UNIVERSITÄT LEIPZIG

Institut für Informatik
Fakultät für Mathematik und Informatik
Digital Humanities

Investigating artist and style mentions in Diffusion Model prompts

Projektarbeit Digital Humanities

vorgelegt von:
Georg Schneeberger

Matrikelnummer:
3707914

Betreuer:
Dr. Andreas Niekler

Abstract

In this work, I investigate the prompting behaviour of 'Stable Diffusion' users in regard to the inclusion of artist names in prompts. I try to disprove the hypothesis that top-mentioned artists are mentioned for their style. I find that there is a correlation between the popularity of an artist and the style references. The prompts of more popular artists resemble the prompts in which they are not mentioned more closely than those of less popular artists.

The conclusion is that more popular artists are less used for their characteristic style but rather are included in prompts for other reasons.

Georg Schneeberger
3707914

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

1. Introduction

For a long time, the creation of computer art was mostly limited to programmers, engineers and scientists, for they had access to computing resources and programming knowledge. Since then, the creation of computer art has become much more accessible with drawing and image editing tools. The recent progress of research in the field of artificial intelligence has led to the development of many image generation tools with different underlying algorithms and purposes. One of these are diffusion models, which are a class of generative models that are able to generate images of high quality conditioned on user input. Services such as 'Stable Diffusion' [1] and 'Midjourney' [2] allow users to generate images and art based on image descriptions, also called the prompt, and other parameters, such as for example a reference image.

This new technology has led to a series of controversies, in August 2022 an image generated by Jason Allen won the Colorado State Fair's digital art competition. It was later revealed to be computer generated, which sparked a discussion about the merit of its creator and artificial intelligence in art [3]. Furthermore, these models were trained on a dataset of images collected from the internet. It is currently unclear if the models' creators have committed cases of copyright infringement. In response, lawsuits have been filed [4], and artists started movements against AI art by posting protest pictures [5].

Nevertheless, these services have become very popular, with many users generating images and sharing them on social media and dedicated discord channels. Many services have dedicated discord servers or websites where the issued prompts and generated images are published. Diffusion model users formed communities on various discords and messageboards like Twitter and Reddit.

1.1. Prompt Design

The process of building prompts to obtain the desired output is called prompt engineering and has been studied for diffusion models [6] and other generative models like ChatGPT [7]. This process requires trial and error, since the models are very complex, and it currently cannot be explained how different parts of a prompt influence the output. Users and researchers discuss the images and the prompts used to generate them; from this exchange and the exploration of these services, the users and researchers [8] noticed patterns in the prompts. For example, the users found that adding specific keywords, such as 'unreal engine' to a prompt, results in a hyperrealistic, 3D rendered image. Furthermore, the usage of artist names can be used to influence the style and content of images [6].

Many users in the AI art community started mentioning the Polish artist Greg Rutkowski in their prompts. This even led to some of these generated images appearing in front of his work in Google searches. The artist's name was used in some early official tutorials [9] and has become a symbol of AI art. This overall popularity of the artist in the scene might be why his name is mentioned so often in prompts, compared to mentioning him for his unique style. This could also be described as following a trend or cargo cult programming [10].

In this project, I will investigate how diffusion model users reference artists in their prompts. I will try to find out if the top-mentioned artists are mentioned to replicate artworks in their characteristic style or if they are instead mentioned for other reasons. I will investigate this by trying to disprove the following hypothesis: Top mentioned artists are mentioned specifically to replicate artworks in their characteristic style.

Answering this and other similar questions might become necessary in future discussions about AI art and copyright infringement concerning artists' work. Furthermore, it might help to understand the diffusion model users and their motivations, which could help build more accurate generative models.

2. Data

2.1. Prompt Dataset

The prompt analysis requires a dataset of user-generated diffusion model prompts. Such a dataset can be collected from discord servers, the services' websites or dedicated prompt and image hosting websites such as 'Lexica Art' [11]. Ideally, the dataset should contain prompts from different diffusion services such as 'Stable Diffusion' and 'Midjourney' to allow for a comparison between them. Differences in prompt structure for different services have already been observed by [12]. A comparison of prompts for different versions of the services could also be possible, given such data.

I will use the 'diffusiondb' dataset that is publicly available on huggingface [12]. It contains a set of 14 million prompt-image pairs with about 1.8 million unique prompts. The prompt-image pairs were collected from the 'Stable Diffusion' discord server and included additional metadata such as the prompt author's hashed username, the date it was created, the image generation seed and more.

