The many faces of Art
What techniques can we use to protect authentic artists from AI-generated art?

Sabina-Mălina Grădinariu[1]
Supervisor: Dr. Anna Lukina[1]

[1]**EEMCS, Delft University of Technology, The Netherlands**

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 23, 2024

Name of the student: Sabina-Mălina Grădinariu
Final project course: CSE3000 Research Project
Thesis committee: Dr. Anna Lukina, Dr. Petr Kellnhofer

Abstract

The advancement of generative models in simulating human creativity has greatly impacted the art world. In this context, artists are concerned about the devaluation of their work, especially considering the questions that appear surrounding authenticity and ownership rights. This study uses DE-FAKE, a state-of-the-art research-based detection tool, to address the crucial problem of separating AI-generated art from human-created works. The performance of DE-FAKE across a wide range of artistic styles will be assessed by carefully examining images chosen from the reputed WikiArt and AI-ArtBench datasets. This serves to highlight the advantages and disadvantages of the system. While DE-FAKE performs well in recognizing AI-generated art in the modern and abstract domains, the results show that it faces significant difficulties in recognizing more realistic styles, highlighting the need for additional development. This study lays a solid foundation for future research and practical solutions for safeguarding artists' intellectual property in the era of AI-generated art.

1 Introduction

Generative models have progressed in their pursuit of mimicking creativity [1], which has caused various effects in the domain of art [2]. While these efforts are impressive, they make artists fear their work is devalued because issues arise about authenticity and uniqueness [3]. AI-created art materializes alongside its human-made counterparts [4], creating confusion about what original work means [5], as ownership rights waver in this new uncharted territory primarily due to copyright issues surrounding generative models [6]. There remains a significant knowledge gap [7] to effectively distinguish these AI-generated images from those created by humans especially when these images could not merely be recreations but potentially novel works [8]. But what does all of these mean for artists and animators? Several posts and articles highlight that AI has become an extra burden for creators [9].

In this context, several AI detectors appeared. The state-of-the-art detectors are of two categories: commercial black-box detectors such as Hive AI Detector (Hive) [10], Optic AI or Not (Optic) [11], Illuminarty [12] and research-based detectors with transparent implementations: take DIRE [13] and DE-FAKE [14].

The recent study [15] provides essential information on the abilities and limitations of these detectors. The best results for unperturbed images are achieved by Hive, producing zero false positives and a 3.13% rate for false negatives, while expert artists had an accuracy of 83%, 20.78% false positives and 14.63% false negatives. The limitation and intriguing finding is that human experts outperform the classifier for detecting perturbed images. Commercial detectors work well on the available generative models but become less effective when exposed to perturbed images to trick detection systems. From our experimental research, and based on [15] the accuracy for Hive is 98.03%, while research-based detectors vary dramatically between 48% to 84%, depending on the art type. This study [15] highlights the need to incorporate analytical tools such as Hive with the subtle understanding of human specialists to achieve the most precise outcomes when identifying AI-generated art. In the context of art, this procedure minimizes the prospect of false positives and negatives while maintaining a comparable detection accuracy.

This paper will analyze techniques that improve the ability to distinguish between authentic human-made and AI-generated art to protect artists. The main ambition is to bridge the lack of knowledge by blending the latest experimental research with an extensive review of related literature. The work presents a practical point of view, which is beneficial for artists, curators, and legal experts as they traverse the changing art world in the era of artificial intelligence by offering guidelines that can easily be implemented.

