2019 5th International Conference for Convergence in Technology (I2CT)
Pune, India. Mar 29-31, 2019

A Review of Generative Image Modeling Techniques

Gayashan Bombuwala
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
email: gayashan.2015047@iit.ac.lk

Guhanathan Poravi
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
email: guhanathan.p@iit.ac.lk

Abstract **Generative image modeling is an area of research in computer vision that many works has been conducted to address various issues that has been encountered in the domain. In this paper, we categorize review several works that has been proposed for generative image modeling. The goal of this paper is to encourage the future researchers to pursue the work through the limitations of existing generative image modeling techniques.**

Index Terms—Generative Image Modeling

I. INTRODUCTION

GENERATIVE modeling is the process of generating new samples by learning the distribution of a given dataset using deep learning techniques. The dataset used in the process, can be in various forms such as images, sentences, sounds etc. When this process is applied to a collection of images, it's known as generative image modeling. There are mainly two categories of generative image modeling techniques, unsupervised techniques and conditionally unsupervised techniques. The table below [Table 1] shows a representation of various branches of techniques that fall under these two categories.

TABLE I
GENERATIVE IMAGE MODELING TECHNIQUES

Unsupervised	Conditionally Unsupervised
Variational Autoencoders	Attribute Label Conditioned
Autoregressive Models	Category Conditioned
Generative Adversarial Networks	Image Conditioned
	Text Conditioned

In the following two section, we will take each of these techniques individually and review several techniques that fall under each branch.

II. UNSUPERVISED GENERATIVE IMAGE MODELING TECHNIQUES

A. *Variational Autoencoders*

Autoencoders [Figure 1] are type of neural networks which consists of two neural networks known as the encoder and decoder. Autoencoders learn the representation of data by compressing it (encoder) and decompressing it back (decoder) it to match the original input data. The compression process can generate an information loss, which is similar to what happens in JPEG or MP3 file compression. A distance function that quantifies the information loss that is generated from the lossy compression is used to aid the learning process of the autoencoder.

In 2014, Kingma *et al.* [1] proposed a new type of autoencoders known as Variational Autoencoders (VAE). Meanwhile, a typical autoencoder is just learning a function representing the data, Variational Autoencoders function by learning the parameters of the representing data's probability distribution. Variational Autoencoders can sample from the distribution and generate new input data samples as they learn the model of data. Rezenede *et al.* [2] proposed a similar approach where a re-parameterization technique was used, in order to maximize the lower bound.

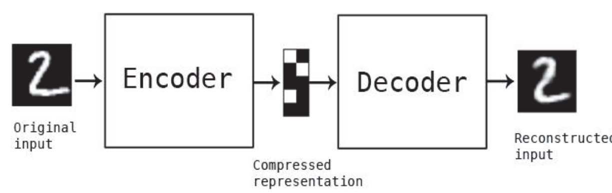

Fig. 1 The structure of an Autoencoder

B. *Autoregressive Models*

Neural Autoregressive Distribution Models (NADE) proposed by Uria *et al.* [3] are a set of neural network architectures which can be applied to solve the problem of unsupervised distribution and density estimation. In their work, they have discussed how the topological structure of pixels in an image can be exploited using a deep convolutional architecture for NADE.

Oord *et al.* [4] proposed another technique known as Pixel Recurrent Neural Networks (Pixel RNN). Pixels in an image are sequentially predicted by Pixel RNNs along the two spatial dimensions. Discrete probability values of the raw pixels are modeled and the complete set of dependencies in

978-1-5386-8075-9/19/$31.00 ©2019 IEEE