ACADEMY OF PERFORMING ARTS IN PRAGUE
FILM AND TELEVISION FACULTY

Film, Television, Photography, and New Media

Photography

MASTER'S THESIS

Creative Applications of Machine Learning in Photography

Alexander Rossa

Thesis advisor: Tomáš Dvořák
Examiner:
Date of thesis defense:
Academic title granted: MgA.

Prague, 2021

AKADEMIE MÚZICKÝCH UMĚNÍ V PRAZE

FILMOVÁ A TELEVIZNÍ FAKULTA

Filmové, televizní a fotografické umění a nová média

Fotografie

DIPLOMOVÁ PRÁCE

Kreativní aplikace strojového učení ve fotografii

Alexander Rossa

Vedoucí práce: Tomáš Dvořák

Oponent práce:

Datum obhajoby:

Přidělovaný akademický titul: MgA.

Praha, 2021

Declaration

I declare that I have prepared my Master's Thesis independently on the following topic:

under the expert guidance of my thesis advisor and with the use of the cited literature and sources.

Prague, date: Signature of the candidate:

Warning

User’s Log

By signing below, the user attests that he/she has used this thesis solely for study purposes, and he/she declares that he/she will always cite the thesis appropriately among the sources used

Name	Institution	Date	Signature

Keywords

artificial intelligence; machine learning; deep learning; neural networks; generative adversarial networks; image synthesis; computational photography; creativity

Klíčová slova

umělá inteligence; strojové učení; hluboké učení; neuronové sítě; generativní adversariální sítě; syntéza obrazu; komputační fotografie; kreativita

Abstract

Artificial intelligence and deep learning hold considerable potential to be utilized for creative work, particularly in photography. This thesis focuses on the subject of machine learning via generative adversarial networks and its potential applications for creatives and artists in this field. In recent years machine learning has improved significantly as a result of big data and leaps in computer processing power. While the recent resurgence of novelty surrounding artificial intelligence has already raised substantial questions and criticisms in regards to its unforeseen consequences on our evolving relationship with technology in general, there is also a growing optimism aided by an increased democratization of access to machine learning. Visual artists from various backgrounds are learning to utilize artificial neural networks to conceptualize solutions for creative problems, as well as to augment and enhance their own work methods. This paper aims to examine the current state of artificial intelligence in photographic practices by exploring its technological innovations, creative tools, and applications. It will also address the problems of machine learning and AI-assisted art, and finally conclude with a prognosis for its future output potential.

Abstrakt

Umělá inteligence a hluboké učení mají značný potenciál pro kreativní práci, zejména ve fotografii. Tato práce se zaměřuje na předmět strojového učení prostřednictvím generativních adversariálních sítí a jeho potenciálních aplikací pro kreativu a umělce v této oblasti. V posledních letech se strojové učení významně zlepšilo v důsledku velkých dat a skoků ve výpočetním výkonu počítače. Zatímco nedávné diskuse obklopující umělou inteligenci již vyvolaly značné otázky a kritiku, pokud jde o její nepředvídatelné důsledky pro náš vyvíjející se vztah s technologií obecně, roste také optimismus, kterému napomáhá zvýšená demokratizace přístupu ke strojovému učení. Výtvarní umělci z různých prostředí se učí využívat umělé neuronové sítě k vytváření koncepcí a řešení kreativních problémů a také k rozšiřování a zdokonalování vlastních pracovních metod. Tato práce si klade za cíl prozkoumat současný stav umělé inteligence ve fotografických postupech zkoumáním jejích technologických inovací, kreativních nástrojů a aplikací. Bude se také zabývat problémy strojového učení a uměním podporovaným umělou inteligencí a nakonec uzavře prognózu jeho budoucího potenciálu.

Table of Contents

List of Abbreviations

AI	Artificial Intelligence
ML	Machine Learning
GUI	Graphical User Interface
ANN	Artificial Neural Network
DL	Deep Learning
CNN	Convolutional Neural Network
CPU	Central Processing Unit
GPU	Graphical Processing Unit
GAN	Generative Adversarial Network
SR	Super-Resolution
VAE	Variational Autoencoder
GPT-3	Generative Pre-trained Transformer 3
NLP	Natural Language Processing

1. Introduction

The purpose of this dissertation is to shed light on the emerging technologies of artificial intelligence (AI) and machine learning (ML) in the context of their creative applications, including photography and other artistic practices. A novel subject, often obscured by its technical complexity, might appear to be out of bounds to anyone without previous experience in computer science and engineering. However, with rapid advances in recent years, machine learning is becoming a much more accessible medium than it used to be, as user-friendly tools and code snippets are continuously being developed and shared by communities across the web. My motivation to write this thesis arose out of my own research into machine learning, from the point where I needed to determine a starting point for the artistic purposes that I wanted to realize. This led me to the discovery of a myriad of applications and ML models which facilitated my learning process and helped me become reasonably versed in this field.

Preceded by a historical overview of the relationship between art and technology and a technical explanation of the fundamentals of ML, this paper examines different approaches, including suggestions of generative models for image synthesis, photographic post-processing, and presents the most accessible ML tools for beginners and intermediate users. Following this, I feature a number of contemporary artists using AI in profoundly creative ways, some of who bring attention to the problems with ML and how its discourse can be framed to think more critically about these technologies. The thesis concludes with an evaluation of the research and a discussion of its current state, followed by my prognosis of its future potential and how creatives can perhaps best use ML to their advantage.

2. Machine Learning

2.1. Historical Background

In order to better contextualize the subject of machine learning, it is helpful to get acquainted with the intricate relationship which art and technology have had throughout the course of history. Tools as rudimentary as a chisel and as complex as computers have allowed human beings to record meaning and define the development of culture over many millennia. Cultural artifacts inherently represent a quality of intelligence and expressiveness that is absent in other lifeforms we share this planet with. With further innovations of human tools, we have been able to save, translate, and retrieve linguistic information with greater dependability and speed. Consistently throughout time, certain periods of human history were marked by an acceleration of technological growth whenever a new invention succeeded in revolutionizing previous methods. More recently, understandings regarding technology have been linked with digital systems of information and networked computing, allowing for greater creative potential. The field of artificial intelligence and machine learning is one such revolution and it is gradually reshaping visual culture and the technological fabric of society today.

The modern term for 'technology' derives from the Greek word for craft 'techne', or art. The contemporary usage of the term however slightly differs from its original meaning. In ancient Greece it was thought to be associated with the held belief that technology should always attempt to imitate nature. This principle known as 'mimesis', often associated with Aristotle, helped to shape thought throughout the 13th century, with philosophers and theological scholars presenting support for this conservative paradigm that artistic creations should reflect nature.[1] In the early Middle Ages, what we consider art today was mainly created for sacramental purposes, with Christian and Gothic paintings, sculptures, and stained glass being produced as objects for religious ornamentation. Creative work in this context was primarily commissioned on the basis of the artist serving a divine will, and individual authorship was not given the

[1] Schummer, J. "Aristotle on Technology and Nature." *Philosophia Naturalis,* vol. 38(1). 2001. pp. 105-120.

same importance it has today.[2] Artistic values identified with personal expressions of creativity didn't emerge until 15th and 16th century Florence, Italy, when more progressive social values such as personal enterprise and liberty were being emboldened by a burgeoning economy. With the cultural transition into the Renaissance, new technological developments established themselves throughout Europe, leading to a transition in art making that would include products of creativity and imagination rather than just the technical crafts used to represent religious and political narratives.[3]

F*igure 1:* Illustration of the camera obscura principle from James Ayscough's "A short account of the eye and nature of vision" (1755, Fourth Edition).

An important technological innovation of this time period was a tool modeled for the purpose of light studies and used as a painting aid, the camera obscura *(Figure 1)*. With many design variations being developed over the centuries, the basic functional principle which made it the precursor of modern photography was the same: it greatly helped improve the visual perspective of drawings and paintings. Consequently, discourse surrounding figuration and literal

[2] Wittkower, R. "Individualism in Art and Artists: A Renaissance Problem." *Journal of the History of Ideas.* 1961. pp. 291-302.
[3] Shiner. L. *The Invention of Art: A Cultural History.* University of Chicago Press. 2001.

representation in visual art saw a transformation of depiction which would also help give rise to a changing relationship between technology, art, and humanism, further drawing in ideas of individual creativity and shaping the intellectual landscape throughout the early modern period.[4]

With industrialization, cultural sentiments regarding technology had begun to shift. The effects of rapid urban expansion and mechanization, along with rationalist dogmas, were increasingly at odds with notions of individuality and personal creative expression. Mechanical operations, directed by scientific practices, were being regarded as predestined processes detached from the freedom of the human spirit.[5] This in large part led to the movement of Romanticism throughout Europe, which aimed to counteract the depersonalizing effects of machination. It did this by rejecting the purely mechanistic worldview in favor of intuition, emotion, and the reverence of natural beauty. This underlying skepticism of technology continued well into the 18th century when technological ideas became more closely joined with labor and the production of goods. Developments of new kinds of machinery made it possible for objects, which had traditionally been made by individual artisans, to be mass-produced in large factories. This led to the gradual replacement of skilled craftsmanship by various divisions of labor and mechanizations which produced commodified objects devoid of unique provenance.

A notable technological instrument of this time was the Jacquard loom *(Figure 2)*, which was a device fitted to a loom that simplified the process of manufacturing textiles with complex patterns. The machine, invented in 1804 by Joseph Marie Jacquard, employed a punched card system holding continuous sequences of analog data that automatically controlled the weaving mechanism, making it possible to produce unlimited varieties of pattern weaving.

---

[4] Shiner. L. *The Invention of Art: A Cultural History.* University of Chicago Press. 2001.
[5] Augustyn, A. et al. “Romanticism.” *Encyclopedia Britannica.* https://britannica.com/art/Romanticism. Accessed March 15, 2021.

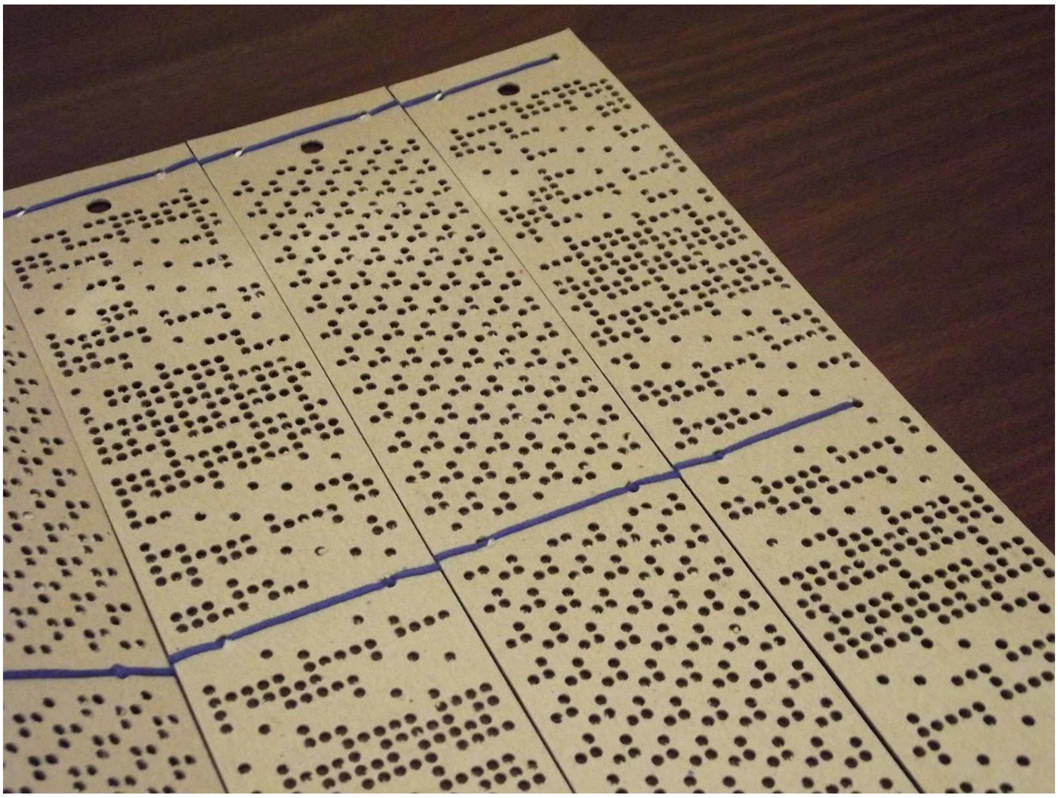

Figure 2: Series of punch cards on the Jacquard hand loom displayed in the McCarl Coverlet Gallery in Pennsylvania, USA.

This device is also considered to have played an important role in the historical development of computer hardware, which would eventually come to influence the invention of the Analytical Engine by Charles Babbage during the 1830s, commonly regarded as the first representation of a general-purpose computer.[6] The Jacquard loom, while significantly increasing the efficiency of textile production, also became a symbol of the cultural disruption felt during the boisterous establishment of industrialization. This was heavily reflected in the Luddites and textile workers of England who were losing their jobs and livelihoods due to this automation of labor, and in response many protested by destroying factory machines.[7]

The invention of photography also became a substantial disrupter during the 19th century, especially in the field of visual art. Some artists, especially those who aligned themselves with the doctrines of Romanticism, rejected photography for

[6] Batchen, G. “Electricity Made Visible”. *New Media, Old Media: A History and Theory Reader.* Edited by Chun; W. Keenan. Routledge. 2006. pp. 27-44.
[7] Bauer, M. *Resistance to New Technology: Nuclear Power, Information Technology and Biotechnology.* Cambridge University Press. 1997. pp. 57-79.

its purely mechanistic functioning and unfiltered depiction of the natural world. French poet Charles Baudelaire, likened it to printing technology and believed that photography inspired lazy and uncreative artworks through a deluded kind of reverence brought about by industrialization.[8] Afterwards, he reiterated his position in his 1864 essay "The Painter of Modern Life" in which he is attributed to first using the term 'modernité' to describe a discontinuation of history from one of constancy to an ever-changing paradigm consumed by the unpredictability of the future and the uniqueness of its own time.[9] His sentiments, influenced by a changing technological landscape, represented an age of remarkable transformation at the time, which was characterized by both enthusiasm and confusion, and the rejection of traditions. This break of established cultural norms welcoming new modes of expressions would ultimately become known as modernism.

