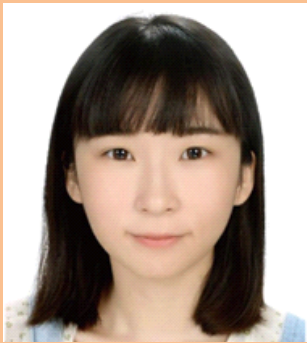# DEEPHAIR: FROM AN ARTWORK TO AN INTERACTIVE APPLICATION

Yu-Hsuan Lo
ms26703369@gmail.com
Yu-Hsuan Lo is a master student at the Department of New Media Art, Taipei National University of the Arts, Taipei, Taiwan. Her research interests are deep learning applications, computer vision applications, and VR/AR applications.

Shih-Wei Sun , Ph. D.
swsun@newmedia.tnua.edu.tw
Shih-Wei Sun is a Professor at Department of New Media Art, Taipei National University of the Arts, Taiwan. He received his Ph.D. in Electrical Engineering from the National Central University, Taiwan. His research interests are computer vision, sensor applications for mobile devices, and applications for human-computer interface.

Recent development of AI technology, especially generative models such as DeepFake and ChatGPT, has garnered concerns in the general public as well as research and governance communities due to the high risk of misuse and negative consequences garnered from unregulated access by unsavory individuals. However, AI technology in itself is a product of technological progress that is not innately harmful. This article introduces the DeepHair project and discusses the positive value that can be brought by DeepFake technology, despite the negative associations it is commonly known for.

Introduction on DeepHair

The project DeepHair [1] is a hairstyle preview system that is proposed based on the DeepFaceLab [2], a leading software for creating DeepFakes. The proposed interactive hairstyle preview system allows a user to stand in front of a camera with an interactive touchable instruction panel to choose the preferred hairstyle video captured from the hairstyle models, as shown in

Fig. 1. A hairstyle preview video will be generated based on the selected hairstyle via the DeepFake method. Compared with the conventional image-based hairstyle preview systems, e.g, GAN-based method [3], and StarGAN method [4], the proposed DeepHair allows users to view the generated hairstyle video with more dynamic coverage from a variety of view angles.
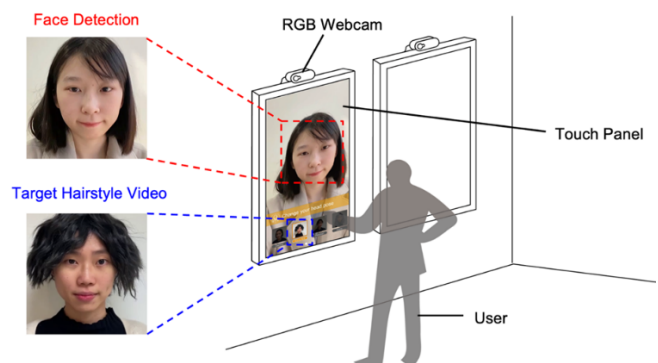

Figure 1: The system diagram of DeepHair [1].

The Artwork <To Be>

Prior to the system development, the project has been explored in a more artistic avenue. The

predecessor of the DeepHair project [1] is the artwork entitled, <To Be, Yu-Hsuan Lo, 2021> as shown Fig. 2. In the artwork, the artist, Yu-Hsuan used a camera to record a short video of herself using OpenCV to detect the facial parts from the frames, and DeepFaceLab [2] was applied to integrate the facial parts into the target videos. As shown in Fig. 2 (a), the target videos consist of four different characters[1]: a boy (right part[2]) , a little girl (upper part[3]) , a teenage girl (left part[4]) , and an office lady (lower part[5]) . The characters represent four different identities that stayed in Yu-Hsuan's mind. Furthermore, on social media platforms, Yu-Hsuan assumed the roles of the four characters with different user IDs. In this artwork, the videos played on the four cubes in Fig. 2 (a) are generated from DeepFake methods. In the central parts, the virtual human subject represents the real Yu-Hsuan, and the four DeepFaked videos are the characters played Yu-Hsuan as the "faked" characters in the social media platforms and game platforms with the respective identities. The videos were generated by the "face swapping" process in DeepFaceLab [2] and displayed on the cubes as illustrated in Fig. 2 (b). In the artwork <To Be>, Yu-Hsuan discussed an in-between relationship between an inside character (the central virtual figure in Fig. 2(a)) and the expected characters in the social world (the images in Fig. 2(b) projected outwardly by the four cubes) of the same person in the real physical world.
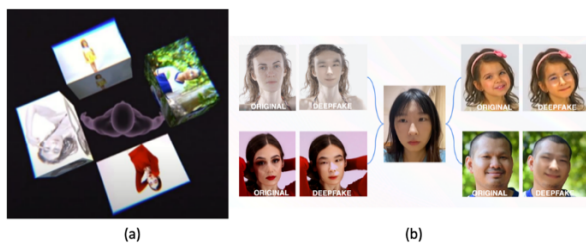

