Abstract

Neural networks have improved significantly in past decades. They are competent to solve complex problems in the field of deep learning and they are capable to manage a large amount of complex data like images, videos and sound. However, the training of neural networks requires a significantly large amount of annotated data, which is not always possible. Machine learning engineers inevitably have to generate synthetic data. Although, the neural networks trained on synthetic data will not able to generalize well on real data. In recent years, an effective technique named domain adaptation has evolved, to address the problem of scarcity of annotated data. The domain adaptation technique can transform data from the source domain to the target domain. For example, domain adaptation techniques like image-to-image translation can be used to transform images of zebras into images of horses and vice-versa. This thesis proposes an image-to-image translation application that aims to reduce the domain gap between synthetic and real data distribution using Cycle-Consistent Adversarial Networks (CycleGANs). The proposed application is used to transform synthetic document images into realistic document images, to overcome the scarcity of annotated real document images. In addition, these generated realistic document images are used to train a classifier to classify similar unlabeled real document images, thereby accelerating the process of labeling images in an unsupervised and automated manner. Experimental results show the generated realistic document images are qualitatively convincing and need improvement quantitatively to match the real data distribution significantly. Such preliminary results show that CycleGAN can solve the problem of data scarcity by generating high-quality images in the target domain. The purpose of this thesis is limited to improving the classification of real document images. Once the rich and sufficient data is generated in the target domain, the performance of the real document image classifier eventually can be improved. This thesis is limited to the study of unpaired image-to-image translation method CycleGAN. The remaining methods and comparisons with them are left for future work. In the future, CycleGAN can be used to generate high-quality realistic images in many tasks, such as handwriting recognition, image classification, image segmentation and object detection.

Keywords: CycleGAN, Generative Adversarial Network (GAN), Domain Gap, Domain Adaptation, Image-to-Image Translation, Data Distributions.