

## The AI Artistry: Unleashing the Power of Generative AI in Image Creation

Ranjith Gopalan, PhD<sup>1\*</sup>, Gopikrishna Kalpinagarajarao, MS<sup>2</sup>

<sup>1</sup>Principal consultant, Cognizant, USA

<sup>2</sup>Product Engineer, Cardinal Health, USA

**Citation:** Gopalan R, Kalpinagarajarao G. The AI Artistry: Unleashing the Power of Generative AI in Image Creation. *J Artif Intell Mach Learn & Data Sci* 2024, 2(4), 1471-1480. DOI: doi.org/10.51219/JAIMLD/ranjith-gopalan/333

**Received:** 12 October, 2024; **Accepted:** 30 October, 2024; **Published:** 01 November, 2024

**\*Corresponding author:** Ranjith Gopalan, PhD, Principal consultant, Cognizant, USA, E-mail: ranjith.gopalan@gmail.com

**Copyright:** © 2024 Gopalan R, et al., This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### ABSTRACT

Recent advancements in artificial intelligence, particularly in computer vision and deep learning, have led to the emergence of numerous generative AI platforms that have the ability to create high-quality artistic media, including visual art, concept art and digital illustrations. These generative AI tools have the potential to fundamentally alter the creative processes by which artists and designers formulate ideas and bring them into fruition. However, the application of these AI-generated image tools in the field of graphic design has not been extensively explored.

The realm of multimedia is being revolutionized by the advent of Generative AI, which is reshaping creative workflows, simplifying content creation and unlocking new avenues for multimedia storytelling. This technology holds the promise of producing enthralling visuals for documentaries from mere historical texts or crafting personalized, interactive multimedia experiences that cater to individual preferences. The influence of generative imaging is palpable, from the high-resolution cameras in our smartphones to the immersive experiences crafted by cutting-edge technologies. This study ventures into the dynamic domain of Generative AI, spotlighting its groundbreaking role in image generation. It delves into the evolution of traditional imaging in consumer electronics and the impetus behind AI integration, which has significantly expanded application capabilities. The research meticulously evaluates the latest breakthroughs in leading-edge technologies such as DALL-E 2, Craiyon, Stable Diffusion, Imagen, Jasper, Night Cafe and Deep AI, gauging their performance based on image quality, variety and efficiency. It also contemplates the constraints and moral dilemmas introduced by this fusion, seeking a harmony between human ingenuity and AI-driven automation. This study stems from its thorough analysis and juxtaposition of these AI platforms, yielding perceptive findings that illuminate their merits and potential enhancements. The conclusion accentuates the transformative power of Generative AI in the sphere of image generation, setting the stage for subsequent research and innovation to further advance and polish these technologies. This paper acts as an essential resource for grasping the present state and future directions of AI-enabled image creation, providing a window into the burgeoning collaboration between human artistry and machine intelligence.

**Keywords:** Gen AI models, Gen AI tools, Variational Autoencoders, Diffusion models, Stable Diffusion, AIML, Image prompt, Medical Imaging.

### 1. Introduction

In the realm of imaging, this technology has unlocked a myriad of opportunities for creative professionals, medical

experts and researchers alike. It is transforming the imaging landscape by empowering creators, customizing user experiences and enhancing accessibility. Generative AI streamlines tasks, produces diverse content variations and

crafts entirely new visuals, allowing creators to concentrate on storytelling and design. It customizes images to match user preferences and promotes inclusivity by generating captions, translating languages, and creating image descriptions. These advancements represent a significant leap forward in how we create and experience visual content. Generative AI in imaging has profoundly impacted various aspects of our lives, heralding a new era of visual content creation and manipulation. Its influence spans multiple domains, from art and entertainment to healthcare and beyond.

The paper presents a thorough examination of Generative AI models' impact on imaging. Key contributions include:

- (1) A detailed analysis of Variational Autoencoders (VAEs),
- (2) Transformers,
- (3) Autoregressive models
- (4) Diffusion models and
- (5) Generative Adversarial

Networks (GANs) and Generative AI Tools Stable Diffusion, Craiyon, Artbreeder, NightCafe, Jasper, BigGAN, StyleGAN, Pix2Pix, Midjourney, IMAGEN, DeepDream, Deep AI and DALL-E 2.

## 2. Methodology

The advent of generative adversarial networks and other generative AI models have enabled the creation of plausible, high-quality images that can serve as a starting point for creative expression. These tools can augment the creativity of human artists and designers by generating novel ideas and concepts, allowing them to explore a wider range of possibilities and push the boundaries of their work. As generative AI becomes more sophisticated, it is poised to play an increasingly important role in the creative industries, potentially transforming the ways in which art and design are conceived and produced.

Generative AI models, such as DALL-E 2, Craiyon, Stable Diffusion and Imagen, have demonstrated their ability to generate diverse and visually appealing images based on textual prompts.

AI's capacity to rejuvenate and colorize ancient photographs is a boon for photographers and historians, making history leap off the page with striking clarity. In the realm of healthcare, generative AI is revolutionizing the field by producing synthetic medical imagery to train diagnostic tools, enhancing the quality of patient treatment significantly<sup>1</sup>.

The methodology includes analyzing different machine learning models for data generation, especially in generative modeling, such as Variational Autoencoders (VAEs), Transformers, Autoregressive models, Diffusion models and Generative Adversarial Networks (GANs). This review aims to understand the unique characteristics, strengths and limitations of each approach, as well as their suitability for various multimedia content generation tasks.

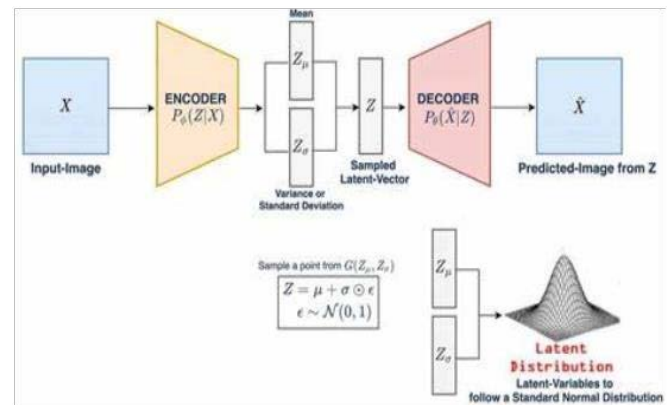
It also includes a comprehensive comparison of AI tools and models designed for image generation or manipulation, such as IIMAGEN, Deep Dream, Deep AI, NightCafe, DALL-E 2, Stable Diffusion, Jasper, Artbreeder, BigGAN, StyleGAN and Pix2Pix. The analysis focuses on the image quality, diversity and efficiency of these models, as well as their potential impact on creative industries and other applications.