Figure 2: The artwork <To Be, Yu-Hsuan Lo, 2021>: (a) The artwork with four different cubes and (b) Yu-Hsuan Lo's facial featured replaces those of the characters in four different videos based on DeepFaceLab [2].

The Transition to an Interactive Application: A Negative or Positive?

During the creation of the artwork <To Be, Yu-Hsuan Lo, 2021>, the news broke the story of a Taiwanese Youtuber "Xiao Yu" who DeepFaked videos [5], specifically using DeepFake technology to impose celebrities' faces over actors' in pornographic videos. Due to this incident, DeepFake technology is widely known as an AI application with a negative undertone. In addition, DeepFaked videos shared over the social media raised issues of invasion of privacy and disinformation [6]. Furthermore, ChatGPT [7] in the text retrieval domain has been developed to use a machine-trained model to provide conversational applications as dialogues from a user to a cloud server. The generated texts are as natural as the sayings of a human being. The situation is similar to the videos generated by the DeepFake [2] method to be viewed by audiences. Not only texts, but also the videos, are hard to be distinguished by the human users whether they are real texts/videos or faked/generated texts/videos. When the technologies are applied in negative ways, this confusing property makes the situation worse. In order to detect such fake texts, Mitchell et al. developed "Detect GPT [8]" whereas, Ciftci et al. developed "FakeCatcher" [9] to detect fake videos. More and more effort has been paid by researchers to detect faked media contents [8] [9] so as to suppress the situation from getting worse.

Despite the negative implication on the DeepFake name, the technology is in fact neutral as demonstrated with the development of the artwork <To Be, Yu-Hsuan Lo, 2021> where it is used as an abstract exploration of the self and the outward characters portrayed to the social world. In particular, we applied light fitting to estimate the features from various angles of views in DeepFake to generate a more natural face swapping effect from the artist Yu-Hsuan to the target video belonging to a character with his/her identity to generate a virtual character in a virtual space. In this context, the DeepFake approach is merely a tool of artistic expression of a user and not an inherently harmful technology.

So, to move on from the negative, we focus on the capabilities and potentials brought upon by the artwork. Since the face swapping from the user (the artist Yu-Hsuan) to a character can be displayed

---

[1] Videos were obtained from the copyright free website, Pixabay.
[2] https://pixabay.com/videos/man-smiling-farmer-happy-young-76888/
[3] https://pixabay.com/videos/kids-dancing-outfits-new-year-12266/

[4] https://pixabay.com/videos/weapons-pistol-depression-stress-45282/
[5] https://pixabay.com/videos/girl-posing-model-portrait-lovely-46948/

naturally, as shown in Fig. 2, such facial swapping has great potential to be developed into applications. We were thinking about: "Is there an application that needs to swap the facial part to a target video?" Thus, after our brainstorming, we came to the scenario of a hair salon. To have a haircut in a hair salon, it is not easy for a customer to estimate the final result of a haircut, and the communication between a customer to a hair stylist is often time consuming if not inaccurate. During a haircut, the customer may be in a constantly anxious state throughout the whole session.

