Investigating Sampler Impact on AI Image Generation: A Case Study on Dogs Playing in the River

Sanjay Deshmukh
D. J Sanghvi College of Engineering, Mumbai, 400056, India

***Abstract*—AI image generation is a new and exciting field with many different uses. It is important to understand how different sampling techniques affect the quality of AI-generated images in order to get the best results. This study looks at how different sampling techniques affect the quality of AI-generated images of dogs playing in the river. This study is limited to a specific scenario, as there are not many images of dogs playing in the river already on the internet. The study used the Playground.ai open-source web platform to test different sampling techniques. DDIM was found to be the best sampling technique for generating realistic images of dogs playing in the river. Euler was also found to be very fast, which is an important consideration when choosing a sampling technique. These findings show that different sampling techniques have different strengths and weaknesses, and it is important to choose the right sampling technique for the specific task at hand. This study provides valuable insights into how sampling techniques affect AI image generation. It is important to choose the right sampling technique for the specific task at hand in order to get the best results. The study also demonstrates the societal relevance of AI-generated imagery in various applications.**

Keywords*—*Artificial Intelligence; image generation; filter; sampler; Euler; Heun

I. Introduction

AI image generation stands as a captivating realm within artificial intelligence, involving the creation of images based on descriptions, prompts, or various inputs [1]. This technology is instrumental across diverse domains, particularly in fields like advertising and web design, where it significantly boosts productivity [2]. AI generators swiftly produce visually appealing content, eliminating the need for intricate editing software and thereby saving both time and costs. Industries like fashion witness substantial benefits as these tools autonomously design clothing and style outfits. Moreover, AI image generators are pivotal in fostering creativity and innovation, generating unique and original art pieces by amalgamating diverse styles and concepts. Their impact extends to several industries, including advertising, architecture, fashion, film, music, and poetry, enriching the creative processes for professionals [3].

The realism achieved by advanced deep learning models in AI image generation is noteworthy, often producing images indistinguishable from those created by humans [4]. However, ethical considerations come to the forefront, especially concerning the generation of images depicting real people in potentially misleading scenarios. Despite the manifold benefits, the utilization of AI image generation technologies should be conscientiously guided by ethical principles.

AI image generation is a rapidly evolving field with numerous applications across various domains, such as advertising, product design, and scientific visualization. As the demand for realistic and high-quality generated images continues to grow, it is crucial to understand the impact of different sampling techniques on the quality and characteristics of the generated outputs. This study aims to investigate the influence of various sampling techniques on the performance of AI image generation models, with a specific focus on a case study involving the generation of images depicting dogs playing in a river.

While AI image generation has made significant strides in recent years, the selection of an appropriate sampling technique can greatly impact the quality, realism, and computational efficiency of the generated images. Different sampling techniques exhibit unique strengths and weaknesses, and their performance can vary depending on the specific use case and requirements. Therefore, it is essential to evaluate and compare the performance of various sampling techniques to identify the most suitable approach for a given application.

The primary research questions addressed in this study are: 1) How do different sampling techniques, such as DDIM, Euler a, DPM, DPM2a, PNDM, Euler, Heun, and LMS, influence the quality and realism of AI-generated images in the context of our case study? 2) What are the trade-offs between response time and image realism for each sampling technique, and how can these trade-offs be balanced to meet specific project requirements? 3) How can the selection of an appropriate sampling technique contribute to the practical applications of AI image generation, particularly in scenarios where readily available reference images are scarce, such as the case study of dogs playing in a river?

The main objectives of this research are: 1) To evaluate and compare the performance of eight different sampling techniques (DDIM, Eular a, DPM, DPM2a, PNDM, Eular, Heun, and LMS) in terms of image quality, realism, and response time. 2) To identify the strengths and weaknesses of each sampling technique and provide insights into the trade-offs between response time and image realism. 3) To propose a framework for selecting the most appropriate sampling technique based on specific project requirements and use cases, with a focus on the case study of generating images of dogs playing in a river. 4) To contribute to the broader understanding of AI image generation techniques and their practical applications, particularly in scenarios where readily available reference images are limited.

By addressing these research questions and objectives, this study aims to provide valuable insights and guidance for researchers, developers, and practitioners working in the field of AI image generation, enabling them to make informed decisions and select the most suitable sampling techniques for their specific projects and applications

There are various AI-powered online tools for generating images. The use of AI in image generation not only saves time but also cuts costs by eliminating the complexities involved in capturing a specific image. AI image generators can be used in many industries for example, advertising a new car. The traditional method of creating a banner involves multiple steps like hiring a photographer, a model, securing a shooting location, getting props and costumes, arranging lighting, and more. Even after these investments, there's no assurance that the real- world picture captured is perfect, often leading to additional iterations and more time and money spent.

On the flip side, using AI for image generation streamlines the entire process. It can generate multiple images within seconds, providing more options and customization. AI also allows for editing specific parts of an image, offering flexibility. The crucial factor in achieving the perfect image with AI lies in selecting the right sampler that determines the final output. Hence, our study aims to establish a platform that sets a standard in selecting samplers for specific cases, such as generating an image of a dog playing in a river.

Given the scarcity of readily available images of dogs playing in rivers, our prompt generates genuinely AI- based images that would be difficult to achieve otherwise. While Photoshop is an alternative for image generation, it requires a highly trained professional, and the results may be influenced by the biases of that professional. Therefore, our decision to focus on the best sampler in the AI image generation domain addresses these challenges and provides valuable insights to the field [5].

II. Different Sampler Under Consideration

We have selected the most popular sampler currently available for free to use and as reported in literature one with the maximum accuracy.

