(x) are the maximum and minimum values of the features, respectively.

iii. Image Augmentation and Noise Removal

Data augmentation techniques are becoming more and more common in other fields as well as being powerful tools for developing computer vision models. By adding variances to the input data, these techniques aim to increase the size of datasets and avoid overfitting. Data augmentation has applications beyond improving overall performance; it can be used to assess how models are robustly trained. The output of the augmented image is passed to the median filtering for noise removal purposes.

iv. Feature Extraction Stage

The grayscale value difference between each image pixel qc and a circle with radius R is used to create a binary pattern known as the Local Binary Pattern (LBP). This is defined from the centre q_c, as in equation (14).

$$LBP_{P,R(q_c)} = \sum_{p=0}^{R-1} S(x)\, 2^p \tag{14}$$

The difference between the intensity level of the central pixel (q_c) and the neighbouring pixel(q_p) Within a circular neighbourhood of radius R and P neighbouring pixels are denoted as $x = q_p - q_c$ this is shown in equation (15).

$$S(x) = \begin{cases} 1 & x \geq 0 \\ 0 & otherwise \end{cases} \tag{15}$$

When a Neighbor is not precisely in the centre of the pixel, interpolation is used. The LBP code is invariant to monotone transformations of image brightness because of the label function *s(x).* Equation (16) illustrates how the histogram of these various features can be utilized to describe the texture and characterize the distribution of LBP patterns.

$$h = \sum_{i=1}^{W} \sum_{j=1}^{H} \delta\left(LBP_{P,R}(i,j) - k\right) \tag{16}$$

The number of LBP patterns, k, satisfies the inequality $0 \leq k \leq d = 2^p$. Here, δ represents the Heaviside function, while W and H refer to the dimensions of the image.

v. Image Regeneration Stage

To train the VAE encoder function, the normalized image output is fed as input to the VAE-3D-SDFCycleGAN. The input image is reduced in size and the z-compressed hidden vector is obtained by passing it through multiple layers in the VAE. For the decoder to reconstruct the input image, the encoder first extracts the mean and standard deviation of each latent variable.

vi. 3D Image Point Extraction

The 3D-SDF receives the output of the trained image and stores the three-dimensional spatial points P∈R^3. are extracted using the closest point distances, k(p)∈R, to the surface points. In this case, the object's P value may be positive (+) or negative (-). Equation (17) defines the distance function.

$$K(x) = \begin{cases} dist(x, \delta\Omega) & if\ x \epsilon \Omega \\ 0, & if\ x \epsilon \delta\Omega \\ -dist(x, \delta\Omega) & if\ x \epsilon \Omega^2 \end{cases} \quad (17)$$

A non-zero volume open set of points with an enclosed boundary is denoted by δΩ, while x is the 3D point and dist is the 3D Euclidean distance. The definition of δΩ is the smallest distance, as given in equation (18), between any point on the boundary and point x.

$$dist(x, \delta\Omega) = \inf_{y \epsilon \delta\Omega} dist(x, y) \quad (18)$$

The sign of point x concerning the boundary $\delta\Omega$ is defined in equation (19).

$$sign(x, \delta\Omega) = \begin{cases} 1 & if\ x \epsilon \Omega \\ 0 & if\ x \epsilon \delta\Omega \\ -1 & if\ x \epsilon \Omega \end{cases} \quad (19)$$

Where the SDF in (13) is rewritten as in equations (18) and (19) as in equation (20).

$$k(x) = dist(x, \delta\Omega) \cdot \text{sign}(x, \delta\Omega) \quad (20)$$

Where Ω be a region in 3D Euclidean space with a boundary denoted by δΩ. The distance from a point x in the region to the boundary δΩ is defined as the smallest distance from x to any point on the boundary. The three-dimensional point extracted from the image is then used as input for the CycleGAN to create artistic images in the following stage.

vii. 3D Image Generation Stage

Using a novel training scheme, we combine our new VAE-3D-SDFCycleGAN system to address the challenge of reconstructing 3D objects from 2D RGB images. A generative, encoder-decoder system called a VAE is used to learn complicated distributions. A latent vector is created by sampling a set of Gaussians parameterised by a vector of means and variances that the encoder creates after observing samples from the target distribution.

The decoder receives this vector and tries to replicate the original sample. By selecting the latent vector from a normal distribution, this encoding system makes generation easier. In addition to the encoder's loss having a regularizing KL divergence term that encourages the Gaussians to resemble conventional normal distributions, both networks are penalized based on the reconstruction error of the sample. The GAN's generator network concurrently uses the VAE's decoder network.

The latent vector is created by sampling the 400-dimensional vector of means of variance that the VAE's encoder creates from an image using Gaussians. The reconstructed object is given to the discriminator to evaluate its validity after the latent vector has been run through the decoder/generator network. Equation (21) presents the loss functions for both the encoder and decoder/generator networks.

$$loss = \| \hat{x} - x \|_2 + KL[N(\mu, \Sigma) \| N(o, I)] \quad (21)$$

Where KL represents the Kullback-Leibler divergence, x is the target sample, x̂ is the generated sample, and μ and Σ represent the means and variances generated by the encoder.

The opposite is done by the decoder. In other words, it creates a high-dimensional space out of the hidden space. To guarantee that the hidden space extracts the greatest amount of data from the dataset space, the decoder's goal is to generate an output equal to its input. The decoder's loss function is shown in equation (22).

$$loss = D(\hat{x}) - \delta \| \hat{x} - x \|_2 \quad (22)$$

Where D is the discriminator, x is the target sample, $\hat{x}$ is the generated sample μ *and* Σ are the means and variance produced by the encoder. This system uses the same discriminator and generator networks as 3D generation, and its encoder network is a basic 5-layer convolution neural network. The generator only learns every five batches during training, but the discriminator and encoder networks are trained at every batch. This final point is crucial to the systems' integration; the system won't converge if the encoder and discriminator aren't trained together at each iteration.

4. Website Development of 3D Artistic Gallery Component

The 3D artistic web-based application gallery interface is shown in Figure 2, this section explains the proposed approach. There are four steps in the proposed system architecture: the image acquisition stage (a), the client side (b), and the server side (e) the application side.

