Using Generative Deep Learning Models to create NFT Artwork

COMP4971C Independent Work Final Report (Spring 2022)

By Mohamed Sobhy, MORSI

Supervised by Dr. David Rossiter

Abstract

This project uses the pre-trained generative deep learning models of Vector Quantized Generative Adversarial Network and Contrastive Language–Image Pre-training (VQGAN+CLIP) to synthesize digital valuables in artwork published as Non-Fungible Tokens. The project involved doing a brief Exploratory data analysis to figure out the general trends in the NFT marketplace between 2017 and 2021. Further, some data cleaning and data processing methods are applied to remove outliers and pick the high-valued NFTs based on some of their characteristics, such as price and likes count. After the data processing stage, a pipeline is created to automate the process of image processing, model loading, and model prediction. Finally, the results are automatically uploaded to the cloud. For demonstration purposes, a collection of 100 NFT images is chosen and published on rarible.com, one of the leading online marketplaces for trading NFTs.

Table of Contents

Table of Figures

Table of Data

Introduction

Recently, humans have changed the way they value things, especially in the last few years. Virtual assets have gained more attention and value. This vast shift was made possible due to the convenience of exchanging virtual items in a decentralized fashion thanks to innovative technologies such as the Blockchain.

Some of the most commonly traded items are Non Fungible Tokens (NFTs), which are blockchain-based tokens where each token represents a unique asset like a piece of art, digital content, or media. An NFT can be thought of as an irrevocable digital certificate of ownership and authenticity for an asset, whether digital or physical. Fundamentally, NFTs allow users to buy and sell ownership of unique digital items, such as drawings, GIFs, memes, songs, collectables, etc, while keeping track of the owner's information on the Ethereum blockchain ledger. This project is investigating whether it is possible to create high-valued NFTs from existing ones by learning their common features and replicating them, analogous to human artists getting inspired by existing paintings before creating their own.

Blockchain

A blockchain is a system of recording information in a way such that it is difficult or impossible to change, hack, or cheat the system. As a database, a blockchain stores information electronically in digital format. Blockchains are best known for their crucial role in maintaining secure and decentralized transaction records for cryptocurrency systems, such as Bitcoin. The cutting-edge innovation of blockchains guarantees the fidelity and security of data records and generates trust without the need for a trusted third party. Those who run blockchain programs, known as the miners, are responsible for maintaining the security of a blockchain with consensus protocols such as proof of work (more computations lead to more rewards). Such decentralization makes the key to blockchain security because no single entity can control the system or modify the data recorded on the blockchain (unless a 51% attack occurs, which is unlikely or nearly impossible for popular blockchains). For this reason, blockchains are widely regarded as secure. **[4]**

An interesting application of blockchains is non-fungible tokens (NFTs), which attract numerous users and investors. Unlike cryptocurrencies such as Bitcoin or Ethereum which can be traded or exchanged for one another, each NFT has a unique digital identifier that makes it impossible for NFTs to be exchanged or for two NFTs to be equivalent. This means they are not mutually interchangeable, and hence not fungible. A non-fungible token can represent the ownership of a digital or physical asset (i.e. digital art or a painting) on a blockchain. A key property of NFTs is that they can be easily traded and sold on a blockchain (mostly on Ethereum now). A representative example is the CryptoKitties where users can buy, own, and sell virtual cats. There are also marketplaces such as OpenSea and Rarible for users to buy and sell NFTs. The marketplace of NFTs creates a new way for artists to sell their artworks easily through the blockchain. **[4]**

Non Fungible Tokens

Since its inception in 2015, non-fungible tokens (NFTs) have developed rapidly, disrupting the art world and other industries, while raking in a significant amount in revenue. Non-fungible tokens (NFTs) are cryptographic assets on a blockchain with unique identification codes and metadata that distinguish them from each other. Unlike cryptocurrencies and fiat money, they cannot be traded or exchanged at equivalency. This differs from fungible tokens like cryptocurrencies, which are identical to each other and, therefore, can serve as a medium for commercial transactions. So far, Ether has been the most often utilised cryptocurrency for transacting NFTs, owing to its dual role as the native asset for the blockchain platform (Ethereum) that hosts the vast majority of decentralised applications and NFTs. **[2]**

The market value of NFTs is rapidly increasing. According to Thursday's NonFungible.com report, which was created with support from L'Atelier BNP Paribas, the total value of all NFT transactions worldwide jumped 21,350% to more than $17 billion in 2021, from $82.5 million in 2020. The 200-fold increase in the NFT market was 100 times as great as the percentage increase in the electric-vehicle market over the same period, the report says. **[14]**

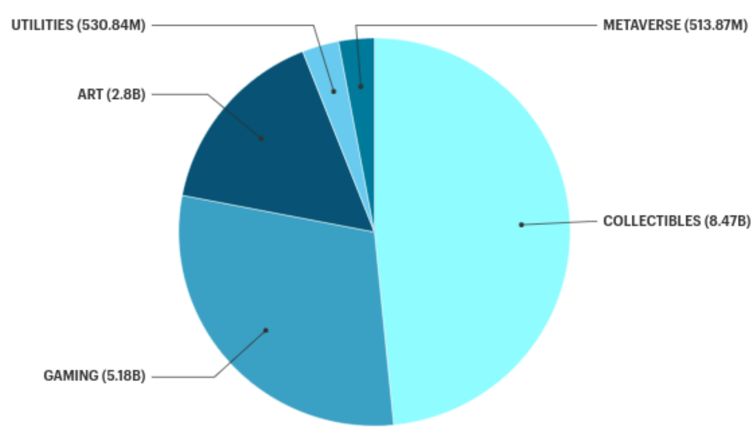

Figure 1: The NFT marketplace in numbers in 2021

SOURCE: L'ATELIER BNP PARIBAS AND NONFUNGIBLE.COM

Figure 1 shows that the total value of Art NFTs is 2.8B USD, making up almost 16% of the market value. **[14]**

