



Article

GANs and Artificial Facial Expressions in Synthetic Portraits

Pilar Rosado ^{1,*} , Rubén Fernández ² and Ferran Reverter ²

¹ Arts and Conservation-Restoration, University of Barcelona, 08028 Barcelona, Spain

² Genetics, Microbiology and Statistics, University of Barcelona, 08028 Barcelona, Spain; rubenfb14@gmail.com (R.F.); freverter@ub.edu (F.R.)

* Correspondence: pilarrosado@ub.edu

Abstract: Generative adversarial networks (GANs) provide powerful architectures for deep generative learning. GANs have enabled us to achieve an unprecedented degree of realism in the creation of synthetic images of human faces, landscapes, and buildings, among others. Not only image generation, but also image manipulation is possible with GANs. Generative deep learning models are inherently limited in their creative abilities because of a focus on learning for perfection. We investigated the potential of GAN's latent spaces to encode human expressions, highlighting creative interest for suboptimal solutions rather than perfect reproductions, in pursuit of the artistic concept. We have trained Deep Convolutional GAN (DCGAN) and StyleGAN using a collection of portraits of detained persons, portraits of dead people who died of violent causes, and people whose portraits were taken during an orgasm. We present results which diverge from standard usage of GANs with the specific intention of producing portraits that may assist us in the representation and recognition of otherness in contemporary identity construction.

Keywords: generative adversarial networks; latent space; synthetic photographic portraits; generative art; facial expression



Citation: Rosado, P.; Fernández, R.; Reverter, F. GANs and Artificial Facial Expressions in Synthetic Portraits. *Big Data Cogn. Comput.* **2021**, *5*, 63. <https://doi.org/10.3390/bdcc5040063>

Academic Editor: Min Chen

Received: 12 September 2021

Accepted: 29 October 2021

Published: 4 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Generative Adversarial Networks

The intersection between computer vision, machine learning, and computational models of the brain are proving to be an extremely fertile field of study in the area of photography. The increase of computational capacity and the use of convolutional neural networks (CNN, a deep neural network architecture currently applied to image-analysis) have favored and accelerated the development of new scenarios, making it possible to analyze large quantities of images. The outcome is that, in recent years, we are seeing a constant growth in the capacity of algorithms to emulate human systems of visual perception.

Using neural networks, in 2014, researcher Ian Goodfellow and his colleagues invented a new generative model of deep learning known as generative adversarial network (GAN) [1]. Radford et al. [2] used better neural network architectures: deep convolutional neural networks for generating faces (DCGAN). GAN works by pitting two neural networks against each other in a competitive training. The generative network tests out combinations of elements from a latent space to generate candidate images to be similar to those in the database. The other network, the discriminator, assesses the candidates put forward by the generative network using the original data in order to decide which faces are genuine and which are artificial. GANs are a powerful generative model and their popularity and importance have led to several reviews with the goal of made comprehensive studies explaining the connections among different GANs variants, and how they have evolved [3–7]. GANs apply in image processing and computer vision, natural language processing, music, speech and audio, medical field, data science, etc. They have been widely studied since 2014 and different variants have evolved. For instance, Isola et al. [8] introduce the so called Pix2Pix model in which implement a conditional GAN to image-to-image translation problems. They are conditional because instead of starting the generator

network from noise, they condition it by using an actual image. Karras et al. [9,10] described a new training methodology for GAN (StyleGAN): the main idea is to grow both the generator and the discriminator progressively, starting from a low resolution and adding new layers to model increasingly fine detail as training progresses. The resolution and quality of images produced by GANs have seen rapid improvement [11] Recently, Cheng et al. [12] proposed an effective method to defense adversarial examples in the field of remote sensing image scene classification, named PSGAN. A training framework was proposed to train the classifier by introducing the reconstructed examples, in addition to clean examples and adversarial examples. Quian et al. [13] propose a novel TSE-GAN with progressive training for co-saliency detection, which can extract multi-level enhanced inter-saliency features with robust semantic consistency and also alleviate the problem of insufficient co-saliency labels. The quality of generated faces by GANs is improved year by year and Karras et al. [14] present new results that pave the way for generative models better suited for video and animation.

There are plentiful GAN methods for face manipulation: Härkönen et al. [15] describes a technique to analyze GANs and create interpretable controls for image synthesis, such as change of viewpoint, aging, lighting, or expression. Liu et al. [16] propose a GAN that involves semantic conditional information of the input by embedding facial attribute vectors in both the generator and discriminator, so that the model could be guided to output elderly face images with attributes faithful to each corresponding input. Broad et al. [17] introduce network bending model that allows for the direct manipulation of semantically meaningful aspects of the generative process. In exploring the limit of how far human expressions can be captured, in this article we have train GANs using a collection of portraits of detained individuals, portraits of dead people who died of violent causes and people whose portraits were taken during an orgasm. In our experiments, we have used DCGAN [2], StyleGAN [10], and StyleGAN2 [11] (see supplemental Appendix A for more details about methodology).

A list of nomenclature used throughout this paper is provided in Table 1 as follows.

Table 1. List of nomenclature used in this paper.

Nomenclature	Referred to
AdaIN	Adaptive Instance Normalization
CNN	Convolutional Neural Networks
CONV	Convolution Layer
DCGAN	Deep Convolutional GAN
FC	Full Connected Layer
FFHQ	Flickr-Faces-High Quality
GAN	Generative Adversarial Networks
GPU	Graphical Processing Unit
NVlabs	NVIDIA Research Projects
PNG	Portable Network Graphics
ProGAN	Progressively Growing GAN
StyleGAN	Style-Based Generator Architecture for GAN
PSGAN	Perturbation-seeking GAN
TSE-GAN	Two Stream Encoder GAN

The remainder of this paper is organized as follows: artificial creativity is discussed in Section 2. Section 3 is devoted to present main results of the artistic projects Prosopagnosia, La Petite Mort, and Beautiful Agony. Section 4 discusses the potential of GAN latent space for artistic purposes and finally, Section 5 exposes the conclusions of our research. Methodological aspects are summarized in supplemental appendices.