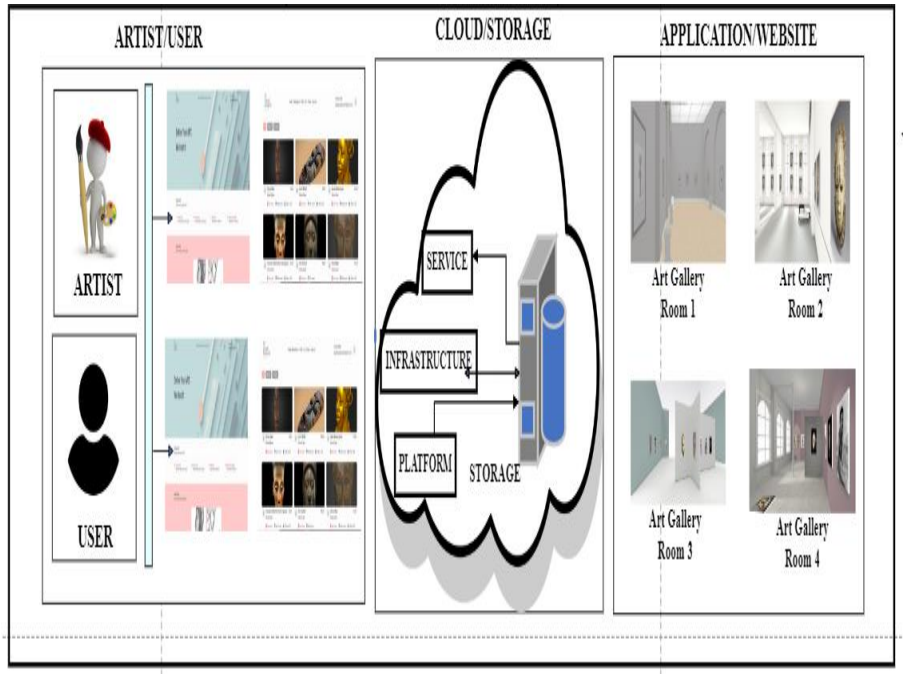

Figure 2. 3D-Art Web-Based Virtual Gallery Framework.

5. Key Technologies

i. Three.js Library

Three.js is a cross-browser JavaScript library and application programming interface used to create and display 3D computer graphics in a web browser using WebGL. Three. Js is an application programming interface and cross-browser JavaScript library that is used to create and display 3D computer graphics in a web browser using WebGL. Three.js is a third-party library written in JavaScript that is part of the webGL library. It is used to abstract the graphics interface, allowing for the rendering of 3D scenes, lighting, and models. The design of the digital virtual exhibition platform discussed in this study can be expedited by utilizing the framework provided by three.js.

ii. The Web Graphical Library (WebGL)

WebGL, also known as the Web Graphics Library, is a collection of industry-standard protocols for 3D drawing. WebGL integrates OpenGL and JavaScript. Concerning web3D technology, webGL technology can address the scene rendering issue. Because the hardware needed for 3D accelerated rendering can be provided by webGL.

iii. Server

On the server side of the website, Nodejs was used as the server language. The API from the front end is used to communicate with the server.

iv. Model & Database

The “model” and “database” are the most important sections in this website development as they implement the art images and provide data to the front end.

MongoDB Atlas database (NoSQL) is used as the database. Expressjs is used as the backend framework, react is used for the frontend framework and Threejs is utilised for the front end of 3D rendering.

v. Overall Website Design

The general framework of the system and the associated Web3D functional framework comprise the overall design planning of a digital virtual exhibition platform built on 3D web technology. There are three main tools for developing the 3D web function subsystem. WebGL and js. To create the Obj file, 3DSmax must first build the 3D model. Next, use three. Js to build web3D's operating domain. The scene is then rendered by combining with WEBGL, and at last, the scene is built. This is done by calling the Obj file via the input statement on the created web page.

Using Web3D technology as a foundation, the interactive tool was created. Xj3D is used to manipulate the scene in the X3D model, and a MySQL database stores data about 2D artworks, like paintings and tapestries. The tool was created in Java and manages scenarios that are described in an X3D file.

vi. Use Case Diagram

This is used to depict the interaction among the elements of the art web-based gallery. It is also used in system analysis and design to identify, clarify, and organise the system requirements of the art web-based gallery. In this research, the main actors of the art gallery are the administrator, system user, customers, and the artists who perform different types of use cases such as managing art, uploading art, uploading art, uploading art, updating the gallery, and managing the gallery. The elements of the UML use case diagram of Art web-based gallery system development are shown in Figure 3.