DDIM (Denoising Diffusion Implicit Models): DDIM is a diffusion model sampler in image generation that works by gradually denoising a latent noise image. It is one of the most popular samplers for image generation. DDIM is known for its high quality and stability [6].

The DDIM sampling process can be described by the following mathematical Eq. (1):

$$xt = x0 + \alpha t * (xt - f\theta(x\{t-1\})) + \eta t \tag{1}$$

where:

xt is the latent image at time step t x0is the latent noise image

αt is a noise schedule that controls the amount of denoising at each time step

fθ (x{t−1}) is the denoising function at time step t-1

ηt is a random noise term

The denoising function f θ(x t 1) is a neural network that is trained to denoise latent images. The noise schedule α tis typically chosen to be a monotonically decreasing function so that the latent image becomes less and less noisy as the time step increases.

To generate a sample image, DDIM starts with the latent noise image x 0. It then iteratively applies the denoising function f θ(x t 1) and adds noise according to the noise schedule α t. This process is continued until the desired time step is reached. The final latent image x tis then decoded to produce the generated image.

DDIM, or Diffusion and Denoising Score Matching, presents notable advantages in the realm of image generation. One of its primary strengths lies in its capability to produce images of high quality and stability, providing a reliable output. Additionally, DDIM operates as an efficient sampler, demonstrating the ability to generate images within a reasonable timeframe. A notable advantage is its ease of training; DDIM does not necessitate adversarial training, simplifying the training process.

However, like any methodology, DDIM is not without its drawbacks. One potential issue is mode collapse, a scenario in which the model tends to generate only a limited subset of possible images, limiting diversity. Another challenge is over-smoothing, wherein DDIM may excessively smooth images, leading to a blurred or unrealistic appearance. In summary, DDIM stands out as a powerful and versatile sampler in the domain of image generation, excelling in quality, stability, and training efficiency. Nevertheless, users must be mindful of its potential limitations, specifically the risk of mode collapse and over-smoothing, to make informed decisions in its application.

LMS (Langevin Monte Carlo Sampler): LMS is a sampler that is based on the Langevin equation, which is a stochastic differential equation that describes the motion of a Brownian particle [7]. LMS is known for its ability to generate high-quality images, but it can be slow and computationally expensive.

The mathematical equation for LMS is as follows:

$$dxt = -\nabla U\,(xt)dt + \sqrt{2Ddt} \tag{2}$$

where:

xt is the latent noise image at time step t

U(xt) is the potential energy function.

D is the diffusion coefficient

The potential energy function in the LMS serves as a metric for gauging the probability of the latent noise image, with LMS employing distinct potential energy functions tailored to specific image generation tasks. This adaptability enables LMS to excel in producing high-quality images across a spectrum of tasks.

Highlighting its strengths, LMS exhibits the ability to generate not only high-quality but also diverse images. However, these advantages come with trade-offs. LMS can be slow and computationally expensive, and its effectiveness relies on a trained potential energy function for each image generation task.

In the broader context, LMS emerges as a robust sampler for image generation, finding applications in tasks like image synthesis, inpainting, denoising, and super-resolution. Beyond image-related tasks, LMS extends its utility to other domains within machine learning, including natural language processing, computer vision, and reinforcement learning.

PNDM (Progressive Noise Diffusion Model): PNDM is a diffusion model sampler that is similar to DDIM, but it is more efficient and can generate higher-quality images at higher resolutions [8]. PNDM works by gradually a latent noise image, just like DDIM. However, PNDM uses a progressive approach, where it starts with the image at a low resolution and then gradually increases the resolution. This approach allows PNDM to generate high-quality images at higher resolutions without sacrificing efficiency.

The mathematical equation for PNDM is as follows:

$$xt+1 = xt + \alpha(xt - f(xt)) \tag{3}$$

xt is the latent noise image at time step t

α is the learning rate

f(xt) is the denoising function

The denoising function is a neural network that is trained to denoise images. PNDM uses a different denoising function for each resolution level. This allows PNDM to generate high-quality images at higher resolutions without sacrificing efficiency.

Here is a more detailed explanation of the PNDM algorithm: Start with a latent noise image, x0.

Select a resolution level, r.

Compute the denoising function, f(x0), at the selected resolution level.

Update the latent noise image, x0, using Eq. (4):

$$x1 = x0 + \alpha(x0 - f(x0)) \tag{4}$$

Repeat the steps until the latent noise image is sufficiently denoised. Increase the resolution level, r, and repeat the steps. Once the latent noise image is sufficiently denoised at the highest resolution level, stop the algorithm. The output of the PNDM algorithm is a denoised image, which can then be decoded into a final image.

PNDM can be slower than other samplers, such as Euler and Heun, especially at high resolutions. PNDM requires a trained denoising function for each resolution level. Overall, PNDM is a powerful sampler for image generation that can produce high-quality and diverse images at high resolutions.

Euler: Euler is a simple and efficient sampler that is often used as a baseline for other samplers. It is known for its speed, but it can produce less realistic images than other samplers [9]. The Euler method works by approximating the solution of the ordinary differential equations (ODE) at a given time step using the following Eq. (5):

$$xt+1 = xt + h * f(xt) \tag{5}$$

xt is the solution of the ODE at the time step t,

h is the step size,

f(xt) is the right-hand side of the ODE,

The Euler method is a first-order method, which means that it is not very accurate. However, it is very fast and efficient, and it can be used to generate approximate solutions to ODEs. To apply the Euler method to image generation, we can use it to sample from the latent space of a diffusion model. The latent space of a diffusion model is a space of high- dimensional vectors that represent images. Diffusion models work by gradually denoising latent noise images. To sample from the latent space using the Euler method, we can start with a random latent noise image and then iteratively update the image using the following equation:

$$xt+1 = xt + h * \nabla x \log p(xt) \tag{6}$$

Start with a latent noise image, x0 the function, f(x0). a noise term "epsilon" from a random distribution. the latent noise image x0. Repeat the steps until the latent noise image is sufficient. The output of the Euler A algorithm is an image, which can then be decoded into a final image.

