June 2022

Poster: Art Installation "Mirror, Mirror" featuring
Deepfake and Neural Style Transfer Technologies [1]

Anastasia VANDEN BERGHE [a] Fatima Zahra FATHI [b] and Sana NOUZRI [b]
[a] *University of Luxembourg*
[b] *AI ROBOLAB, University of Luxembourg*

Abstract. The robotic art installation "Mirror, Mirror" is a result of a creative collaboration between an artist and an engineer. The aim of this project is to transform a person's image and background in order to prompt them to examine their version of self and the surrounding reality and make them aware of the developments in the field of social robotics and Artificial Intelligence.

Keywords. Art, Artificial Intelligence, Deepfake, Neural Style Transfer, Social Robotics

1. Introduction and objectives

The inspiration behind the interactive robotic art installation "Mirror, Mirror" is a classic folk tale of "Snow White," published in 1812 in "Grimm's Fairy Tales." In today's world, social media plays the role of the proverbial magic mirror that shows reality. With the help of this art piece, the artist hopes to test the public's fundamental assumption that reflections never lie to us. In this project, the screens serve as magical mirrors, that transform the visitor's reflection with Artificial Intelligence (AI) using computers, sensors, and actuators that respond to audience interactions and are being influenced by the participant's input. The project is displayed at the University of Luxembourg's AI ROBOLAB projects during ESCH 2022 at the AIART Pavilion.

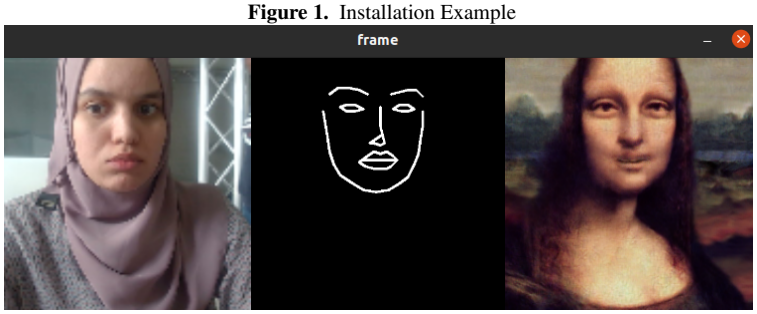
Through our installation, we hope to give visitors a sense of what it's like to be part of the creative art process by turning deep video using neural style transmission into a work of art based on a real painting created by an artist. This experience aims to create a deeper connection to art and make it a fluid and enjoyable experiment for the viewer. In addition, we seek to bring social awareness about the development of deepfake technology to the public. We hope to spark a debate about the role of the new technology in people's daily lives by showing how, using a properly trained model, a program can take an image of a person and convert it into a deepfake video or an NST altered image in only a few seconds. Those developments can have long-lasting effects on all aspects of society, whether it is in the areas of education, psychology, medicine, arts, or many others.

[1] Luxembourg University, AI ROBOLAB

2. Installation description

For the art installation “Mirror, Mirror”, we used three screens each representing a different aspect of the project. The first scenario, titled “Living portraits” was constructed using a convolutional neural network, a type of AI that analyses and interprets images similar to the human brain. These digital copies were previously created using a technique known as generative adversarial network (GAN), in which an AI attempts to create a convincing image. As a result, the viewer becomes a living portrait that they can control with their facial expressions, as presented in figure 1. In the second scenario, the “Deepfake model”, a video sequence is used in which an object in a source image is animated in response to the motion of a driving video. A person’s image is taken and transformed into an animated video that is subsequently shown to them. The last scenario, “Neural style transfer”, demonstrates an optimization technique used to take two images – a content and a style reference image (in this case a painting by a painter) – and blend them together so that the output image resembles the content image but in a style of the reference image.

Figure 1. Installation Example



3. Technical aspect

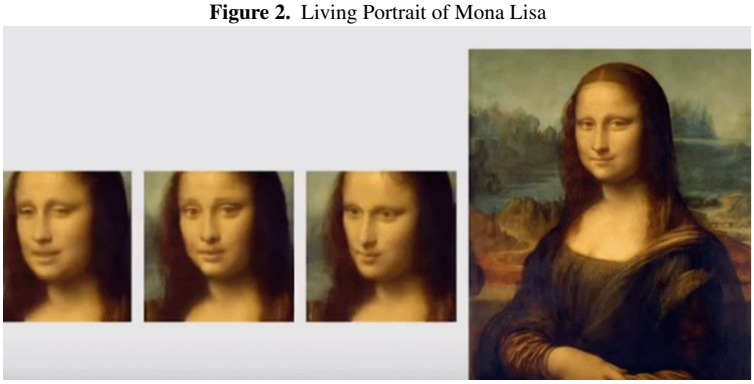
In this section, we will discuss three main techniques employed in this project: Living portraits, deepfake, and NST. Firstly, we will discuss the technicalities of living portraits regarding AI. In the second part, we will address the deep fake mechanism and its own architecture, then we will discuss the NST mechanism, and finally the project architecture.