Building on previous research, the primary research question driving this work is the following: *What techniques can we use to protect authentic artists from AI-generated art?* This leads to several relevant sub-questions, from focusing on the tools available for AI detection to adapting detection and analysis methods from computer vision computational photography and digital forensics. To specifically tailor these methods to the art domain, we will assess how well they have been used to identify AI-generated content. Furthermore, considering any potential ethical ramifications, the research assesses how the solutions proposed affect the process of art curation, sale, and authentication. This piece aims to help close the existing knowledge gap by providing a clear pipeline that creators and experts could put into practice, inspired by the comprehensive literature review and experimental research. The ultimate goal is to provide these stakeholders with the means to thrive in the changing art scene. The research findings and conclusions should be the foundation for further investigations, guaranteeing that the line between AI-generated art and human art will always be distinct and verifiable by proposing a pipeline that artists and curators can follow. This pipeline includes glazing techniques [16], advanced AI detection tools, ongoing monitoring, and collaboration with technologists. The outcomes demonstrated the need for additional detection algorithm improvement, by showing that DE-FAKE was better at identifying AI-generated art in modern and abstract styles than in realistic ones. In the following chapter, the background and related work will be discussed (Section 2). Next, the methodology and the proposed solution in the form of a pipeline, will be presented (Section 3). Afterwards, the experiment and its results will be shown (Section 4). Section 5 will cover responsible research (Section 5), and Section 6 will include the discussion (Section 6). Finally, the conclusion will be provided in Section 7.

2 Background & Related Work

This paper will start by explaining generative models such as GANs and diffusion models, then presenting what tools for detection are out there and introducing glazing as a protection mechanism against neural style transfer.

Table 1: Summary of AI Detection Tools

Tool	Type	Description
Hive AI Detector (Hive) [10]	Commercial	Accuracy of 98.03%, but its metrics/methods are not public [7].
Optic AI or Not (Optic) [11]	Commercial	90.67% accuracy in [7].
Illuminarty [12]	Commercial	Has high false positive and false negative rates [7].
DIRE [13]	Research-based	A tool for identifying general diffusion-generated images, using the distribution differences between diffusion model outputs and real images [7].
DE-FAKE [14]	Research-based	Uses a binary classifier, a 2-layer perceptron, to identify AI-generated images [7].

2.1 Generative models

Generative models are a class of groundbreaking machine learning models that were developed to generate new resembling data after learning the patterns of a dataset [17]. Their applications span a variety of fields, including biology, languages, astronomy, and art [18]. One of the first primitive models was ELIZA, a text chatbot [19]. Nowadays, Generative Adversarial Models (GANs) and diffusion models are two prominent models that have greatly impacted the art field.

Generative Adversarial Networks (GANs) is a deep learning architecture for supervised and unsupervised learning that retains complex high-dimensional distribution over the provided dataset [20]. This can be used to create new life-like images from text prompts. Multiple GANs were created for image generation: Midjourey, DALL-E 3.

Diffusion Models were inspired by non-equilibrium thermodynamics, the central idea around them is destroying the training data by gradually adding Gaussian noise and recovering the data by reversing the process. In the context of image generation, some of the leading models are Stable Diffusion XL and Adobe Firefly.

2.2 AI Detection Tools

Various AI detection tools emerged as an answer to the growing AI industry. The state-of-the-art detectors are of two categories: commercial black-box detectors and research-based detectors. Overall, significantly better results are obtained in detection and minimising the false positives and negatives by the commercial detectors (Table 1).

2.3 Artistic Style Protection

The goal of a recently proposed method called Glazing is to shield artists from having their distinctive styles copied by AI art generators such as MidJourney and Stable Diffusion without permission. The Glaze tool has to be used before sharing art online, it can subtly alter an art piece's appearance, by adding "style cloaks" to an artist's work [16]. These changes are nearly invisible to the human eye, but they have a substantial impact on how AI models interpret and pick up on the original artistic style. The main goal of Glazing is to deceive AI models into learning an erroneous representation of the artists' signature style during the training phase. As a result, any attempt to create artwork that mimics that style will fall short of capturing its genuine spirit and uniqueness. Glazing offers a promising solution for protecting creative intellectual property, and it was created in close cooperation with professional artists and validated through extensive user studies involving over 1000 artists.

2.4 Related Work

Scarce research was done specifically examining how AI-generated art differs from real artwork. Instead of customizing detectors or training datasets for the artistic domain, efforts have mainly focused on addressing the misclassification of regular images and deepfakes. Nonetheless, the research by [15] offers insightful details about the capability to accurately categorize images from an artistic dataset composed of seven distinct styles and corresponding generative AI outputs from five models. One key conclusion, drawn from this research, demonstrates how expert professionals are good at identifying AI-generated art. Experts achieved an astounding 83% accuracy rate, while amateurs found it difficult to distinguish between the two categories. Notably the best-performing detector demonstrated an even higher accuracy of 98.03%. On the other hand, methods to stop generative models from exploiting and mimicking the distinctive styles of artists, are slowly emerging. These developments are shifting the balance of power, entrusting artists to protect their intellectual property from being co-opted for training AI models without consent.