Euler A: Euler A also known as Ancestral Euler, is a diffusion model sampler that is similar to Euler, but it is more efficient and can generate more diverse images [10]. It works by gradually a latent noise image, but it uses an ancestral sampling scheme that allows it to explore a wider range of possible image configurations. The mathematical Eq. (7) for Euler A is as follows:

$$xt+1 = xt + \alpha(xt - f(xt) + \epsilon) \tag{7}$$

xt is the solution of the ODE at time step t f (x) is the right-hand side of the ODE

h is the step size

Euler A is more efficient than Euler and can generate more diverse images. Euler a is more stable than other samplers, such as the Langevin Monte Carlo Sampler (LMS). Euler a stable to generate high-quality images. Euler a can be slower than other samplers, such as Euler, especially at high resolutions Euler A requires a trained function. Overall, Euler is a powerful sampler for image generation that can produce high-quality and diverse images.

Heun: Heun is a numerical method used for approximating the solutions of ordinary differential equations (ODEs) [11]. It is an advancement over Euler's method, providing more realistic images at the cost of increased computational complexity. Represented as a second-order method, Heun's approach involves predicting the solution at the next time step using Euler's method and refining this prediction with the midpoint method. The iterative process continues until the desired accuracy is achieved. The mathematical equation for Heun's method is as follows:

$$k1 = f(xt) \tag{8}$$

$$k2 = f(xt + hk1) \tag{9}$$

where:

xt is the latent noise image at time step t h is the step size

∇xlogp(xt) is the gradient of the log-probability of the latent noise image at time step t

where:

$$xt+1 = xt + 0.5h(k1 + k2) \tag{10}$$

The gradient of the log probability of the latent noise image can be computed using the diffusion model. By iteratively updating the latent noise image using the Euler method, we can generate a variety of different images. The quality and diversity of the generated images will depend on the step size and the number of iterations. The Euler method is very fast and efficient. The Euler method is easy to implement. The Euler method is not very accurate. The Euler method can be unstable for large step sizes. Overall, the Euler method is a simple and efficient sampler for image generation. It is not the most accurate sampler, but it is very fast and easy to implement.

While Heun is more accurate and stable than Euler, it comes at the expense of higher computational demands. The method is relatively straightforward to implement and remains stable across a broad range of step sizes. Despite being surpassed by more sophisticated numerical techniques in terms of accuracy, Heun serves as a reliable baseline method for scenarios where computational resources are limited, making it a pragmatic choice for approximating ODE solutions when striking a balance between accuracy and efficiency is crucial.

DPM2 (Denoising Diffusion Probabilistic Model 2):

Denoising Diffusion Probabilistic Model 2 (DPM2) stands out as a diffusion model sampler celebrated for its stability and capacity to yield high-quality images [12]. While sharing similarities with DDIM, DPM2 excels in efficiency and the generation of superior images, particularly at higher resolutions. The underlying mechanism involves iteratively denoising a latent noise image using a distinct denoising function. The core equation of DPM2 integrates this denoising process with guidance, encompassing the learning rate (α), denoising function (f(xt)), and guidance function (g(xt)).

$$xt+1 = xt + \alpha(xt - f(xt) + g(xt)) \tag{11}$$

where:

xt is the latent noise image at time step t

α is the learning rate

f (xt) is the denoising function

g(xt) is the guidance function.

These neural networks are crucial components, where the former refines image denoising, and the latter guides the process toward an intended output image. DPM2's algorithm unfolds in steps, initiating with a latent noise image and progressing through resolution levels, applying denoising and guidance functions. This process iterates until the latent noise image achieves sufficient denoising, and the algorithm progressively increases the resolution level until the highest is reached, concluding the generation of a denoised image. The decoded result becomes the final image.

DPM2's merits include its enhanced stability compared to other samplers like Euler and Heun, its prowess in generating high-quality images, and its ability to align generated images with a specified guidance image. However, there are drawbacks; DPM2 might be slower than alternative samplers, especially at high resolutions, and necessitates trained denoising and guidance functions for each resolution level.

DPM2a (Denoising Diffusion Probabilistic Model 2 ancestral): DPM2a (Denoising Diffusion Probabilistic Model 2 ances- tral) emerges as a noteworthy variant of the DPM2 diffusion model sampler, renowned for its capacity to generate more diverse images [13]. While sharing a fundamental resemblance with DPM2, DPM2a distinguishes itself through an ancestral sampling scheme, enabling exploration of a broader spectrum of potential image configurations.

$$xt+1 = xt + \alpha(xt - f(xt)) + \beta(xt - xt-1) \tag{12}$$

where:

xt is the latent noise image at time step t

α is the learning rate

β is the ancestral sampling rate

f (xt) is the denoising function

The mathematical formulation for DPM2a involves the latent noise image (xt), learning rate (α), ancestral sampling rate (β), and the denoising function (f(xt)), shared with DPM2. The ancestral sampling rate governs the influence of the latent noise image's previous state, with higher values enhancing image diversity but potentially slowing down the algorithm and introducing instability.

