

DDPM による画像生成の実装

情報メディア創成学類 202312705 長野由依

2025 年 6 月 19 日

1 はじめに

今回の課題では Denoising Diffusion Probabilistic Models (以下、DDPM) を用いて、MNIST データセットに対する画像生成を試みた。DDPM は、画像生成タスクにおいて GAN を凌駕する性能を示す拡散モデルの一つであり、ノイズを段階的に除去することで高品質な画像を生成するモデルである。

2 モデルの概要

DDPM は、元画像に対して少しずつガウスノイズを加える「前向き拡散プロセス」と、学習によってこの逆過程を再構成する「逆拡散プロセス」から構成される。

学習時には、入力画像 x_0 に対して t ステップ分のノイズを加えた x_t とその時刻 t を与え、ノイズ ϵ を予測するニューラルネットワークを学習する。損失関数には平均二乗誤差 (MSE) を用いた。

3 実装と実験設定

PyTorch を用いて実装を行った。主な設定は以下の通りである。

- データセット：MNIST (28×28 グレースケール画像)
- モデル：4 層の CNN
- 拡散ステップ数： $T = 200$
- エポック数：5
- 最適化：Adam (学習率 $1e-3$)

学習には約 10 分程度を要し、画像の生成には逆拡散を 200 ステップに分けて実行した。

4 結果と考察

学習後にノイズから生成された画像の例を図 1 に示す。

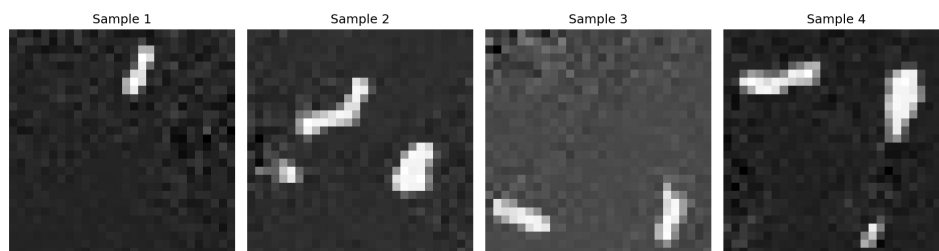


図 1 生成された MNIST 画像 (4 枚)

生成された画像には数字らしきが見られるが、まだ形が不明瞭である。モデルの精度向上のために、CNN の層を増やす、高解像度への拡張、学習エポック数を増やすなどの改善を試みた。

5 あきらめた点とその理由

より高精度な画像生成を目指してモデルの構造を深くし、エポック数を増やすなどの改善を行おうとしたが、実行時間が大幅に伸びてしまった。特に CPU 環境では 1 エポックあたり 5 分以上かかり、実験の反復が困難となったため、今回は簡易モデルによる生成にとどめることとした。

参考文献

- [1] Jonathan Ho, Ajay Jain, and Pieter Abbeel.
Denoising Diffusion Probabilistic Models.
NeurIPS 2020. <https://arxiv.org/abs/2006.11239>