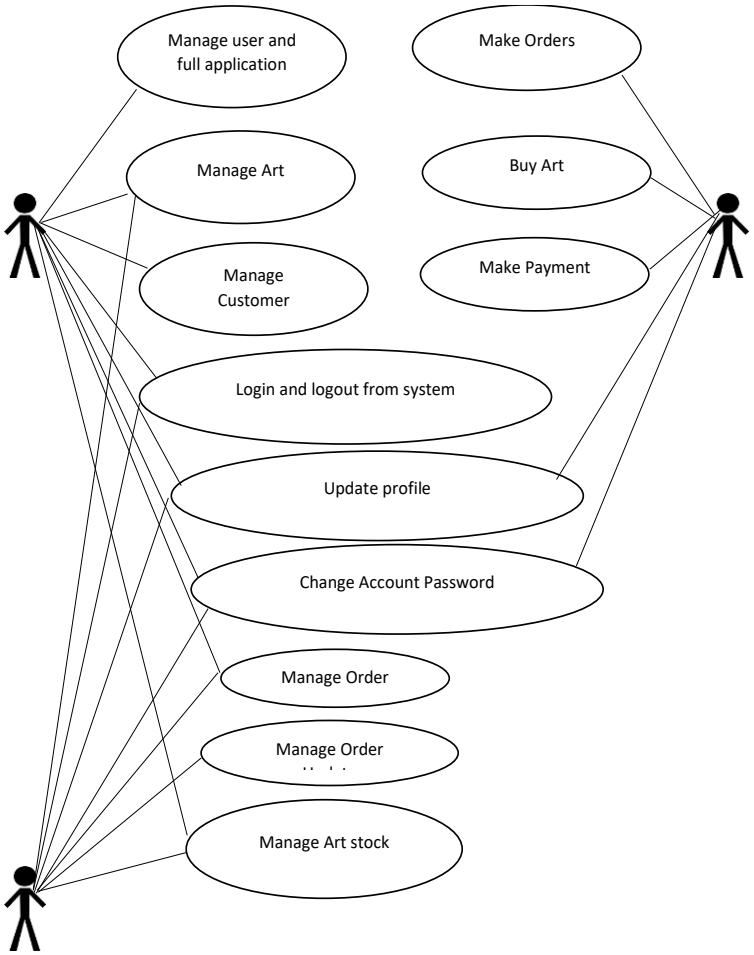

Figure 3. Use Case Diagram for Art Web-Based Gallery.

6. Evaluation Metrics

The proposed model is evaluated with two different evaluation schemes:

i. Qualitative Evaluation Scheme

To demonstrate how well the used model performs in terms of image sharpness, the image enhancement is visually examined for qualitative assessment. The following qualitative evaluation metrics are used in this study: Fréchet Inception Distance (FID), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Inception Score (IS), Chamfer Distance (CD) and Learned Perceptual Image Patch Similarity (LPIPS).

ii. Quantitative Evaluation Scheme

The questionnaire is assigned using a Google form and sent to the users to evaluate the performance of the system.

D. *Experimental Evaluation and Results*

This section showcases various experiments conducted to generate historical artistic images using the suggested methodology. The configuration, parameter settings, and simulation of the experiment are covered in the next section.

1. Experimental Configuration and Parameter Setting

The Tensor Flow library was installed independently on a Python Google Collaboratory computer, and the training and testing procedures were conducted on a computer equipped with a GPU frequency of 2.5 GHz. On the publicly accessible African Coco mask and CelebFace datasets, experiments were conducted both quantitatively and qualitatively. Each dataset was validated using specific techniques, as detailed in Section 2.2. Throughout the simulation, ten thousand frames were sampled. The epoch number is 250, the training batch size is 1500, and the generator and discriminator learning rates are 0.0001. During training, stability and speed are increased by the values selected for the training rate and batch iterations. The next section provides a detailed description of the experiments.

2. Hyperparameter Selection for VAE-3DSDFCycleGAN

During the implementation, the VAE-3D-SDFCycleGAN model parameter values were set to the following as shown in Table I.

Table 1. Hyperparameter used for training vae-3D-SDFCycleGAN implementation

	Values
Discriminator hidden size	128
Leaky _alpha	0.01
Discriminator Learning rate	0.0001
Generator Learning rate	0.0001
Epochs	250
Batch Size	1500

3. Experiment 1: A Qualitative Evaluation of the Proposed Model and the Existing GANS on Coco Africa Mask Dataset

This study aims to showcase how various image generation techniques affect the COCO African Mask dataset's quality. The images produced using 3D-SDF-CycleGAN and other chosen methods can be seen on the right side of Figure 4.

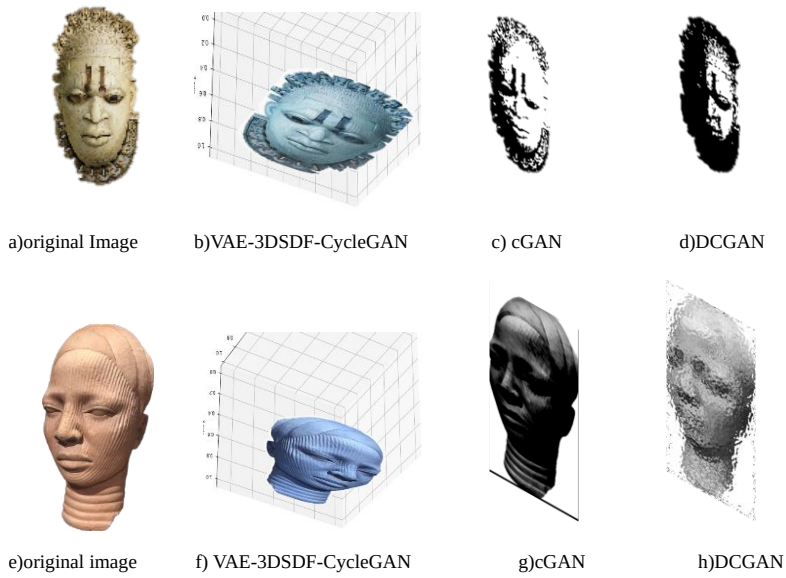

Figure 4. Qualitative Evaluation of 3D Artistic Image Generation on the Proposed Method and Other Selected GAN Methods.

Figures. 4(a), (e) shows the original artistic image, Figures. 4(b), and (f) represent the results of generating a three-dimensional generative image using the proposed method and Figures. 4(c), and (g) are images generated using the cGANs method. Finally, Figures. 4(d), and (h) show the images generated by the DCGAN method. When compared to other implementation methods, VAE-3DSDFCycleGAN produced clearer three-dimensional images, according to a visual inspection of the results generated by the proposed results.

Additionally, we demonstrate that the suggested model outperforms the DCGAN and cGANs techniques employed in our implementation in terms of producing 3D image quality that is higher, more comprehensive, and more accurate. The corresponding training loss values of the chosen models with the suggested model are shown in Table II.

