IT UNIVERSITY OF COPENHAGEN

May 14, 2024

Bachelor Project

Course ID: BIBAPRO1PE

XGANART

Authors:

Anne-Marie Rommerdahl annro@itu.dk
Daria Damian dard@itu.dk
Ida Barkou Vilstrup ivil@itu.dk

Supervisors:

Djordje Grbic djgr@itu.dk
Imke Grabe imgr@itu.dk
Louie Søs Meyer lomm@itu.dk

Computer Science Department,

IT University of Copenhagen

Contents

Abstract

In recent years, Generative Adversarial Networks (GANs) have emerged as a powerful tool for generating realistic images. Despite their effectiveness, the opaque nature of these models has hindered their integration into human-AI collaborative environments. This bachelor project addresses the challenge of enhancing the transparency and explainability of GANs through the development of the XGANArt interface. Using the theoretical work of previous XAI research, XGANArt builds upons the foundational work of the GANPaint project [Bau et al., 2018]. XGANArt allows users to interact with different layers of a GAN, providing insights into the inner workings of the model. By exposing the hidden layers of convolutional neural networks, XGANArt aims to facilitate a deeper understanding of the image generation process and fosters human-AI interactions. User testing indicates that XGANArt improves comprehension of GANs, with participants showing an increased ability to understand the technical workings of the model. The XGANArt project highlights the importance of transparency and explainability in generative AI systems and suggests that interfaces like XGANArt can make GANs more accessible and comprehensible, thus enhancing the usability and reliability of generative models in co-creative processes.

1 Introduction

In recent years, there has been a significant increase in the use and popularity of generative models. These models, particularly Generative Adversarial Networks (GANs), have showcased remarkable proficiency in synthesizing realistic images, facilitating various applications across various domains, such as healthcare, education, security, scientific research, etc. However, in many applications of Machine Learning (ML), the inner workings of these generative models often remain shrouded in mystery, resembling what is commonly referred to as a 'black box'. This term relates to the opacity or lack of transparency in understanding how these models function internally.

This lack of transparency inherent to generative models poses a considerable challenge, particularly in scenarios where human interaction and interpretability are crucial. Even though information about the inner workings of models in ML is available, these

sources are often very technical and require time and effort to understand. Therefore, in order to facilitate interactions between the end-user and the model, one must confront the challenge of ensuring an understanding of *how* the ML model created its artifact(s) [Bau et al., 2018], so that understanding can be gained without extensive research.

The inability to comprehend the underlying processes through which these models generate outputs/artifacts impedes their integration into collaborative and co-creative endeavors involving humans and AI systems [Wei Xu and Gao, 2023]. However, to our knowledge, there are currently no interfaces that comprehensively tackle this problem, and the existing AI generation of artifacts does not provide an understanding of the inner workings. Existing interfaces such as GANPaint from the GANDissect project [Bau et al., 2018] offer valuable insights into certain aspects of GAN operations, yet they often fall short of providing a holistic and intuitive understanding that is accessible to non-experts. Interpretability methods, for instance, the GANDissect method, allow users to visualize and modify the feature maps of a GAN to understand what each neuron in the network does, but it requires a significant level of technical knowledge to be effectively used. Similarly, GANPaint provides a way to edit images generated by GANs interactively, yet it does not fully explain the underlying mechanisms of these modifications. Recognizing this critical gap, this project sets out to build out an interface that builds on deciphering methods of the 'black box' nature of generative models and enhancing their explainability to facilitate meaningful human-AI interaction in the co-creative processes. Our project poses the research question: How can the transparency and explainability of generative models, particularly Generative Adversarial Networks (GANs), be enhanced through the development of an interface that exposes the hidden layers of a model and enables users to browse them - using the GANDissect model?

To address this, our project aims to develop the explainable interface - XGANArt - that unravels the intricacies of generative models, shedding light on the latent mechanisms driving image generation. This interface intends to allow users to engage with and comprehend the inner workings of these models in a transparent and intuitive manner. This endeavor is underpinned by the field of Mechanistic Interpretability, which seeks to elucidate how semantic features are encoded and manipulated within the hidden layers

of neural networks [Olah, 2022]. At its core, our research aims to explain the intricate processes involved in image generation by bridging the gap between the complex inner workings of generative models and human understanding. By creating an interface that exposes the hidden layers of convolutional neural networks (CNNs) - the backbone of many state-of-the-art generative models - to users, we aim to provide individuals with the ability to understand and potentially influence the image generation process - demystifying the 'black box'.

2 Related Work

In this section, we delve into the existing body of research that underpins our project. We begin by introducing the foundational concepts in Generative Adversarial Networks (GANs) and the seminal works that have influenced our approach. Additionally, we explore theoretical frameworks such as Human-AI interaction, Explainability, and Interaction patterns, which inform the development of our project. These topics are explored to gain an understanding of how humans interact with AI and how this knowledge can be used to improve explainability.

2.1 Generative Adversarial Networks

Generative Adversarial Networks (GANs), initially proposed by Goodfellow et al. (2014), have emerged as a powerful tool for generating realistic images by learning from complex data distributions. This transformative technology has found applications across diverse domains, including computer vision, natural language processing, and healthcare [Liao and Varshney, 2022]. Notably, Bau et al. (2018) contributed to the understanding of GANs' underlying features and mechanisms, providing opportunities for subsequent advancements. Our project builds upon the pioneering work of Bau et al. (2018) in "GANPaint," which introduced an innovative image editing tool leveraging GANs. By directly manipulating neuron activations, users could paint visual concepts onto images, demonstrating the creative potential of GAN-based interfaces. The project adopts the Progressive GAN architecture proposed by Karras et al. (2018) and trains the model on diverse scene datasets, enabling the generation of high-resolution, realistic

images. The implementation and training process will be described in a later section 3, of our paper.

By building on the work of the GANDissect project, our study aims to provide users with insights into the underlying processes of the generative model. Through the further development of the GanPaint interface, we intend to expose hidden layers of the convolutional neural network and enable users to delve deeper into the inner workings of the model, unraveling the mechanisms behind their generation process. Our project desires to foster a more transparent and collaborative human-AI interaction. By shedding light on the workings of GANs and providing users with insights into their operation, we strive to enable users to interact more meaningfully with these models and harness their creative potential in collaborative endeavors.

2.2 Human-AI Interaction and Co-creativity

Human-Computer Interaction (HCI) is a growing field that explores the collaborative interplay between humans and artificial intelligence systems. HCI focuses on developing an empirical understanding of the user, which lies at the base of our project endeavors. Understanding HCI yields insights crucial for designing interfaces that cater to human needs and preferences [Fawcett, 2021].

Previous research shows that interaction between humans and AIs fosters creativity, not just supports it. Human-computer interaction can support creativity through human-computer cooperation during idea production and using creative enhancement techniques[Lubart, 2005][Wu et al., 2021]. A computer can nurture the creative ideas of a human designer, "tackling nontrivial tasks through an indirect, non-linear, creative approach", fostering co-creativity [Yannakakis et al., 2014].

However, a significant challenge lies in ensuring transparency and explainability in AI systems, particularly in the context of creative collaboration. AI-based decisions are often not intuitive; for many non-technical users, a generative machine-learning model is like a "black box", leaving the users in the dark about the processes happening behind it. Previous research proposes a possible solution to this problem in the form of eXplainable AI (XAI), which refers to the ability of AI systems to provide understandable explanations for their decisions and actions [Wei Xu and Gao, 2023]. We intend

to use such HCI techniques in order to provide the user insights into what the machine is thinking and explain how it works.

2.3 Explainability

As AI systems advance in capability, they simultaneously become increasingly challenging for users to comprehend. The research area of explainability aims to bridge the gap between the interfaces presented to users and the inner workings of machines. [Zhu et al., 2018] states that "there is no established definition of explainability and how to measure it". This is true in the sense that there is no shared definition of explainability. However, since the release of the Zhu et al. (2018) article, several papers have been published which present different definitions of explainability. [Liao and Varshney, 2022] defines an explanation in regards to explainability as; "An explanation is a presentation of (aspects of) the reasoning, functioning and/or behavior of a machine learning model in human-understandable terms." [Nauta et al., 2023] defines explainability as all technical means that share the goal of "making AI understandable by people." This definition does not seek to define what an explanation is but rather defines the goal of explainability and thus accepts any means that will reach this goal. In this project, we choose to define explainability as encompassing both of these definitions. We wish to give an explanation which is a presentation of the reason of a machine learning model with the goal of making AI understandable by people.

Various research projects have already made efforts to further the understanding of the inner workings of machine learning models. One such project [Olah, 2022], is using mechanistic interpretability to reverse engineer neural networks. It is here argued that the central part of mechanistic interpretability is to reach an understanding of individual neurons and activations. As such, mechanistic interpretability can help to offer a very technical understanding of machine learning models.

It is, however, also important to consider the use case in which an AI will be used, as the above would be too technical for some groups of users. A simpler interface might not offer the complexity needed for a specific user. Similarly, an interface that offers explanations for each operation of a ML model could cause information overload depending on user needs. Thus, both of these scenarios would counter the goal of

explainability [Zhu et al., 2018]. As such explainability must be implemented in a way that is beneficial to the user.

This is the challenge of explainability, which is one of the core issues facing interfaces based on machine learning today. Our project seeks to address this challenge by facilitating greater transparency and interpretability in the operation of generative models, thereby fostering more meaningful human-AI interaction.

2.4 Interaction patterns

As mentioned, one could consider the possible use cases in which AI will be used by a human to design an intuitive user experience. The framework proposed by [Grabe et al., 2022] identifies different actions that make up the interaction between the AI and the human agent, as well as four different human-AI interaction patterns - of which the "Explore"-pattern will be our focus.

The Explore-pattern is signified by a process wherein the human iteratively *adapts* artifacts that are created by sampling from a model. To adapt is to make changes to existing artifacts. The AI can then create new artifacts based on the changes, allowing the human agent to control the process as they explore the possible designs. Therefore, the Explore-pattern aligns with our project's emphasis on enabling users to try and manipulate generative models in order to explore how the model works.

2.5 Current Existing Interfaces

Several existing applications focus on human-AI interaction, offering varying degrees of functionality and explainability. Examples include CG-GAN [Zaltron et al., 2020], and research, such as [Yosinski et al., 2015] and [Olah, 2022]. While some applications prioritize user functionality and intuitiveness, they often lack transparency regarding the inner workings of the underlying AI models. Conversely, our project aims to bridge this gap by exposing the hidden layers of the model to users. Similarly to our project, GANs have also been used in other generative systems such as seen in CG-GAN [Zaltron et al., 2020]. The CG-GAN approach showed that building generative systems based on GANs helped even non-expert users to create target images far more accurately than when using existing interfaces built on more traditional methods.

By using GANs for image generation interfaces like CG-GAN, the user was able to manipulate and evolve images instead of relying on the traditional methods of being able to select pre-defined features, which made for an intuitive user experience. While CG-GAN effectively utilizes GANs to facilitate user interaction and image generation, it lacks mechanisms for users to comprehend the underlying workings of the model. Unlike our project, which aims to expose hidden layers of the model, CG-GAN does not provide users with insights into how the AI generates images or the specific processes occurring within the neural network. As a result, users are unable to gain a deeper understanding of the AI's decision-making process or the factors influencing image generation, limiting their ability to interpret or modify the output effectively.

While these interfaces demonstrate the potential of AI-driven creativity, they often lack transparency and user-friendly explanations of model behavior. Our project aims to approach this gap by providing users with intuitive tools to interact with and understand generative models effectively.

3 Theoretical Understanding of GANs

In this technical section of our project paper, we delve into the intricate architecture and functioning of GANs based on the research presented in "GAN Dissection: Visualizing and Understanding Generative Adversarial Networks" by [Bau et al., 2018]. This study provides an analysis of the individual units within GANs, exploring their specific roles and the substantial effects these units have on the visual outputs generated by the networks. Building upon this, our project extends the dissection techniques to further enhance the understanding and manipulation of GANs. We employ interpretability methods, including feature visualizations, to demystify the often opaque processes within these networks. Our contributions aim to explain the internal workings of GANs, focusing on making the generative models more transparent and understandable, thereby addressing critical challenges in the field of generative modeling.

3.1 Neural Networks Fundamentals in GANs

At the heart of a GAN are two distinct neural networks termed the generator and the discriminator, as shown in Figure 1. These networks engage in a continuous adversarial process, with the generator creating data instances and the discriminator evaluating them. The generator aims to produce data indistinguishable from genuine data, while the discriminator strives to distinguish between real and generated data [Bau et al., 2018].

- Generator: This component of a GAN takes a sample from latent space as input and then adds random noise and transforms it into a data output that mimics the real data in structure and content. The transformation is achieved through a series of layered computations, each passing through numerous neurons.

- Discriminator: Operating as a binary classifier, the discriminator assesses the authenticity of the data received, outputting a probability that the data is real (as opposed to synthetic). It is trained to maximize its accuracy in identifying real and synthetic data.

3.2 Neural Networks Summary

Linear algebra is integral to understanding how neural networks operate. The operations within a GAN's neural network components are composed of linear transformations and non-linear activations that manipulate the input data into the output [Géron, 2019].

1. Matrix Multiplication: Each layer in a neural network can be conceptualized as performing a matrix multiplication. The inputs to each neuron are multiplied by a weight matrix and then summed up with a bias vector. The matrix multiplication aspect is crucial because it allows the network to learn complex patterns from the data by adjusting the weights during training. Notably, both the weight matrix and the bias vector are tunable parameters, meaning they can be iteratively optimized during the training process to minimize the loss function and improve the network's performance.

$$\mathbf{z}^{(l+1)} = \mathbf{W}^{(l)}\mathbf{a}^{(l)} + \mathbf{b}^{(l)}$$

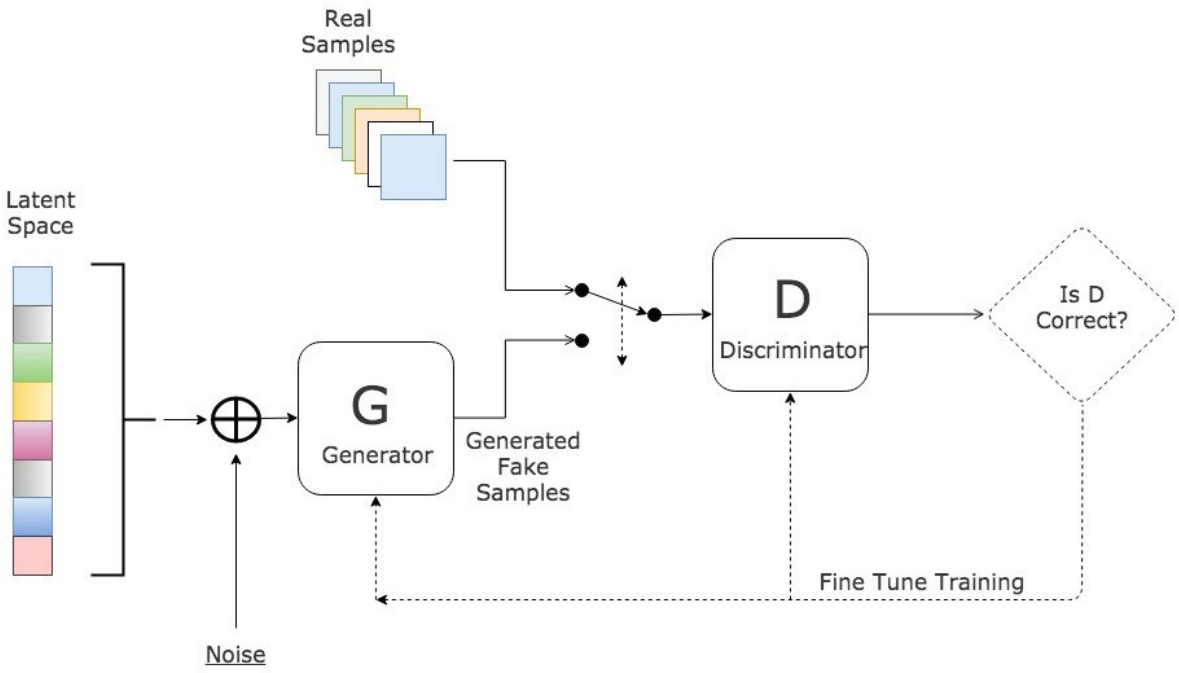


Figure 1: The structure of a Generative Adversarial Network (GAN). Reprinted from Scene Level Image Classification: A Literature Review by Chavda, S., Goyani, M. Neural Process Lett 55, 2471–2520 (2023) [Chavda and Goyani, 2023].

Here, $\mathbf{z}^{(l+1)}$ represents the input to the next layer, $\mathbf{W}^{(l)}$ is the weight matrix for layer $\mathbf{l}$, $\mathbf{a}^{(l)}$ is the activation from the previous layer, and $\mathbf{b}^{(l)}$ is the bias.

2. Activation Functions: After the linear transformation, an activation function is applied to introduce non-linearity. This allows the network to learn and represent more complex non-linear patterns.

$$\mathbf{a}^{(l+1)} = \sigma(\mathbf{z}^{(l+1)})$$

In this formula, σ denotes the activation function. One such activation function is the Leaky Rectified Linear Unit (Leaky ReLU), which addresses some of the issues of the standard ReLU function by allowing a small, non-zero gradient when the unit is not active. For Leaky ReLU, the function σ is defined as:

$$\sigma(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0.01x, & \text{otherwise} \end{cases}$$

This formulation of σ transforms the linear combination of inputs into an output vector that acts as input to the next layer or serves as the final output of the neural

network. By allowing a small slope (e.g., 0.01) when x is less than zero, it maintains a gradient flow during the training process, which can help mitigate the vanishing gradient problem encountered in traditional ReLU functions.

3.3 Training Process

Training a GAN involves adjusting the parameters of both the generator and the discriminator to minimize their respective loss functions. Both networks have loss functions designed to measure how well they are performing their tasks shown in Equations 1 and 2 (e.g., how well the discriminator distinguishes real images from fakes and how convincing the generator's images are). The training is performed using backpropagation, which efficiently computes gradients using the chain rule of calculus, essential for updating the weights in the network:

The process can be broken down into two phases for the two different networks:

- **Generator Training**: The generator creates images that are passed to the discriminator. The objective of the generator is to generate images that are indistinguishable from real images, deceiving the discriminator. The loss function for the generator (L_G) can be defined as:

$$L_G = -\log D(G(z)) \tag{1}$$

  Here, $G(z)$ represents the output of the generator given a noise variable z, and $D(G(z))$ represents the discriminator's estimate of the probability that the generated image is real. The generator G tries to minimize this loss function, which corresponds to maximizing the error of the discriminator.

- **Discriminator Training**: The discriminator improves its ability to distinguish real-world data from synthetic data generated by the generator. This is done by adjusting its own weights based on its success in classifying real and generated images correctly. The loss function for the discriminator (L_D) can be expressed as:

$$L_D = -\left[\log D(x) + \log(1 - D(G(z)))\right] \tag{2}$$

In this equation, x represents real data, and $D(x)$ is the probability estimated by the discriminator that the real data is indeed real. Similarly, $D(G(z))$ is the probability estimated for the fake data generated by G. The discriminator D aims to maximize this function, effectively trying to minimize its classification errors for both real and synthetic images.

Each network trains by attempting to outperform the other. The generator seeks to fool the discriminator, while the discriminator tries to become better at distinguishing genuine from fake inputs. This adversarial process leads to improved quality of the generated outputs and more effective discriminative performance.

3.4 Specific Model Architecture Used

The [Bau et al., 2018] adopts the Progressive GAN (PGGAN) architecture proposed by [Karras et al., 2018] and trains on diverse scene datasets, promising the generation of high-resolution, realistic images.

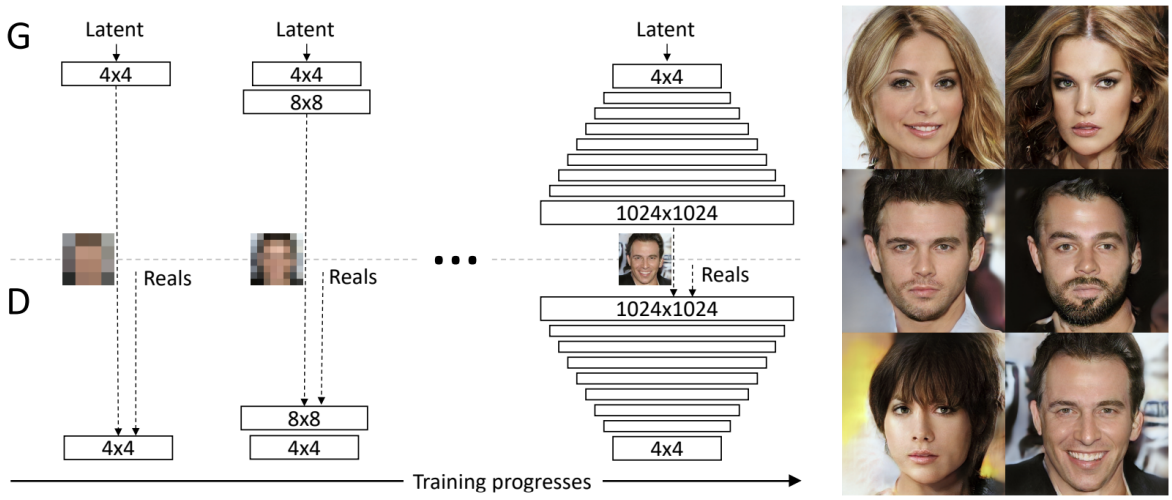


Figure 2: Training of PGGAN, with generator (G) and discriminator (D). Reprinted from Progressive Growing of GANs for Improved Quality, Stability, and Variation by Tero Karras, Timo Aila, Samuli Laine, Jaakko Lehtinen p3 [Karras et al., 2018].

The PGGAN approach starts the training process with low-resolution images, typically beginning at 4x4 pixels. This initial simplicity allows the networks to focus on capturing the most fundamental aspects of image distribution. As training progresses, the model scales up by incrementally increasing the resolution of the generated images,

up to 1024x1024 pixels as proposed by [Karras et al., 2018] in Figure 2. Each layer captures and synthesizes features at different scales.

Initially, both G and D networks are configured to consist of a small number of layers (as low as two convolutional layers) and are gradually upgraded to more complex structures (e.g 14 layers in total) as the resolution increases. The model gradually adds layers to networks (both the generator and discriminator) that introduce higher-resolution details. This technique helps in stabilizing the training process and enables the generation of high-resolution images

The PGGAN employs Leaky ReLU as its activation function throughout most of the network. This choice is motivated by the function's ability to allow small, non-zero gradients when the unit is inactive, which helps prevent neurons from dying out during training—a common problem in deep neural networks [Maas, 2013]. In the final layer of the generator, a linear activation function is utilized to map the learned complex representations to the desired output format.

3.4.1 Normalization Techniques:

- Pixelwise Normalization is implemented following each convolutional layer in the generator. This method normalizes the feature vectors of each pixel to unit length, thereby stabilizing the variance of the features across the network and preventing the scale of the gradients from blowing up.

- Minibatch Standard Deviation is used in the discriminator to incorporate batch diversity. GANs sometimes produce images that are less varied than those in the training set [Saxena and Cao, 2021]. To address this, the proposed technique by [Karras et al., 2018] starts by calculating the standard deviation for each image in the batch. This calculation covers all channels and all pixel values of the image.

  Once the standard deviation for each image is computed, the mean of these standard deviations is calculated across the entire batch. The resulting value, a single scalar, is then duplicated across all examples in the batch and each pixel value. This creates a new channel, which is identical for every image.