## 3. Comparative Analysis of Machine Learning Models for Generative Data Modeling

The development of powerful generative models has been a significant driver in the advancement of Generative AI. These models, such as Variational Autoencoders, Transformers, Autoregressive models, Diffusion models and Generative Adversarial Networks, have demonstrated remarkable capabilities in generating diverse and high-quality multimedia content.

### 3.1. Variational Autoencoders

Variational Autoencoders (VAEs) are a type of generative model that combines the principles of autoencoders and variational inference. They are used to generate new data samples that are similar to the training data. VAEs consist of two main components (**Figure 1**): an encoder which maps the input data to a latent space and a decoder which reconstructs the input from the latent representation. VAEs can generate high-quality images, but they may struggle with capturing complex, fine-grained details in the output<sup>2</sup>.



**Figure 1:** Image flow (Encoder and decoder flow) of VAEs.

### Key Concepts

- » **Latent Space:** A lower-dimensional space where the input data is represented.
- » **Reparameterization Trick:** A technique used to allow backpropagation through the stochastic sampling process.
- » **Loss Function:** Combines reconstruction loss (how well the output matches the input) and KL divergence (how well the learned distribution matches the prior distribution).

VAEs can be used to generate complex images by learning the underlying distribution of the training images and then sampling from this distribution to create new images. This is particularly useful in applications like image synthesis, data augmentation and anomaly detection.

Please see below the program screen shot run for generating/optimizing model for image generation.

Program contains following process (**Figure 2**).

### 3.2. Transformers

The Transformer model is a type of neural network architecture that has revolutionized natural language processing (NLP) and other fields. Transformers use self-attention mechanisms to capture long-range dependencies in the input data, allowing them to generate coherent and contextual output.

**1. Additional Layers in Encoder and Decoder:**

- The encoder now has three hidden layers with dimensions `hidden_dim1`, `hidden_dim2`, and `hidden_dim3`.
- The decoder mirrors the encoder with three hidden layers.

**2. Learning Rate Scheduler:**

- A learning rate scheduler is added to adjust the learning rate during training. The learning rate is reduced by a factor of 0.5 every 10 epochs.

**3. Measure Total Execution Time:**

- The total execution time is measured using `time.time()` before and after the training loop.

**4. Hyperparameter Tuning:**

- Batch Size:** Increased to 128 for faster training.
- Learning Rate:** Adjusted to `1e-3` for potentially better convergence.
- Number of Epochs:** Increased to 50 for more training iterations.

Figure 2

```

from AI in Image Creation > Variational Autoencoders > variational_autoencoders.py > VAE > forward
import torch
import torch.nn as nn
import torch.optim as optim
from torchvision import datasets, transforms
from torch.utils.data import DataLoader
import matplotlib.pyplot as plt
import time

# Define the VAE Model with additional layers
class VAE(nn.Module):
    def __init__(self, input_dim, hidden_dim1, hidden_dim2, hidden_dim3, latent_dim):
        super(VAE, self).__init__()
        self.encoder = nn.Sequential(
            nn.Linear(input_dim, hidden_dim1),
            nn.ReLU(),
            nn.Linear(hidden_dim1, hidden_dim2),
            nn.ReLU(),
            nn.Linear(hidden_dim2, hidden_dim3),
            nn.ReLU(),
            nn.Linear(hidden_dim3, latent_dim * 2) # Output both mean and log-variance
        )
        self.decoder = nn.Sequential(
            nn.Linear(latent_dim, hidden_dim3),
            nn.ReLU(),
            nn.Linear(hidden_dim3, hidden_dim2),
            nn.ReLU(),
            nn.Linear(hidden_dim2, hidden_dim1),
            nn.ReLU(),
            nn.Linear(hidden_dim1, input_dim),
            nn.Sigmoid()
        )

    def reparameterize(self, mu, logvar):
        std = torch.exp(0.5 * logvar)
        eps = torch.randn_like(std)
        return mu + eps * std

    def forward(self, x):
        x = self.encoder(x)
        mu, logvar = x.chunk(2, dim=-1)
        z = self.reparameterize(mu, logvar)
        return self.decoder(z), mu, logvar

# Define the loss function
def loss_function(recon_x, mu, logvar):
    BCE = nn.functional.binary_cross_entropy(recon_x, reduction='sum')
    KLD = 0.5 * torch.sum(1 + logvar - mu.pow(2) - logvar.exp())
    return BCE + KLD

# Prepare the Data
transform = transforms.ToTensor()
train_dataset = datasets.MNIST(root='./data', train=True, transform=transform, download=True)
train_loader = DataLoader(train_dataset, batch_size=128, shuffle=True) # Adjusted batch size

# Initialize the Model, Optimizer, and Device
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
model = VAE(input_dim=784, hidden_dim1=512, hidden_dim2=256, hidden_dim3=128, latent_dim=25).to(device) # Adjusted hidden dimensions
optimizer = optim.Adam(model.parameters()) # Adjusted learning rate
scheduler = optim.lr_scheduler.StepLR(optimizer, step_size=10, gamma=0.5) # Learning rate scheduler

# Measure total execution time
start_time = time.time()

# Train the Model
model.train()
for epoch in range(50): # Increased number of epochs
    train_loss = 0
    for batch_idx, (data, _) in enumerate(train_loader):
        data = data.view(-1, 784).to(device)
        optimizer.zero_grad()
        recon_batch, mu, logvar = model(data)
        loss = loss_function(recon_batch, data, mu, logvar)
        loss.backward()
        train_loss += loss.item()
    optimizer.step()
    scheduler.step() # Adjust the learning rate
    print(f'Epoch {epoch + 1}, Loss: {train_loss / len(train_loader.dataset)}')

# Measure total execution time
end_time = time.time()
print(f'Total execution time: {end_time - start_time} seconds')

```

Figure 3: VAEs Program

```

# Generate New Images
model.eval()
with torch.no_grad():
    z = torch.randn(64, 20).to(device) # Sample from the latent space
    sample = model.decoder(z).cpu() # Decode the latent space to generate images
    sample = sample.view(64, 1, 28, 28) # Reshape the generated images

# Visualize the generated samples
fig, axes = plt.subplots(8, 8, figsize=(8, 8))
for i, ax in enumerate(axes.flatten()):
    ax.imshow(sample[i].squeeze(), cmap='gray')
    ax.axes.set_title('')
plt.show()

```

Program run for 50 epochs with loss function data reduced to 104.60 (Figure 4).

```

PS C:\Git\multiplePythonUtilities> & C:/Users/ranji/AppData/Local/Programs/Python/Python311/python.exe "c:/Git/multiplePythonU
Epoch 39, Loss: 105.16439806315104
Epoch 40, Loss: 105.12601534830729
Epoch 41, Loss: 104.84185893554688
Epoch 42, Loss: 104.82359287109375
Epoch 43, Loss: 104.79058326822917
Epoch 44, Loss: 104.7607018391927
Epoch 45, Loss: 104.7131174641927
Epoch 46, Loss: 104.73386534830729
Epoch 47, Loss: 104.67914702148437
Epoch 48, Loss: 104.66407495117187
Epoch 49, Loss: 104.62502623697917
Epoch 50, Loss: 104.60633662109375
Total execution time: 548.26 seconds

```

Figure 4: Execution Result.