Table 2. Generator and Discriminator Loss Values with different epoch values

Epoch	**cGANs**		**DCGAN**		**Proposed model VAE-3D SDFCycleGAN**	
	Discriminator loss (x10)	Generator loss (x10)	Discriminator loss (x10)	Generator Loss (x10)	Discriminator loss (x10)	Generator Loss (x10)
50	0.232	0.267	0.311	0.334	0.218	0.22
100	0.226	0.299	0.331	0.347	0.353	0.29
150	0.453	0.291	0.352	0.335	0.319	0.25
200	0.323	0.283	0.351	0.341	0.321	0.21
250	0.433	0.224	0.377	0.369	0.306	0.21

Table II displays the training loss iterations for each model. It is evident from these that the suggested model has lower content loss values than the other

models, indicating that the generated image content is consistent with the original image. The generator and discriminator loss curves from the GAN training procedure are displayed in Figure 5. These curves shed important light on the GAN model's convergence and performance, clarifying the learning dynamics and stability attained during the training rounds.

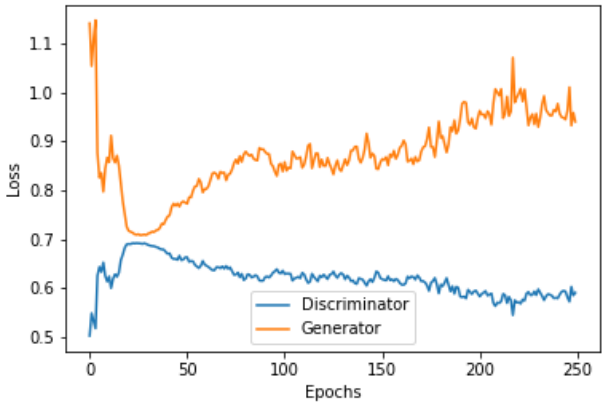

Figure 5. Loss of Generator and Discriminator Values in the VAE-3D-SDFCycleGAN Training Process

The graph shown in Figure 5 provides insights into the convergence and stability of the model, which is one of the study's goals. This graph depicts the model successfully generating real images and the discriminator can't tell apart (the higher the generator loss function and the lower the discriminator loss function the better the system).

4. *Quantitative Evaluation of Proposed and Selected GAN Methods on COCO Africa Mask Datasets*

In addition, the performance of each method is evaluated by quantitative evaluation in terms of the FID, IS, SSIM, PSNR, MSE, LPIPS and CD scores. The summarized results generated images are shown in Table III.

Table 3. Performance Evaluation of GAN Baseline Methods

Model	FID	IS	PSNR	SSIM	MSE	CD	LPIPS
ArtGAN[20]	22.8	8.40	19.76	0.56	-	-	1.346
LRGAN[25]	-	7.17	-	0.59	80.22	-	-
MSGAN[26]	28.44	-	17.78	-	73.11	-	-
PROGAN[27]	-	17.8	23.1	0.57	59.75	-	-
SNGAN[28]	22.81	13.5	27.2	0.65	65.12	-	-
Proposed model	10.03	18.5	33.1	0.66	55.12	0.313	0.24

Table III presents an additional analysis of the generated images' performance to FID, IS, SSIM, PSNR, MSE, CD, and LPIPS. Based on the outcome displayed in Table III, the suggested model has an FID score of 10.03 and an IS score of 18.52. When the FID and IS scores are lower, the images generated by the suggested method resemble the original input image more. The generated image is like the original input image when compared to it using the suggested method,

which yields an SSIM score of 0.662, a CD of 0.313, a PSNR score of 35.15, and an LPIPS of 0.244. Furthermore, the suggested approach produced an MSE of 55.12.

5. Benchmarking Proposed Model with Other Existing Techniques in Literature

The comparison of the proposed model with other approaches in terms of FID, IS score, PSNR, SSIM, MSE, CD and LPIPS score is shown in Table IV.

Table 4: Comparative of Proposed Model with Baseline Models

Models	**Performance**						
	FID	**IS**	**SSIM**	**PSNR**	**MSE**	**CD**	**LPIPS**
cGANs	17.43	11.11	0.613	24.13	84.78	0.833	1.985
DCGAN	19.22	14.22	0.524	22.32	88.01	0.922	1.376
Proposed Model	10.03	18.52	0.662	33.15	55.12	0.313	0.244

From Table IV, the proposed model shows significantly better performance on the Coco African Mask datasets with FID of 10.03, IS of 18.52, PSNR of 33.15, SSIM of 0.66, MSE of 55.12, Chamfer Distance of 0.313 and LPIPS of 0.98 compared to other models. It's worth noting that the higher the IS, PSNR and SSIM of the model the better image quality generated, and the lower the FID the better the structural features that the proposed methodology exhibits when compared with selected baseline techniques.

6. Experiment 2: Quantitative Evaluation of the Designed Application

i. Login/Sign up Feature

This login feature shows where artist and user can create their username and password to sign off to the art gallery. The user details are stored in the database, After the sign the user can use the registered details to log into the art gallery, the capture of the application is shown in Figure 6.

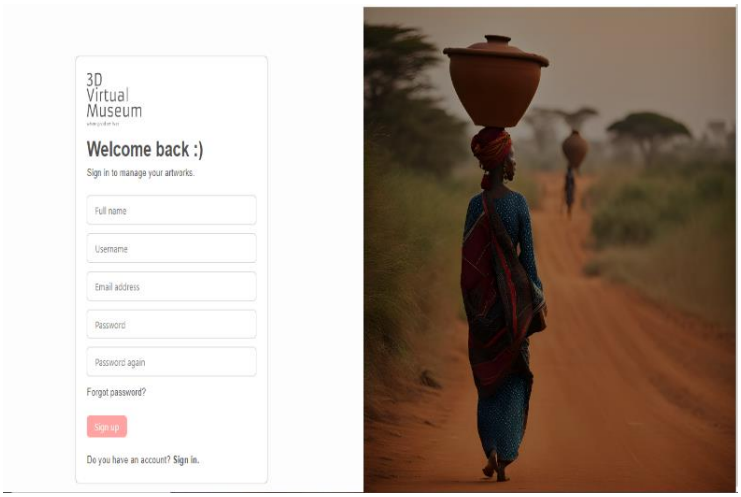

Figure 6. Login Graphical User Interface for 3D-art Web-Based Virtual Gallery.

ii. Home Page Feature

Here the details regarding the art are displayed in terms of the information that is needed. On this page, there are features such as gallery, chart, profile, and sign-out at the upper menu. A screen capture of the home page is shown in Figure. 7.

