PseudoGAN.
A conceptual software for generative art

Prof. Umberto Roncoroni, MSc, PhD
Department of Design, Universidad Peruana de Ciencias Aplicadas, Lima, Perú
umbertoluigironcoroni@gmail.com
umberto.roncoroni@upc.pe

Abstract

Introduction

Generative Artificial Intelligence solutions provide astounding audiovisual results in digital art, illustration, videogames and more [1, 2]. But these achievements come with many challenges.

The first problem with generative AI is that issues arise from the lack of a proper connection between digital and traditional culture, the clash between Romantic aesthetics, and the drawbacks of AI systems, including biased training data, flawed model assumptions, and lack of transparency [3, 4, 5]. Furthermore, many esthetics, artistic, scientific, educational, and ethical question are still unresponded [6, 7]. These questions are related to computational creativity, computer science and digital divide issues. Leaving apart phylosophical positions, developing the pseudoGAN software project, I was interested in three specific topics: a) the aesthetic and educational limits of generative AI such as GANs; b) the drawbacks of the standardized solutions offered by IA monopolies; c) the importance of original technology development and Open Source Software, making it available for use, modification, and distribution in artistic and educational contexts.

The hipotesis is that achieving an inclusive, sustainable, and creative use of digital media, in general but especially for generative art requires the development of original proprietary digital technology, grounded in cultural identities, without black boxes and the biases of data models. The main task was to emulate machine learning and neural networks' results, and the unpredictable morphological interpolations made popular some years ago by Google DeepDream (Fig. 1). Why to do all this? Basically, I do not like the idea of AI doing all the intersting work of artistic research and leaving the artstic experience to machines.

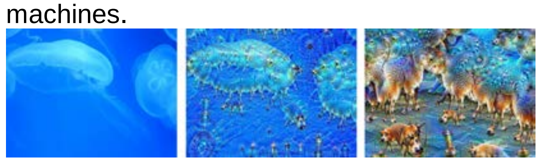

Figure 1. Google DeepDream interpolations. Source: https://en.wikipedia.org/wiki/DeepDream

This research is based on a literature review, software design, and artistic practice, and it highlights the importance of the humanities, ethno-computation, and analogical processes and materials. Starting from the images generated by AI algorithms such as GAN, we investigated innovative solutions to interpolate images or generate morphing processes. To do this, during the research project were experimented alternative algorithms using only two source images, simple well known image processing techniques, without arcane mathematics or bigdata references. This application is a sort of conceptual software: even if it can be used for practical applications, such as creative image processing or video effects, it enlightens concepts about Artificial Intelligence art, creative software, digital art, inclusive education, and multicultural technological development.

1. PseudoGAN's software design principles

In the first place, my point is that the substantial contributions to artistic theory and practice are not images, animations or other audiovisual artworks, but software design and coding. The second important concept here is that the important learning process is the user's learning process, not the machines. This can guarantee real creative autonomy and originality, since the artist-user stays in full control of the process and its development. On the other hand, there are interactivity issues that should be revised, following Habermas communicative action statements [8]. Thus, I followed these guidelines to match technical and communicative action requirements (*in italic*):

a) Clearly organize the interactive workflow (*achieving mutual understandings*).
b) Explain every step of the process, withoyut the need to open help windows (*feedback*).
c) Help the user to avoid possible computational mistakes, like divisions by zero, bad protocols, or improper data formats (*intelligibility*).
d) Solve interactivity issues accomplishing, like transparency of knowledge with Open Source Software (*sharing knowledge, democracy*).
e) Simplify the algorithms and resource demands, working only with 2 images a a time, no big data, no statuistics, and almost no math (*intelligibility*).

I will enter into development details in the following chapters.

2. Interface and functions

Form the point of view of interface design and communicative action, the goal was to expose the processing sequence and making intuitive the connections between parameters. Standard interface componentes are designed to be similar to neural networks' nodes and weights (fig. 2), but clearly explained, parametrized, and visualized (communicative actions' transparency and equality requirements).