Therefore, we proposed DeepHair [1], a Deepfake technology for swapping the facial part of a customer in a salon to a target video with a hairstyle model. In this circumstances, issues of invasion of privacy and disinformation are not touched, and the generated video is a demonstration of a positive implication, that is as a medium for a customer and a hair stylist to have common understanding and effective discussions regarding the hairstyles. The obtained hairstyle preview video can release the pressure of misunderstanding from both the customer and the hair stylist in a salon and promote favorable communications. Hence, the proposed DeepHair [1] system is an exhibition of a positive use case of the DeepFake technology.

The DeepHair System

In the developed DeepHair [1] system, a customer in a hair salon can select a video with a hairstyle model (the video with a human subject at the bottom-left of Fig. 3) from our system. Next, a customer can use a camera to record a small period of video with her face (the top-left of Fig. 3) as the input data. The hairstyle is then transferred by the adapted DeepFaceLab [2], on to each frame of the customer video. An example of the target losses measured among different epochs are shown on the top-right of Fig. 3, and the corresponding generated deepfake frames are shown on the bottom-right of Fig. 3.

As a prototype, a demo system of DeepHair [1] was developed and showcased to the authors and participants in ACM Multimedia Asia 2022, so as to provide a practical interactive user experience. Figure 4(a) shows the process where a user can use her finger to select the target video with a hairstyle. When recording a short video as the input to the system, a user is asked to rotate her head to different directions, as shown in Fig. 4(c)-(d). Finally, the face swapping results from the input short video to the hairstyle model video are obtained, and the hairstyle preview from different angles can be observed as shown in Fig. 5. This hairstyle preview video can then be used a point of discussion between a customer and a hair stylist in a hair salon.
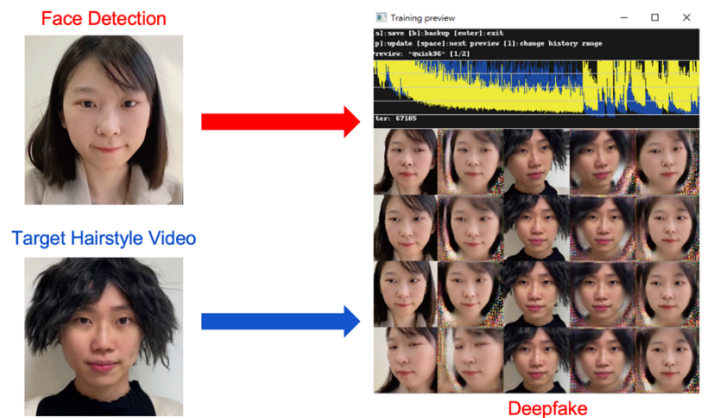

Figure 3: Hairstyle transfer by adopting DeepFaceLab [1] [2].
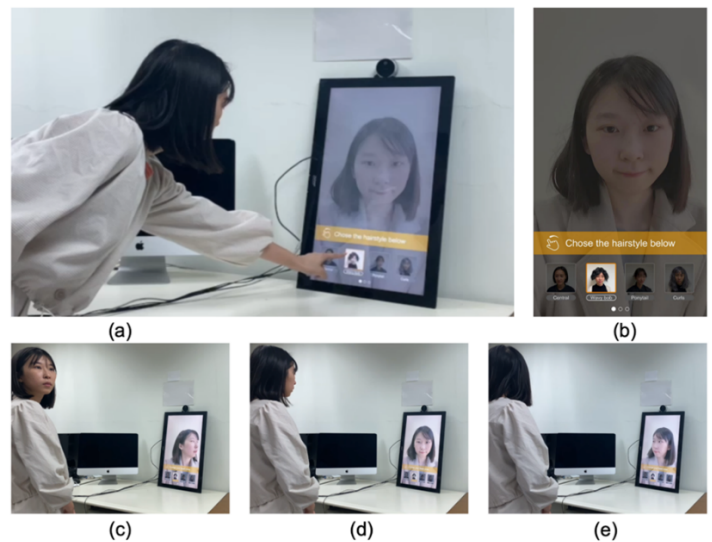

