

指示文に基づく高品質脳MR画像生成手法の初期検討

中津颯太[†], 西牧慧[‡], 戸張柊也[‡], 彌富仁[†]

[†]法政大学 理工学部 応用情報工学科 [‡]理工学研究科 応用情報工学専攻

概要

MRIの蓄積された医用画像情報は診断支援などの2次利用が期待され、機械学習技術を用いた様々な技術が開発されているが、良質な学習データの確保に課題がある。生成モデルの一種であるGANは近年医用画像生成に利用されているが、生成できる画像の多様性に制限があった。本研究では、拡散モデルであるStable DiffusionにDreamBoothによる追加学習の枠組みを適用し、テキストによる指示が可能な新しい脳MR画像生成法の初期検討を行った。加えて、そのモデルと生成画像の有効性を、生成画像に対するモデルの尤度計算とCNNにより構築された識別器のアルツハイマー病の予測能による評価を行った。

1 背景

医用画像の診断支援において、MRIは重要な役割を担っており、アルツハイマー病や神経疾患、脳腫瘍の診断など幅広く使用されている。機械学習技術の急激な進歩とともに診断支援技術も近年著しい進展を遂げているが、解析に使用されるモデルを学習するには大量の医用画像が必要となる。しかし、医用画像の収集は患者のプライバシー等多くの課題があるため、大量にデータを共有することが困難である。患者のプライバシー保護を確保しつつ合成データ生成によりデータを適切に増やす事ができれば、様々な応用でデータ不足に起因する問題の低減が期待できる。

これまで近年の実践的な生成モデルの1つである敵対的生成ネットワーク(GAN)はリアルな画像の生成に成功し、医用画像の生成にも貢献している。

洗練されたGANは、高品質な画像生成が可能であるが、指定したクラスの学習データをもとに生成が行

Preliminary study on the generation of high-quality brain MR images based on instruction

Sota Nakatsu[†], Kei Nishimaki[‡], Shuya Tobari[‡] and Hitoshi Iyatomi[†]

[†]Applied Informatics, Faculty of Science and Engineering, Hosei University
184-8584, Tokyo, Japan

{sota.nakatsu.2c@stu., kei.nishimaki.6b@stu,
shuya.tobari.7b@stu., iyatomi@}hosei.ac.jp

われるため、画像の多様性は限られていた。一方、近年ではより多様性に富んだ拡散モデルが大きく注目を集めている。特にLatent Diffusion Models (LDM) [1]は、cross attentionの活用によりテキストから画像を生成するという点においてもGANにはない大きな長所を有し、またそれまでに学習した他クラスの画像も活用することで多様な画像の生成を可能にしている。

LDMを基に大規模なデータセットで学習しテキストから画像を生成する拡散モデルがStable Diffusion (SD)である[2]。また、追加学習としてモデルに対して新しい概念を学習させるDreamBooth (DB)は、画像とテキストの関係を直接学ぶ仕組みのため、他の追加学習枠組みより学習した画像の形状とコンセプトを保持した画像の生成を可能にしている[3]。

本稿では、SDにDBを活用し、学習用の脳MR画像として活用することを目的とした健常者とアルツハイマー病患者の脳MR画像の生成を試み、効果的な方法、指示文を模索した結果を報告する。生成した画像の評価には、カーネル密度推定を用いた尤度ならびに、別途構築した機械学習モデルによる分類能を用いた。

2 高品質脳MR画像生成

2.1 データセットと前処理

本研究では、脳MR画像の生成のために、軽度認知障害やアルツハイマーの研究を目的として収集された3次元脳MR画像のデータセットであるADNI2のAlzheimer's disease (AD)とCognitively normal (CN)の脳画像をそれぞれ300枚と400枚採用した。前処理として深層学習に基づく脳領域の抽出処理(skull stripping)[4]によって頭蓋骨の除去とMR画像の回転と平行移動による体積、角度の補正を行った。前処理を施した脳画像の冠状面の中心のスライスを切り取り、解像度を1mm間隔、画像を160×160にリサイズし、DBを活用したSDへの追加学習に用いた。

