Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

My first and last thesis
Subtitle Subtitle Subtitle

Master's Thesis

Eckhart Immerheiser
Department of ...

Advisors: Egon Hasenfratz-Schreier
Supervisor: Prof. Dr. Luc van Gool

January 10, 2020

Abstract

The abstract gives a concise overview of the work you have done. The reader shall be able to decide whether the work which has been done is interesting for him by reading the abstract. Provide a brief account on the following questions:

- What is the problem you worked on? (Introduction)
- How did you tackle the problem? (Materials and Methods)
- What were your results and findings? (Results)
- Why are your findings significant? (Conclusion)

The abstract should approximately cover half of a page, and does generally not contain citations.

Acknowledgements

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Give an introduction to the topic you have worked on:

- *What is the rationale for your work?* Give a sufficient description of the problem, e.g. with a general description of the problem setting, narrowing down to the particular problem you have been working on in your thesis. Allow the reader to understand the problem setting.
- *What is the scope of your work?* Given the above background, state briefly the focus of the work, what and how you did.
- *How is your thesis organized?* It helps the reader to pick the interesting points by providing a small text or graph which outlines the organization of the thesis. The structure given in this document shows how the general structuring shall look like. However, you may fuse chapters or change their names according to the requirements of your thesis.

1.1 Focus of this Work

1.2 Thesis Organization

Chapter 2

Related Work

Describe the other's work in the field, with the following purposes in mind:

- *Is the overview concise?* Give an overview of the most relevant work to the needed extent. Make sure the reader can understand your work without referring to other literature.
- *Does the compilation of work help to define the "niche" you are working in?* Another purpose of this section is to lay the groundwork for showing that you did significant work. The selection and presentation of the related work should enable you to name the implications, differences and similarities sufficiently in the "discussion" section.

2.1 Diffusion Models

2.1.1 Forward Diffusion Process

$$\begin{aligned} q(x_t \mid x_{t-1}) &= \mathcal{N}(\sqrt{1-\beta_t} x_{t-1}, \beta_t I) \\ q(x_t \mid x_{t-1}) &= \mathcal{N}(\sqrt{\alpha_t} x_{t-1}, (1-\alpha_t) I) \\ q(x_t \mid x_0) &= \mathcal{N}(\sqrt{\bar{\alpha}_t} x_0, (1-\bar{\alpha}_t) I) \\ \bar{\alpha}_t &= \prod_{i=0}^{t} \alpha_t \end{aligned} \tag{2.1}$$

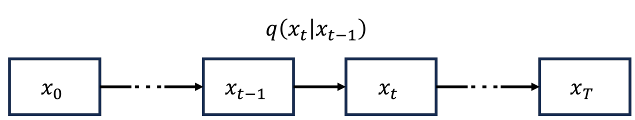

Figure 2.1: Forward Diffusion Process: An image is iteratively destroyed by adding normally distributed noise, according to a schedule. This represents a Markov process where the transition probability $q(x_t|x_{t-1})$.

2.2 Excursion to Bayes

Before getting started we need to quickly define the terms used in the next section, since they all stem from Bayesian statistics. The Bayesian theorem can be written like this:

$$p(z|x) = \frac{p(x|z)p(z)}{p(x)} \tag{2.2}$$

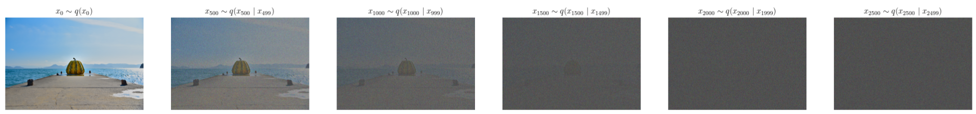

Figure 2.2: Example of Iterative Image Destruction through Forward Diffusion Process: The indices give the time step in the iterative destruction process, where β was created according to a linear noise variance schedule (5000 steps from in the 0.001 to 0.02 range and picture resolution of 4016 by 6016 pixels).

It is implicitly assumed here that p is a probability density function over two continuous random variables x and z. The formula holds in general, but in generative machine learning we usually assume that z represents a random variable in latent space (unobserved) from which we will eventually sample to generate new samples, whereas x is the random variable that represents the training images (observed space). Using above described ordering, the four terms in this formula use distinct names:

$p(z|x)$ is called the *posterior*

$p(x|z)$ is called the *likelihood*, since it gives the literal likelihood of observing an example x when choosing the latent space to be a specific z.