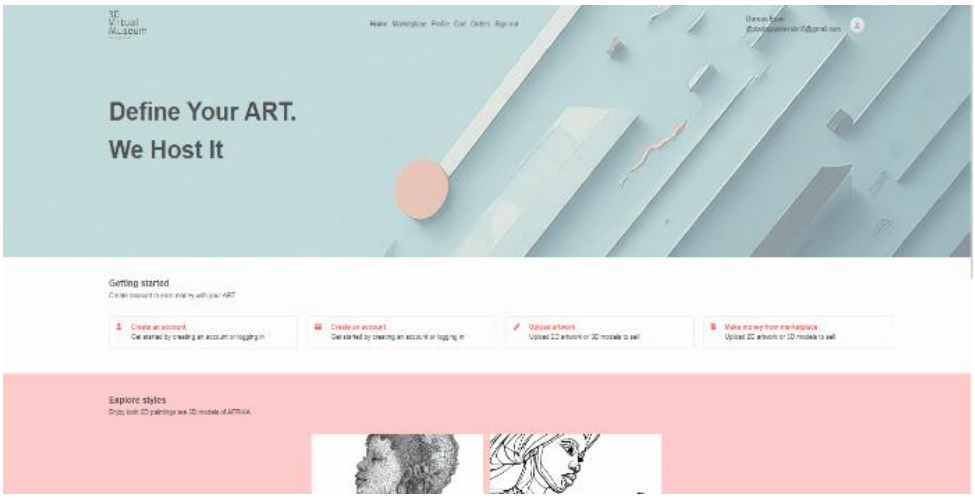

Figure7. Home Page Interface for 3D-art Web-Based Virtual Gallery.

iii. Artist Gallery Feature

The art gallery features the uploaded image by the artist as shown in Figure 8.

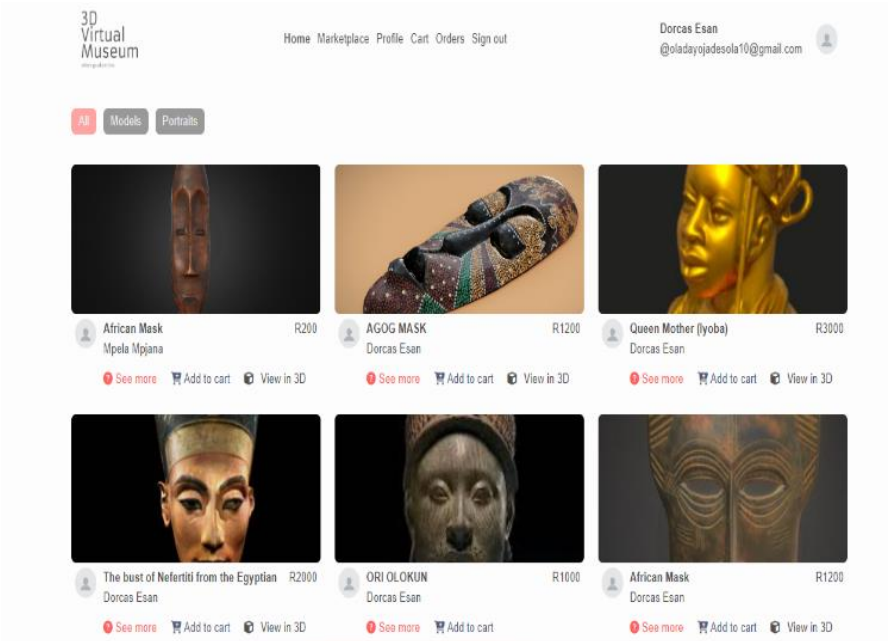

Figure 8. Artwork Showcase Interface Gallery Interface for 3D-art Images.

The uploaded images contain the name of the images uploaded and the selling price of each image. The user can click on any image and add them to the chart basket.

iv. 3D-Virtual Gallery Feature

This feature consists of different room sections where the artist displays the different images as shown in Figure 9. These images can be viewed in 3-D dimensions to show detailed information about the image in the gallery. Furthermore, at this gallery the information that clearly explains the image is included in this 3D virtual gallery for understanding of the users.

Figure 9. 3D-Artwork Exhibition Room.

7. Experiment 2.2: Qualitative Evaluation of the Designed Applications on User Satisfaction Using Questionnaire.

This section presents the results of the study based on the data collected from the questionnaires given to participants and the corresponding responses are presented in subsequent sections:

1. Demographic

Q1: *What is your Gender?*

From the 40 questionnaires distributed to both genders (female and male), 75% were female and 25% were male as shown in Figure 10.

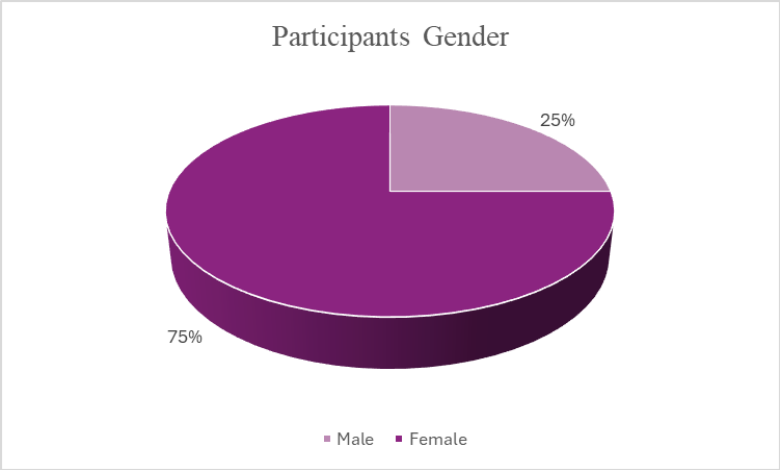

Figure 10. Total Number of Genders Participated in the Questionnaire.

Q2: *What is your level of study?*

Figure 11 shows the percentage of the participant's level of education used in the study.

