Tutorial 9: Deep Learning for Image Transformation

Image transformation tasks, where the input and output are both images, play a fundamental role in many different applications, including edge detection, semantic segmentation, and colorization of black and white images. Convolution-based operators have traditionally been used for low-level image processing, however, with the advent of recent deep learning techniques, they have gained applicability to a wide range of problems. The key idea is the fully convolutional network which consists of stacking many convolution operators together, along with non-linear activation functions, which allow representing arbitrary image transformation functions and processing images of any resolution.

Although the concept of the fully convolutional network is simple, there are many issues that have to be overcome for their successful usage in practice, such as vanishing gradients and insufficient spatial support of the model. Models suffering these problems will not converge during learning, or show low performance during evaluation, as they will be unable to learn from the training data of a specific task. Recent trends to overcome these issues include encoder-decoder architectures, batch normalization, atrous convolutions, and residual learning, and must be tailored to the specific application at hand.

In parallel to developments on fully convolutional networks, generative adversarial networks have shown that it is possible to train neural network to do tasks that were previously difficult such as image generation from unsupervised training data. The core of generative adversarial networks, adversarial training, can also be used for image transformation tasks and have shown great results in image completion and other generation-related tasks. Furthermore, using two image transformation networks in tandem with a reconstruction loss and adversarial training, it is possible to train image transformation networks from weakly-supervised data, where explicit training pairs are not available, and is known as cycle-consistency.

In this tutorial, we first review the basics of the modern convolutional neural network models and how to train them. Then we illustrate how they are used to process images, using several research examples. We will review the latest research advances and trends in the art of processing and generating images using deep learning and discuss the challenges in deploying these methods in practice.

Speakers:

Hiroshi Ishikawa
Waseda University, Japan

Edgar Simo-Serra
Waseda University, Japan