Software

The project used python 3 and PyTorch running on Google Colaboratory (https://colab.research.google.com/) as a working environment to access Google computing resources; the GPU used is **NVIDIA Tesla K80**. CUDA is a software platform that pairs with Nividia GPU hardware, which speeds up computations using the parallel processing power of Nvidia GPUs. It has sped up most of the computations involved when running the model in the project.

Some of the python libraries used are pandas for tabular data manipulation, and Seaborn which is a Python visualization library based on matplotlib, to create plots that provide visual insights into the dataset distribution.

OpenCV is a library of programming functions mainly aimed at real-time computer vision. It is used for the Exploratory Data Analysis section to display images, and for image preprocessing in a later stage.

Method

Generative Adversarial Networks are a type of deep learning generative model that can achieve startlingly photorealistic results on a range of image synthesis and image-to-image translation problems.

The project used Vector Quantized Generative Adversarial Network and Contrastive Language–Image pre-training (VQGAN+CLIP), which was pre-trained on the wikiart dataset to create NFT artwork based on a dataset of some of the famous NFT artwork.

Data

This project used a dataset called NFT art collection 2021 which is collected from various NFT showrooms. The dataset contained various forms of artistic NFT, such as photos, GIFs, and videos. The scope of the project is to synthesize image art, so only the photos portion of the dataset is used.

The dataset structure:

Column	Description
title	Title of the art piece.
name	Name of the art piece
creator	Creator of the art piece
art_series	Name of the collection the art piece is a part of
price	Price of the art piece in the given symbol
symbol	Currency the art piece is sold in
type	Type of art
likes	Amount of likes the art piece got
nsfw	Label if the art piece is safe for work (Example of not safe for work is sexual art pieces)
tokens	Amount of art pieces for sale
year	Year the author created the art piece
rights	1 = Private 3 = Limited production
royalty	Unused
cid	IPFS hash
path	Path to the art piece.

Table 1: Dataset categories

Exploratory Data Analysis

This section is an analysis of the dataset aimed at summarizing its underlying structure, pointing out its key feature, and detecting outliers or anomalies.

