Navigating Subjectivity in AI-Generated Photography: The Quest for Ethics and Creative Agency

Paula Gortázar
University of Westminster, London, UK

Abstract

This study identifies alternative models for the production of AI-generated images to those currently used by mainstream AI platforms. Based on primitive computational art processes, these systems allow designers to gain greater control over the final visual result, while avoiding potential issues with intellectual property theft and breach of privacy. The article starts by analysing the level of artificiality that might be effectively attributed to each part of the creative process involved in the development of AI-generated images. It then moves on to discuss the extent to which individual actors intervene in each of these stages, with the aim of untangling the different subjectivities that would ultimately shape the meaning of AI-generated photographs, and identifying existing AI models able to offer practical solutions for every designer to retain creative autonomy over their AI creations.

Key Words: AI images, AI-generated photography, synthetic photography, expanded photography, virtual photography, photography theory

INTRODUCTION

The term ‘synthetic’ has quickly gained traction as a way to describe AI-generated photographs. The use of this adjective seems to operate as a synonym of ‘artificial’, and thus, in principle, is considered perfectly valid to describe the products of AI generative systems. It can be argued, however, that all photographs produce an artificial representation of reality, which can then be synthetically modified using a range of post-production processes. This is true whether the photograph is created through analogue or digital means. Crucially, Oxford’s definition of ‘artificial’ states that it refers to something ‘made or constructed by human skill, esp. in imitation of, or as a substitute for, something which is made or occurs naturally’ (*OED Online* 2023). If we accept that a photograph is already a synthetic representation of something natural, then AI-generated imaging processes are simply reiterating an artificial procedure that has already been performed through the production of the billions of photographs that exist in image datasets. What is most noteworthy about Oxford’s definition here is the emphasis on artificiality being something that is essentially ‘constructed by human skill’.

In 2022, artist Chin-Wei Liang, then a postgraduate student at the University of Westminster, was the first in the photography department to produce a Masters project entirely through the use of AI-generative technology. His series, *Untitled*, challenged the ambiguity of language often used by art critics (Figure 1). By feeding abstract descriptions of artworks that had been previously published by art critics through a written prompt to the AI program 'DALL-E', the system produced a series of images that held little resemblance to the original works to which the descriptions referred. Liang exhibited the prompt and AI-generated images side by side to illustrate the disparity between common language (that which is generally used to label and tag images in a dataset) and the complex linguistic constructions often employed by art critics, highlighting how the inaccessibility of art meaning to the wider public can be attributed, in part, to the distance between these two modes of language.