Figure 5

### Key Components of the Transformer Model

- » **Self-Attention Mechanism:** This is the core innovation of the Transformer. It allows the model to weigh the importance of different words in a sentence when encoding a particular word. This mechanism helps the model understand context more effectively than previous models like RNNs or LSTMs.
- » **Encoder-Decoder Structure:** The Transformer consists of an encoder and a decoder, each made up of multiple layers. The encoder processes the input sequence and generates a set of encodings, which are then used by the decoder to produce the output sequence.
- » **Multi-Head Attention:** Instead of having a single attention mechanism, Transformers use multiple attention heads.

This allows the model to focus on different parts of the input sequence simultaneously, capturing various aspects of the context.

- » **Feed-Forward Neural Networks:** Each layer in both the encoder and decoder contains a fully connected feed-forward network, which processes the attention outputs.
- » **Positional Encoding:** Since Transformers do not have a built-in sense of the order of words (unlike RNNs), they use positional encodings to inject information about the position of each word in the sequence.

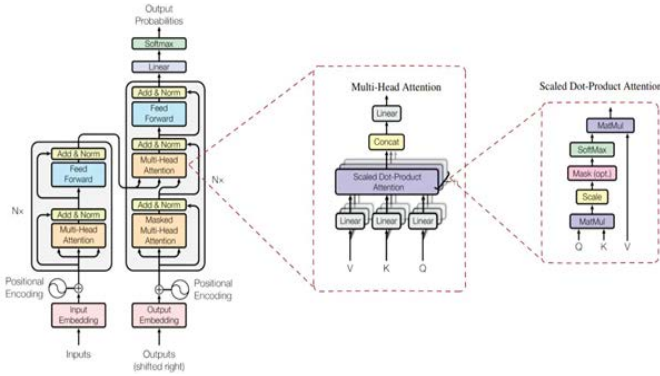


Figure 6: explains input embedding and out embedding flow.

### 3.3. Autoregressive models

Autoregressive models are a class of generative models that generate data sequentially, where each new sample is predicted based on the previously generated samples. These models, such as PixelRNN and PixelCNN, can produce high-quality images by learning the underlying distribution of the training data and then sampling from this distribution to create new samples.

Autoregressive models are versatile tools used across various fields for predictive purposes. Professionals employ these models in numerous ways, such as forecasting future stock prices, estimating annual earthquake occurrences, analyzing protein sequences in genetics, projecting patient health outcomes, tracking symptom progression over time, monitoring the spread of diseases in animals and predicting patterns in circadian rhythms<sup>3-5</sup>.

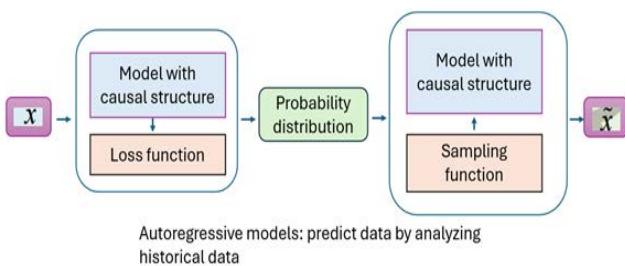


Figure 7: Explains autoregressive model flow

### 3.4. Diffusion models

Diffusion models are a category of generative models that are trained to create data by inverting a diffusion process. This process incrementally introduces noise into the data, which the model is then trained to remove, allowing it to generate new data samples. Notably successful in producing high-quality images, diffusion models are generally more stable during training as they do not depend on adversarial methods. Their versatility extends to various data types, such as images, audio and text, and they can be tailored to diverse domains, grounded in well-established principles of statistical physics and probability theory.

### How Diffusion Models Work

- **Forward Diffusion Process:** In the forward process, a clean image is gradually corrupted by adding Gaussian noise over several time steps. This process is designed to be reversible.
- **Reverse Diffusion Process:** In the reverse process, the model learns to denoise the image step-by-step, starting from pure noise and gradually refining it to produce a clear image. The model is trained to predict the noise added at each step, allowing it to reverse the corruption process.

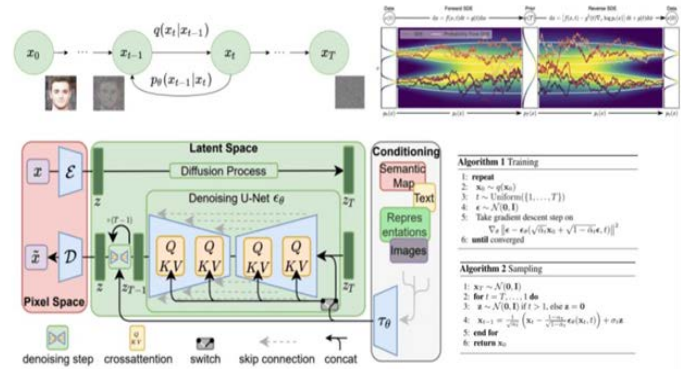


Figure 8: explains diffusion model and workflow.

Please see the execution result of diffusion model having 8 epoch cycle and result in loss function value 0.00202 (Figure 8).

```
pythonUtilities/Generative AI in Image Creation/diffusion_models/diffusion_model.py
Files already downloaded and verified
Files already downloaded and verified
Epoch 1, Loss: 0.005840946175009012
Epoch 2, Loss: 0.0026167775829136373
Epoch 3, Loss: 0.002355782651975751
Epoch 4, Loss: 0.002215686437338591
Epoch 5, Loss: 0.0021370093840360643
Epoch 6, Loss: 0.002079256727322936
Epoch 7, Loss: 0.002063074046820402
Epoch 8, Loss: 0.0020203044574707745
```

Figure 9: explains the Diffusion model execution result.