The DPM2a algorithm commences with a latent noise image, progresses through the selection of an ancestral sampling rate, computes the denoising function, and iteratively updates the latent noise image. This process repeats until the image achieves sufficient denoising, concluding the algorithm. Advantages of DPM2a include its capacity to generate more diverse images than DPM2 while still maintaining high-quality and relative stability. However, drawbacks include potential slowness compared to DPM2 and the requisite of a trained denoising function.

Fig. 1 illustrates the conceptual framework of our proposed system designed to investigate the impact of different samplers on a case study prompt—"a dog playing in a river." Utilizing a range of sampler filters, including DPM2a, DDIM, PNDM (PMS), Euler, Euler a, Heun, and LMS, our system incorporates a dedicated validation unit. This unit assesses both response time and reality score, with the aim of minimizing response time and maximizing realization score. In our study, an ideal realization score is defined as 1, a benchmark only achieved by the DDIM sampler. Furthermore, the generated images corresponding to the specified text prompt are systematically archived for potential future investigations.

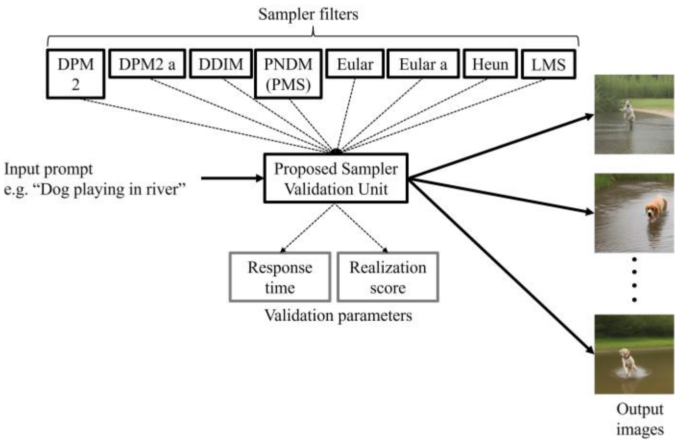

Fig. 1. Concept diagram of the proposed system for exploring sampler impact.

III. LITERATURE REVIEW

The field of artificial intelligence (AI) has seen significant contributions from various researchers across different domains. In (Hosny et al., 2018), a comprehensive understanding of AI methods, especially those related to image-based tasks, is established [14]. Moving from general AI to more specific applications, (Pereira et al., 2020) aims to conduct a thorough analysis of the necessity to integrate tumor information with other lung structures for the advancement of Computer-Aided Diagnosis (CADs), anticipating an impact on targeted therapies and personalized medicine [15]. The study by (Ibrahim et al., 2021) introduces a novel methodology for developing a detailed performance understanding of machine learning benchmarks [16]. While AI's benefits in marketing and advertising are well-documented, (Jeffrey, 2022) investigates the perceptions of Generation Z regarding AI in marketing. This research delves into levels of awareness, understanding, concerns about data privacy, and worries about psychological profiling, stereotyping, and manipulation [17].

(Nasari et al., 2022) undertake a performance comparison between Graphical Processing Units (GPUs) and Intelligence Processing Units (IPUs) by running training benchmarks of common AI/ML models [18]. In another exploration, (Cheng et al., 2023) assess the potential of GPT-4 in various branches of biomedical engineering, addressing challenges like ethical concerns and algorithmic biases [19]. (Alqahtani et al., 2023) contribute to the ongoing discussion about AI's role in education and research, emphasizing its potential to enhance outcomes for students, educators, and researchers [20]. Shifting focus to the realm of 3D object generation, (Sun et al., 2023) present UniG3D, a dataset addressing the limitations of existing 3D object datasets [21]. Lastly, (Tan et al., 2023) introduce DifFSS, the first work leveraging the diffusion model for Few-Shot Segmentation (FSS) tasks [22]. The landscape of AI research is dynamic and encompasses a wide array of applications, from medical diagnosis to marketing perceptions and educational enhancements, highlighting the need for interdisciplinary considerations and ethical frameworks, as underlined by influential works like (Joyce, 2010) [23].

While previous studies have made significant contributions to the field of AI image generation, there are certain limitations that our proposed approach aims to address. One notable limitation is the lack of comprehensive analysis on the impact of different sampling techniques on image quality and realism, particularly for specific use cases or prompts. Additionally, most studies focus on general image generation tasks, overlooking the unique challenges and requirements of generating images for specific scenarios, such as dogs playing in the river. Our proposed approach tackles these limitations by conducting a thorough investigation of various sampling techniques and their effects on image quality and realism, specifically for the case study of generating images of dogs playing in the river. Furthermore, we provide a systematic framework for selecting the most appropriate sampler based on project requirements, enabling more informed decision-making in AI image generation tasks.

IV. METHODS

Eight distinct sampling filters constituted the crux of our experimentation, including DDIM (Denoising Diffusion Implicit Models), Euler a, DPM, DPM2a, PNDM (PMS), Euler, Heun, and LMS. Parameters for evaluation encompassed execution time (t) and the generation of realistic images. To maintain consistency, we employed specific experimental settings random seed "137927237," 25 iterations, prompt guidance set to 7, and fixed image size at 512x512 pixels. Each generative operation yielded four images, and our analysis focused on the realism of these images and the time investment for their creation.

Beyond performance metrics, our study concentrated on the utility of the generated images. We aimed to pinpoint the optimal sampler for the specific scenario of dogs playing in a river, a relatively unconventional but visually engaging context. The versatility of these images was considered for applications ranging from advertising dog-related products to enhancing the appeal of websites and desktop backgrounds. The proposed system, inclusive of the eight distinct filters, underwent rigorous validation using the Proposed Sampler Validation Unit, comparing the output against ground truth images. Evaluation metrics included response time, realization score, and validation parameters tailored to each filter.