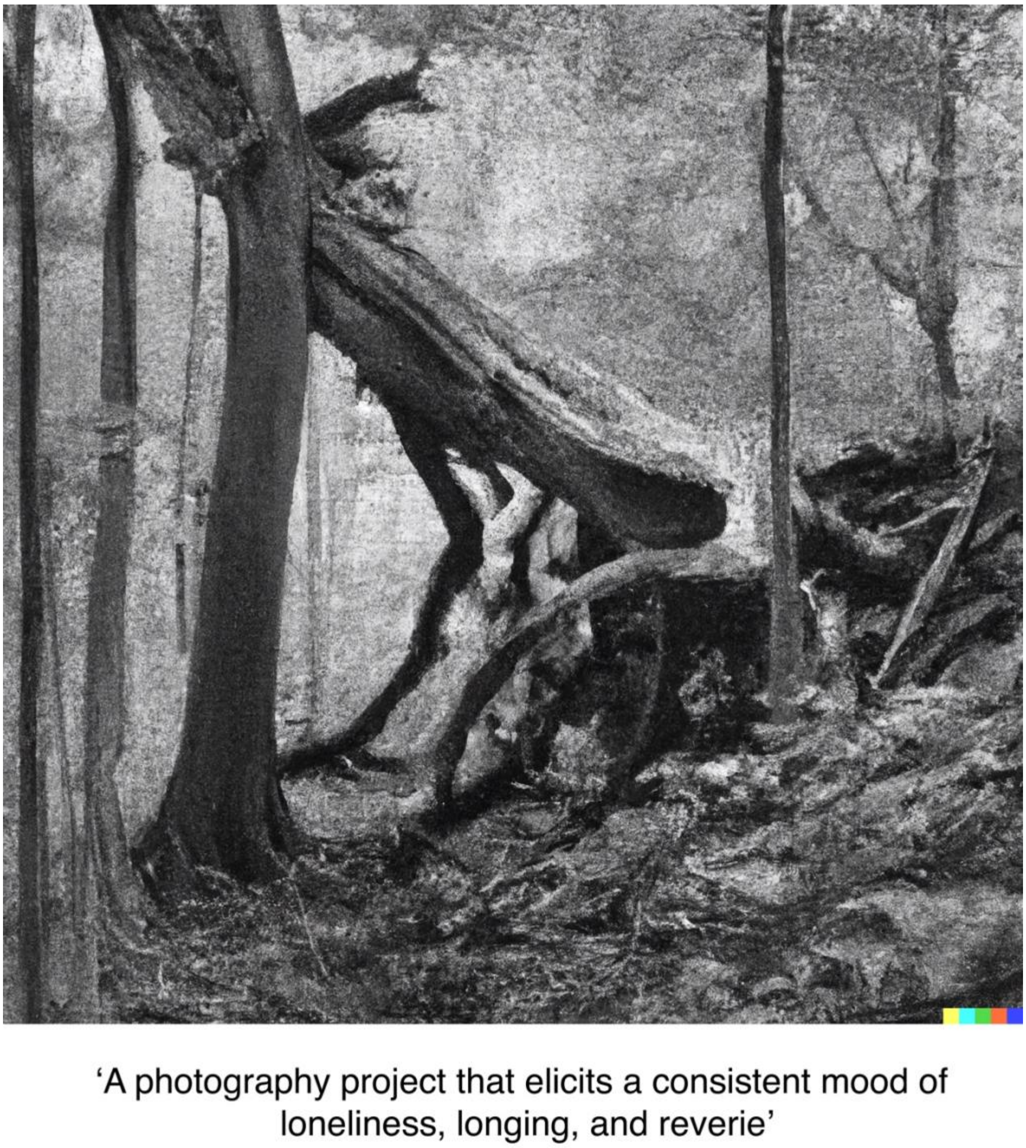

Figure 1, Chin-Wei Liang, from the series *Untitled*, 2022, AI image on paper, courtesy of the artist.

Beyond the rather plausible conceptual approach used by Liang, the exhibition of his work at Ambika P3 Gallery in London sparked a great deal of attention. While some expanded practices, such as CGI images and in-game photography, had long been accepted as constituting a valid form of photographic image making, including AI-generated photographs in a photography exhibition was clearly a novelty open to controversy. The AI hype, however, soon dominated the photography debate, and as is often the case with any revolutionary way of making, most users – creators and critics alike – moved at speed from positions of ignorance and suspicion to embracement and excitement.

The technology behind AI art production systems, however, is by no means new. The functioning of such systems is fundamentally based on technologies that have been around for several decades. While early experiments on machine learning and computer vision date back to the mid-1950s and ’60s respectively (Fjelland 2020: 2), the earliest examples of generative art were created as early as 1956 by Ben Laposky (Mckim 2021: 460). By the mid-1960s, the computational medium had already become the latest trend amongst a number of contemporary artists, such as Vera Molnar, whose *Machine Imaginarie* was used to produce abstract geometric forms following creative instructions in the form of computer code (Roe-Dale 2019: 18). Fast-forward to the early 2000s – creative coding programs such as Processing made it much easier for computational artists to create complex forms of generative artwork through the use of custom-made datasets. During the last decade, advancements in the field of computer vision and machine learning, alongside the availability of significantly larger image datasets provided by social media platforms and the recent advancements in computational hardware allowed for current forms of AI-generated imaging programs to be widely accessible to internet users (Citenic and She 2022: n.pag.). Following the launch of various AI art production platforms in 2022, such as Midjourney or Stable Diffusion, image creators no longer needed to know how to code, use a camera or master any postproduction software in order to create appealing images. Unsurprisingly, the possibility of making photo-realistic images with such ease created high levels of unrest amongst photography enthusiasts, who pointed at the urgency of establishing clear ‘rules of the game’, appealing at a ‘moral need’ to indicate when an image had been created through AI means.[1]

The question of whether AI-generated photographs can be considered photographs at all, or at least in part, is a debate that will surely attract photography theoreticians for years to come. However, at this stage, and perhaps as a precondition to resolving the ontological debate around this new form of image-making, it appears necessary to examine the creative nature of the different processes involved in the production of such images. While current forms of artificial intelligence are not capable of generating new knowledge in a human sense (Fjelland 2020: 2), as this paper will discuss, humans remain the main protagonists in the production of AI images. The starting point would thus be to question the level of artificiality that might be effectively attributed to each part of the process involved in the development of such images. This includes the creation of individual photographs that form the image dataset, the verification, labelling and categorisation of such photographs, the design of generative algorithms, the production of final images through prompt writing and other reference images, and the processes of communication and reception of its results. This article discusses the extent of human intervention in each of these stages, in order to untangle the different subjectivities that would ultimately shape the meaning of AI-generated photographs. Following this analysis, the study identifies an alternative production model based on primitive computational art processes, through which image creators might retain creative autonomy over their visual results, while avoiding potential issues with intellectual property theft and breach of privacy.