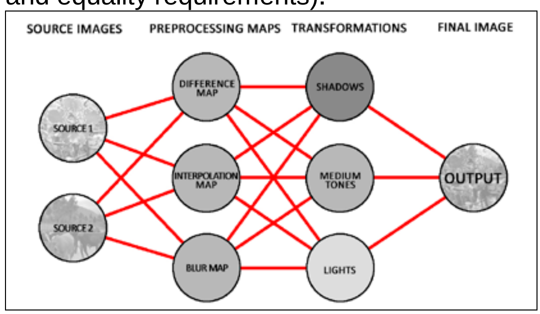

Figure 2. PseudoGAN neural network metaphor example. Work in progress.

To do this, the pseudoGAN software architecture was split into different modules and the elaboration process split into three main steps, organized in the proper sequence in the interface: preprocessing, morphing processes, and

postproduction.
I took great care to provide effective feedback to the user, this is necessary also when developing and experimenting with the software development.
I choose to forget the best effciency, like speed, in favor of the legibility of code to facilitate editing, improvements and experimentation. The writing of code is an important part of multiauthoring and of the generative artistic experience.

Figure 2. Interface design basic design and concepts. Frome left to right: input and output functions, preprocessing, processing and postproduction. Still in development.

2.1 Preprocessing

The first task of the process is to prepare images that will be used as masks in different ways in the mixer module (fig. 3). The preprocessing process consists of three steps: 1) create a difference map that is a grayscale image mapping the differences between pixels of two source images, 0 means the pixels are equals, 255 100% different; 2) create the interpolation map blending the source images, values are assigned using percentuals of the sum of the values of the two source images, when the sum is 0, gets 100% of the first source image, when 128, 50% of source image 1 and 50% of source image 2, when 255 100% of source image 2. Pixels between these values gets different percentuals; 3) create the blur map, that is a sort of antialising of the two source images. More maps can be added using different criteria. These maps are used as references for mixer’s parameters modulation and act like neurons in a neural network, the learning process consists in adjusting interactively their weights (fig. 2).

Fig. 3. From left to right: source images, difference map, interpolation map, blur map.

2.2 Transformations and morphing processes

The basic concept of the morphing process is to interpolate a selection of source images’ similar pixels, filling the spare pixels with some interpolation between the 2 source images. This is done for shadows (Fig 3a) and lights (Fig 3b). Similar pixels expand in the bitmap like a wet into wet effect of watercolor paintings. Then lights are added to shadows (fig 3c) using different algorithms and parameters’ values. Finally, a gray map (pixels around 110-138 brightnmess values) is used to refine details in the middle tones.
Maps and shadow-light interpolation have interactive menus to change the algorithms’ behaviors. In this way is possible to generate different versions of the same process (functions can be programmed to generate animations, video effets and more). These imagens will be used in the mixer to get the final effect.

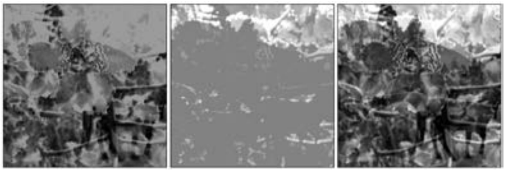

Fig. 3. From left to right: source images, difference map, interpolation map, blur map.

2.3 Transformations and morphing processes

The basic concept of the morphing process is to interpolate a selection of source images' similar pixels, filling the spare pixels with some interpolation between the 2 source images. This is done for shadows (Fig 3a) and lights (Fig 3b). Similar pixels expand in the bitmap like a wet into wet effect of watercolor paintings. Then lights are added to shadows (fig 3c) using different algorithms and parameters' values. Finally, a gray map (pixels around 110-138 brightnmess values) is used to refine details in the middle tones.

Maps and shadow-light interpolation have interactive menus to change the algorithms' behaviors. In this way is possible to generate different versions of the same process (functions can be programmed to generate animations, video effets and more). These imagens will be used in the mixer to get the final effect.