The calculated realization score offered a comprehensive measure of system performance, considering the diverse set of filters. Output analysis involved the generation of eight distinct images corresponding to each filter, exemplifying how different filters processed the input prompt. This realization score facilitated nuanced comparisons of performance.

The system, designed to address a spectrum of image processing tasks, from classification to segmentation and denoising, demonstrated versatility. Beyond its immediate applications, the proposed system hinted at the potential for future developments in image processing filters and algorithms. In essence, our methodology ensured a meticulous exploration of samplers, combining performance evaluation with practical considerations in the realm of image generation.

The screenshot in Fig. 2 provides an overview of the Playground AI interface, showcasing a range of accessible options. These include the ability to rate generated images, engage with the community feed featuring images from other users, and utilize the canvas-like Board for image generation and editing. Users can import existing images for further editing, organize their Board using columns, and experiment with different styles of image generation such as Euler, Heun, DPM2, and more. The interface incorporates features like DOIM for diffusion model image generation, PNDM (PLMS) for probabilistic neural diffusion model generation, and Euler as a method for solving differential equations. Filters can be applied, and users can exclude specific details from images. The text prompt, an integral component, guides image generation, while the Generate option brings the vision to life. A Private Session feature ensures the privacy of generated images. Playground AI emerges as a versatile tool, offering a spectrum of options for users to create images ranging from realistic to abstract, thereby establishing itself as a potent resource for image generation and editing.

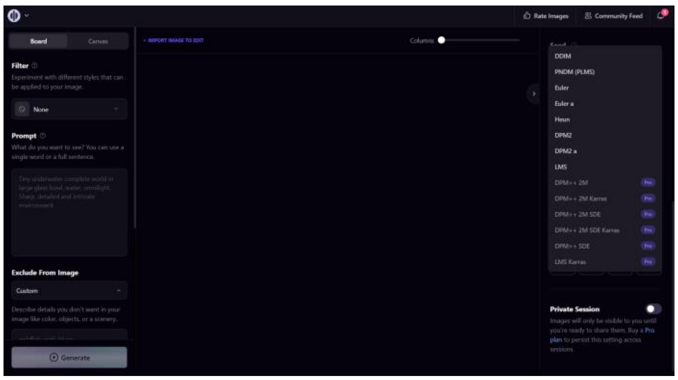

Fig. 2. Screenshot of the playground AI interface with all options visible.

The proposed system, inclusive of the eight distinct filters, each filter underwent rigorous validation using the Proposed Sampler Validation Unit, comparing the output against ground truth images. Evaluation metrics included response time, realization score, and validation parameters tailored to each filter. The calculated realization score offered a comprehensive measure of system performance, considering the diverse set of filters. Output analysis involved the generation of eight distinct images corresponding to each filter, exemplifying how different filters processed the input prompt. This realization score facilitated nuanced comparisons of performance.

In Fig. 3, the input command, exemplified by”Dog playing in River,” is provided as the initial input. The subsequent step involves the selection of a sampler filter from a set of eight options using the algorithm”Select sampler filter.” Once the sampler is chosen, the DiffusionNet CGAN (Conditional Generative Adversarial Network) is employed. The resulting output image is stored to facilitate validation. A check is implemented to verify whether all filters have been adequately tested; otherwise, the operation is halted. This process ensures a systematic approach to testing various sampler filters and capturing their output for thorough validation.

The ability to exclude specific details from images during the filtering process is a powerful feature of our proposed system. This capability is grounded in the principle of selective attention, which allows the model to focus on the most relevant aspects of the input prompt while disregarding unnecessary or distracting elements. By excluding specific details, the system can generate images that are more aligned with the intended subject matter, reducing visual clutter and enhancing the overall coherence and clarity of the output. The exclusion of specific details is particularly useful in scenarios where the input prompt may contain extraneous information or when the desired output requires a certain level of abstraction or stylization. For instance, in our case study of generating images of dogs playing in a river, excluding irrelevant background details could result in a more focused and visually appealing representation of the subject matter. Furthermore, the ability to exclude specific details can be leveraged to mitigate potential biases or undesirable elements that may be present in the training data or the input prompt. By carefully filtering out such elements, the generated images can better reflect the intended message or concept.

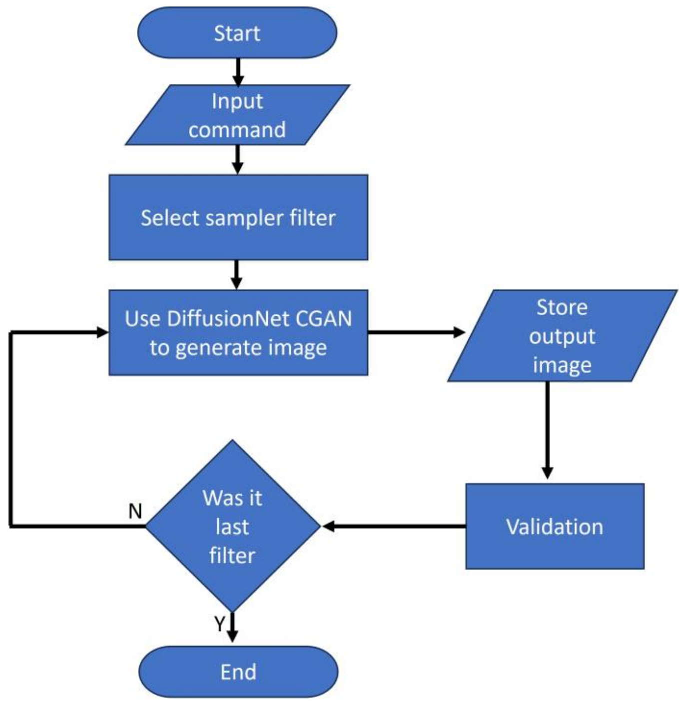

Fig. 3. Flowchart depicting the proposed mechanism for investigating the influence of samplers on AI generation.

