Inpainting – white paper

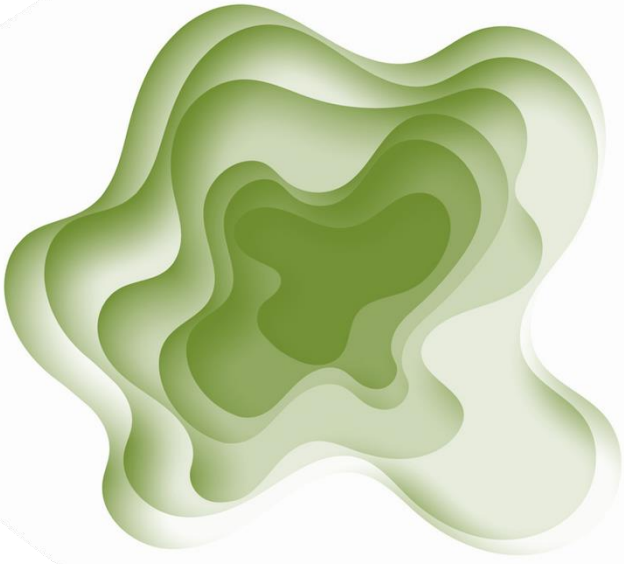
Feb 2022

saiwa

simple artificial intelligence web application

Image inpainting refers to one of the challenging problems in image processing that is concerned with filling missing regions of an image. Texture synthesis-based methods have been one of the dominant solutions to this problem where gaps are filled using surrounding known regions. These methods assume that the missing regions are repeated somewhere in the image. For non-repetitive regions a general knowledge of the image content is required. This is achieved using advances in deep learning and convolutional neural networks, where texture synthesis and the overall knowledge of the image are combined in a twin encoder-decoder network. The two convolutional parts are trained together to predict missing regions.

With the saiwa image inpainting service we provide a simple interface to one of the most interesting deep neural network based inpainting methods: generative image inpainting network, with the support of contextual attention and gated convolution which is called **DeepFill v2** [1, 2]. In DeepFill v2 the baseline network consists of two generators and two discriminators which is a typical network architecture for deep image inpainting. Then the two main properties of this network are: (1) the **contextual attention layer** that is designed to employ features of distant special known regions to predict features of local missing regions [3]; and (2) **gated convolution** as a learnable partial convolution to handle irregular masks.



INPAINTING

Figure 1 shows an overview of DeepFill v2 network. DeepFill v2 is a two-stage coarse-to-fine network. The inputs of the coarse generator consist of the masked image, the mask and user sketch guided image (optional) to estimate a coarse version of missing regions. Then, this coarse version is passed to the second refinement generator network. Here, contextual attention (CA) layer is used in this refinement network which is actually an extra branch that reflects the contributions of all known regions to the unknown region based on an attention score (for details please refer to [3]). For the discriminator, PatchGAN and spectral normalization to each standard convolutional layer of the discriminator is employed. For more details and also a fast and straightforward

In addition to a simple user interface for applying this interesting algorithm, saiwa also provides a simple and handy tool to generate masks. Figure 2 shows a screen shot of the mask generator. Figure 3 shows a few results of filling gaps in images using saiwa image inpainting services. Note that you may fill multiple regions in an image simultaneously without repeating the algorithm.