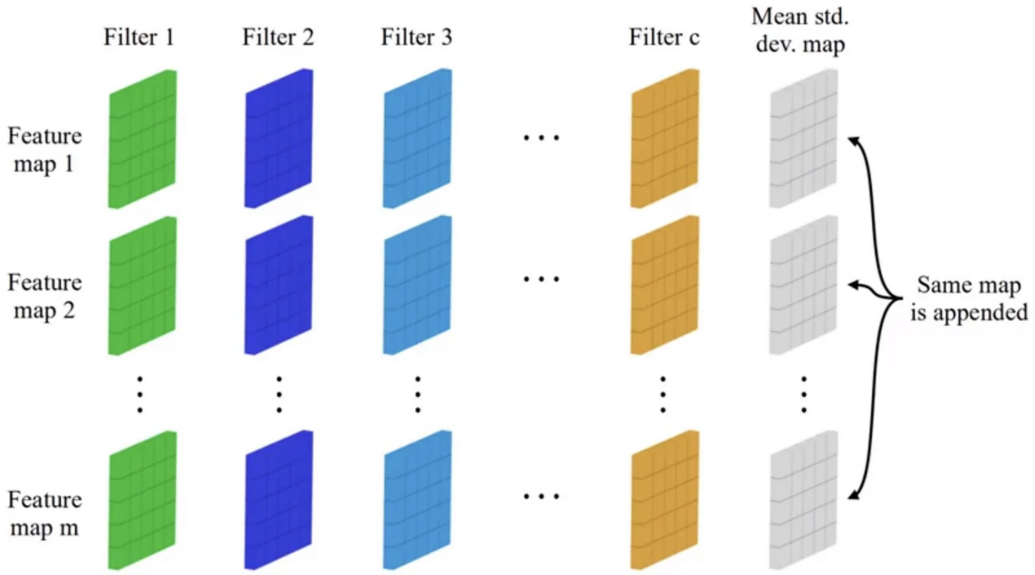


Figure 3: Mean std. dev. map is appended at the end of each output [Deviation, 2022].

Finally, this new channel is appended to the existing channels of each image as shown in Figure 3. This addition helps the discriminator better assess the diversity of the batch, thereby encouraging the generation of more varied images.

3.4.2 Weight Initialization and Scaling

PGGAN's weights are initialized from a centered normal distribution, with a small variance, to maintain a diverse and non-uniform range of initial features. During training, learning rates are equalized across layers by scaling the weights. For example, before performing a convolution with `f` filters of size `(k, k, c)`, the weights of those filters are scaled as shown below:

$$W_f = W_i \times \sqrt{\frac{2}{k \times k \times c}}$$

In this way, they ensure that every weight is in the same dynamic range, and then the learning speed is the same for all weights. Without this scaling, some layers might learn faster and dominate the process, which can lead to a less effective model overall.

The architecture also incorporates advanced features such as skip connections and feature concatenations, which are used to improve the flow of gradients during back-propagation and to enhance feature diversity across layers, respectively. These tech-

niques are particularly valuable in the deeper layers of the network, where they help to maintain the quality and coherence of the generated images [Bau et al., 2018].

The described architectural design enables PGGANs to efficiently generate high-quality, high-resolution images by gradually learning to add finer details while maintaining training stability and model scalability.

3.5 Understanding GANs through Dissection

[Bau et al., 2018] introduces a methodological framework for dissecting GANs, identifying what and how particular 'units' within the generator network influence the generation of complex image features. This process involves segmenting the network into interpretable units that correspond to specific features in the output images, such as trees, doors, etc., and analyzing how these units interact with each other and the wider image context. By manipulating these units, researchers can alter the presence of these features in generated images, improving the quality and versatility of the output.

3.6 Methods

The [Bau et al., 2018] study uses PGGANs to generate images from the LSUN dataset [Yu et al., 2016], which includes various scene categories (e.g. bedrooms, churches, etc.). The paper focuses on dissecting these models to identify and analyze specific units (neurons) within the generator that are responsible for generating discernible objects and features in the images. The researchers employ several methods shown in Figure 4:

(a) Dissection: Identifying which units are responsible for specific features in the generated images. This process begins by mapping the network's units to specific object classes based on their activation patterns across various images. Using a segmentation-based approach, the study identifies units whose activation correlates strongly with particular visual features, such as trees or doors. This correlation is quantified using Intersection over Union (IoU) metrics, comparing the units' activations to labeled segments of images.

(b) Intervention: The next step involves an automated process to manipulate these

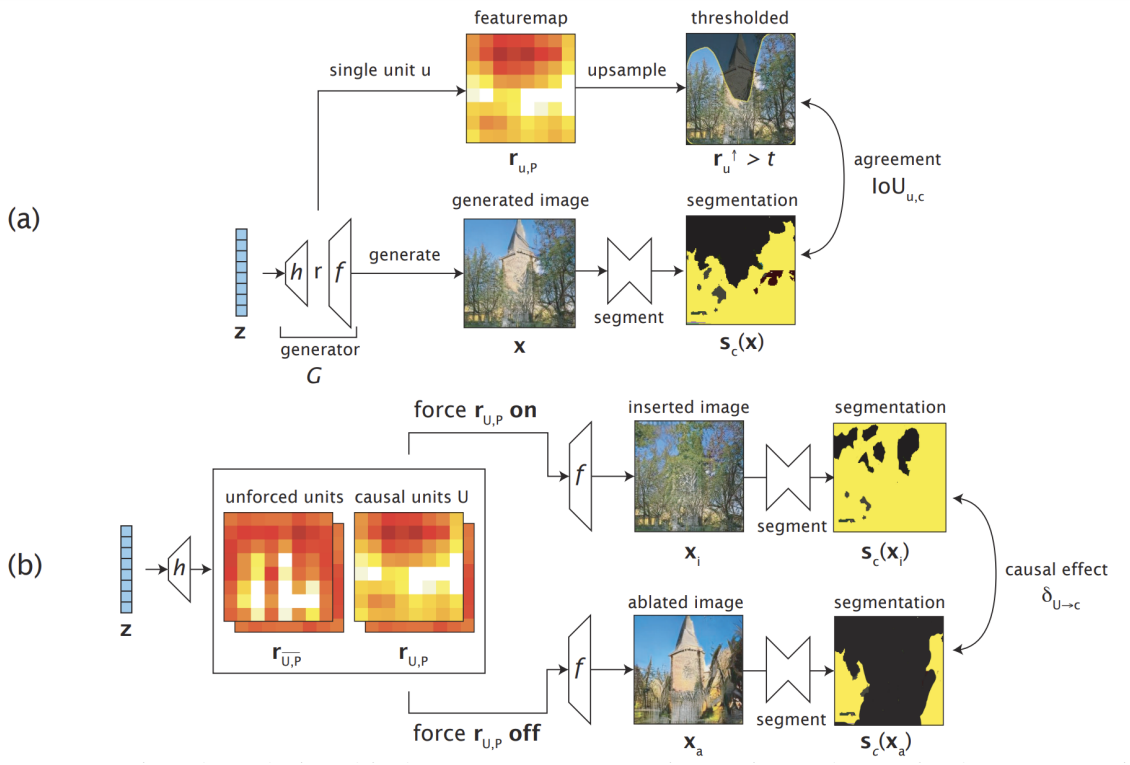


Figure 4: "Measuring the relationship between representation units and trees in the output using (a) dissection and (b) intervention. **Dissection** measures agreement between a unit u and a concept c by comparing its thresholded upsampled heatmap with a semantic segmentation of the generated image $s_c(x)$. **Intervention** measures the causal effect of a set of units U on a concept c by comparing the effect of forcing these units on (unit insertion) and off (unit ablation). The segmentation s_c reveals that trees increase after insertion and decrease after ablation. The average difference in the tree pixels measures the average causal effect. In this figure, interventions are applied to the entire featuremap P, but insertions and ablations can also apply to any subset of pixels $P \subset P$" [Bau et al., 2018].

identified units within the neural network. This is done to assess their causal impact on the presence of objects within the generated images. Using a system that can selectively activate or deactivate certain units, the study explores how specific changes in unit activity affect the visual content of the outputs. The system measures automatically the causality of each unit's contribution to the final image by the presence or absence of the object when the specific units are turned on or off.

(c) Visualization: The units of interest are visualized to understand their influence on the generated images. This involves modifying the units' activations and observ-

ing the changes in the output images, which provides insights into how different features are synthesized by the network, as shown in Figure 5.

Figure 5: ": Comparing representations learned by progressive GANs trained on different scene types. The units that emerge match objects that commonly appear in the scene type: seats in conference rooms and stoves in kitchens. Units from layer 4 are shown. A unit is counted as a class predictor if it matches a supervised segmentation class with pixel accuracy > 0.75 and IoU > 0.05 when upsampled and thresholded. The distribution of units over classes is shown in the right column " [Bau et al., 2018].

3.6.1 IoU metric

Let $r_{u,P}$ denote the one-channel $h \times w$ feature map of unit u in a convolutional generator, where $h \times w$ is typically smaller than the image size. We are interested in whether a specific unit $r_{u,P}$ encodes a semantic class such as "tree". For image classification networks, [Bau et al., 2018] observed that many units can approximately locate emergent object classes when the units are upsampled and thresholded.

In this spirit, we select a universe of concepts $c \in C$ for which we have a semantic segmentation $s_c(x)$ for each class. We then quantify the spatial agreement between the unit u's thresholded feature map and a concept c's segmentation with the following intersection-over-union (IoU) measure:

$$IoU_{u,c} \equiv \frac{E_z\left[(r^{\uparrow}_{u,P} > t_{u,c}) \wedge s_c(x)\right]}{E_z\left[(r^{\uparrow}_{u,P} > t_{u,c}) \vee s_c(x)\right]}, \text{ where } t_{u,c} = \arg\max_t \frac{I(r^{\uparrow}_{u,P} > t; s_c(x))}{H(r^{\uparrow}_{u,P} > t, s_c(x))},$$

and where $\wedge$ and $\vee$ denote the intersection and union operations, respectively, and $x = G(z)$ denotes the image generated from z. The one-channel feature map $r_{u,P}$ slices the entire feature map $r = h(z)$ at unit u. As shown in Figure 8, we upsample $r_{u,P}$ to the output image resolution as $r^{\uparrow}_{u,P}$.

The expression $(r^{\uparrow}_{u,P} > t_{u,c})$ produces a binary mask by thresholding the $r^{\uparrow}_{u,P}$ at a fixed level $t_{u,c}$. $s_c(x)$ is a binary mask where each pixel indicates the presence of class c in the generated image x. The threshold $t_{u,c}$ is chosen to be as informative as possible by maximizing the information quality ratio I/H (using a separate validation set); that is, it maximizes the portion of the joint entropy H, which is mutual information I [Wijaya et al., 2017].

This methodological approach contributes to the understanding of GANs and offers a systematic framework to dissect and manipulate the neural representations within these networks. By clarifying how certain units within a GAN influence specific elements of the output images, this research paves the way for more interpretable and controllable generative models.

4 XGANArt

In this section, we will describe the finished interface, starting with the theory that informed the decisions made regarding the final design. Section 4.1 describes the XAI Question Bank and how it helped us point out which user question types we are trying to answer, followed by section 4.2, which describes how the 'Explore' interaction pattern is used to answer said questions. Finally, section 4.3 will present the implemented features of our interface.

As our project is centered around explainable AI (XAI), our interface is named XGANArt — Explainable GANArt. This name reflects our goal of creating an interface that allows the user to be creative with the power of GANs while offering explainability. The idea for this project is centered around image generation but is not tied to a specific context. In other words; there was little focus on a user group with specific needs in regards to a situation they might use the interface in - such as work or school. However, still a user group with a technical background. The chosen approach is, therefore, a

more general one, aiming to provide the user with a basic and high-level understanding of GANs and neural networks. By testing the effect of our approach in regards to explainability, we can identify features and ideas that proved effective as well as areas of improvement, opening up further research for this approach to XAI.

The GANDissect project in its entirety is the code, the connected paper, as well as a website, which explains the project and GANs. A demo (GANPaint) is hosted, which allows the user to use the results of a dissected project to manipulate an image by activating a set of neurons in a GAN. The version that can be locally hosted from the source code also has an 'expert'-mode. The expert mode allows the user to change the current selected layer to draw on but offers no explanation about the inner workings of the model - such details can be found in either the source code or in the material on their website, for example their published paper.

XGANArt is specifically an expansion of the GANPaint demo, utilizing the already existing interface, which lets people draw and add to that. When expanding upon the demo, we face two challenges in regards to XAI. The first is the question of user needs - who is this built for and why? What needs does the user have? The second is how one can go about actually meeting the user's needs. Specifically, this question has garnered a lot of attention over the last few years when it comes to XAI [Liao et al., 2020]. After all, figuring out the needs of a user is pointless if one doesn't attempt to figure out how to meet those needs as well.

4.1 XAI Question Bank

[Liao et al., 2020] tackles the challenge of user needs for explainability by approaching it as presenting answers to the questions that users might have. Their study presents an XAI Question Bank, as well as XAI methods mapped to user type questions, which provides insight into the user need for different types of explainability for AI and ideas on how to address these needs.

The XAI question bank and the XAI methods were used as inspiration for XGANArt. The decision of which of the XAI explanation methods to use as inspiration was influenced by several things. As we are to expose the inner workings of a neural network to the user, we have the opportunity to give a high-level explanation of the model,

before narrowing down focus to be on the hidden layers and the individual predictions within.

The chosen XAI explanation methods can be seen highlighted in Figure 6. It should be noted that these were used as inspiration, not as strict guidelines to be followed.

Category of Methods	Explanation Method	Definition	Algorithm Examples	Question Type
Explain the model (**Global**)	Global feature importance	Describe the weights of features used by the model (including visualization that shows the weights of features)	[41, 60, 69, 90]	**How**
	Decision tree approximation	Approximate the model to an interpretable decision-tree	[11, 47, 52]	**How**, Why, Why not, What if
	Rule extraction	Approximate the model to a set of rules, e.g., if-then rules	[26, 93, 102]	**How**, Why, Why not, What if
Explain a prediction (**Local**)	Local feature importance and saliency method	Show how features of the instance contribute to the model's prediction (including causes in parts of an image or text)	[61, 74, 83, 85, 101]	**Why**
	Local rules or trees	Describe the rules or a decision-tree path that the instance fits to guarantee the prediction	[39, 75, 99]	**Why**, **How to still be this**
Inspect counterfactual	Feature influence or relevance method	Show how the prediction changes corresponding to changes of a feature (often in a visualization format)	[8, 33, 36, 51]	**What if**, How to be that, How to still be this
	Contrastive or counterfactual features	Describe the feature(s) that will change the prediction if perturbed, absent or present	[27, 91, 100]	**Why**, **Why not**, **How to be that**
Example based	Prototypical or representative examples	Provide example(s) similar to the instance and with the same record as the prediction	[13, 48, 50]	**Why**, How to still be this
	Counterfactual example	Provide example(s) with small differences from the instance but with a different record from the prediction	[37, 55, 66]	**Why**, **Why not**, How to be that

Figure 6: The XAI methods mapped to user question types, with relevant explanation methods highlighted [Liao et al., 2020].

Each explanation method maps to certain question types. Section 4.3 describes how the design of our interface is aimed at explaining the relevant question types seen in Figure 6. Question types such as "How to be that" and "How to still be this" require an explanation of the decision boundaries of the model. What causes a prediction to change, and how will it change? In our case, since example-based explanations were recommended in the paper, we decided to answer these questions by utilizing the features of GANPaint.

4.2 Explore Pattern

[Grabe et al., 2022] presents four human-AI interaction patterns to support co-creativity. Of the four, the 'Explore'-pattern is the one that best describes GANPaint, as GANPaint allows users to adapt the artifacts that have been created by sampling from the GAN.

By adapting the artifact, the user can explore the latent space - and this can be done in different directions, where each direction represents a different design. GANDissect identifies semantic features in the model, and allows the user to draw with the features.

In this manner, GANPaint allows for exploration in the latent space with regards to the semantic features.

As the user explores the latent space, they are essentially exploring the decision boundaries of the model. By drawing with different features in different places on the image, the user can gain an understanding of how their actions change the image. They'll also be able to see which changes are required of them for the model output the desired result, or which changes they can make before the image changes in an unexpected manner. An example of this is the situation where the user attempts to draw a door in the sky, but the model is unable to draw the door at such a location. The user can then try to draw doors in different places in order to figure out where the feature can be successfully drawn.

4.3 Design

It should be noted that some decisions in regard to design were affected by technical limitations, which are described in detail in section 5.

The inner workings of a GAN are intricate and complex, and trying to explain a GAN in its entirety can easily lead to 'information-overload' for the recipient. In order to narrow the scope of project, the interface focuses on showing and explaining how the changes by the user affect the artifact, and does not explain how the model was trained.

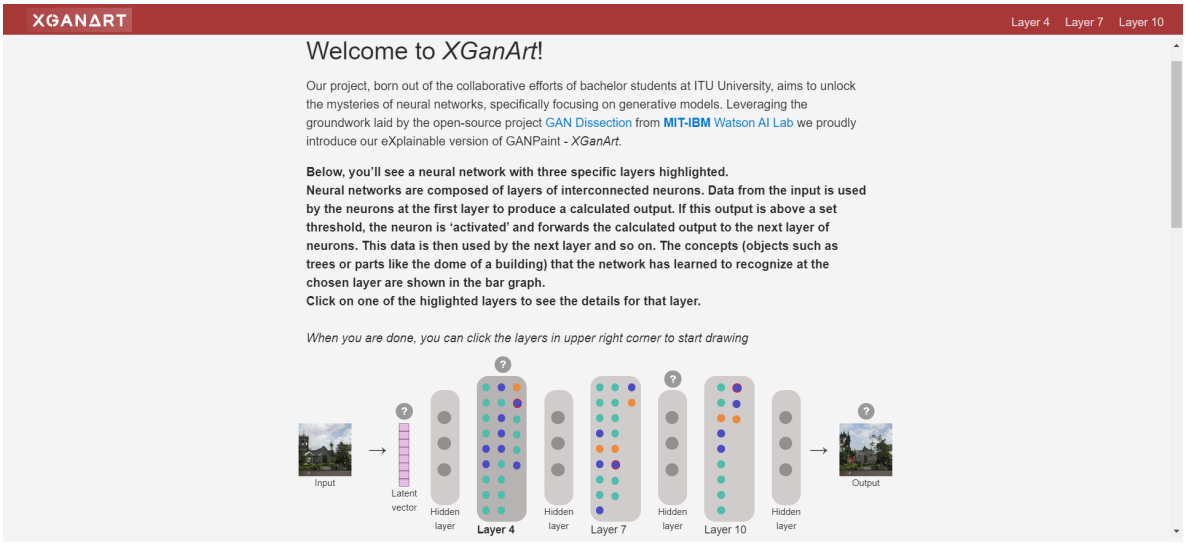


Figure 7: The front page visuals.

Our interface incorporates visual elements as well as interactive elements to further the user's understanding. The front page presents the user with a short text about the functionality of a neural network, as well as a visualization of the network (Figure 7). The hidden layers of the network are shown, with three specific layers highlighted (Figure 8.)

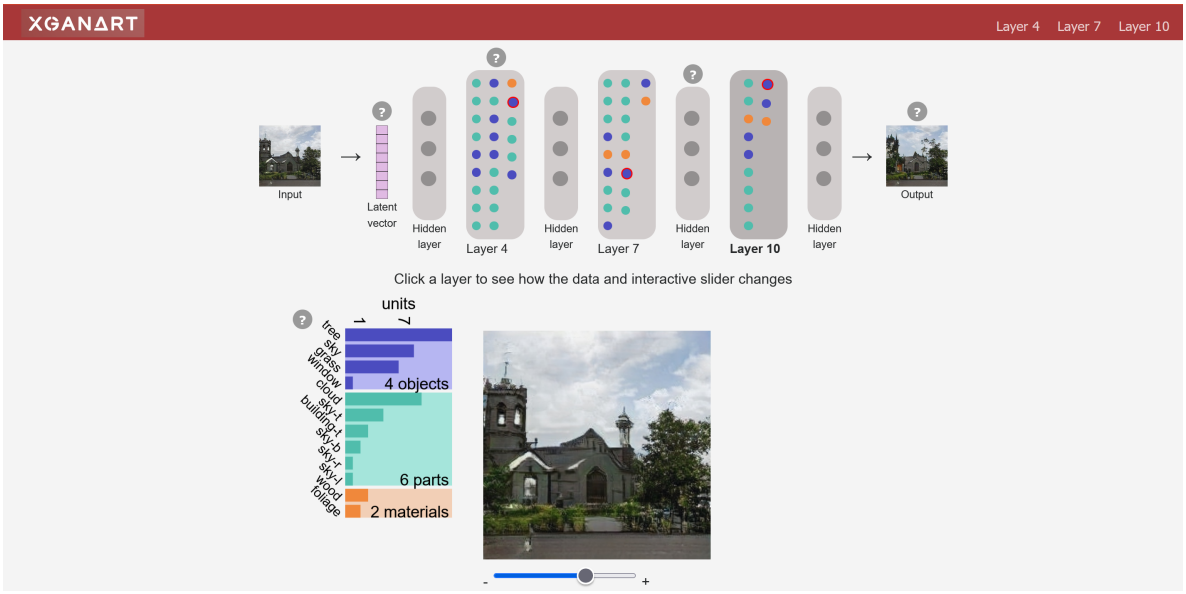


Figure 8: The hidden layers visualization.

Dissection was performed on these layers, and the identified semantic features are presented to the user in both the figure of the network as well as the connected bar graph. The user can choose one of the highlighted layers to see the results for that layer. Each dot (Figure 9) in a layer represents a semantic feature, also represented in the bar graph. When a user hovers over any of the dots in the layers, the name of that specific feature is shown in a tool-tip above the dot.

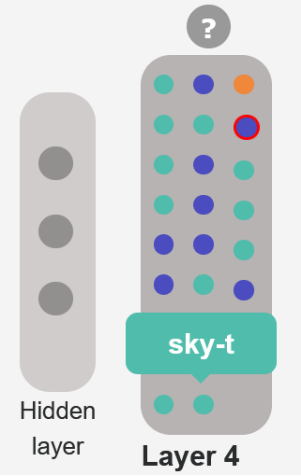


Figure 9: The feature 'sky-t' represented for layer 4.

The user is also presented with an image with a slider (Figure 10). The slider can be used to manipulate the 'tree'-feature, by altering the neurons associated with that feature in the chosen layer. This will be explained further in Section 5. This tool allows for real-time visual feedback on how changes in neuron activations affect the generated images.

If the user desires more in-depth explanations of the specific elements on the site, they can hover over the question marks (Figure 11). For further explanation of the technical background behind the project, the user can also use the collapsible placed at the bottom of the front page. These contain text explaining the different technical areas of the project.

Of the question types from the XAI question bank, the components of the front page seek to explain the inner workings of the neural network (How), using text, visuals and the interactive slider. These also contain high-level explanations for how the neurons in the network are activated and what effect that has on the resulting artifact (Why, Why not).

From the front page, the user can navigate to a page associated with one of the highlighted layers (layer 4, 7 or 10 - see figure 12), where they will be presented with a page closely resembling GANPaint, but with some changes. Firstly, expert-mode was enabled to show the user the history of changes they've made on the image.