Fig. 4(a) shows the effect of the sampler (filter) type used for AI-based image generation and the corresponding response time. The response time is the amount of time it takes for the model to generate an image from a given text prompt. The sampler type has a significant impact on the response time. The ‘PNDM’ and ‘Euler a’ samplers are the fastest, followed by the DPM, LMS, and Heun filters. The slowest sampler is the Euler type. Fig. 4(b) shows the sampler type used in AI image generation versus reality score of the generated images. The reality score is a measure of how realistic the generated images are. The sampler type also has a significant impact on the reality score. The DDIM sampler generates the most realistic images, followed by other samplers. The DPM2 and Euler sampler generate the least realistic images.

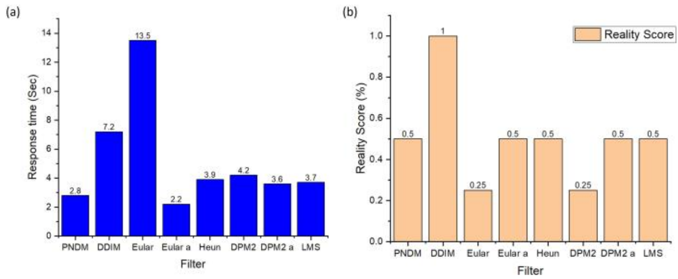

Fig. 4. Results obtained with proposed method (a) Filter type in AI image generation vs response time. (b) Filter type in AI image generation vs reality score in (%).

In general, there is a trade-off between response time and reality score. If you need to generate images quickly, you should use a faster sampler, such as Euler a. However, if you need to generate realistic images, you should use a slower filter, such as DDIM. If you need to generate realistic images at a moderate speed, you could use the PNDM sampler. Ultimately, the best sampler type for the project will depend on specific needs and requirements.

The images in Fig. 5 show a clear difference in the quality and realism of the generated images, depending on the sampler used. The PNDM and Euler a samplers produce the fastest results, but the images are also the least realistic. The DDIM and Euler samplers produce more realistic images, but they are also slower. The PNDM sampler strikes a balance between speed and realism. The dog in the PNDM image is blurry and the details are not very sharp. The water also looks unrealistic. The DPM image is sharper than the PNDM image, but the dog is still blurry in some places. The water looks more realistic, but it is still not perfect. The DPM2a image is sharper than the DPM image and the dog is no longer blurry. The water looks even more realistic. The DDIM image is the sharpest and most realistic image of all. The dog and the water are both very well-rendered. The Euler image is also very blurry and less realistic. The dog is poorly-rendered, and the water looks unnatural. The 'Euler a' image is much better than the Euler image, but it is slightly less sharp. The Heun image is not as sharp as the Euler a or DDIM images, but it is still more realistic than the DPM images. The LMS image is the blurry and least realistic image of all.

Fig. 5. Output images from each type of sampler for the prompt "dog playing in the river".

V. Results

The research underscores the significance of sampler selection in shaping the outcomes of AI image generation projects. The contextual analysis of a case study, where different samplers were evaluated for generating images of dogs playing in the river, illustrates the project-specific nature of this decision-making process. In this section, we discuss implications and findings that arise from the exploration of different samplers in the context of AI image generation. This study involved eight distinct sampling filters, each contributing unique characteristics to the generated images. The key factors evaluated were response time and the realism of the images, providing a nuanced understanding of the trade-offs involved in selecting a sampler for specific projects.

The DDIM sampler emerged as a consistent frontrunner in terms of image realism for this particular case study of dog image generation that plays in the river. This could be because of the iterative refinement approach of DDIM that sets it apart, enabling the generation of highly realistic images. This makes DDIM a compelling choice for projects where authenticity and visual fidelity are paramount, such as scientific studies or applications demanding a high level of image realism. On the other end of the spectrum, the PNDM sampler demonstrated a remarkable balance between response time and image realism. Its efficiency, coupled with the ability to produce realistic images, positions it as a versatile option suitable for a broad range of projects. The findings emphasize the importance of considering project requirements and objectives when selecting a sampler.

For instance, the Euler a sampler, with its combination of speed and reasonable realism, might be ideal for a marketing campaign requiring quick generation of high-quality images. The nuanced understanding of each sampler's strengths and weaknesses provides a practical guide for selecting the most appropriate sampler based on the specific requirements of a given project. This research lays the foundation for informed decision-making in the rapidly evolving field of AI image generation.

Our research has the potential to impact a diverse range of fields, including advertising, product design, and even dog training. For example, advertisers could leverage our findings to generate more engaging and effective marketing campaigns. Product designers could use our insights to create more realistic and appealing product prototypes. Additionally, dog trainers could utilize our research to develop more effective training methods.

Table I presents a comparison of different filters used in image processing along with their corresponding validation parameters and the number of images processed. The validation parameters include response time, realization score, and a calculated validation parameter. Response time indicates the time taken by each filter to process the images, with lower values being preferable as they indicate faster processing times. Realization score measures the effectiveness of each filter in achieving the desired outcome, with higher scores indicating better performance. The validation parameter is calculated using a formula based on the realization score and response time, providing an overall assessment of filter performance. The table enables the evaluation and comparison of filters based on these parameters, facilitating the selection of the most suitable filter for image processing tasks.