The dataset's authors [12] acknowledge there might be a bias in the prompt collection, as the prompts were collected from the official 'Stable Diffusion' Discord server. This server might have a disproportionate amount of AI art enthusiasts and might not be representative of novice users. Given that the data consists solely of 'Stable Diffusion' prompts, the findings might not apply to other diffusion services.

2.1.1. Dataset Details

Each of the 14 Million entries consists of an image_name, prompt and other metadata 5.1. The image_name can be used to find the image generated by the prompt and other parameters. The prompt string is the main focus of this thesis and will be analysed in the following sections.

The prompts were issued from the 6th of August 2022 to the 20th of August 2022 by 10351 unique discord users. On average, each user issued 175 prompts. However, the standard deviation is very high at 346. 442 users only issued one prompt, while the most active user submitted 14059 unique prompts. The dataset creators [12] analysed the prompts with regard to prompt length, amount of specifier clauses and the most common tokens. Instead of tokenizing the prompts using punctuation marks, the researchers used the tokenizer provided in 'Stable Diffusion v1.4' [13]. Specifier clauses are parts of prompts separated by one of the three delimiters {,;|} 5.1.

2.2. Artist Dataset

As with other modern AI systems, it is not apparent how the input prompt to a diffusion model influences the output. Style studies investigate this relationship between prompts and the generated images. Researchers typically examine the impact of keywords on the resulting images by prompting the model with different keywords and analysing the changes in the generated images.

I will use the dataset [14] as a reference dataset of keywords, that significantly impact the output images. The dataset contains artist names that have been shown to influence the output images. Furthermore, the creators of the dataset also characterised this influence by assigning tags to the artists. This list was created for the 'Stable Diffusion' service. The publishers of [14] created style references for other diffusion models such as 'Midjourney' and 'Disco Diffusion' as well. There are other datasets available online [15], however this dataset was by far the most comprehensive and detailed.

Several artists are mentioned in the dataset multiple times, by their real names and pseudonyms. For example, the artist 'Jean Giraud' is also listed as 'Moebius' and 'Mœbius'. I removed these duplicates from the dataset and added a pseudonyms column 5.2.

2.3. Styles Dataset

To reference existing art styles, I will use the 'List of art movements' from Wikipedia [16]. This list contains a total of 192 art movements, including movements around from 1000 AD to the present day. The list is not limited to European or Western art, for example, it includes the art movements 'Ukiyo-e' from Japan and 'Samikshavad' from India. The list also includes several art movements that emerged through computers and modern technology, such as 'Digital art' and 'Art Photography'.

3. Methods

I want to disprove that the top-mentioned artists are mentioned for their style. This would mean that they are less used for their unique style but are added to prompts for other reasons. When artists are not mentioned for their own unique style, the prompts in which they are mentioned should closely resemble other prompts in which they are not mentioned. Artists included for reasons other than their style should essentially be independent of the rest of a prompt.

There are many different ways of measuring the similarity between prompts. Since we focus on the analysis of styles in prompts, I will use direct references to styles to quantify the prompts.

3.1. Preprocessing

The primary dataset contains 1.8 million prompts and metadata as mentioned in 2. The first step in the analysis is to detect the mentioned artists and styles in the prompts, the results are saved to separate files for the following steps.

3.1.1. Detecting artists

I use the large set of artist names and pseudonyms from the dataset 2.2. Each prompt is analysed, the mentioned artists are extracted via an exact match of their full name or pseudonym. The amount of mentions of an artist across the entire dataset is counted and indicates the artist's popularity.

The first iteration of this preprocessing step did not include the analysis of pseudonyms. After analysing the resulting data, I found that many artists were mentioned by their real names and pseudonyms. I thus decided to include pseudonyms in the analysis.

3.1.2. Detecting styles

Similarly to the extraction of artists, the styles are extracted from the prompts via exact match. As a reference dataset of styles, I used the list 2.3 of 192 styles. 185 of the 192 styles were found in the prompts.

3.2. Characterizing the artists

To characterize the usage of an artist in the prompts, all prompts where an artist is mentioned are extracted. This group of prompts will be called the artist corpus. The set of prompts that do not contain that artist are called the negative corpus. In all cases, the negative corpora are smaller than the artist corpus since no artist occurs in more than 50% of the prompts.