2. Artificial Creativity

In our days, as every minute goes by, we take millions of images and pour them into the net, like a river with an endless water-supply. With them, we supply the trace of our

shared patterns. The difficulty of rescuing images from this torrent and gaining access to their visualization in an intelligible manner can, in fact, make them invisible. Deep learning technologies [18], through automatic image categorization tasks, make these enormous visual files more comprehensible. Generative models, feeding on these data sets, have achieved an unprecedented level of realism in the generation of synthetic images.

Deep learning is based on the idea of automatic learning by means of given examples. The algorithm arising from the sample builds a model capable of predicting the underlying rules and modifying it when predictive errors arise; thus, in an iterative manner, these systems can learn in an extremely precise way because the method is capable of extracting relevant patterns for the requested task. The power of deep learning is due to the capacity of these methodologies to unassistedly discover the visual descriptors without needing to use predefined ones. In 2019, Lev Manovich [19] pointed out that this aspect of automatic learning is in fact something new in the history of computer-art; a computer that is capable, on its own, of learning the structure of the world. Indeed, Manovich himself analyzes the various strategies for applying artificial intelligence to the art of the present time, in an attempt to clarify whether it generates a type of art we could genuinely call 'new'. We are encountering proposals of art generated by intelligent algorithms that have the purpose of simulating examples from art history, as well as other standpoints where the process of production still requires the intervention of a human author to make choices and control the system at some point. Although recognizing creativity in machines or in humans is no simple task, for some years now authors such as Margaret Boden [20] are studying the way in which artificial intelligence can help us to understand and explain creativity in humans. Berns et al. [21] analyze GAN artists' practices establishing bridges between generative deep learning, art, and computational creativity.

3. Results

3.1. Prosopagnosia Project

This type of technological innovations awakens some concern as they lead us to perceive the question of identity as something problematical, from the moment that computers become able to replicate features, such as our face, that are the very definition of our individuality. In this sense, for some time, Pilar Rosado and Joan Foncuberta have been developing a project called Prosopagnosia, related to facial recognition. Both artists have their work facilities at Roca Umbert Factory of the Arts in Granollers, Barcelona. This neighborhood situation allowed them to find shared concerns about art and technology that fueled their collaboration on different projects.

Prosopagnosia is a complaint that affects perception and memory, so that it is difficult or impossible to recognize faces. The project began with the creation and publication of a book under the same title [22] that consisted of a 21 foot-long (7 meters) leporello. On the "recto" side are original portraits of people who were well-known in the 1920s, found in the files of journalist Alonso Bonet, from the daily paper "La Prensa", distributed in Gijón between 1921 and 1936 [23]. It consists of a curious collection of very varied newspaper cuttings which they used when they had to publish an article about some celebrity (Figure 1).

On the "verso" side, with a glossy finish suggestive of the screens of electronic devices, appears a sequence of portraits generated from these photographs by means of the DCGAN [2] (see Appendix A.1). The book captures the wonderful sequence of learning of the computer. In the creation series shown in Figure 2 (the time-line should be read from top to bottom and left to right) we can see the learning process of the algorithm. It starts by generating images with the pixels ordered at random but, little by little, the system discovers the pattern associated with the faces and it improves the creations until it obtains photographically convincing images of people who never existed. Establishing an analogy with the morphogenesis of a living being, observing the failed images produced evokes for us important steps in the history of art, such as expressionism, cubism, surrealism, and abstraction.

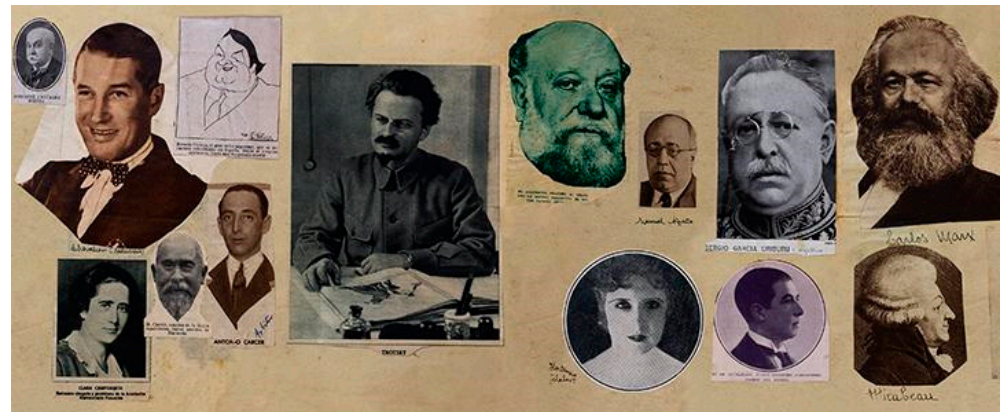


Figure 1. *Prosopagnosia* book (fragment). On the “recto” side are original portraits of people who were well-known in the 1920s, found in the files of journalist Alonso Bonet. Reprinted with permission from ref. [22]. Copyright © 2019 RM Ediciones.



Figure 2. The “verso” side of *Prosopagnosia* book (fragment) shows a sequence of portraits generated by means of the DCGAN. Reprinted with permission from ref. [22]. Copyright © 2019 RM Ediciones.

The result of this process is a new collection of photo-realistic images of historical characters from the 1920s who never existed. As we can see in Figure 3, the final images are, by and large, photographically convincing (we can remember the stunning Mike Tyka’s *Portraits of Imaginary People* presented at Ars Electronica in 2017 [24]).