The features available for drawing now reflect the features identified on the specific layer. This was changed from the original expert-mode that only presented the user

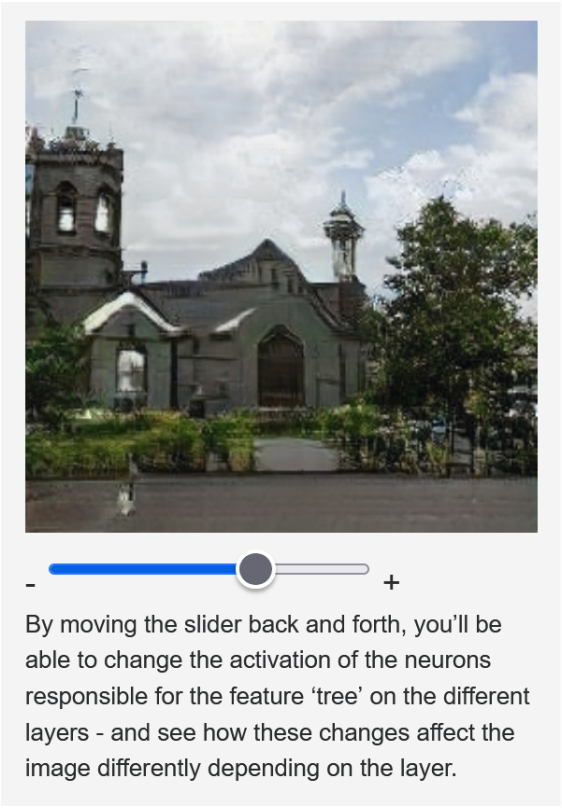


Figure 10: Activation Slider

with a small set of pre-chosen features. Some of the identified features only had a few neurons associated, and could therefore be described as having a weak presence on that layer. They were therefore left out in order not to clutter the page.

To enhance user experience and intuitive interaction with XGANArt, we have implemented a color-coding scheme that categorizes the features available for drawing based on their roles as "objects," "parts," or "materials." The color of each category corresponds to the colors seen in the bar graph on the front page. This categorization not only simplifies the selection process for users but also aids in quickly identifying the type of features they intend to manipulate.

To relate this back to the XAI Question bank, the painting interface changes the image artifact sampled from the model when the user makes a change in the activations of their chosen feature by painting, answering the question types of "What if", "How to be that", "How to still be this", as explained in section 4.2.

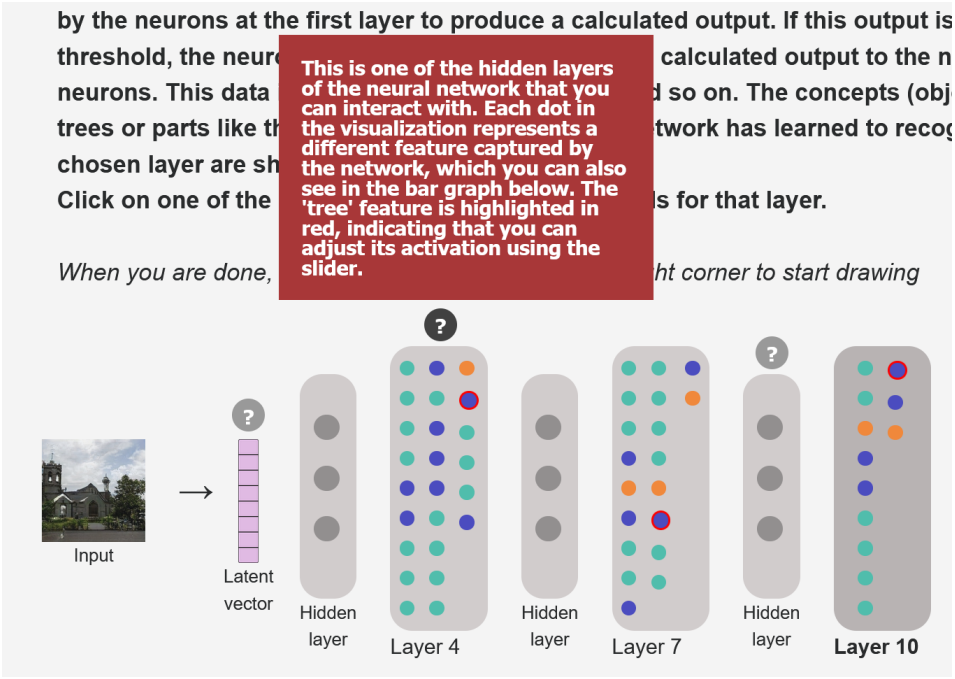


Figure 11: Tool-tips explanations

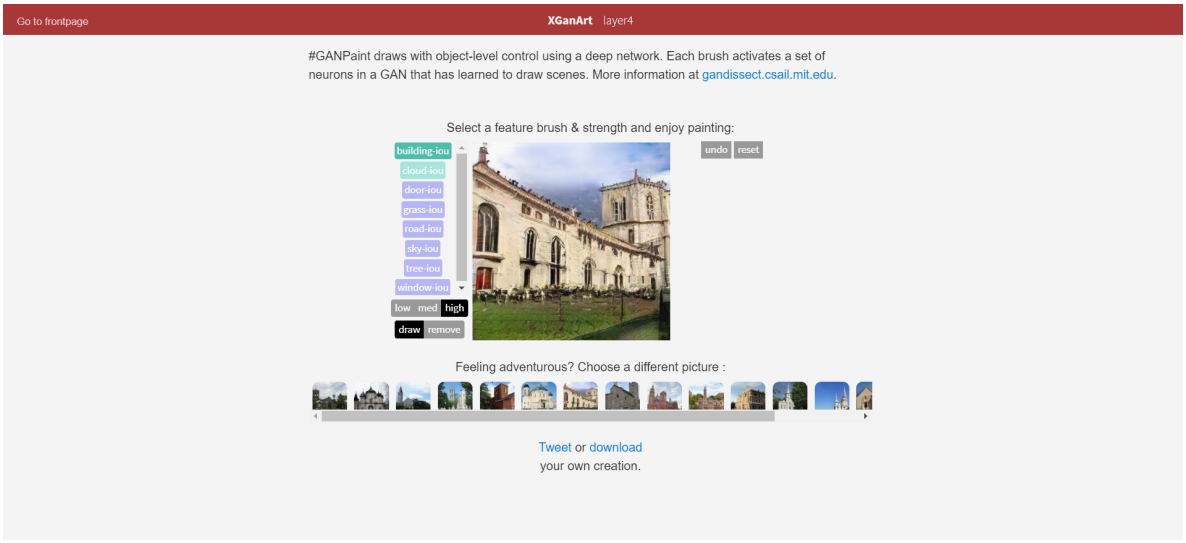


Figure 12: The drawing application - layer 4

Screenshots of the whole interface are available in the appendix A.1

5 Implementations

The XGANArt application is built upon the GANDissect-project, which is open source code and is available on Github. Due to complications running the code caused by

compiler issues, XGANArt is set up to run in a docker container.

The GANDissect project has a rather comprehensive tech stack, but the relevant things for our project were CUDA, Python, Javascript, HTML, and CSS. Our front end builds upon the existing GANDissect front end and is therefore set up using Javascript, HTML, and CSS, as changing to another framework would be time-consuming, and no specific benefits a framework might offer us were considered important enough. As we expanded the software, we wanted to follow the architectural pattern already present in the GANDissect project—the Model View Controller pattern.

5.1 Model View Controller Pattern

Architectural patterns are a way to reuse knowledge about software systems - specifically, how to set up the software. A specific pattern can provide certain benefits, such as better performance or security, maintainability, etc. The Model View Controller (MVC) pattern is one such architectural patterns.

The MVC pattern separates the web application into three components: model, view, and controller - Each component has its own responsibilities. In this way, the system data can be separated from the parts of the system aimed at presentation and interactions. There are some differences in the ways communication takes place in this pattern, depending on the implementation [Sommerville, 2016][Fowley, 2011]. One way is as follows; The controller is responsible for handling requests from the client, and communicates with model and view. The model is responsible for data logic and interacting with the system data. If the client requests something from the data, the controller must forward this request to the model, which then gathers the data. The model may manipulate the data before returning the result to the controller. The controller then uses the view to create a visual representation of the data, which is afterward shown to the client [Kumar, 2021]. The MVC pattern has the benefit of "separation of concerns", with changes in one of the components not interfering with the others.

The application of the XGANArt project follows the MVC pattern. See Figure 13. The View is composed of HTML, Javascript and CSS, while the controller is the Python-file which launches the applicaton (server.py located in the netdissect directory).

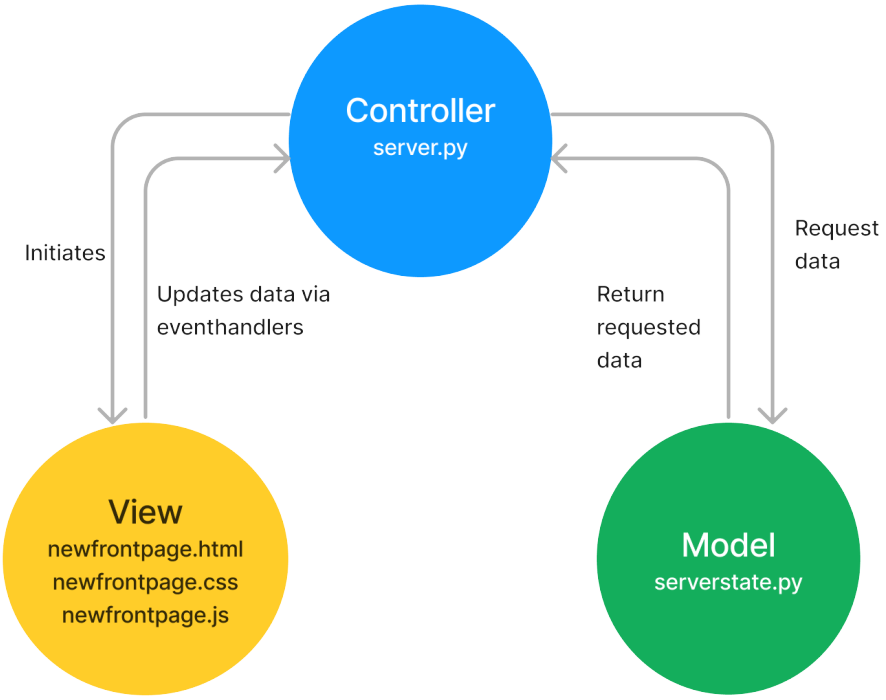


Figure 13: The MVC pattern as implemented in XGANArt.

Serverstate.py, also located in netdissect, is the model that accesses the data from previous dissections, which are stored in the dissect folder. When the client accesses the front page of XGANArt, the Javascript (view) sends an API request to the server (controller), requesting information from the dissection of the "churchoutdoor" project. Serverstate (model) is asked by the server to fetch and return the data. This data is then presented to the user by the View as a representation of the neural network, as well as showing the bar graphs for each layer.

The image that the user can draw on was implemented in the class "GanPaintView" in the GANDissect-project. The same class is used for the slider, with minor modifications that cause the whole image to be changed depending on the value of the slider - as opposed to only changing where the user draws on the image. When the user changes the value of the slider, this value is added to the values of the top ten neurons responsible for the "tree" feature. The image is then processed with these values, and the newly generated image is returned and shown to the user.

5.2 Features for Drawing

The drawing interface of XGANArt presents the user with some, but not all, features identified at their chosen layer which they can use to draw. As described in section 4.3, some features had a weaker presence on a layer and were left out of the available selection of features. The collection of features was sorted according to how many neurons were responsible for each feature. Then, the harmonic mean of the collection was calculated, and the features with an associated number of neurons below the harmonic mean were left out. For each layer there were often a very small amount of features that had large amounts of neurons, with the other features having low amounts. Using the average, therefore, proved to remove more features than we would like. The harmonic mean was used instead to counter this, as large outliers will have less impact on the result [Brownlee, 2020].

5.3 Limitations

'Code smells' is a term often used to describe code that is written in such a way that makes the code difficult to work with and may (if it hasn't already) introduce problems down the line [Martin Fowler, 2018]. What makes codes 'smelly' is considered somewhat subjective, but certain things are found to often contribute to the smell. Examples of this are duplicated code and long methods. The code of the GANDissect project could easily be described as smelly. It contains huge swathes of out-commented code, methods that are declared and functional but are never used, as well as multiple instances of duplicate code. Code smells can be dealt with by refactoring the code. However, refactoring is not a small or quick task, especially for a project of this size. For XGANArt the decision was made to refactor only small amounts of the code, as creating the actual product had higher priority.

The decision was also influenced by one other thing; In the project, there exist two folders, 'client' and 'client_dist.' The files in 'client', such as the Typescript files, have been transpiled into the files found in 'client_dist', in order to create the web application. Ideally, changes to the application would have been implemented in the files in 'client' and then re-transpiled. However, many of the dependencies required re-transpile the files from 'client' to 'client_dist' are deprecated. One of these dependencies is a GitHub

repository, which no longer exist. As such, the decision was made to directly implement changes in 'client_dist'.

5.4 Repository Setup and Data Handling Limitations

The XGANArt interface repository is hosted on GitHub. It can accessed by visiting the following URL: GitHub Repository Link. This repository contains all necessary files and instructions for setting up and using the XGANArt interface. To ensure smooth operation, please follow the detailed steps provided in the README file.

Due to the substantial size of the data folders used in this project, it is not feasible to include them directly in this repository. In order to get the data necessary for hosting XGANArt, follow the instructions in the README to download the datasets and run the dissection on the 'churchoutdoor' project.

6 User Testing

We performed user testing in order to test how well our interface offered improved explainability and to identify possible areas of improvement. The type of user testing that we used is known as qualitative user testing which specifically focuses on gathering insights of a user's experience with an interface. In turn, this can reveal both potential issues with an interface and also opportunities for future work [Moran, 2019].

To conduct the test, we had a user group of seven people that performed in-person testing. All interview participants were students from the IT University of Copenhagen. Doing user testing with this user group, who all have a technical background, would therefore provide insight into what the user's previous experiences were with AI and whether their experience could be enhanced by using XGANArt. Specially this would be done by comparing the user's explanation of what was going on from a technical point of view at the beginning of the user test before using XGANArt, to their understanding at the end of the user test after using XGANArt. This would also show whether XGANArt had managed to answer the 'how' question from the XAI question bank mentioned earlier in section 4.1. In other words, this would show whether XGANArt successfully explains how the underlying machine learning model works. To see

the full script of the user test see B.1.

Each person consented to both screen recording and voice recording of their session. The tests started with the following questions about the user's background:

- How old are you?
- What are you currently studying? (study program - bachelor/master's - semester)
- Have you used generative AI before e.g. ChatGPT, DALL-E, Midjourney or similar?
- If yes, how many times do you use it during a week (for instance last week)?
- What do you use it for?

They were then asked the following questions and asked to rate their answers on a scale from one to five:

- What is your knowledge of machine learning/AI? Where 1 is "I don't know what it is", and 5 is "I have taken a course/researched the technical aspects of machine learning myself"
- Are you interested in understanding how AI/machine learning works? Where 1 is "No, not at all" and 5 is "I am very interested in this topic"

After the introduction questions, the user was presented with various tasks to perform while thinking out loud. The first tasks B.1.2 were concerning the original GAN-Paint interface. Their tasks here were both to explain their overall impression of the site and also to explain what they believed was happening behind the scenes.

The user was then presented with the front page of XGANArt B.1.3, where they were asked to think aloud as they formed their first impression of the site and later when they used the interactive features on the site. Our intention here was to let the user play around with the site and figure out the interactive features by themselves. This was done to see how obvious it would be for the user that the components were interactive. The front page had the intention of explaining the inner workings of the

GAN model. As such, asking the user to try to explain what was going on when using the hidden layers graph, slider and the text presented throughout the page, was done to see how well our interface had answered the question types 'how', 'why' and 'why not' from the XAI question bank, as these aim to explain the global model.

We then asked the user to navigate to layer 4, the painting interface of XGANArt A.2, share their first impression and play with the painting functionality. This would indicate how well our interface answers the XAI question bank questions relating to a user's understanding of a local instance of the model: 'How to be that', 'how to still be this' and 'what if'. The user was then asked to do the same for layer 10, with the intention of seeing if the user noticed a difference when working on the different layers. By doing this, we aimed to utilize the XAI question bank category 'Inspect counterfactual' as the user would be presented with a contrasting part of the model compared to what they saw in layer 4. The user was then asked to explain what they believed was happening behind the scenes, just as they were asked earlier when using GANPaint, to see if their understanding had changed.

Lastly, the user was asked wrap-up questions about their overall experience, and whether they had any advice or improvement, they would like to add. Some of the participants gave us some remarks about the user test itself after they had participated, which meant that we changed the user test a bit for the last participant. The previous participants would have liked more time to sit with the application and felt they would have had an easier time going over the application and absorbing the presented information if they were sitting alone. Therefore, the user 7 was left alone in a room to sit and was given more time to look through XGANArt.

6.1 Results

6.1.1 Front page

When doing tasks on the front page, the users generally did not read the text before they started playing with the interactive features on the site.

A recurring issue for the users was not realising the click-ability of many of the functions on the front page. For the neural networks graph, four out of seven users did not realise that it was clickable, explaining that it was not obvious to them that they could switch

User Nr.	Age	Programme of Study	Semester	Knowledge of Machine Learning (1-5)	Interest in Machine Learning (1-5)
User 1	23	BSc Data Science	2nd	1	4.5
User 2	21		6th	3	2
User 3	20		4th	5	5
User 4	21		4th	5	5
User 5	25	MSc Computer Science	2nd	4	5
User 6	23	BSc in Software Development	6th	5	5
User 7	24		6th	5	5

Table 1: Summary of Users' ML Knowledge and Interest

between the different layers. After being asked to do tasks where they had to switch between different layers in the graph and use the slider, six out of seven users pointed out that they would have liked every 'dot' inside of the layers to be clickable and for it to decide which feature the slider was inserting or ablating.

Overall, five out of seven of the users who had rated their machine learning knowledge as four or higher immediately understood the neural network graph and could explain each part of it. The users that had rated their machine learning knowledge below four did not know what the neural network graph represented. Since the users did not intuitively read the text presented above the graph, they did not gain an understanding of the neural network graph throughout the user test in terms of understanding the global model. This is concluded based on the fact that these users could not explain what was going on from a technical point of view at the beginning of the user test, and their explanations had not changed by the end of the test. During the usability test, only four out of the seven users utilized the tool-tip question marks placed around the page. They pointed out that the colour of this functionality was too similar to the hidden layers and should be highlighted more for them to notice and use it.

6.1.2 XGANArt Painting Interface

When asked to do tasks on the XGANArt painting interface, the users all related it to the GANPaint interface they had seen earlier, but they also all noticed differences between the two interfaces. The first impression of five out of seven users was to point out that they liked the new colour scheme of the interface. When asked about the

left-side buttons of the interface, all the users noticed the new colours, but none of the users saw the relation between the colour scheme and the feature-categories in the bar graphs on the front page. Three users also mentioned that they did not understand the meaning of 'iou' that was now added to each button. This briefly confused some users when they were asked to draw a tree, as they could only find 'tree-iou'. One user recommended that we implement the tool-tip question marks from the front page to explain both the meaning behind the buttons' colours and the meaning of '-iou'. Overall, the users liked the new functionalities of the XGANArt painting interface such as the ability to see the history of what you had drawn as well as being able to paint on multiple different layers.

6.1.3 General Understanding

When asked to explain what they believed was going on behind the scenes when doing tasks on the original GANPaint, five out of seven participants related the experience to a PhotoShop-like experience. When asked to elaborate on this, the users believed that they were drawing on top of the existing base-image with a picture of a feature, for example, a picture of a tree. The other two users understood that they were drawing with features trained by a neural network instead of references for images. In common for both groups was that none of them related GANPaint to being a specific layer of the Machine Learning model.

When asked the same question at the end of the user testing, after they had experimented with XGANArt, all seven users noticed a difference between drawing after using the navigation buttons 'layer 4' and 'layer 10'. The description of this difference varied from user to user with the most common explanation was that layer 4 seemed to draw with 'whole objects', while layer 10 seemed to draw 'textures and colours' that looked like a filter over the existing image. This indicates that all users gained an understanding of the local features of the model, even if some of them did not gain an understanding of the global model.

However, noticing a difference between drawing on the different layers did not mean that every user understood what a layer was. Four out of the five users who had rated their knowledge of Machine Learning at four or above understood that clicking the layer

navigation buttons redirected them to a representation of a specific layer in a neural network. The rest of the users did not understand the neural network graph and did not understand that the difference between painting on layer 4 and painting on layer 10 was a result of different layers in a neural network.

7 Discussion

This project aimed to enhance the transparency and explainability of generative models, particularly Generative Adversarial Networks (GANs) by developing the interface XGANArt, which exposes the hidden layers of a model. Our goal was to increase the transparency of the GANDissect model to make it more accessible and understandable to users.

To do this, we developed the interface XGANArt based on the open-source code project GANPaint by the MIT-IBM Watson AI Lab. XGANArt enables users to interact with different layers of the GAN and includes an interactive front page that helps users comprehend the technical workings of the underlying model. By engaging users with the layers directly involved in image generation, XGANArt aimed to provide a hands-on understanding of the model's internal processes.

In this discussion, we will discuss whether we succeeded in the above as well as the influence of the processes used throughout the project.

7.1 Enhancing User Comprehension of GANs through the XGANArt Interface

Through the results of the conducted user test, we indicate that the implementation of our interface enhances the explainability and understanding of GANs. Users showed a higher understanding of the inner workings of the ML model when interacting with the XGANArt interface, based on the fact that they were able to explain the technical workings of the model with higher accuracy after using XGANArt.

The majority of the participants had little to no understanding of the original GANPaint application, believing at first that it was an image-editing tool, e.g., PhotoShop. The users believed that they were drawing on top of the existing base image with a

feature that was blended into the background. After experiencing and interacting with XGANArt all of the users showed a noticeable improved understanding of the inner workings of the model. The most important perception showcased was the understanding of the role of different layers in image generation. Users quickly noticed the differences in drawing on different layers, having a high-level understanding that early layers, like layer 4, draw "with whole objects", while later layers, such as layer 10, allow you to draw with "textures and color" as explained by user 3 (see appendix B.5).
Judging by the users' hesitance to read through the written explanations at the top of the front page and by how their attention was focused mostly on the visual and interactive components, it could be argued that incorporating further interactive elements would benefit user understanding. Given that the painting application follows the 'Explore'-pattern, letting the user learn through exploration, it would make sense to incorporate elements, such as interactive components, that the users are naturally drawn to and are willing to engage with without being told to.

7.2 Relevance of the results

Our research on the XGANArt interface implemented the existing approaches in the field of XAI to enhance user interaction with Generative Adversarial Networks (GANs). This interface aimed to address critical concerns about the 'black box' nature of such technologies by making the internal mechanisms of GANs more understandable. This initiative aligns with the field of explainable AI (XAI), which aims to bridge the gap between advanced technological processes and user comprehension.

Previous literature has emphasized the importance of transparency in AI systems, highlighting that a better understanding of AI processes is essential for human-AI interaction. For instance, studies by Bau et al. (2018) and Olah (2022) have discussed how exposing the inner workings of neural networks can demystify AI decisions, making these systems more palatable and accountable to users. By allowing users to interact with and visualize different layers of a GAN, the XGANArt interface directly contributes to these objectives, enabling users to see firsthand how alterations in input parameters can affect output in real time, which enhances their understanding of AI systems.

Moreover, the XGANArt interface not only aims to enhance user interaction with

GANs but also delves into the potential of AI tools to support creative and collaborative endeavors. By providing a tool that allows for the manipulation of neural network layers and observing the results, XGANArt aims to encourage further research into designing for XAI, potentially leading to new uses of GANs across various domains. This aspect of our research highlights its relevance and potential impact in fostering human-AI co-creativity by allowing a user to switch between various layers of a GAN to explore the differences in co-creative output at different stages of a neural network.

Additionally, our findings also contribute to discussions around the usability and design of AI interfaces. The variation in usability perceptions based on user background (technical vs. non-technical) highlights the need for designing AI tools that are accessible to a diverse range of users. This insight is vital for the development of future AI applications that aim to reach a broader audience, ensuring that AI benefits are not confined to technically proficient users but are available to all segments of society.