The artist corpus and the respective negative corpus are computed for every artist. Artists without mentions in the dataset and artists that do not have any style mentions in their corpus are excluded from the analysis. The artist corpus and the corresponding negative corpus are used to compare the different artists.

The mentioned styles in the artist corpus and negative corpus are grouped and counted. The counts are normalized to proportions using the total amount of style mentions to enable comparisons between the two corpora.

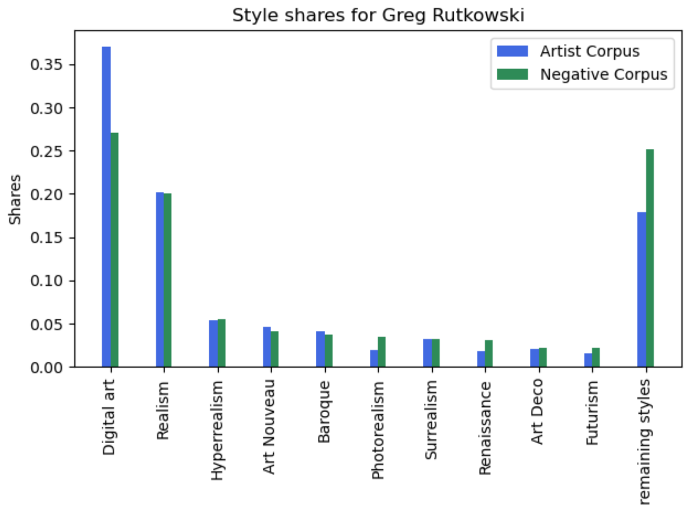

Abbildung 3.1.: Example Style distribution for the 10 most popular styles and remaining styles

3.3. Comparing Style distributions

Artists with similar corpus and negative corpus have a smaller influence on the prompts they are used in. These artists are independent of the styles mentioned in the prompts. I use the Bray-Curtis dissimilarity to evaluate the similarity between the two corpora. With these dissimilarity values, we can examine if there is a relation between the popularity of an artist and the artist's independence of the styles mentioned in the prompts.

3.3.1. Bray-Curtis dissimilarity

For comparing the style distributions, I will use the Bray-Curtis dissimilarity. It comes from Biology and quantifies the compositional dissimilarity between two sites, based on counts. In this case, we are comparing the style distributions of the artist corpus i with the negative corpus j. The dissimilarity value is 0 for identical distributions and 1 for distributions without common elements. It is defined as:

$$BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j}$$

Since we express the style mentions as proportions, the sizes S_i and S_j are always one. The forluma is simplified to:

$$BC_{ij} = 1 - C_{ij}$$

C_{ij} is the sum of the lesser values for styles in common between the compared prompts:

$$C_{ij} = \sum_{s \in Styles} min(i_s, j_s)$$

Figure 3.2 shows the calculation of a dissimilarity value. The red bars represent the shares $min(i_s, j_s)$, that the style distributions of the artist corpus and negative corpus have in common. The summation of all the red bars equates to C_{ij}.

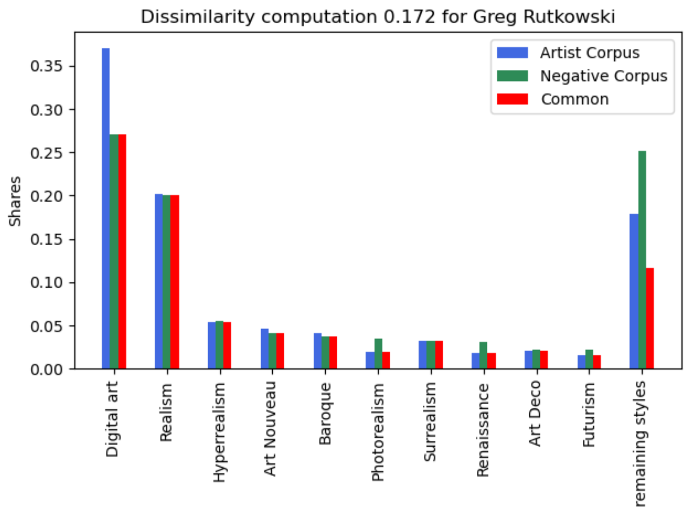

Abbildung 3.2.: Example Dissimilarity calculation

3.3.2. Evaluating the observed dissimilarity

The observed dissimilarity values will be analysed with Spearman's rank correlation coefficient. This correlation coefficient is chosen since it is appropriate for comparing two ranked variables and does not require the variables to satisfy conditions such as being normally distributed. In this case, the coefficient indicates whether the artist's popularity rank correlates with the style distributions' dissimilarity. The coefficient produces an r_s value between -1 and 1, where 1 indicates a perfect positive correlation, 0 indicates no correlation and -1 indicates a perfect negative correlation.