The basic idea behind diffusion models is rather simple. They take the input image  $x_0 \times x_0$  and gradually add Gaussian noise to it through a series of  $T$  steps. We will call this the forward process. Notably, this is unrelated to the forward pass of a neural network. If you'd like, this part is necessary to generate the targets for our neural network (the image after applying  $t < T < T$  noise steps).

Afterward, a neural network is trained to recover the original data by reversing the noising process. By being able to model the reverse process, we can generate new data. This is the so-called reverse diffusion process or, in general, the sampling process of a generative model.

### 3.5. Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs) are a class of machine learning frameworks designed by Ian Goodfellow and his colleagues in 2014. GANs consist of two neural networks, the generator and the discriminator, which are trained simultaneously through adversarial processes. The generator creates data that mimics real data, while the discriminator evaluates the authenticity of the generated data.

Today Generative AI has profoundly transformed the field of imaging. It leverages advanced machine learning techniques to create, enhance, and manipulate images in ways that were once considered the realm of science fiction. This transformative

technology is centered around the development of algorithms and models that can autonomously generate images, modify existing ones, or even fill in missing information within images.

```

100 real_loss = adversarial_loss(discriminator(real_imgs), valid)
101 fake_loss = adversarial_loss(discriminator(gen_imgs.detach()), fake)
102 d_loss = (real_loss + fake_loss) / 2
103 d_loss.backward()
104 optimizer_D.step()
105
106 print("[epoch: %d / epochs: %d] Generator Loss: %d, Discriminator Loss: %d,
107       Item: %d" % (epoch, epochs, g_loss.item(), d_loss.item(), item))
108
109 # Save generated images every 10 epochs
110 # epoch % 10 == 0
111 gen_imgs = gen_imgs.view(gen_imgs.size()[0], *img_shape)
112 save_image(gen_imgs.data[:25], "%images/epoch.png", nrow=5, normalize=True)

```

```

Epoch 0/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 1/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 2/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 3/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 4/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 5/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 6/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 7/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 8/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Epoch 9/10 | Generator Loss: 100.0 | Discriminator Loss: 0.0028105804813094437
Total execution time: 715.56 seconds
PS C:\vscode\multiplespythonutils>

```

Figure 10: explains the GANs execution result.

### 4. Comprehensive Comparison of Gen AI Tools for Image Generation and Manipulation

Following an overview of fundamental Generative AI models and their impact, we now focus on specific technologies in image synthesis from text descriptions. Generative AI models represent a diverse range of technical approaches and applications in the field of text-to-image generation. From state-of-the-art models like DALL-E 2 and Imagen to accessible tools like NightCafe and Stable Diffusion, each model offers unique strengths and capabilities that cater to unique needs and use cases. Please see the strength and application of these Gen AI tools in the following categories.

#### 4.1. State-of-the-Art

These models represent the cutting edge of text-to-image generation, pushing the boundaries of what's possible.

##### 1. DALL-E 2

- **Strengths:** High-quality, diverse image generation with detailed and coherent outputs.
- **Applications:** Creative content generation, advertising, design and research.

##### 2. Imagen

- **Strengths:** Produces highly realistic images with accurate semantic content.
- **Applications:** Research, creative industries and content creation.

#### 4.2. Technical Diversity

These models showcase a range of technical approaches, providing a comprehensive understanding of the different techniques driving the field.

##### 1. Deep AI

- **Strengths:** Grounded in GANs, offering a different technical approach compared to transformer-based models.
- **Applications:** Artistic creation, research, educational tools, and creative projects.

##### 2. BigGAN

- **Strengths:** High-quality, high-resolution image generation

with diverse outputs.

- **Applications:** Research, high-quality image synthesis, creative industries and academic studies.

#### 3. StyleGAN

- **Strengths:** High-quality image generation with detailed control over style and features.
- **Applications:** Art creation, design, research and commercial projects.

#### 4. Pix2Pix

- **Strengths:** Versatile image-to-image translation with practical applications.
- **Applications:** Image editing, artistic creation, research and educational tools.

#### 5. DeepDream

- **Strengths:** Unique artistic effects and visualizations.
- **Applications:** Art creation, visual effects, educational tools and creative experimentation<sup>6</sup>

#### 4.3. Accessibility

These models include both open-access options and research-focused models, providing insights into both cutting-edge advancements and user-friendly tools.

##### 1. NightCafe

- **Strengths:** User-friendly interface with multiple model options for diverse artistic styles.
- **Applications:** Creative projects, personal use, educational purposes and community engagement.

##### Stable Diffusion

- **Strengths:** High-quality outputs with a focus on accessibility and community contributions.
- **Applications:** Creative content, research, community-driven projects and opensource development.

#### 4.4. Strengths and Applications

These models have distinct strengths and are known for their specific applications.

##### 1. Jasper

- **Strengths:** High-quality text generation that complements image generation tasks.
- **Applications:** Content creation, marketing, automated writing and customer service.

##### 2. Artbreeder

- **Strengths:** Interactive and collaborative platform for generating and evolving images.
- **Applications:** Art creation, character design, collaborative projects and personal use.

This grouping provides a clear understanding of the generative AI models based on their state-of-the-art status, technical diversity, accessibility, strengths and applications. Each model offers unique capabilities that cater to different needs and use cases in the field of text-to-image generation<sup>7-9</sup>.

### 5. Results

Here we are talking about comprehensive comparison (**Figure 10**) from the various execution result on mentioned Gen AI models in terms of parameters like Technical Aspects, Performance and Robustness, Customization and Control, Ethical and Accessibility and User Experience and Handling of these technologies based on the following parameters.