Figure 4: System interaction, (a) choosing a target hairstyle video, (b) the screen shot of the system UI, (c) rotating the head to show the left face, (d) rotating the head to show the middle face, (e) rotating the head to show the right face [1].
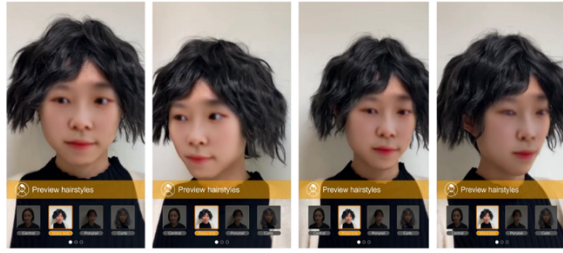
Figure 5: Hairstyle previewing result from different viewing angles in a generated DeepFake video [1].

Now and the Future

Currently, DeepHair [1] is a pilot study of integrating AI technologies to create artworks as well as practical positive uses in the future. In terms of technical implementations, the adoption of DeepFace Lab [2] for the DeepHair [1] still requires significant computational power in order to generate the DeepFake preview video, from the current 15 minutes to a more manageable period of time. Cloud computing or distributed DeepFake processing approaches can be developed in the future to support a system in this aspect. In addition, the DeepHair [1], actually performs 2D frame-to-frame swapping, therefore, 3D processing is expected to be more effective. By scanning the 3D point clouds of a user's face, we expect the outcome of the face to be more natural, and the 3D model generation may allow users to interact with the generated 3D content more efficiently. Furthermore, the proposed hairstyle preview system [1] can also be deployed into web services in the future. For example, the virtual try-on glasses of OWNDAYS [10] (a Japanese eyeglass company), and the virtual makeup system of L'Oréal Paris [11]. The use of built-in camera of the computers of the users to provide an AR-based user experience of virtual-on holds great potential. In the future, we expect a hairstyle preview system can be used in hair salons to be used on portable tablets, pads, and even smartphones as a testament that DeepFake technology can be used for good.

References

[1] Lo, Y. H., & Sun, S. W. (2022) DeepHair: a DeepFake-based Hairstyle Preview System. In Proceedings of the 4th ACM International Conference on Multimedia in Asia, 41, 1 - 2.

[2] Perov, I., Gao, D., Chervoniy, N., Liu, K., Marangonda, S., Umé, C., Jiang, J., RP, L., Zhang, S., Wu, P., & Zhang, W. (2020). DeepFaceLab: Integrated, flexible and extensible face-swapping framework. arXiv preprint arXiv:2005.05535.

[3] Zhu, P., Abdal, R., Femiani, J., & Wonka, P. (2021). Barbershop: GAN-based image compositing using segmentation masks. ACM Transactions on Graphics (TOG), 40(6), 1-13.

[4] Nakamae, Y., Wang, X., & Yamasaki, T. (2020). Recommendations for Attractive Hairstyles. In Proceedings of the 2020 Joint Workshop on Multimedia Artworks Analysis and Attractiveness Computing in Multimedia, 19-24.

[5] Benack, E. (2021). Taiwanese Youtuber uses deepfake tech to make celebrity porn, Radio Taiwan International. Retrieved February 28, 2023, from https://en.rti.org.tw/news/view/id/2006342

[6] Chesney, B., & Citron, D. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review, 107, 1753.

[7] OpenAI. (2022). ChatGPT: Optimizing Language Models for Dialogue. OpenAI. Retrieved February 28, 2023, from https://openai.com/blog/chatgpt

[8] Mitchell, E., Lee, Y., Khazatsky, A., Manning, C. D., & Finn, C. (2023). DetectGPT: Zero-Shot Machine-Generated Text Detection using Probability Curvature. arXiv preprint arXiv:2301.11305.

[9] Ciftci, U. A., Demir, I., & Yin, L. (2020). Fakecatcher: Detection of synthetic portrait videos using biological signals. IEEE transactions on pattern analysis and machine intelligence. (Early Access)

[10] OWNDAYS. (n.d.) Virtual Try-On. Retrieved February 28, 2023 from https://www.owndays.com/vn/en/try-on

[11] L'Oréal Paris. (n.d.) Virtual Makeup Try-On. Retrieved February 28, 2023 from https://www.lorealparisusa.com/virtual-try-on-makeup