The way in which the GAN, which are two neural networks interacting together to generate something completely new, seems like a vision, like a miracle; they are capable of converting random noise into a photographic portrait without having received any previous information of high-level figurative concepts such as “head”, “ears”, or “eyes”. For the generative system to be successful it is also requested not to reproduce images equal to those in the training group, but to make new proposals. It is revealing to make a detailed study of the strategies these algorithms use to learn. In the procedure, a generative neural network and another discriminative network are interconnected and work together. In the learning process of the discriminator net, it is fundamental to label the real images as ‘real’ and the false images as ‘fake’. However, at a certain moment, the evolution of the system will be boosted by the use of deceitful strategies that ‘surprise’ the discriminator system and constrain it to make corrections of what had already been learned before. These strategies consist of labeling the generated images as if they were ‘real’; again, this trick enables the generator system to make the necessary corrections and thereby turn out portraits of greater likeness to those provided as the model.

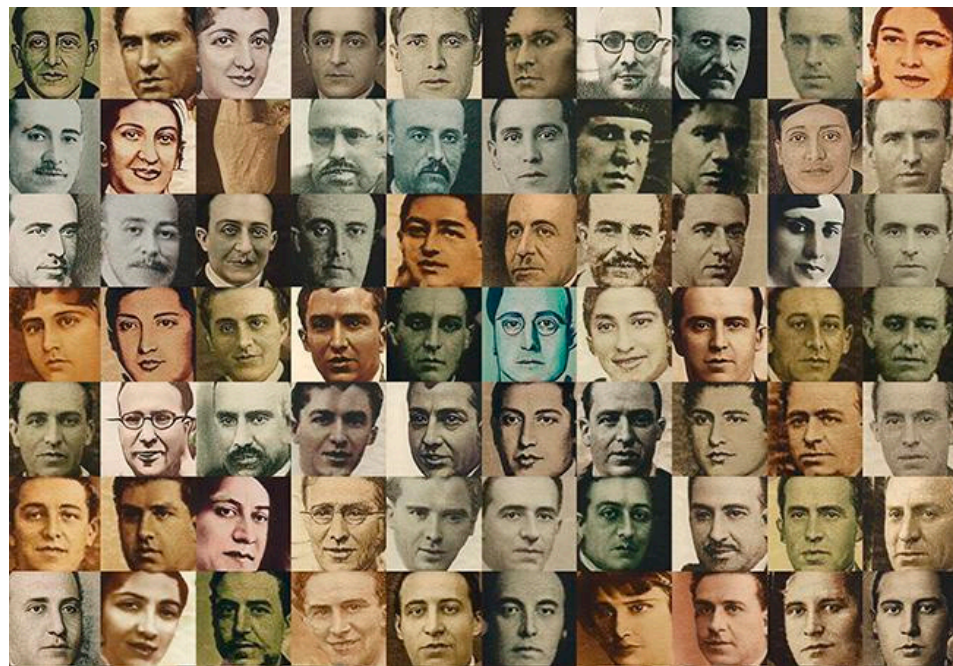


Figure 3. New collection of photo-realistic images of historical characters from the 1920s who never existed, generated by means of the StyleGAN. Reprinted with permission from ref. [22]. Copyright © 2019 RM Ediciones.

This iterative generative learning system still has some challenges to overcome. When confronted with an insufficient sample, the model collapses and is unable to bring forth new proposals, generating the same image over and over again. (Figure 4). The GAN can also behave in an unstable fashion, in which case it unlearns, and with each iteration, the portraits get worse. This brings to our attention that artificial intelligence still has some weaknesses that make us empathize with it when we verify that “to err is not only human” and that “only those who do nothing make no mistakes”. However, error continues to be the gateway to learning.

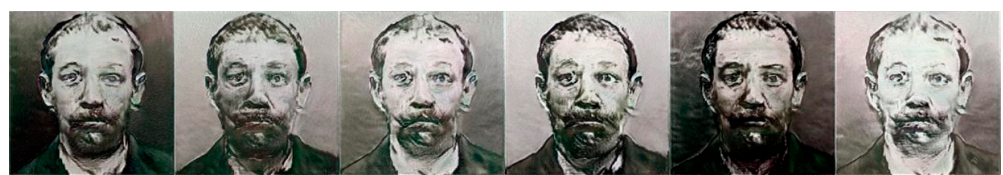


Figure 4. When the StyleGAN confronted with an insufficient sample, the model collapses and is unable to bring forth new proposals, generating the same image over and over again.

We have also had access to a collection of photographs of detained individuals taken by Alphonse Bertillon [25], a French biometrics researcher who worked for the Paris police and implemented methods of anthropological individuation. In 1880, he created a system known as bertillonage (“bertillonization”), based on the anthropometric measurements of the head and hands. The method was based on the fact that the bones of adult humans do not change and are different in each individual. Once the measurements of the detained individuals had been registered, he was classified and identified. The system proved its efficacy and, from 1882 onward it was included among the methods of the Paris police, followed soon by the authorities in other nations. However, its reliability was not complete, as the measurements were subject to human error. The system was criticized by other contemporary criminologists, such as Francis Galton, who confronted the shortcomings of bertillonization with the virtues of classifying by fingerprints. In Figure 5, we can see the generated portraits obtained by the StyleGAN (see Appendix A.2) from the collection of

Bertillon's portraits. It is clear that the algorithm does not only capture the morphology of the face but also the colors and texture of images in those times.



Figure 5. The generated portraits obtained by the StyleGAN from the collection of Bertillon's portraits. Reproduced with permission from J. Fontcuberta and P. Rosado.

If we had to wait for a random pixel ordering process to produce the image of a face by chance, we could wait for all of eternity without it happening. The purpose of StyleGAN is to detect patterns in the initial images (real portraits) and be able, by a process of automatic learning, to generate new images (false portraits) that turn out to be indistinguishable from the originals.