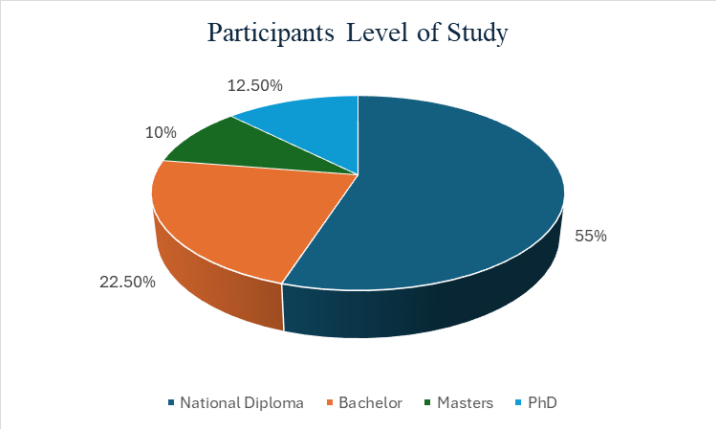

Figure 11. Participant Level of Study.

In comparison to other levels of study, 55% of the participants have a National Diploma, 12.5% of the participants have a PhD degree, a total of 22.5% have a bachelor's degree, and 10% of the participants are master's degree holders.

2. User Satisfaction towards using the Art web-based Gallery Applications

To assess the level of user satisfaction with the use of an art web-based gallery, the following questions were asked.

Q3: *Have you used any art gallery before?*

The participants were asked in the questionnaire if they had any prior experience with any art gallery. The responses gotten from the participants are shown in Figure 12.

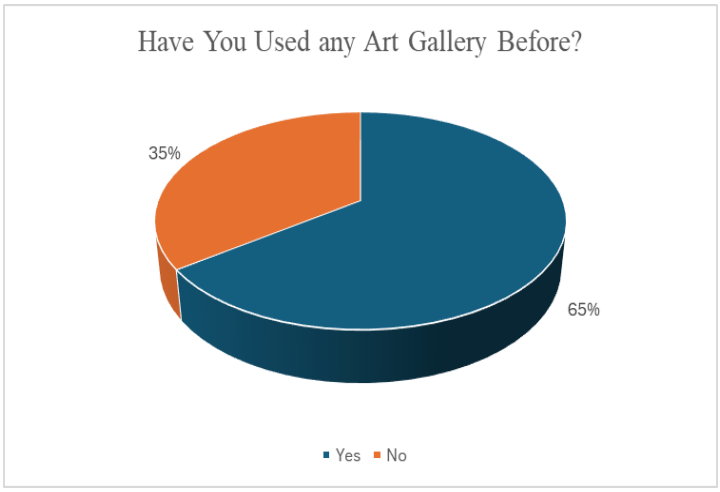

Figure 12. Respondent's prior use of art gallery.

Figure 12, the percentage of participant's responses to previous usage of any art gallery. It can be seen that 65% of the participants have no prior experience in using any art gallery and 35% of the questionnaire partakers have used an art gallery previously.

Q4: *Does the art gallery allow you to view artistic image exhibitions in 3-dimension?*

Here, the respondents were asked if the website allows them to view artistic images in a 3-dimensional view. The response distribution is shown in Figure. 13.

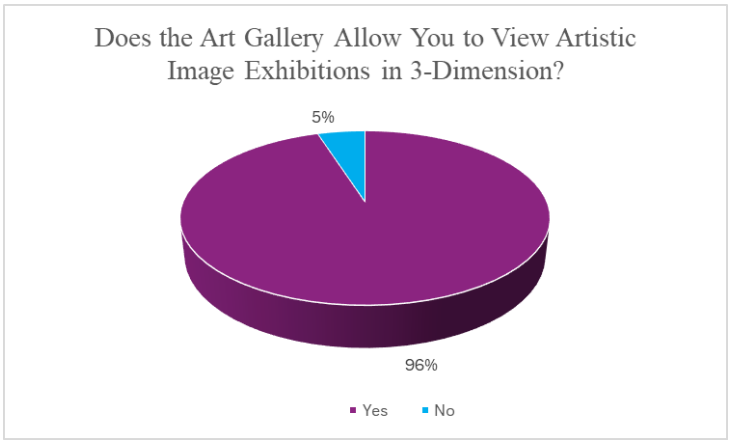

Figure 13. Respondents' response to viewing images in 3-dimensional.

From Figure 13, it is observed that 95% of the respondents confirmed that they were able to view the images in 3-D while 5% of the participants were unable to view images in 3D.

Q5: *How is your interaction with the artistic gallery?*

From the questionnaire, the participants were asked about their experience in interacting with the new web-based artistic gallery. The responses of the participants to the questionnaires are shown in Figure. 14.

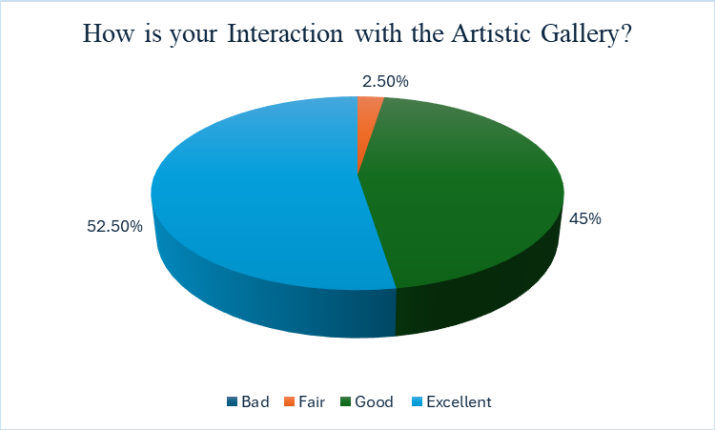

Figure 14. Participants' satisfaction rate in using the art gallery applications.

Figure 14 shows that over half of the participants (61.7) said their interaction with the gallery was excellent, 24% agreed their interaction was very good, 10 said it was good and none said it was fair or poor.

Q6: *Is the art gallery interface user-friendly?*

The respondents were asked about their satisfaction with the art gallery interface and whether it is friendly or not. The participant's answer is shown in Figure 15.