WE ARE ALL HUMANS IN THE MAKING

Until the 2010s, computer vision researchers worked with a limited number of images that were either obtained from the public domain or were specifically created to serve the purpose of a given dataset and its training objective. Once collected, these images had to be manually labelled and verified by humans, for them to be later classified and arranged into different categories according to their content and/or style. The image dataset was then used to train machines to perform a variety of functions aimed at achieving accurate visual recognition (Sluis 2022: 47). It soon became clear, however, that for machines to learn how to see like humans, they would need to be exposed to a comparable number of images that humans encounter in their first few years of their life. According to Fei Fei Li, creator of the popular dataset ImageNet, this would equate to hundreds of millions of pictures by the time a child turns three (Fei Fei Li 2015).

As the number of images uploaded to the internet continued to grow in the first decade of the millennium, collecting large amounts of photographs was an easily achievable task. In contrast, labelling such large numbers of items turned into a labour-intensive job. Paid at very low rates per verified photograph, professional image labellers, known as turkers, were hired by multinational companies like Amazon to provide a series of responses when confronted with a given image. Some of these recognition tasks were descriptive and required the identification of its content through the choice of a specific noun and adjective, while others were simply based on affirmative or negative answers. The images turkers were presented with came from a variety of online sources, including social platforms like Flickr. The resulting image datasets aimed at offering a large variety of content, in order to offer a picture of the world as close as possible to that experienced by human sight (Zylinska 2020: 110). As explained by Katrina Sluis in her text 'The networked image after Web 2.0: Flickr and the "real-world" photography of the dataset' (2022), in order to achieve this mission, the amateur snapshot soon became the preferred 'photo-style', since it appeared to have the ability to represent a diverse and 'non-aestheticised' picture of the world. As Sluis points out, 'for computer vision researchers, the snapshot enters the dataset as an *unmediated* product of a visual stimulus' (Sluis 2022: 52, emphasis added). Her discussion is in part based on a paper that accompanied the launch of Flickr's dataset YFCC100Ms. As its creators explained, the dataset included 'a diverse collection of complex real-world scenes, ranging from 200,000 street life-blogged photos by photographer Andy Nystrom, to snapshots of day to day life, holidays and events', and they defended that 'our dataset has been curated to be comprehensive and representative of real world photography, expansive and expandable in coverage…' (Thomee et al. cited in Sluis

2022: 49). But while Flickr's researchers aimed at building a dataset containing a picture of the world that was both objective and diverse, it is questionable whether objective choices could have been effectively made throughout the different stages of the dataset building process.

The idea that a snapshot – or indeed any other type of photograph – might capture a neutral representation of reality, appears completely at odds with the long-established assumption that a subjective view is unavoidably at play every time the shutter release is pressed, whether this is placed in a smartphone, a photographic camera, or programmed to function automatically in any other image-capturing device. This subjective view is always permeated with the individual's values and desires, which would ultimately shape the ideological connotation of their resulting image. As Victor Burgin observes 'the structure of representation, the eye and the base which captures it, is intimately implicated in the reproduction of ideology (we speak of a "point of view", a "frame of mind")' (Burgin 1986: 16). While certain images, snapshots or not, might appear non-aestheticised, they offer a point of view that is unique in the representation of a given experience and will therefore always be charged with subjectivity. Besides, as explained by William Davies in his article 'The reaction economy', following the proliferation of online social media platforms, 'spontaneity is now carefully staged, constantly looking for immediate reactions and engagement, while being wrapped into some sort of 'strategic authenticity' (Davies 2023: n.pag.). Indeed, most snapshots uploaded to social media platforms appear to have been given a considerable thought, even if, as Davies suggests, some users tend to make them appear as spontaneous and aesthetically unmediated.

Beyond the lack of neutrality of the visual narratives contained within the image dataset, the labelling and categorisation process further contributes to the subjectivisation of the collection. While human verifiers might spend very little time identifying the characteristics of each image, their responses are never objective, since, to use John Berger's words, these are always informed by individual 'ways of seeing' (Berger 1972). To overcome this issue, researchers might argue that turkers are based in a variety of geographic locations. Their different cultures might thus imply diverse values and visual influences, which as a result can permeate the dataset with a variety of 'ways of seeing' similar content. As explained by Joanna Zylinska, however, while it is true that turkers come from different territories, the data shows that over 93% of them are based in only five countries (United States of America, India, Canada, Great Britain, Philippines and Germany), with the United States alone accounting for 75% of those workers (Difallah et al. cited in Zylinska 2020: 102).The limited geographic scope of most turkers' bases evidences a lack of cultural diversity and thus poses serious questions about the neutrality of the labelling process. As a result, any claimed objectivity of the image dataset becomes clearly compromised. Besides, as Burgin reminds us in his chapter 'Seeing sense', 'regardless of how much we may strain to maintain a "disinterested" aesthetic mode of apprehension, an appreciation of the "purely visual" when we look at an image is instantly and irreversibly integrated and collated with the intricate psychic network of our knowledge' (Burgin 1986: 64). It would thus appear highly unlikely for these turkers to offer any objective identification, or neutral description, of the subjects represented in the images with which they are confronted. Thus, when pairing each image with their choice of label, they would be unavoidably adding an additional layer of subjectivity to the training collection.