In the early 20th century, German philosopher Walter Benjamin begins to discuss in his 1936 essay "The Work of Art in the Age of Mechanical Reproduction" the growing shift of cultural attitudes in relation to the rise of convenience in the production of reproducible media. For Benjamin the mechanical reproduction of images was inextricably linked with human experience and how we perceive visual artworks. He critiqued the reproduction of art of having lost its aesthetic authority and authenticity due to the lack of 'aura', a quality of uniqueness he claimed only the original material form of an artwork could possess.[10] In respect to traditional art media such as painting and sculpture, this indeed reflects a certain logic since the representational quality of a physical art object is very much tied to its presence in time and space. But in regards to photography, and particularly cinema, where the reproduction is at the same time also its exhibitive purpose, the quality of the work is not necessarily diminished, but rather only transferred. This becomes much more apparent when considering purely digital artworks on the internet today, where there is no longer a clear conceptual distinction between original and reproduction.[11] Many artists, especially of the

---

[8] Baudelaire, C. *The Mirror of Art, Critical Studies.* Translated by Mayne, J. Doubleday Anchor Books. 1956. pp. 231-232.
[9] Baudelaire, C. *The Painter of Modern Life.* Translated by Charvet, P. E. Penguin UK. 1972.
[10] Benjamin, W. *The Work of Art in the Age of Mechanical Reproduction.* Translated by Zohn, H. Schocken Books. 1969.
[11] David, D. *The Work of Art in the Age of Digital Reproduction, An Evolving Thesis: 1991-1995.* MIT Press. 1995.

pictorialist and avant-garde movements across Europe, would turn to embrace the reproductive image technology of photography for its experimental value and innovative means of production. Another notable example is the movement of Italian Futurism which radically rejected past traditions and advocated for a new dynamic relationship towards technology. Cultural attitudes such as this helped usher a wider recognition of the role of new media, such as animation and cinema, into the realm of art.

As technology progressed into the mid 20th century it would continue to profoundly change every facet of modernized society. Notable advancements such as the internal combustion engine and semiconductor materials used in electro-magnetic transmitters and receivers such as the radio further accelerated society forward in time. However after two world wars, and especially the detonation of two atomic bombs by the United States over Japan in 1945, it became clear that technology could also present an existential threat to human survival. Deeply rooted sentiments of distrust towards technology did not subside until predictable trends would begin to emerge. Much of public attitudes were reshaped by the availability of consumer products. Convenience, but also the rise of a flourishing middle class and new economic opportunities, helped justify the narrative that technology was there to primarily improve the quality of peoples' lives.[12]

It wasn't until the eventual dawn of digital computing that this notion of predictable progress manifested itself into a physically measurable form. Transistors and the integrated circuit substantially lowered the cost of electronics production while also simultaneously guaranteeing an exponential rise in computational power *(Figure 3).*[13] Through manufacturing improvements, the number of transistors on a circuit could predictably double about every two years. This would become known as Moore's law, named after Gordon Moore who observed this trend in 1965. For artists working with new media, this meant that a consistent adaptation of their practice was required, as the developing landscape of computation was now fitted to this new technological determinism.

[12] Halliwell, M. *American Culture in the 1950s.* Edinburgh University Press. 2007. pp. 4-12.
[13] Scace, R. I. "Electronics, The History of Electronics, Integrated Circuits." *Encyclopedia Britannica.* https://britannica.com/technology/electronics/The-semiconductor-revolution#ref233776. Accessed April 4, 2021.

Figure 3: Jack Kilby's original hybrid integrated circuit from 1958. This was the first integrated circuit, and was made from germanium.

Early forms of digital art emerged mainly out of scientific research in the field of cybernetics during the 1960s. Much of the required equipment at the time was not yet accessible to the general public, with most of the work being made at research labs in the form of scientific and mathematical visualizations. A. Michael Noll, who worked at Bell Telephone Laboratories, was an influential pioneer of the beginnings of digital art. He used computers to create artistic patterns and helped formalize the use of algorithmic processes in the creation of digital arts *(Figure 4)*.[14]

[14] Noll, A. M. "Computers and the Visual Arts." *Design Quarterly,* no. 66/67. 1966. pp. 64-71.

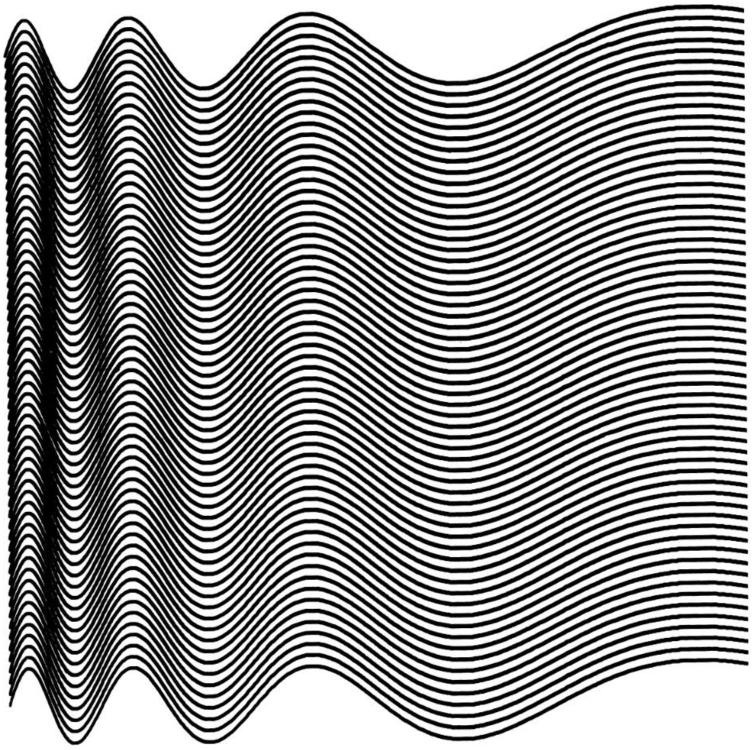

Figure 4: "Ninety Parallel Sinusoids With Linearly Increasing Period" by A. Michael Noll (1967). This computer drawing is based on a painting by Bridget Riley (1964).

Computer scientists such as Noll, who were one of the first to experiment with the digital algorithmic image, eventually exhibited their works of computer art in a gallery which opened the door for computer assisted visual art to enter the art world. "Cybernetic Serendipity", which was held at the Institute of Contemporary Art in London in 1968, was an international show that included many of the first digital artists such as Nam June Paik, John Whitney, and Charles Csuri. It was the first of its kind to open up discourse into the relationship between modern technology and art. The show made a strong public impression and thus became an influential moment in the history of computer art and new media.[15] Much of this development in the aesthetics of computer art can also be attributed to the new modes of expression found in conceptualism, which questioned the function and utility of traditional art. Artistic trends such as process art and especially minimalism turned out to be very compatible with this new medium as well. Additionally, comprehensive media theories such as those developed by theorist

[15] Usselmann, R. "The Dilemma of Media Art: Cybernetic Serendipity at the ICA London." *Leonardo,* vol. 36(5). 2003. pp. 389-396.

Marshall McLuhan influenced the awareness of how new forms of media control can shape human associations and actions.[16]

A significant step in computing began in the 1980s once personal computers powered by microprocessors had become widely available to the consumer market. These devices were no longer mainframe computers that required operation by a technical team, but were designed for individual interactive use. Companies such as Apple and Commodore were some of the first to appeal to the creative vocations of the public with multimedia applications consisting of accessible graphical-user interfaces (GUI).[17] In 1990, Photoshop, the first commercial computer graphics editor was released by Adobe, which soon became the standard tool for image and photo editing. In the meantime, a decentralized media landscape known as the internet began to establish itself around the globe, changing the fundamental ways in which people could communicate and share media. This caused power to shift away from traditional media channels such as television and newspapers and into the hands of online enterprises. By the mid 2000s, companies such as Facebook, Google, and Twitter gradually began seizing dominance over internet communications in the form of social media. While these centralized platforms provided users with unprecedented access to public discourse and novel opportunities to share ideas freely, these corporations would in turn also work to maximize their advertising revenue by harvesting private user data and manipulating site algorithms so as to increase user engagement.[18] For artists, these new digital landscapes still provide an effective means for creative collaboration and self-promotion, allowing them to directly engage with their audiences. The internet also layed a new groundwork for how art and media are consumed, further disrupting creative traditions and established cultural norms.

---

[16] McLuhan, M. *Understanding Media: The Extensions of Man.* Signet Books. 1964.
[17] Myers, B. A. “A Brief History of Human-Computer Interaction Technology.” *Interactions,* vol. 5(2). 1998. pp. 44-54.
[18] McDavid, J. “The Social Dilemma.” *Journal of Religion and Film,* vol. 24(1), article 22. 2020.

Figure 5: An eight-rack pod of Google's liquid-cooled TPU version 3 servers for artificial intelligence workloads. (Image: Google).

The field of machine learning is the most recent development in this line of technological history. While its ideas and theoretical applications aren't particularly new, and can be traced back to network theories conceived in the previous century, it wasn't until the maturation of the internet in the 2010s when data availability converged with the computational processing power necessary to be able to practically implement them on a large scale *(Figure 5)*. As of this writing, artificial intelligence plays an impactful yet often invisible role in the functioning of society. Machine learning technologies are currently being heavily utilized in virtually all major industries today, including financial services, government, health care, retail, entertainment, communications, energy, transportation, and many others. However, despite its practical and commercial success, AI still remains a young field with many underexplored research opportunities.[19] Given its current trajectory, it is reasonable to suggest that machine learning is likely to be one of the most transformative technologies of the 21st century, although exactly how it will manifest itself is very difficult to predict. It is likely that this will depend on how society is able to maximize its benefits while preferably reconciling it with sound ethical values.

[19] Jordan, M. I. Mitchell, T. M. "Machine Learning: Trends, Perspectives, and Prospects." *Science,* vol. 349(6245). 2005. pp. 255-260.

In summary, technology has been a substantial source of inspiration for artists throughout time. The angst that accompanies it is often associated with the uncertainty of an ever-changing techno-cultural landscape. For many the invention of new tools has provided innovative means to communicate one's ideas, emotions, and personal attitudes through creative expression, but often at the expense of authenticity. The digital age has propelled an already accelerating evolutionary momentum into a fixed technological determinism which forces artists to continuously adapt and depend on new systems of media production. Based on this, it is feasible to state that the rise of future media technologies will continue to disrupt traditional conventions, while also offering access to new creative possibilities. But this will always be accompanied with unpredictable cultural consequences and challenges.

2.2. Technical Overview

Since the focus of this thesis is on the creative applications of machine learning in photographic and artistic practices, I do not intend to dive too deep into the technicalities and mathematics of its underlying operations. However, due to the very technical nature of this medium one cannot neglect this aspect entirely. Therefore, I will proceed to introduce the subject of machine learning and highlight some of its main principles. This allows us to build a general foundation of understanding that will later be useful in our interpretations of the creative uses of ML models, more specifically in the visual context of generative imaging.

Basically, machine learning is defined as a field of study in computer science which aims to improve the performance of a machine or algorithm on a certain task through experience.[20] The term itself was originally coined by researcher Arthur Samuel back in 1959. It has many relationships to other disciplines such as mathematics, statistics, and software engineering. More recently, scientists have made a number of significant advances particularly in artificial intelligence, which is an emerging computer technology that deals with the question of how to apply data and general knowledge in order to replicate intelligent behavior in machines. These two fields are very similar, but differ slightly in the way that they intrinsically function. Whereas ML is generally programmed to learn and predict outcomes based on passive observations such as probabilistic analysis, networked AI plays a more active role in interpreting the data, learning to take actions and interact with its environment in a way that helps it to maximize the success of a specific task.[21] Because the terms machine learning and artificial intelligence are so closely connected, they are often used synonymously. As a rule of thumb, AI is commonly regarded as an umbrella term for machine learning technologies which exhibit a more generalized form of intelligence, one that is fitted for solving a task in a manner that could be expected from humans. There are numerous types of machine learning processes, but for the sake of practicality they can be framed within three major recognized categories: supervised learning, unsupervised learning, and reinforcement learning. It is to

[20] Mitchell, T. *Machine Learning.* McGraw Hill. 1997.
[21] Grewal, D. S. “A Critical Conceptual Analysis of Definitions of Artificial Intelligence as Applicable to Computer Engineering.” *Journal of Computer Engineering,* vol. 16(2). 2014. pp. 9-13

note that these processes are not always mutually exclusive. Oftentimes ML models can incorporate more than one learning paradigm depending on how a particular problem is phrased.

Supervised learning is perhaps the most common form of machine learning. Essentially, it uses labelled data to map approximations onto input variables. It can be likened to teaching a child with flash cards. Example data is paired with a label and fed into the algorithm, allowing it to learn the classification for each input. Feedback is given based on whether the predicted answer is right or wrong. The algorithm over time learns to approximate the relationship between the examples and their labels. This enables a well-trained model to make accurate class predictions from new inputs it has not been trained on. A typical example of a machine learning application that is built on supervised learning techniques is spam detection in email systems.[22] Facial recognition, which determines the identity of people in images, as in surveillance video or on social media, is another example of supervised ML.

When a learning process does not feature any labels, and a model is instead being fed a huge amount of data for organizing, it is regarded as unsupervised learning. What is intriguing about this particular method is that here an algorithm is able to discover hidden patterns and groupings in data that a human couldn't otherwise detect. This is an especially important aspect given the fact that the bulk of existing data in the world is not labelled. With a model capable of taking enormous amounts of unlabelled data and making sense out of it, it not only becomes a valuable tool for research but it also garners appeal from many commercial industries. Buying habits and other consumer histories collected in databases are perpetually being entered into unsupervised ML systems to help companies identify clustered segments of market behavior. Recommendation systems such as those seen on e-commerce or video streaming sites are the more visible examples of this type of application.[23]

[22] Christina, V. et al. "Email Spam Filtering Using Supervised Machine Learning Techniques." *International Journal on Computer Science and Engineering,* vol. 2(9). 2010. pp. 3126-3129.
[23] Isinkaye, F. O. et al. "Recommendation Systems: Principles, Methods, and Evaluation." *Egyptian Informatics Journal,* vol. 16(3). 2015. pp. 261-273.

Reinforcement learning is somewhat different from the previous two types in which one can readily correlate their functions with the existence of labelled data. Reinforcement learning instead operates more heavily on behavioral science. A fitting analogy would be to imagine a mouse in a maze that needs to learn how to effectively navigate an unknown environment in order to reach a reward such as a piece of cheese. Here there are two main elements at play: an agent and environment which are both connected by a feedback loop of signals. The aim of reinforcement learning is to allow the algorithm to make a set of actions and learn from resulting mistakes within an environment that presents the goal of obtaining a certain reward. The conditions of the environment continuously reinforce the agent to maximize its performance in order to reach its goal. Applications can be as mundane as teaching an AI to play a video game, to more complex functions such as robotics training, industrial simulation, or resource management. A notable example would be the incentivization of networked machines to use less electricity by balancing power needs as efficiently as possible.[24]

Artificial Neural Networks (ANN) are a novel development in AI research which is inspired by the way that biological neural systems in the brain process information *(Figure 6).* The idea can be traced back to 1943, when neuroscientists Warren McCulloch and Water Pitts proposed a theory which attempted to demonstrate how logic and computation could be used to understand neural activity and mental functions.[25] The tendency to anthropomorphize ANNs is tempting, and it's important to stress that ANNs differ from biological neural networks in a number of ways.[26] Nonetheless, mimicking complex neurological processes through logical systems has proven to be a very effective mechanism for developing solutions to complex mathematical problems, ones which require computational capacities that go beyond conventional rule-based systems or statistical methods. Prime examples where ANNs are frequently used are in computer vision and speech recognition.