$p(z)$ is called the *prior*, since it exposes information on z before any conditioning.

$p(x)$ is called the *evidence*, since it encompasses our actual observations.

One of the most straightforward examples of a generative model where we search for such a latent space representation of our distribution over the training examples, is the Variational Autoencoder (VAE) [4]. The name of the VAE stems from the Autoencoder, a network that tries to recreate its output through a bottleneck and thereby learning a compressed representation of the data. [5] It bears similarity to other dimension reduction methods like Principal Component Analysis (PCA) and therefore was first published under the name *Nonlinear principal component analysis*. The *variational* part in the VAE stems from the fact that it tries to reduce the data not into an arbitrary low dimensional latent space, but into a latent parameterized distribution (usually i.i.d multivariate Gaussian). This distribution is sampled in the forward pass (therefore *variational*, since we use a stochastic layer) and reproducing the input is now not a feasible loss function, but maximizing the likelihood is. Maximizing the likelihood $p(x|z)$ from above means that we want to tune the parameters of this latent distribution such that the produced output is “likely” an example that could come from the original distribution. Training generative models such as a VAE or also a GAN is usually either done with *Evidence Lower Bound* as the loss, or with an additional network and an *adversarial loss*. [3] Both examples will be further explained in the next sections.

2.3 Loss Functions

In generative machine learning we would want our model to learn the distribution that generated out training examples. Often this distribution is conditioned on some description (e.g. text) or on the corruption process in our case, where we use generative models to solve inverse problems. Assuming our original data distribution (of images) is $p(x)$, then we try to find a parameterized variational machine learning model ($q_\theta(x)$) that will closely match the data distribution.

In order for this $q_\theta(x)$ to be trained we need a differentiable loss function that expresses “closeness” in a distributional sense. The usual approach to this is to use the Kullback-Leibler (KL) divergence.

2.3.1 Kullback-Leibler Divergence

2.3.2 Wasserstein Distance

A different approach to comparing the similarity of distributions is the Wasserstein metric, successfully used in the Wasserstein GAN. [1].

Chapter 3

Materials and Methods

The objectives of the "Materials and Methods" section are the following:

- *What are tools and methods you used?* Introduce the environment, in which your work has taken place - this can be a software package, a device or a system description. Make sure sufficiently detailed descriptions of the algorithms and concepts (e.g. math) you used shall be placed here.
- *What is your work?* Describe (perhaps in a separate chapter) the key component of your work, e.g. an algorithm or software framework you have developed.

Chapter 4

Experiments and Results

Describe the evaluation you did in a way, such that an independent researcher can repeat it. Cover the following questions:

- *What is the experimental setup and methodology?* Describe the setting of the experiments and give all the parameters in detail which you have used. Give a detailed account of how the experiment was conducted.

- *What are your results?* In this section, a *clear description* of the results is given. If you produced lots of data, include only representative data here and put all results into the appendix.

Chapter 5

Discussion

The discussion section gives an interpretation of what you have done [2]:

- *What do your results mean?* Here you discuss, but you do not recapitulate results. Describe principles, relationships and generalizations shown. Also, mention inconsistencies or exceptions you found.
- *How do your results relate to other's work?* Show how your work agrees or disagrees with other's work. Here you can rely on the information you presented in the "related work" section.
- *What are implications and applications of your work?* State how your methods may be applied and what implications might be.

Make sure that introduction/related work and the discussion section act as a pair, i.e. "be sure the discussion section answers what the introduction section asked" [2].

Chapter 6

Conclusion

List the conclusions of your work and give evidence for these. Often, the discussion and the conclusion sections are fused.

References

[1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. *Wasserstein GAN*. 2017. arXiv: 1701.07875 [stat.ML].

[2] R.A. Day and B. Gastel. *How to Write and Publish a Scientific Paper*. Cambridge University Press, 2006.

[3] Ian J. Goodfellow et al. *Generative Adversarial Networks*. 2014. arXiv: 1406.2661 [stat.ML].

[4] Diederik P Kingma and Max Welling. *Auto-Encoding Variational Bayes*. 2022. arXiv: 1312.6114 [stat.ML].

[5] Mark A. Kramer. "Nonlinear principal component analysis using autoassociative neural networks". In: *AIChE Journal* 37.2 (1991), pp. 233–243. DOI: https://doi.org/10.1002/aic.690370209. eprint: https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.690370209. URL: https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.690370209.

Appendix A

The First Appendix

In the appendix, list the following material:

- Data (evaluation tables, graphs etc.)
- Program code
- Further material