Data source: https://www.kaggle.com/vepnar/nft-art-dataset

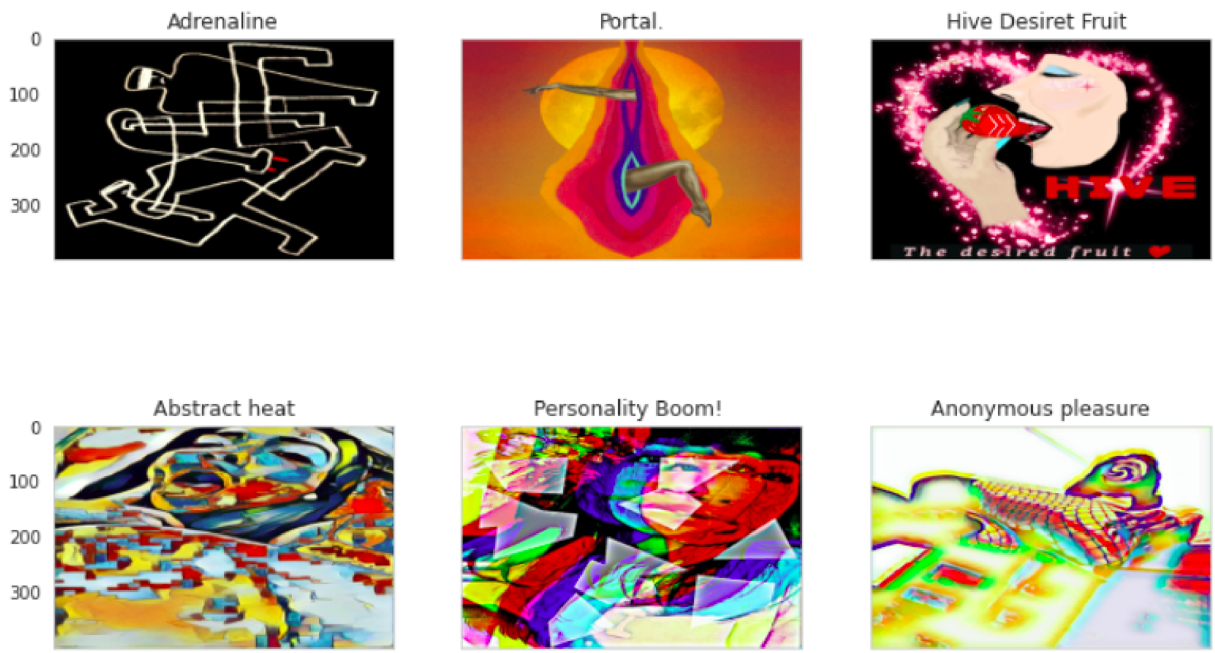

Figure 2: Examples of some image art pieces with their titles

Deciding a time frame for analysis

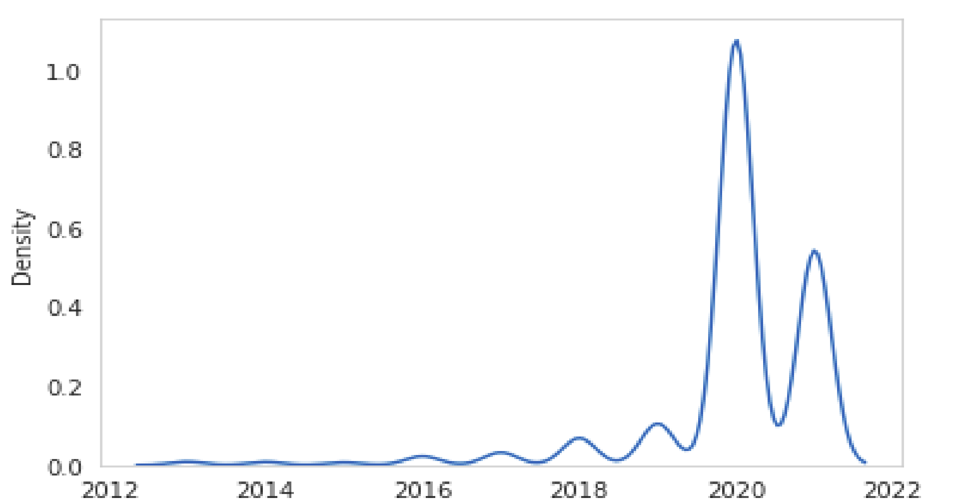

Figure 3: A density distribution of the creation year of the art pieces in the dataset between 2012 and 2021

The plot shows an upward spike in the number of art pieces created in 2020, followed by a smaller peak in 2021. As most art pieces were created between 2017 and 2021, this section will focus on analyzing this time frame.