The proliferation of social media platforms followed by the incremental upload of photographs to the internet and the interactions users make with such images by tagging, describing, liking and sharing their photographs, provided a vast amount of publicly available content, which datasets like CLIP made a good use of to train their generative models further (Cetinic and She 2022: n.pag.). What used to be a paid job delivered by dozens of thousands of turkers world-wide became increasingly available for free every time social media users uploaded visual content or interacted with that shared by fellow users. Both uploaded images and their accompanying texts, whatever their depicted subject might be, are currently up for grabs for companies such as LAION, whose latest dataset, released in March 2022, contains over five billion image-text pairs (Bridle 2023: n.pag.). This reckless appropriation of existing culture, alongside the exploitation of digital labour provided by internet users around the world, was made possible thanks to the entrepreneurial vision of Silicon Valley gurus who convinced millions of citizens world-wide that staying connected and sharing their experiences online could improve most aspects of their lives. In doing so, they generated highly effective surveillance systems based on computer vison technologies, where the users' passions, desires, fears and frustrations are permanently tracked to sell us all an appropriate consumable, service or political party that could help us solve our problems, whether those are material, physical or, in most cases, purely sentimental. If this were not profitable enough, advancements in computer vision and AI research have allowed social media platforms to further expand their business portfolio through the creation of the largest image datasets known to date. In doing so, the training collections contained in such datasets were filled with the aesthetic regimes and ideologies of its users, whose uploaded content (photographs, tags or comments) follows a specific, subjective purpose according to whatever personal objective they might hold. After all, this is precisely how the owners of social media platforms expect us all to behave.

Although some media theoreticians like Joanna Zylinska, have claimed that the increasing role of machines in image-making processes and the resulting cultural and technical algorithms they produce make 'most people's wedding photographs, holiday snapshots and Instagram feeds look uncannily similar' (Zylinska 2020: 107), it is clear that for their creators each depiction remains unique, as they unfold their particular aspirations onto the public realm of the internet through carefully thought-out aesthetic choices. And while it is evident that the appearance of such images often holds clear similarities (see for example Penelope Umbrico's *Sunsets from Flickr* series, 2006–ongoing), this is not so much caused by the increasing intervention of automated machine functionalities but rather by a shared 'regime of vision' pertaining to a given culture that tends to be present across certain territories (here an increasingly global one) during a limited time in history (Berger 1972).

Finally, the attribution of a certain dataset to specific purposes of machine learning constitutes the ultimate subjective choice, one that risks causing serious prejudices, and which can potentially bring detrimental consequences to its targeted subjects. Given that machines can only be trained using the specific items included in their allocated image dataset, these must collectively contain enough truthful information to provide the machine with a comprehensive range of the visual possibilities it may encounter, so that when required, the machine can trigger optimum automated responses.

An interesting example of a faulty image dataset was used by Volvo during a series of experiments aimed at training autonomous vehicles to detect animals on the road. As researchers explained, their dataset contained images of all sorts of available animals in Sweden, including a range of both domestic and wild animals of various colours and sizes. Whenever the car identified a still or moving animal in front of it, the computer vision performed through the vehicle's camera triggered an automated response that made it stop. What the dataset had not accounted for were the different ways in which animals move in other parts of the world. Kangaroos, for instance, move forward by making very high jumps. As a result, when tested in Australia, Volvo's camera moved from detecting the kangaroo to immediately thinking that there was no longer an animal in front of the car. The animal was still there, of course, only it was now a metre above the ground. The jumping motion meant the computer was moving from 'seeing' to 'not seeing', simply because it had not been trained to 'acknowledge' that certain animals jump, rather than walk or run (Zhou 2017: n.pag.).