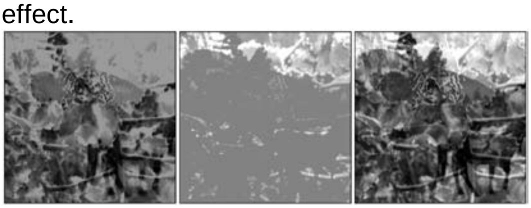

Fig. 3. From left to right: shadows and lights interpolation, union of shadows and lights interpolations.

2.4 Mixing

When interpolations are finished, the user has a set of different images that can be remixed like tracks in electronic music production. Using the mixer module, it is possible to equalize gray levels or colors, compress values, use masks (difference map, blur map, etc.), and apply filters with interactive parameters that can be animated modifing the effect parameters values in time. Frames will be automathically generated and saved in an Animation folder.

The user can apply neural networks feedback concepts adjusting the parameters (weights) or changing the order of the function in the interpolation sequence, until the right is achieved

2.5 Results

As explained before, the project's goal is to develop an application to create an image from source images which parts melts together to generate unexpected forms, like generative AI or GAN do using databases of thousands of images. Obviously, the final effect depends on the characteristics of the source images, AI does not understand the meaning of the images or of their parts; in fact, it is using statistic analysis that generative Ais can "select" the best matching forms and, adjusting the weights of the nodes (the "learning" process super or unsupervised), combine them realistically [9].

On the contrary, I purposedly use only two source images, and decided to tune the parameters manually and without complex mathematics [reference]. This solution puts the artist in command of the weight tuning process [10]; the learning process then becomes something creative, since the user must experiment to match images that can be completely different, as shown if figure 4

Figure 4. Details of source images and of the interpolation result. It is a nice surprise if images match in some ways.

A useful trick is to work some interpolations using a low-resolution copy of the images; this not only speed up the process but can improve the effects of the filters in many ways, like the sumi-e effect of the topo left image of figure 5.

The interface design must provide enough flexibility and variety to let the user experiment freely and switch easily between images, maps and filters. The interpolations' sequences and functions' parameters can be changed anytime interactively; thus, the user can create his own processing scripts.

Figure 5. Gray scale and color images using different resolutions and sources. R, G and B channels are processed in the same way of grayscale values. The RGB's process is still under development.

3. Discussion and conclusions

To finish, I will resume the ideas and findings of the research developed in the previous sections, and terminate with some general considerations about software devlopment, artistic practice, and generative art, computational creativity, education, and digital culture in general. It seems important to mention that academic literature seems not too much interested in cultural or esthetic issues of AI, with few exceptions [11, 12, 13].

Anyway, as said in the introduction, the starting point was to analyze and hack GANs processes, to get more computational creativity's freedom, under two main assumptions: first, that using AI the artist is the AI (or it's programmer or computer scientist), and not the user [10]; second, that, observing GANs results, it is clear that these are standardized combinations of existing materials, which in the long run always repeate the same formal characteristics.

Starting from the technical point of view, the differences with generative AI and GANs are the followings:

a) Almost everything that generative AI and GANs do can be succeeded using standard image processing techniques.
b) PseudoGAN's effect is not always realistic, since the interpolation statistics is just between two images, while machine learning checks thousands of images to catch the best matches between pixels.
c) Parameters feedback and backpropagation between functions (like the weights of the nodes of the neuarl network) are in the user's hands, to trigger and stimulate his creativity.

d) Algorithms are transparent, no blac boxes of any kind, so interactivity is real, not just smashing buttons or playing with prompts.

e) Instead of generic and global references base on bigdata, pseudoGAN's users can acces local and contextualized information and processes; these assets include local cultural and technological assets, and artistic and natural environments references; these can be embedded in the algorithms and inspire the users' scripts.

f) Concerning future developments and research, I will do more research to improve the interface and software architecture, refine color interpolations, develop more interpolations options, experiment new algorithms for image processing, like textures, adding more pixels selection patterns, not just single pixels as done in this pseudoGAN version, and finally develop more macros and animation tools.