7.3 What makes a good user test?

However, it is also important to point out the possible faults in the user test that was conducted. Since the user test was conducted together with a facilitator, users might feel a need to rush things compared to if they were doing things by themselves. By asking a user to complete specific tasks, the way the user explores the application could be seen as artificial and, therefore, not offer the same results compared to if the user explored the application naturally. There might also be a difference in how a user explores an application when they are participating in a user study compared to if they naturally decided to use the application.

Moreover, we consider the time constraints of our user test, which can often limit the depth and breadth of feedback that can be gathered during user testing. In the case of XGANArt, the restricted duration might have prevented users from exploring all features thoroughly, leading to a potentially superficial understanding of the interface. This could skew the test results towards initial impressions rather than a comprehensive evaluation.

The participant (User 7) who got more time to look over XGANArt and sat alone in a room also remarked that they would still have liked more time to look over the

interface.

It could also be pointed out that the introduction questions asked at the beginning of the user test could lead to inaccurate answers. These questions are all based on self-assessment since the users must rate themselves on a scale of one to five on various parameters. A user's self-assessment of understanding may not always align accurately with their actual knowledge compared to results obtained through a structured and official understanding test. Such formal evaluations provide a more reliable basis for comparing user outcomes, offering a standardized measure of comprehension that can highlight discrepancies between perceived and actual understanding.

Additionally, the predominance of participants with a technical background in ML introduces further bias, particularly influencing the perceived ease of use and the educational value of the interface. Such participants are likely to find the interface more intuitive and less challenging compared to users with a low understanding of ML. As most of the test users are already somewhat familiar with the concepts underlying generative models, this discrepancy can lead to an overestimation of the interface's usability and effectiveness.

Lastly, it must be mentioned that the user testing was conducted during the last two weeks of the project. As such, the user test has not been used to further develop the application but primarily to gain an understanding of the current state of the interface. Here, one could argue that one or multiple spaced-out user tests earlier in the process could have iteratively improved the application based on the users' input.

7.4 Strengths and Weaknesses of our project

In addition to the insights gained in the above section for what makes a good user test, it is also important to consider other aspects of the project with the same constructive criticism. Our research into the XGANArt interface has both notable strengths and limitations that influenced its outcomes and the validity of our gained insights.

Strengths:

1. Approach: One of the primary strengths of our research lies in its approach to enhancing the explainability and transparency of GANs through an interactive

interface. This approach is based on methods from the field of XAI, which aims to provide a better understanding of the often opaque processes within AI systems. Specifically these methods include the XAI question bank and the "Explore"-pattern.

2. Practical Application: The development of the XGANArt interface offers practical, hands-on experience with complex AI models. This approach not only aids in understanding but also engages users in a way that theoretical explanations cannot, making the learning process both engaging and effective.

3. Enhanced User Engagement: By allowing users to interact directly with the layers of GANs and observe the impacts of their modifications in real time, the research facilitates a deeper understanding and a more intuitive grasp of these complex models. This direct engagement is important for such explainability tools, especially in fields where abstract concepts can be challenging to convey.

Limitations:

1. Source code limitations: A lot of time was lost struggling with technical issues when setting up the project at the beginning. As the GANDissect project is 4 years old, many of its dependencies were deprecated. Furthermore, there were issues with the compiler, which forced us to spend time setting up the project within a Docker container instead.

2. Technical Limitations: The interface currently only allows interaction with certain layers and neurons of the GAN, which may not fully represent the model's capabilities. Expanding the range of accessible layers and neurons could provide a more comprehensive tool and richer insights into the model's operation. However, the full model used in this project consists of thirteen layers. To include all thirteen layers in the interface would require running a dissection of each layer of the model. Each dissection of a layer requires multiple gigabytes of free disk space and can take several hours to run, depending on the GPU running it. As such, it was decided that a full dissection of the model was beyond the scope of this project. This incident highlights the issue of the growing complexity of these

'black-box' models for the field of XAI. As the models become more complex, the task of giving an explanation of a model in its entirety also becomes more complex and demanding.

As this project would not dissect the full model, it was instead decided to dissect three layers from both earlier and later stages of the neural network. As such the project should still show a clear difference between the categorization of the features of the layers. Ideally, however, the project would contain all layers of the model to allow for the most accurate representation of the neural network.

3. Feedback Implementation: During our project, we did not include a cycle for implementing user feedback to iteratively improve the tool. Including such a feedback loop in future studies could enhance the tool's design and functionality based on direct user experiences. If the user test had been conducted some weeks earlier, the first implementation phase would include the most frequently given feedback from the participants, as will be described in the following. The first changes would consist of changes to the styling to highlight the functionality of certain existing elements. At the front page, we would make it clear that the layers in the hidden layer graph is clickable. This would be done by dynamically darkening the colour of a layer elements when a user hovers over it. We would also add a text element saying 'click me' with an arrow pointing to one of the layers as well as a drop shadow border to the currently selected layer. To make the tool-tip question marks on the front page more noticeable, we would change their colour to a bright blue. This would make them contrast with the grey elements of the hidden layer graph. Similarly to highlighting the clickability of the different layers, we would also add an arrow pointing to one of the question marks with a text element saying 'hover me'. At the painting interface, we would add a question mark like those on the front page explaining the colour coding of the buttons - in other words - their connection to the front page. We would also remove the '-iou'-naming from each of the buttons in the painting interface as these confused users, and were mainly there for technical reasons. Beyond just visual changes, other changes would relate more to the actual functionality of certain parts of the interface on the front page. Firstly, this would consist of being able to use

the slider with more features than just the 'tree' feature. Specifically, the slider should be able to use each of the different features represented in the bar graph of each layer. This would also entail making every 'dot' representing a feature in the hidden layer graph clickable, which would be how the user could pick the feature they intend to insert or ablate in the slider. These changes would cover the most frequent comments that were given by participants of our user test. After implementing them, it might then have been relevant to conduct another user test to see the effects of these changes.

These limitations, together with the results of the user test, had a direct impact on the research outcomes. For instance, the predominance of technically proficient users might have resulted in overly positive feedback regarding the interface's usability, which may not hold true for a broader audience. Similarly, addressing these limitations in future research cycles will be crucial for developing a more robust and universally accessible educational tool.

However, the user test findings support the claim that the XGANArt implementation overall improved both the explainability and the transparency of GANs for the seven participants of the user study. By allowing these users to observe and manipulate layer-specific effects, the interface facilitated a deeper engagement with and understanding of the model's internal mechanics.

The enhancements of XGANArt not only deepened the understanding of the internal mechanisms of GANs for the majority of the seven test participants but also showed how methods from XAI can be applied in practical scenarios.

Our contributions propose to make a step forward in making GAN technology more understandable, thereby expanding the potential uses and users of these powerful machine learning models. Our work aims to highlight the importance of transparency and control in artificial intelligence, paving the way for further innovations in the field.

The project met and, in some aspects, exceeded our initial expectations by successfully demonstrating that interface design can impact users' understanding of complex AI models. However, there remains a gap between user engagement and a complete understanding of GAN processes, indicating areas for further research.

8 Conclusion

With the growing interest in AI, and the continued research into explainable AI (XAI), there have been more and more proposed ideas of how to implement explainability. This project has provided a comprehensive exploration of the explainability and transparency of Generative Adversarial Networks (GANs). We have utilized one of the most influential works from the body of literature concerning XAI, namely the XAI question bank, as inspiration for the design and implementation of the XGANArt interface. Likewise, the decision to incorporate interactive features was highly inspired by the 'Explore'-pattern, a pattern describing the interactions between humans and AIs in GANs. By facilitating direct interaction with the GAN's hidden layers and enabling users to modify neural network activations, it was indicated that XGANArt enhanced the understandability of GANs based on the self-assessed understanding of the participants in the user testing performed in the project. Our research potentially suggests that interfaces such as XGANArt can help increase the transparency of complex models such as the GANDissect model.

Through user testing, we demonstrated that the XGANArt interface not only indicated to have improved the test participants' understanding of how GANs operate but also suggested that the participants gained a more intuitive grasp of the machine learning processes involved. From the user test, we also identified areas of improvement, such as visual design choices to make the site more intuitive and using more interactive features to explain the inner workings of the model. These outcomes align closely with the definition of explainability established for this project as giving an explanation which is a presentation of the reason of a machine learning model, with the goal of making AI understandable by people. Even so, the outcomes also encourage future work to explore the effects of XGANArt.

8.1 Future Research

The development and initial testing of the XGANArt interface have laid a strong foundation for further exploration and improvement in the field of explainable AI and human-AI interaction. Here are several avenues for future research that can expand

upon the current work:

In future research endeavors, a priority should be implementing the feedback from the XGANArt user test. Integrating suggestions from the users is essential as it helps refine the tool's functionality, enhance its usability, and ensure it meets the real-world needs of its audience. By actively responding to user input, we can create a more intuitive and effective tool that addresses specific challenges users face and encourages continuous engagement and learning.

Our interface attempted to tackle the issue of lack of explainability by exposing the hidden layers of the neural network to the user. This allowed the user to take a peek into the output of some hidden layers of the model and tweak the values. However, our interface only allows the user access to three manually picked layers. Furthermore, any changes made on a specific layer do not carry over to the output for a later layer accessible to the user. Creating an interface that makes all hidden layers accessible to the user and allows for modifications across successive layers can provide a more comprehensive tool and richer insights into the model's operation, enhancing explainability, as well as the opportunity for co-creation between the AI and the human user.

Moreover, further investigation should be yielded into the adaptability of XGANArt to various AI models and neural network architectures. Such flexibility in the application would not only broaden the scope of the tool but also enhance its practicality across diverse scientific and industrial domains. Additionally, the potential to customize these interfaces to align with specific user requirements represents a significant advancement. Tailoring tools like XGANArt could enable users—from students learning foundational concepts to professionals tackling complex datasets—to modify the interface's complexity and functionality to suit their unique contexts and goals. This adaptability would make the tool more accessible and effective, encouraging deeper engagement and understanding of AI systems.

Future efforts should aim to expand the accessibility of such tools to a wider audience and further refine the interaction mechanisms to cover more aspects of GAN functionality. By continuing to bridge the gap between AI operation and user comprehension, we can enhance the co-creative potential of AI. This project has laid a solid foundation for further exploration into user-friendly AI systems, setting the stage for

more innovative and inclusive advancements in the field of artificial intelligence.

References

[Bau et al., 2018] Bau, D., Zhu, J.-Y., Strobelt, H., Zhou, B., Tenenbaum, J. B., Freeman, W. T., and Torralba, A. (2018). Gan dissection: Visualizing and understanding generative adversarial networks.

[Brownlee, 2020] Brownlee, J. (2020). Arithmetic, geometric, and harmonic means for machine learning.

[Chavda and Goyani, 2023] Chavda, S. and Goyani, M. (2023). Scene level image classification: a literature review. *Neural Processing Letters*, 55(3):2471–2520.

[Deviation, 2022] Deviation, M.-B. S. (2022). Medium.

[Fawcett, 2021] Fawcett, A. (2021). Introduction to human-computer interaction design principles. *educative*.

[Fowley, 2011] Fowley, M. (2011). *Patterns of Enterprise Application Architecture*. Addison-Wesley.

[Géron, 2019] Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Incorporated.

[Grabe et al., 2022] Grabe, I., González-Duque, M., Risi, S., and Zhu, J. (2022). Towards a framework for human-ai interaction patterns in co-creative gan applications.

[Karras et al., 2018] Karras, T., Aila, T., Laine, S., and Lehtinen, J. (2018). Progressive growing of gans for improved quality, stability, and variation.

[Kumar, 2021] Kumar, N. (2021). How the model view controller architecture works – mvc explained. *freeCodeCamp*.

[Liao et al., 2020] Liao, Q. V., Gruen, D., and Miller, S. (2020). Questioning the ai: Informing design practices for explainable ai user experiences. *CHI*.

[Liao and Varshney, 2022] Liao, Q. V. and Varshney, K. R. (2022). Human-centered explainable ai (xai): From algorithms to user experiences. *arXiv*.

[Lubart, 2005] Lubart, T. (2005). How can computers be partners in the creative process: Classification and commentary on the special issue. *International Journal of Human-Computer Studies*, 63(4):365–369. Computer support for creativity.

[Maas, 2013] Maas, A. L. (2013). Rectifier nonlinearities improve neural network acoustic models.

[Martin Fowler, 2018] Martin Fowler, Kent Beck, J. B. W. O. D. R. E. G. (2018). *Refactoring: Improving The Design Of Existing Code*. Addison-Wesley.

[Moran, 2019] Moran, K. (2019). Usability testing 101.

[Nauta et al., 2023] Nauta, M., Trienes, J., Pathak, S., Nguyen, E., Peters, M., Schmitt, Y., Schlötterer, J., Keulen, M. V., and Seifert, C. (2023). From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable ai. *ACM Computing Surveys*.

[Olah, 2022] Olah, C. (2022). Mechanistic interpretability, variables, and the importance of interpretable bases.

[Saxena and Cao, 2021] Saxena, D. and Cao, J. (2021). Generative adversarial networks (gans): Challenges, solutions, and future directions. *ACM Comput. Surv.*, 54(3).

[Sommerville, 2016] Sommerville, I. (2016). *Software Engineering*. Pearson.

[Wei Xu and Gao, 2023] Wei Xu, Marvin J. Dainoff, L. G. and Gao, Z. (2023). Transitioning to human interaction with ai systems: New challenges and opportunities for hci professionals to enable human-centered ai. *International Journal of Human–Computer Interaction*, 39(3):494–518.

[Wijaya et al., 2017] Wijaya, D. R., Sarno, R., and Zulaika, E. (2017). Information quality ratio as a novel metric for mother wavelet selection. *Chemometrics and Intelligent Laboratory Systems*, 160:59–71.

[Wu et al., 2021] Wu, Z., Ji, D., Yu, K., Zeng, X., Wu, D., and Shidujaman, M. (2021). Ai creativity and the human-ai co-creation model. *Human-Computer Interaction*.

[Yannakakis et al., 2014] Yannakakis, G. N., Liapis, A., and Alexopoulos, C. (2014). Mixed-initiative co-creativity. In *International Conference on Foundations of Digital Games*.

[Yosinski et al., 2015] Yosinski, J., Clune, J., Nguyen, A., Fuchs, T., and Lipson, H. (2015). Understanding neural networks through deep visualization.

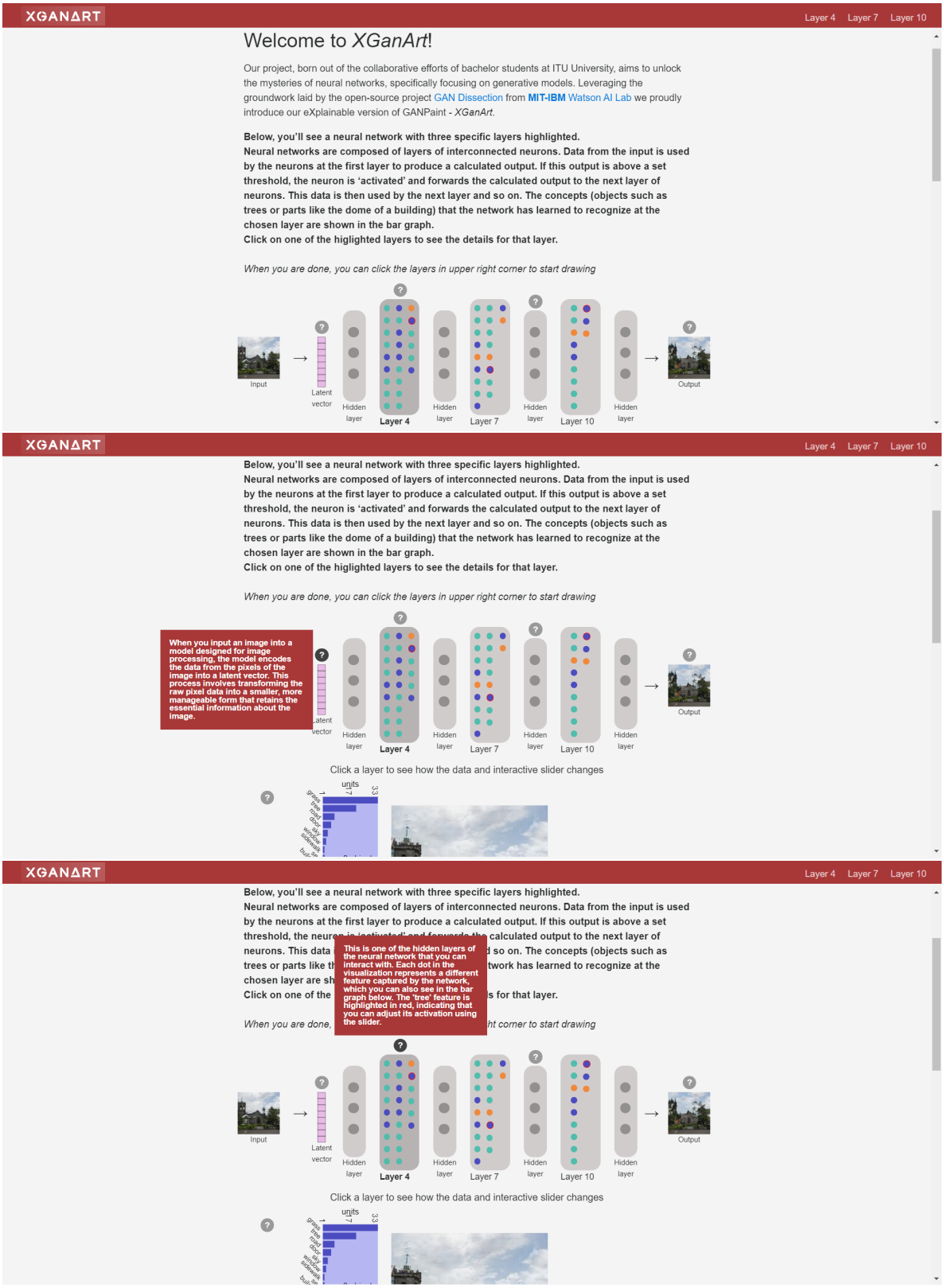
[Yu et al., 2016] Yu, F., Seff, A., Zhang, Y., Song, S., Funkhouser, T., and Xiao, J. (2016). Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop.

[Zaltron et al., 2020] Zaltron, N., Zurlo, L., and Risi, S. (2020). Cg-gan: An interactive evolutionary gan-based approach for facial composite generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 2544–2551.

[Zhu et al., 2018] Zhu, J., Liapis, A., Risi, S., Bidarra, R., and Youngblood, G. M. (2018). Explainable ai for designers: A human-centered perspective on mixed-initiative co-creation. In *Proceedings of the IEEE Conference on Computational Intelligence and Games*.

A Appendix A

A.1 Interface Screenshots



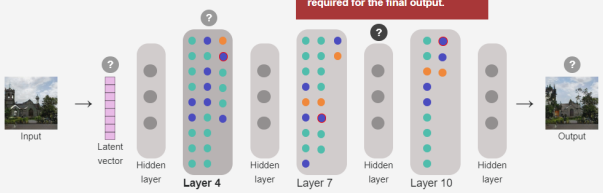
XGANART

Layer 4 Layer 7 Layer 10

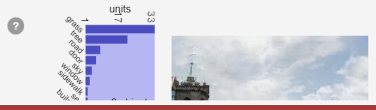
Below, you'll see a neural network with three specific layers highlighted.
Neural networks are composed of layers of interconnected neurons. Data from the input is used by the neurons at the first layer to produce a calculated output. If this output is above a set threshold, the neuron is 'activated' and forwards the calculated output to the next layer of neurons. This data is then used by the next layer and so on. The concepts (objects such as trees or parts like the dome of a building) that the network has learned to recognize at the chosen layer are shown in the bar graph.
Click on one of the higlighted layers to see the details for that layer.

When you are done, you can click the layers in upper right corner to start drawing

All layers inbetween the input- and output-layer are 'hidden layers'. As data flows through the hidden layers, the network learns to recognize features. Each hidden layer can extract different types of features from the input, gradually building a more detailed understanding of the data, required for the final output.



Click a layer to see how the data and interactive slider changes



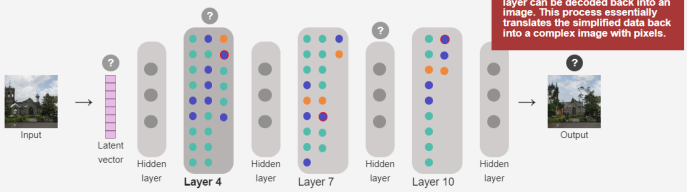
XGANART

Layer 4 Layer 7 Layer 10

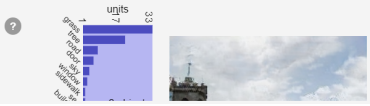
Below, you'll see a neural network with three specific layers highlighted.
Neural networks are composed of layers of interconnected neurons. Data from the input is used by the neurons at the first layer to produce a calculated output. If this output is above a set threshold, the neuron is 'activated' and forwards the calculated output to the next layer of neurons. This data is then used by the next layer and so on. The concepts (objects such as trees or parts like the dome of a building) that the network has learned to recognize at the chosen layer are shown in the bar graph.
Click on one of the higlighted layers to see the details for that layer.

When you are done, you can click the layers in upper right corner to start drawing

The data from the last hidden layer can be decoded back into an image. This process essentially translates the simplified data back into a complex image with pixels.