To check the significance of the correlation, we will use the p value of the Spearman's rank correlation coefficient. This p value is for a hypothesis test, where the null hypothesis states that the two variables are not correlated.

4. Results

4.1. Data Exploration

About 45% of the 1.8 Million prompts mention at least one artist and only 13% of all prompts mention a style. 8% of all prompts mention both a style and an artist, which means that 61% of prompts with style mentions also mention at least one artist. We thus have to note that we are in essence only using 8% of all prompts in the dataset for the artist characterizations and subsequent dissimilarity analysis.

A total of 2104 artists qualify for the artist characterization and dissimilarity analysis since they were mentioned in prompts that included style mentions.

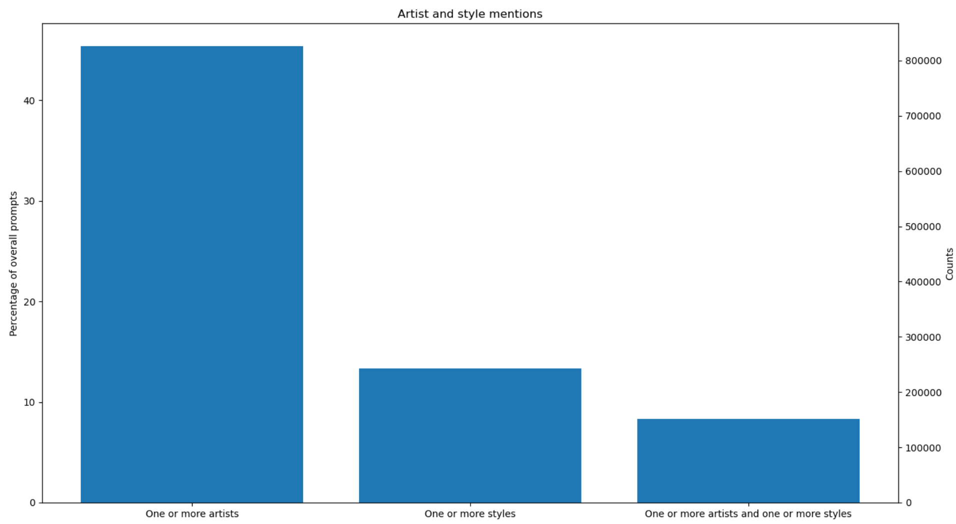

Abbildung 4.1.: Prompts with artist and style mentions in the dataset

4.1.1. Artist Mentions

Out of the 3477 artists from the artist dataset, 2825 were mentioned in the prompt dataset. The amount of mentions of the individual artists vary drastically. It appears to follow a Zipf distribution, similar to words in a corpus. The top 10 artists with the most mentions are shown in 4.3. They make up a large share of the total artist mentions. The most mentioned artist 'Greg Rutkowski' appears in about 10% of all prompts.

4.1.2. Style Mentions

185 of all 192 styles in the list are found in the dataset. Similarly to the artist mentions, they appear to follow a Zipf distribution.