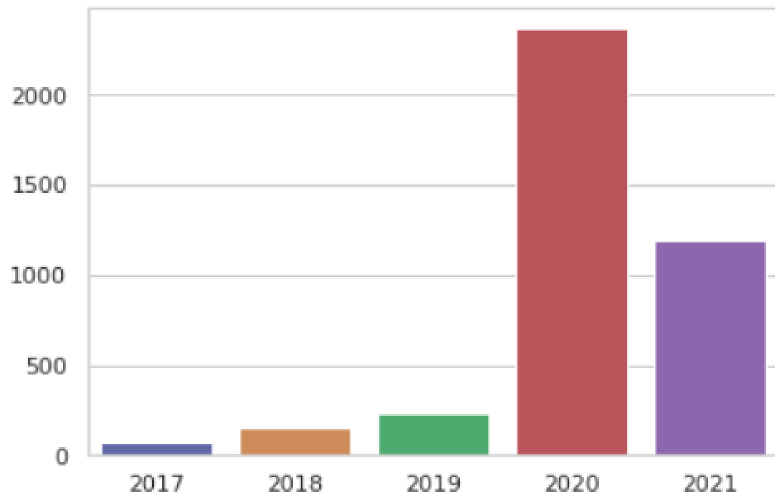

Figure 4: Bar plot for the year of creation of NFTs in the dataset **[8]**

Types of Artwork

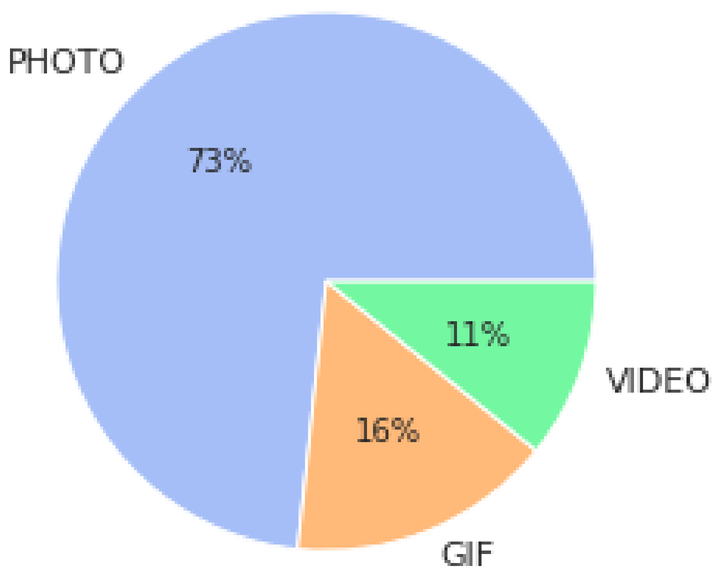

Figure 5: A pie plot of the categories **[8]**

Price Statistics

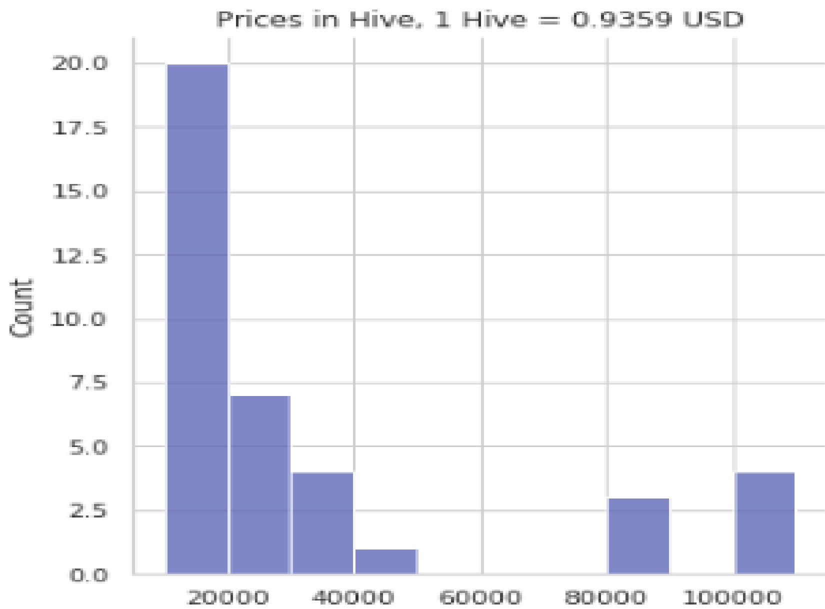

Figure 6: A distribution plot for the listed prices **[8]**

Year	Number of Art pieces	Average Price	Median Price
2017	69	205.0	75.0
2018	149	230.4	66.0
2019	231	828.9	50.0
2020	2368	1775.4	70.0
2021	1196	716.0	50.0
Total	4013	1320.8	60.0

Table 2: Prices Statistics for years 2017-2021

average price	2006.9 coins
Median Price	60.0 coins
The standard deviation	58981.4

Table 3: Summary of prices central tendency

The table shows a clear discrepancy between the average and median prices, which is mainly caused by the outliers. A more realistic overview of the prices is achieved by dropping the top 1% and lowest 1% of art pieces in terms of the price.

Year	Number of Art pieces	Average Price	Median Price
2017	66	138.5	72.5
2018	145	135.2	66.0
2019	224	229.6	50.0
2020	2314	185.0	70.0
2021	1156	145.1	50.0
Total	3905	173.1	60.0

Table 4: Prices Statistics for years 2017-2021 after trimming by 1%

The new averages after dropping some outliers are more consistent with the median price. Choosing to trim the 1% lowest and highest values is arbitrary; increasing the percentage from 1% decreases the gap between the average and median prices. Such a big gap indicates the presence of outliers in the dataset.

Year	Number of Art pieces	Average Price	Median Price
2017	60	110.1	77.5
2018	129	100.9	66.0
2019	207	106.3	50.0
2020	2121	124.6	70.0
2021	1023	89.6	50.0
Total	3541	112.3	60.0

Table 5: Prices Statistics for years 2017-2021 after trimming by 5%

We conclude that removing the top 5% and lowest 5% is enough to ensure the data is consistent. Removing such outliers is essential since their listed prices do not reflect their actual values, while the project focuses on the common characteristics of high-valued NFTs and not special cases.

Data Processing

Two types of processing are applied to the dataset: automated and manual processing. Firstly, using OpenCV, all images are resized into 512x512 for consistency. Second, images of low quality and images that contain much text are manually eliminated to increase model stability. Finally, the dataset is divided into 5 subcategories based on their context. All images that do not fit into one of the categories are excluded for simplicity.

Category	Image count
Animals	270
Faces	383
People	454
Objects	100
Scenery	313

Table 6: Image count in each subcategory after all the processing

Generative Deep Learning Models

Generative Adversarial Networks (GANs)

Generative adversarial networks (GANs) are algorithmic architectures that use two neural networks, pitting one against the other (thus the "adversarial") in order to generate new, synthetic instances of data that can pass for real data. They are used widely in image generation, video generation and voice generation. **[12]**

The first neural network is referred to as the generator. It generates new data instances, while the other network, called the *discriminator*, evaluates them for authenticity. In other words, the discriminator decides whether each instance of data that it reviews belong to the actual training dataset or not. The goal of the discriminator, when shown an instance from the true dataset, is to recognize those that are authentic. Meanwhile, the generator is creating new, synthetic images that it passes to the discriminator. **[12]** The goal of the generator is to generate as realistic samples as possible, and the goal of the discriminator is to identify images coming from the generator as fake and the ones coming from the dataset as real.

