# Encoding in Style: Class-Specific Generation and Translation

Dana Cohen[1], Raja Giryes[1], and Hayit Greenspan[1]

[1] Faculty of Engineering, Tel-Aviv University, Tel-Aviv, Israel

There have been great advancements in deep learning methods in recent years, primarily owing to large scale datasets. A fundamental problem in employing these methods in the medical field is the lack of labeled data as well as a severe class imbalance that exists in many datasets. Image synthesis and image-to-image translation introduce ways for adding labeled data that can be used to train and improve existing deep learning methods. StyleGAN proposed by Karass et al. [1], yields state-of-the-art results in unconditional generative image modeling. StyleGAN generates high-resolution images and is unique since it exposes novel ways to control the image synthesis process.

In this work, we propose an encoder, to embed real medical images into the StyleGAN latent space to enable controlling of the synthesized images and thus enlarging small-scale datasets. We exploit StyleGAN's ability to synthesize images of high-resolution to create new realistic medical images and broaden the StyleGAN's control over the synthesis by training an encoder to learn the disentangled latent representation of the data. This allows the encoding of any image to the latent space and the controlled manipulation of the generated images to generate class-specific images or translate between two modalities.

The data used for the training of StyleGAN consists of 11 computed tomography (CT) scans taken from the publicly available training dataset of the LiTS challenge [2]. Only a small portion of the images in the database was used (11 out of 130 CT scans) to demonstrate the training of the StyleGAN on a small dataset.

Our framework is an autoencoder comprised of a ResNet50 based encoder and a StyleGAN generator which serves as the decoder. The high-quality reconstruction of the images and the disentanglement in the latent space is achieved by training the encoder with generated images along with their corresponding latent representations. This provides an unlimited data source for training and forces the embedding to a disentangled latent space.

The combination of the StyleGAN traits along with the encoder allows for many potential applications to enlarge datasets in the medical field. The generator architecture of the StyleGAN can be used to control the generated class or translate between two modalities by embedding the images to the latent space and utilizing the "Style Mixing" feature of the StyleGAN, Fig. 1. By controlling the input to each resolution in the generator we can control the features in the output images. It can be inferred from the results that the coarse styles determine whether the output image will contain a tumor or not and the fine styles control the "modality style" of the image. Therefore, by inputting a latent vector of an image with or without a tumor in the coarse styles we can generate additional images belonging to the desired class. Additionally, the style of the output image will be determined by the modality of the image used in the fine styles and the anatomy of the image used for the rest of the styles.

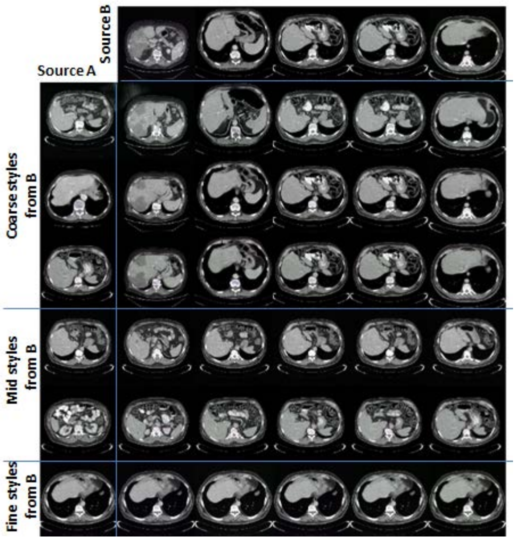


**Fig. 1.** An example of the style mixing effect with abdominal CT images. Two sets of images were generated (sources A and B) and the rest were generated by combining a specified set of styles from the two sources. Coarse styles represent coarse spatial resolutions of $4^2$–$8^2$, middle styles - $16^2$–$32^2$, and fine styles - $64^2$–$512^2$. The output image is influenced by each source according to the styles copied from them.

The disentanglement in the latent space which arises from the StyleGAN architecture can be exploited to create class-specific images. By encoding the images, classifying them in the latent space, and learning the direction which corresponds to each class, specifically tumor/no tumor, we can alter the latent vector in the desired direction and control the class being generated by StyleGAN, as presented in Fig. 2. This has the potential of improving classification problems by adding data to each class and thus enlarging small-scale datasets and reducing the class imbalance.

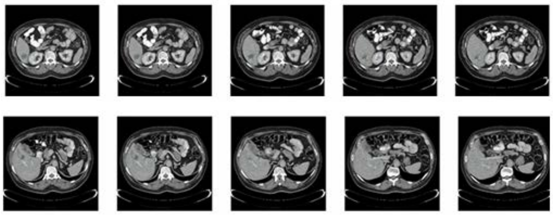

**Fig. 2.** Adding a tumor to an image by moving towards the direction of the tumor in the latent space (from right to left).

# References

1. Karras, T., Laine, S., Aila, T. A style-based generator architecture for generative adversarial networks. Computer Vision and Pattern Recognition (CVPR),(2019).
2. Christ P, Ettlinger F, F G, Lipkova J, Kaissis G. LiTS - Liver Tumor Segmentation Challenge (2017).

ACK: This work is funded partly by The Ministry of Science Grant on Deep Learning for Medical Imaging