Model	Technical Aspects	Performance and Robustness	Customization and Control	Ethical and Accessibility	User Experience and Handling
<b>Stable Diffusion</b>	<ul style="list-style-type: none"> <li>• Uses diffusion models</li> <li>• Open-source</li> <li>• Requires significant computational resources</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality image generation</li> <li>• Robust against noise</li> <li>• Good for various styles</li> </ul>	<ul style="list-style-type: none"> <li>• Highly customizable</li> <li>• Fine-tuning possible</li> <li>• Control over diffusion steps</li> </ul>	<ul style="list-style-type: none"> <li>• Open-source, accessible</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• User-friendly interfaces</li> <li>• Requires setup for local use</li> </ul>
<b>Craiyon</b>	<ul style="list-style-type: none"> <li>• Transformer-based</li> <li>• Open-source</li> <li>• Lightweight</li> </ul>	<ul style="list-style-type: none"> <li>• Moderate quality</li> <li>• Limited robustness</li> </ul>	<ul style="list-style-type: none"> <li>• Limited customization</li> <li>• Predefined prompts</li> </ul>	<ul style="list-style-type: none"> <li>• Free and accessible</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• Easy to use</li> <li>• Web-based interface</li> </ul>
<b>Artbreeder</b>	<ul style="list-style-type: none"> <li>• GAN-based- Web-based</li> <li>• Uses collaborative filtering</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality artistic images</li> <li>• Robust for creative tasks</li> </ul>	<ul style="list-style-type: none"> <li>• Customizable through sliders</li> <li>• Collaborative control</li> </ul>	<ul style="list-style-type: none"> <li>• Accessible</li> <li>• Ethical concerns with content ownership</li> </ul>	<ul style="list-style-type: none"> <li>• User-friendly</li> <li>• Intuitive</li> </ul>
<b>NightCafe</b>	<ul style="list-style-type: none"> <li>• Uses various models (VQGAN, CLIP, etc.)</li> <li>• Web-based</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality artistic images</li> <li>• Robust for creative tasks</li> </ul>	<ul style="list-style-type: none"> <li>• Customizable through prompts and settings</li> </ul>	<ul style="list-style-type: none"> <li>• Accessible</li> <li>• Ethical concerns with content ownership</li> </ul>	<ul style="list-style-type: none"> <li>• User-friendly</li> <li>• Intuitive interface</li> </ul>
<b>Jasper</b>	<ul style="list-style-type: none"> <li>• Transformer-based</li> <li>• Web-based</li> <li>• Focus on text-to-image</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality text-to-image</li> <li>• Robust for various prompts</li> </ul>	<ul style="list-style-type: none"> <li>• Customizable through text prompts</li> </ul>	<ul style="list-style-type: none"> <li>• Accessible</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• User-friendly</li> <li>• Intuitive int</li> </ul>
<b>BigGAN</b>	<ul style="list-style-type: none"> <li>• GAN-based</li> <li>• Requires significant computational resources</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality, high-resolution images</li> <li>• Robust for various styles</li> </ul>	<ul style="list-style-type: none"> <li>• Customizable through latent space manipulation</li> </ul>	<ul style="list-style-type: none"> <li>• Limited accessibility</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• Requires technical expertise</li> <li>• Not user-friendly for non-experts</li> </ul>
<b>StyleGAN</b>	<ul style="list-style-type: none"> <li>• GAN-based</li> <li>• Requires significant computational resources</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality, high-resolution images</li> <li>• Robust for various styles</li> </ul>	<ul style="list-style-type: none"> <li>• Highly customizable</li> <li>• Fine-tuning possible</li> </ul>	<ul style="list-style-type: none"> <li>• Limited accessibility</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• Requires technical expertise</li> <li>• Not user-friendly for non-experts</li> </ul>
<b>Pix2Pix</b>	<ul style="list-style-type: none"> <li>• GAN-based</li> <li>• Requires paired training data</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality image-to-image translation</li> <li>• Robust for specific tasks</li> </ul>	<ul style="list-style-type: none"> <li>• Customizable through training data</li> </ul>	<ul style="list-style-type: none"> <li>• Limited accessibility</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• Requires technical expertise</li> <li>• Not user-friendly for non-experts</li> </ul>
<b>MidJourney</b>	<ul style="list-style-type: none"> <li>• Uses proprietary models</li> <li>• Web-based</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality artistic images</li> <li>• Robust for creative tasks</li> </ul>	<ul style="list-style-type: none"> <li>• Customizable through prompts and settings</li> </ul>	<ul style="list-style-type: none"> <li>• Accessible</li> <li>• Ethical concerns with content ownership</li> </ul>	<ul style="list-style-type: none"> <li>• User-friendly</li> <li>• Intuitive interface</li> </ul>
<b>Imagen</b>	<ul style="list-style-type: none"> <li>• Transformer-based- Proprietary</li> <li>• Focus on text-to-image</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality text-to-image</li> <li>• Robust for various prompts</li> </ul>	<ul style="list-style-type: none"> <li>• Customizable through text prompts</li> </ul>	<ul style="list-style-type: none"> <li>• Limited accessibility</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• User-friendly</li> <li>• Intuitive interface</li> </ul>
<b>DeepDream</b>	<ul style="list-style-type: none"> <li>• CNN-based- Open-source</li> <li>• Artistic transformations</li> </ul>	<ul style="list-style-type: none"> <li>• High-quality artistic transformations</li> <li>• Robust for creative tasks</li> </ul>	<ul style="list-style-type: none"> <li>• Customizable through layer selection and parameters</li> </ul>	<ul style="list-style-type: none"> <li>• Open-source, accessible</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• Requires setup for local use</li> <li>• User-friendly for technical users</li> </ul>
<b>Deep AI</b>	<ul style="list-style-type: none"> <li>• Uses various models</li> <li>• Web-based</li> </ul>	<ul style="list-style-type: none"> <li>• Moderate quality</li> <li>• Robust for various styles</li> </ul>	<ul style="list-style-type: none"> <li>• Limited customization</li> <li>• Predefined prompts</li> </ul>	<ul style="list-style-type: none"> <li>• Accessible</li> <li>• Ethical concerns with misuse</li> </ul>	<ul style="list-style-type: none"> <li>• Easy to use</li> <li>• Web-based interface</li> </ul>

**Figure 11:** presents a comprehensive comparison of Gen AI models used for image creation.

### 6. Visual analysis and inference

This section is talking about practical comparison of visual images generated.

We undertake a detailed comparative analysis of four distinct models: Stable Diffusion, Craiyon, Artbreeder and NightCafe. Chosen for their broad adoption, varied technological methodologies and distinctive features, our aim is to rigorously assess and compare each model’s performance and artistic prowess. This will be achieved by testing them against six carefully curated and demanding case scenarios, each designed to cover a broad range of visual content. This approach ensures a comprehensive evaluation of the models’ abilities. The assessment criteria will focus on image quality, consistency, artistic expression, and the precision of converting textual prompts into corresponding visual representations.

The six distinct case scenarios (Flying car, Crowd face, Joyful elephants, A robot welding, Sunrise at mountain lake, Cozy rustic kitchen) (**Figure 11**) were chosen for the analysis because they represent a broad spectrum of visual content.

This thorough comparative analysis is designed to illuminate the strengths and weaknesses of each model, as well as their appropriateness for various artistic and practical endeavors. By assessing their capabilities in demanding situations, we provide artists, developers and researchers with the necessary insights to choose an image synthesis technology that best fits their unique creative or functional goals.

Please see the comprehensive comparison of the four models across our six test scenarios. This analysis offers valuable insights into the distinct strengths and limitations of each model, empowering users to select the most suitable technology for their specific needs and objectives, whether they are pursuing creative, practical or research-oriented tasks.