Exploring the limit of how far human expressions can be captured, we have put the StyleGAN to the test with a collection of 600 portraits of dead people who died of violent causes, extracted from a Mexican newspaper devoted to issues and incidents. In Figure 6, we can see how the system generates artificial portraits in which it is capable of capturing the essence of the last expression of the subjects, many of them with closed eyes, their mouths half open.

3.2. Beautiful Agony and La Petite Mort Projects

The nature of the human senses is analogical, meaning that they are in a constant state of variation. This is precisely the reason why words are so often insufficient to give a precise explanation of certain concepts. For example, the word 'orgasm' describes the sudden discharge of the tension built up during the cycle of sexual response, which varies in men and women, as it does with age and other circumstances. In the phase of arousal there is a point of no-return in which the will ceases to function and we are left in the hands of our biological algorithms. For some people, increase in their blood pressure and heart beat produces an effort of such magnitude that they may even lose consciousness for a brief period of time. The French language has an expression to describe such experiences for which the term "orgasm" is no longer sufficient: "la petite mort".



Figure 6. Generated portraits of corpses obtained by the StyleGAN. Reproduced with permission from J. Fontcuberta and P. Rosado (*La Petite Mort* project).

We have also used the StyleGAN in the *Beautiful Agony* project [26] to explore facial expressions during such sensorial moments. The face of a person experiencing an orgasm, instead of gazing sweetly into the eyes of their partner, takes on a brutish character. Researchers report that the gestures most frequently displayed by people when they are reaching sexual climax are regularly associated to facial expressions such as closed eyes, tightened jaws, wrinkled brow, and separated lips. In this specific moment there is an activation of many regions of the brain that are also activated when we feel pain, which explains why many of these expressions appear to be more of pain than pleasure. Our experiment of image generation proves that StyleGAN also reveals a pattern in the expressions of people in such states of ecstasy (Figure 7).



Figure 7. Generated faces of people experiencing an orgasm obtained by the StyleGAN. Reproduced with permission from J. Fontcuberta and P. Rosado (*Beautiful Agony* project).

As we were saying, the French expression “la petite mort” attempts to overcome the limitations of language in referring to a specific state in which the concepts of “death” and “orgasm” are indeed so close as to be almost indistinguishable. If we train the StyleGAN

with a set of photographs of faces of dead bodies and of faces at the moment of orgasm, the generative algorithm turns out images that reveal the patterns that pleasure shares with death. We could say that words leave empty spaces that generative neural networks fill with possibility Figure 8 shows a sample of *La Petite Mort* project [27].

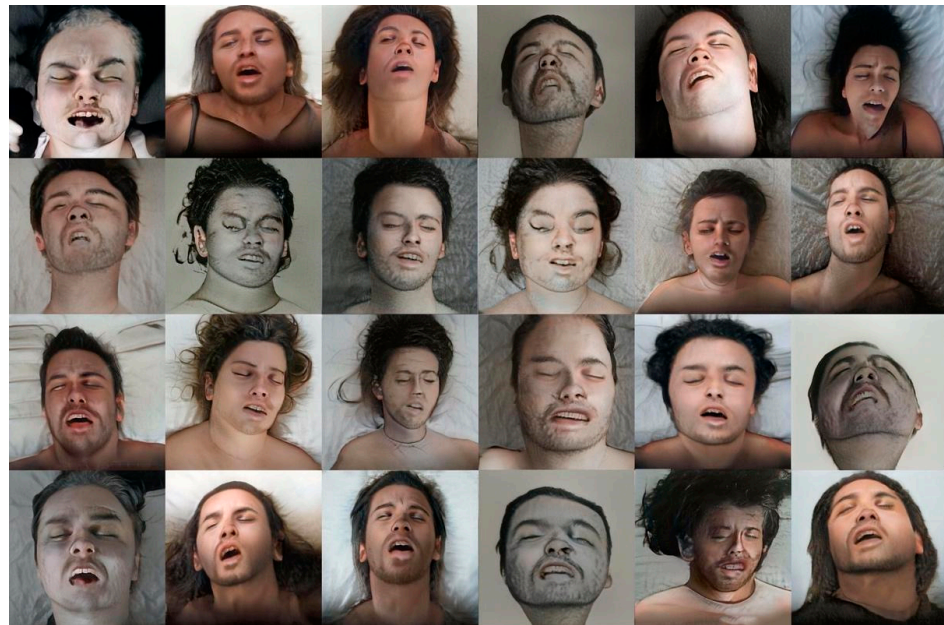


Figure 8. Images generated by the StyleGAN trained with faces of dead bodies and of faces at the moment of orgasm. Reproduced with permission from J. Fontcuberta and P. Rosado (*La Petite Mort* project).

4. Discussion

At present, with the excuse of social stability, security, and the control of cyberspace, the individual's rights and freedom are often overstepped by an excessive gathering of biometric information. Facial recognition algorithms trained to subsequently make decisions or carry out actions run the risk of having a bias and implementing discrimination unless great care has been invested in their design and the choice of data that their training is based on. Regarding 'artificial creativity', one of the challenges that must be solved is the risk of no longer being able to distinguish what is real from what has been created artificially (so-called 'deep fakes'). We cannot sidestep the dangers involved in the misuse of these technologies for the classification of human beings, but the fact is that intelligent algorithms are able to detect the visual patterns that human expressions are based on.