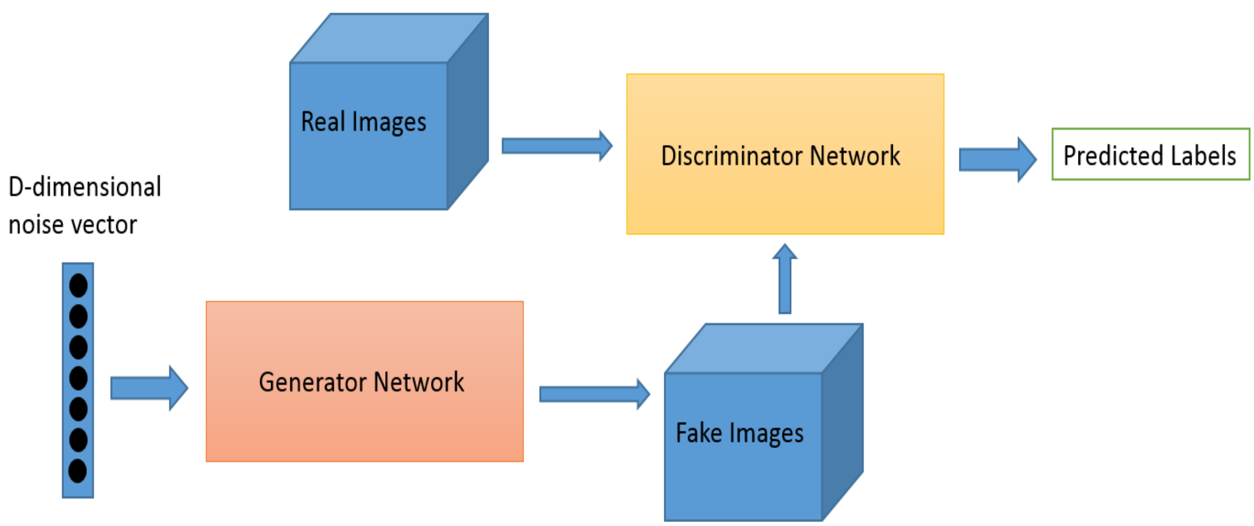

Figure 7: Generative adversarial neural networks architecture. **[12]**

VQGAN+CLIP

VQGAN and CLIP are two separate machine learning algorithms that can be used together to generate images based on a text prompt. VQGAN is a generative adversarial neural network that is good at generating images that look similar to others (but not from a prompt).CLIP is another neural network that is able to determine how well a caption (or prompt) matches an image. The two algorithms were combined in various forms by AI-generated art enthusiasts like Ryan Murdock and Katherine Crowson. **[15]**

This method can be used to either synthesize an image from random noise based on the text prompt or take an initial image and morph other images into it based on the text prompt. This project used the categories name as the text prompt, which ensures that each image will be affected by other images from the same subcategory or a related context in general.

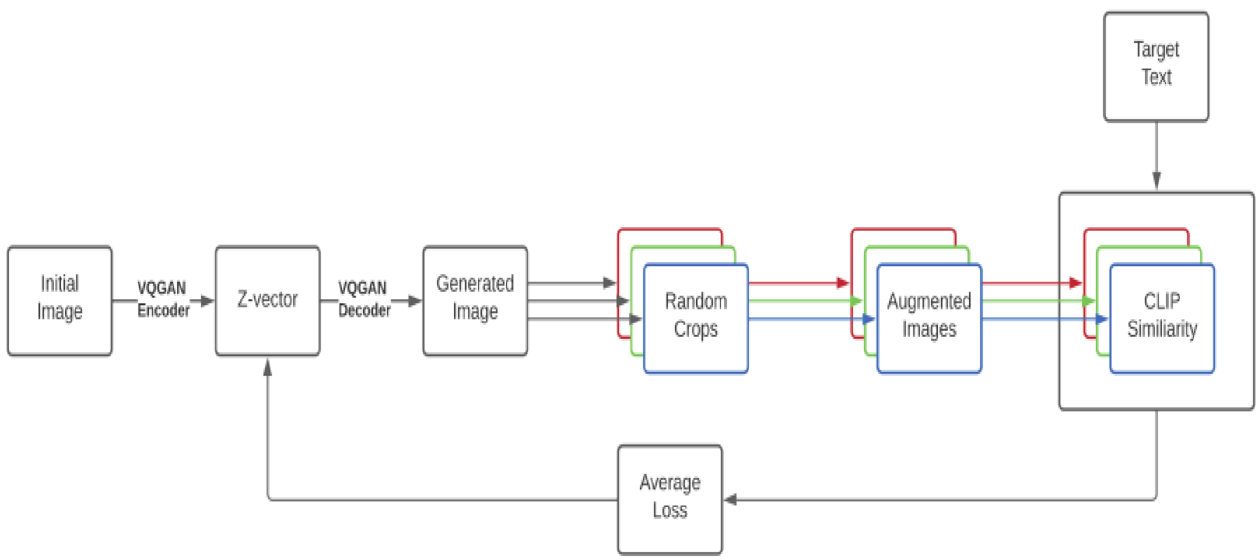

Figure 8: VQGAN+CLIP Architecture **[7]**

Model Training

The VQGAN+CLIP model is pre-trained on the wikiart dataset, which is a collection of paintings from 195 different artists. It contains 42129 images for training and 10628 images for testing. **[17]** Wikiart is suitable for the project since the NFTs in the dataset are mostly art pieces, which means the images of both datasets used (wikiart and the NFT collection) have similar characteristics. The pre-trained model is used to synthesize new samples, but instead of initializing the model with random noise. We used images from the NFT dataset. This approach ensures the model applies the representation it learned from wikiart dataset to the NFT dataset.

Parameter	Description	Assigned value
height	Height of output images	512
width	Width of output images	512
initial_image	The initial image from the NFT dataset	One image from the NFT dataset per iteration
seed	The seed used for randomization	-1
max_iterations	Maximum number of training steps the model should make on each image	20

Table 7: Model Parameters

During each iteration, the model takes one image from the NFT dataset and outputs gradually changes it over 20 steps (since we assigned 20 to max_iterations). All the outputs are automatically uploaded to the Cloud to avoid progress loss if the runtime is interrupted. The next two pages are examples of the outputs during two different iterations of the model.