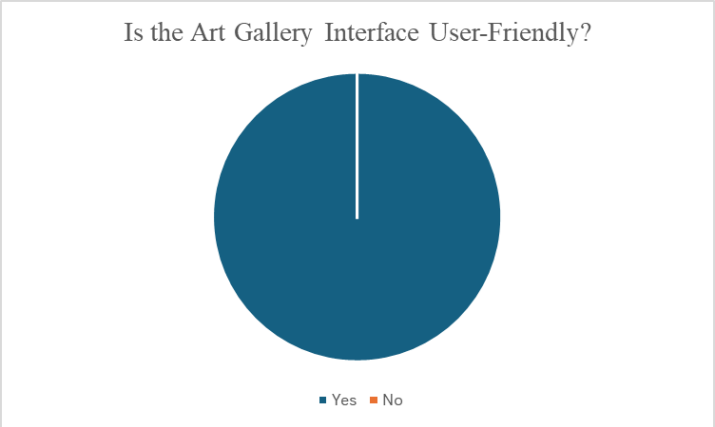

Figure 15. Participants' responses related to the gallery applications interface.

Figure 15 shows that all the participants said the art gallery was user-friendly.

Q7: *Is uploading artistic images an easy task?*

The question regarding the ease of uploading artistic image content was asked and the response of the respondents is shown in Figure 16.

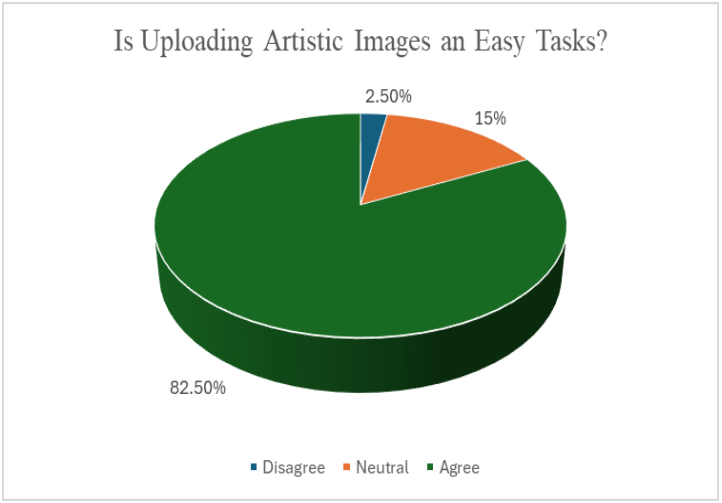

Figure 16. Respondents' ease of use of the applications for uploading tasks.

From Figure 16, 55% of respondents argue that they are easy to use artistic web-based galleries for uploading artistic image content, 30% of the respondents agree that uploading image content is hard to administer and only 5% of participants disagree with the statement.

Q8: *On a scale of 1-10 rate, how easy is the use of the artistic web-based gallery?*

To assess the ease of use of the artistic web-based gallery for image viewing, image uploading, downloading etc., the question regarding how easy or convenient the use of the artistic web-based gallery application was asked, and the response is shown in Figure 17.

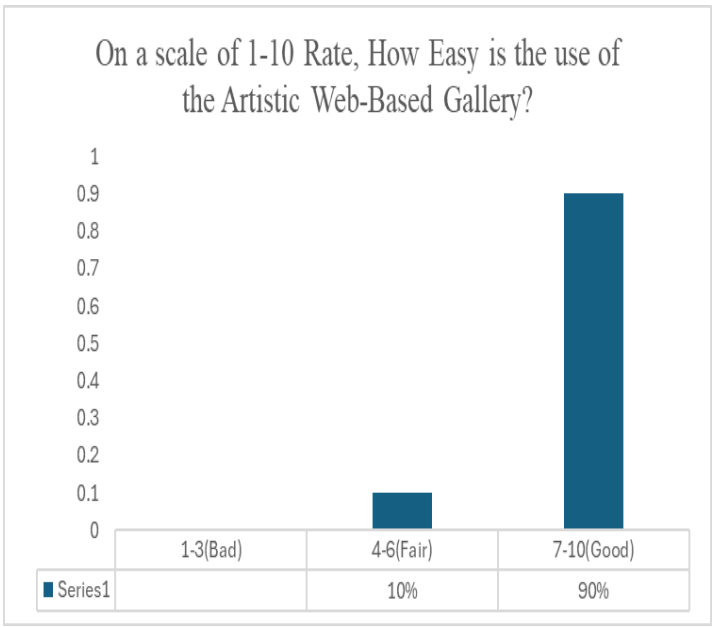

Figure 17. Respondents' rating of artistic gallery applications.

From the 40 questionnaires given, 90% of the respondents rated the art gallery excellent (7-10) and indicated that the platform was easy to use, 10% (4-6) responded that the platform is good to use. Also, no one indicated that the platform is difficult to use. This result indicates that the use of the platform was easy and convenient for users to view and interact with different artistic images.

E. Conclusion

Virtual exhibitions have been considered to have the potential to present artistic work in digital form to people regardless of distance and time. This study has presented the development of an artistic web-based gallery inclined on GANs to address the issue of the artistic web which suffers from being too passive and lacks in-depth interactivity to virtually keep people meaningfully engaged with an exhibition which consequently leads to user lack of satisfaction. Different experiments were conducted using the proposed VAE-3DSDFCycleGAN on the Coco African Mask art dataset. The result of the generated images of the proposed method was compared with other baseline methods and the results show that the proposed method generated images like the original image compared to other GANs methods used in the implementation for comparison.

The web-based application was developed to allow users to view the artistic work in different viewing directions and appreciate the quality of the artwork. A questionnaire is administered to the user and the feedback of their evaluation is analyzed as shown in Figures. 10-17. The participants' feedback shows that users were satisfied with using the virtual art gallery compared to being physically present at the museum. Also, the participants were satisfied with the interface of the application and the three-dimensional view of the art images generated.