3 Methodology & Contribution

The main research question is: What techniques can we use to protect authentic artists from AI-generated art? The hypothesis is that measures like glazing and sophisticated AI detection tools, such as Hive [10] or DE-FAKE [14], can significantly improve the capacity to discriminate between real and artificial intelligence-generated art, protecting the integrity and authenticity of artists creations. To address this hypothesis I employed several methods: data collection, sampling, investigating detection tools, using evaluation metrics, and conducting a literature review. It was a methodological choice to include a wide variety of styles to assess the effectiveness of AI detection methods across a range of visual characteristics and techniques, while capturing the rich tapestry of expression. This decision was made to guarantee that the detection model is reliable and adaptable.

3.1 Dataset & DE-FAKE

The dataset was created by employing a stratified random sampling technique, adhering to the standards of rigorous scientific inquiry. From the WikiArt database [21] [22], eight artistic styles were used: Cubism, Impressionism, Action Painting, Realism, Baroque, Abstract, Fauvism, and Expressionism. This was decided based on the availability of at least 1000 images per style category. This cutoff point guaranteed a large enough sample size for accurate statistical analysis. To minimize any potential bias resulting from subjective selection, a random sample of one thousand images was selected for each chosen style. Using random sampling guarantees representativeness and makes it easier to extrapolate results to a larger population, representing the various artistic styles. The value was selected to ensure a statistically significant sample size for trustworthy analysis and to match the AI-generated dataset images, which have 1000 images for each style for both Latent Diffusion and Stable Diffusion.

The set of generated images was obtained from [7]. This dataset ensures a diverse representation of AI-generated art because it covers a varied spectrum of artistic styles and two prominent generative models (LD and SD). LD Baroque, SD Baroque, LD Art Nouveau, LD Expressionism, SD Expressionism, LD Realism, and SD Realism were experimented with. The investigation concentrated on the DE-FAKE detector, an open-source, cutting-edge system for categorizing AI-generated art, as the best open-source alternative to the proprietary Hive. Due to Hive being unavailable for scientific use, we had to orientate ourselves towards a public detector. The selection of DE-FAKE was driven by its extensive use in the research community and its stated superior performance, when compared to other research-based detectors. Extensive experimental evaluations were carried out on the curated datasets to evaluate DE-FAKE's accuracy, success rate, and false positive and negative rates. This empirical approach permits reproducibility and objective assessment by being aligned with scientific best practices.

3.2 Literature review

A thorough literature review was conducted to place the research in the larger framework of artistic style protection mechanisms and AI-generated art detection [1] [3] [6] [7]. The foundation of this research was the work of [15], and its references were carefully examined to capture a variety of perspectives on the subject. To add to the study's theoretical foundation and provide multiple viewpoints resources, impact analysis surveys and technical documentation were consulted [2] [5] . This comprehensive strategy follows accepted scientific norms of objectivity, reproducibility, and theoretical foundation, guaranteeing the validity and reliability of the results. It combines empirical dataset analysis, experimental evaluations, and a thorough literature review.

Before creating this pipeline, to comprehend the state of AI-generated art detection and artist safety protocols, a comprehensive assessment of the literature was conducted. The review emphasized the significance of employing detection tools and protective techniques to preserve the integrity of human-created art. It also underlined the importance of continuing research and development to enhance detection algorithms and take advantage of generative model development capabilities. This analysis adds value by offering a practical and transparent process that curators and artists can use to safeguard their creations. The pipeline aims to reduce the risk of AI-generated mimicry and ensure that the intrinsic value of human-created art is preserved by combining glazing techniques with cutting-edge AI detection tools. This strategy not only provides a workable answer but also lays the groundwork for further studies that will increase the effectiveness of these protective measures.