Figure 9: First sample iteration of the model

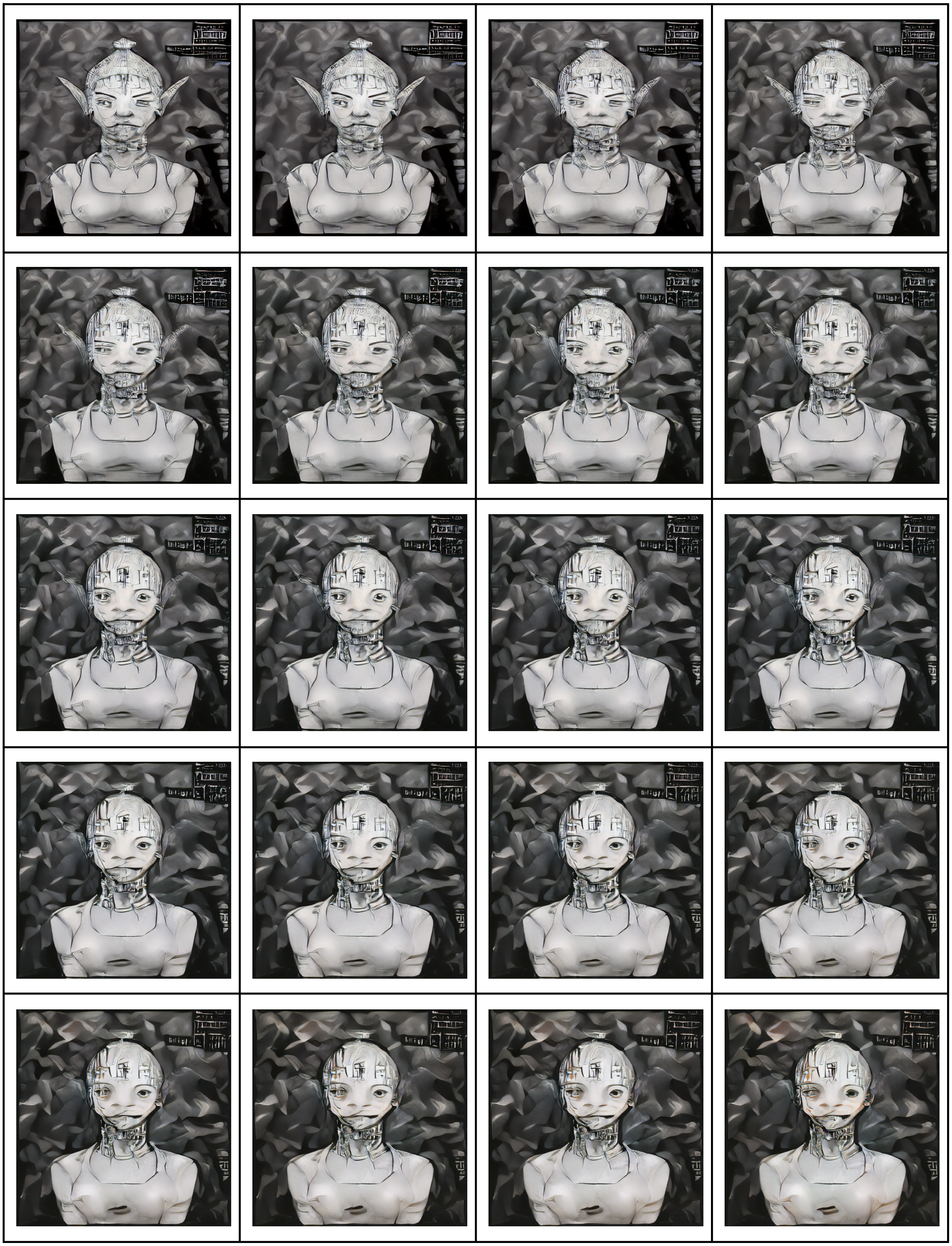

Figure 10: Second sample iteration of the model

Results Evaluation

There is no single definitive measure that determines the intrinsic value of NFTs. Instead, several factors are taken into consideration to assess the created value qualitatively and quantitatively. Some of them are fully subjective, while others are more objective. Some of the following measures are used in this project.

Human Interpretation

During each iteration of the model, an NFT is created with 20 slightly different versions. Human interpretation is used to assess their quality and manually choose the best one. The criteria for selection is choosing the image with the highest quality, and least amount of artifacts while capturing some features of the original art piece.

Objective Artistic Evaluation

There are many objective characteristics of good art on the technical level, unlike the human interpretation, this is not about selecting the most appealing art piece from a set of candidates, rather it is about judging the output based on objective qualities such as: the amount of details, colour choices, degree of realism, …etc. This is not a measure of art novelty in any means. It can be thought of as a measure of how well the output is following some common art principles. Many principles can be borrowed from aesthetics to reach a fair judgement of the results. However, it was not implemented in this project since it involves a human expert critic.

Neural Network Critique

Since the results were created using a Generative Adversarial Network, which included a generative network and a discriminative network, the latter was optimized to be an art critic during model training. The output of the discriminative network is a percentage that represents how likely the sample is real according to the network. Given that the network is well optimized, such a percentage could quantitively measure the overall value of the results, simply because the more the network thinks the output is real, the more it is likely to be a good piece of art. Another approach is to use a different neural network that was solely trained for art critique.