It may be a complex matter to differentiate between the expression of an alleged 'delinquent' and a 'celebrity', or between that of a 'corpse' and that of someone 'experiencing an orgasm'. Seeking an answer to how these processes crystallize in the face, we have put GANs to the test, exploring the limits of the capture of these expressions. The *Prosopagnosia* project addresses the ontology of photography by generating portraits of people who do not exist. The images created are the result of calculations, which sidesteps the classical binomial of eye and camera. The GANs base their generative task on the capacity to transform the visual characteristics of objects in the world into numbers. This is no easy task, for which we have had to develop extremely sophisticated algorithms; these obviously have their limitations; they offer us the possibility of describing semantic concepts by means of functions that unfold a manner of numerical gradient that overcomes the representational limitations associated to describing reality in other ways, such as natural language. The GANs detect the patterns that support the forms, generating a latent space (see Appendix A.3) that can be explored by following trajectories that bring within reach the infinite variations of formal possibilities there are between one number

and another. We can watch a continuous video sequence of images between death and orgasm generated by the model from the latent space, in the link [28]. Frontiers between categories dissolve, which is clear in the images of faces that illustrate this article. To finish, we would like to quote a much-used cliché, generally with a critical tone: “They are making us into numbers” (quite often, indeed, the comment refers to computational algorithms). In a numerical context, many of the traces we leave as we wander the world—words, sounds, images and others—are put on the same level. Computers can detect latent aspects shared by very different fields and give us an alternative point of view from which to think our complex universe.

The designers and programmers of robots are giving them advanced interaction systems such as natural language, body language, the capacity to memorize situations, locations and people, and algorithms are being developed that emulate a variety of human emotions like sadness, joy, fear, etc. In these circumstances, we may well ask whether, beyond invigilating us, intelligent algorithms can help us to speculate on a possible future where androids may even emulate our basic instincts. At this point, however, perhaps the crucial question is not if androids will someday be able to surrender to the lovely agony of an orgasm. The key question is how we can use this new language, this new way of thinking concepts in the form of numerical gradients, offered to us by the latent space of possibilities generated by GANs, to expand the limits of our present way of thinking, fundamentally based now on linguistic structures. Instead of answers, the images generated give us a new perspective on otherness. GANs determine an essential identity and generate hybrid predictions, speculations regarding genders and categories that natural language would be unable to explain easily. Latent space constitutes an extraordinary tool with which to extend our communicative skills, adapting them to visual concepts better than is possible with textual language.

5. Conclusions

GANs models are not completely autonomous creative systems, they are not able to formulate an intention and assess their own results. The dataset the GAN is trained on is the key to its creativity. In this way, by feeding a pre-trained StyleGAN with only 600 images of different types of facial expressions, we have achieved convincing synthetic images. Our research reinforces future uses of StyleGAN in artistic practice, without requiring complex face editing algorithms.

If we focus on learning for perfection, we will miss other interesting outcomes for artistic creation processes. In this article, we investigated the potential of GAN’s latent spaces to encode human expressions, highlighting creative interest for suboptimal solutions rather than perfect reproductions, in pursuit of the artistic concept. From the perspective of artistic creation, the application of GAN technology opens up new formal and conceptual imagery. We present results which diverge from standard usage of GANs with the specific intention of producing portraits that may assist us in the representation and recognition of otherness in contemporary identity construction.

Author Contributions: Conceptualization, P.R.; Methodology, R.F. and F.R.; Software, R.F.; Validation, P.R., R.F. and F.R.; Writing—review and editing, P.R.; Supervision, P.R., R.F. and F.R.; Funding acquisition, P.R. and F.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC was funded by the Ministry of Science, Innovation and Universities, grant number PID2019-104830RB-I00 and by the Government of Catalonia, grant number 2017 SGR 622.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request due to privacy restrictions.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Appendix A.1. Generative Adversarial Networks

Training a GAN is a far from trivial task because it requires a very large number of images to arrive at a correct adjustment of lots of parameters, and the computational cost must be added to this.

On the basis of the original GAN [1], different versions, better adapted to image analysis, were devised. Radford et al. [2] proposed a new type of neural network that also combines antagonistic training with convolutional networks that improved any of the already existing systems. These improvements consist of a more stable training process and the ability to generate quality images while being simpler in its structure.

The training of the DCGAN is conducted in two successive processes that are iterated. At the beginning, the candidate images produced by the generator are basically just random noise. At this point, the discriminator's task could hardly be easier, because the real images are labelled as 'real' and the generated images as 'fake', and are, as we said, just noise. In the first process (the parameters of the discriminator having been learned and fixed) the generator then starts to label all the false images it produces as 'real', though they still bear little or no resemblance to the genuine images. The discriminator now begins to reject candidates it evaluates as 'fake', while the generator tries to trick it by labelling them as 'real'. To go beyond this situation in this first process the DCGAN can only readjust the parameters of the generator (the parameters of the discriminator are fixed, of course) to produce false images that more closely resemble the real ones. In this way, the generator has already made some progress in learning how to generate real images.

In the second process the discriminator receives correct information, with the generated images tagged as 'fake' and real images from the database labelled as 'real', so that the discriminator can also learn and adjust its parameters. At this stage the discriminator's task of evaluation is somewhat more difficult than in the beginning because the quality of the false images has improved a little. The DCGAN now comes back into action and tries to fool the discriminator by labeling false images as 'real', as described in the previous paragraph. The processes are repeated until the discriminator is incapable of distinguishing generated images from genuine ones (Figures A1 and A2).