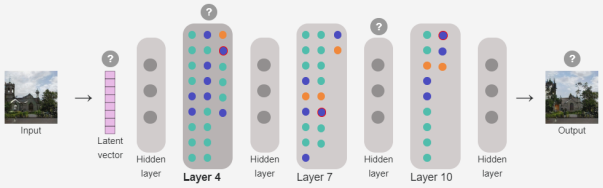
Click a layer to see how the data and interactive slider changes



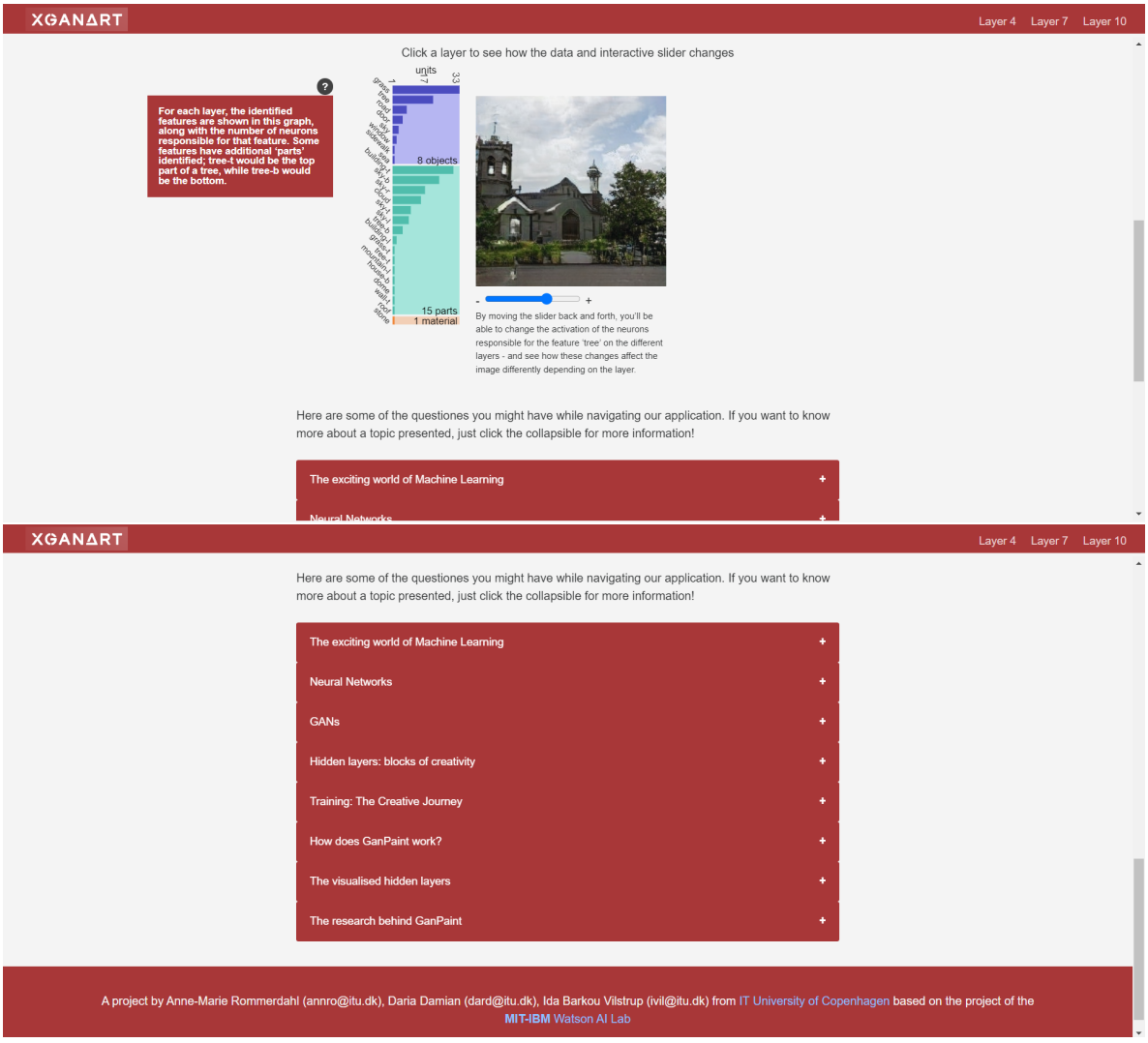
XGANART

Layer 4 Layer 7 Layer 10

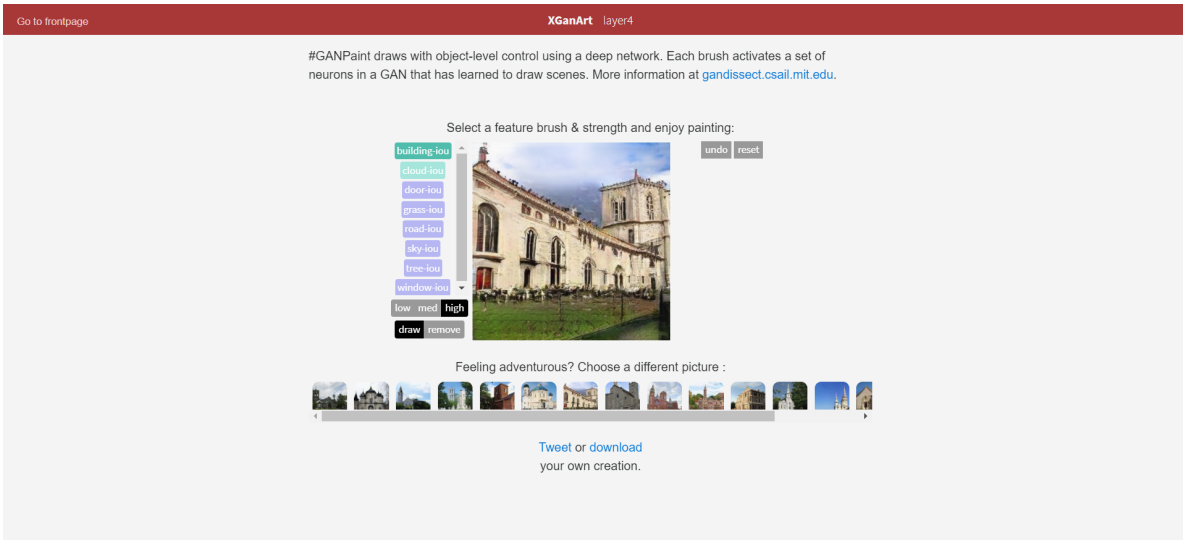
When you are done, you can click the layers in upper right corner to start drawing

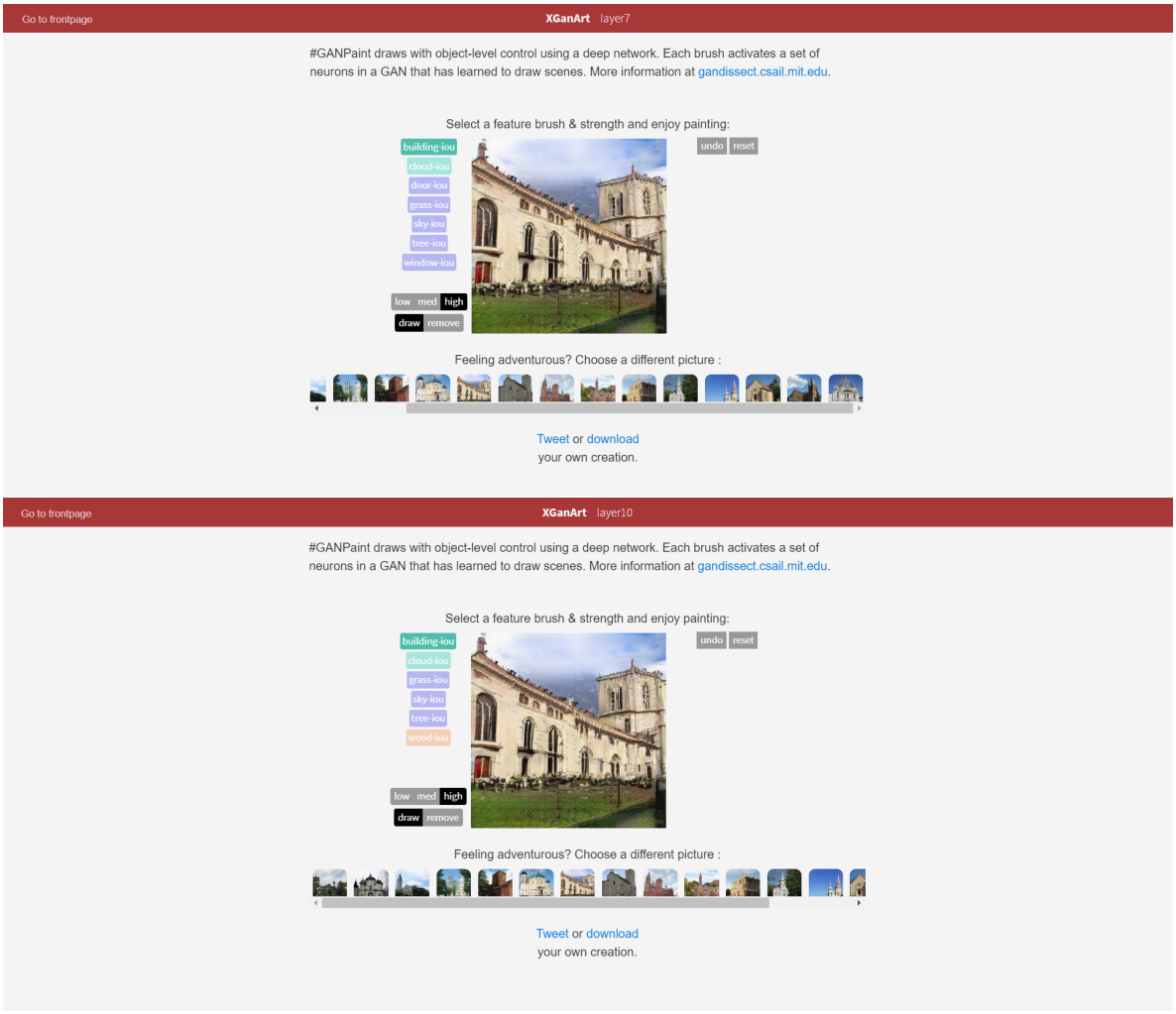


Click a layer to see how the data and interactive slider changes



A.2 GANPaint Interface Reworked





B Appendix B

B.1 User Test

Welcome and thank you for joining today!

We have been working on an application. We're currently asking people to try it out and share their opinions. We want to learn about your impressions and how it works for you. This will help us to improve the application going forward. The session should take about 15-20 minutes.

As you use the site, try to think out loud: To say what you're looking at, what you're trying to do, and what you're thinking. I will be helping you by asking some additional questions.

We're testing the application, not you. Please don't worry that you're going to hurt

our feelings. We're doing this to understand what improvements could be made to the application for future work, so we need to hear your honest reactions.

If you have any questions as we go along, just ask them. I may not be able to answer them right away, but I will make sure to answer them at the end of our session.

I would also like to ask you if it's okay that we voice-record this whole session and screen record the interactive parts of it. The recordings are only for internal documentation purposes.

Do you have any questions before we start?

B.1.1 Intro Questions

Before we start, I have a few introduction questions:

- How old are you?
- What are you currently studying? (study program - bachelor/master's - semester)
- Have you used generative AI before e.g. ChatGPT, DALL-E, Midjourney or similar?
- If yes, how many times do you use it during a week (for instance last week)?
- What do you use it for?

Next we will ask a few questions where your answers should be on a scale from 1-5.

- What is your knowledge of machine learning/AI? Where 1 is "I don't know what it is", and 5 is "I have taken a course/researched the technical aspects of machine learning myself"
- Are you interested in understanding how AI/machine learning works? Where 1 is "No, not at all" and 5 is "I am very interested in this topic"

B.1.2 Original GANPaint interface

Now, we are going to show you the original GanPaint interface.

- Describe in your own words what you see on this page.

- What was your first impression when looking at it? Anything that caught your attention?

— Let the user look over the page for a few seconds —

- What do you think the buttons on the left side of the image mean? What might their purpose be?
- Now, try drawing something - anything you'd like. (Give the user a little time - around 30 seconds - to do this however they want)
- Now, what do you think the buttons mean?
- Try to explain to me what you think is happening behind the scenes as you're drawing.

B.1.3 Front page of XGANArt

Now we are going to show you our version - XGANArt. We will first show you the frontpage we developed before you go to the painting application.

- What was your first impression when looking at it? Anything that caught your attention?

— Let the user look through the page for a minute —

- What would you think the below interface does? (The graph etc.)

Ok, I will now ask you to do a few simple tasks. Please remember to think aloud as you go along.

- Have a look at the interface - graph etc. (If they do not think aloud ask: What do you think it represents?)
- Try out the slider (Give the user about 15 sec)
- Play with the graph - what happens on the site and what do you think it means?
- Choose layer 10 and play with the slider again

B.1.4 XGANArt painting page

Let's now look at the application's XGANArt part. Please click on layer 4 in the top right corner.

- What was your first impression when looking at it?
- Anything that caught your attention?
- How do you think this page relates to the front page?

I will now ask you to do a few simple tasks. Please remember to think aloud as you go along.

- Draw a tree
- Go back to the frontpage, then go to layer 10
- Draw a tree (if the user doesnt say anything - ask them about their thoughts. Did they notice a difference when drawing on this layer compared to layer 4?)
- FREE PLAY - You can now take a few minutes to go to any layer/page you'd like and draw whatever you'd like.

Now, try to explain to me what you think is happening behind the scenes as you're drawing.

B.1.5 Wrap-up questions

- What was your overall impression of XGANArt?
- Was there a clear 'progression' through the site? Were you confused about anything at some point?
- Do you have any comments for us? Something that could be improved, or other things?

Thank you very much for your time. It's been helpful and we are sure that we can use your feedback.

B.2 User Test Interviews

B.3 Interview 1

Date: 05/05/2024

Interviewer: Okay, perfect. Before we start, I have a few introduction questions. How old are you?

Interviewee: I am 23 years old.

Interviewer: And what are you currently studying?

Interviewee: I'm studying Software Development at the IT University of Copenhagen.

Interviewer: And what semester are you?

Interviewee: Sixth semester.

Interviewer: And have you used generative AI before?

Interviewee: Many times.

Interviewer: Yeah, and how many times a week would you say you use it?

Interviewee: Oh, that's a tough one because I'm doing a project for my bachelor's, which is straight up about generative AI and large language models. So I'm probably doing at least 50 generations a day.

Interviewer: Or a week. So I guess that is what you use it for, right?

Interviewee: Yes.

Interviewer: Okay. And now I will ask you a few questions, but your scale should be between 1 to 5. So, what would you say your knowledge of machine learning/AI is? One is, I don't know what that is. And five is, I have taken a course in it/researched it myself.

Interviewee: Currently, that's a four.

Interviewer: Yeah. Okay, that's a four. Are you interested in understanding how AI machine learning works?

Interviewee: Very much.

Interviewer: Okay, so now I will show you just the original GANPaint interface of the open-source project we started with. So, just from this, what you see, can you describe what you see?

Interviewee: I see a grey website with an image of a castle and a bunch of buttons I can use. There are also some descriptions.

Interviewer: You can read them if you like.

Interviewee: It was released in 2019 for image editing and generative networks. To draw object-level control using a deep network. Each brush activates a set of neurons with a GAN. I'm not familiar with that and learned to draw scenes. More information over at some cool website. So I can imagine I can just swap between images here. All of them just happen to be castles for whatever reason. Yeah, I mean, from what I can gather, we seem to have some filters about drawing, removing, undoing, and resetting. I'm a bit familiar with at least Photoshop so I would assume that those are the tools that we can use to manipulate the image in some way. And we also seem to have some buttons along the side with labels such as sky, tree, and door. I would sort of assume those would be akin to a paintbrush or something where you can use, say, if you select tree and draw then you can draw trees or you can do tree and remove and remove trees from the image. I would assume that's something like that.

Interviewer: So now, you can just use the site for a bit, as much time as you want to. Interviewee: I... I need to see if they have a selection for the door. But as far as I can tell, it's very difficult for me to tell when something is a door. There we go, that's something. Make a bigger door, even bigger, there we go, beautiful. Honestly, that's probably my favorite feature right now.

Interviewer: Perfect. Okay, so did that change your understanding of what the book is?

Interviewee: No, it did absolutely what I expected it to do, although the effect was very underwhelming.

Interviewer: Can you try to explain what you think is happening behind the scenes?

Interviewee: Oh, that's a tough one. I'm not too familiar with image-generated AI, but I would assume that the way that a general generative AI works is you have a bunch of neurons that all have certain values, and you give them a value input, and then they're going to produce a certain output based off of the data that it's getting. And so here the input is an instruction to draw or remove something combined with a keyword along with sky, door, or grass and then it's going to pass this information through

combined with the area that I have selected, and it's using either image recognition to analyze the image, to have a preconceived notion as to what the individual pixels are categorized as, and so when I select a certain amount of pixels, then it's going to try and move them from one category to another and try and manipulate the image based off of previously trained images that it has used to categorize images beforehand. I have no question about image generation, but yeah. It is an assumption, I have no idea.

Interviewer: Okay, perfect. Okay, now we'll move on, and I will show you our version, which is called XGANArt, and we'll first show you the front page, which we developed before you then go to the actual appliaction.

Interviewee: Okay.

Interviewer: Alright, so that is here.

Interviewee: So, you can scroll down and tell me what your first impression is.

Interviewee: I'm very overwhelmed by a bunch of text just filling up out front, I guess. I don't have to read that right now, but I'll introduce it. This is an illustration of how it works, or how layers work. That's pretty neat. A bit of a description as to how the model works. The three days that you're talking about it's like the first day is to obtain the case, the second day is to read the case, and the third day is to write the thing. Yeah, so this is sort of the layered illustration of how a machine-learning model works. And yeah, that's something I am sort of familiar with, the layered structure. Moving the slider back and forth, you'll be able to change the activation of neurons responsible for the feature tree on different layers. You can see these changes affect the image differently depending on the layer.

Interviewer: That's neat.

Interviewee: You can completely remove any sort of notions of tree. Each of the identified features is shown in the graph, along with the number of neurons responsible for the feature. So, what does that mean? I'm a bit confused as to what this graph is supposed to tell me, although I will assume it is the values of the base image, I think it would be pretty cool to see the values change as you are moving the slider back and forth because I think that would make a lot more sense and sort of thing.

Interviewer: Here are some questions you might have on navigating our application

if you want to know more about the topic presented. So if you want to go to Layer 7.

Interviewee: These individual buttons. So if I'm in layer 7 and I press door now, that does nothing. Unless... wait... does that mean I can... no. I am very confused as to, like, since... I mean, I guess the individual nodes within the layer are supposed to just illustrate that at this point in the model we're considering values of stone, grass... Then you have... Ah, okay, I guess it doesn't make sense to illustrate it that way, but it is a bit confusing to have them labeled as buttons, because that would infer that there's a difference between me pressing a button A in layer 7 versus pressing a button B in layer 7. But it doesn't seem to have any difference. But for the sake of showing the individual values or nodes within the layer, that makes sense.

Interviewer: Can I then ask you to use the slider again on layer 4 first? And then maybe turn it all the way up?

Interviewee: Okay.

Interviewer: And then can I get you to layer 10? And then use the slider again and do the same thing?

Interviewee: So this value changes the activation of the neurons, tree on different nodes, so this one is, this one should have a tree, this one should have a tree node, it probably does not, no it does not, oh wait it does, it has this tree node, what is the other tree node? The other ones have. They also just have tree associations. I would probably assume that if I go back to layer 4, do this, go to layer 7, do that, and layer 10, do that, hmm, that's very puzzling. It obviously has something to do with how the model has been trained on the data, but this is absolutely where my knowledge of the field comes short, because I am absolutely not aware as to what the differences are. I would assume it is either one of the layers is more associated with certain characteristics of a tree and the other layers are more associated with other characteristics, or it's the aggregate characteristics of a tree that are then split up across the different layers, but it's a bit tough to tell from the image.

Interviewer: Okay, now you can go to layer 4 in the top corner so this is the part of the application we actually call XGANArt and the other part is the front page. So what is your first impression?

Interviewee: I mean, it's very similar to the first thing. Although I do like the color-

coding of the tools on the side, I am still not entirely sure why they're color-coding or what the categories of color-coding essentially mean.

Interviewer: Alright, perfect. How do you think this page relates to the front page?

Interviewee: I would assume that since we're clicking on layer 4, then we can manipulate it using a very small bit of the model instead of the entirety of it.

Interviewer: Now, I will ask you to do a few tasks and think it out. Can you draw a tree?

Interviewee: I don't know, maybe. I don't think I succeeded on that, to be fair.

Interviewer: You can draw some more if you want to.

Interviewee: I can try and do this. We do have resets so what if I try and do this? Oh yeah, so this is removed, it's not that's not it.

Interviewer: But then now can I ask you to go back to the front page and then go to layer 10? And now try to draw a tree. You can find the same picture if you want to.

Interviewee: No, I think it's not going to make too much of a difference, but we can try this one. Yeah! Oh, that's just going to be the green fielder, isn't it?

Interviewer: So, are you noticing a difference between the layer you had before and this one?

Interviewee: So this layer seems to be taking existing pixels in the photo and just applying a green layer of mush over it and the previous layer, layer 4, which had a really tough time actually producing anything it seemed to exaggerate existing pieces of what the model might consider to be a tree.

Interviewer: Now I'll just give you some time where you can play around with everything you want to. Like also front page, these pages, whatever you want really.

Interviewee: That is very cool. I do want to check out layer 7 real quick. So if I do the low tool, I'm a bit, I'm actually a bit confused as to what either of these is gonna do, but I'm gonna assume that this is gonna be like a four pixelated cloud if I select sky or cloud and blue is that gonna be a... Yeah, it's a bit, what do you say, rugged, but if you do it with high, it does seem to be, I don't know, it's a bit tough to tell the exact difference between high, medium, and low. If you do a tree on this layer with medium, this seems to sort of, it does seem to produce something, but I wouldn't exactly call it a tree, but it does seem to be some shrubbery, at least. Does that tweet link actually

work?

Interviewer: I think it does, I don't have a Twitter though.

Interviewee: Dana does. But you don't actually have a Twitter account for it? There are also people I was called X, but you know.

Interviewer: But okay, yeah, so now after you've been at the original GANPaint, you've been at the front page, you've been here, do you, can you now explain to me what you think is happening behind the scenes? Like, have you, or is there any difference in what you, what you would think is happening?

Interviewee: I don't think I gained too much more insight than I already knew before, if I'm being frank, because my knowledge is just very general and it's not very specific as to exactly how the values are being used in the model, so it's a bit rough, but I think that I got a as to how one keyword can be split across multiple different layers and that each layer is going to be responsible in some way of interpreting like a certain aspect of a keyword for the images.

Interviewer: And then I just have some wrap-ups, so what would be your overall impression?

Interviewee: Overall impression of the entire thing?

Interviewer: Yeah, well, actually not the very first image.

Interviewee: I mean, the website is alright. It definitely doesn't, it's more of a 2005 website than a 1998 website, so that's something. The images are very, very low quality. I would like to see either a bigger sample size or a higher-quality image. I think the ability to choose between different layers to operate on the drawing tools is very cool. But I do think I would have liked to see the base version from the first version again, where you can just do whatever you want with whatever tool you want, because on this front page, I can only exaggerate one element instead of getting to draw on it again. I feel like that's what's missing. But other than that, it seems to be alright. And I do like this layered illustration, it is pretty cool. The only thing that could make it better is if you would have, say, an illustration of exactly what a, if you have the input and the keyword alongside it, and you could have, say, an illustration of what it would look like after layer 1, after layer 7, and after layer 10. That would probably make this layered illustration here a lot more effective in my mind. For example, if you have the

keyword tree, then you would have like, oh, this is... So basically, it would be like the end results of the slider. Like you would have an illustration of, okay, you would have this, and then you would have this, and then you would have this.

Interviewer: Yeah, exactly.

Interviewee: That's one thing. I think that would make it a lot more clear as to what would happen and you have it all congregated into one singular page and one image. That would be very effective.

Interviewer: Nice, okay. Was there a clear progression through the site or were you confused about anything?

Interviewee: I was absolutely confused by having the layered buttons like this, and then you have the exaggerated layer up in the top right because since they're named the same, I would assume they go to the same place, but apparently, they don't. So, I don't know how I would actually go about solving that. Maybe I would name the top right button something else, but I am a bit clueless as to how I would do it. Except maybe you could take this element from this page on the slider page and move it to the same page.

Interviewer: Into the same view?

Interviewee: Yes, exactly. That would make a bit more sense in my mind.

Interviewer: So, you've already added some things that could be improved. Do you have anything else?

Interviewee: Nah, I think that was everything I could think of at least.

Interviewer: Perfect, so thank you for your time, it's very helpful and we can use your feedback.

B.4 Interview 2

Date: 05/05/2024

Interviewer: So welcome. Thank you for joining us today.

Interviewee: Anything I should know or am I going in completely blind?

Interviewer: You're going in completely blind.

Interviewee: Okay.

Interviewer: Okay, so before we start, just a few introduction questions. How old are you?

Interviewee: I'm 20.

Interviewer: All right, and what are you currently studying? What's your study program, and what semester are you in?

Interviewee: So I'm in the fourth semester, and I'm studying for my bachelor's in data science here at ITU.

Interviewer: Yes. So things like general generative AI like ChatGPT, Dally, Mid Journey. Have you used anything like that?

Interviewee: Yeah, yeah.

Interviewer: How many Would you say you use it regularly? How many times a week?

Interviewee: Maybe not that much, but what I say regularly, maybe from time to time. So I don't use it like every single day or every few days. But when I have a question, I really use it and I like debate with it and try to get a lot of information out of it.

Interviewer: Alright. And then we're going to ask some questions where your answers should be a scale of 1 to 5. We'll answer what 1 to 5 means. So what would you say, like, how is your general knowledge of machine learning slash AI? Where a 1 would be "I don't know what that is". And a 5 is "I have taken a course or I have researched it myself".

Interviewee: Okay, well, I have literally taken a course here and I've also taken a course which was not here So I'd say five on that scale But I wouldn't say like I'm an expert or anything But if your scale is five is I have taken a course then five.