While the allocation of a faulty image dataset probably delayed Volvo's research into autonomous driving, it was unlikely to carry any further consequences for the subjects involved. More detrimental results can be achieved, for instance, when training machines for law enforcement purposes. Facial recognition technologies used in those fields have proved to be especially weak at reading the features of people of colour, which as a result allows for repeated mistaken interventions to be performed on innocent black citizens. This is not due to the fact that darker skin is harder to read, but rather because the image datasets used for such purposes do not contain a sample that is comprehensive enough to train machines to accurately read dark faces. According to Alex Najibi, Associate Researcher at Harvard University, various advocacy groups have engaged with lawmakers, educating on racial literacy in face recognition and demanding accountability and transparency from *the producers* of training datasets (Najibi 2020: n.pag., emphasis added). But rather that *the producers* themselves, it is perhaps the qualified (political) decision made by *the allocators* of those datasets, in charge of assigning racialised training collections to law enforcement programs, that are more specifically contributing to sustained discriminative behaviours against black individuals and those of colour.

GENERATIVE SUBJECTIVITIES

The technology by which AI-generative programs such as DALL-E and Stable Diffusion produce images is based on a system called Diffusion Model. This model starts with a single noise vector and gradually compares the images in the training collection with the instructions prompted by users into the system, either through text or another reference image. The model progressively refines its results until it considers that the content and style provided by the user has been matched to the best of the system's abilities. One of the most common datasets used by Diffusion Model is CLIP, which was created by OpenAI to generate what they call a 'CLIP-guided diffusion process'. This is one of the largest image datasets available as it contains a number of secondary datasets, such as COCO, Flikr30K, Open Image and ImageNet, alongside a range of images from a variety of online sources that are publicly available but which the creators of CLIP have not yet provided specific details on (Salvaggio 2023). Interestingly, ChatGPT, a text generator also created by OpenAI, has taken care to defend itself forefront from the risks of generating biased visual results through their image generator product, DALL-E. When asked about this possibility, the response generated by ChatGPT read 'by including images from a wide range of sources, the dataset aims to cover a broad range of concepts and objects, and to minimize biases that may arise from any single source' (2023) [2].

While it is clear that using a larger sample of items in a training collection can help minimise the risk of biased results, given that this content arrives mostly from 'the wilderness' of the internet, both its images and accompanying tags, comments and labels, are likely to contain large amounts of discriminative content against minority groups, which would unavoidably feed the algorithm within the Diffusion Model. Indeed, it does not take much to test how biased the system is. Soon after DALL-E was launched, I conducted a test in the system by introducing a number of prompts to explore the different representations the program was likely to produce around the concept of 'motherhood'. To avoid being over-prescriptive, the prompts tested contained only a few words. One of the first requests read 'a mother exhausted after giving birth'. Astonishingly, in the generated images, these new mothers were often represented like animals, lying inside stables or in the middle of the wilderness, with physical features that in most cases referenced black women and women of colour. A similar test, this time reading 'a mother suffering from depression after giving birth', brought up very different results. In this case, those new mothers seemed to be recovering from labour inside a hospital room and almost all of them were white Westerns or Northern Asian female. The second test suggested that post-natal depression was probably something that was mostly discussed in wealthier countries with a majority of white females, which would account for a larger number of medical images and stock photos depicting the condition through the inclusion of female subjects. More troubling was the fact of seeing women with darker skin depicted as if they were giving birth outdoors or in farms, when these mothers were defined as 'exhausted' following labour. It is hard to think what training data would allow the Diffusion Model to come across such regrettable representation, or whether it might have been the Diffusion Model itself (the generative algorithm) that had arrived to such result by means of 'defining and redefining' a given visual pattern. Either way, the test made it evident in a matter of seconds that the combined used of CLIP dataset alongside OpenAI's generative algorithm was highly problematic when it came to perpetuating racist stereotypes.[3] As Lev Manovich anticipated in his book *AI Aesthetic*, the automated connections made by AI 'will play a large part in professional cultural production' (Manovich 2019: 8). Four years on from Manovich's forecasting, there is little question that this is likely to be the case. As image-making processes become increasingly automatised across the creative industries, and unless significant improvements are soon implemented into these systems, the resulting mass production of visual content through AI means will unstoppably expand discriminative stereotypes.

FIXING THE SUBJECTIVE BUG

There is so far no doubt that AI generative systems have arrived to stay, just like there is no doubt that they will become more refined and powerful over time, allowing creators to produce high-quality images able to meet their design objectives at a much lower cost and over a shorter period of time. But in order for these creators to match their business or artistic needs, and beyond the technological advancements that are still required to produce high-resolution images, designers would also need to gain much greater creative control over the content of their final result. As we have seen, AI-generative programs currently available employ the (often unpaid) labour of several anonymous actors whose subjective decisions, implemented throughout the different steps of the AI chain, influence the final look of every generated image. This suggests that in order for designers to gain full control of the creative process, they would need to own each of these stages, or at least be able to closely oversee what occurs in each phase, and have the ability to intervene should any re-direction be needed at some point in the process.