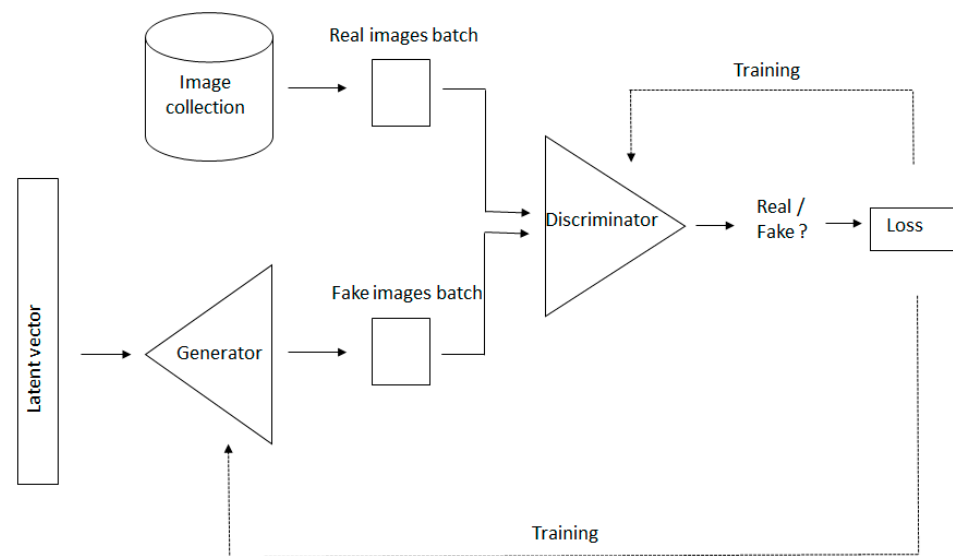


Figure A1. We show a DCGAN scheme. The basic components of a DCGAN are two convolutional neural networks: a generator that synthesizes new images from scratch, and a discriminator that takes images from both the training data and the generator's output and predicts if they are 'real' or 'fake'.

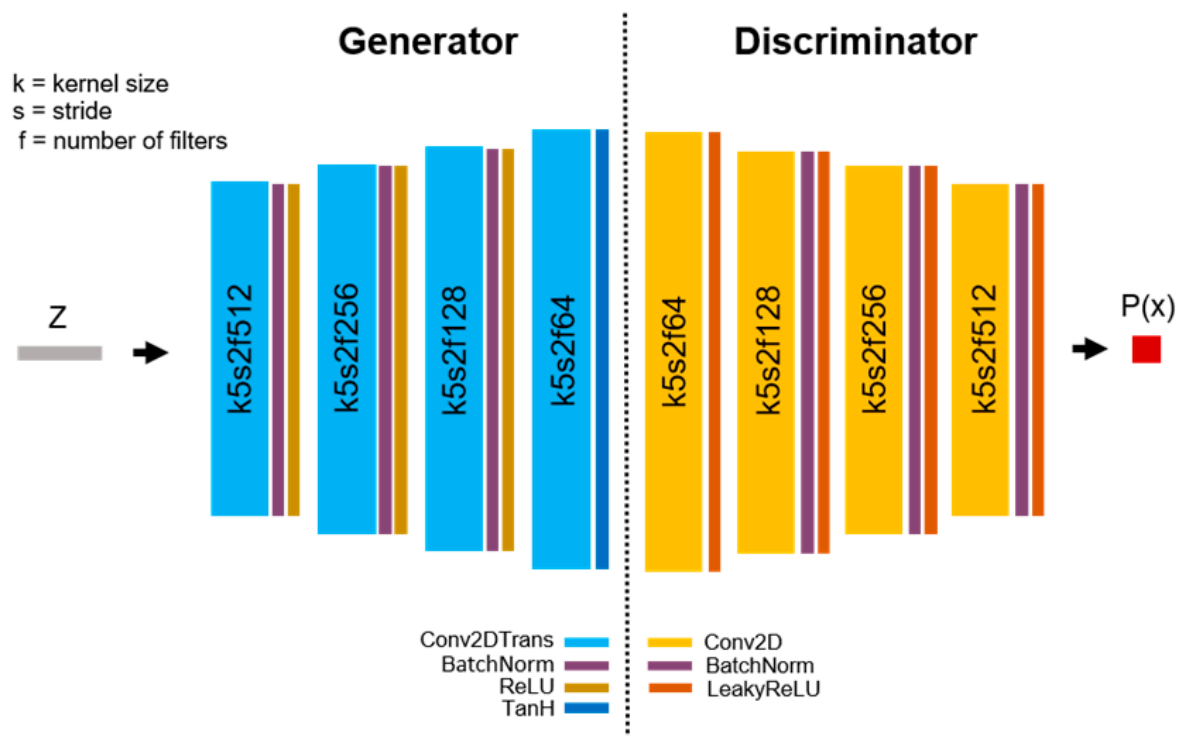


Figure A2. We see a common convolutional DCGAN architecture: four stacked convolutional layers in both the generator and discriminator. The input z denotes a vector in the latent space, the output P denotes the probability that the image x is real or generated.

Appendix A.2. StyleGAN Training

In 2019, Karras, Laine, and Aila [9–11] described a new training methodology for GAN (StyleGAN): the main idea is to grow both the generator and the discriminator progressively, starting from a low resolution and adding new layers to model increasingly fine detail as training progresses.

The StyleGAN [10] enhances ProGAN [9]. Specifically, it introduces an update to the Generator, keeping the same type of Discriminator. ProGAN generates high-quality images but its ability to control specific features of the generated image is very limited. Features are entangled and therefore attempting to tweak the input, even a bit, usually affects multiple features at the same time.

The improvements are focused on the generator (Figure A3). In StyleGAN, the generator is made up of two networks: the mapping network and the synthesis network. The mapping network reduces entanglement by mapping latent vectors to intermediate vectors. Intermediate representation is learned using full connected layers (FC).

With network synthesis, the style is transferred, integrating at each resolution level, the intermediate vector with the output of the convolution (CONV) in the previous level. An adaptive normalization (AdaIN) module is applied to transfer the encoded information created by the mapping network, into the generated image. This defines the visual characteristics at each level of resolution. Furthermore, StyleGAN authors provide network synthesis with explicit noise inputs, as a direct method to model stochastic details. Noise inputs consist of single-channel images based on uncorrelated Gaussian noise. They are fed as an additive noise image to each layer of the synthesis network. The single noise image is injected to all feature maps.