From the methodological point of view, we consider important to mention:

a) Generative AI artworks are oftent formal excercises with little impact or even negative impact on real contexts, due to the limitations wwe have discussed. On the contrary, developing propietary solutions can be significant in specific environments, because references relate to their cultural identities, and the solutions developed can be shared as educational processes designed for the specific artistic research context [12].

b) The development of propietary and original digital technology is paramount for educational, digitasl divide and digital neocolonization issues [14]. It reduces the power of digital monopolies.

c) Generative AI appears extremely flexible and unpredictable. This is an improper attribution, that is valid only speaking of the artwork as an object, that in the light of postmodern esthetics, overproduction, language saturation and so on is antiquated. And, considering processes and concepts, where hipertextual autorship develops, the AI's esthetic framework is repetitive and outdated.

And finally, I will pose some conceptual issues that are important to mention for education and artistic commitments. Some authors, critics or developers have doubts about the possibility that IA could find esthetic solutions that human can`t imagine, but agree that AI can help study and develop human creativity [11] is false, since the model of IA is the human brain, its algorithms and criteria are human. Considering that, as Plato said, an imitation is always less than the original, AI's models are dangerous. For this reason, pseudoGAN software was developed keeping in mind the human artistic practice and creativity (even if in the scope of a specialized and small solution).

Considering all this, a fundamental question arises: why train and use AI when we can train and learn ourselves? The complexity of human artist experience is something a lot more significant and creative, that AI technologies are stealing this from us. The abuse of IA reduces and turns off the efforts to undertake original research in generative art, and explore different solutions

References

[1] Wolfram, S. (2023). Will AIs Take All Our Jobs and End Human History—or Not? Well, It's Complicated… *Stephen Wolfram Writings*. https://writings.stephenwolfram.com.

[2] Cetinic, E., & She, J. (2021). Understanding and Creating Art with AI: Review and Outlook. arXiv:2102.09109. https://doi.org/10.48550/arXiv.2102.09109.

[3] Larson, E. (2021). The myth of artificial intelligence. Belknap Press.

[4] O'Neil, C. (2016). Weapons of math destruction. Crown.

[5] Taddeo, M.R. (2023). L'impatto e l'accelerazione dell'Ai nella nostra vita tra scienza, etica e tecnologia. https://lutech.group/it/lutech-talks-milano-2023-intelligenza-artificiale

[6] Leonard, N. (2021). Emerging Artificial Intelligence, Art and Pedagogy: Exploring Discussions of Creative Algorithms and Machines for Art Education. Digital Culture & Education, 13(1), pp. 20-41.

[7] Yusa, I. M. M. ., Yu, Y. ., & Sovhyra, T. . (2022). Reflections on the use of artificial intelligence in works of art. Journal of Aesthetics, Design, and Art Management, 2(2), 152-167. https://doi.org/10.58982/jadam.v2i2.334

[8] Habermas, J. (1990). Moral consciousness and communicative action. MIT Press.

[9] Wyse, L. (2019). Mechanisms of artistic creativity in deep learning neural networks. En Proceedings of the International Conference on Computational Creativity. Charlotte, NC, EUA, junio de 2019.

[10] Sofian, A. e Ippolito, J. (2019). Can Artificial Intelligence make art without artists? Ask the viewer. Arts, 8(35), 245-253. doi:10.3390/arts8010035

[11] Boden, M. A. (2009). Computer models of creativity. AI Magazine 30(3): pp. 23–34.

[12] Mazzone, M. y Elgammal A. (2019). Art, creativity, and the potential of artificial intelligence. Arts, 8(1) 26. doi:10.3390/arts8010026

[13] Morris, J., & Paris, L. (2022). Rethinking arts-based research methods in education: Enhanced participant engagement processes to increase research credibility and knowledge translation. *International Journal of Research and Method in Education*, 45(1), 99–112.

[14] Ponzanesi, S. (2021) Postcolonial intellectuals: new paradigms. Postcolonial Studies, 24:4, 433-447, DOI: 10.1080/13688790.2021.198