Interviewer: And are you, in general, interested in understanding how AI and machine learning work? Where one is a no, not at all, and five is I'm very interested. Um, I mean, by your five being I'm very interested I would say five, but it's not like my number one interest in life.

Interviewee: So by a normal scale, maybe I would say four But by your scale, it sounds like a five. Yeah, so I'm not sure.

Interviewer: So now we're going to show you an original demo from another project,

and it is this one. In your own words, what do you see on the beach?

Interviewee: So it says painting with generative models, and I see a picture of I don't know what it is maybe a cathedral I have no clue what this is some sort of old building maybe religious I really don't know I don't recognize the building maybe a church maybe a cathedral maybe I have no clue what it is but it's some building yeah okay so it says like the feature brush is straight that I enjoy painting so I'm just gonna say I want I want some grass and I want to draw some grass. Oh, my grass is... Oh, yeah, that was pretty cool. That was pretty cool. Okay, so I just selected the grass brush and originally it looked kind of funky because it looked like turquoise, like cyan color. And I just put it there and then it magically put in some grass where there was like some road, not road, but like some like walking space before which wasn't grass. So that's pretty cool. Okay, I'm gonna put a door where there was a window and see what happens. Oh, yeah, that kind of looks like a door Yeah, although it's kind of big. Yes. Okay. Now I drew again and gave you another door Just what actually looks like a proper door So and now I'm gonna put some sky where there was some grass see how funky that ends up looking

Interviewee: That's that looks very weird I'm not gonna lie the sky on grass looks pretty funky Oh, no, you know, it's actually properly sky. It looks like we're falling here Oh, oh, no, everything is getting weird. I totally didn't do this by making horrible life choices Yeah, now I just have a building in the middle of the screen surrounded by sky everywhere Now I'm gonna bring in a brick randomly floating in the sky And now it looks very very weird. Some things are going on So can I ask you, like based on the first couple of seconds when you looked at this website like what was your impression of it anything goes um i think the background color is very weird but this is just me picky like it's like it's like a light gray it's a very weird color anyways i'm just gonna look at another picture

Interviewer: Now, just to reiterate, those buttons on the side of the picture, what do you think they mean?

Interviewee: Um, what sort of things do you want to insert in the picture.

Interviewer: Now that you've played around with it a bit, just anything goes again, like can you try to explain to me what you think is happening behind the scenes? But

if you don't know, you can say, I don't know, right? Like just try. Yeah, anything. Whatever you're thinking.

Interviewee: I mean, I'm honestly not sure. Yeah, I don't think anything from what I've learned is necessarily applicable here, honestly.

Interviewer: That's fine. Yes, we're going to move on to the next one. We're going to show you our project, and we have a front page to start with. And before you do anything too much, just again, what's your first impression? anything that caught your attention.

Interviewee: This page looks more polished than the other one. It looks like some sort of explanation for the thing I tried before. Yeah, and we're giving you some time here so you can go through it. Okay.

Interviewee: Right, I like that there's a frequently asked questions sort of part. that's nice.

Interviewer: Alright, so now I would like to ask you, if you could scroll up a little bit, this part of it (showing the neural network graph), what do you think it does, like what's the intention behind it? graph You mean like... What do you think about it?

Interviewee: I mean, it's showing like the input, which I suppose is like the image before you add things to it, and then the hidden layers, which here you are showing some of them.

Interviewee: Yeah, it's like it had something to do with the slider, but I don't see anything change when I move it, so I'm not sure what it was supposed to change. I mean, it's changing the pictures, but I don't see anything else on this side being changed.

Interviwer: Okay then I will ask you to do something, scroll up a little bit. Okay. And then you have these three layers, you're currently sort of like on layer four, try to click another layer.

Interviewee: Oh, you could select them. Okay, I didn't think of doing that. To me, it kind of looked like just a static image, and then I tried like moving the slider to see if it changed something because it was saying something in the text, but like I didn't think of clicking another layer. Yeah, I thought the slider would affect this graph, like the bars, because I also looked at the layers and I didn't see any change, but I didn't

think of clicking it because I thought it was just an image after I didn't see any change after moving the slider.

Interviewer: Then as the last part, can you choose layer 10? Okay. Then play with the slider again a little bit.

Interviewee: I remember something out loud. I honestly still don't see anything change other than the image, I'm not sure what I'm supposed to... Oh, wait a minute. Wait, what?

Interviewee: Wait, it had help buttons. I'm stupid. Anyways. Okay.

Interviewee: Yeah I did notice the picture I just thought that the point was some of the other things changing.

Interviewer: Okay now we'd like you to please go here, click on layer four, and again, the first impression of this,

Interviewee: Yeah it looks nice, like it looks the same style as the pages from before

Interviewer: How do you think this page relates to the front page?

Interviewee: Well, I clicked layer four to get here, so I suppose it somehow relates to that. All right now, yeah, you can maybe pick the tree and try to draw something. Okay

Interviewee: Yeah, it works, it inserted some trees in a pretty weird way, but it works. Now, please go back to the front page. Oh, you can actually see a history of everything, that's cool. Yeah, okay, so what do you say? Go back to the front page, please, and then navigate to layer 10. Okay. And then pick the tree again and draw something. Okay. and remember to think about okay so oh this looks very different like this time it doesn't really look like trees it looks like just some sort of tree texture.

Interviewer: All right, and then if you'd like we you can do some free play if you'd like where can do whatever you want on the side go to whatever page and just fiddle around with it a bit for like one or two minutes all right and remember to speak out loud whatever you're noticing.

Interviewee: okay, I guess I'll go to the page I haven't been to yet, which is layer seven, and get the tree again. Okay, the tree is waiting, But I'm not having the tree selected, okay, that's okay.

Interviewee: Yeah, it's not really adding a tree on this one either. I guess, kind of.

Interviewee: I suppose it's because we're only using certain parts of the model, maybe, since the pages are called layer 4, layer 7, and layer 10, but the one which is the actual page with the fig, it's very polished, but on the ones with the layers, it looks like it's adding like different components of a tree but not like a full tree okay and then there's a bunch of different pictures I don't think that's gonna be very interesting it's gonna be the same thing but with a different picture back to the front page.

Interviewee: Yeah, as I said they like the like frequently asked questions sort of section the thing I would like make more clear is what sort of things you are supposed to see changing and what you're supposed to focus your attention on this like interactable part because I didn't find this intuitive at all yeah so yeah that would be my overall feedback but it looks nice like I think it's pretty fun and the front page is also really nice maybe as well I would change this page around a little bit like the main page just to look a little bit more polished and maybe be on the same style as the main page and stuff like that. But that's just design, it doesn't really matter.

Interviewer: One question for this is then, when you've looked through our site now and when you've made a change to the picture, like, the same question as when you looked at the very first demo, like, can you try to explain to me what is happening behind the scenes?

Interviewee: Yeah, I mean it looks like with some sort of neural network from this page.

Interviewer: Okay, cool. So, in general, what was your overall impression of XGA-NArt? So that's this page and these layers, not the grey page. Okay. What was your overall impression?

Interviewee: It's pretty nice; it's just that this demo is a bit counterintuitive, but I think the design is nice, and the idea is interesting. Maybe a bit more writing explaining how it works on the page and stuff like that.

Interviewer: Where would you place this explanation, so it makes sense?

Interviewee: Maybe like a separate like page, like where you have layer 4, layer 7, layer 10 have like a different page, which is like how it works and then have like a big explanation, maybe.

Interviewer: Okay, then next question. Do you feel there was like a clear progression

through the site? Like were you confused about anything at some point?

Interviewee: I guess it's a bit confusing, like those layers 4, 7, and 10, and what specifically you want to show with them.

Interviewer: Do you have any comments on this, something that could be improved or anything else?

Interviewee: I guess just the things I've already said.

Interviewer: Well, then, we got nothing more for you. Thank you very much for your time. Yeah very helpful.

B.5 Interview 3

Interview Date: 05/14/2024

Interviewer: Welcome and thank you for joining today. So, do you have any questions before we start?

Interviewee: I don't think so. It all seems pretty reasonable, so yeah, I'm ready.

Interviewer: Okay, perfect. So, before we start, I have some introductory questions. How old are you?

Interviewee: I am currently 24.

Interviewer: What are you currently studying?

Interviewee: I am studying Software Development and I'm currently writing my Bachelor's.

Interviewer: So you're in your last semester?

Interviewee: Yeah, last semester, sixth semester.

Interviewer: Have you used generative AI before, such as ChatGPT, Dolly, Mid Journey, or similar?

Interviewee: [redacted name], my bachelor partner, and I are currently working on trying to implement AI on systems with very few resources, so yeah, we've done a lot of work on AI.

Interviewer: Okay that's good. So my next question would be how many times do you use it during the week?

Interviewee: Almost every day currently.

Interviewer: You answered my question: Do you only use it for these purposes, or do you also use it in your personal life?

Interviewee: Both in my personal life and when studying because I feel like it's not a good thing to get results from, but it's a good tool to use to achieve some results or get towards some kind of point.

Interviewer: Okay, perfect. So next, I will ask you a few questions, and your answer should be on a scale from 1 to 5. And I will tell you what that scale is. So, what is your knowledge of machine learning AI? Where 1 is I don't know what it is and 5 is I have taken a course and researched it myself.

Interviewee: I've had AI as a course, so I guess 5.

Interviewer: Are you interested in understanding how AI works? Where one is no, not at all, and five, I am very interested in this topic.

Interviewee: Five. I love AI. That's why I've taken this course and why I'm writing my bachelor's on it.

Interviewer: Perfect. Okay. Now, if you can move a bit closer so you're not in the sun. and yeah my touchpad is a bit messy so if it's easier for you to utilize the mouse yeah okay so now I will start by showing you this is the original GANPaint interface it's a project that was released by MIT in 2019 so it's not something we've done at all it's just how the project looked like in 2019 and And I would like you to look at it and describe in your own words what you can see on this page.

Interviewee: A picture and some options for objects or natural environment things. I'm guessing from everything that is some kind of editing AI or generative AI where you can pick something and then you can draw on it and it will try to draw the object.

Interviewer: So, what's your first impression when looking at it, anything that caught your attention?

Interviewee: The UI is pretty simple, there's not a lot put into the page layout, but the simplicity makes it easy to navigate for testing purposes.

Interviewer: Would you like to, for a few seconds, I don't know, 30 seconds or so, to play with it?

Interviewee: Sure.

Interviewer: I don't know if you can see the screen properly, but let me know if you

want to change that.

Interviewee: That's fine, I can see.

Interviewer: Yeah, so it's not totally perfect in generating things apart from each other, it's not great at trying to layer things behind and in front of each other, because it seems like it has some sort of natural layering where, at least with the trees, it's trying to front, like put them in front, but that fucks up some of the generating from some of the other items that you try to draw. It sort of gets pixels in between, it seems like, and it boxes things in, because I can definitely see the edges on the tree that I'm currently trying to draw behind or in front of.

Interviewer: Maybe a helping question, so what do you think the buttons on the left side of the image mean?

Interviewee: These? Exactly. I mean, simply put, I think that is the option that you're trying to draw. So, for example, if I select tree and begin to draw, it will try to draw a tree. If I say grass and I draw it, it draws grass, door, sky, cloud, so forth, yeah, so forth. Pretty straightforward.

Interviewer: Okay, so now if you could tell me in your own words what do you think is happening behind the scenes as you are drawing on the picture?

Interviewee: I am guessing that the picture is trying to blend and find some sort of total object. So for example, if I choose tree and I try to draw a tree, it will try to find some reference to a tree that fits about the size or resizes something to whatever you draw, and then tries to blend in the background or the foreground to whatever that is drawn through some algorithm, but I don't know what algorithm you're using.

Interviewer: No, that's okay. That's completely okay. That's very well. Okay. So this was, again, the original campaign from MIT, and now I will show you our project, and we're going to start with the front page we've developed. And before you get into reading and everything, what's your first impression when looking at it?

Interviewee: The front page? Yes. It's way more thought out. It has color, which the other one didn't have. I haven't read the text, but it seems to have some introductory text, which is always nice when you're trying to learn or try out something new. It has a, what seems to be, a descriptory model to show how the behind -workings of the generation works. It is definitely a step up from just the bland side that was the other

one.

Interviewer: Okay, so if you have the time to go a bit through the text and just read a bit through it and play a bit with our interface.

Interviewee: I might be biased because I also know a lot about AI, but the descriptive text makes a lot of sense to me. But that could also be because I've worked a lot with it, so I don't know how it would be for someone who hasn't. We're also taking that into consideration. We're testing this on multiple people with different backgrounds. That's why also we had the introductory question. So it helps us a lot if you give your honest opinion. Like, don't think you're biased, just, you know, tell us your opinion.

Interviewer: I'm currently looking at the slider that activates the neurons for the trees. I'm guessing this affects the weights. It's fun to see how affecting the weights of something that would affect the picture in a lifetime is a great way to showcase how it works.

Interviewee: You have some almost question and answers like things in the bottom, that's also always a nice thing to have.

Interviewer: We're not going into that, but it's nice that you're looking forward to it. Okay, so you got to look a bit at the interface and probably read the first part of the text. So now I'm going to ask you to have a look at this graph, and if you tell me in your own what you think it represents, just this part of the hair?

Interviewee: That's the neural network. It has the input, the different layers in between, which we usually call the hidden layers, and then the output layers at the end, which usually represent some... The input, in this case, is a picture. Often it represents it as some kind of matrix with a lot of colors on it. It puts it in, it has some sort of, it does some sort of change to it. I'm guessing this is where you change the weights of what I can see in the input and output.

Interviewer: So yeah, it does some sort of manipulation of what is on the picture and then it spits out an output.

Interviewee: Okay, so you said this looks like a matrix, right? No, I'm guessing the input and output are matrixes because this is a neural network. Yeah, because neural networks are represented by nodes and edges, with nodes being the things that are inside layers, And they all have edges to the next layer.

Interviewer: Okay, perfect. So, what do you think one layer represents? Like you said, these are different layers, right? So what do you think this represents in this graph? Hmm that I wouldn't know a lot. You can see there are a lot of circles inside. Yeah, I'm guessing the circles of the nodes You have not represented the edges here, but that would also fill out a lot of, it would be, it would fill out a lot of edges because nodes and neural networks have edges to every other node in the next layer, but yeah, I'm guessing the circles are nodes. Is it unintuitive to go over the, you know, there is a question mark? And then, you know, there are some explanations, but I feel like you haven't looked at those. Is that counterintuitive?

Interviewee: I don't think it's counterintuitive. I think if I were to use more time, I would have... And I was trying to learn more about it, I would probably use these. I think you could probably give them another color than grey, the same color that you used for the notes. It gives a little bit of association between them. They would stand out more if they weren't grey, I think.

Interviewer: Thank you for the suggestion. And now, if we go down a bit, that's perfect. Now, after we have the slider, I would love for you to play a bit with it again.

Interviewee: As I said, I still think it is affecting the different weights of the nodes in the, which probably is the nodes that represent the trees.

Interviewer: And now with layer seven, please.

Interviewee: Here, you can definitely see a difference between the fourth and last layers. The first one is very focused on building outwards from the trees, while this one is more focused on the whole picture, turning the whole picture green if you move the slider.

Interviewer: And now with layer 10, please.

Interviewee: Yeah, okay. Now, I would like you to do what you did before and choose layer 4 in the top right corner, yes. This is the application called Scan Art. Again, what is your first impression of looking at it?

Interviewee: It has, it's sort of as simplistic as the first one, Though white background works better than grey background in this case. It has more descriptive buttons, but I don't think a start user would know what IOU would mean on these buttons. You also have the low, medium, and high buttons down here, but again, a starting

user would not know what low, medium, and high were in this case. I think these would benefit from example -wise the question marks that you had in the last or some descriptive text.

Interviewer: So how do you think this page relates to our front page that you saw before?

Interviewee: It relates well. They have the same color scheme other than the light blue which I don't think I saw on the front page.

Interviewer: Do you correlate it with somewhere here?

Interviewee: Not the light blue, other than maybe the notes.

Interviewer: And if you go to layer four again, do you think it's similar, it relates to the colors?

Interviewee: You can see the relation, but you probably wouldn't have thought much about it if you didn't think hard about it.

Interviewer: Okay. Makes sense. Okay. Now we're going to play a bit with it, and I would like you to draw a

tree. Yes.

Interviewee: Yeah?

Interviewer: Yes. And then if we go back to the front page, and then layer 7 in the top right corner, let's make layer seven and if you would be nice to draw a tree again

Interviewee: you can come out there yeah okay and now we do the same with layer 10 nice okay i don't know how much you can see because of the screen but I can sort of see. Yeah, okay. So, what are your thoughts after this process of drawing on each layer?

Interviewee: The fourth layer is definitely trying harder to put in the whole tree with more context, but does not care a lot about what it is manipulating in the background. Whereas the higher you go in the layer, for example, here in the 10th layer, it's not trying to generate the whole tree, but it tries to keep a lot of the background. So it is more like it colors and textures the things that are already on the picture than it does drawing a whole tree.

Interviewer: Yeah, I think that's a great answer. And now you can do whatever you want on the page, also on the front page and the layers pages. And I would love you

to free play a bit with it for about a minute or so. and to speak out loud, if it would be okay for you. It's like thinking out loud about what you're doing.

Interviewee: I'm mostly interested in trying to, so now going back to layer four and I'm trying to, I'll just try to draw a cloud over what is the building. On layer four, I'm expecting the building to almost disappear when I draw a cloud over it. It doesn't disappear, but it changes a lot. And if I try to take a tree and draw in front of the building it will definitely try to draw a tree and manipulate the foreground a lot but if I go back and go to layer 10 and I try to draw a cloud over the building it should not yeah it colors the building it more warps what is the building and does not try to blend it as much so it's almost more like a paintbrush than is a generative tool that takes in a lot of blur from the whole building, or the whole picture, which shows a lot of difference between how the layers work on the picture.

Interviewer: Nice. Would you like more time to explore? I also have some of the last questions.

Interviewee: You can ask some last questions.

Interviewer: Okay, so what's your overall impression of XGANArt, our application?

Interviewee: I think it's nice. It is, of course, a little bit limited to what the options of generating are, which therefore is around seven, eight? Eight options. So, of course, it is not as powerful as, say, the bigger picture generation software like Stable Diffusion and such, which take prompts instead of drawing. So that's like the strong suit of this. It's the fact that you draw on it instead of giving prompts. I think it works well for the purpose that it does, especially on some of the lower layers, but it gets a bit wonky the higher you go in layers. Yeah. But I think the overall layout of the site is nice and I think the text is descriptive.

Interviewer: Was there a clear progression through the site, and were you confused at anything at any point?

Interviewee: I think the progression is good. I would say when it comes to progression through the site, I think it's sad that when you enter a layer, there is no direct path to another layer. So when I click layer 4 here it does not have an option to go forward to layer 7 or layer 10. I have to go back to the main page to then go to layer 10. I think most of it is very intuitive. I also mentioned this, but the options of low, medium and

high and the IOU written on the different options is not very descriptive, especially if you don't initially understand what it means. You could sort of guess your way to low, medium, or high, being the strength of the brush you're using, but in essence, you wouldn't know directly from trying to interact with it. But otherwise, do you have any comments for us? Something that could be improved?

Interviewee: I don't think I have any comments other than the things that I've just mentioned. I would take what I've just mentioned as feedback to things that could be improved.

Interviewer: Perfect. Thank you very much for your time.

Interviewee: No problem.

B.6 Interview 4

Date: 05/05/2024

Interviewer: Alright, perfect. So, actually, before we start using the application, I have a few intro questions. So, how old are you?

Interviewee: I'm 20, going 21.

Interviewer: Oh, okay, fine. Okay, so what are you currently studying?

Interviewee: I'm currently studying for the Bachelor of Data Science.

Interviewer: And what semester are you?

Interviewee: I'm in the fourth semester.

Interviewer: And have you used generative AI before, such as ChatGPT, DALL·E, Midjourney, or similar?

Interviewee: Yes, I have used ChatGPT, I think, mostly.

Interviewer: How many times a week would you say you use it?

Interviewee: I don't use it every week, more so often during exam season.

Interviewer: So that's what you use it for, it's for schoolwork?

Interviewee: Yes, when I cannot answer, when I cannot ask a human, I ask ChatGPT.

Interviewer: Yeah, okay, perfect. So now I'll ask a few questions where your answer should be a scale from one to five. So first, what is your knowledge of machine learning/AI? Where one is, I don't know what that is, and five is, I have taken a course

about it or I've researched it myself.

Interviewee: So, five is I have taken a course and the maximum is ten, right?

Interviewer: No, five. The maximum is five.

Interviewee: Oh. Well, I have definitely taken a course in machine learning, so I will say five, but I'm not so good at programming yet.

Interviewer: But you have, like, an understanding of it?

Interviewee: Yes, I do have an understanding.

Interviewer: Very nice. And are you interested in understanding how it works, where one is, no, I'm not interested at all, and five is, I'm very excited about it?

Interviewee: Totally 5.

Interviewer: Totally 5?

Interviewee: Yes.

Interviewer: Perfect. So now, let's get to the more fun stuff. So first, this is the original GANPaint interface, which is open-source code that we work from. So this is just their code, which I will ask you to go through and tell me what you are thinking about it. So first, can you describe what you're seeing on this page?

Interviewee: Yes, so I'm seeing a web page of the application GANPaint painting with generative models. I see a picture of a building, and I see some buttons and more pictures at the bottom.

Interviewer: Yeah, okay, perfect. So what's your, what was your first impression looking at this, like anything that caught your attention?

Interviewee: The interface is not very modern. It's very old-fashioned.

Interviewer: Yeah, a bit old-fashioned.

Interviewee: However, it is very clean and it's easy to read.

Interviewer: So, can I ask you, what do you think these buttons mean?

Interviewee: I feel like they're keywords, like what are the main objects on the picture. I mean, I'm not sure, but yeah, keywords.

Interviewer: Keywords.

Interviewee: Although can I touch it or just move on?

Interviewer: Yeah, you can try to use it.

Interviewee: Okay, so they are buttons then maybe there to add some objects to

the picture.

Interviewer: You can try to use it; you can try to play a bit around with it.

Interviewee: It's my time.

Interviewer: It's your time, yeah.

Interviewee: Okay, that was interesting.

Interviewer: Yeah? Okay, I'm done.

Interviewer: Okay. So now, what do you think these buttons mean?

Interviewee: They're objects that you can add to the picture that are going to be artificially adjusted to look more natural. Interviewer: Yeah, perfect. So now, can you try to explain to me what you think is happening behind the scene as you're drawing something?

Interviewee: What I think is happening is you're trying to add some object and maybe the AI knows what grass looks like, what types of grass can be added and based on the shape that you're drawing, it tries to add grass in the most natural-looking way.

Interviewer: So now I'm going to show you our version, which we call Egg Scanner. So before showing you the actual application, we also developed the front page. So, we will start at that piece here.

Interviewer: So, now I will... So, what's your first impression looking at this? You can scroll down a bit.

Interviewee: What do you think this part does?

Interviewee: I think... Oh, sorry. I think this is a graph of the different layers of a neural network. And it's very understandable for me, but I'm a data science student, so...

Interviewer: But it's easy for you to understand.

Interviewee: It is, yes.

Interviewer: Can you explain what you do understand in the graph?

Interviewee: Yes, so I can see the image as an input. And then latent vector, I guess it will be the flattened image. Then it goes into the first hidden layer. And I don't completely understand what layer 4 is exactly. And layer 7 and 10 when we have the hidden layers in between. So this is something that is new for me but after the last hidden layer we get the output which is the new picture.

Interviewer: Okay, so maybe the layers 4, 7, and 10 are the objects that the users add, is what I'm thinking.

Interviewer: Okay so now I will ask you to do a few simple tasks, so please remember to think aloud as you try to go through them. So first, have a look at the interface and you can move your mouse around and different things.