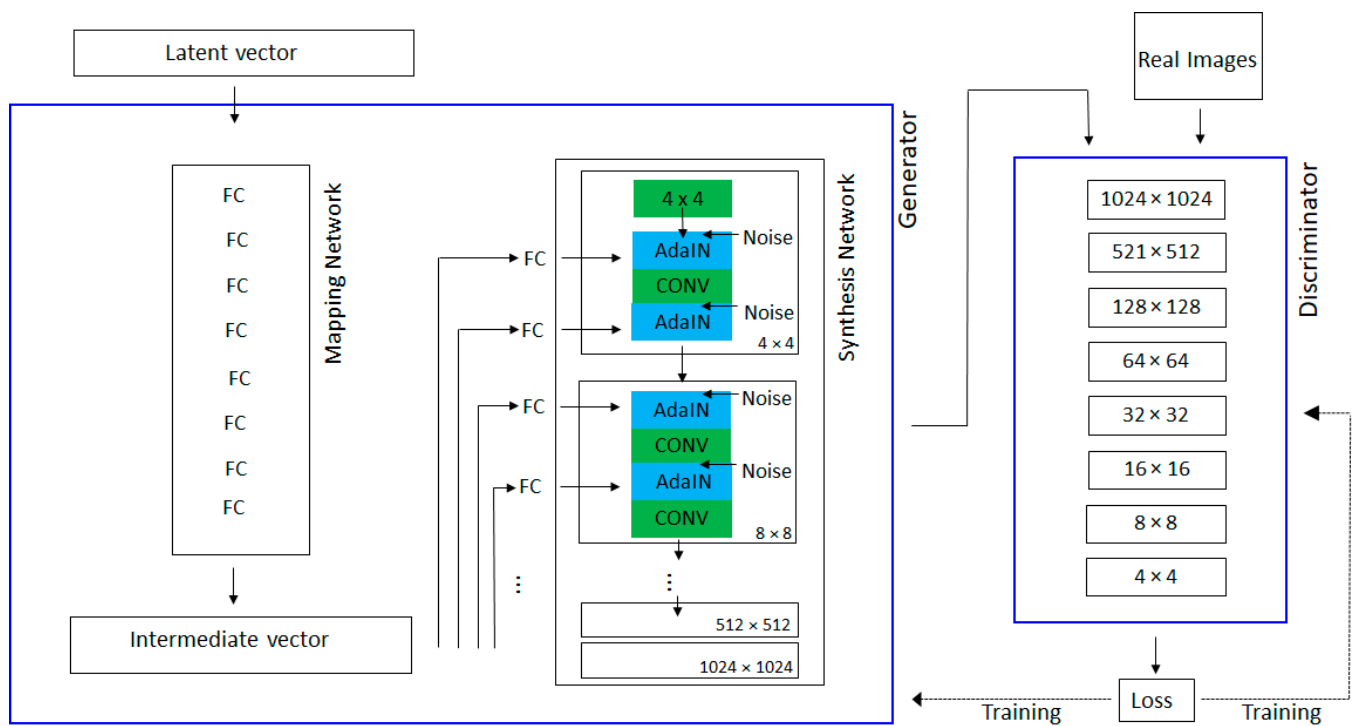


Figure A3. StyleGAN scheme. The StyleGAN generator's is made up of two networks: the mapping network and the synthesis network. The discriminator network exhibits a common architecture having stacked convolutional layers.

To obtain the generated portraits of this article, we have feed a StyleGAN pre-trained by NVlabs on the Flickr-Faces-High Quality dataset (FFHQ). FFHQ is a high-quality image dataset of human faces. The dataset consists of 70,000 high-quality PNG images at 1024×1024 resolution and contains considerable variation in terms of age, ethnicity, and image background. It also has good coverage of accessories such as eyeglasses, sunglasses, hats, etc.

We have continued training the pre-trained StyleGAN with 600 portraits from each collection under study for one week using eight Tesla V100 GPUs. It is implemented in TensorFlow and the repository can be found at this link: <https://github.com/NVlabs/stylegan> (accessed on 2 July 2020).

Appendix A.3. Latent Space Representation of Our Data

After training, we obtained a function that maps a point in the latent space to the high-dimensional space of all images (Figure A4). The latent space representation of our data contains all the important information needed to represent the features of our original data. Each point of the latent space is the high-dimensional representation of an image. The power of representation learning is that it actually learns which features are most important for it to describe the given observations (images) and how to generate these features from the raw data. We can sample points from this latent space to generate 'new' data.

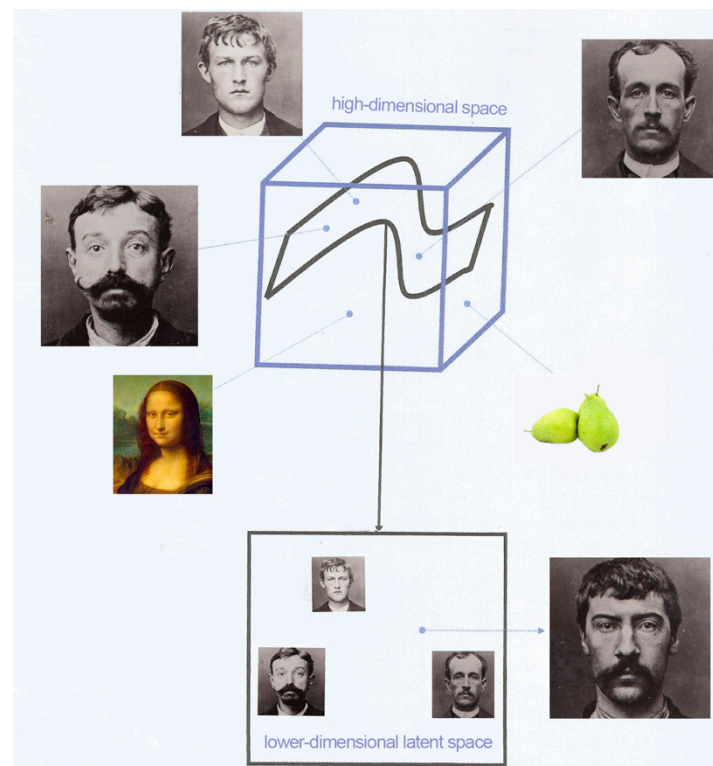


Figure A4. Cube represents the high-dimensional space of all images; representation learning tries to find the lower-dimensional latent space on which particular kinds of image lie (for example, Bertillon's portraits).