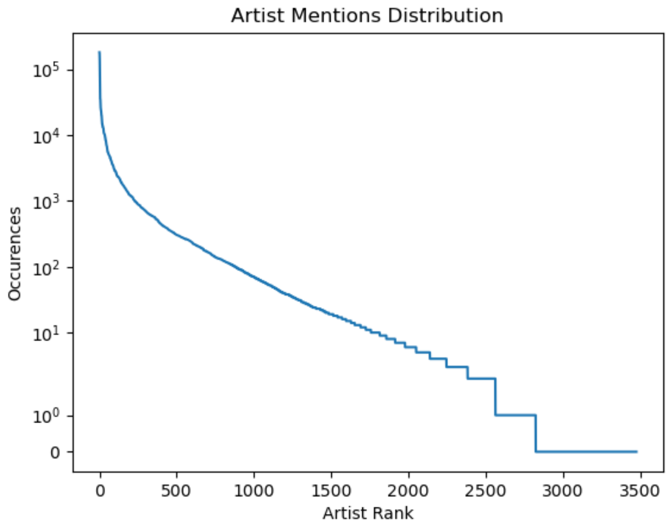

Abbildung 4.2.: Artist mentions appear to follow a Zipf distribution

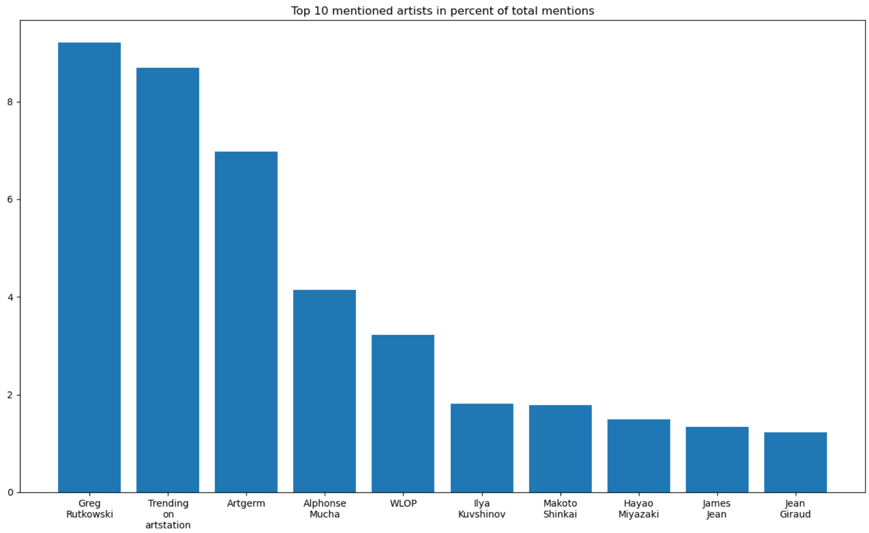

Abbildung 4.3.: Top 10 artists with the most mentions in the dataset

4.2. Artist Characterization

The dissimilarity computations for the artists with the highest and lowest dissimilarity are shown below. The artist with the highest dissimilarity is 'João Artur da Silva'. This artist is only mentioned in 2 prompts and only has a single style mention. 'João Artur da Silva' has a dissimilarity value of 1, meaning there is no overlap between the style proportions in the corpus and negative corpus. 'João Artur da Silva' is ranked at position 2417 in the dataset by mentions. Any artist with such low amounts of artist and style mentions will necessarily have a high dissimilarity value since such low amounts of mentions are not able to express the more complex distribution of the negative corpus.