---

[24] Chemingui, Y. et al. "Reinforcement Learning-based School Energy Management System." *Energies,* vol. 13(23), 6354. 2020.
[25] McCulloch, W. Pitts, W. "A Logical Calculus of Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics,* vol. 5(4). 1943. pp. 115-133.
[26] Nagyfi, R. "The Differences Between Artificial and Biological Neural Networks." *Medium.* https://towardsdatascience.com/the-differences-between-artificial-and-biological-neural-networks-a8b46db828b7. Accessed April 12, 2021.

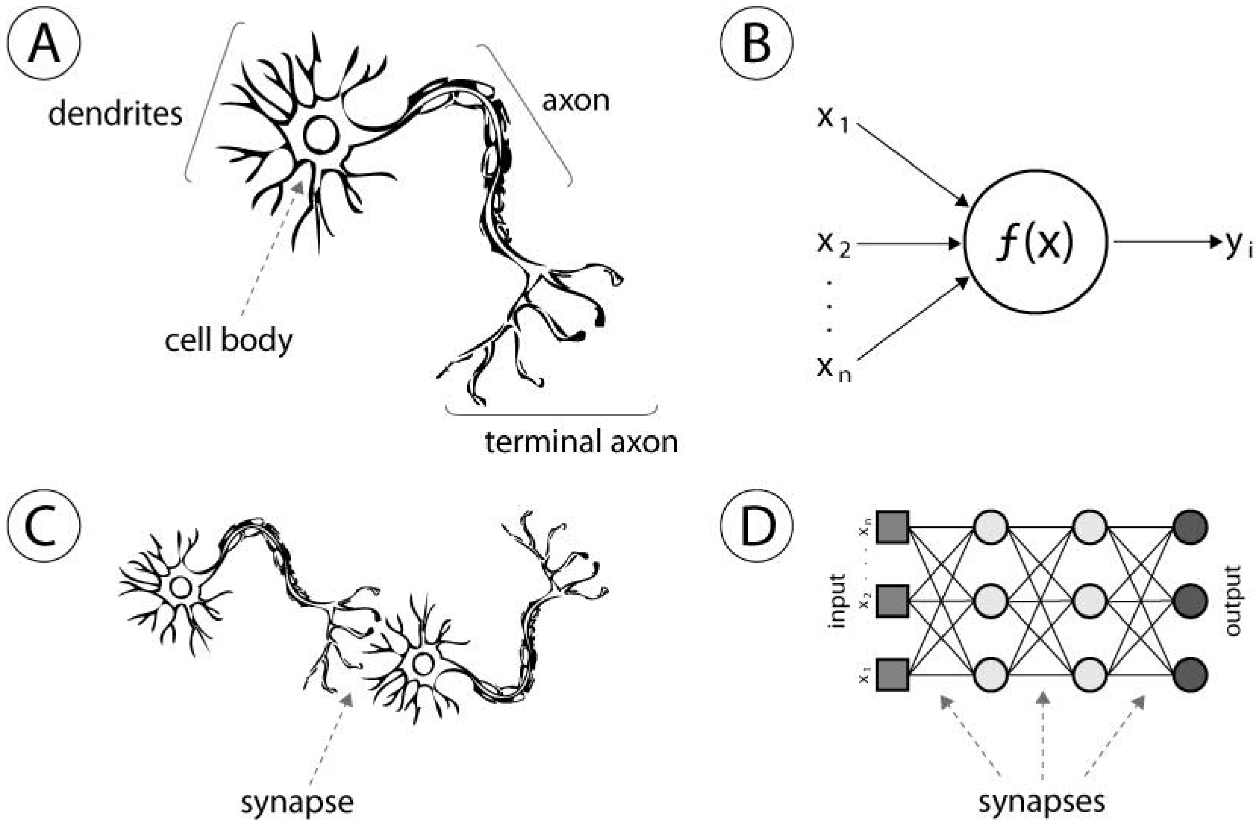

Figure 6: Visual analogy of biological neurons and artificial neural networks with diagram A and C representing an individual neuron and synapses in the human brain, and B and D showing their logical equivalent in machine learning.

ANNs, similar to other forms of ML, are designed to continuously improve on a given task over time. What sets an ANN apart is its interconnected structure of hidden layers and ability to backpropagate calculations through the network *(Figure 7).* It does this by adjusting the weights of input variables. For instance, when an input is given and the network makes a random prediction based on its previous experience, it can then backpropagate and adjust parameters based on how close the output was to the prediction. It uses the sum of weights of all inputs from the previous layer in order to pass it onto the next. In effect, this process goes on until the network converges to a point where no further training improves the output value of the model. After a model has been trained, it can then use new input variables and propagate them forward, using its learned experience to make predictions for entirely new inputs.[27]

[27] Dongare, A. D. et al. "Introduction to Artificial Neural Network." *International Journal of Engineering and Innovative Technology,* vol. 2(1). 2012. pp. 189-194.

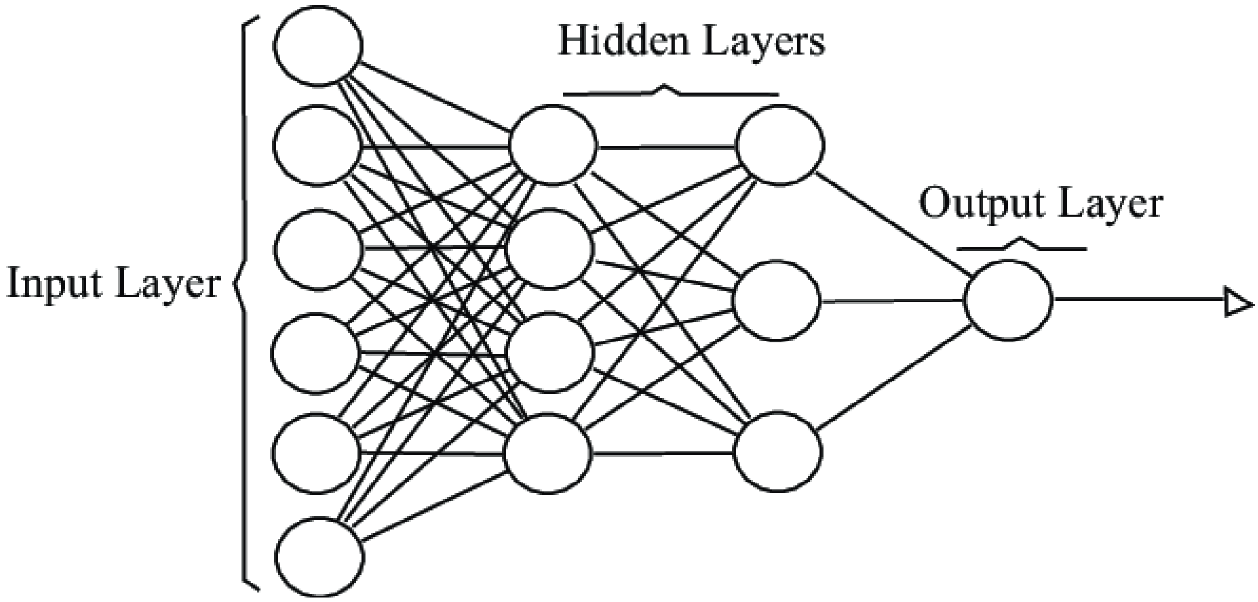

Figure 7: A typical structure of an artificial neural network.

Generally speaking, the more complex the task, the deeper the hidden layer structure needs to be. Deep learning (DL) is a term more recently used to describe self-optimising neural networks with architectures consisting of many more hidden layers than typical ANNs. Advancements in computer vision, image, video, and natural language processing are attributed to recent improvements in DL. There are a number of different types of networks, but one of the most relevant architectures being used are Convolutional Neural Networks (CNN). Due to their ability to recognize patterns in visual data, they are commonly used in image and facial recognition, video analysis, and other areas which require intricate object classification and detection methods.[28] Early versions of CNNs were used back in the 1990s for recognizing handwritten digits on bank checks or reading zip code numbers on mail,[29] but their development and scaling potential quickly bottlenecked due to the limitations of central processing units (CPU). Much of AI research in fact stalled during that time, and wouldn't find the means to revive itself until the early 2010s when the availability of large data sets combined with highly-powered graphical processing units (GPU) brought

[28] O'Shea, K, Nash, R. "An Introduction to Convolutional Neural Networks." *arXiv preprint arXiv:1511.08458*. 2015.
[29] Denker, J. S. et al. "Neural Network Recognizer for Hand-written Zip Code Digits." *Advances in Neural Information Processing Systems*. 1989. pp. 323-331.

significant performance breakthroughs to the field of deep learning, and thus marked a turning point in the history of machine learning as a whole.[30]

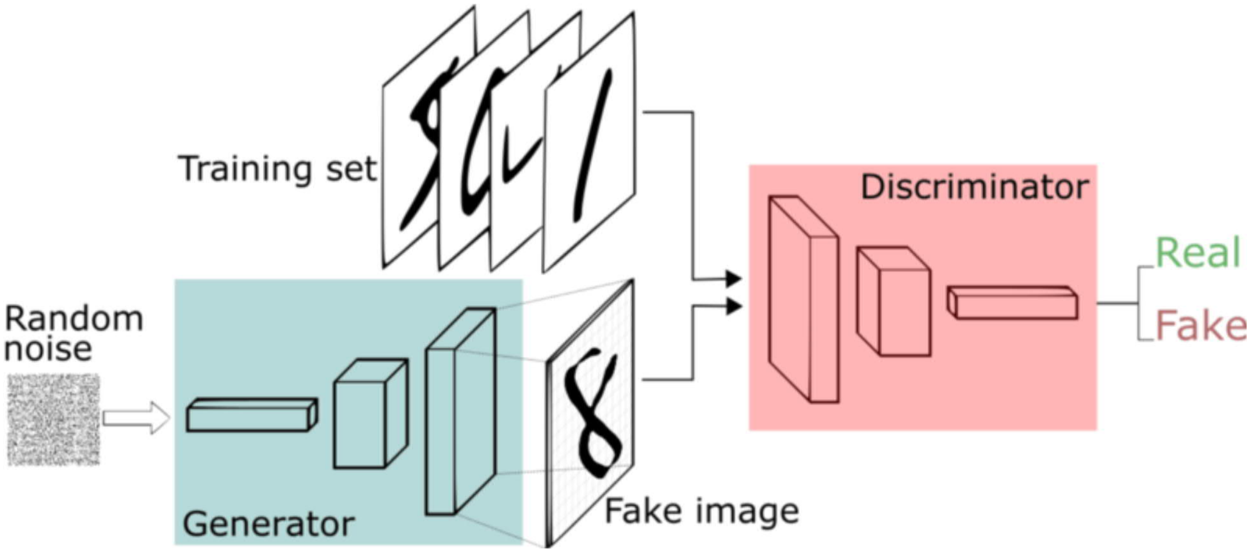

Figure 8: The basic framework of a Generative Adversarial Network.

A few years later Generative Adversarial Networks (GAN) arrived on the scene, first proposed in 2014 by Ian Goodfellow and researchers at the University of Montréal.[31] A GAN is an unsupervised machine learning algorithm that is made up of a combination of two competing neural networks, a generator (G) and discriminator (D). One can think of a GAN as a game of cat and mouse, where a counterfeiter G attempts to pass on fake images, and a judge D tries to determine the authenticity of them. This kind of dynamic results in a zero-sum game, where the maximization of the loss function in both networks leads them to continuously learn how to outsmart each other.[32] GANs are notably successful for their ability to generate photorealistic images which can often be indistinguishable from real images. This type of network architecture lies at the forefront of current research efforts to generate high-fidelity, diverse images with models that progressively learn from data.[33] There are a number of different ways of applying GANs, some of which I will cover in the next chapter.

[30] Alom, M. Z. et al. “The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches.” *arXiv preprint arXiv:1803.01164.* 2018.
[31] Goodfellow, I. J. et al. “Generative Adversarial Networks.” *arXiv preprint arXiv:1406.2661.* 2014.
[32] Creswell, A. et al. “Generative Adversarial Networks: An Overview.” *IEEE Signal Processing Magazine,* vol. 35(1). 2018.
[33] Brock, A. et al. “Large Scale GAN Training for High Fidelity Natural Image Synthesis.” *arXiv preprint arXiv:1809.11096.* 2018.

To conclude, the technical capacities of machine learning, particularly deep learning, have progressively advanced within the last decade thanks to improvements in algorithms, the availability of data, and significant performance increases in computer hardware. Research in the field is garnering attention not only for its scientific and commercial applications, but also for its creative potential using novel computer vision models and architectures. With this, it is today also becoming easier for creative practitioners to gain access to the forefront in this area of research, which has been primarily facilitated by the affordability of faster computers, access to cloud computing, and open sources of code and information.

3. Creative Applications

The different ways in which one can creatively use generative models of GANs is elegantly illustrated in the work of Lenka Hamosova, whose workshops explore the tools of synthetic media and possible futures of AI. In her 2020 project "Collective Vision of Synthetic Reality", she presents a set of educational cards which represent and briefly detail the most influential deep learning models *(Figure 9)*. I've chosen to reference this project for the reason that it provides an approachable way to get acquainted with specific use cases of ML. The card deck is divided into seven categories, in which she defines the most accessible AI models for image, video, and sound synthesis, text-based models, models for image recognition, post-processing, and models that are useful in media workflow.[34]

Figure 9: Cards designed by Lenka Hamosova which represent the currently most available AI models for generating synthetic media. (Image: Hamosova).

[34] Hamosova, L. "Collective Vision of Synthetic Reality." *Medium.* 2020. https://medium.com/@lenkahamosova/collective-vision-of-synthetic-reality-4ed02c01aaef . Accessed April 16, 2021.

Due to the wide scope of different deep learning methods, one will naturally gravitate towards one or a few models over others. In my own artistic practice and for the purpose of this thesis, I have found image synthesis and visual post-processing tasks to be the most relevant areas of study for the photographic medium. Therefore I will present some examples of models which fall into these two categories, and explain how they function and what they can be utilized for.

3.1. Image Synthesis

Image synthesis here is defined as the creation of new photographs and other digital images through the generative processes of an artificial neural network. The first step to building any GAN model is to collect enough numbers of training data to teach the neural network. The training process depends on specific augmentations and hyperparameters in order to achieve good results, but what is always an essential factor in determining the efficacy of a model's output is the size and quality of the image data set.[35] Depending on the scale of the model one wants to build, this can be achieved from the ground up with time and practice. Alternatively, it is possible to train new data sets on top of already existing models which can save one a lot of time.[36] This method, called transfer learning, is very often used in constructing new GANs, and there are some well-developed models, which we are able to build upon, that have already been pre-trained on very large datasets of images.