Research on artistic image generation using GANs and the development of its applications is still in its infancy stage. However, preliminary pieces of evidence have been obtained in this study that interactive art web-based gallery applications on GANs have potential applications in the creativity of artistic works. In future, further development related to information media design and the development of a wider range of art content could be investigated.

F. Acknowledgement

The authors would like to thank the Department of Computer Systems Engineering for their financial support.

G. References

[1] R. A. Widjono, "Analysis of User Experience in Virtual Art Exhibition During Pandemic," *International Conference of Innovation in Media and Visual Design (IMDES 2020),* vol. 502, pp. 93-101, 2020.

[2] N. Antara and S. SEN, "The Impact of COVID-19 on the Museums and the Way Forward for Resilience," *Journal Of International Museum Education,* vol. 2, no. 1, pp. 54-61, 2020.

[3] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Computers & Education,* vol. 147, no. 103778, 2020.

[4] L. Cheng, J. Xu, and Y. Pan, "Investigating User Experience of VR Art Exhibitions: The Impact of Immersion, Satisfaction, and Expectation Confirmation," *Informatics,* vol. 11, no. 2, p. 30, 2024. [Online]. Available: https://www.mdpi.com/2227-9709/11/2/30.

[5] V. Zakaryan, "Going Virtual: Why Art Gallery Business Needs Online Exhibitions," ed, 2021.

[6] M. Wang, "Application of Multimedia Arts in the Exhibition Industry in the Computer Era," *Advanced in Control Engineering and Information Science,* vol. 15, pp. 3164 – 3168, 2011, doi: 10.1016/j.proeng.2011.08.594.

[7] R. H. Sayeed, Toby, " State of the Art Non-Photorealistic Rendering (NPR) Techniques," *Theory and Practice of Computer Graphics* 2021.

[8] G. Coiffier, P. Renard, and S. Lefebvre, "3D Geological Image Synthesis From 2D Examples Using Generative Adversarial Networks," *Frontier Water,* vol. 2, 2020, doi: https://doi.org/10.3389/frwa.2020.560598.

[9] H. A. Alhaija, A. Dirik, A. Knorig, S. Fidler, and M. Shugrina, "XDGAN: Multi-Modal 3D Shape Generation in 2D Space," *NNN,* pp. 1-14, 2022, doi: doi:10.5244/C.36.NNN.

[10] E. R. Chan *et al.*, "Efficient Geometry-aware 3D Generative Adversarial Networks," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 16102-16112, 2022, doi: 10.1109/CVPR52688.2022.01565.

[11] R. J. Spick, S. Demediuk, and J. A. Walker, "Naive Mesh-to-Mesh Coloured Model Generation using 3D GANs," *In Proceedings of Australasian Computer Science Week (ACSW'20). ACM,* pp. 1-6, 2020, doi: https://doi.org/10.1145/1122445.1122456.

[12] M. Umair, L. N. M. Tissen, and A. P. O. Vermeeren, "3D Reproduction of Cultural Heritage Artifacts Evaluation of Significance and Experience," *Studies in Digital Heritage,* vol. 5, no. 1, pp. 1-29, 2019.

[13] Y. Liu, "Evaluating Visitor Experience of DIgital Interpretation and Presentation Technologies at Cultural Heritage Sites: A Case Study of the Old Town, Zooying," *Built Heritage,* vol. 4, no. 14, pp. 1-15, 2020.

[14] M. Roussou and A. Katifori, "Flow, Staging, Wayfinding, Personalization: Evaluating User Experience with Mobile Museum Narratives," *Multimedia Technologies and Interaction,* vol. 2, no. 32, pp. 1-31, 2018.

[15] J. Li, J.-W. Nie, and J. Ye, "Evaluation of Virtual Tour in an Online Museum: Exhibition of Architecture of the ForbiddenCIty," *PLoS ONE,* vol. 17, no. 1, 2022.

[16] L. Meng, Y. Liu, K. Li, and R. Lyu, "Research on a User-Centered Evaluation Model for Audience Experience and Display Narrative of Digital Museums," *Electronics,* vol. 11, no. 1445, pp. 1-18, 2022.

[17] M. Yu, "Analysis of the Quantitative Impact of Virtual Reality Technology on Visual Communication Art Design," *Applied Artificial Intelligence,* vol. 37, no. 1, 2023.

[18] E. Zidianakis *et al.*, "The Invisible Museum: A User-Centric Platform for Creating Virtual 3D Exhibitions with VR Support," *Electronics* vol. 10, no. 63, pp. 1-38, 2021.

[19] P. D and K. V.D, "Variational Autoencoders for Data Augmentation in Clinical Studies," *Applied Science,* vol. 13, no. 8793, 2023, doi: doi.org/10.3390/app13158793.
[20] K. Liu and G. Qiu, "Lipschitz constrained GANs via boundedness and continuity," *Neural Computing and Applications,* vol. 32, pp. 18271-18283, 2020.
[21] J. Y. Zhu, T. Park, P.Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," *IEEE International Conference on Computer Vision,* pp. 2242-2251, 2017.
[22] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," *arXiv 2020, arXiv:1511.06434,* 2020. [Online]. Available: Available online: https://arxiv.org/abs/1511.06434 (accessed on 29 May 2020).
[23] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," *arXiv.1411,1784[cs.LG],* 2014.
[24] D. Victor, "COCO-AFRICA: A Curation Tool and Dataset of Common Objects in the Context of Africa," *Conference on Neural Information Processing, 2nd Black in AI Workshop,* 2018.
[25] Y. J, K. A, and B. D, "Lr-Gan: Layered Recursive Generative Adversarial Networks for Image Generation," 2017.
[26] A. Karnewar and O. Wang, "MSG-GAN: Multi-Scale Gradients for Generative Adversarial Networks," *arXiv:1903.06048v4 [cs.CV] 12 Jun 2020,* pp. 1-18, 2020.
[27] L. Tran, X. Yin, and X. Liu, "Disentangled Representation Learning Gan for Pose-Invariant Face Recognition," 2017.
[28] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida., "Spectral Normalization for Generative Adversarial Networks," 2018.