As explained earlier in the text, the technology used by AI systems is not radically new. Generative models have been used by computational artists for decades, with programs like Processing, launched in 2001, which made it easier for artists to build their works through creative coding. These computational artists often use datasets to create their practice. The items in their collections may include all sorts of media, such as sound, video, photographs, and can be used to produce generative outputs in a multiplicity of formats. The program thus allows artists to be involved in every step of the generative process, from the creation of the dataset (including the items it contains, its labels and categories), to the design of the generative model, its postproduction and communication process. They may choose, however, to use someone else's dataset if this choice aids the conceptual approach of their work, but even in those cases, computational artists would always have the possibility of overseeing each stage of the production should they want their 'single subjectivity' to dominate the entire generative process.

With the recent launch of AI-generative platforms to produce visual works, most of the generative process has been handed over to third parties. Aside from offering a highly advanced generative algorithm, the main 'improvement' provided by mainstream AI-generative systems is the vast number of items contained in the dataset on which the models are trained. This content, however, has the potential of generating serious conflicts in relation to IP theft and breach of privacy laws. Even when the items in the dataset have been made publicly available (often through obscure online usage agreements) their images cannot be automatically used for commercial purposes or shared in ways that its content might be publicly recognisable. While it is true that AI-generative models are designed to avoid those problems through the use of complex processes aimed at generating entirely new content (producing a composite rather than a straight collage) it has already been proved that 'data leakage' can be a very palpable threat. In a recent paper by Nicholas Carlini et al., titled 'Extracting Training data from Diffusion Models', the authors found that data could sometimes be retrieved from original datasets. From photographs of individual subjects to company logos, researchers in this study were able to extract over a thousand items through the use of Diffusion Models, something that, according to the authors, was not possible with primitive generative systems such as GANs (Generative Adversial Networks) (Carlini et al. 2023: n.pag.). The research thus raises another argument about the need to own control of the different stages of AI generative processes. This seems particularly crucial when producing photo-realistic images, where the subjects depicted within the training dataset could potentially become recognisable in the final generated result.

Unfortunately, mainstream AI generative platforms, such as Midjourney and Stable Diffusion, are not currently able to meet privacy and IP protection standards, while leaving designers with very little creative control over the final result, which, as explained, is also likely to be permeated with biased meaning layered by a multiplicity of anonymous contributors during the different steps of the generative process. This is not to say that AI generative systems are not useful for creative purposes. These programs have clearly demonstrated an enormous potential and will surely be used across a number of disciplines in the creative industries during the coming years, but before this can be done both accurately and ethically, creators must devise responsible ways to put the technology in use.

Figure 2, Generated Photos, screenshot showing a range of AI-generated faces, 2023.

An interesting example of how the above might be achieved is the tailored process developed by Generated Photos. The company, funded in 2019, offers their own generative AI system for the production of photo-realistic representations of human beings. Their image dataset was created entirely by their team members, who took over 30,000 pictures of professional models from different ages and ethnical backgrounds. The portraits were all produced in studio under similar lighting conditions, enabling the generated image composite to look as natural as possible (Figure 2). These employees were also in charge of labelling and categorising the images according to their physical characteristics. Finally, a team of its own developers designed a custom-made GANs model. Their generative system was then embedded into user-friendly software where clients could carefully craft the look of the final subject, selecting different parameters. As the system is used, it not only learns from the original training data shot in studio but also from the results it constantly delivers, allowing it to progressively remove more flaws and inconsistencies in its generated images (Generated Photos 2023: n.pag). The approach of Generated Photos is set to resolve most of the ethical, legal and creative issues discussed earlier. Privacy issues disappear with the consent of the original models whose representations are included in the dataset. The possibility of IP theft is also removed from the equation, since all images contained in their dataset are produced in studio by photographers from Generated Photos. Meanwhile, a greater control over the aesthetic result is achieved by designing the rest of steps of the AI process (labelling, categorising and generative coding). As it occurs with OpenAI's generative system, those steps are also produced by different individuals, likely to permeate the generative result with their own subjective decisions. In the case of Generated Photos, however, these people are not anonymous actors but known company workers set to perform under a single creative direction; a fact that is likely to unify the aesthetic choices made throughout the labelling process. As for the final user – that is, the client prompting generative instructions – they can be certain that they would not be incurring in IP theft, nor any offence against privacy rights, while enjoying the increased creative control that is made possible thanks to the fact that all items within the dataset have been produced under similar technical conditions and labelled to meet the client's 'generative needs'.