One such model is StyleGAN, which was developed and presented by researchers of Nvidia Corporation in late 2018. It is most commonly known for its impressive results in generating an unlimited number of portraits of fake human faces, oftentimes convincingly *(Figure 10)*. Since early 2020, it has evolved into an updated version called StyleGAN2 which has shown to improve image quality over its predecessor. It was trained using over 70,000 images of portraits of people on the image sharing site Flickr, called the Flickr-Faces-HQ data set.[37]

[35] Shamsolmoali, P. et al. "Image Synthesis with Adversarial Networks: A Comprehensive Survey and Case Studies." *Information Fusion.* 2021.
[36] Zhao, M. Cong, Y. Carin, L. "On Leveraging Pretrained GANs for Generation with Limited Data." *International Conference on Machine Learning*, PMLR. 2020.
[37] Karras, T, Laine, S, Alia, T. "A Style-Based Generator Architecture for Generative Adversarial Networks." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.

Figure 10: Random images generated with StyleGAN2.
https://thispersondoesnotexist.com/

One typically hears about synthetic images and media in the context of their potential misuses. Facial recognition, image data scraping, and deep fakes, in which a person in an existing image or video is replaced by someone else's likeness, are some main examples of this, and these can certainly be a cause for concern.[38] What I would like to emphasize here though is that StyleGAN2, which was trained primarily to generate human faces, is a tool that can be remodeled to serve constructive creative visions. The fact that this neural network's training is so comprehensive means that it already has the built-in ability to distinguish basic visual features such as lines, shapes, edges, colors, gradients, and textures. This capacity proves useful in image synthesis because these features will always be a prerequisite for producing complex visual outputs. This is something which can therefore be positively regarded as a technical advantage for thoughtful, creative ML work.

[38] Citron, D. K. Chesney, R. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *Calif. L. Rev. 107.* 2019.

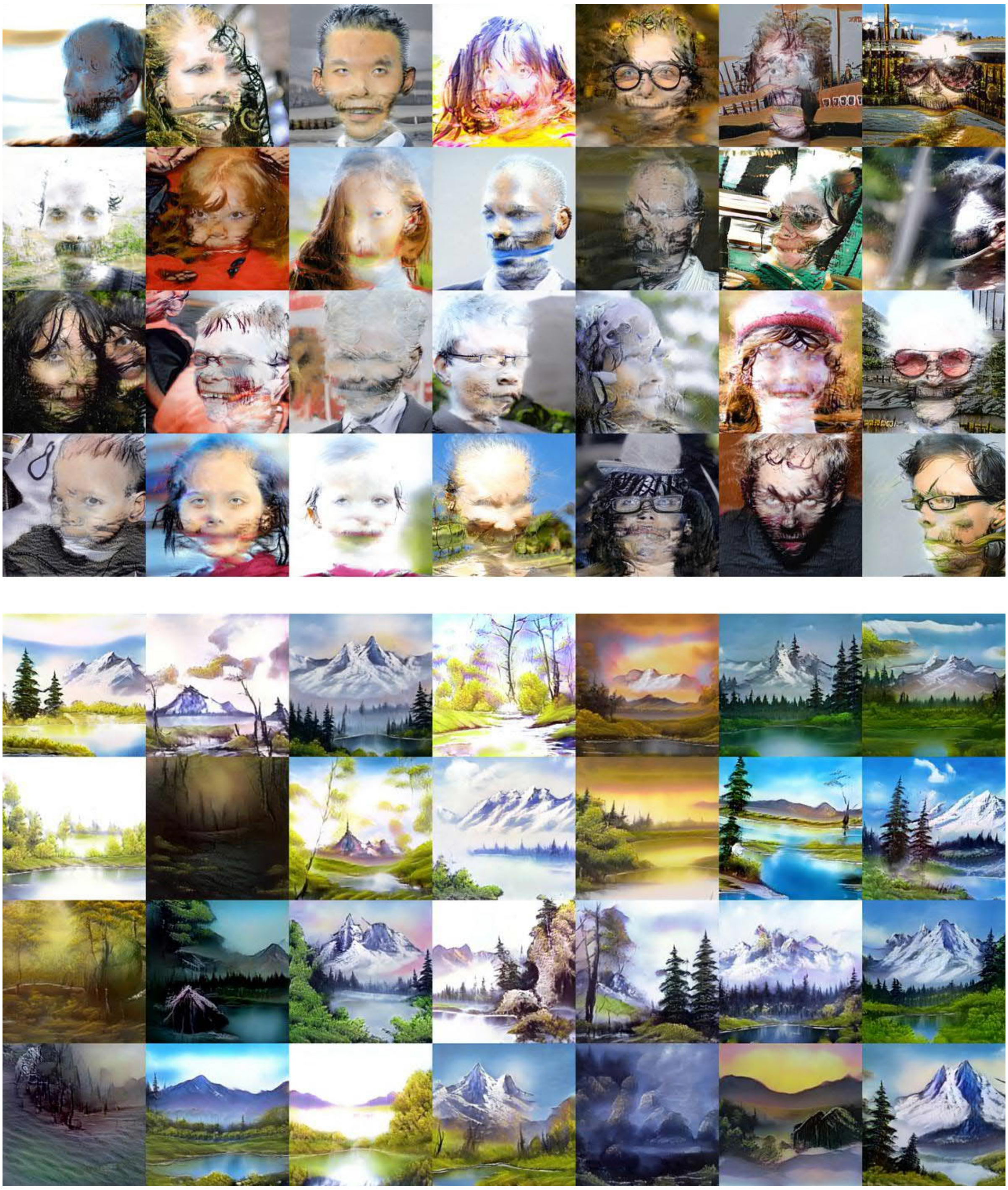

Figure 11: “Bot Ross” (2021). StyleGAN2 trained on a custom Bob Ross data set. Initial training checkpoint after 500 steps (above) and at completion of training after 5000 steps (below).

To provide an example, during the course of my research, I trained a model on StyleGAN2 using an image data set of paintings by Bob Ross, the American painter, art instructor, and television host of his series “The Joy of Painting”, which aired from 1983 to 1994 *(Figure 11)*. I manually compiled 250 individual images of his paintings and mirrored them horizontally to create a combined total of 500 images for my data set.[39] Since a GAN learns from every new area of pixel information, mirroring is an effective method to increase one's training data. However one must keep in mind that the neural network will learn from every data point it is given, therefore it is important to consider the logic behind manipulating the inputs. In my specific case, vertical mirroring would've caused the predominant landscape features of sky, mountains, and land to invert, which would severely affect the output. Mirroring horizontally however, differentiated the images just enough so that they did not fundamentally change the appearance of a natural landscape. This experiment illustrates just how well a neural network is able to continue incorporating new input variables from any given point of its training progress, even if the outputs of the previous model don't visually match the new objective. Due to its advanced pre-training, StyleGAN2 is considered to be one of the most suited architectures to use for training new GANs that can efficiently produce high-quality images.

Another pre-trained model for image synthesis worthy of mention is BigGAN. What BigGAN does differently from StyleGAN is that it pools together a wide variety of image classifications and scales up the batch size and number of model parameters resulting in generations of novel high-fidelity images (Figure 12).[40] The data set used to train this network is sourced from ImageNet, first published in 2009. It is a large visual database designed to provide researchers with image data for training large-scale object recognition models. ImageNet's data set consists of over 14 million manually annotated images organized into a hierarchy of over 20,000 categories called synsets, with typical categories such as "indian elephant", "reflex camera", or "grocery store" containing a number of images which can all be machine-read.[41]

---

[39] Auer, F. "Bob Ross Paintings." *TwoInchBrush.* 2019. https://www.twoinchbrush.com/all-paintings. Accessed February, 4, 2021.
[40] Brock, A. et al. "Large Scale GAN Training for High Fidelity Natural Image Synthesis." *arXiv preprint arXiv:1809.11096.* 2018.
[41] Deng, J. et al. "ImageNet: A Large-scale Hierarchical Image Database." *2009 IEEE Conference on Computer Vision and Pattern Recognition.* 2009.

Figure 12: Class-conditional image samples generated by BigGAN.

Another notable aspect of BigGAN, besides its scalability, is a kind of regularizing technique called 'truncation' that can be applied to the generator. Truncating allows one to finely control the balance between the fidelity and variety of images that the network produces. This is best visualized by thinking of all the possible images the network can potentially generate not as pre-rendered images but rather as mathematical vectors that exist as coordinates in a multi-dimensional latent space. Every definable vector within this space is a condensed data point that is able to generate one specific image when decoded. The further away two points in this space are, the more different the resulting two images will be from one another. Truncation here refers to the shortening of distance between latent image points, with shorter distances producing images with less variety, but in return more consistency and higher fidelity.[42]

At this point, I deem it necessary to also mention some key technical limitations in regards to these generative models. Even though they are capable of visually producing high-quality representations of data, there are a few output conditions which limit the scope of certain implementations. First would be concerning the aspect ratio. Common GAN models output only in square formats. This is because there is a default requirement for uniform dimensions for generating

[42] Zhang, M. "BigGAN: A New State of the Art in Image Synthesis." *Medium.* 2018. https://medium.com/syncedreview/biggan-a-new-state-of-the-art-in-image-synthesis-cf2ec5694024. Accessed April 21, 2021.

consistent images that can be repeatedly downscaled by a factor of two. If aspect ratios in the training data were to vary from each other, it would in effect deform the spatial processing of information through a manifold of varying data points. While it is technically possible to engineer GANs which allow diverse input sizes, it is still considered to be a proof-of-concept which only shows that it is feasible, yet still requires more research to improve the quality of images.[43]

Figure 13: Increasingly realistic synthetic faces generated by variations on GANs. In order, the images are from papers by Goodfellow et al. (2014), Radford et al. (2015), Liu and Tuzel (2016), and Karras et al. (2018).

Secondly, there are limitations in regards to output resolution size. Nvidia's progressively growing GAN architecture published in 2018 proposed an initial scaling of images which could add up in layers to generate image resolutions as high as 1024x1024. There are however some problems when attempting to go higher than this. One issue lies in how easily the discriminator network of a GAN is able to distinguish real and fake images when the random output of the generator network is too large, causing inhibitions in the training process. Another hurdle relates to raw computational costs, including the larger GPU memory requirements needed for larger outputs. To contrast these difficulties, researchers of GAN architectures have removed some of these obstacles in favor of shorter training times and increased model stability.[44] In the near future, one can reasonably expect resolutions to increase, as is evidenced by recent developments. Whereas generative models can now produce synthetic images

[43] Kendrick, C. et al. "Anysize GAN: A Solution to the Image-warping Problem." *arXiv preprint arXiv:2003.03233.* 2020.
[44] Karras, T. et al. "Progressive Growing of GANs for Improved Quality, Stability, and Variation." *arXiv preprint arXiv:1710.10196.* 2017.

that are close to indistinguishable from photographs today, only a few years ago the images they produced were crude and of much lower output resolution (Figure 13).[45]

[45] Brundage, M. et al. “The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation.” *arXiv preprint arXiv:1802.07228.* 2018.

3.2. Post-Processing

Image editing has long been aided by algorithms built into computer graphics software. Since the revival of AI technologies in the past decade, we can see these algorithms gradually being enhanced by machine learning innovations. Usual image post-processing tasks such as sharpening, denoising, enlarging, and masking are steadily being improved by ML, benefiting sectors of e-commerce, marketing, media production, individuals, designers, photographers, and visual artists alike.[46] Because there are many generative models I could discuss that can be used to post-process visual media via machine learning, I will only mention a couple of models here which I have experimented with in my own practice. Later in this chapter, I introduce the tools that can be utilized for applying models, and which also include access to a wider scope of AI-assisted post-processing tasks for varying use cases.

Resampling digital images is a routine practice in image editing because outputs can serve different end uses, such as printing or web publishing. Typically, when one is exporting an image to be uploaded to the internet a smaller image resolution is preferred; whereas in printing, saving to a larger file size in a higher resolution yields better results. Scaling a raster image, such as a digital photograph, uses a process called interpolation which distorts the image from one pixel grid to another. By using approximations of a pixel's intensity based on the values of neighboring pixels, the simplest mode of interpolation is able to use this known data to estimate pixels at unknown points. However, resampling an image to a size that is larger than its original resolution causes several issues.

Because the interpolation algorithm is attempting to create new pixels where previously none existed, data degradation manifests in the form of images visibly appearing softer, fuzzy, pixelated, with less overall sharpness and contrast.[47] In creative professions, upsampling images using interpolation is commonly frowned upon due to this resulting loss of image quality. Instead, it is recommended during prepress to resize the output to a smaller DPI value which

[46] Amato, G. et al. "AI in the Media and Creative Industries." *arXiv preprint arXiv:1905.04175*. 2019.
[47] Acharya, T. Tsai, PS. "Computational Foundations of Image Interpolation Algorithms." *Ubiquity,* 2007 (October), article 4. 2007.

doesn’t alter the image resolution, but still adjusts the file accordingly to make a larger print.

Figure 14: Zoomed crop of original compressed image (top left), enlarged 2x using Bicubic Smoother in Photoshop (top right), Preserve Details 2.0 in Photoshop (bottom left), and Super-resolution in Gigapixel AI by Topaz Labs.

With machine learning technology, new techniques have since been developed which can train deep learning networks for upscaling images. These are referred to as super-resolution algorithms (SR). The problem these models try to solve is how to increase resolution from compressed images while still maintaining structure and fine details. A variety of SR methods have been proposed, such as prediction-based, edge-based, statistical, patch-based, or sparse representation methods.[48] What SR models have in common, like other generative neural networks, is that they need to be trained on data sets. Some of these provide low-resolution and high-resolution pairs, which the network learns from, while others only provide high-resolution images. The results of using SR to resample digital images often appear visibly better than those of traditional interpolation algorithms *(Figure 14)*. There still remain a number of challenges to further improve the quality of SR methods, but it goes to show how DL networks are continuously providing novel image enhancing tools for creative workflows. And the industry is responding, with Adobe just recently beginning to incorporate their own SR algorithm into their 2021 releases of Camera Raw, Lightroom, and Photoshop.[49]

Another significant use of AI post-processing can be applied to the field of digital photograph restoration, which works with a digital copy to restore the original appearance of a physical photograph that has been damaged by natural, man-made, or environmental causes. Evidence of dirt, scratches, and chemical changes on a photograph can make this particular practice a laborious affair. AI-based photo restoration tools allow one to fix certain defects such as spots, scratches, restore missing parts of the image, and even add colors to grayscale photos, in a shorter amount of time than compared to manual retouching techniques *(Figure 15)*.

Figure 15: Image restoration results produced by "Bringing Old Photos Back to Life."