Interviewer: And see if that, does that give you any, like what's your reaction to different things?

Interviewee: Okay, so maybe those dots representing neurons represent different objects, different types of sky, different types of grass, door, wall, and then we have again trees and again grass. Now can you, if you see this pipe, can you try to use the slider?

Interviewee: Oh wow!

Interviewer: What are you seeing?

Interviewee: I see that the amount of tree changes. We get more and more vegetation.

Interviewer: So now this is the graph for layer 4, right? Can I ask you to click on layer 10 and then try to use the graph again?

Interviewee: It gets blurry when I put it back so much as a real image. It's like the camera that took the picture is broken. Now that the slider moves up and further away it again gets a bit weird.

Interviewer: So now instead of having just more vegetation as in layer four, it's like the picture gets added some filters.

Interviewee: Now, you can also play around a bit with all the different parts of this interface before going up here. But yeah, just look at the front page and see.

Interviewer: And tell me what you're seeing. Interviewee: So I see the different units, and we have wood, stone, foliage, which are materials, then trees, sky, which are parts, some object, window, door, and in layer 7 when they move the slider it's like the style of the picture changes. We have less and less greenery when the slider is close to minus, and then the dimension grows also gets more foliage before like it's completely enclosed so I guess in layer 7 it controls the amount of greenery or something like this, in layer 4 I think it's kind of the same but the house maybe is... no, still layer 7 and 4

look very similar to me at first glance. I think it's layer 10 that is most different and makes the most interesting changes to the image. But I cannot guess the connection yet.

Interviewer: Is there any other part of the interface you want to explore?

Interviewee: Yes. I want to see if there are any more buttons.

Interviewer: Yeah.

Interviewee: So for each layer, the identified features are shown in this graph. Interviewer: Okay. What does that mean?

Interviewee: Well, it makes me think that the different neurons correspond to different parts of an object, like one neuron, 3T, corresponds to the upper part, and 3B, if it's the bottom, then we can clearly see them in the graph the neural network. So now let's move on, and can I get you to click on layer four up in the top right corner? So up here.

Interviewer: So what is your first impression of this site?

Interviewee: This site, well it looks a bit weird on the sides and a bit too sharp the image. At first, it looks very normal, but when you look closely, it's a bit wonky.

Interviewer: And how do you think this page relates to the front page, so the page you wanted just before? How it relates?

Interviewee: I mean, apart from being generated by the same application, I cannot really guess.

Interviewer: Okay, perfect. So now I will ask you to do a few more tasks and remember to think aloud as you go. So can I ask you to draw a tree?

Interviewee: To draw a tree?

Interviewer: Yeah.

Interviewee: Well, when you tell me to draw a tree, I will just press the tree button as I did on the previous layer. Here. and then I'll try to add one here on the side and I'll draw it from top to bottom and hope the AI knows what I mean.

Interviewer: Yeah okay and can I now ask you to go back to the front page and then go to layer 10 up in the top right corner?

Interviewee: And can you now draw a tree again?

Interviewee: Okay, I again look for the tree, but on the side. And I'll put it here

when it's more empty space. And here is tree-like structure.

Interviewer: So, do you notice anything different?

Interviewee: I feel like the trees on this are definitely worse than the trees on the first picture and the colors are also not very... the colors do not go well with this type of picture. So I don't know.

Interviewer: Realistic.

Interviewee: Yes, it's less realistic.

Interviewer: Okay, so now I will, you can just do some free play. So you can either stay in here, you can go to the front page, you can play with anything you want. Yeah, just do whatever you want, free play.

Interviewee: I think I like this, oh, it's a new one.

Interviewer: Is that supposed to happen? Was I here before?

Interviewee: Yeah, what do you think happened?

Interviewee: So, I'm pretty sure I wasn't there before, so maybe it just generated a new picture with the same prompt, because it again looks like a cathedral or some type of church, so I guess just created a new picture with the same prompt.

Interviewer: Yeah I feel like adding a bit more grass.

Interviewee: I guess yeah you can play around more if you want anything you want.

Interviewer: Don't tell me that it's dangerous.

Interviewee: No, it's very helpful for us, but you can also just let me know when you're tired of playing around.

Interviewee: I hope you have a lot of free time.

Interviewer: Oh, we do, or you know it's bachelor time.

Interviewee: Okay I think I'm done.

Interviewer: Okay, perfect. So, can you now try to tell me what you think is happening behind the scenes when you're drawing something? If it has differed from when I first asked you that question. It's okay if it hasn't.

Interviewee: Honestly, I don't know, I think I'll stick with my previous opinion that choosing different objects from the side buttons activates different neurons in the neural network that will generate those objects in the picture.

Interviewer: Okay, so now I'll just do some wrap-up, so like what was your overall

impression of not the first page I showed you, the gray one, but the stuff I showed you afterwards?

Interviewee: Um, I'm not sure what you mean. Interviewer: Okay, like your overall impression of not this page, but this page and also the front page. How many people would you say?

Interviewee: Three? Yeah, both pages are very clean, very nice. Design is very understandable without being overwhelming. So I quite like it. I'm not really sure if in those graphs, layer 4 is an additional layer that's not part of the hidden layer, or it's just like a peak inside the hidden layer. So that's what I'm not really sure about. So that's a bit confusing.

Interviewer: Yeah. Do you think there was like a clear progression through the site? Interviewee: I mean I did not see them until you pointed them out so maybe just make them a little brighter color because with everything like all the buttons and interesting stuff happening here I did not even think about looking at the top corner yeah I think otherwise I mean those questions what do you think changed they were a bit unclear for me because I I'm not sure what actually.

Interviewer: But then it's good for us to know that if you are like, oh, nothing, then yeah, if that makes sense.

Interviewee: But yeah.

Interviewer: Thank you so much.

Interviewee: You're welcome. It was very helpful.

Interviewer: Of course. Thank you so much for helping us.

B.7 Interview 5

Date: 05/05/2024

Interviewer: So welcome, and thank you for joining us today. Do you have any questions before you start?

Interviewee: Yeah. What's this award like about?

Interviewer: You're going in blind.

Interviewee: I'm going in blind. So that's... I cannot tell you.

Interviewer: Okay, so when you say don't ask or ask questions, any questions you...

Interviewee: Well, during the interview or like... Like that's a question I cannot ask you.

Interviewer: All right. Answer you right now.

Interviewer: So, before we start with the actual application, I will ask you some introduction questions. Okay. How old are you?

Interviewee: I am 24, turning 25 in a few days.

Interviewer: Okay, happy early birthday.

Interviewee: Thank you.

Interviewer: What are you currently studying?

Interviewee: I'm studying computer science on my fourth semester.

Interviewer: So, master's?

Interviewee: Yeah.

Interviewer: Okay, perfect. Have you used AI before? For example, ChatGPT, DALL·E, Midjourney or similar?

Interviewee: Yeah, for a few years I've used it.

Interviewer: And how many times do you use it during a week, for example?

Interviewee: Like how many times I ask them a prompt or how much I in time use it?

Interviewer: I would say more like in time, time-wise, yeah.

Interviewee: Okay, I probably use it a few hours.

Interviewer: A few hours, okay. Every day?

Interviewee: Yeah, almost.

Interviewer: Okay. And what's the purpose? Like what do you use them for?

Interviewee: Mostly inspiration and like rewording of text.

Interviewer: Next I will ask you a few questions where your answer should be on a scale from 1 to 5. I'm sorry. OK, so I will tell you what the scale is for each question. What is your knowledge of machine learning AI? Where one is, I don't know what it is, and five is, I have taken a course, researched it myself. On a scale of one to five.

Interviewee: Four.

Interviewer: OK. Are you interested in understanding how AI machine learning works? Where one is, no, not at all, and five, I'm very interested in this topic.

Interviewee: A five.

Interviewer: Okay, nice. Okay, now we're gonna take the practical part and I would like you to look at this page. This is the original GANPaint interface released in 2019 by MIT, so it's not our project, it's the original project. And I would like you to describe in your own words what you see on this page?

Interviewee: Yeah, it's very plain. Some text that describes what it does. And you know my touchpad is not very good, that's why I have the mouse, the placement is not very good, but... Yeah, it's very simple. I have a few buttons to click on.

Interviewer: So what are your first impressions looking at it? Anything that catches your attention?

Interviewee: Not really that I want. I want to try all the buttons and see what it does, but I think it's very plain. So, nothing really catches my attention, but usually, I mean, the big feature in the middle certainly does catch my attention, the canvas.

Interviewer: So what do you think the buttons do?

Interviewee: On the left side of the... I assume it will... Yeah, I'm actually not quite sure what it does.

Interviewer: Would you like to play with it or you know how to play with it?

Interviewee: It's like if you select a button and you try drawing with them.

Interviewer: Ah, I see, okay. Maybe you can also try drawing on the other side of the page.

Interviewee: I'm going to make the tree very large.

Interviewer: Yeah, okay, I see. So, now, what do you think the patterns do for me?

Interviewee: So, it selects a feature that I can paint with on the canvas.

Interviewer: And if you can, in your own words, describe what you think is happening behind the scenes while you draw.

Interviewee: I assume that it...

Interviewer: There's no wrong answers. You can say anything you would like.

Interviewee: Yeah, then I'm not quite sure what it does behind the scenes.

Interviewer: Okay, that's it, that's okay.

Interviewer: Okay, now I'm going to show you our version of GANPaint, what we developed. So this is the front page we developed ourselves. So if you just look a bit

at it, what is your first impression? Anything that catches your attention?

Interviewee: Yeah, I mean, I love...

Interviewer: You can also go into scroll mode.

Interviewee: Oh, okay. So, I see an introduction to the website, which tells me a lot about what the website is doing. There's some text in bold, which I assume is very important to read, so I'm gonna read that.

Interviewer: You can take your time for a minute or so to go through the interface and maybe read some text. Thank you. Don't feel any pressure. As I said, I'm not in a hurry, so take your time.

Interviewee: Okay.

Interviewer: When you're ready, you can also think out loud, because I would love to hear your thoughts when you are going through it.

Interviewee: Yeah, I mean, currently I'm just quite interested in looking at the figure, which explains what input goes through and what's happening throughout the entire process up until the output. There are some question marks that explain each step. And it's, I think, like for an introduction part, a lot of the words might need like a bit more explanation, so the wording of things always... yeah, it's technical and like it's... you're already supposed to know what encoding of data means or... yeah.

Interviewer: Oh, and then I can zoom down and look at more facts about machine learning.

Interviewee: I think it's interesting if you really want to know more, right?

Interviewer: Yeah, yeah. It's after you play with the tools, you will probably go down there.

Interviewer: So, can I ask you a few additional questions, helping questions?

Interviewee: Sure.

Interviewer: So, if you look at the interface in this part, what do you think this graph represents? Like, you said it already, but more if you can repeat that, because, yeah. So we take it step by step. Just this part.

Interviewee: Yeah, I assume it... What it represents... I assume it represents the behind the scenes of how the image goes from changing depending on the input.

Interviewer: I don't know. I'm not quite sure how I'm supposed to use it, I think.

Interviewer: That's okay, just tell me your honest opinion. It's completely okay if you don't understand how it works. You can tell me what you don't understand and how you think it should be, you know? Like, just think out loud.

Interviewee: Yeah. Maybe I'm... I'm... I'm just not quite sure how I am supposed to engage with it, I think.

Interviewer: Each dot in the visualization represents a different feature captured by the network, which you can also see in the bar graph below. The tree feature is highlighted in red indicating that you can adjust its activation using the slider.

Interviewee: So using the slider I'm able to put a lot of trees or not.

Interviewer: But I can select other features I think that would be fun.

Interviewee: Yeah, that would be the perfect world, but for now it's just trees.

Interviewer: And then in the next one

, the tree is still selected? And I assume the same thing is going to change? Do you see, is it the same thing?

Interviewee: Like if you play with it on each layer? There is a lot of trees? And then if you go again to layer 4 and you play again, do you see any difference? And maybe layer 10, try it also.

Interviewee: Oh my god, that looks horrible. I enjoy layer 4 much more than the other layers.

Interviewer: Can you tell me what you see in layer 10?

Interviewee: Yeah, in layer 10 it's much more distorted and even the color of the picture changes. And in layer 7 the picture is more like distorted but in a tree kind of way.

Interviewer: So what do you think happens when you play with the slider?

Interviewee: So, then I assume that the more that I put towards the minus, the trees will disappear and when I put it towards the plus, more trees will appear. And I assume the more layers that go through, it goes through the more dominant the trees will be?

Interviewer: There is no wrong answer, so whatever you think is right.

Interviewer: Maybe we can try another experiment, so we still play with the layers, if you go to layer 4. Can you look a bit at this application and what's your first

impression again of looking at it?

Interviewee: It's the same as GANPaint, the first one, so I assume it, and I can choose between low, medium and high. Not sure what that means, but I can choose between it.

Interviewer: Then whenever I draw something, I can see what I've drawn and undo those if I don't like whatever change I've had.

Interviewer: Do you think this page relates to our front page in any way?

Interviewee: Yeah, maybe it's like...

Interviewer: It's... Oh, it relates in what way?

Interviewee: In any way, that's the question. Do you see how it relates to the front page?

Interviewee: Yeah, I mean, I assume this is... This is when, if I changed or when I pressed on layer 4, I was able to see the change. So I assume that now I can just repaint in layer 4 instead of having to use a slider.

Interviewer: Exactly. And now if we make an experiment, can you draw a tree for me? You can also choose any picture you want, right?

Interviewee: I can choose any picture?

Interviewer: You see here. It looks like you can pick your own picture. And if you draw a tree for me, anywhere.

Interviewee: There you go. Multiple trees.

Interviewer: Nice. And then, if we go to the front page.

Interviewee: Yeah.

Interviewer: And we can see that there are a lot of pictures.

Interviewee: 10.

Interviewer: Up here?

Interviewee: Yeah, exactly.

Interviewer: And draw a tree again, different picture, and then draw a tree again.

Interviewee: I'm gonna choose the same one.

Interviewer: So what are your thoughts? Do you notice any difference?

Interviewee: Yeah, it's not really a tree. It's just green blur.

Interviewer: Exactly. So, what do you think is happening behind the scenes and

how do you think it works, seeing this, playing with it a bit?

Interviewee: Yeah, I don't think I know. So, I'm unsure how to answer the question.

Interviewer: That's completely fine, you made some great points that I'm gonna tell you about in these experiments. But then what would be your overall impression of Egg Scanner, like our application?

Interviewee: I think it's quite fun to be able to play with the different kind of layers. And then you can see like, for example, that layer 10 couldn't create a tree that, for example, layer 4 could create. Now I'm just testing out layer 7 to see what it comes up with.

Interviewee: I'm just going to try and reset. Sometimes it works, sometimes it doesn't, I guess. There's a tree behind there somewhere. But yeah, I think it's quite fun to play around the different layers and see how different paint styles will actually be the outcome.

Interviewer: Was there a clear progression through the site? Were you confused at any point? Was that something unintuitive somewhere?

Interviewee: Yeah, I think the links for the different layers could be better. Maybe you could even put it on the boxes, so you could go to that specific layer. Because I feel like it was kind of forgotten up here. Maybe it should be in the middle or somewhere where you actually notice that they're there. I haven't read all of these facts, but I'm sure they are very informative. They're kind of like a little bit forgotten down there, I feel like. Maybe they could be also somewhere else, like maybe if there were some hyperlinks up in the text that took me down to the boxes, I think that would be... So if the user wanted to know more about neural networks, they could click on neural networks and then it will go down to the neural networks. But else, I think it's very nice and fast, not slow.

Interviewer: But is there anything else that confused you, it was mostly that, I don't want to lead you in a way but you told me about that you don't really understand what's happening behind it, like what confused you in that way?

Interviewee: I feel like it was the way I was supposed to go through the website, so I feel like that surprised me the most. I didn't know I was able to click on different layers and see what actually happens on the slider, but I guess it's because I only skimmed

the text and didn't like thoroughly read it.

Interviewer: So, yeah, I think that's mostly that. Do you have any other comments or suggestions for us? Something that can be improved?

Interviewee: You also made a lot of good suggestions, so if you have any additional ones, that's completely fine.

Interviewee: I think it would be awesome to be able to select different features than trees. I think that would be a lot of fun too. But other than that, I don't think I have any extra comments. Yeah, that's okay. I think I should just read on the page before I use it. I'm very like, oh, I just want to try stuff out.

Interviewer: Yeah, that's understandable.

Interviewee: It's fun. Because it looks fun to play with, right? So I think if I went into the page with another mindset of wanting to learn, then I'll definitely read more about the page. If I was told to study the space, I would go in with a different mindset.

Interviewer: That's perfect. Okay, thank you so much for your time. I hope you enjoyed this.

B.8 Interview 6

Date: 05/14/2024

Interviewer: Thank you for joining today. Do you have any questions before we start?

Interviewee: No, not really.

Interviewer: All right, so first of all, some introduction questions. How old are you?

Interviewee: 23.

Interviewer: Right. And what are you currently studying? What's your program and what semester are you on?

Interviewee: I'm currently in my second semester, and I'm studying Data Science.

Interviewer: Bachelor?

Interviewee: Yes.

Interviewer: So have you used generative AI before, like ChatGPT or DALL-E or Midjourney or something similar?

Interviewee: Yes, I have, ChatGPT.

Interviewer: And would you say you use it regularly? How many times a week?

Interviewee: I haven't counted, but I could estimate that almost every day.

Interviewer: And what do you use it for?

Interviewee: I ask for consultation. I ask ChatGPT so I can consult with it and I ask questions like, "Is my method correct? What would you do?"

Interviewer: So this would be in regards to schoolwork?

Interviewee: Yeah, to schoolwork and projects and stuff.

Interviewer: So I'm going to ask you two questions, which you have to answer on a scale from 1 to 5. The first question is: What is your knowledge of machine learning or AI? A 1 would be, "I have no clue," and a 5 would be, "I have taken a course or researched it myself to a point where I feel like I have a good understanding."

Interviewee: From 1 to 5? 1.

Interviewer: 1. So you don't know what it is?

Interviewee: No.

Interviewer: Okay. Are you interested in understanding how AI or machine learning works, where a 1 is a no, not at all, and a 5 would be I'm very interested?

Interviewee: Can I say 4.5?

Interviewer: Okay, cool. Now I'm going to show you a demo for the GANPaint interface. I'm going to start screen recording and then here, if I could navigate to it, please.

Interviewee: Okay.

Interviewer: Here we go. All right. So look at this for a few seconds, like, what do you see here and what's your impression of it?

Interviewee: An old building, I see, and it's probably not well kept, like renovated, like other buildings that I've seen that are well kept. This, what else do I see? It looks like a castle or like a haunted castle or house from a Halloween movie. Yeah, these buttons on the side of the picture, the buttons, what do you think? About the... This is the tree, girl.

Interviewer: Also, you're allowed to do stuff if you'd like.

Interviewee: I don't know what a dome means. What does that mean?

Interviewer: Well, I mean, a dome, in general, is like...

Interviewee: Ah, is it the... Yeah, the building. Right. Ah, these words clearly describe the elements from the picture. A tree... The grass... Wait, I actually don't see the grass, but okay. A door, the sky, clouds, I don't see the clouds either. Yeah, bricks, yeah, the buildings.

Interviewer: Okay, so why do you think they are there?

Interviewee: They, I think they were just put randomly.

Interviewer: Okay, so what if I asked you to draw something on the picture?

Interviewee: Okay.

Interviewer: Please do.

Interviewee: Okay.

Interviewer: What do you think is happening?

Interviewee: I just, like, I don't know. I just morphed the image. I thought I wanted to draw some clouds, but on the left side, I see a weird-looking fish. That is staring at me, weirdly, yeah, a sad fish staring at me, and I really wanted to make a cloud, but I couldn't, two clouds, that's okay, now it looks uglier.

Interviewer: You don't think you can click on some of these?

Interviewee: Yeah, I can. Oh, wait a minute. Ah, okay. So, of course, if you click on draw and cloud and then draw, try to draw a cloud, no. Okay, you know what I thought? That the cloud, that every little word, the tree, grass, door, cloud, they are there to help you draw a cloud if you want to draw a cloud. So I thought that it would be like a white painting over the picture, but no, it's just morphing.

Interviewer: No problem.

Interviewee: But if I try to... Yeah, it's just morphing it. I thought it would be like for the grass, like a green painting, but no, it's not. So those buttons, the tree, and the grass don't do anything other than morph the picture.

Interviewer: Okay, so the last question for this specific part of it is like, when you're drawing something, what do you think happens behind the scenes?

Interviewee: Oh, maybe, yeah, the pixels changed the positions. Maybe the ones in the zeros, maybe the picture had more masks, like in Photoshop, with more layers, I mean layers, not masks. And if I would use some of the words, like these words, and draw something, then there's a layer for each of these words, maybe.

Interviewer: Okay, and...

Interviewee: And can I do this because maybe it's like... Yes, yes, like this is exactly what I meant, yes, this is exactly what I meant. So like a picture for every word, like a layer, it's like layer over layer, like picture over picture, and that I can see right now.

Interviewer: Then I'd like to move on to the next part. We're going to show you our application, and to start with, we have this sort of front page. If I scroll through it like this, just first impression, anything that comes to mind?

Interviewee: It's colorful. It has some nice graphs or drawings, which I like to look at.

Interviewer: Yeah, but now you can look for it, feel free, click on whatever you want, read whatever you want. I'll give you some time.

Interviewee: Ah, neural network. I don't like neural network.

Interviewer: Also, remember to think out loud.

Interviewee: Oh yeah, okay. Because I actually thought that this looked like a neural network, these things, the image processing drawings, because I tried to document for my project for data science this semester about the neural networks, and it was so difficult, in my opinion. I just gave up, but yeah, it just looked like neural network structures and it is, and what else should I say? I don't know.

Interviewer: You can also take some time to look through it if you have any thoughts to share.

Interviewee: Okay, is it useful if I just say that I don't like the look of the website? It's kind of...

Interviewer: But tell me what it is that you don't like about the website.

Interviewee: Yeah, if you can describe... The title is not centered. It's like whole texts without bullet points. I don't like that. It looks ugly. Some texts are in bold, some are not in bold. Yeah, I like the pictures. I like this.

Interviewer: What else?

Interviewee: Wow, I like this, yes, I like interactive tools on the website. I really like it. I

like this kind of things, if you want to learn more about the topic, I really like this because it explains the neural networks what they are. Yes, exactly this, exactly

that. I like the graphs. They're really pretty. This one is well-structured. I like these explanations, and there's not too much explanation here, so it's readable, which is a bit longer.

Interviewer: I'd like you to do something.

Interviewee: Yes.

Interviewer: Can you make sure you read this part?

Interviewee: Should I read this part?

Interviewer: Yeah. Should I read it out loud?

Interviewer: No, just read it. I'll read it now.

Interviewer: And take your time.

Interviewee: Is there something in your throat? What part is that in your throat? Where are you going? Now you start looking through the graph, what are you thinking as you're doing it?

Interviewee: Okay, so I'm done reading, but honestly it's just like...

Interviewer: If you want to look at the graphs, don't let me stop you. Just tell me what you're thinking as you go.

Interviewee: So these three are the highlighted layers or graphs with random words in it. And I didn't really understand the text. It was kind of like messy and it could have been more clearly explained. So I didn't understand anything from the text. And then by clicking on this, it highlights it like layers, it's like, yeah, I still don't understand anything.

Interviewer: Okay. Yeah. Can I ask you then to click on either layer 7 or 10?

Interviewee: Yeah.

Interviewer: And then try to, like, what do you see happens below?

Interviewee: Okay, so this thing here changes with either layer I click on, and it has like a different number of objects, parts, materials. So these are the materials of... Of course, parts, maybe elements in the picture.