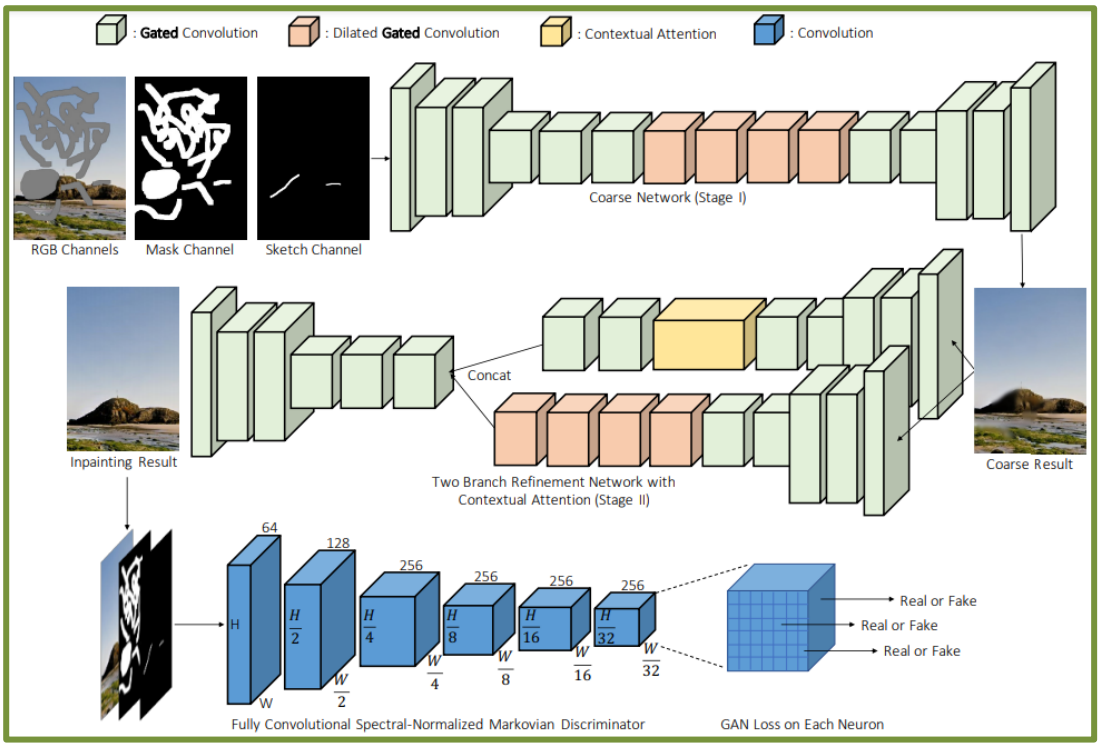


Figure 1. Overview of DeepFill v2 network architecture with gated convolution, dilated gated convolution, contextual attention and normal convolution layers (printed from [1])

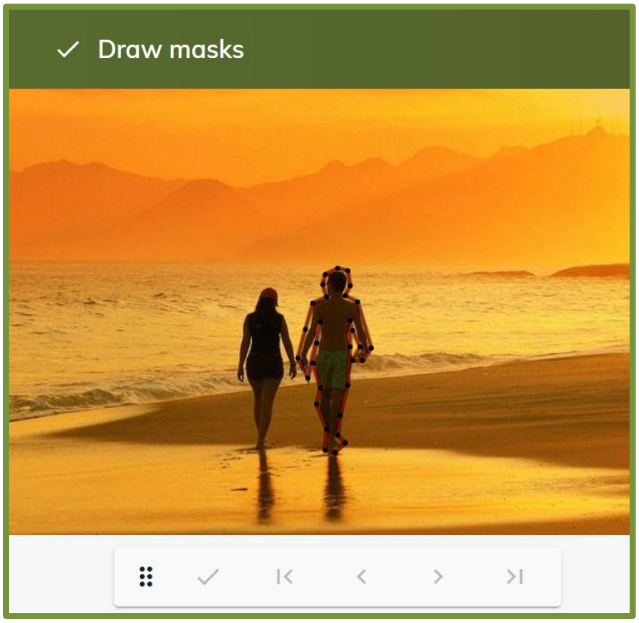


Figure 2. saiwa mask generation tools.

Figure 3. saiwa inpainting results; (from left to right) input image, mask and result image after filling the masked regions.

References:

[1] Yu, Jiahui, et al. "Free-form image inpainting with gated convolution." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.

[2] Yu, Jiahui, et al. https://github.com/JiahuiYu/generative_inpainting. 2019

[3] Yu, Jiahui, et al. "Generative image inpainting with contextual attention." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[4] Chu-Tak Li, "A Practical Generative Deep Image Inpainting Approach. Review: Free-Form Image Inpainting with Gated Convolution", 2020.

Computer Science, Tech. Rep., 2016.

info@ saiwa.ai

720 Guelph Line Burlington, ON L7R 4E2 Building

+15148131809

www.linkedin.com/company/saiwa

www.instagram.com/saiwa.ai