References

1. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. *Adv. Neural Inf. Process. Syst.* **2014**, *2*, 2672–2680.
2. Radford, A.; Metz, L.; Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. Available online: <https://arxiv.org/pdf/1511.06434.pdf> (accessed on 20 August 2021).
3. Wang, K.; Gou, C.; Duan, Y.; Lin, Y.; Zheng, X.; Wang, F.Y. Generative adversarial networks: Introduction and outlook. *IEEE/CAA J. Autom. Sin.* **2017**, *4*, 588–598. [CrossRef]
4. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [CrossRef]
5. Hong, Y.; Hwang, U.; Yoo, J.; Yoon, S. How generative adversarial networks and their variants work: An overview. *ACM Comput. Surv.* **2019**, *52*, 1–43. [CrossRef]
6. Wang, Z.; She, Q.; Ward, T.E. Generative Adversarial Networks: A Survey and Taxonomy. Available online: <https://arxiv.org/pdf/1906.01529.pdf> (accessed on 15 August 2021).
7. Gui, J.; Sun, Z.; Wen, Y.; Tao, D.; Ye, J. A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications. Available online: <https://arxiv.org/pdf/2001.06937.pdf> (accessed on 25 August 2021).
8. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 5967–5976.
9. Karras, T.; Aila, T.; Laine, S.; Lehtinen, J. Progressive Growing of GANs for Improved Quality, Stability and Variation. Available online: <https://arxiv.org/pdf/1710.10196.pdf> (accessed on 10 August 2021).
10. Karras, T.; Laine, S.; Aila, T. A Style-Based Generator Architecture for Generative Adversarial Networks. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 4396–4405.
11. Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; Aila, T. Analyzing and improving the image quality of StyleGAN. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 8107–8116.
12. Cheng, G.; Sun, X.; Li, K.; Guo, L.; Han, J. Perturbation-Seeking Generative Adversarial Networks: A Defense Framework for Remote Sensing Image Scene Classification. *IEEE Trans. Geosci. Remote. Sens.* **2021**, 1–11. [CrossRef]
13. Qian, X.; Cheng, X.; Cheng, G.; Yao, X.; Jiang, L. Two-Stream Encoder GAN with Progressive Training for Co-Saliency Detection. *IEEE Signal Process. Lett.* **2021**, *28*, 180–184. [CrossRef]

14. Karras, T.; Aittala, M.; Laine, S.; Härkönen, E.; Hellsten, J.; Lehtinen, J.; Aila, T. Alias-Free Generative Adversarial Networks. Available online: <https://arxiv.org/pdf/2106.12423.pdf> (accessed on 8 August 2021).
15. Härkönen, E.; Hertzmann, A.; Lehtinen, J.; Paris, S. Ganspace: Discovering Interpretable GAN Controls. Available online: <https://arxiv.org/pdf/2004.02546.pdf> (accessed on 3 August 2021).
16. Liu, Y.; Li, Q.; Sun, Z. Attribute-aware face aging with wavelet-based generative adversarial networks. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 11869–11878.
17. Broad, T.; Leymarie, F.F.; Grierson, M. Network bending: Expressive manipulation of deep generative models. In *Artificial Intelligence in Music, Sound, Art and Design*; Romero, J., Martins, T., Rodríguez-Fernández, N., Eds.; Springer: Cham, Switzerland, 2021; pp. 20–36.
18. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2017; p. 645.
19. Manovich, L. Defining AI Arts: Three Proposals. Available online: <http://manovich.net/content/04-projects/107-defining-ai-arts-three-proposals/manovich.defining-ai-arts.2019.pdf> (accessed on 1 September 2021).
20. Boden, M.A. *Inteligencia Artificial*; Turner Noema: Madrid, Spain, 2017.
21. Berns, S.; Colton, S. Bridging Generative Deep Learning and Computational Creativity. Available online: <http://computationalcreativity.net/iccc20/papers/164-iccc20.pdf> (accessed on 9 October 2021).
22. Fontcuberta, J.; Rosado, P. *Prosopagnosia*; RM Verlag: Barcelona, Spain, 2019.
23. Checa, A. *Prensa y Partidos Políticos Durante la II República*; Universidad de Salamanca: Salamanca, Spain, 1989.
24. Miller, A.I. *The Artist in the Machine: The World of AI-Powered Creativity*; The MIT Press: Cambridge, MA, USA, 2019; pp. 90–91.
25. Ruiza, M.; Fernández, T.; Tamaro, E. Biografía de Alphonse Bertillon. Available online: <https://www.biografiasyvidas.com/biografia/b/bertillon.htm> (accessed on 15 December 2020).
26. Fontcuberta, J.; Rosado, P. Beautiful Agony. In *Par le Rêve. Panorama 23*; Kaepelin, O., Ed.; Lefresnoy: Tourcoing, France, 2021; pp. 80–81.
27. Rosado, P. La Petite Mort: Simulación artificial de la expresión de un orgasmo. In *Cuerpos Conectados. Arte, Identidad y Autorrepresentación en la Sociedad Transmedia*; Baigorri, L., Ortuño, P., Eds.; Dykinson: Madrid, Spain, 2021; pp. 79–88.
28. Fontcuberta, J.; Rosado, P. Prosopagnosia, Link to the Video on Vimeo. Available online: <https://vimeo.com/495021778> (accessed on 27 July 2020).