3.1. Living Portraits

According to Live Science, generative adversarial networks, or GANs, can be used to make video footage or deepfakes [1]. A GAN is made up of two AI systems that work together to create content that looks disturbingly identical to the real thing [2]. The generator, for example, makes a replica of an image from new data. The discriminator is another mechanism that determines if the data is real or fraudulent. Both technologies will learn to make images that are extraordinarily lifelike and impossible to identify from the originals over time [1].

The major technique used in this scenario to construct photorealistic talking heads is direct (warping-free) synthesis of video frames using adversarially-trained deep convolutional networks (ConvNets) [3]. With this strategy, we must train enormous networks, with tens of millions of parameters for each talking head in both the generator and discriminator. To develop a new personalized talking head model, as presented in figure 2, these systems require a several-minute-long video or a big dataset of portrait images, as well as hours of GPU training [4]. While this work is less than that required by systems that build photo-realistic head models using advanced physical and optical modeling, it is nevertheless valuable in our artistic settings, where we want users to be able to design their own individualized head models with as little effort as possible.

Figure 2. Living Portrait of Mona Lisa



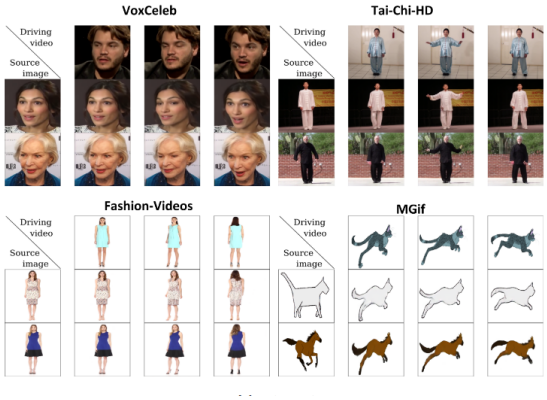
3.2. Deepfake

The process of picture animation requires a video sequence in which an object in a source image is animated in response to the motion of a driving video [5]. To make a deepfake video, it is necessary to have the availability of a ground truth pose and form data during training. Our system addresses this need without requiring any annotations or prior knowledge of the item to be animated [5]. Our approach can be used with any object of this class after being trained on a set of movies portraying objects of the same category (e.g. faces, human bodies), as presented in figure 3. We do this by employing a self-supervised approach to dissociate appearance and motion information [6]. We utilize a representation consisting of a set of learned key points and their local affine transformations to allow complex motions. The appearance taken from the source image and the motion produced from the driving video is combined in a generator network to mimic occlusions that occur during target motions [6].

3.3. Neural Style Transfer

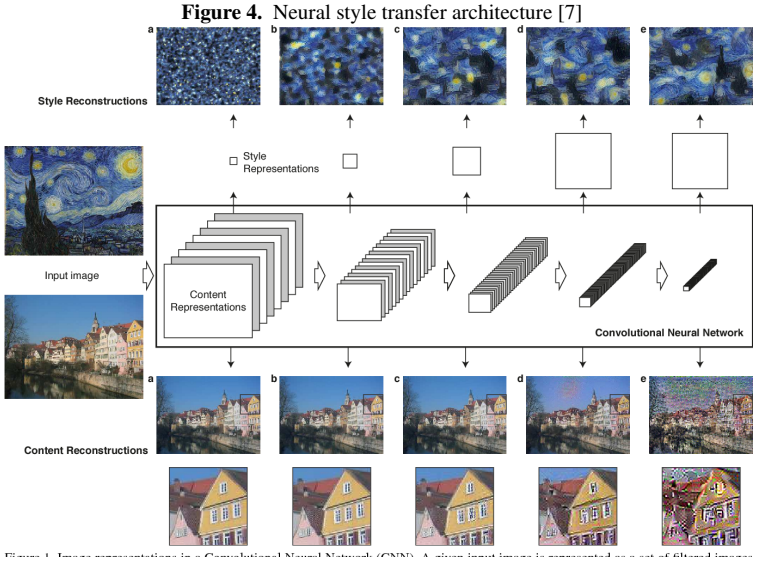
The key technique that makes NST possible is the convolutional neural network (CNN) [7]. We will discuss here the major techniques of doing NST on images, and then briefly examine one way of extending NST to video. Distinguishing content from style in natural photos is already a complicated task to solve. Deep convolutional neural networks, on the other hand, have recently made significant progress, resulting in strong computer

Figure 3. Deepfake models [6]



vision systems that can learn to retrieve high-level semantic features from real pictures [8]. CNN, as shown in figure 4, learn to extract high-level image content in standard image features that illustrate across data sets and even to other visual information processing tasks. such as texture recognition and artistic style classification when trained with enough labeled data on specific tasks [7].