TABLE I. FILTER AND CORROSPONDING VALIDATION PARAMETERS

Filter	Validation parameter			Number of images
	Response time	*Realization score*	*Validation Parameter*	
PNDM	2.8	22	50.4	250
DDIM	7.2	39	7.2	250
Eular	13.5	11	391.5	250
Eular a	2.2	23	37.4	250
Heun	3.9	21	74.1	250
DPM2	4.2	12	117.6	250
DPM2 a	3.6	22	64.8	250
LMS	3.7	21	70.3	250

There is a clear difference in the quality and realism of the generated images, depending on the sampler used. The PNDM and Euler a samplers produce the fastest results, but the images are also the least realistic. The DDIM and Euler samplers produce more realistic images, but they are also slower. The PNDM sampler strikes a balance between speed and realism. The dog in the PNDM image is blurry, and the details are not very sharp. The water in the images generated also looks unrealistic. The DPM image is sharper than the PNDM image, but the dog is still blurry in some places. The water looks more realistic, but it is still not perfect. The DPM2a image is sharper than the DPM image, and the dog is no longer blurry. The water looks even more realistic. The DDIM image is the sharpest and most realistic image of all. The dog and the water are both very well-rendered. The Euler image is also very blurry and less realistic. The dog is poorly rendered, and the water looks unnatural. The 'Euler a' image is much better than the Euler image, but it is slightly less sharp. The Heun image is not as sharp as the Euler a or DDIM images, but it is still more realistic than the DPM images. The LMS image is the blurriest and least realistic image of all.

Our research has several potential future implications. First, it could inspire the development of new samplers that offer even better performance in terms of response time, image realism, or both. Second, it could lead to the development of new AI image generation tools that incorporate our findings to make them more user-friendly and effective. Third, it could inform the development of new applications for AI image generation in a wider range of fields.

We are excited to see how our research is used to advance the field of AI image generation and its applications in the future. PNDM is more efficient than DDIM and can generate higher- quality images at higher resolutions. PNDM is more stable than other samplers, such as Euler and PNDM is able to generate diverse images.

Our proposed model introduces several innovative contributions to the field of AI image generation, particularly in the context of evaluating and selecting appropriate samplers for specific use cases:

1) Comprehensive Sampler Evaluation Framework: Our study presents a systematic and holistic framework for evaluating the performance of various samplers in AI image generation. By considering a diverse set of eight samplers, including DDIM, Euler a, DPM, DPM2a, PNDM, Euler, Heun, and LMS, we provide a comprehensive understanding of their strengths, weaknesses, and trade-offs in terms of response time and image realism.

2) Case Study-Driven Approach: Our research adopts a novel case study-driven approach, focusing on the specific scenario of generating images of dogs playing in a river. This unconventional yet visually engaging context allows us to evaluate the samplers' performance in a real-world setting, providing practical insights that can inform decision-making processes for diverse applications.

3) Proposed Sampler Validation Unit: A key innovative aspect of our work is the introduction of the Proposed Sampler Validation Unit. This unit systematically validates the output of different samplers against ground truth images, employing a combination of quantitative metrics (response time and realization score) and qualitative analyses. This robust validation approach ensures a thorough assessment of the generated images, enabling informed selection of the most suitable sampler for a given task.

4) Versatile and Extensible Framework: Our proposed model is designed to be versatile and extensible, capable of addressing a wide range of image processing tasks, from classification and segmentation to denoising. Additionally, the framework lays the foundation for future developments in image processing filters and algorithms, fostering continued innovation and improvement in the field of AI image generation.

By presenting these innovative contributions, our research not only advances the understanding of sampler impact on AI image generation but also provides a practical and adaptable framework for researchers, developers, and practitioners to leverage in their respective domains.

VI. CONCLUSIONS

AI image generation is a rapidly evolving field with a wide range of potential applications. However, the quality and realism of generated images are highly dependent on the sampler type used. This paper presents a comprehensive study on the impact of eight distinct samplers on AI image generation, namely DPM, DPM2a, DDIM, PNDM, Euler, Euler a, Heun, and LMS.

Our findings reveal a nuanced landscape where response time and image realism form a delicate balance. Samplers such as PNDM and Euler a offer the fastest response times, making them ideal for projects where expeditious output is essential. Conversely, the Euler sampler, albeit slower, demonstrates superior performance in terms of image realism. In terms of image realism, the DDIM sampler consistently outperforms all others. This is attributed to its unique sampling approach, which iteratively refines the generated image to achieve greater realism. As such, the DDIM sampler is the best choice for projects where image authenticity is paramount.

The PNDM sampler strikes a balance between response time and image realism. It is faster than the DDIM sampler but still produces realistic images. This makes it a versatile option for a wide range of projects. Our research underscores the importance of carefully selecting the appropriate sampler for each project. For instance, in the context of our case study on images of dogs playing in the river, we identified that the Euler a sampler would be the best choice for a marketing campaign that requires quick generation of high-quality images. Conversely, the DDIM sampler would be the better choice for a scientific study that requires highly realistic images of dogs playing in the river.

DECLARATIONS

A. *Ethics Statements*

All authors ensure that the manuscript fulfills the following statements:

1) This material is the author's original work, which has not been previously published elsewhere.

2) The paper is not currently being considered for publication elsewhere.

3) The paper reflects the author's own research and analysis truthfully and completely.

4) The paper properly credits the meaningful contributions of co-authors and researchers.

B. *Availability of Data and Materials*

Data and code will be made available, on reasonable request, to the corresponding author.

C. *Funding Statement*

This research has no funding associated with it.