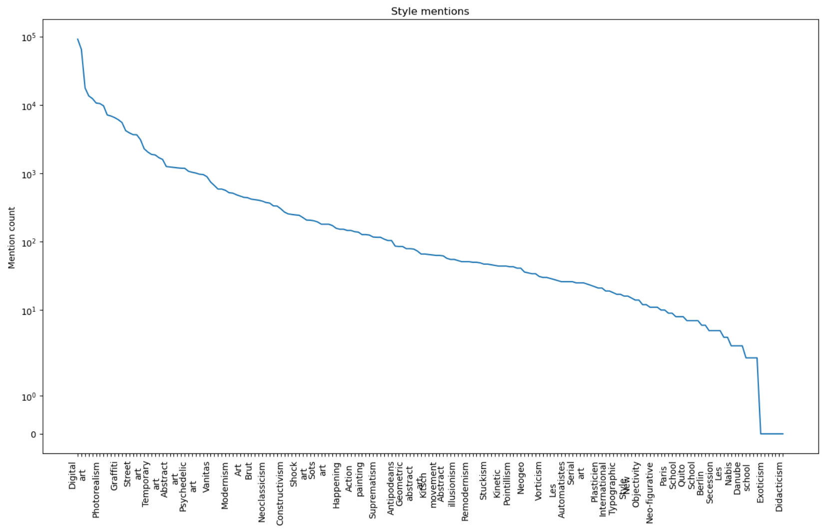

Abbildung 4.4.: Style mention count distribution

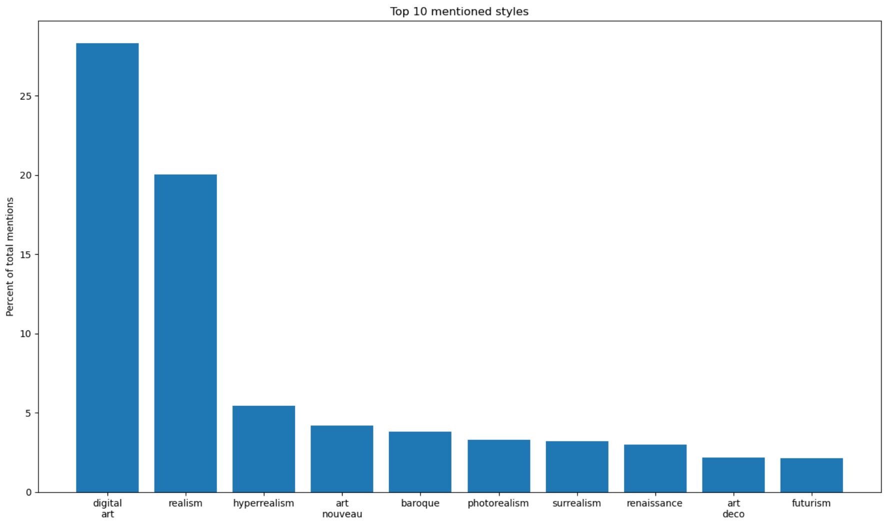

Abbildung 4.5.: Top 10 styles with the most mentions in the dataset

The artist with the lowest dissimilarity value is 'Stanley Lau', also called 'Artgerm'. 'Stanley Lau' is the third most popular artist by mentions, he is an illustrator and produces art for giants in the entertainment and gaming industry.

4.3. Dissimilarity correlation results

When plotting the dissimilarity values of artists sorted by their popularity, we can notice there is a relation between popularity and dissimilarity. The dissimilarity tends to increase with lower artist popularity, this relationship does not look very strong, however. The Spearman's rank correlation coefficient produces an r_s value of 0.459, indicating a positive correlation between the popularity of an artist and the dissimilarity between the artists' corpus and negative corpus. This means less popular artists tend to have a higher dissimilarity value. The p value of the hypothesis test is

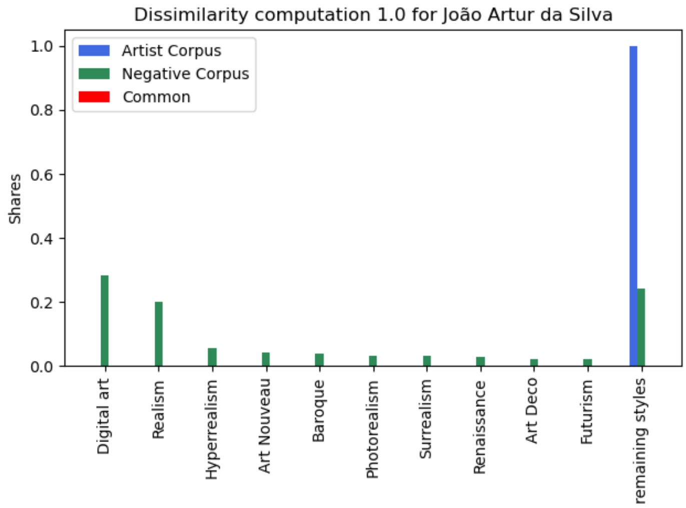

Abbildung 4.6.: Artist characterization for the artist with the highest dissimilarity

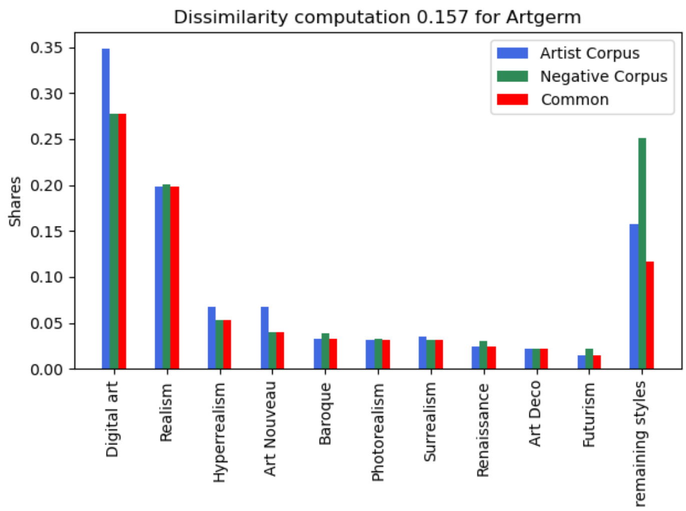

Abbildung 4.7.: Artist characterization for the artist with the lowest dissimilarity

7.867e-181, this means the null hypothesis can be rejected with a significance level of 0.05. We can thus conclude that a correlation most likely exists between the two variables.