Interviewer: And the graphics?

Interviewee: Can I ask you to click on layer 10?

Interviewee: Yeah.

Interviewer: And then try to play around with the slider a bit.

Interviewee: This one, as you did earlier. Is this the tint? This looks like the tint when you edit an image in Photoshop or on your phone with a photo editor. It looks like tint. This is the pinkish. Okay.

Interviewer: Can I ask you to click on layer 4 up here?

Interviewee: Yeah.

Interviewer: So immediately again, first impression, anything that comes to mind by looking at this image?

Interviewee: I just think the same thing about the first image that I saw like five, ten minutes ago. It has like a neural network thingy going on and that processes this whatever things they are.

Interviewer: So now, yeah, you can draw a little bit.

Interviewee: It kind of looks the same. Oh wow, okay, now I can see the masks that I've made.

Interviewer: I'd like to ask you to do something specific. First, reset, and then I'd ask you to choose the tree and then try to draw something.

Interviewee: Okay, I think more trees are popping into the picture because it's a function for the tree.

Interviewer: Yeah?

Interviewee: Yeah.

Interviewer: And now I'd like you to go back to the front page.

Interviewee: Yes.

Interviewer: And then you're going to layer 10 up here.

Interviewee: Yeah.

Interviewer: And then I'd like you to do the same thing. Draw some trees.

Interviewee: Oh, now it's just making it greener. Maybe the color of the trees? Of the leaves from the trees? Yes. It's just making it greener. The shapes of the trees are not popping into the image. It's just the color. It's not any tree.

Interviewer: Can I try with cloud?

Interviewee: Yeah, actually now we're gonna have like a bit of time where it's essentially free play. So you can go to the front page or you can go wherever you want on the website, play around here, go to the front page or go to another layer. You just

have a bit of time. But you do whatever you want and test things out. And again, if you have any comments or thoughts while you're doing it, speak them out loud, please.

Interviewee: While I'm trying here is to put all the features into this one image to just prove myself to myself that these functions, or like, yeah, these functions or features, these are for the color, and the previous ones from the other layer was for the texture. If you want, we can go to front page and have... Oh, is it this one, right? We went on layer four, but you can try layer seven. Let's try layer seven, which is kind... Let's see what this one does.

Interviewee: Okay, this one masks the image better. I think that the other two layers aren't just bad at drawing and masking. I think this has also the color and the texture and the feeling of the cloud or like the sky, of course. This is the cloud. Let's try the cloud. Yeah, this has both of what layer 4 and layer 10 had separately, like individually, this has it both. When I click on low, this doesn't do much.

Interviewee: Can I click on this one?

Interviewer: Sure.

Interviewee: Yeah, allow.

Interviewer: Sorry, that's my pop-up blogger.

Interviewee: Okay. Uh-huh. This one says exactly what it does.

Interviewee: Okay, here it works way better than on the website I was at. This one really works very good. Yeah, the one was unclear and I struggled to understand what was going on, but this website explains it very well. It's like a better version of...

Interviewer: But what makes it better, if you can pinpoint anything?

Interviewee: The explanations are simple to understand for everybody, even for me. And the text is just less text, more images and examples, and better structured. We have a lot of examples here, and we also have like, if you would like to go deeper into the research, it shows clear where you can click on. On your website, it was like random capital letters, small letters, everything combined. Yes.

Interviewer: All right, then I am gonna ask you one sort of like last question related to specifically, you know, what you've been doing. It's the same one I asked you when you looked at the demo. Can you try to explain to me what happens when you draw?

Interviewee: When do I draw?

Interviewer: Yes, like what happens behind the scenes?

Interviewee: On your demo website, right?

Interviewer: Yeah, the one where you...

Interviewee: Yeah, the previous one.

Interviewer: What happens when I draw? Well, the image changes depending on the features that I use, such as grass or sky. It changes in texture and color, but honestly, it doesn't do a good job.

Interviewer: That's fine.

Interviewee: And it also probably explains what the neural networks do for image processing.

Interviewer: Yeah, that's more the front page.

Interviewer: Okay, well, that's most of this stuff, so I just have a few quick wrap-up questions. And then I'll let you have some cake and go. So you sort of already said it, but just to reiterate, what was your overall impression of our project? So, the websites where it was like white and red compared to the demo and then the other site you navigated to?

Interviewee: Like your project, or like the looks of your website?

Interviewer: Both things, anything.

Interviewee: Okay, the looks of the website. Very ugly, unorganized, and unprofessional. But I think that was on purpose. Yes, and this one... No, yeah, the last website that I clicked on and that I saw was like the better version, and I like that better, and I explained why. And your project seems very interesting, and I like expressing my thoughts and I like giving feedback. But I would not... I would never use your website.

Interviewer: Yeah, I guess maybe you've already answered this as well, but again, was there a clear progression for the sites? Was there anything you were confused about at some point?

Interviewee: On your website?

Interviewer: Yeah.

Interviewee: Yes, I was so confused at the beginning. I needed to figure out what was going on and probably I haven't even figured out maybe half of what was going on there. It was very hard to understand. The explanations were there but were not doing

much, were not helping much. If I would have gone on your website, just search on Google for neural network, and I would reach your website, I would, within maybe one minute, close the tab because I would not understand anything. I would go to another website to explain what the neural networks do.

Interviewer: Okay, then the very last question. Do you have any comments for us? Like something you feel could be improved that you've not already said, or anything else?

Interviewee: Well, the looks of the website

could be improved just like the last one that I saw, the better version. If I saw that there can be, that there is a better version, then you can also make it better, like to look like the last website. And yeah, just make it less confusing, put more explanations, clearer images, work on your masking and the features and the layers, and that's about it.

Interviewer: The last question is sort of more of a personal one. Do you feel like you would have understood it better if you had more time and if you were not sitting beside you monitoring you, but you just sat with it in your own time and looked through it? Do you feel like you would have a better chance of understanding it?

Interviewee: If I had to do that?

Interviewer: Yeah, if you had to. If you were forced to gunpoint.

Interviewee: Okay, then yes, then yes. Also, I would fear if I went to the gunpoint. But definitely, of course. And I would also to better understand your website and what's going on there, I would also use other external websites and information external information from Google to understand what's going on.

Interviewer: Okay.

Interviewee: To supply. Maybe, oh, what's sky, yo, tree, yo, what is this, yo, so I can understand better the feature and what it does and what it works with. Does it work with the background, with the texture, with the color?

Interviewer: Yeah.

Interviewee: Okay.

Interviewer: Any other comments?

Interviewee: No comment at all.

B.9 Interview 7

Date: 05/14/2024

Interviewer: If you have any questions through it, you can ask them, but I cannot guarantee that I will be able to answer before the end. So, are there any questions before we start?

Interviewee: I don't know anything about this project. Should I know anything before I see the website?

Interviewer: Not quite. I mean, I can tell you it's about machine learning and sort of about the use of it, but I will actually just have you test it out yourself. Okay, so I'm gonna start screen recording now. The first thing that's gonna happen is I will show you this website, and you're allowed to as soon as you want, just go and start moving the mouse around, do whatever you like. Oh, actually, that's my bad. I had a few questions to ask you before. So, how old are you?

Interviewee: 21.

Interviewer: Okay, and what are you currently studying?

Interviewee: I'm studying a Data Science Bachelor programme, 6th semester.

Interviewer: So, have you used generative AI before like ChatGPT or DALL·E or Midjourney?

Interviewee: Yes.

Interviewer: Okay, and would you say you use it somewhat regularly? If so, how many times a week?

Interviewee: I do use it regularly. I don't know how many times, but every second day.

Interviewer: And what do you use it for?

Interviewee: Well, I use it sometimes just for research, even for academia, but also in real life. Like, I don't know, I'm sometimes asking about travel stuff, or I ask about like cooking for the... we had an assignment, but also then, in general, I know some substitutes and anything really. Yeah, I am a really pro user. I have ChatGPT 4. I know it's bullshit, but I can use it. I like using it. I think it's great.

Interviewer: And I have two questions for you that you're gonna be answering on a scale from one to five. The first question is: What is your general knowledge of machine

learning and AI? Where one is like I have no idea, and five would be like I have intricate knowledge like I've taken a course or I've researched it myself to the point where I feel very confident.

Interviewee: Machine learning or AI? I would say three because I know I have the course, but I wouldn't know exactly what's behind AI or how it works. I take things for granted. Oh, it knows how to answer my questions, good, great. I don't need to know, I do not care how it works.

Interviewer: Okay, the next question is: Are you interested in understanding how AI and machine learning work? Where one is no, not at all, and five is I'm very interested.

Interviewee: Okay, I'm gonna put a 2.

Interviewer: Okay, alright. And now we're actually gonna do this thing. So if you're gonna look at this page, just immediately, the first impression.

Interviewee: Okay, no, okay. I know I have to give normal answers as I know... So no, but I think it's really like it looks like a beginner site that we would need to do in middle school or high school to create the site. The background is awful. The fonts also, it's a very basic-looking site. And I don't know exactly what's happening because I haven't read it yet.

Interviewer: Yeah, but that's what I'm gonna ask you to do now. You're gonna have, you're gonna be given some time. I want you to sort of, at your own pace, actually take a closer look at it, like what's the text saying, what is the front page. This is what I'm having you look at right now, I guess.

Interviewee: 2019, wow. Okay, I would have thought that it's older, although I know generative models are not that old, but it's... Wow. Okay, it's one of the first tools to allow image editing. Interesting. Okay, so select a feature, brush and select, and strength and range.

Interviewer: So what do you think the buttons beside the picture mean, like these ones?

Interviewee: When I choose the object that I want to paint, then it just does it for me. Draw another tree here. Oh, wow. Okay, that is pretty cool. Let's see a door, random door here. Okay, so it kind of has to make sense where you put the door because otherwise, I see this doesn't really put it. Sky, if I want to put sky down here.

It cannot comprehend. Maybe it has to have some kind of sense because it just moved this down. It doesn't really let me put bricks in the sky.

Interviewer: Can I ask you, then? You can say anything here, but what do you think is happening behind the scenes when you're drawing? Like, how does this actually work?

Interviewee: I'm imagining that it's kind of like Photoshop. Like I don't know exactly how Photoshop works, but I know it has different tools that allow you to delete some parts of the photo and then... I'm thinking something like that. Like it's just reiterating... I don't know what's the word for it.

Interviewer: Rendering?

Interviewee: Rendering. Okay. Over and over the same image, just adding more stuff.

Interviewer: Cool, now I'm gonna have you look at something else.

Interviewee: Oh, I could do undo actually, which is very cool. So I'm guessing it's not lost. You can undo it but not redo it. No. Once you undo it, it's gone. Okay, and reset. Where did I start from? Where was I again? All right, and now I can just choose. And this is all the pictures. There's no more than this. Okay.

Interviewer: Then I'd like you to have a look at something else. So again, look this over for only a few seconds. You can also scroll, and just...

Interviewee: Should I read everything?

Interviewer: No, look, we're going for a first impression first, so look at it for only a few seconds and then just anything that comes to mind.

Interviewee: Okay, so I see it's kind of trying to explain how it gets from the input to like a certain output and these are the layers that what's happening underneath, which is really cool that this was figured out how it works.

Interviewer: Now I'm gonna do something, but it's a lot of information.

Interviewee: Yeah.

Interviewer: But I'm gonna do something now. I'm actually gonna leave the room for maybe a few minutes and I want you to look through this in however much detail you feel like. I will give you some time. I'll leave the room and then again this is the recording so anything you think, please think out loud. But I will be back in some time, but it's just so you have peace and quiet to look at it. I'll be back in a moment.

Interviewee: Okay.

Interviewee: I feel like the other one was just doing, like, you figure it out, and I had explanations, but I feel like I was so wrong in what I said. I didn't think of it out of that box, but it makes so much sense. Okay, I do like the slide by. I like that. Nice feature. Here are some of the questions you might have. Navigate the art division if you want to know more. Just click the glass. It's like a world of machine. Not nearly done. Okay, maybe I should be more interested in what is behind on the yard. I just got so tired of the way I used them. Like I used them, but they just keep giving and giving and giving on.

Interviewee: Paradox truly.

Interviewee: I think I have 6, do I have? Okay, I have 4. Ah, what? Okay. What did I put here? Oh, I like the history. Okay. Oh, and that, oh, okay, so you can learn some circles. That's so far so good. Sorry for the swelling. Alright now, oh and then download your creation and like that. I wonder if that is like what you did or it was already like also in the original one

. Okay, layer 7. Okay, so let's see the difference, low, medium, very high. What is actually the difference if I didn't leave enough? I need to look more into it then remove. What if I remove the clay? Okay, it doesn't work like that. I have to remove what I did. I just messed everything up. I mean it's a nice thing to play, but... Okay, let's try again. What if... Do we do... No. Right?

Interviewee: Okay, wait. Draw. Okay, no, I think we don't need to do much. Oh, yes. The lizard. Hmm, nice. Go to front page, we have layered fan, style, fan. Oh, okay, I wonder if it's. The thing is, I don't know what am I supposed to see, like, I'm just fascinated by it but since I don't know exactly what it was, maybe it would be nice if I had the expertise to know how it should look like.

Interviewee: Yes, so I wanted to ask exactly what is the difference between low, medium, and high. Because I draw a line of each. I mean, it's clear the difference when I like try to draw it next to each other, but I haven't read maybe that part exactly where they ask.

Interviewer: Can you remember that question?

Interviewee: Yes.

Interviewer: Because then I'll answer it at the end.

Interviewee: Okay.

Interviewer: Alright, so useful for us now?

Interviewee: Yes, pretty much so.

Interviewer: Now I want to ask you about this one. Did you notice you can click on it?

Interviewee: Yes. And I think it's very cool.

Interviewer: What happened when you clicked on the different ones?

Interviewee: I can see... Not the buttons, but like the grey thing.

Interviewee: Yeah, yeah, yeah. I saw that. It changes, it explains, like the number of objects, parts, and materials for each layer and how it's added.

Interviewer: And have you tried going on layer 10 and messing with the slider?

Interviewee: Mm-hmm, I did that.

Interviewer: Alright. Then, I would like you to navigate up here and click on layer 4.

Interviewee: Mm-hmm.

Interviewer: Now you see this again, first impressions, but I think just also for my own sake I'd still like to ask you the question like so you've tried drawing on layer 4 and on all of them.

Interviewee: I tried, but...

Interviewer: You're on layer 4 right now.

Interviewee: Yes.

Interviewer: Do you remember if you went on layer 10 and did you notice a difference in how things looked?

Interviewee: I mean, it was also different images, but yeah. It was a difference. Well, I didn't think much of it, but yes. I mean, yeah, I'm sure. But wasn't there also... Or the number of these pictures were everywhere? Building... I don't know. On layer 4 I have more features. I don't mind.

Interviewer: Alright, so you're asking the buttons on each layer page, they change?

Interviewee: Yeah.

Interviewer: The number? One, two, three, four.

Interviewee: Yeah, they do change.

Interviewer: Interesting. How do you think pages like these in layer 4 and layer 10 are connected to the front layer? Like, do you feel there's any sort of connection?

Interviewee: Anything that's...

Interviewer: To the front layer?

Interviewee: Yeah, the front page, the one with the text and...

Interviewee: Ah, it's very intuitive, it's good. They are connected very intimately. Like, when you go to the front page, your eyes, it would be cool if you had, like, the eye tracker thingy to see where my eyes go when I see a new website. But immediately when you see this, you look at the colors because that pops up, and then you see what's up here. Big text, multiple pages. Okay. I think it's very good. But I thought that maybe when I go on 4, I will also see here 7 and 10. You can do those links, if it's not too much. But I like also that, oh, go to the front page and then go back. That's also cool.

Interviewer: Cool. Then I have only a few more questions. So, like, now that you've looked through this, can you, in your own words, explain to me what happens when you draw on a picture?

Interviewee: What happens when I draw?

Interviewer: And maybe without looking too much back on the text.

Interviewee: Yeah. No, I wasn't reading. I was trying to think how well I should say it. So it tries to add that specific object on top of the previous layer, and you don't need to be correct at all, so don't look at me for, like, am I doing right you just have to explain however you understand it. So I do understand that there is an input picture, and then there are these multiple layers that you draw on as you choose this object. It's very cool that you can also, on the original you couldn't see your history or on one of the layers... Oh yeah, no, here you can like if you draw multiple things; you can choose to keep which ones you like. I think that's very cool, but yeah, I don't know. It's a very cool feature, but the functionality... Yeah, I was thinking about the functionality did you add it, or was it also in the original that you can download? I guess it's nice for a thing that you want to play around with, but I wouldn't necessarily use it as an artist or as an animal. In my free time, I don't know, like it's... Maybe generative AI is

the painting, maybe it is the future but why would you want to create... I don't know why would you want this to, from a normal picture, to create like to create a false... I don't really get it. Why would you want another version of this picture, like with the door and the sky?

Interviewer: Well, I mean I guess you could ask the same question to Photoshop users, but I guess, yeah, if you create like an album or like a gallery of them, right? But the functionality, like how the image is changed by your input, you sort of got to, like, okay, you draw on layers and this, and then, and then that's it.

Interviewee: I mean you can also remove them.

Interviewer: Yeah.

Interviewee: Okay, but then, guess yeah, I think yeah you draw on layers. That's what's my understanding.

Interviewer: That's fine it's just because we suddenly switched over to talking about other stuff. Okay, so now I have three wrap-up questions, and we're done. So, using this front page in this part, what was your overall impression of it? Again, this is not a part of our project, so you only focus on like this stuff.

Interviewee: Yeah, no I like it. It explains, I mean I can't say oh it explains so well because I didn't, I tried to read as much as I could while you were gone, but I'm also not the kind that immediately what I did also locks into my brain, therefore, I cannot explain it very well to you right now but if I was to study this website would be very helpful instead of just going into the old one, and okay you figure it out yeah you do you in this one. I don't know if this should be any longer or shorter. As I said, I think I read one, maybe, or two. Hmm, skimmed through. Okay. I would just, yeah, think that you did your research very well in these three-part questions.

Interviewer: Okay, next question is like, was there a clear progression through the site? Now I left you alone to sit with it, but like, at any point where you had to do something or go somewhere, were you like, or when you were reading something, were you left confused or did you feel like you sort of understood?

Interviewee: I sort of understood, I wasn't confused, it wasn't necessary. I just yeah maybe okay I was a bit bored with the... what you said you're gonna explain me after exactly what's with the low, medium and high.

Interviewer: Oh yeah those buttons.

Interviewee: But yeah, except for that, it's a very inclusive site. As I said, if you want to add those buttons here, it's not necessary. But it is more easy to just, oh, two.

Interviewer: Okay, last question is, do you have any comments, any advice, or anything you feel could be improved about this?

Interviewee: I think it follows like this type of site; it has like those kind of themes that, I don't know. When you look for documentation for different libraries and such in science programming, this is kind of the website that you are looking for. I don't think it should be, I don't know, some wow design, I don't know, what interaction, what else can you do more? I think it's really good as it is. I don't think I would. Maybe if the button is bigger, but I'm not blind, I can see the buttons, I don't know. Maybe to stand out more, but again, that's like, I don't know, personal preference, maybe. Yeah. I don't know. And that's also design stuff and this one's yeah because uh sometimes like it took me I don't know those two three seconds to be like oh my god this is so cool that you can actually like see everything see the names on the button.

Interviewer: Yeah all right. Well we're basically done with these questions but you had some questions for yourself. The first thing was what these intensity buttons are at low, medium, and high. Okay, so I mean I am just keeping on recording in case you drop any gold nuggets I can use but the user test is officially over, just so you know. But this is because what's happening behind the scenes here is actually that, you know, we have a neural network with these neurons in them and the project we're building upon was actually some sort of code that could sort of figure out like okay, this neuron is responsible for this specific feature, like all these neurons are tied to like a tree. So what you're doing when you're either drawing or removing is actually that you are changing the values of these neurons on that specific layer because at every layer they've calculated a certain result so what you're essentially doing is just like well I'm just gonna either add to that result like make it more visible or make it less visible like you know take some from the result and these low medium and high intensity are essentially like you're not changing the values of absolutely every unit or neuron that is tied to a feature but depending on the intensity there's like a selected amount so with the high intensity you're taking like the top 15 neurons that are most responsible

for that feature and changing their value, but if you go with a medium intensity then you're taking the top 10 neurons and then low feature or intensity you're taking the top 4 neurons so you're changing the top neurons values but you're changing the values of less neurons depending on the intensity.

Interviewee: It would have been cool if you had been, okay? But that's super extra. Maybe in the future, you can have this and also update it to upload your own picture.

Interviewer: Do you think there already exists such a thing?

Interviewee: Probably. Possibly.

Interviewer: We actually did try to do something like that but actually all of these pictures you see here are generated by the code and we try to find a picture and put it up and then try to do the dissection, but since we're building on another project there was a lot of problems with like getting the code to work so we had to give it up. And that's also sort of another reason why you can from this layer not go directly to another layer because we had to think about like you know you're only actually changing the result at layer four it then goes for all the layers and then the end result is what you see before you but we didn't want to make the user confused that like okay you're changing stuff on layer four and then like with those results will then affect what you see if you go into the site for layer seven. They're not connected in that way. We tried to do that, but we couldn't figure it out, so our supervisor was like, okay, just make it clear that each of these layers is... when the user is doing something with them, they're separated like what you're showing. What you're doing on the page for layer four will not be represented at all in the page for layer seven or the page for layer ten.

Interviewee: So now on layer 4 this is only changing the last 4 neurons.

Interviewer: At the top 4 neurons.

Interviewee: The top, that's the end, the top 4.

Interviewer: Yeah. This is the top 10.

Interviewee: 15. Medium is 10, 15 is high. Then the same for layer 7 and for layer 10.

Interviewer: Yeah.

Interviewee: Yeah so our project was essentially about like trying to it's sort of about like explainability that's what this whole site yeah and like if you're interested in

this why not yeah and our supervisor had the idea that because there's a lot of articles to talk about like oh explainability is good and like you should answer these questions for the user, blah, blah, blah. But there's not a lot of them that talk about like how to do that in practice. So our supervisor had the idea, okay, but well, if we have a neural network and since we also have this project to build on, what if we try to show the hidden layers that the user usually doesn't have access to and try to show them what happens and maybe if possible let them do changes at those layers but of course we hit a lot of technical difficulties so we can scale it back a lot.

Interviewee: So this is the final product?

Interviewer: Yeah.

Interviewee: Nice.

Interviewer: But is it fine that this, that I can move it here? It's not written.

Interviewee: No, that's specifically because of my Firefox browser. For some reason, I can click anywhere and it would be like oh you're writing here but you're not and I also have a pop-up blocker so like allow and where would you oh this is the page for the original project where there's like a lot of information and there's also that's also where you can go back to the demo let's see why painting with a gun is interesting a computer could draw a scene in two ways it could compose the scene out of objects it knows, or it could memorize an image and apply one just like it.

Interviewee: Yeah, so they're talking about like, oh it doesn't learn to copy images or it doesn't learn to copy like the actual sort of objects, and apparently it copies like the actual object. That's what allows you to draw like with a tree in some cases.

Interviewee: I'm just fascinated by it, but I wouldn't think to know, to ask about the math behind it, but it's so cool when you do know it, it's cool.

Interviewee: I also think for our page, what we were also going for is that our thing is not really aimed at a specific group with a specific purpose in mind like we're not aiming at professional advertisements, creators, or something.

Interviewee: But you want to publish it, no?

Interviewer: No?

Interviewee: You don't want to?

Interviewer: Well, I'm not sure what you mean by publish.

Interviewee: I mean to put it on the website.

Interviewer: You cannot?

Interviewee: I don't think we're going to do that. Right now, we're just local hosting it. But it's more high-level, so we intentionally try to make it a bit high-level and not give you too many details.

Interviewee: Good. You shouldn't. That's good.