3.3 Pipeline

Several crucial steps are included in the proposed pipeline to guarantee the authenticity and preservation of artwork. Before posting digital art online creators are advised to apply a process called glazing to their work. Next, it is critical to use cutting-edge AI detection tools to confirm the authenticity of the artwork. Although the Hive AI Detector is a black box detection tool, experimental research has shown it to be the most successful at identifying AI-generated images and human artwork. Curators and artists can use it for free, but not much is known about how it operates. The DE-FAKE algorithm was extensively evaluated, its findings will be disclosed later in the paper. The pipeline proposed is as follows:

1. **Glazing of Artworks:** Artists should apply glazing techniques to their digital artworks before sharing them online. This involves adding intentional, subtle variations to the artwork that are designed to disrupt AI models' ability to learn and replicate the artist's style accurately.

2. **Use of AI Detection Tools:** Before finalizing and showcasing artworks, curators should use advanced AI detection tools, such as the Hive AI Detector or the DE-FAKE algorithm, to verify the authenticity of the pieces.

3. **Ongoing Monitoring and Verification:** Using the recent AI detection technologies curators and artists should routinely check if their collections are authentic. Continuous verification keeps the artworks authentic, while stopping AI-generated pieces from being added to collections that were made by humans.

4. **Collaboration with Technologists:** Artists and curators are encouraged to collaborate with technologists to stay informed about the latest advancements in AI detection and protection techniques.

By implementing this pipeline, curators and artists can better protect their works from AI-generated mimicry and ensure that the intrinsic value of human-created art is preserved. This proactive approach safeguards individual artists' intellectual property rights, while contributing to the broader effort of maintaining authenticity in the art world amid the rise of AI technologies.

4 Experimental Setup and Results

In this study, we evaluated the performance of the DE-FAKE AI detection algorithm using a curated dataset, including human-created art from the WikiArt database [21] and AI-generated art images from AI-ArtBrench [7]. The focus was assessing how the algorithm classifies multiple artistic styles (Table 2).

4.1 Experimental Setup

Two primary datasets were used for the experimental evaluation: the WikiArt and the AI-ArtBench dataset, both open source. The WikiArt, a dataset that has over 80,000 images spanning 27 different artistic styles, is considered the most complete art database. Its inclusion ensured a comprehensive representation of various artistic movements (Modern, Cubism, Renaissance, Impressionism, Fauvism, Mannerism, Art Nouveau/Modern, Ukiyo-e, Baroque, and Color Field Painting) and techniques, enabling a thorough assessment of the detection algorithm's performance across a broad spectrum of styles. The AI-ArtBench dataset from Kaggle complemented the human-created art images by providing a diverse collection of AI-generated artworks. Specifically, this dataset contains images generated by two influential generative models: Latent Diffusion (LD) and Stable Diffusion (SD), across ten styles. Each style has 1,000 images, and an equal number was sampled from the corresponding styles in the WikiArt dataset to ensure a balanced representation. While the research behind this dataset's curation [7] is yet to be published, its diverse composition and inclusion of human-created and AI-generated art made it a valuable resource for this study.

A stratified random sampling approach was employed to select the exact number of images for the human art from WikiArt to guarantee balanced representation and mitigate potential biases. A random sample of 1,000 images was drawn for each of the selected artistic styles, resulting in an equal distribution of 1,000 images per category. This sampling ensured statistical significance and aided the generalization of the findings to the broader population of the respective artistic styles.

The detection tool chosen for this study was DE-FAKE, a state-of-the-art research-based detector specifically designed to classify AI-generated images. By employing DE-FAKE, this study aimed to leverage cutting-edge detection capabilities and contribute to the ongoing efforts to distinguish AI-generated art from human-created works. To evaluate the performance of the DE-FAKE algorithm three key metrics were employed: accuracy, false positives, and false negatives. Accuracy measured the percentage of correctly identified images, providing an overall assessment of the detector's performance. False positives referred to instances when human-created art was incorrectly classified as AI-generated, while false negatives represented cases when AI-generated art was mistakenly identified as human-created. These metrics quantified the algorithm's effectiveness and highlighted potential areas for improvement, contributing to a comprehensive understanding of its strengths and limitations.