CONCLUSION

There is no question that the process of production, distribution and reception of traditional photographs is embedded with subjectivity at every step. As different actors intervene in this chain (photographer, retoucher, editor, gallerist, public), different layers of meaning are built into the photograph, attributing to it a set of possible connotations. But when a photograph is generated through AI, the accumulation of subjectivities is even greater. The different interventions move from the photographer and the labeller/tagger/commentator to the creator of the dataset, the designer of the generative algorithm, the prompt writer and finally to whatever means of distribution is used to communicate the final generated result. While a 'traditional photographer' might have certain room to oversee the decisions made by the additional actors who intervene in the communication chain of their work, for the creators of AI-generated photographs, liberating the resulting images from the accumulation of subjectivities embedded throughout the generative process seems an impossible task. This is especially true if the image dataset is formed by publicly available items that have been anonymously labelled and collected by a given corporation or institution that is oblivious to the user's needs.

The generative models used in primitive computational art practices offer some practical solutions aimed at gaining full ownership over the production of AI images. Computational artists have traditionally been in charge of all stages involved in the generative process. From the creation of the dataset to its organisation and algorithmic design, these artists can aim to have total control over their final creative result. In the recent, but growing, field of photo-realistic AI production, this was precisely the approach followed by Generated Photos. Despite its being a labour-intensive process, the company has devised a product that can be sold to a large variety of industries, such as gaming, advertising or education, with the certainty that they are complying with both legal and ethical standards, while minimising the risks of some of AI's biggest ethical barriers attributed to use of biased training collections. The expansive work performed by the team members of this company demonstrates that AI generative systems have not necessarily arrived to make the creative process easier, nor eliminate large parts of the workforce. When done properly, these systems require several professionals with a variety of skills, all of whom need to have a sound understanding of the operational frameworks involved in each step of the generative process. This type of twentieth-first century digital creators, able to work fluidly across different technologies, are what Alan Warburton has called 'soft subjects'. As opposed to the idea that automation makes the creative process easier, the author argues that the lack of 'granular control' offered by automated processes, alongside the multiplicity of software that is constantly emerging and evolving, means that 'competent' digital creators are more needed than ever. This competency, explains Warburton, can only be achieved by an operational understanding of key production frameworks, which, once mastered, could be adapted and extrapolated to any technology, allowing digital creators to use the different tools in each software – hack them and crush them – in ways that could never be achieved through automated functions alone (Warburton 2023: 118).

While it would be unrealistic to ask any individual AI creator to have active control over every step of the generative process, it seems only reasonable to expect AI generative platforms to do so. This would ultimately require these companies to meet IP costs and be responsible for the privacy of every subject depicted in their image datasets. Likewise, in order to minimise the risks of bias, they should be required to comply with rigorous ethical procedures in the both production and distribution of datasets, as well as the algorithmic design itself. Once implemented, such improvements could provide final users with AI programs where their visual creations might be produced with greater control, not only in relation to the aesthetic result but, most importantly, through newly made representations able to retain both a legal and ethical approach. Through a process that combines human creative abilities with endless algorithmic potential, such systems may offer a highly efficient yet sustainable model for the production of AI images in the decades to come.

REFERENCES

'artificial' 2(023), *OED Online*, 3r ed., Oxford: Oxford University Press, https://www.oed.com/view/Entry/11211?redirectedFrom=Artificial&. Accessed 23 March 2023.

Berger, John (1972), *Ways of Seeing*, London: British Broadcasting Corporation and Penguin Books.

Bridle, James (2023), 'The stupidity of AI', *The Guardian,* 16 March, https://www.theguardian.com/technology/2023/mar/16/the-stupidity-of-ai-artificial-intelligence-dall-e-chatgpt. Accessed 20 March 2023.

Burgin, Victor (1986), *The End of Art Theory*, London: Macmillan.

Carlini, Nicholas, Hayes, Jamie, Nasr, Milad, Jagielski, Matthew, Sehwag, Vikash, Tramèr, Florian, Balle, Borja, Ippolito, Daphne and Wallace, Eric (2023). 'Extracting training data from diffusion models', Cornell University, 30 January, https://arxiv.org/abs/2301.13188. Accessed 2 April 2023.

Cetinic, Eva and She, James (2022) 'Understanding and creating art with AI: review and outlook', *ACM Transactions Multimedia Computing, Communications and Applications*, 18:2, pp. 1–22, https://dl.acm.org/doi/abs/10.1145/3475799. Accessed 20 March 2023.

Davies, William (2023), 'The reaction economy', *London Review of Books*, 45:5, https://www.lrb.co.uk/the-paper/v45/n05/william-davies/the-reaction-economy Accessed 15 April 2023.

