
進捗報告

1 研究発表会資料草稿

研究発表会資料の草稿を github にアップロードしました。生成画像の Inception Score は測定中なので、実験結果以降の章は未完成です。測定結果により、生成画像例や評価指標の表など今現在の暫定版から差し替える可能性があります。

また、テストデータのキャプション入力による生成画像に対する FID に関しては測定可能か検討中です。AttnGAN 学習時の訓練データにない画像と比較してよいのか、または、キャプションが異なる画像と比較してよいのか、という 2 点の問題があります。

来週のゼミで発表練習をさせていただくことになるかと思いますので、その際には研究発表会資料及びスライドの草稿完成版をお見せできるかと思います。

2 バッチサイズ 2 での学習

バッチサイズ 2 のモデルは、結論として学習がうまく進んでいませんでした。

この結果を失敗として残念に思う気持ちもありますが、もともと「AttnGAN の学習時の情報量は不足しているのではないか」という仮説をもとに、今まで研究で色々つけ足してきたということもあるため、バッチサイズ 2 で学習が成功していたら自分の研究が根本から全否定されてしまいかねないという懸念もあり、逆に安心した気持ちも無きにしも非ずです。

また、GAN のバッチサイズに関するちゃんとした学術論文は調べた限り見つかりませんでした。GAN のバッチサイズは不安定さと多様性の表裏一体のトレードオフにかかわる問題になってくるため最適解は存在せず、GAN のモデル構造、画像生成の目的や用途、生成画像に修正を加えるか否か、動作環境等、複雑な要因によって変わるものかと考えられます。

ただし、AttnGAN を含めた text-to-image 系列の GAN のバッチサイズに関して議論しているものは、論文だけでなく qiita にも存在しないと思われるため、「バッチサイズ 2 で失敗した」という事実を検証したというのは、無意味ではないと考えています。

3 生成画像の共有

SAMPLING_11_24.zip に圧縮し、ギガファイル便にアップロードしました。

以下、url とハイパーリンクです。

<https://31.gigafile.nu/0124-jbadf93a5f1070d00e61b7f7ee1df357e>

SAMPLING_11_24.zip