4.2 Results

The DE-FAKE algorithm's performance varied significantly across different art styles. I reviewed the accuracy rate, false positives and false negatives and AI detection success rate (ADSR) as the main metrics for the performance. The images can be classified as 0 (generated) or 1 (real), and one knows the absolute truth about the images. Figure 1 compares the detection results for Latent Diffusion and Stable Diffusion generated art, while Table 2 illustrates the detection results for the various types of art.

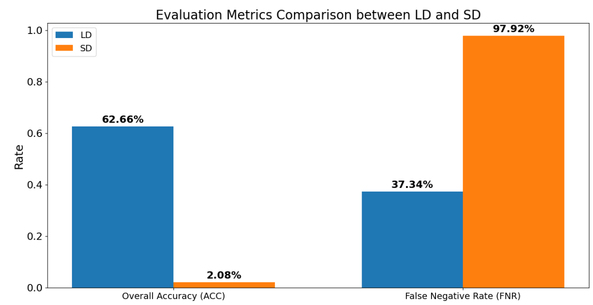

Figure 1: Detection results for Latent Diffusion and Stable Diffusion generated art

For **realistic styles** such as Baroque and Impressionism, the algorithm had a low accuracy in detecting the real art pieces, approximately 49% resulting in high rates of false positives. The Baroque results disclose an accuracy of 61.5% and an AI detection rate of 78%. For overall detection in Baroque, of both generated and real images, there is an accuracy of 69.75%, the algorithm is better at detecting AI-generated images than at classifying real art. Similarly, for Impressionism, the accuracy is 58.5% in detecting fake images. An interesting remark is that Expressionism has a mere accuracy of 43% while Abstract Expressionism has an accuracy of 72%. Once again, the more abstract style is more effortlessly put in the correct category.

For **modern and abstract styles** such as Modern Art, Cubism, Color Field Painting, the algorithm, had a high accuracy ranging from 80% to 94%. This is an interesting result, contrasting with the accuracy of more realistic, life-like styles such as Baroque and Impressionism. Intuitively one can suggest that the algorithm performed better with classes exhibiting distinctive patterns characteristic of genuine art, which are challenging to replicate.

Generated Art Latent Diffusion vs Stable Diffusion detection rate is very different depending on the generation type. LD has an accuracy rate of 78% for Baroque, 84 % for Art Nouveau, 43% for Expressionism, 47.5% for Realism, and 58.5% for Impressionism. This results in an overall accuracy of 62.2%. The algorithm struggles to detect AI images generated through SD. For all artistic styles, the accuracy of AI detection, for SD generated images, is under 20%, as seen in Figure 1.

Table 2: Summary of DE-FAKE Algorithm Results for LD Images

Art Style	Accuracy for Generated Art (%)	Accuracy for Real Art (%)
Abstract Expressionism	-	72
Baroque	78	61.5
Cubism	-	73
Expressionism	43	73
Impressionism	58.5	42
Realism	47.5	49
Fauvism	-	69
Art Nouveau	84	41
Ukiyo-e	-	50
Color Field Painting	-	97

4.3 Summary of Results

We will go into detail only for the results obtained for the Latent Diffusion images as the Stable Diffusion detection showed flawed results overall. The experimental results indicate that DE-FAKE is more effective at detecting AI-generated art in abstract and modern styles than realistic styles. A detailed analysis showed differences in the detection rates between the types of generation. The algorithm performed satisfactorily with LD images, but the accuracy significantly dropped with SD images, which were often misclassified. This suggests that the detection algorithm can benefit from further refinement to improve accuracy across all studied art categories.

5 Responsible Research

We adhered to the highest ethical standards and principles of responsible research conduct. The following were considered:

Ethical Implications: The application of AI to the creation of art raises several ethical issues. A significant risk to artists and creators is the potential for AI systems to imitate distinctive artistic styles without permission or payment. Developing strategies to protect artists' intellectual property rights and the integrity of their creative expressions served as the driving force behind my research. We tried to mitigate the detrimental effects of AI on the art world and preserve the inherent value of true human creativity by looking into a thorough analysis of methods to distinguish AI-generated art from human-created pieces accurately.