Monetary Value

The most obvious way to assess the monetary aspect of the project is to list the results for sale in an NFT marketplace and consider the revenues over a specific timeframe compared to the original pieces. Rarible is an NFT online marketplace built on the Ethereum BlockChain, and it was chosen for this purpose because of its new lazy minting option. Lazy minting allows creators to sell their NFTs on the market without paying gas fees, by processing the transactions needed to store the information on the Ethereum BlockChain only when the NFT is purchased, so the NFTs would be kept off-chain, and only be minted when necessary. Once the NFT is purchased, the minting process is done normally.

A collection of 100 images of the results is to be listed for bids on Rarible, to gain better insights on how much money buyers are willing to pay for them, which is more accurate information than listing them at a fixed price.

This measure complements the previous objective artistic measures, since on a fundamental level, it is enough to consider two factors: the intrinsic value of the art by assessing how much an art piece is objectively worth, and the market value which could be affected by many factors like supply and demand or other external factors, hence it is more subjective yet more practical to use.

Possible Further Optimizations

The scope of this project was to investigate popular NFTs in a specific time frame (between 2017 and 2021). A more comprehensive is achieved by automatically scraping the information from popular NFT marketplaces, and directly feeding the high-valued ones into the model. The following flowchart is a demonstration of the proposed optimizations..

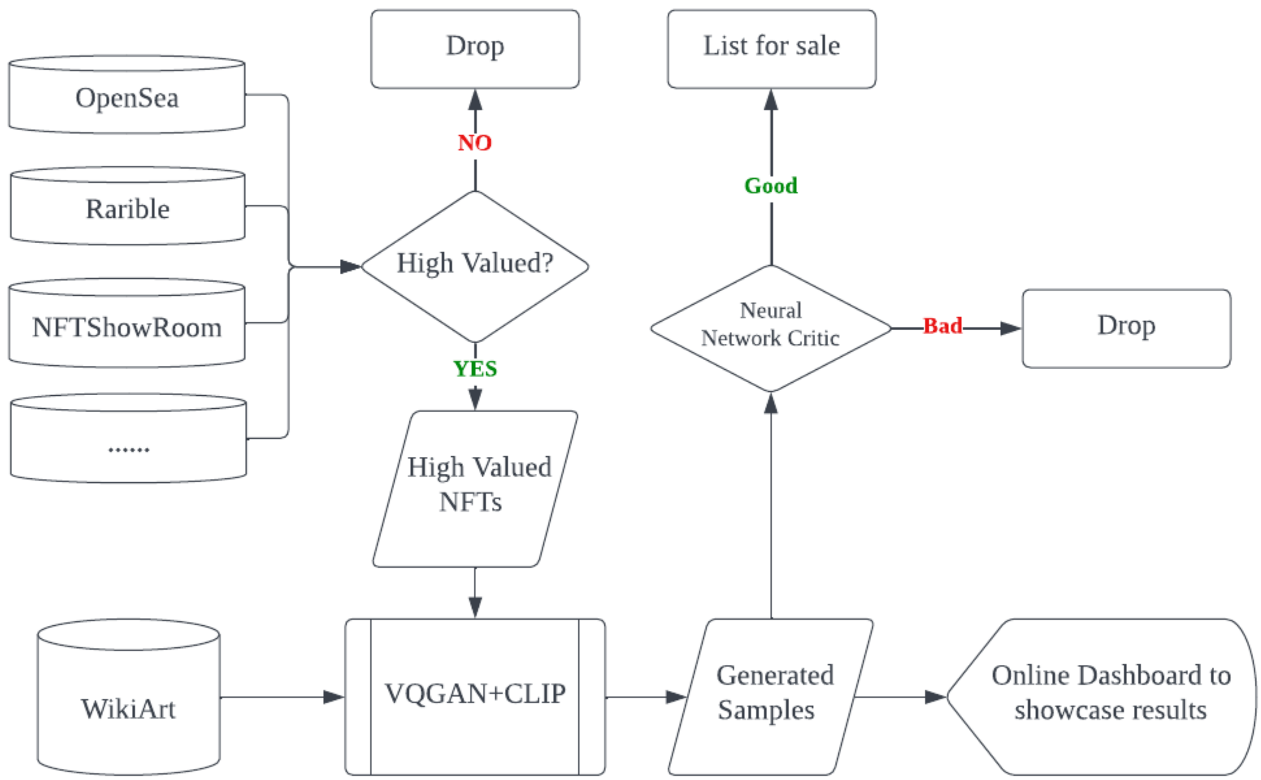

Figure 11: Project Flow Chart with some possible further optimizations

The image quality results could be further improved using ERSGAN, which is a machine learning technique that can be trained to upscale low-resolution images into higher resolutions. It is trained by comparing low-resolution images to their high-resolution counterparts and can enhance and can super-enhance 500x500 images to 1746x1746 without quality loss.

References

[1] Areeb Q. M., Imam R., Fatima N., & Nadeem M. *AI Art Critic: Artistic Classification of Poster Images using Neural Networks,* 2021 International Conference on Data Analytics for Business and Industry (ICDABI), 2021, pp. 37-41, doi: 10.1109/ICDABI53623.2021.9655869.