Figure 4. Neural style transfer architecture [7]



4. Robotic Art

The concept of computational art is defined as the study and simulation of natural human creative behavior, by means of computer programs [9]. AI, or the recreation of human creativity by a computer, is used to apply this notion to musical, artistic, paintings, and literary creations. Robots can create art under this premise. Although it is impossible

Figure 5. Neural style transfer output

to foresee how AI will progress in the future, an increasing number of experts believe that man and machine will form a new relationship based on collaboration rather than dominance perceiving AI as a tool, much like a paintbrush [9].

5. Conclusion

The world around us is constantly changing as a result of continuous technological advancements in the field of AI. The impact of these breakthroughs has a long-lasting effect on all parts of our society. We would like to address some of these critical advancements by displaying them to the general public through our robotic art installation ”Mirror, Mirror.”

Although deepfakes evoke mostly negative connotations, they can play a positive role in our society. While empirical research on the subject is rather scarce, some studies indicate that watching their own doppelganger can lead to the creation of false memories and can cause a person to believe that they actually performed an activity that they did not [10]. For example, one can create a deepfake of a person who is afraid of public speaking. Upon seeing themselves giving a lecture in front of large groups, they might feel more confident in doing that themselves. There was also an increase in exercising after persons viewed that that activity had a positive health implication for their doppelganger [11]. Therefore, a deepfake of a person can also be used in promoting healthier lifestyles, where persons upon seeing a more physically fit version of themselves, might be prompted to make more health-conscious decisions.

NSTs, in the meanwhile, allow the viewer to be at the heart of the creative process as an active participant which can potentially have implications on a person’s creativity. If physical participation in a creative process boosts an individual’s creativity, NSTs can be used in various fields which require creative processes, such as applied arts, education, architecture, marketing, communication, and many other fields.

Historically, the arts have served as a reflection of society, a mirror. As civilization evolves and new technologies are introduced, such changes can be seen in art. We use deepfake techniques and NST in our project to give the audience a unique experience through interaction with screens, bringing attention to developments in the AI field while also inviting the public to participate in a larger discussion about its societal implications.

References

[1] Gamillo, E. (2021). New A.I. Tool Makes Historic Photos move, Blink and smile. Retrieved December 6, 2021, from `https://www.smithsonianmag.com/smart-news/ai-program-deep-nostalgia-revives-old-portraits-180977173/`

[2] Hui, J. (2018). GAN — What is Generative Adversarial Networks GAN? Retrieved December 6, 2021, from https://jonathan-hui.medium.com/gan-whats-generative-adversarial-networks-and-its-application-f39ed278ef09

[3] Zakharov, E., Shysheya, A., Burkov, E., Lempitsky, V. (2019). Few-Shot Adversarial Learning of Realistic Neural Talking Head Models. Retrieved December 6, 2021, from https://arxiv.org/pdf/1905.08233.pdf

[4] Hyeongwoo, K., Garrido, P., Tewari, A., Weipeng, X., Thies, T., Nießner, M., Perez, P., Richardt, C., Zollhofer, M., Theobalt, C. Deep. (2018). Video portraits. arXiv preprint arXiv:1805.11714.

[5] Nguyen, T., Nguyen, C., Nguyen, D., Nahavandi, S. (2019). Deep Learning for Deepfakes Creation and Detection. Retrieved December 6, 2021, from https://arxiv.org/pdf/1909.11573.pdf

[6] Siarohin, A., Lathuilière, S., Tulyakov, S., Ricci, E., Sebe, N. (.n.d). First Order Motion model for Image Animation - NeurIPS. Retrieved December 6, 2021, from https://proceedings.neurips.cc/paper/2019/file/31c0b36aef265d9221af80872ceb62f9-Paper.pdf.

[7] A. Gatys, L., S. Ecker, A., Bethge, M. (2015). A Neural Algorithm of Artistic Style. https://arxiv.org/abs/1508.06576

[8] A. Gatys, L., S. Ecker, A., Bethge, M. (.n.d.). Image style transfer using Convolutional Neural Networks. Retrieved December 6, 2021, from https://www.cv-foundation.org/openaccess/content$_$cvpr$_$2016/papers/Gatys$_$Image$_$Style$_$Transfer\protect\@normalcr\relax$_$CVPR$_$2016$_$paper.pdf.

[9] Can machines create art? (.n.d.). Retrieved December 6, 2021, from https://www.iberdrola.com/culture/computational-creativity-robot-art

[10] Segovia KY, Bailenson JN. Virtually true: children's acquisition of false memories in virtual reality. Media Psychology 2009; 12:371–393.

[11] Fox J, Bailenson JN. Virtual self-modeling: the effects of vicarious reinforcement and identification on exercise behaviors. Media Psychology 2009; 12:1–25.