Data Usage and Consent: Everything used in this study, including the AI-ArtBench dataset from Kaggle and the pictures from the WikiArt database, are from publicly available sources. The use of these datasets was compliant with the terms of service of each provider when applicable. We also ensured our analysis did not violate the original content creators' rights.

Reproducibility and Transparency: We kept documents describing the procedures and algorithms used in this investigation because we recognize how critical reproducibility and transparency are to scientific research. The public nature of the datasets, sourced from reliable sites like Kaggle and WikiArt, makes it possible for other researchers to build on and duplicate our findings. We provide an in-depth description of the code and algorithms, including a reference to how the DE-FAKE detector is implemented.

We strived to demonstrate the highest standards of responsible research practices while making a significant contribution to the understanding of AI-generated art by adhering to these ethical guidelines and using exacting methodological approaches.

6 Discussion

In the future, more work should be done on improving detection algorithms, until they can handle realistic art forms more effectively. The results of this study point to a critical area for development, showing that DE-FAKE is more accurate at identifying AI-generated art in modern and abstract domains than realistic styles. Through further optimization of these algorithms, we can improve their efficacy in all artistic domains, guaranteeing the conservation of the inherent worth and genuineness of art produced by humans. Furthermore, continuous cooperation with artists is essential to ensure that the solutions developed are not only theoretically sound but also practically applicable and effectively address the concerns of the art community.

Views on the practical difficulties artists confront and the particular subtleties of their work that need to be protected can be gained from artists' insights and comments. A mutually beneficial relationship between technological advancements and artistic expression could be fostered by this collaborative approach, which guarantees that the solutions meet the demands and expectations of the art world seamlessly. The results of this study confirm other research that has shown how difficult it can be to distinguish AI-generated art from human-created pieces. One proactive way to address these issues is to use methods such as glazing and sophisticated tools like DE-FAKE. However, to keep up with the rapid advancements in generative AI technologies, there is still a pressing need for more complex and flexible detection techniques. Developing relationships with companies such as Hive, which have produced efficient AI detection tools, is one realistic step forward. Researchers and companies that create commercial tools can work together to improve the overall strength of AI detection techniques, aiding in the protection of artists' intellectual property.

In summary, although the existing detection tools show respectable abilities, more research and cooperation are needed to improve their precision and generalizability across a range of artistic mediums. We can build a more solid and efficient framework for safeguarding the authenticity of works of human creation by incorporating artist insights, utilizing cutting-edge technologies and forming strategic alliances with tool developers who sell their products. This all-encompassing strategy will protect the intrinsic worth of art while also promoting within the art world a smooth evolution in step with technological breakthroughs, fostering a mutually beneficial relationship between innovation and creativity.

7 Conclusions & Future Work

Our study has considered the DE-FAKE AI-detection algorithm and investigated the results it delivered across a wide range of artist styles. The results disclosed its proficiency in identifying AI-generated images within Modern and Abstract Art, scoring an accuracy between 70% and 84%. However, its performance dropped when applied to more realistic styles, such as Baroque and Renaissance, with an accuracy rate of around a mere 40%.

The pipeline offers curators and artists a viable way to protect human-made art against imitations produced by artificial intelligence. The pipeline provides a tool for safeguarding the authenticity and inherent worth of human art in the age of artificial intelligence. This is done by combining glazing methods with cutting-edge AI detection technologies and collaborating with technology experts. Moreover, given the ethical implications and potential repercussions of AI on the art community, future work must continue to actively collaborate with artists, ensuring that the solutions developed are both theoretically sound and also practically applicable to addressing their concerns.

Future research efforts should prioritize enhancing detection algorithms to reflect the nuances and subtleties inherent in realistic art styles. Furthermore, exploring the integration of complementary detection methodologies and tools, drawing from digital forensics and computational photography, could pave the way for a more robust and comprehensive solution. This study is a pivotal stepping stone for ongoing efforts to safeguard authenticity and the intrinsic value of human-created art amidst the relentless advancements in artificial intelligence technologies.