[2] Bosun, O. (2022, February 4). *NFT marketplaces and the impending rise of Fiat currency for purchases*. Yahoo! Finance. Retrieved April 14, 2022, from https://finance.yahoo.com/news/nft-marketplaces-impending-rise-fiat-093011185.html

[3] Büsken, D. (2020, April 21). 6 characteristics of good art – what makes a drawing good? BüeskenArt. Retrieved May 9, 2022, from https://bueskenart.com/characteristics-of-good-art/

[4] CHEN, Q. (2022, February 24). Blockchain and NFTs explained. STEM@HKUST. Retrieved April 27, 2022, from https://stem.hkust.edu.hk/insights/blockchainandNFTsexplained

[5] Chello, A. (2021, December 1). *The internet of assets (IOA)*. Medium. Retrieved April 28, 2022, from https://chelloandrea.medium.com/the-internet-of-assets-ioa-d9ce104cc1a2

[6] Conti, R., & Schmidt, J. (2022, April 8). *What is an NFT? non-fungible tokens explained. Forbes*. Retrieved April 14, 2022, from https://www.forbes.com/advisor/investing/cryptocurrency/nft-non-fungible-token/

[7] Crowson, K., Biderman, S., Kornis, D., Stander, D., Hallahan, E., Castricato, L., and Raff E. *VQGAN-CLIP: Open Domain Image Generation and Editing with Natural Language Guidance.* Retrieved from: https://doi.org/10.48550/arXiv.2204.08583

[8] Exploratoray Data Analysis: https://www.kaggle.com/code/mohamedsobhi777/nft-eda

[9] Gautam, T. (2021, July 5). *Vqgan: Taming Transformers for high-resolution image synthesis*. Analytics Vidhya. Retrieved April 29, 2022, from https://www.analyticsvidhya.com/blog/2021/07/understanding-taming-transformers-for-high-resolution-image-synthesis-vqgan/

[10] Goodfellow, Ian J., Pouget-Abadie, Jean, Mirza, Mehdi, Xu, Bing, Warde-Farley, David, Ozair, Sherjil, Courville, Aaron C., and Bengio, Yoshua. *Generative adversarial nets*. NIPS, 2014.

[11] Hare, V. (2021, November 17). *NFT vs. Tokenization. Secure Any Sensitive Data Element.* Retrieved April 30, 2022, from https://www.tokenex.com/blog/nft-vs.-tokenization

[12] Nicholson, C. *A beginner's guide to generative adversarial networks (gans).* Pathmind. Retrieved April 29, 2022, from https://wiki.pathmind.com/generative-adversarial-network-gan

[13] Project GitHub Repository: https://github.com/mohamedsobhi777/COMP4971C---Independent-Study-Project

[14] Quiroz-Gutierrez, M. (2022, March 10). *NFTs had a great 2021, but will they hold their value in 2022?* Fortune. Retrieved April 30, 2022, from https://fortune.com/2022/03/10/bored-apes-cryptopunks-jolt-nft-market-to-billions-in-sales/

[15] Russell, A. (2021, August 15). *How to use VQGAN+CLIP to generate images from a text prompt- a complete, non-technical tutorial.* Medium. Retrieved April 27, 2022, from https://medium.com/nightcafe-creator/vqgan-clip-tutorial-a411402cf3ad

[16] Waskom, M. *Statistical Data Visualization*. seaborn. Retrieved April 27, 2022, from https://seaborn.pydata.org/

[17] WikiArt visual art encyclopedia: https://paperswithcode.com/dataset/wikiart

Appendix: Glossary of Related Terminology

The internet of assets (IoA)	Refers to the direct and provable ownership of internet assets.
Tokenization	Dividing up and selling off shares of assets ranging from real estate to artwork.
Blockchain	A blockchain is a growing list of records, called blocks, that are securely linked together using cryptography.
Wallet	Software that holds public and/or private keys for cryptocurrency transactions.
Smart Contract	A smart contract is a self-executing contract with the terms of the agreement between buyer and seller being directly written into lines of code. The code and the agreements contained therein exist across a distributed, decentralized blockchain network.
Proof of Stake	PoS asks users to prove ownership of a certain amount of cryptocurrency (their “stake” in the network) in order to be able to participate in the validation of transactions).
Proof of Work	A piece of data (the proof) that requires significant computation to find.
DEX	Open marketplaces for ETH and other tokens. They connect buyers and sellers directly.
Gas Fee	Price for making a transaction on a blockchain.
Ethereum	Ethereum is a decentralized, open-source blockchain with smart contract functionality.

Ether	The native cryptocurrency of the Ethereum platform.
Polygon	Polygon is a blockchain that provides scalable, secure, and instant transactions with Ethereum currencies like ETH, USDC, and DAI, offering low gas fees and high speeds without sacrificing security.
Allowlist	A method of pre-registering for an NFT drop to gain access to purchasing.
Minting	The process of creating an NFT on the blockchain.
Metadata	All necessary and unique data that define an NFT
Opensea	The leading NFT marketplace on Ethereum network
MetaMask	An NFT wallet in the Ethereum ecosystem
GANs	Generative Adversarial Networks are deep-learning-based generative models.

Table 8: Terminology Glossary