### 7. Flying Car

- **Stable Diffusion:** The images were visually striking and captured the futuristic theme well, with minor issues in rendering some elements like flying cars. The cityscape was well-rendered with a futuristic look.
- **Craiyon:** Good overall quality with detailed cityscapes but occasional distortions in flying cars. The flying cars were present but sometimes appeared slightly distorted or out of proportion.
- **Artbreeder:** The images were aesthetically pleasing and creative, capturing the essence of a futuristic cityscape, but with a more artistic rather than realistic approach. Artistic and visually appealing images with well-integrated but stylized flying cars.
- **Nightcafe:** Nightcafe exhibited the poorest performance among the models. The cityscape lacked detail and coherence. The flying cars were either missing or poorly rendered, often blending into the background.












Idea	Prompt	Stable Diffusion	Craiyon	Artbreeder	NightCafe
Flying car	A futuristic cityscape with flying cars				
Crowd face	Generate a realistic image featuring 4-5 individual faces of people in a crowded stadium, showcasing their expressions while cheering and shouting with enthusiasm				
Joyful elephants	Generate a vivid and realistic image of a lush jungle landscape teeming with joyful elephants and vibrant, colorful animals, set amidst a dense forest filled with flowering trees				
A robot welding	A robot inside a dimly lit, cluttered workshop, sparks flying from a welding operation, computer screens displaying schematics, the robot's body covered in welding residue				
Sunrise at mountain lake	A serene mountain lake at sunrise, with mist rising from the water, surrounded by towering pine trees and snow-capped peaks, a small wooden cabin with smoke coming from the chimney, and a family of deer grazing nearby.				
Cozy rustic kitchen	A cozy, rustic kitchen with a roaring fireplace, wooden beams on the ceiling, a table set with freshly baked bread and a pot of stew, a cat curled up on a chair, and sunlight streaming through a window with flower boxes.				

Figure 12: shows a comparison of images generated by AI models, created using specific prompts.

## 8. Crowd Face

- **Stable Diffusion:** Poor performance with lack of detail and coherence in both faces and background. The expressions were not well-captured, often appearing unnatural or missing entirely.
- **Craiyon:** Vibrant and detailed images with realistic expressions and a coherent background, capturing the prompt effectively. The images were visually striking and captured the prompt well, with minor issues in rendering some elements.
- **Artbreeder:** Artistic and visually appealing images with well-captured expressions but more stylized than realistic. The background was present and integrated well with the faces, though it had an artistic rather than realistic look.
- **Nightcafe:** Good overall quality with detailed faces and expressions, though occasional distortions were present. The background of the stadium was present but sometimes lacked detail and coherence.

## 9. Joyful Elephants

- **Stable Diffusion:** Poor performance with lack of detail and coherence in both the landscape and the animals. The elephants and other animals were poorly rendered, often blending into the background or appearing unnatural.
- **Craiyon:** High-quality images with detailed and vivid jungle landscapes, well-rendered elephants, and vibrant animals. The background was detailed with dense foliage and flowering trees, adding to the overall realism.

- **Artbreeder:** Vibrant and detailed images with realistic jungle landscapes, joyful elephants and colorful animals, capturing the prompt effectively. The background was detailed and coherent, with dense foliage and flowering trees adding to the overall realism.
- **NightCafe:** Artistic and visually appealing images with well-integrated but stylized animals and a creative jungle landscape. The elephants and other animals were wellintegrated into the scene, though they sometimes appeared more stylized than realistic.

## 10. A robot Welding

- **Stable Diffusion:** Poor performance with lack of detail and coherence in both the workshop and the robot. The images were far from realistic, with a lack of clear structure and detail in both the workshop and the robot.
- **Craiyon:** High-quality images with detailed and realistic workshop environments, wellrendered robots and welding operations. The images were vivid and realistic, capturing the essence of a robot welding in a cluttered workshop.
- **Artbreeder:** Artistic and visually appealing images with well-integrated but stylized robots and a creative workshop environment. The robot and welding operation were wellintegrated into the scene, though they sometimes appeared more stylized than realistic.
- **Nightcafe:** Vibrant and detailed images with realistic workshop environments, accurate robot and welding operations, capturing the prompt effectively. The robot and welding operation were depicted accurately, with sparks flying and welding residue visible on the robot's body.

### 11. Sunrise at Mountain Lake

- **Stable Diffusion:** Not biased performance with small lack of detail and coherence in both the lake and the surrounding environment. The images were far from realistic, with a lack of clear structure and detail in both the lake and the surrounding environment.
- **Craiyon:** High-quality images with detailed and realistic lake environments, well rendered surrounding elements, and integrated cabin and trees. The images were vivid and realistic, capturing the essence of a serene mountain lake at sunrise.
- **Artbreeder:** Artistic and visually appealing images with well-integrated but stylized elements and a creative lake environment. The images were aesthetically pleasing and creative, capturing the essence of a serene mountain lake at sunrise with a more artistic approach.
- **Nightcafe:** Vibrant and detailed images with realistic lake environments, accurate surrounding elements, and well-integrated cabin and deer, capturing the prompt effectively. The towering pine trees and snow-capped peaks were depicted accurately, adding to the overall realism.

### 12. Cozy Rustic Kitchen

- **Stable Diffusion:** High-quality images with detailed and realistic kitchen environments, well-rendered fireplace, wooden beams, table setting, cat and sunlight. The table set

with freshly baked bread and a pot of stew was detailed and realistic.

- **Craiyon:** The images were vivid and realistic, capturing the essence of a cozy, rustic kitchen. The cat curled near to fireplace and the sunlight streaming through the window with flower boxes were well-integrated into the scene.
- **Artbreeder:** Artistic and visually appealing images with well-integrated but stylized elements and a creative kitchen environment. The cozy, rustic kitchen ambiance was well-captured with realistic lighting.
- **Nightcafe:** Vibrant and detailed images with realistic kitchen environments, accurate fireplace, wooden beams, table setting, cat, and sunlight, capturing the prompt effectively. The table setting with bread and stew was present but had an artistic rather than realistic look<sup>10-15</sup>

### 13. Discussion

We all talk about how generative AI models have the potential to revolutionize the creative industry, what are the general model algorithms and most useful Gen AI solution in the market and compared images created by it. We also need to discuss challenges as well.

The challenges associated (Figure 12) with including Generative AI (GAI) in imaging are multifaceted and require a comprehensive approach. Let us List those and discuss them in detail in the below table.