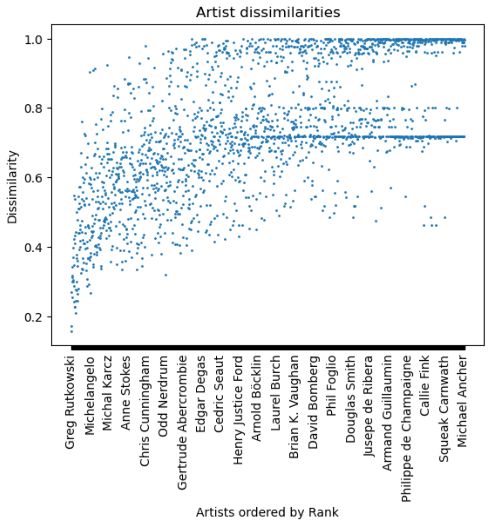

Abbildung 4.8.: Dissimilarity values of artists sorted by their popularity

5. Discussion

The results of the correlation of artist rank and dissimilarity show a relation between an artist's popularity and style references. More popular artists' prompts resemble the prompts in which they are not mentioned more closely than the prompts of less popular artists. More popular artists are essentially more independent of style mentions than less popular artists. With the analysis the original hypothesis can be rejected. Thus the most popular artist are not mentioned to specifically replicate artworks in their characteristic style. They are rather included for other reasons, to a large part.

The analysis delivers clear results, however during this process some limitations and areas of improvement were identified. I believe these limitations would be worth investigating in the future.

5.1. Areas of improvement

Further analysis and filtering of artists and prompts could prove useful and might even effect the results of the dissimilarity analysis. I found three reasons for removing certain artists and prompts before carrying out the dissimilarity analysis:

- Some artists' mentions were caused by a few or even just a single user. The popularity and style distributions for these artists might not be representative of the broader community.
- There are prompts that mention a lot of different styles. The maximum amount of style mentions in a single prompt is 30. These outliers could skew the results of the analysis.
- Many artists have a high dissimilarity value by virtue of having a low amount of style mentions. These artists could also be disqualified for the analysis, similarly to artists without style mentions.

Other methods for characterizing the artists could also be explored, such as the use of topic modelling.

Literatur

[1] Accessed: 2023-03-10. URL: https://huggingface.co/spaces/stabilityai/stable-diffusion.

[2] Accessed: 2023-03-10. URL: https://www.midjourney.com/.

[3] Sarah Kuta. „Art Made With Artificial Intelligence Wins at State Fair“. In: (2022). URL: https://www.smithsonianmag.com/smart-news/artificial-intelligence-art-wins-colorado-state-fair-180980703/.

[4] James Vincent. „Getty Images is suing the creators of AI art tool Stable Diffusion for scraping its content“. In: (2023). URL: https://www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit.

[5] Greg Rutkowski. *NO TO AI GENERATED IMAGES*. 2022. URL: https://www.artstation.com/artwork/g8E2xZ.

[6] Sam Witteveen und Martin Andrews. *Investigating Prompt Engineering in Diffusion Models*. 2022. DOI: 10.48550/ARXIV.2211.15462. URL: https://arxiv.org/abs/2211.15462.

[7] Jules White u. a. *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT*. 2023. DOI: 10.48550/ARXIV.2302.11382. URL: https://arxiv.org/abs/2302.11382.

[8] Vivian Liu und Lydia B. Chilton. *Design Guidelines for Prompt Engineering Text-to-Image Generative Models*. 2021. DOI: 10.48550/ARXIV.2109.06977. URL: https://arxiv.org/abs/2109.06977.

[9] Melissa Heikkilä. *This artist is dominating AI-generated art. And he's not happy about it.* 2022. URL: https://www.technologyreview.com/2022/09/16/1059598/this-artist-is-dominating-ai-generated-art-and-hes-not-happy-about-it/.

[10] *Cargo cult programming*. https://en.wikipedia.org/wiki/Cargo_cult_programming. Accessed: 2023-03-10.

[11] *Lexica Art*. https://lexica.art/. Accessed: 2023-03-02.

[12] Zijie J. Wang u. a. „DiffusionDB: A Large-Scale Prompt Gallery Dataset for Text-to-Image Generative Models“. In: *arXiv:2210.14896 [cs]* (2022). URL: https://arxiv.org/abs/2210.14896.