Difallah, Djellel, Filatova, Elena and Ipeirotis, Panos (2018), 'Demographics and dynamics of mechanical turk workers', *Proceedings of WSDM 2018: The Eleventh ACM International Conference on Web Search and Data Mining*, Marina Del Rey, CA , USA , 5–9 February, pp. 135–143, https://doi.org/10.1145/3159652.3159661Accessed 15 February 2023

Fjelland, Rajnar (2020), 'Why general artificial intelligence will not be realized', *Humanities and Social Sciciences Communications*, 7:10, n.pag, https://www.nature.com/articles/s41599-020-0494-4, Accessed 10 March 2023.

Generated Photos Official Website (2023), 'FAQ', https://generated.photos/faq. Accessed 19 April 2022.

Li, Fei Fei (2015), 'How we are teaching computers to understand pictures', *TED Talk*, https://www.ted.com/talks/fei_fei_li_how_we_re_teaching_computers_to_understand_pictures. Accessed 25 January 2023.

McKim, Joel (2021) 'Oscillons and cathode rays: photographic hybrids in early computer art', *Photographies*, 14:3, pp 459–447, https://www.tandfonline.com/doi/full/10.1080/17540763.2021.1959387. Accessed 3 April 2023.

Manovich, Lev (2019), *AI Aesthetics,* Moscow: Strelka Press, http://manovich.net/index.php/projects/ai-aesthetics. Accessed 3 January 2023.

Najibi, Alex (2020), 'Racial discrimination in face recognition technology', Harvard University', 24 October, Harvard Unversity, https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/. Accessed 16 January 2023.

Roe-Dale, Rachel R. (2019) 'Vera Molnar's exploration of the unimaginable', *Math Horizons*, 26:4, pp. 18–21, https://www.tandfonline.com/doi/full/10.1080/10724117.2019.1568081. Accessed 4 February 2023.

Salvaggio, Erik, official website (2023), 'Diffusion, CLIP & LAION: flowers blooming backward into noise', https://www.cyberneticforests.com/ai-images. Accessed 3 March 2023.

Sluis, Katrina (2022), 'The networked image after Web 2.0: Flickr and the "real-world" photography of the dataset", in A. Dewdney and K. Sluis (eds.), *The Networked Image in Post-Digital Culture*, London: Routledge, pp. 41–59-

Umbrico, Penélope Official Website (2006), 'Sunsets from Flickr, 2016 – ongoing' http://www.penelopeumbrico.net/index.php/project/suns-from-sunsets-from-flickr/. Accessed 2 January 2023.

Warburton, Alan (2023) 'Soft subjects. hybrid labour in media software', in A. Dewdney and K. Sluis (eds.), *The Networked Image in Post-Digital Culture*, London: Routledge, pp. 114–131.

Zhou, Naaman (2017) 'Volvo admits its self-driving cars are confused by kangaroos', *The Gurdian*, 1 July, https://www.theguardian.com/technology/2017/jul/01/volvo-admits-its-self-driving-cars-are-confused-by-kangaroos. Accessed 1 February 2023.

Zylinska, Joanna (2020), *AI Art. Machine Visions and Warped Dreams*, London: Open Humanities Press, http://www.openhumanitiespress.org/books/titles/ai-art/. Accessed 2 February 2023.

Biography

Dr Paula Gortázar is a researcher and visual artist. Her research practice investigates recent developments of AI systems and extended reality technologies, and how these are exponentially changing the possibilities in which photographic images might be created and used within, and beyond, virtual space. Her work has been published in several academic journals, including *Photography and Culture*, *Third Text* and *Fotocinema*, and presented at multiple international conferences across the world. Her artistic practice has been widely exhibited in Canada, the United States of America, Germany, the United Kingdom and Spain, including venues like The

Photographers' Gallery, and published in different international media, such as *EXIT* magazine and *The British Journal of Photography*. She currently lectures in photography at the University of Westminster in London.

---

[1] See, for example, the controversy that arose during the Sony World Photography Awards in 2023, when one of the awardees, Boris Eldagsen, recognised that his image had been entirely created by AI and should thus not be allowed to compete in a photography competition. The full rejection discourse, alongside his personal plea to set 'the new rules of the game' can be read here: https://www.theguardian.com/technology/2023/apr/17/photographer-admits-prize-winning-image-was-ai-generated.

[2] This answer was generated during a 'conversation' between the author, Paula Gortázar and ChatGPT, which took place on the 3 March 2023

[3] Note that I have deliberately chosen not to publish the generated results of these tests in this article. Firstly because, despite not having any real referent, the representations appeared highly disrespectful towards the concept of motherhood, and secondly because any image published online might eventually be included in an image dataset, thus further contributing to the stereotypes generated through these artificial representations.