[48] Wang, Z. Hoi, S. C. H. "Deep Learning for Image Super-resolution: A Survey." *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2020.
[49] Chan, E. "From the ACR team: Super Resolution." *Adobe Blog,* Creativity, March 10, 2021. https://blog.adobe.com/en/publish/2021/03/10/from-the-acr-team-super-resolution.html. Accessed April 22, 2021.

“Bringing Old Photos Back to Life”, created by researchers from the University of Hong Kong in collaboration with Microsoft Research Asia, is a recent deep learning computer vision model developed for this particular purpose. The main issue is that degradation in vintage photos is rather complex and because there exists a domain gap between the collected data of synthetic images and real old photographs, conventional supervised learning methods aren’t able to effectively generalize for this specific task. To solve this problem, the researchers worked to close this gap by developing a type of domain translation network which operates on two variational autoencoders (VAE). Compared to GANs, VAEs are generative models which are particularly well-suited for adapting data to determine probabilities within the latent space of the network, rather than purely generating new data. Qualitative comparisons presented in the research paper demonstrate that their model restores both unstructured and structured degradation significantly better than other ML methods.[50]

To conclude this section, image synthesis and AI-assisted post-processing tasks are only some of the generative models available which are relevant to the field of image making. Significant improvements in image generating, enhancing, and editing capabilities are increasingly leading to the adoption of DL neural networks by creative industries. These innovations provide considerable benefits for creatives and visual artists who can access and utilize these networks to experiment with ideas, augment their workflows, and save time during post-production. Given the growth trajectory of ML technology in recent years, it is safe to say that artificial neural networks will continue to play new and integral roles in systems of image production, including photography.

[50] Wan. Z, et al. “Bringing Old Photos Back to Life.” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 2020

3.3. Creative Tools

A major hurdle of beginning to work with machine learning is that, despite its creative applications, it essentially is a field of computer and data science that still remains highly technical. Programming knowledge, particularly of the coding language Python, mathematical skills, as well access to powerful GPUs are necessary for engineering ML algorithms. These prerequisites often present themselves to creatives as a technical barrier, causing AI to appear as a domain only reserved for data scientists and programmers. However over the course of the past few years, this gap has been closing by means of an opening of dialogue between coders and creatives, enabling the sharing of technical knowledge through the internet. Furthermore, the organization and development of ML tools in the form of accessible web apps and GUIs has helped democratize access to machine learning, thus leading to growing optimism and enthusiasm about ML amongst creative communities. In this section, I will mention some of the AI-powered tools for designers, artists, and other creators that can be readily utilized.

Founded in 2018 by a New York-based startup, RunwayML has since positioned itself as a machine learning platform for creators, artists, designers, and filmmakers, providing libraries of pre-trained algorithms for diverse needs. It features models for image synthesis and analysis, stylization, text generation, green screens, video post-processing, and animation, including community created models. The tool runs under a polished and functional GUI similar to workspaces of other creative software programs, and operates as both a web-app and desktop application *(Figure 16)*. The distinguishing feature of RunwayML is that it requires no coding knowledge to use. This makes it a very accessible tool to users who have previously never worked with machine learning.

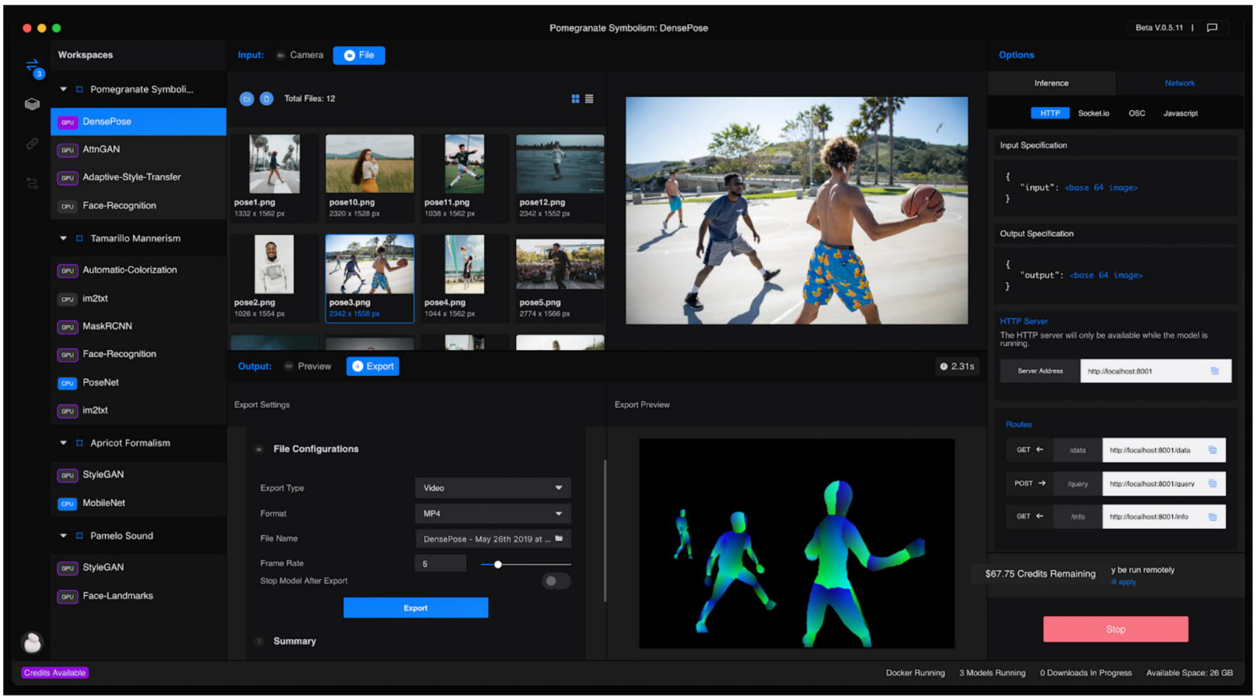

Figure 16: A screenshot of the interface of RunwayML.

All the image synthesis and post-processing methods I've explained in the previous sections are integrated in this software. Additionally, RunwayML also offers the means to train GANs using one's own custom data set. So for instance, if one wanted to train a neural network to generate artificial photographs of brutalist architecture, one could do this by collecting real images, uploading it to the platform as a data set, and then training the model on StyleGAN2.[51] The question of how large a data set needs to be in order to train a GAN that will produce quality outputs is an important one to ask. Generally speaking, the more data that is compiled the better; however the emphasis on quantity must also be coupled with the rationale behind image choice. This means that one needs to consider the style and content that is logical for the outcome intended to be produced, because GANs will only perform as well as the combined quality of the data with which it is fed with. In my experience, using a data set of a few hundred images already produces acceptable results but it is often recommended to use much more, in fact thousands, especially when combined with longer training times. In regards to setting up a training process

[51] Studio Kupol. "Brutalist Architecture Generator." *RunwayML.* 2019. https://app.runwayml.com/models/Kupol/Brutalism_Generator. Accessed April 23, 2021.

or running models, RunwayML offers an online learning resource with guides that are helpful to new users.[52]

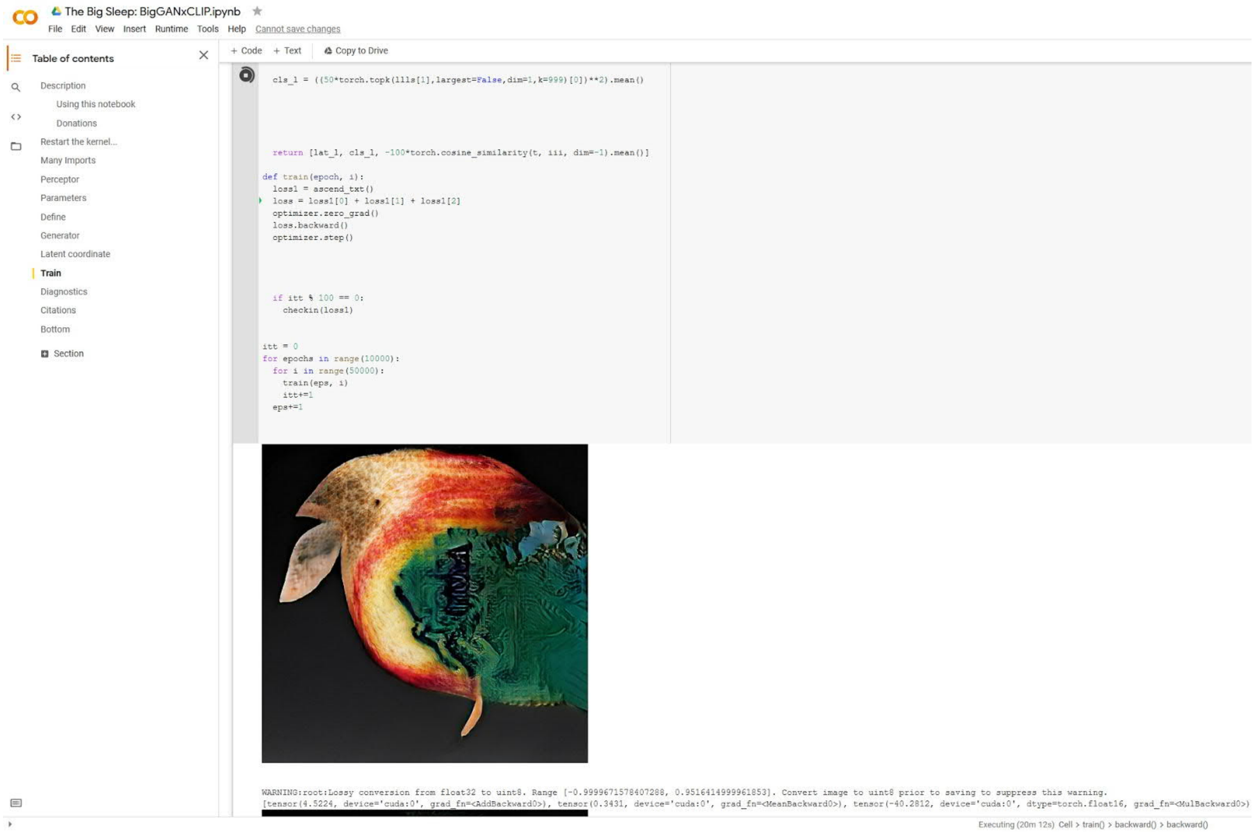

Figure 17: Screenshot of Google Colab running "The Big Sleep" notebook created by Ryan Murdock, running a text-to-image translation model.

While RunwayML provides a good starting point for beginners, there will eventually come a point when one has become acquainted with the limits of the application's potential. Derrick Schultz, artist and educator who regularly teaches online classes on how to creatively use generative machine learning, recommends Google Colaboratory (also known as Colab) as the next tool to adopt after getting comfortable with RunwayML.[53] Colab is a cloud service based on Jupyter Notebooks for interactive data science applications, providing a runtime that is fully configured for deep learning. Having less of a GUI than that of RunwayML, it consists mainly of an operational terminal interface that also serves as a code editor with note taking capabilities *(Figure 17)*. Functionally, it

[52] Valenzuela, C, et al. "Train an Image Generation Model". *RunwayML.* https://learn.runwayml.com/#/create/train-image-generation. Accessed April 23, 2021.
[53] Schultz, D. "What is Google Colab? (A Gentle Introduction/Explainer)." *YouTube,* uploaded by Artificial Images, February 24, 2021. https://youtu.be/b7s-NKmOEpQ. Accessed April 24, 2021.

works similar to a Google Docs object, which can be shared and also allows multiple users to collaborate on the same notebook.

What greatly sets Colab apart from RunwayML, in addition to the aforementioned characteristics, is that it can process different file types such as audio formats, offers the option of a free service, and provides access to powerful GPU hardware comparable to that of a mainstream workstation or Linux server equipped with 20 physical cores.[54] While the learning curve of Google Colab is certainly higher, it does offer creators with an intermediate skill level in ML a more sophisticated tool set, one that comes at a better price-performance ratio when compared to using RunwayML or renting a private GPU server. Lastly, I want to emphasize again the closing of the gap between segmented areas of technical knowledge enabled by information sharing, social media, and an increase of collaborations between data scientists and creative practitioners. Due to this growing relationship, there is today less of a requirement to be proficient in coding, and many of the functions necessary to operate Python-based machines can be searched on the internet or learned through engaging in dialogue with relevant online communities. As this trend continues, I predict it will only further lower entry barriers and help democratize access to ML for creatives and the general public alike.

Another application called Artbreeder, formerly known as Ganbreeder, is a collaborative art tool developed by Joel Simmons for discovering and synthesizing images. It is an open-source project and web app largely built on the creative contributions of its users. The ML engine used on this platform is BigGAN, which due to its image classification structure provides an intriguing functionality which progressively produces visual novelty. It does this by 'breeding' images, or more precisely by adjusting the ImageNet classifications via weights and then feeding them forward, thus allowing the progressive blending of different image categories into new images. A byproduct of this amalgamation of parameters is that the images take on a perceptual phenomenon known as 'visual indeterminacy.' Coined by artist and writer Robert Pepperell, it describes artwork that appears at first to be coherent and realistic, but defies consistent

[54] Carneiro, T. et al. "Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications." *IEEE Access,* vol. 6. 2018. pp. 61677-61685.

spatial interpretation.[55] Aaron Hertzmann of Adobe Research, furthermore takes on this subject in connection with machine learning, and describes how indeterminacy is a key feature of much modern representational art, but notably art created by GANs *(Figure 18)*. In his paper, Hertzmann suggests that indeterminacy is a consequence of powerful, but imperfect AI image synthesis models that are forced to combine general classes of objects, scenes, and textures without context, and are as a result unable to produce visual specificity through complex image-class combinations.[56]

Figure 18: Indeterminate images: (Left) The "impossible trident," (middle) "Succulus," 2015 by Robert Pepperell, oil on canvas painted specifically to be indeterminate, (right) a synthetic image generated with Artbreeder.

As we have seen, GAN models such as StyleGAN and BigGAN can produce realistic images, but these images are typically not presented as artwork. Most images produced by GANs however, which are shown in an artistic context, do exhibit this peculiar quality of indeterminacy and visual ambiguity. According to Hertzmann, there appears to exist an 'uncanny ridge' that finds its way into contemporary GAN imagery, causing much of the novelty to inadvertently turn into fatigue, as progressive user outputs continuously become more homogenous. In the following section, I will address this key issue with GAN art, and mention some artists who are attempting to deal with ML algorithms in ways which go beyond simple input-output mechanisms that are easily mimicable.

[55] Pepperell, R. "Seeing Without Objects: Visual Indeterminacy and Art." *Leonardo*, vol. 39(5). 2006. pp. 394-434.
[56] Hertzmann, A. "Visual Indeterminacy in GAN Art." *Leonardo*, vol. 53(4). 2020. pp. 424-428.