Aspects	Sub aspects	Solution Approach	Future Work
Data Privacy and Security	<ul style="list-style-type: none"> <li>• Anonymization</li> <li>• Data Encryption</li> <li>• Access Control</li> </ul>	<ul style="list-style-type: none"> <li>• Anonymization Techniques (e.g., Differential privacy)</li> <li>• Encryption Protocols (e.g., AES-256)</li> <li>• Role-Based Access Control (RBAC)</li> </ul>	<ul style="list-style-type: none"> <li>• Advanced Anonymization</li> <li>• Quantum Resistant Encryption</li> <li>• Automated Access Control</li> </ul>
	Data Leakage	Utilize federated learning and secure aggregation techniques to prevent data leakage.	Enhance secure multi-party computation (SMPC) for data privacy.
	Security Risks	Implement model encryption and secure model serving frameworks to safeguard AI models.	Research AI model watermarking techniques for model validation.
Technological Challenges	<ul style="list-style-type: none"> <li>• Model Complexity</li> <li>• Computational Resources</li> <li>• Scalability</li> </ul>	<ul style="list-style-type: none"> <li>• Model Optimization (e.g., pruning, quantization)</li> <li>• Cloud Computing</li> <li>• Distributed Computing (e.g., Apache Spark)</li> </ul>	<ul style="list-style-type: none"> <li>• Efficient Algorithms</li> <li>• Edge Computing</li> <li>• Hybrid Model</li> </ul>
	Resource Intensity	Optimize model architectures with hardware accelerators like TPUs and GPUs.	Investigate low-power and energy efficient model architectures.
	Scalability	Design distributed generative AI frameworks to efficiently scale for real-time deployment.	Explore edge AI deployment for on-device image generation.
Privacy and Consent	<ul style="list-style-type: none"> <li>• User Consent</li> <li>• Data Minimization</li> <li>• Transparency</li> </ul>	<ul style="list-style-type: none"> <li>• Consent Management Systems</li> <li>• Data Minimization Practices</li> <li>• Transparency Reports</li> </ul>	<ul style="list-style-type: none"> <li>• Dynamic Consent</li> <li>• Automated Compliance</li> <li>• User Education</li> </ul>
	Privacy Violations	Apply federated learning techniques to ensure data remains decentralized and private.	Create homomorphic encryption techniques to enable secure data sharing.
Ethical Concerns	<ul style="list-style-type: none"> <li>• Bias and Fairness</li> <li>• Misuse Prevention</li> <li>• Accountability</li> </ul>	<ul style="list-style-type: none"> <li>• Bias Mitigation Techniques</li> <li>• Usage Policies</li> <li>• Audit Trails</li> </ul>	<ul style="list-style-type: none"> <li>• Fairness Metrics</li> <li>• Ethical AI Frameworks</li> <li>• Collaborative Governance</li> </ul>
	Misuse	Create deep neural network-based classifiers to verify content authenticity and detect deepfakes.	Advance deepfake detection using GAN based counter-GANs.
Security Threats	<ul style="list-style-type: none"> <li>• Adversarial Attacks</li> <li>• Model Theft</li> <li>• Data Poisoning</li> </ul>	<ul style="list-style-type: none"> <li>• Adversarial Training</li> <li>• Model Watermarking</li> <li>• Data Validation</li> </ul>	<ul style="list-style-type: none"> <li>• Robustness Research</li> <li>• Secure Model Sharing</li> <li>• Automated Threat Detection</li> </ul>
	Cybersecurity	Utilize differential privacy techniques during AI model training to safeguard against model inversion attacks.	Examine the security of AI models within federated learning environments.
	Authentication Issues	Apply anti-spoofing methods, such as liveness detection, in facial recognition systems.	Investigate AI-resistant methods for biometric authentication.
Legal and Regulatory Challenges	<ul style="list-style-type: none"> <li>• Compliance</li> <li>• Intellectual Property</li> <li>• Liability</li> </ul>	<ul style="list-style-type: none"> <li>• Legal Frameworks</li> <li>• Compliance Tools</li> <li>• IP Management</li> </ul>	<ul style="list-style-type: none"> <li>• Regulatory Sandboxes</li> <li>• Global Standards</li> <li>• Liability Models</li> </ul>
	Intellectual Property	Create AI model licensing frameworks and implement smart contracts to verify ownership.	Develop decentralized copyright registries utilizing blockchain technology.
	Regulation	Establish regulatory sandboxes for AI research, development, and deployment.	Collaborate internationally to standardize AI governance protocols.
Verification and Trust	Deepfakes Detection	Create multi-modal deepfake detection systems that leverage both audio and video analysis.	Explore explainable AI methods for detecting deepfakes.

Figure 13: presents challenges in Generative AI (GAI) imaging and probable solutions.



To summarize, the integration of Generative AI into imaging presents a multitude of challenges spanning ethical, privacy, security, legal and technological aspects. Tackling these issues necessitates a comprehensive strategy that includes technical solutions, continuous research, interdisciplinary and international cooperation.

#### 14. Review Comments

This paper is meticulously reviewed by Mr. Shantanu Sengupta, Senior Director - Projects at Cognizant, whose invaluable insights and expertise significantly enriched the quality and depth of this study. His critical assessment of the research methodology, analysis and the presentation of the latest advancements in AI image generation, as well as his thoughtful suggestions for enhancing the discussion on moral dilemmas and potential improvements of AI platforms, have been instrumental in refining the final manuscript.

#### 15. Future Directions

The advancement of artificial intelligence has revolutionized artistic practices, leading to the creation of dynamic AI art communities dedicated to fostering collaboration and innovation. As research into generative image creation deepens, it's crucial to consider the future trajectory of these communities. Trends suggest a growing focus on interdisciplinary collaboration, uniting artists, technologists and academics to push the limits of creativity. This fusion of varied insights is anticipated to produce more nuanced artistic works and spur the development of novel techniques surpassing conventional approaches.

Moreover, the increasing accessibility of AI tools promises to make the creative process more democratic, allowing a wider array of people to engage with AI art communities. Progress in intuitive interfaces and open-source platforms will enable beginners and hobbyists to delve into generative AI, creating a welcoming space for education and idea exchange. This evolution will not only foster a sense of community but also propel innovation through shared creativity. These focused collectives will support concentrated research and trials, significantly contributing to the larger AI art community. By dedicating efforts to specific uses of generative AI, researchers aim to discover new methods and enhance existing ones.

As technology evolves, it is essential to remain proactive in addressing these issues and driving innovation in the field. Future improvements for generative AI in multimedia include:

**AI-powered Story Weaving:** Generative AI tools can assist multimedia developers in crafting their stories by generating various concept art sketches from high-level story descriptions or suggesting visually captivating multimedia that aligns with the narrative's emotional tone. This enhances both agility and creativity in the development process.