D. *Data Availability*

Data will be made available at reasonable request to the authors.

E. *Supplementary Information*

Not applicable.

F. *Ethical Approval*

All the ethics approval was taken by an institutional review board or equivalent ethics committee.

G. *Author Contributions*

Conceptualization was done by Sanjay Deshmukh (SD). The experimental design was done by SD. All the experiments were performed by SD. The manuscript draft was prepared by SD. Data analysis and graphics designing were done by SD.

H. *Conflicts of Interest or Competing Interests*

The authors declare that there is no conflict of interest or competing interests.

REFERENCES

[1] S. Tanugraha, "A Review Using Artificial Intelligence- Generating Images: Exploring Material Ideas from Mid- Journey to Improve Vernacular Designs," Journal of Artificial Intelligence in Architecture, vol. 2, no. 2, pp. 48–57, 2023.

[2] S. O. Abioye, L. O. Oyedele, L. Akanbi, A. Ajayi, J. M. D. Delgado, M. Bilal, O. O. Akinade, and A. Ahmed, "Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges," Journal of Building Engineering, vol. 44, p. 103299, 2021.

[3] N. Anantrasirichai and D. Bull, "Artificial intelligence in the creative industries: a review," Artificial intelligence review, pp. 1–68, 2022.

[4] D. Schraml, "Physically based synthetic image genera- tion for machine learning: a review of pertinent litera- ture," Photonics and Education in Measurement Science 2019, vol. 11144, pp. 108–120, 2019.

[5] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learn- ing applications and challenges in big data analytics," Journal of big data, vol. 2, no. 1, pp. 1–21, 2015.

[6] J. Song, C. Meng, and S. Ermon, "Denoising diffu- sion implicit models," arXiv preprint arXiv:2010.02502, 2020.

[7] S. Kumar, B. Paria, and Y. Tsvetkov, "Constrained sam- pling from language models via langevin dynamics in embedding spaces," arXiv preprint arXiv:2205.12558, 2022.

[8] D. Ryu and J. C. Ye, "Pyramidal denoising diffusion probabilistic models," arXiv preprint arXiv:2208.01864, 2022.

[9] Q. Zhang and Y. Chen, "Fast sampling of diffu- sion models with exponential integrator," arXiv preprint arXiv:2204.13902, 2022.

[10] A. M. Ben-Amram, "The Euler path to static level- ancestors," arXiv preprint arXiv:0909.1030, 2009.

[11] H. R. Rezazadeh, M. Maghasedi, and B. Shojaee, "Nu- merical Solution of Heun Equation Via Linear Stochastic Differential Equation," Journal of Linear and Topological Algebra, vol. 1, no. 02, pp. 83–95, 2012.

[12] S.-Y. Chou, P.-Y. Chen, and T.-Y. Ho, "How to backdoor diffusion models?" in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 4015–4024.

[13] D. Watson, J. Ho, M. Norouzi, and W. Chan, "Learning to efficiently sample from diffusion probabilistic models," arXiv preprint arXiv:2106.03802, 2021.

[14] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. Aerts, "Artificial intelligence in radiology," Nature Reviews Cancer, vol. 18, no. 8, pp. 500–510, 2018.

[15] T. Pereira, C. Freitas, J. L. Costa, J. Morgado, F. Silva, E. Negra˜o, B. F. de Lima, M. C. da Silva, A. J. Madureira, and I. Ramos, "Comprehensive perspective for lung cancer characterisation based on ai solutions using ct images," Journal of Clinical Medicine, vol. 10, no. 1, p. 118, 2020.

[16] K. Z. Ibrahim, T. Nguyen, H. A. Nam, W. Bhimji, S. Far- rell, L. Oliker, M. Rowan, N. J. Wright, and S. Williams, "Architectural requirements for deep learning workloads in hpc environments," in 2021 International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS), 2021, pp. 7–17.

[17] T. R. Jeffrey, "Understanding Generation Z Perceptions of Artificial Intelligence in Marketing and Advertising," Advertising & Society Quarterly, vol. 22, no. 4, 2021.

[18] A. Nasari, H. Le, R. Lawrence, Z. He, X. Yang, M. Krell,A. Tsyplikhin, M. Tatineni, T. Cockerill, and L. Perez, "Benchmarking the Performance of Accelerators on Na- tional Cyberinfrastructure Resources for Artificial Intel- ligence/Machine Learning Workloads," pp. 1–9, 2022.

[19] K. Cheng, Q. Guo, Y. He, Y. Lu, S. Gu, and H. Wu, "Exploring the potential of GPT-4 in biomedical engi- neering: the dawn of a new era," Annals of Biomedical Engineering, pp. 1–9, 2023.

[20] T. Alqahtani, H. A. Badreldin, M. Alrashed, A. I. Al- shaya, S. S. Alghamdi, K. bin Saleh, S. A. Alowais, O. A. Alshaya, I. Rahman, and M. S. A. Yami, "The emergent role of artificial intelligence, natural learning processing, and large language models in higher educa- tion and research," Research in Social and Administrative Pharmacy, 2023.

[21] Q. Sun, Y. Li, Z. Liu, X. Huang, F. Liu, X. Liu, W. Ouyang, and J. Shao, "UniG3D: A Unified 3D Object Generation Dataset," arXiv preprint arXiv:2306.10730, 2023.

[22] W. Tan, S. Chen, and B. Yan, "Diffss: Diffusion model for few-shot semantic segmentation," arXiv preprint arXiv:2307.00773, 2023.

[23] D. Joyce, "Photography and the image-making of inter- national justice," Law and Humanities, vol. 4, no. 2, pp. 229–249, 2010.