Abstract

論 文 の 内 容 の 要 旨

論文題目　Learning to Generate Images from Uncertain Data
（不確実データからの画像生成の学習）

氏　　名　金子　卓弘

Image generation is a task of producing or reproducing data that are indistinguishable from the real data. In computer vision and machine learning, this has been intensively studied owing to its high potential for various applications, such as image colorization, super resolution, image inpainting, photographic image synthesis, and photo editing. Recently, deep generative models, such as generative adversarial networks (GANs), variational autoencoder (VAEs), autoregressive models (ARs), and Flow-based models (Flows), have emerged as a powerful framework that provides clues to solving this problem. Owing to their promising results, they have attracted special attention in the research field, and diverse studies, ranging from basic to practical, have been conducted.

However, a persistent issue is that the previous studies assume that it is possible to obtain large-scale high-quality images or accurately labeled images, which are often laborious or impractical to collect in a real-world scenario. To deal with this issue, in this thesis, we address the problem of learning to generate images from uncertain data. In particular, we tackle this problem for two fundamental tasks in image generation, i.e., unconditional image generation and conditional image generation. To advance this research, we first categorize the uncertainty in terms of operation and dependency between uncertain and latent ideal data. Then, we propose three novel extensions of GANs for unconditional image generation (Part I) and two new variants of GANs for conditional image generation (Part II).

Part I describes methods for learning to generate images from uncertain images. In particular, we propose noise robust GAN (NR-GAN) for noisy images; blur robust GAN (BR-GAN) for blurred images; and compression robust GAN (CR-GAN) for compressed images. In these models, we introduce a noise generator, kernel generator, and compression parameter generator for NR-GAN, BR-GAN, and CR-GAN, respectively, and train them

with a non-degraded image generator. We also propose distribution, transformation, and architectural constraints to facilitate the performance. Furthermore, we introduce blur, noise, and compression robust GAN (BNCR-GAN) as the combination of BR-GAN, NR-GAN, and CR-GAN. In the experiments, we demonstrate that NR-GAN, BR-GAN, CR-GAN, and BNCR-GAN can achieve noise robust, blur robust, compression robust, and degradation robust image generation, respectively, without having the complete knowledge of the image degradation.

Part II presents methods for learning to generate images from uncertain labeled images. Especially, we provide label-noise robust GAN (rGAN) for random or class-dependent label noise and classifier's posterior GAN (CP-GAN) for instance-dependent label noise. In rGAN and CP-GAN, we respectively incorporate a noise transition model and a classifier's posterior into typical conditional extensions of GANs to deal with label noises. Through extensive experiments, we show that rGAN and CP-GAN can accomplish label-noise robust image generation and class-distinct and class-mutual image generation, respectively.

Part III discusses the applications of the models described in Part I and Part II. Towards learning to generate images from real-world degraded images, we introduce image signal processing BNCR-GAN (ISP-BNCR-GAN), which is an extension of BNCR-GAN that incorporates ISP. We further extend it to neural image signal processing BNCR-GAN (NISP-BNCR-GAN) to learn ISP in a data-driven manner. Additionally, we show the generality of rGAN and CP-GAN by applying them to multi-domain image-to-image translation.

We conclude this thesis with a summary of our contributions, summarization and generalization of our ideas, and remarks on future directions. In the future, we hope that our findings and guidelines facilitate the construction of a generative model in various real-world scenarios where only uncertain data are available for training.