**Interpretable AI in Multimedia:** Creating explainable AI techniques tailored for the multimedia domain enables users to comprehend the reasoning behind AI-generated content.

This approach promotes trust and transparency.

**Technological Advancements:** Future efforts should aim to enhance the realism, quality, scalability, robustness, and semantic coherence of multimedia. Investigating the integration of various generative AI models (such as VAEs, GANs, Transformers and Autoregressive Models) can spur innovation. Additionally,

advancing data compression techniques will facilitate the efficient storage and transmission of multimedia content.

**Countering Deepfakes and Disinformation:** Research findings can guide the creation of effective deepfake detection techniques specifically designed for the multimedia field. This enables users to critically assess multimedia content and reduce the spread of misinformation.

**User Behavior Analysis:** Creating generative AI models that dynamically adjust to user preferences in real-time has the potential to transform multimedia content delivery. By offering personalized and customized content, we can ensure that each user gets exactly what they desire, enhancing their engagement and satisfaction.

Image synthesis technology has the potential to transform numerous fields. Here are some key areas it can influence:

**Materials Science:** AI can facilitate virtual testing and optimization of new material properties before they are physically created, speeding up innovation. Additionally, it plays a vital role in simulating material degradation over time, enabling non-destructive testing, preventative maintenance, and enhancements in infrastructure safety.

**Medical Imaging:** Image synthesis can improve early disease detection by analyzing medical scans, like mammograms, to spot subtle anomalies, resulting in better patient outcomes.

Additionally, it enables doctors to simulate treatment effects on a patient's condition, allowing for personalized treatment plans and enhancing surgical results through AI-generated 3D models of organs and tissues.

**Space research:** AI can enhance planetary imaging by analyzing telescope data to eliminate noise and improve images of distant planets, uncovering essential atmospheric details and the potential for life. Furthermore, AI can create simulations of Martian landscapes, supporting mission planning and astronaut training.

#### 16. Conclusion

The intersection of technology and the arts within the multimedia sphere has long fascinated researchers. The advent of Generative AI marks a pivotal shift in this narrative, fundamentally altering the production of multimedia and the collaborative nature of creation. Generative AI has been a game-changer in image generation, expanding the horizons of creativity and partnership. It has not only enhanced human creative capabilities but also made the content creation process more efficient, encouraging new forms of collaboration between AI systems and human artists. The paper thoroughly examines cutting-edge Generative AI models, such as Stable Diffusion, Craiyon, Artbreeder, NightCafe, Jasper, BigGAN, StyleGAN, Pix2Pix, Midjourney, IIMAGEN, DeepDream, Deep AI, and DALL-E 2, evaluating their effectiveness in terms of image quality, diversity, interpretability, and computational efficiency. This integration of AI with image generation delves into text-to-image synthesis, where AI converts textual descriptions into compelling visual art. Nonetheless, this advancement also reveals a complex web of ethical dilemmas and challenges.

The future of AI art communities is poised for significant transformation, driven by interdisciplinary collaboration, increased accessibility, specialization, and ethical considerations.

As researchers navigate this dynamic landscape, their contributions will play a vital role in shaping the trajectory of generative image creation. By embracing these future directions, AI art communities can enhance their impact on the broader artistic ecosystem, redefining the role of technology in creative expression. The ongoing dialogue between researchers, artists and technologists will be essential in unlocking the full potential of AI as a transformative force in the world of art.

## 17. References

1. Oniani D, Hilsman J, Peng Y, Poropatich RK, Pamplin CJC, Legault L, Wang Y, et al. From Military to Healthcare: Adopting and Expanding Ethical Principles for Generative Artificial Intelligence. *Cornell University* 2023;6(1):225.
2. <https://arxiv.org/abs/2306.02781>
3. <https://www.businessperspectives.org/index.php/journals/investment-management-and-financial-innovations/issue-393/forecasting-stock-returns-on-the-amman-stock-exchange-do-neural-networks-outperform-linear-regressions>
4. [https://link.springer.com/chapter/10.1007/978-0-387-21606-5\\_1](https://link.springer.com/chapter/10.1007/978-0-387-21606-5_1)
5. Singh S, Madan TK, Kumar J, Singh AK. *Stock Market Forecasting using Machine Learning: Today and Tomorrow* 2019.
6. Batra K, Kavidayal M. *Animating Still Images*. Cornell University 2022.
7. <https://doi.org/10.1109/access.2024.3397775>
8. Cao Y, Li S, Liu Y, Yan Z, Dai Y, Yu PS, Sun L. A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT. *Cornell University* 2023.
9. <https://doi.org/10.1109/airc57904.2023.10303174>
10. <https://doi.org/10.1155/2022/3302700>
11. <https://iopscience.iop.org/article/10.1088/1755-1315/452/1/012050>
12. Interior renders Tutorial, 2020.
13. <https://dl.acm.org/doi/10.1145/383259.383306>
14. Kulshreshtha P, Lianos K, Pugh B, Jiddi S. *Layout Aware Inpainting for Automated Furniture Removal in Indoor Scenes*. Cornell University 2022.
15. Srinivasan PP, Mildenhall B, Tancik M, Barron JT, Tucker RP, Snavely N. *Lighthouse: Predicting Lighting Volumes for Spatially Coherent Illumination*. Cornell University 2020.
16. <https://arxiv.org/abs/2306.00080>
17. <https://arxiv.org/abs/2007.15129>
18. <https://ci.nii.ac.jp/ncid/BA24808561>
19. <https://arxiv.org/abs/2012.14092>
20. <https://arxiv.org/abs/2306.04542>
21. <https://arxiv.org/abs/2306.01795>
22. <https://arxiv.org/abs/2306.04141>
23. <https://arxiv.org/abs/1508.06576>
24. <https://arxiv.org/abs/2101.04812>
25. <https://arxiv.org/abs/1803.04469>
26. <https://dl.acm.org/doi/10.1145/3478513.3480485>
27. Mathys, M., Willi, M., Graber, M., & Meier, R. (2024, March 18). *Synthetic Image Generation in Cyber Influence Operations: An Emergent Threat?* Cornell University 2024.
28. <https://ieeexplore.ieee.org/document/9634015>
29. <https://arxiv.org/abs/2101.08629>
30. <https://arxiv.org/abs/2301.09515>
31. <https://arxiv.org/abs/1811.03390>
32. <https://onlinelibrary.wiley.com/doi/10.1155/2022/4956839>
33. <https://www.atlantis-press.com/proceedings/icelaic-19/125934179>
34. <https://arxiv.org/abs/2302.02398>