References

[1] H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P.-A. Heng, and S. Z. Li, "A survey on generative diffusion models," *IEEE Transactions on Knowledge and Data Engineering*, 2024.

[2] F. B. Council. (2024) The impact of artificial intelligence on the art world. Accessed: Feb. 20, 2024. [Online]. Available: https://www.forbes.com/sites/forbesbusinesscouncil/2024/02/02/the-impact-of-artificial-intelligence-on-the-art-world/?sh=66a590e5532d

[3] İ. E. Gülaçtı and M. E. Kahraman, "The impact of artificial intelligence on photography and painting in the post-truth era and the issues of creativity and authorship," *Medeniyet Sanat Dergisi*, vol. 7, no. 2, pp. 243–270, 2021.

[4] K. Roose. (2022) The ai artists of the future are here. Accessed : May 27, 2024. [Online]. Available: https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html

[5] M. Boden, "Authenticity and computer art1 2," *Digital Creativity*, vol. 18, pp. 10 – 3, 2007.

[6] Z. Epstein, S. Levine, D. G. Rand, and I. Rahwan, "Who gets credit for ai-generated art?" *iScience*, vol. 23, 2020.

[7] R. Silva, "Ai-artbench," https://www.kaggle.com/datasets/ravidussilva/real-ai-art, 2023, accessed : 2024-05-17.

[8] A. Elgammal, B. Liu, M. Elhoseiny, and M. Mazzone, "Can: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms," 2017.

[9] L. Clarke. (2022) When ai can make art – what does it mean for creativity? [Online]. Available: https://tinyurl.com/creativity-and-art

[10] Hive, "Hive ai detector," 2023, https://thehive.ai/demos.

[11] Optic, "Ai or not," 2023, https://www.aiornot.com/.

[12] Illuminarty, "Is an ai behind your image?" 2023, https://illuminarty.ai/.

[13] Z. Wang, J. Bao, W. Zhou, W. Wang, H. Hu, H. Chen, and H. Li, "Dire for diffusion-generated image detection," 2023.

[14] Z. Sha, Z. Li, N. Yu, and Y. Zhang, "De-fake: Detection and attribution of fake images generated by text-to-image generation models," 2023.

[15] A. Y. J. Ha, J. Passananti, R. Bhaskar, S. Shan, R. Southen, H. Zheng, and B. Y. Zhao, "Organic or diffused: Can we distinguish human art from ai-generated images?" 2024.

[16] S. Shan, J. Cryan, E. Wenger, H. Zheng, R. Hanocka, and B. Y. Zhao, "Glaze: Protecting artists from style mimicry by text-to-image models," 2023.

[17] G. Harshvardhan, M. K. Gourisaria, M. Pandey, and S. S. Rautaray, "A comprehensive survey and analysis of generative models in machine learning," *Computer Science Review*, vol. 38, p. 100285, 2020.

[18] K. Sankaran and S. P. Holmes, "Generative models: An interdisciplinary perspective," *Annual Review of Statistics and Its Application*, 2022.

[19] H.-Y. Shum, X.-d. He, and D. Li, "From eliza to xiaoice: challenges and opportunities with social chatbots," *Frontiers of Information Technology & Electronic Engineering*, vol. 19, pp. 10–26, 2018.

[20] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Processing Magazine*, vol. 35, pp. 53–65, 2017.

[21] "Wikiart," https://www.wikiart.org/, 2010, accessed : 2024-05-10.

[22] Steubk, "Wikiart," https://www.kaggle.com/datasets/steubk/wikiart, 2022, accessed : 2024-05-10.

Acknowledgements

I want to give credit to the people who helped in the process that led to this moment, finishing my bachelor thesis. I want to thank my supervisor teacher, Anna Lukina, who oversaw my progress in the past few months. I am also grateful for my great student counsellor, Jitske van der Laan, who helped me throughout my bachelor's.

Some powerful women in my high school education, my mathematics teachers who inspired me to push my limits and always strive for excellence, guided my journey at TU Delft.

On another note, I would like to thank my family, who encouraged me to continue even when the situation was uncertain. Thanks to you: mom, dad, Sofia, my dear grandparents, and you, my amazing friends (my support system), we got here. Look at us now!