Abstract

Many computer vision tasks could be formulated as image feature learning tasks where the objective is to perform image restoration or image synthesis. Image restoration is an ill-posed problem aiming to restore an image from a corrupted observation. The image synthesis task is related to synthesizing a new image while preserving various properties such as image context and object structure.

The deep learning framework to perform image restoration and image synthesis without using paired samples for training was not explored until 2017. Recently, we have witnessed various research frameworks that are able to perform such tasks without using paired samples. By paired samples, we mean a training dataset consisting of pairs of source and target images, e.g., a clean and noisy image pair for image denoising.

Deep feature learning without using prior examples has practical applications for situations where one has limited computational resources, the training samples are difficult to collect, and one needs image output having no influence through bias from the training samples.

Initial frameworks that do not use paired samples for training were proposed for image restoration tasks such as denoising and super-resolution. The key observation was the network structure itself works as an implicit image prior to perform image restoration as in Deep Image Prior (DIP). We propose a general framework ($\mathcal{MED}$) to study how the structure of the network influences the quality of restoration. Our framework allows various network structures by modifying the network components such as skip links and the composition of the encoder-decoder subnetworks. These handcrafted network structures illustrate how the construction of untrained networks influence the following image restoration tasks: denoising, super-resolution, and inpainting.

Later, the deep internal learning-based single image GAN framework (InGAN and SinGAN) was proposed for the image synthesis task. Inspired by InGAN, we propose a single image GAN framework Deep Contextual Internal Learning (DCIL) for image restoration and image synthesis tasks. There are two interesting challenges here. (1) What aspects of the network would help in the GAN framework for image generation when there is a lack of feature learning from the training samples? (2) What should be the structure of the loss function when the source image and the target image are not aligned and do not have spatial correspondence? Our DCIL framework investigates the challenges above. It uses the internal learning of image context to perform various tasks. We perform image resizing application in the following setups: classical image resizing using super-resolution, a challenging image resizing where the low-resolution image contains noise, and content-aware image resizing using image retargeting.

The challenge in learning deep features without prior examples is to learn the semantics of the scene. We proposed a Deep Contextual Feature Learning framework (DeepCFL) to preserve and synthesize contextual features when performing image restoration and image synthesis. The contextual features are simply high dimensional vectors representing the semantics of the given image. DeepCFL is a single image GAN framework that learns the distribution of the context vectors from the input

image. We show the performance of contextual learning in various challenging scenarios: outpainting, inpainting, restoration of randomly removed pixels, and object synthesis.

We also propose a framework Deep Object-based Style Transfer (DeepObjStyle), where the task is to synthesize a new image that incorporates content features from the content image and style features from a style image while minimizing the content mismatch due to feature correlation in content and style images. Another interesting problem we address is the image enhancement of corrupted images to improve image quality. Our Deep Internal Learning for Image Enhancement framework (DILIE) shows the generalizability of deep learning methods for diverse applications such as hazy and noisy image enhancement.

Our novel strategies for image restoration, image synthesis tasks, and the related methods have shown that training deep IR models without using training samples (*i.e.*, source and target image pairs) can be performed and applied to solve several low-level vision problems efficiently.