Denoising Diffusion Probabilistic Models for Image Inpainting

Pranav Talluri

2022/23

Declaration of Originality

Proforma

Contents

1 Introduction

TODO

1.1 Motivation

Deep generative models ”enabled scalable modeling of complex, high-dimensional data including images, text, and speech” - https://deepgenerativemodels.github.io/

”GAN models are known for potentially unstable training and less diversity in generation due to their adversarial training nature. VAE relies on a surrogate loss. Flow models have to use specialized architectures to construct reversible transform.” - https://lilianweng.github.io/posts/2021-07-11-diffusion-models/

DPMs - parameterised markov chains trained to denoise data inspired by non-equilibrium statistical physics

DDPMs - DPMs trained on a weighted variational bound designed according to a connection between DPMs and denoising score matching with Langevin Dynamics

DDPMs offer a solution. . .

Applications of DDPMs. . .

Image inpainting. . .

1.2 Previous Works

2 Preparation

PrepWork...

- logistic distribution
  - normal dist but with heavy tails
  - "increases the robustness of analyses based on it compared with using the normal distribution"
- PixelCNN++
  - openai implementation of PixelCNN
  - tractable likelihood
  - "model fully factorizes the probability density function on an image x over all its sub-pixels"
  - modification to PixelCNN - discretised logistic mixture likelihood rather than softmax
  - modification - conditions on whole pixels rather than rgb vals
  - modification - downsampling to encourage long range dependencies
  - modification - shortcut connections
  - modification - standard binary dropout to prevent overfitting
- DDPMs (modified PixelCNN++)
  - replaced weight normalisation to group normalisation - simplicity
  - 4 resolution levels for 32x32 and 6 for 256x256
  - two convolutional residual blocks per resolution level
  - "self-attention blocks at the 16x16 resolution between the convolutional blocks"
  - dropout rate set by sweeping over values
  - linear beta schedule
  - random horizontal flips
  - Adam rather than RMSProp
  - batch size is 128 for CIFAR
  - EMA set to 0.9999 decay factor
- ResNets
  -

Related works...
Maths behind DDPMs...
Alternative methods of image generation...
Alternative methods of image inpainting...

3 Implementation

4 Evaluation

5 Conclusion

Appendicies

Index