3.4. Artistic Uses

Machine learning has shown great potential for producing images with high fidelity, complexity, and abstraction. However, a potential pitfall of much of AI art is that due to its instructional mechanisms, outputs of similar aesthetic quality can be replicated by anybody with access to the same model and data set to execute them. ML for artistic purposes thus has the side effect of creating a large quantity of images that collectively become too consistent, homogenous, and effectively uninteresting. Though because current developments in ML are happening so quickly, a lot of art using GANs is new enough that the art world doesn’t understand it well enough to properly evaluate it. A prime example of this is the auction of “Portrait of Edmond Belamy” published by French art collective “Obvious”, which succeeded at selling their GAN artwork through Christie’s for $432,500 in 2018 *(Figure 19).*[57]

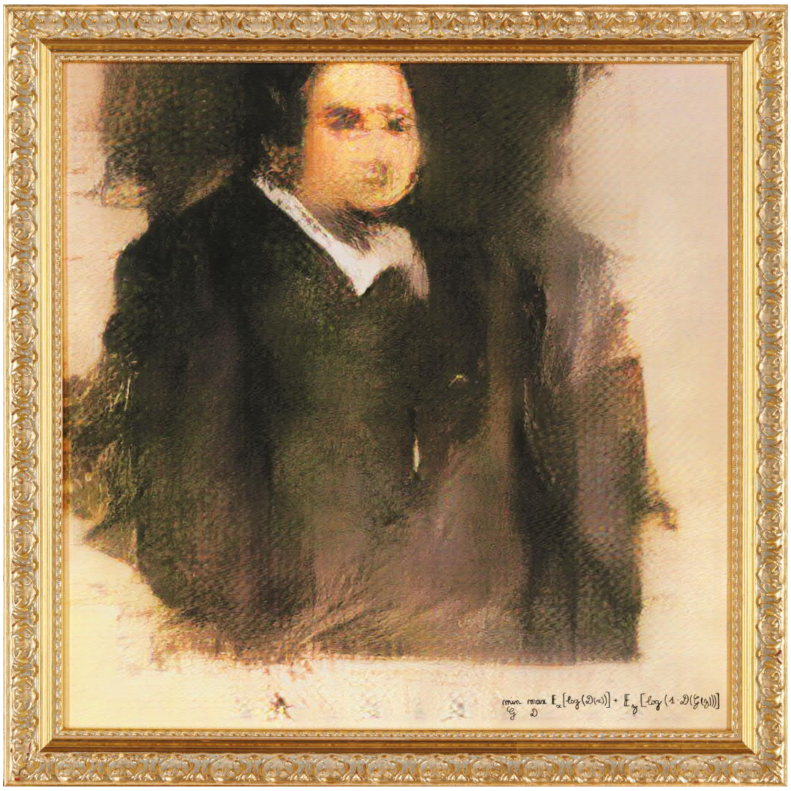

Figure 19: Edmond de Belamy, from La Famille de Belamy. Generative adversarial network print on canvas, 2018. Signed with GAN model loss function in ink by the publisher, from a

[57] Cohn, G. “AI Art at Christie’s Sells for $432,500.” *New York Times.* 2018. https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html. Accessed April 24, 2020.

series of eleven unique images, published by Obvious Art, Paris, with original gilded wood frame (70 x 70 cm).

At the time, many in the art community took issue with Christie's selection of Obvious, because it appeared that their piece represented a visual artifact that was solely created by the automatisms of a machine. This sentiment curiously reflects back on similar historical opinions I referenced earlier, which were held against newly emerging forms of technology, when critics such as Baudelaire claimed that photography inspired lazy and uncreative artworks. Important here would be to consider the fact that the print was hand-signed with a mathematical representation of the network's loss function, thereby alluding to Obvious' awareness of the ontological authorship of the work. At the same time, by taking into consideration the developments ML has made in recent years, one could certainly argue that this particular artwork is nothing remarkable by today's standards. A similar aesthetic could easily be reproduced in RunwayML, for instance by using Peter Baylies' StyleGAN2 model trained on images of historical paintings found on WikiArt.[58] But because this particular art object was created back in 2018, and AI-assisted art at that time was not as well-known and accessible as it is today, I prefer to give the work credence; while on the other hand its monetary worth can certainly be debated. For better or for worse, in the business of the art world, sometimes it just comes down to who sold their work first for the highest amount.

Chasing new technological advancements as a way to differentiate ML art often rewards speed, money, and computational power over raw creativity. While new technology is indeed exciting, it appears that the staying power of synthetic media calls for more than just state of the art capabilities. For GAN art to grow and mature, it appears to require laterally different creative approaches that inhibit the emergence of homogeneity and kitsch. This can potentially be realized by implementing a combination of methods, styles, and conceptual ideas that encourage new ways of seeing. Here, I will present some contemporary artists who are working with GANs and talk about their creative solutions which go beyond mere technological prowess.

[58] Baylies, P. "StyleGAN2 pbaylies fork". *GitHub.* 2020. https://github.com/pbaylies/stylegan2. Accessed April 25, 2021.

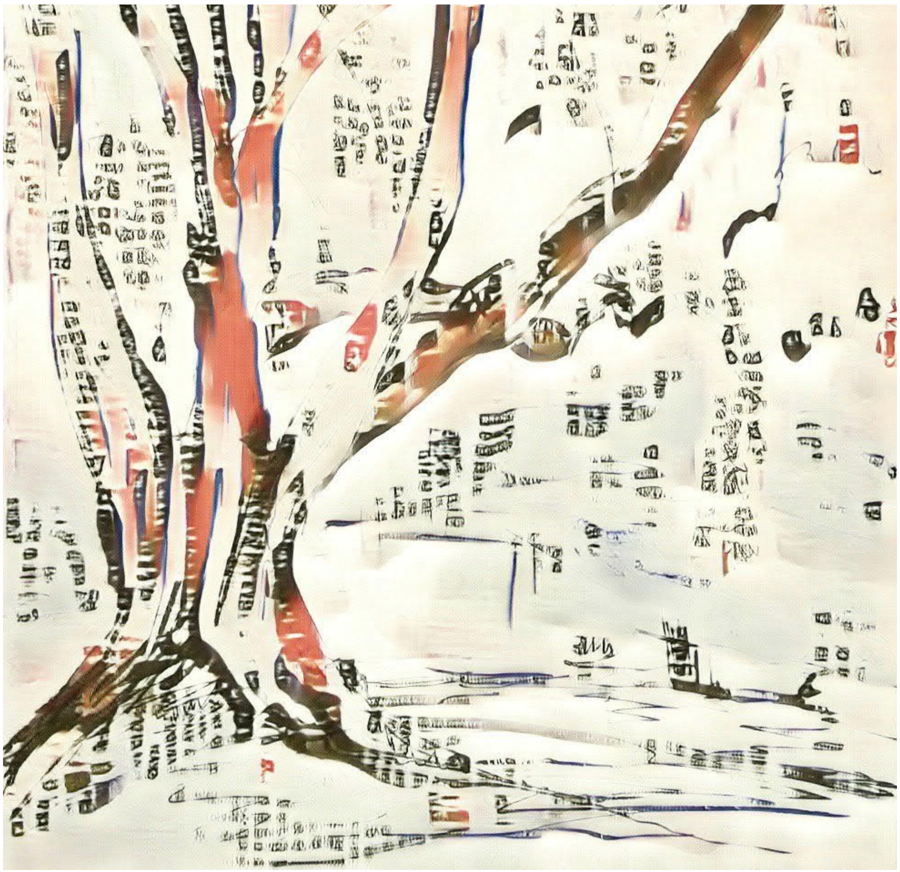

Figure 20: “Troops of Tourists Come for April Flower-viewing Oh, They’re Sparrow-men”, 2018 by Helena Sarin. GAN model trained on blooming trees and a book of Haiku.

A way to express a personal touch can be seen in the work of Russian visual artist and software engineer Helena Sarin, who focuses on training her own models with hand-crafted data sets using images of her drawings, paintings, and photography *(Figure 20)*. She frequently works with CycleGAN, which is an unsupervised image-to-image translation model, to generate her synthetic data sets. Image-to-image translation makes it possible to transfer visual data from one unlabeled image to another. Simple examples to help visualize this mechanism would be transferring the black and white stripes of a zebra onto a brown horse, transforming a photograph taken at night into day, or a picture of winter into summer.[59] Her method of image synthesis resembles techniques similar to collage, where fragments of collected images are composed into a unified work through juxtaposition. Jason Bailey, founder of the art and tech publication Artnome, comments on Sarin’s work by comparing her style to that of early Cubist collage works by Picasso and Braque, connecting them to her

[59] Zhu, JY. et al. “Unpaired Image-to-image Translation Using Cycle-consistent Adversarial Networks.” *Proceedings of the IEEE International Conference on Computer Vision.* 2017.

limited color palette and the way that she works with GANs by fracturing images and recombining their elements algorithmically to form new perspectives.[60] What Sarin's work effectively reveals is that custom-building data sets using one's own work and generating assemblages of them through GANs are one way of not only introducing more novelty into ML art, but also helping to maintain its uniqueness by keeping its visual outputs tied to the individual inputs and intentions of the creative author.

Berlin-based painter Roman Lipski, in similar fashion, also uses custom data sets, but his approach is quite different compared to that of Sarin. Lipski works in a kind of interactive partnership with his AI model, training the neural network to recreate his own paintings, and using the resulting visual outputs as inspirations for new paintings *(Figure 21)*. These in turn are fed back into the model ad infinitum. Working in this way over the course of several years, his project "Unfinished" has established what he describes as a constant intellectual dialogue which has fully transformed his artistic language and left him with exploring uncharted ways to express himself.[61] Lipski's methodology can be understood as interacting with an artificial muse, a system powered by ML which provides visual feedback and inspiration but one that also challenges the artist's creative process, all while the work perpetually transforms, at times unpredictably, in new stylistic directions.

[60] Bailey, J. "Helena Sarin: Why Bigger isn't Always Better with GANs and AI Art". *Artnome,* November 26, 2018. https://www.artnome.com/news/2018/11/14/helena-sarin-why-bigger-isnt-always-better-with-gans-and-ai-art. Accessed April 26, 2021.
[61] Verbist, E. "Roman Lipski - How Art Meets Artificial Intelligence". *ArtDependence Magazine,* November 9, 2018. https://www.artdependence.com/articles/roman-lipski-how-art-meets-artificial-intelligence-ai/. Accessed April 27, 2020.

Figure 21: “Unfinished 2”, 2016 by Roman Lipski. Acrylic on canvas. 217 x 355 cm.

Just like painters, artists of different media backgrounds are also attempting to see things differently through computer vision. London-based artist Memo Akten works with ML to create images, sounds, experimental films, and performances. The 2017 project “Learning to See”, consists of a real-time, interactive AI system, trained on a number of image data sets that makes predictions based on live camera input. Both a video work and an interactive installation, the model is part of an ongoing development with the aim of enabling continuous, meaningful human control of GANs, specifically for creative expression *(Figure 22)*. As the system processes the visual inputs of the camera, it reconstructs new images which resemble the input in composition, but has their overall form and content transformed based on determinations of the various data sets by the network.[62]

[62] Akten, M. Fiebrink, R. Grierson, M. “Learning to See: You are what you see.” *ACM SIGGRAPH 2019 Art Gallery.* 2019. pp 1-6.

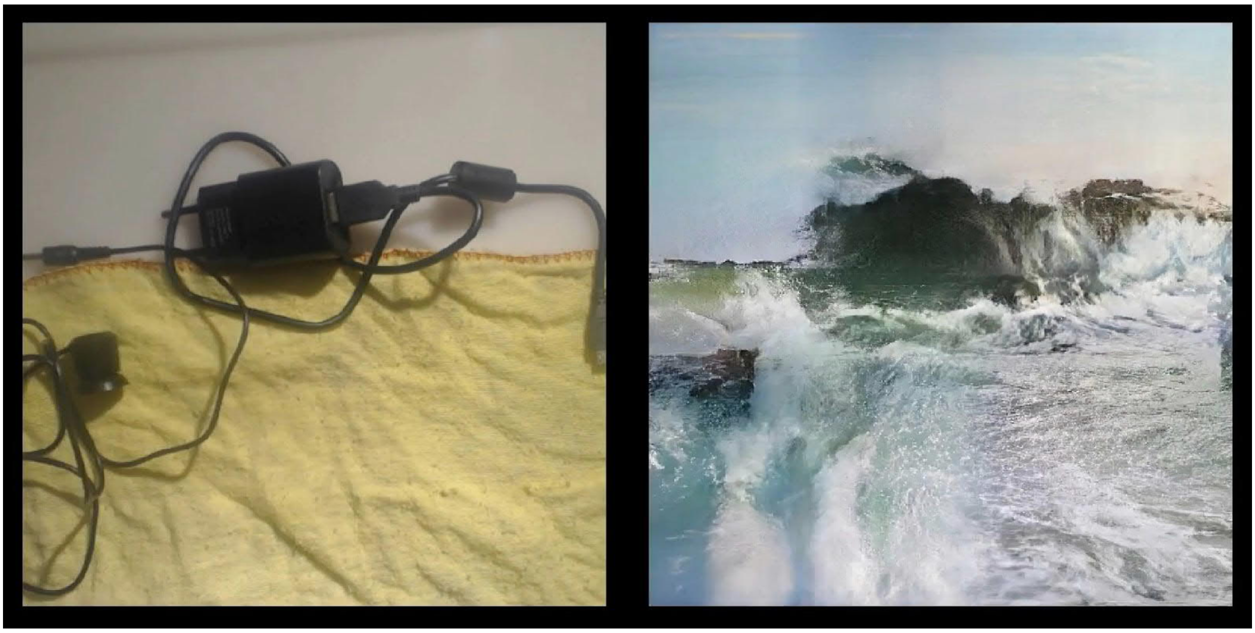

Figure 22: A frame from “Learning to See” by Memo Akten, using a model trained on ocean waves. (Left) Live image from camera, (Right) output from ML software.

This particular work amplifies in a sensitive way the learned bias in artificial neural networks, demonstrating how critical the training data is to the predictions that a particular model will make. It also underlines the huge potential that DL has in many exhibitory applications, including real-time visual presentations and live performances, which can be perceived as a visual analog to playing a musical instrument. ML art here reveals the level of interaction that is possible with neural networks, while simultaneously raising awareness around the idea that what AI reflects back to the observer is often sourced from our internal divisions and biases that have been programmed into the computational medium.

Figure 23: “From ‘Apple’ to ‘Anomaly’” by Trevor Paglen. Installation view at the Curve, Barbican. September 2019 - February 2020.