[13] Robin Rombach u. a. „High-Resolution Image Synthesis With Latent Diffusion Models“. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Juni 2022, S. 10684–10695.

[14] https://twitter.com/sureailabs https://twitter.com/proximasan https://twitter.com/KyrickYoung https://twitter.com/sureailabs https://twitter.com/KyrickYoung. *Image Synthesis Style Studies Database (The List)*. Project Webpage https://proximacentaurib.notion.site/proximacentaurib/parro zone-74a5c04d4feb4f12b52a41fc8750b205. 2023. URL: https://docs.google.com/spreadsheets/d/14xTqtuV3BuKDNhLotB_d1aFlBGnDJOY0BRXJ8-86GpA.

[15] *Stable Diffusion artist list*. 2023. URL: https://stablediffusion.fr/artist.

[16] *List of art movements*. https://en.wikipedia.org/wiki/List_of_art_movements. Accessed: 2023-03-11.

Erklärung

Ich versichere, dass ich die vorliegende Arbeit mit dem Thema:

„Investigating artist and style mentions in Diffusion Model prompts“

selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe, insbesondere sind wörtliche oder sinngemäße Zitate als solche gekennzeichnet. Mir ist bekannt, dass Zuwiderhandlung auch nachträglich zur Aberkennung des Abschlusses führen kann.

Leipzig, den 15.03.2023

Georg Schneeberger

GEORG SCHNEEBERGER

Separation of Duties

Moritz Weinrich and I worked together on this topic until beginning of March 2023. Up to that point we shared our work and met up to discuss the progress. During the initial discussion with Dr. Andreas Niekler and Dr. Thomas Efer we found the dataset [12] and decided to use that for our project. Subsequently, Moritz researched mappings of artists to art-styles. Moritz found the two datasets [14] and [16]. During that time, I mainly analysed the [12] dataset and the distibutions of artists present in the dataset.

Appendix

5.1. Code

The code of this project is available publicly at https://github.com/geschnee/dh_project The read-me file describes how to get the code running and which files produced the results of this project. Scripts that did not contribute to the final results are also available in the repository.

5.2. Artist Dataset

The artist dataset [14] was modified to fit the needs of this project. The modified dataset is available in the GitHub repository https://github.com/geschnee/dh_project

5.3. Tables and Figures

A cavalier knight approaching a castle, future, render of dreamy beautiful landscape, fantasy dreamy, artgerm, large scale, details, vintage, photo, hyper realistic, ltra realistic, photography, unreal engine, high detailed, 8 k

Abbildung 5.1.: An example of a prompt with several specifier clauses [12]

Column Name	Type	Description
image_name	string	Image UUID filename.
prompt	string	The text prompt used to generate this image.
part_id	uint16	Folder ID of this image.
seed	uint32	Random seed used to generate this image.
step	uint16	Step count (hyperparameter).
cfg	float32	Guidance scale (hyperparameter).
sampler	uint8	Sampler method (hyperparameter). Mapping: 1: 'ddim', 2: 'plms', 3: 'k_euler', 4: 'k_euler_ancestral', 5: 'k_heun', 6: 'k_dpm_2', 7: 'k_dpm_2_ancestral', 8: 'k_lms', 9: 'others'.
width	uint16	Image width.
height	uint16	Image height.
user_name	string	The unique discord ID's SHA256 hash of the user who generated this image. For example, the hash for xiaohk#3146 is e285b7ef63be99e9107cecd79b280bde 602f17e0ca8363cb7a0889b67f0b5ed0. 'deleted_account' refer to users who have deleted their accounts. None means the image has been deleted before we scrape it for the second time.
timestamp	timestamp	UTC Timestamp when this image was generated. None means the image has been deleted before we scrape it for the second time. Note that timestamp is not accurate for duplicate images that have the same prompt, hypareparameters, width, height.
image_nsfw	float32	Likelihood of an image being NSFW. Scores are predicted by LAION's state-of-art NSFW detector (range from 0 to 1). A score of 2.0 means the image has already been flagged as NSFW and blurred by Stable Diffusion.
prompt_nsfw	float32	Likelihood of a prompt being NSFW. Scores are predicted by the library Detoxicy. Each score represents the maximum of toxicity and sexual_explicit (range from 0 to 1).

Tabelle 5.1.: Data Fields descriptions [12]