

Splicing ViT Features for Semantic Appearance Transfer

(No. T4-2192)

Principal investigator

Tali Dekel

Faculty of Mathematics and Computer Science
Department of Computer Science and Applied Mathematics

Overview

Artificial intelligence (AI) solutions are vastly available for applications such as object segmentation, image detection, visual creation, writing assistance, and many more. There is a continuous quest to improve photograph editing software and existing AI systems. A new technology developed by Dr. Tali Dekel and her team can controllably edit and generate high-definition (HD) chimeric images based on semantically related image pairs in a self-supervised and pre-trained manner, by using the features of a Self-Distillation with no labels Vision Transformer (DINO-ViT) as perceptual losses. Furthermore, this technology outperforms competitive technologies by transforming and finalizing an artificial naturally-looking image. The system is not limited to specific structures or appearances, as it can process animals, landscapes, and inanimate objects, providing an advanced editing tool based on cutting-edge AI technology for multiple usages.

The Need

The demand for achieving controllable editing and generation capabilities of real and artificial images is constantly increasing. Popular text-to-image AI algorithms currently do not offer automated auxiliary features for image edition. This results in multiple text-to-image iterations by the users until the desired outcome is achieved. In photo editing applications, the editing process is time-consuming and often demands unique skills from the user.

The Solution

This technology potentially provides an additional layer of image processing for AI-generated and real images. With this algorithm, users can further process images according to their preferences by providing the image appearance they desire (Fig. 1). The algorithm will automatically detect the structure in the source image and match it to the supplied appearance image. The outcome is a high-quality naturally-looking image.

Â

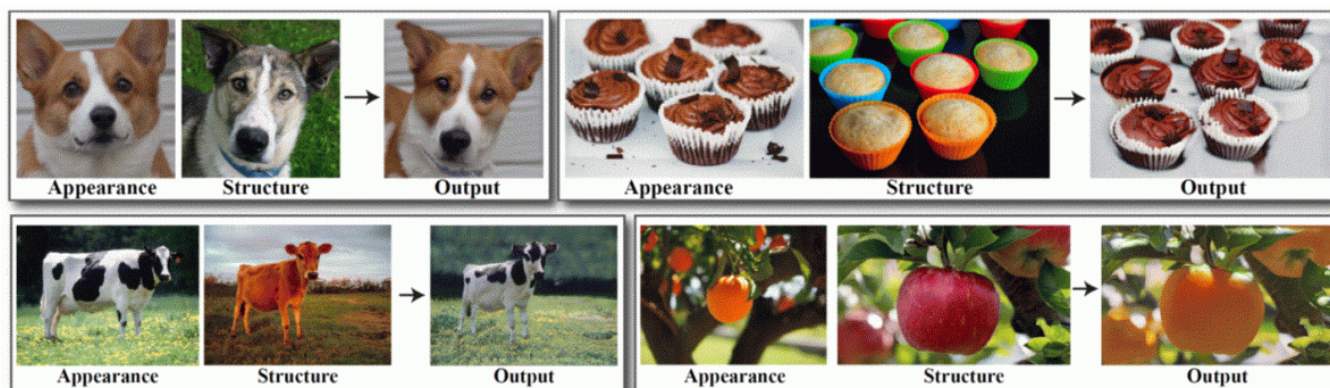


Figure 1: Examples of technology usage. Given two input images (a source structure image and a target appearance image), the algorithm generates a new image in which the structure of the source image is preserved while the visual appearance of the target image is transferred in a semantically aware manner. (Adopted from Tumanyan, N. et al.)

Technology Essence

The technology leverages a self-supervised, pre-trained, and fixed DINO-ViT model that can automatically detect structures in a source image and "paint" them with the appearance of a semantically related object of a target image. In opposed to AI-based image processors, this technology uses DINO-ViT's features as perceptual losses. The generator is trained only on a single input image pair, without additional information (e.g., segmentation/correspondences) and adversarial training. Thus, the technology can be applied across various objects and scenes and generate high-resolution results [1]. This technology eventually enables controllable editing and generation capabilities that alternative platforms currently fail to deliver.

Applications and Advantages

Applications

- A feature for AI systems that can create realistic images with different appearances
- An add-on for image processing software to modify HD photos
- An add-on for social media applications for the real-time modification of photos

À

Advantages

- A fast and precise editing tool
- Creation of chimeric HD realistic images
- The system operates on multiple objects, not limited to a specific set of structures

Development Status

The ViT algorithm is developed, demonstrated, and published.

Market Opportunity

The global photo editing App market was valued at \$293M in 2022 and is projected to reach \$402M by 2030 at a CAGR of 3.57% [2].

The global image recognition market was valued at \$36B in 2021 and is projected to reach \$177B in 2030 at a CAGR of 18.3% [3].

References



[1] N. Tumanyan, O. Bar-Tal, S. Bagon, and T. Dekel, "Splicing ViT Features for Semantic Appearance Transfer," pp. 10738-10747, Jan. 2022, doi: 10.48550/arxiv.2201.00424.

[2] "Photo Editing App Market Size, Share, Trends, Opportunities & Forecast," <https://www.verifiedmarketresearch.com/product/photo-editing-app-market/> [1] (accessed Jan. 01, 2023).

[3] "Image Recognition Market Size, Share, Trends, Opportunities & Forecast," <https://www.verifiedmarketresearch.com/product/image-recognition-market/> [2] (accessed Jan. 01, 2023).

Patent Status

PCT Published: Publication Number: WO2023126914