Recent work by artist Trevor Paglen, known for disclosing the politics of vision technology in the context of government surveillance and the ethics of drone warfare,[63] also explores the political and cultural biases revealed by machine learning. For his massive 2019 installation “From ‘Apple’ to ‘Anomaly’”, Paglen created a vast mosaic of 35,000 images displayed along the wall of the Curve Gallery at the Barbican Centre in London, depicting snap-shot sized photographs clustered around specific keywords. The photographs were all taken from ImageNet, the database mainly used as the source for training ML technologies, and collectively present the problematic relationship between images and class labels, in what Paglen himself describes as an extended homage to Magritte’s “Treachery of Images” for the age of machine learning.[64] Here the specific appeal to Paglen lies in the politics of categorization. While certain parts of the tableau appear benign, as for example with “pizza” or “honeycomb”, other sections

[63] Parks, L. Kaplan, C. (Eds.). *Life in the Age of Drone Warfare.* Duke University Press. 2020.
[64] Paglen, T. “From ‘Apple to ‘Anomaly’ (Pictures and Labels)”. *Paglen Studio.* 2020. https://paglen.studio/2020/04/09/from-apple-to-anomaly-pictures-and-labels-selections-from-the-imagenet-dataset-for-object-recognition/. Accessed April 27, 2021.

labeled “bad persons”, “alcoholics”, or “garbage heap” reveal a more sinister side of how certain stereotypes and judgments about humanity are programmed into the visual memory of AI models. The message of Paglen’s piece appears not only to be about the machine-learned biases which can act themselves out when deployed in real world settings, but also refers to the invisible nature of computer vision itself; namely that certain images taken by us no longer exist to pass through our organic senses, but rather through the oculus of machines, inherently inaccessible to human eyes.[65]

German artist Mario Klingemann also deals with the question of the meaning of photography through the eyes of a machine *(Figure 24)*. His multimedia work, which encompasses a wide scope of genres such as generative and evolutionary art, glitch art, data classification, visualization, and robotic installations, mainly explores how creatively repurposing images and systems can reveal to us hidden qualities about them. His particular interest lies in investigating the presence of visual anomalies in synthetic images, artifacts of a kind of chaotic materiality which make up the representative medium of AI, just like the film grain of a photograph, or the jpg fabric in its digital version.[66] Coining the term ‘neurography’, Klingemann directly compares the medium of ML to that of photography, where instead of framing images in the real world, one is generating and fixing virtual latent spaces in search for new perspectives.[67] With this he also suggests the idea that images in themselves are primarily based on systems consisting of codes and signals, whether it’s through the probabilistic determinations of visually literate algorithms, the interpolation of pixel data on digital camera sensors, or for that matter, visual perception generated by biological systems in the brain. Although one should be careful not to confuse this correlation entirely with fact, it does on the one hand bring about a certain questioning of how we are to regard our assumptions and convictions about photography in the emerging contexts of AI and synthetic imaging. Here, artwork made using ML not only has the power to stimulate our senses, but also

---

[65] Clark, T. “Trevor Paglen, From ‘Apple’ to ‘Anomaly’ (Pictures and Labels).” *British Journal of Photography 7888.* 2019.
[66] Barale, A. “The Ruin and the Artifact: Walter Benjamin and AI Art.” *Pensiero: Rivista di filosofia: LIX, 2.* 2020.
[67] Chatel, M. “In Focus: Mario Klingemann.” *Medium,* October 5, 2018. https://medium.com/digital-art-weekly/in-focus-mario-klingemann-783533ec91fe. Accessed April 27, 2021.

potentially can lead us to reevaluating our personal temperaments and fixed notions about the world. This effect of art on a spectator does not solely depend on technological efficacy and aesthetic display, but rather on the creative interplay of media technologies together with the conceptual ideas and actions of a critically-thinking artist who makes use of them.

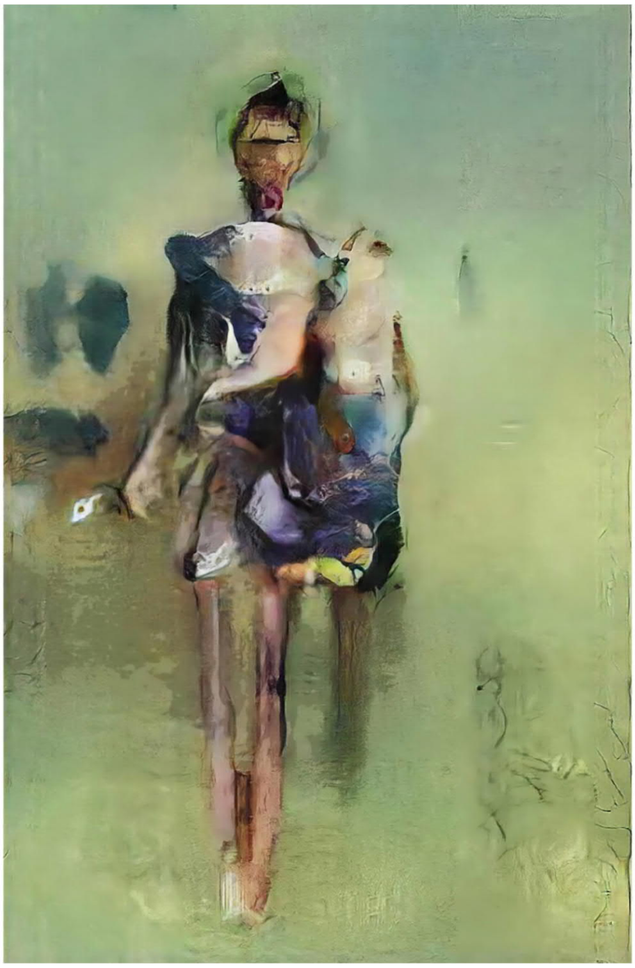

Figure 24: "Imposture Series - Do Not Kill the Messenger", 2017 by Mario Klingemann. AI trained on postures of stick figures, "painting" a neural network's vision of the human form.

4. Conclusion

For visual artists, photographers, and other creative practitioners, machine learning and artificial neural networks hold considerable potential for productivity, creative experimentation, and augmentation. In particular, GANs have shown remarkable capabilities of synthetisizing images; though within certain technical limitations. Developments in AI are progressing at a faster rate than ever before, facilitated by increased data availability, substantial leaps in GPU performance, and growing scientific and academic research. Therefore one can confidently predict the current state of generative imaging to change significantly within this decade. With this, creative tools assisted by DL algorithms will likely also evolve, with companies such as Topaz Labs, Runway AI, Google, and Adobe already offering machine learning applications and software integrations for creative end users. A wide variety of artistic uses of AI within the disciplines of photography, painting, sculpture, digital and interactive art demonstrate a large number of possibilities for exploring the aesthetics of synthetic media, for considering our relationships with digital technology, and calling attention to think critically about automation, big data, and computer vision. In this concluding chapter, I'd like to point out some key developments and issues I am currently observing, and take the opportunity to reflect on my own sentiments regarding creative AI.

The current state of AI imaging made a significant stride in 2020 with the release of Generative Pre-trained Transformer 3 (GPT-3) by the San Francisco-based for-profit research laboratory called OpenAI. GPT is a language model that uses deep learning to artificially produce human-like text. Also referred to as natural language processing (NLP) or neural NLP, the workings of this latest model is capable of generating a quality of texts so high that they have become virtually indistinguishable from those written by a human.[68] Its 2019 predecessor GPT-2, trained on 1.5 billion ML parameters, was already able to generate text on a level which mimicked that of humans but struggled with long passages of text, often causing repetition and nonsensical outputs. GPT-3 in comparison has a massive

[68] Sagar, R. "OpenAI Releases GPT-3, The Largest Model So Far". *Analytics India Magazine.* 2020. https://analyticsindiamag.com/open-ai-gpt-3-language-model/. Accessed April 30, 2021.

capacity of 175 billion machine parameters.[69] The reason I mention this model is because its advancement is already fundamentally changing the way images, or stories for that matter, can now be synthesized using pre-trained language prediction and image representation, thus making text generation one of the biggest trends in machine learning today.

Figure 25: AI-generated images by DALL-E using the text prompt “an armchair in the shape of an avocado. an armchair imitating an avocado.” (Image: OpenAI).

Another development connected to NLP is an AI program called “DALL-E”, also created by OpenAI, which can generate images from textual descriptions using 12 billion parameters (scaled down from GPT-3). By processing words and, more importantly, their relationships within sentences, DALL-E has the ability to visually render a variety of types of linguistic logic. These include object attributes, drawing multiple objects in a scene, visualizing perspective and three-dimensionality, internal and external structures, inferring contextual details, combining unrelated concepts, illustrations, visual reasoning, and geographic as well as temporal knowledge.[70] Input prompts such as “a photo of a phone from

---

[69] Brown, T. B. et al. “Language Models are Few-Shot Learners.” *arXiv preprint arXiv:2005.14165.* 2020.
[70] Ramesh, A. et al. “DALL-E: Creating Images from Text.” *OpenAI.* 2021. https://openai.com/blog/dall-e/. Accessed April 30, 2021.

the 1930s” or “an armchair in the shape of an avocado” are just some examples released by OpenAI that demonstrate its capabilities *(Figure 25)*.

One can indeed marvel at the successes of deep neural networks and how they have led to such amazing results in completing convoluted tasks and outperforming benchmarks. However, as James Bridle points out in his book *New Dark Age,* increasing computational complexities also leads to further clouding our understanding of the real world, by making us realize just how unknowable it really is.[71] As previously mentioned, AI is especially prone to implicit biases, such as racial discrimination in facial recognition, caused unknowingly by the training data bias of the human engineers who construct the supposedly universal, all-encompassing data sets.[72] In a review of Paglen’s 2017 exhibition of “Behind These Glorious Times”, Alexander Strecker presents the thought experiment of a self-driving car faced with the decision of distinguishing between two children running out into the road, one white and one black. What will the computer see? If it happens to “see” the black child as a small animal, then its choice, according to the machine, will be indubitably clear.[73] Strecker goes on to challenge us, the programmers and creators, asking how we will make room to have these debates about our own visual stereotypes, while being faced with the awareness that AI by itself does not have the capacity for self-reflection or ethical evaluations.

All technology is in some way a double-edged sword, only with AI its unintended consequences are much more pervasive and global. That too goes for its environmental impacts. Generative machines trained on abundant amounts of data logically require exceptionally large computational resources, and this in turn contributes greatly to the warming climate[74]. Determining the precise energy consumption of AI is however no easy task, as not all ML models are born equal.

---

[71] Bridle, J. *New Dark Age: Technology and the End of the Future.* Verso Books. 2018.
[72] Simonite, T. “The Best Algorithms Struggle to Recognize Black Faces Equally.” *Wired,* July 22, 2019. https://www.wired.com/story/best-algorithms-struggle-recognize-black-faces-equally/. Accessed April 30, 2021.
[73] Strecker, A. “An Urgent Look at How Artificial Intelligence Will See the World.” *Lens Culture.* https://www.lensculture.com/articles/trevor-paglen-an-urgent-look-at-how-artificial-intelligence-will-see-the-world#slideshow. Accessed April 30, 2021.
[74] Strubell, E. Ganesh, A. McCallum, A. “Energy and Policy Considerations for Deep Learning in NLP.” *arXiv preprint arXiv:1906.02243.* 2019.

Also, information regarding the energy grid which cloud servers are connected to are rarely publicly available, while many other factors contribute indirectly to emissions such as resource mining and manufacturing. Efforts are being made to help users of AI technology become more aware of their carbon footprint. Tools such as the “Machine Learning Emissions Calculator”, which was developed out of a scientific study to help quantify CO2 emissions of machine learning, can be used to estimate one’s power consumption when operating ML models.[75] One should be reminded however that it has in the meantime become an accepted standard of practice for big industries to delegate their own environmental and ethical responsibility onto consumers.[76] This is a norm that I feel should be deeply questioned and pushed back against, without ignoring one’s own personal contributions.

For artists, what does it mean to work with ML? As a fundamental consideration, I think it’s necessary to realize that by working with neural network architectures, for whichever creative purpose, one is also choosing to participate in the monopolized operations of big technology and data corporations. According to Joanna Zylinska, there can often be a tendency of AI artists to adopt a “worryingly uncritical instrumentalism”, which she describes as a pacifying attitude towards the processes of ML that ultimately prevents artists from pursuing any serious questioning of computer vision art.[77] While at the same time I don’t wish to fault the initial enthusiasm that arises when one first discovers for themselves the creative potential of machine learning. I can only state, from my own personal experience and limited amount of research, that the aesthetic thrills of driving a new generative ML model tend to wane eventually. There naturally comes a point where one needs to shift down gears, perhaps even pause for a moment, to consider the potential implications of the medium one is interfacing with, its invisible political influences and unforeseen anthropocenic impacts. I therefore encourage anybody interested in machine learning and artificial intelligence to give it a spin and experiment with it. The next

[75] Lacoste, A. et al. “Quantifying the Carbon Emissions of Machine Learning.” *arXiv preprint arXiv:1910.09700*. 2019.
[76] Kaufman, M. “The Carbon Footprint Sham.” *Mashable.* 2020. https://mashable.com/feature/carbon-footprint-pr-campaign-sham. Accessed April 30, 2021.
[77] Zylinska, J. *AI Art: Machine Visions and Warped Dreams.* Open Humanities Press. 2020. pp. 75-86.

step in the learning process will surely come, at the latest once gas runs out and requires refilling again. Let us then start seriously thinking about where exactly this gas comes from and, for the sake of sustainable solutions, make conscious efforts to stay out of our creative comfort zones. After all, the medium is a way, not the destination. And depending on who is driving, technology can either be a faithful servant or a dysfunctional vehicle.

List of References

Acharya, T. Tsai, PS. “Computational Foundations of Image Interpolation Algorithms.” *Ubiquity,* 2007 (October), article 4. 2007.

Akten, M. Fiebrink, R. Grierson, M. “Learning to See: You are what you see.” *ACM SIGGRAPH 2019 Art Gallery.* 2019. pp 1-6.

Alom, M. Z. et al. “The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches.” *arXiv preprint arXiv:1803.01164.* 2018.

Amato, G. et al. “AI in the Media and Creative Industries.” *arXiv preprint arXiv:1905.04175*. 2019.

Auer, F. “Bob Ross Paintings.” *TwoInchBrush.* 2019. https://www.twoinchbrush.com/all-paintings. Accessed February, 4, 2021.

Augustyn, A. et al. “Romanticism.” *Encyclopedia Britannica.* https://britannica.com/art/Romanticism. Accessed March 15, 2021.

Bailey, J. “Helena Sarin: Why Bigger isn’t Always Better with GANs and AI Art”. *Artnome,* November 26, 2018. https://www.artnome.com/news/2018/11/14/helena-sarin-why-bigger-isnt-always-better-with-gans-and-ai-art. Accessed April 26, 2021.

Barale, A. “The Ruin and the Artifact: Walter Benjamin and AI Art.” *Pensiero: Rivista di filosofia: LIX, 2.* 2020.

Batchen, G. “Electricity Made Visible”. *New Media, Old Media: A History and Theory Reader.* Edited by Chun; W. Keenan. Routledge. 2006. pp. 27-44.

Baudelaire, C. *The Mirror of Art, Critical Studies.* Translated by Mayne, J. Doubleday Anchor Books. 1956. pp. 231-232.

Baudelaire, C. *The Painter of Modern Life.* Translated by Charvet, P. E. Penguin UK. 1972.

Baylies, P. “StyleGAN2 pbaylies fork”. *GitHub.* 2020. https://github.com/pbaylies/stylegan2. Accessed April 25, 2021.

Bauer, M. *Resistance to New Technology: Nuclear Power, Information Technology and Biotechnology.* Cambridge University Press. 1997. pp. 57-79.

Benjamin, W. *The Work of Art in the Age of Mechanical Reproduction.* Translated by Zohn, H. Schocken Books. 1969.

Bridle, J. *New Dark Age: Technology and the End of the Future.* Verso Books. 2018.

Brock, A. et al. “Large Scale GAN Training for High Fidelity Natural Image Synthesis.” *arXiv preprint arXiv:1809.11096.* 2018.

Brown, T. B. et al. “Language Models are Few-Shot Learners.” *arXiv preprint arXiv:2005.14165.* 2020.

Brundage, M. et al. “The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation.” *arXiv preprint arXiv:1802.07228.* 2018.

Carneiro, T. et al. “Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications.” *IEEE Access,* vol. 6. 2018. pp. 61677-61685.

Chan, E. “From the ACR team: Super Resolution.” *Adobe Blog,* Creativity, March 10, 2021. https://blog.adobe.com/en/publish/2021/03/10/from-the-acr-team-super-resolution.html. Accessed April 22, 2021.

Chatel, M. “In Focus: Mario Klingemann.” *Medium,* October 5, 2018. https://medium.com/digital-art-weekly/in-focus-mario-klingemann-783533ec91fe. Accessed April 27, 2021.

Chemingui, Y. et al. “Reinforcement Learning-based School Energy Management System.” *Energies,* vol. 13(23), 6354. 2020.

Christina, V. et al. “Email Spam Filtering Using Supervised Machine Learning Techniques.” *International Journal on Computer Science and Engineering,* vol. 2(9). 2010. pp. 3126-3129.

Citron, D. K. Chesney, R. “Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security.” *Calif. L. Rev. 107.* 2019.

Clark, T. “Trevor Paglen, From ‘Apple’ to ‘Anomaly’ (Pictures and Labels).” *British Journal of Photography 7888.* 2019.

Cohn, G. “AI Art at Christie’s Sells for $432,500.” *New York Times.* 2018. https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html. Accessed April 24, 2020.

Creswell, A. et al. “Generative Adversarial Networks: An Overview.” *IEEE Signal Processing Magazine,* vol. 35(1). 2018.

David, D. *The Work of Art in the Age of Digital Reproduction, An Evolving Thesis: 1991-1995.* MIT Press. 1995.

Denker, J. S. et al. “Neural Network Recognizer for Hand-written Zip Code Digits.” *Advances in Neural Information Processing Systems*. 1989. pp. 323-331.

Deng, J. et al. “ImageNet: A Large-scale Hierarchical Image Database.” *2009 IEEE Conference on Computer Vision and Pattern Recognition.* 2009.

Dongare, A. D. et al. “Introduction to Artificial Neural Network.” *International Journal of Engineering and Innovative Technology,* vol. 2(1). 2012. pp. 189-194.

Goodfellow, I. J. et al. “Generative Adversarial Networks.” *arXiv preprint arXiv:1406.2661.* 2014.

Grewal, D. S. “A Critical Conceptual Analysis of Definitions of Artificial Intelligence as Applicable to Computer Engineering.” *Journal of Computer Engineering,* vol. 16(2). 2014. pp. 9-13

Halliwell, M. *American Culture in the 1950s.* Edinburgh University Press. 2007. pp. 4-12.

Hamosova, L. “Collective Vision of Synthetic Reality.” *Medium. 2020.* https://medium.com/@lenkahamosova/collective-vision-of-synthetic-reality-4ed02c01aaef. Accessed April 16, 2021.

Hertzmann, A. “Visual Indeterminacy in GAN Art*.” Leonardo*, vol. 53(4). 2020. pp. 424-428.

Isinkaye, F. O. et al. “Recommendation Systems: Principles, Methods, and Evaluation.” *Egyptian Informatics Journal,* vol. 16(3). 2015. pp. 261-273.

Jordan, M. I. Mitchell, T. M. “Machine Learning: Trends, Perspectives, and Prospects.” *Science,* vol. 349(6245). 2005. pp. 255-260.

Karras, T. et al. “Progressive Growing of GANs for Improved Quality, Stability, and Variation.” *arXiv preprint arXiv:1710.10196.* 2017.

Karras, T, Laine, S, Alia, T. “A Style-Based Generator Architecture for Generative Adversarial Networks.” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.

Kaufman, M. “The Carbon Footprint Sham.” *Mashable.* 2020. https://mashable.com/feature/carbon-footprint-pr-campaign-sham. Accessed April 30, 2021.

Kendrick, C. et al. “Anysize GAN: A Solution to the Image-warping Problem.” *arXiv preprint arXiv:2003.03233.* 2020.

Lacoste, A. et al. “Quantifying the Carbon Emissions of Machine Learning.” *arXiv preprint arXiv:1910.09700*. 2019.

McCulloch, W. Pitts, W. “A Logical Calculus of Ideas Immanent in Nervous Activity.” *Bulletin of Mathematical Biophysics,* vol. 5(4). 1943. pp. 115-133.

McDavid, J. “The Social Dilemma*.*” *Journal of Religion and Film,* vol. 24(1), article 22. 2020.

Mitchell, T. *Machine Learning.* McGraw Hill. 1997.

McLuhan, M. *Understanding Media: The Extensions of Man.* Signet Books. 1964.

Myers, B. A. “A Brief History of Human-Computer Interaction Technology.” *Interactions,* vol. 5(2). 1998. pp. 44-54.

Nagyfi, R. “The Differences Between Artificial and Biological Neural Networks.” *Medium.* https://towardsdatascience.com/the-differences-between-artificial-and-biological-neural-networks-a8b46db828b7. Accessed April 12, 2021.

Noll, A. M. “Computers and the Visual Arts.” *Design Quarterly,* no. 66/67. 1966. pp. 64-71.

O’Shea, K, Nash, R. “An Introduction to Convolutional Neural Networks.” *arXiv preprint arXiv:1511.08458*. 2015.

Parks, L. Kaplan, C. (Eds.). *Life in the Age of Drone Warfare.* Duke University Press. 2020.

Paglen, T. “From ‘Apple to ‘Anomaly’ (Pictures and Labels)”. *Paglen Studio.* 2020. https://paglen.studio/2020/04/09/from-apple-to-anomaly-pictures-and-labels-selections-from-the-imagenet-dataset-for-object-recognition/. Accessed April 27, 2021.

Pepperell, R. “Seeing Without Objects: Visual Indeterminacy and Art*.” Leonardo*, vol. 39(5). 2006. pp. 394-434.

Ramesh, A. et al. “DALL-E: Creating Images from Text.” *OpenAI.* 2021. https://openai.com/blog/dall-e/. Accessed April 30, 2021.

Sagar, R. “OpenAI Releases GPT-3, The Largest Model So Far”. *Analytics India Magazine.* 2020. https://analyticsindiamag.com/open-ai-gpt-3-language-model/. Accessed April 30, 2021.

Scace, R. I. “Electronics, The History of Electronics, Integrated Circuits.” *Encyclopedia Britannica.* https://britannica.com/technology/electronics/The-semiconductor-revolution#ref233776. Accessed April 4, 2021.

Schultz, D. “What is Google Colab? (A Gentle Introduction/Explainer).” *YouTube,* uploaded by Artificial Images, February 24, 2021. https://youtu.be/b7s-NKmOEpQ. Accessed April 24, 2021.

Schummer, J. “Aristotle on Technology and Nature.” *Philosophia Naturalis,* vol. 38(1). 2001. pp. 105-120.

Shamsolmoali, P. et al. “Image Synthesis with Adversarial Networks: A Comprehensive Survey and Case Studies.” *Information Fusion.* 2021.

Shiner. L. *The Invention of Art: A Cultural History.* University of Chicago Press. 2001.

Simonite, T. “The Best Algorithms Struggle to Recognize Black Faces Equally.” *Wired,* July 22, 2019. https://www.wired.com/story/best-algorithms-struggle-recognize-black-faces-equally/. Accessed April 30, 2021.

Strecker, A. “An Urgent Look at How Artificial Intelligence Will See the World.” *Lens Culture.* https://www.lensculture.com/articles/trevor-paglen-an-urgent-look-at-how-artificial-intelligence-will-see-the-world#slideshow. Accessed April 30, 2021.

Strubell, E. Ganesh, A. McCallum, A. “Energy and Policy Considerations for Deep Learning in NLP.” *arXiv preprint arXiv:1906.02243.* 2019.

Studio Kupol. “Brutalist Architecture Generator.” *RunwayML.* 2019. https://app.runwayml.com/models/Kupol/Brutalism_Generator. Accessed April 23, 2021.

Usselmann, R. “The Dilemma of Media Art: Cybernetic Serendipity at the ICA London.” *Leonardo,* vol. 36(5). 2003. pp. 389-396.

Valenzuela, C, et al. “Train an Image Generation Model”. *RunwayML*. https://learn.runwayml.com/#/create/train-image-generation. Accessed April 23, 2021.

Verbist, E. “Roman Lipski - How Art Meets Artificial Intelligence”. *ArtDependence Magazine,* November 9, 2018. https://www.artdependence.com/articles/roman-lipski-how-art-meets-artificial-intelligence-ai/. Accessed April 27, 2020.

Wan. Z, et al. “Bringing Old Photos Back to Life.” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 2020

Wang, Z. Hoi, S. C. H. “Deep Learning for Image Super-resolution: A Survey.” *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2020.

Wittkower, R. “Individualism in Art and Artists: A Renaissance Problem.” *Journal of the History of Ideas.* 1961. pp. 291-302.

Zhang, M. “BigGAN: A New State of the Art in Image Synthesis.” *Medium.* 2018. https://medium.com/syncedreview/biggan-a-new-state-of-the-art-in-image-synthesis-cf2ec5694024. Accessed April 21, 2021.

Zhao, M. Cong, Y. Carin, L. “On Leveraging Pretrained GANs for Generation with Limited Data.” *International Conference on Machine Learning*, PMLR. 2020.

Zhu, JY. et al. “Unpaired Image-to-image Translation Using Cycle-consistent Adversarial Networks.” *Proceedings of the IEEE International Conference on Computer Vision.* 2017.

Zylinska, J. *AI Art: Machine Visions and Warped Dreams.* Open Humanities Press. 2020. pp. 75-86.

List of Figures

Figure 8: The basic framework of a Generative Adversarial Network.
https://www.freecodecamp.org/news/an-intuitive-introduction-to-generative-adversarial-networks-gans-7a2264a81394/

Figure 9: Cards designed by Lenka Hamosova which represent the currently most available AI models for generating synthetic media. (Image: Hamosova).
https://hamosova.com/Collective-Vision-of-Synthetic-Reality

Figure 10: Random images generated with StyleGAN2.
https://thispersondoesnotexist.com/

Figure 11: "Bot Ross" (2021). StyleGAN2 trained on custom Bob Ross data set. Initial training checkpoint after 500 steps (above) and at completion of training after 5000 steps (below).
https://app.runwayml.com/models/alexrossa/Bob-Ross-Paintings

Figure 12: Class-conditional image samples generated by BigGAN.
https://medium.com/syncedreview/biggan-a-new-state-of-the-art-in-image-synthesis-cf2ec5694024

Figure 13: Increasingly realistic synthetic faces generated by variations on GANs. In order, the images are from papers by Goodfellow et al. (2014), Radford et al. (2015), Liu and Tuzel (2016), and Karras et al. (2018).
https://arxiv.org/abs/1802.07228

Figure 14: Zoomed crop of original compressed image (top left), enlarged 2x using Bicubic Smoother in Photoshop (top right), Preserve Details 2.0 in Photoshop (bottom left), and Super-resolution in Gigapixel AI by Topaz Labs.
https://pixabay.com/photos/cat-funny-cat-cute-cat-cute-4262034/ (Pixabay License)

Figure 15: Image restoration results produced by "Bringing Old Photos Back to Life." https://arxiv.org/abs/2004.09484

Figure 16: A screenshot of the interface of RunwayML.
https://www.theverge.com/2019/7/10/20682307/ai-machine-learning-easy-to-use-models-creatives-runway-ml

Figure 17: Screenshot of Google Colab running “The Big Sleep” notebook created by Ryan Murdock, running a text-to-image translation model.

Figure 18: Indeterminate images: (Left) The “impossible trident,” (middle) “Succulus,” 2015 by Robert Pepperell, oil on canvas painted specifically to be indeterminate, (right) a synthetic image generated with Artbreeder.
http://direct.mit.edu/leon/article-pdf/53/4/424/1881765/leon_a_01930.pdf

Figure 19: Edmond de Belamy, from La Famille de Belamy. Generative adversarial network print on canvas, 2018. Signed with GAN model loss function in ink by the publisher, from a series of eleven unique images, published by Obvious Art, Paris, with original gilded wood frame (700 x 700 mm).
https://www.christies.com/lot/lot-edmond-de-belamy-from-la-famille-de-6166184/?from=salesummery&intObjectID=6166184&sid=18abf70b-239c-41f7-bf78-99c5a4370bc7

Figure 20: “Troops of Tourists Come for April Flower-viewing Oh, They’re Sparrow-men”, 2018 by Helena Sarin. GAN model trained on blooming trees and a book of Haiku.
https://twitter.com/NeuralBricolage/status/988197444210102277

Figure 21: “Unfinished 2”, 2016 by Roman Lipski. Acrylic on canvas. 217 x 355 cm. https://www.artificialmuse.ai/mission-1

Figure 22: A frame from “Learning to See” by Memo Akten, using a model trained on ocean waves. (Left) Live image from camera, (Right) output from ML software.
https://www.memo.tv/wpmemo/wp-content/uploads/2017/07/msa_gloomysunday_003.jpg

Figure 23: “From ‘Apple’ to ‘Anomaly’” by Trevor Paglen. Installation view at the Curve, Barbican. September 2019 - February 2020.
https://www.lensculture.com/articles/trevor-paglen-from-apple-to-anomaly

Figure 24: “Imposture Series - Do Not Kill the Messenger”, 2017 by Mario Klingemann. AI trained on postures of stick figures, “painting” a neural network’s vision of the human form.
https://www.artsy.net/artwork/mario-klingemann-imposture-series-do-not-kill-the-messenger

Figure 25: AI-generated images by DALL-E using the text prompt “an armchair in the shape of an avocado. an armchair imitating an avocado.” (Image: OpenAI).
https://openai.com/blog/dall-e/