2.2 方法

ADNI2データセットのAD脳画像を学習する際は”AD, a MRI brain, ADNI2”という指示文を、CN

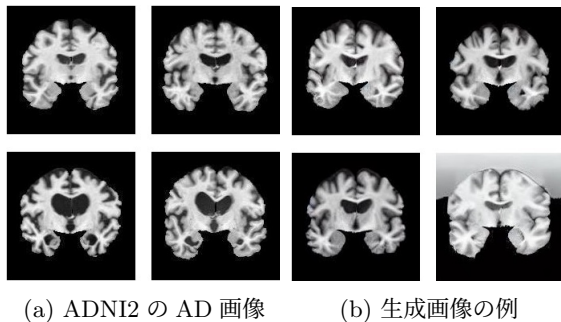


図 1: 使用画像と生成画像の一例

脳画像を学習する際は”CN, a MRI brain, ADNI2”という指示文を学習プロンプトとして入力し DB にて学習を行う。本実験では、より良い画像を取得できる条件を模索するため、学習プロンプトから”a MRI brain”を除いた場合の条件を加えてモデルを学習し、それぞれ画像を 100 枚生成した。

2.3 評価

生成された画像の質の評価のため、本実験では同じ条件で作成した 100 枚の脳 MR 画像の視覚的評価ならびに、2つの定量評価を行った。(1) 学習データと生成した AD と CN の画像のカーネル密度推定に基づく、生成画像の対数尤度を計算し平均値を求めた。(2) 学習データを元に AD と CN の 2次元 CNN による判別器を構築し、その診断能の評価を行った。この時、明らかに生成に失敗している画像は、(1) の尤度推定で明らかに低い値を示すため容易に検出できるため除外すると共に、その割合を記録した。

3 結果

表 1 に生成画像に対する評価結果を示す。この結果からどの条件においても AD に識別された画像の枚数は極端に少なくなっていることが分かる。また利用したデータセットと作成した生成画像の一例を図 1 に示す。また、外れ値として処理を行ったのは”a MRI brain”プロンプトを用いずに作成した時に生成された 1 枚のみである。

表から読み取れる尤度の値は、どの条件下でも CN データセット画像に対する対数尤度が高い傾向にあることが分かる。また、CNN による生成画像の AD/CN 分類でもどちらの条件下も識別結果は CN に偏っている。しかし、プロンプトに”a MRI brain”を除いた条件での生成画像の方が比較的 AD に分類されている枚数が多くなっている。また、対数尤度から誤分類されたと見なせる画像枚数は AD として生成した場合には

極端に高くなっている。外れ値として処理した画像に対しては大幅に尤度が大きくなっていることが分かる。

表 1: 対数尤度と識別器による評価結果

生成画像	AD 尤度	CN 尤度	AD 予測	CN 予測	誤分類 (尤度)
AD	-423.95	-412.64	3	97	98
CN	-423.82	-412.03	2	98	0
AD(brain なし)	-426.56	-418.90	7	92	97
CN(brain なし)	-426.08	-415.96	5	95	3
外れ値	-507.05	-502.40	—	—	—

4 考察とまとめ

(1) の結果から全ての条件において CN 画像に対する尤度は AD 画像に対する尤度より高くなっている。また、対数尤度から生成した画像を分類し、誤分類された枚数も AD として生成した場合が極端に多くなっている。これらは 2つの症例を学習する際すべての条件で”ADNI2”など共通のプロンプトが存在しており、脳画像の情報が混同し、よりデータ数の多い CN の特徴に偏ったものと考えられる。(2) の結果からプロンプトに”a MRI brain”を除いたものは比較的 AD に識別されることが多かった。これは SD のモデルが”brain”の画像を事前に学習を行っており、CN 画像に近い画像が多かったことで全体的に CN に偏ったものの、この条件下ではその影響が低減されたと考えられる。以上より、本実験における条件下では、AD の生成画像を出力することが難しいことが分かった。今後の展望として AD 画像を生成する工夫をさらに模索し、医療現場の診察に足るデータセットの拡張を目指す。

謝辞

本研究は JSPS 科研費 21K12656 の助成を受けた。

参考文献

- [1] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models, 2022.
- [2] Robin Rombach, A. Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10674–10685, 2021.
- [3] Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, and Kfir Aberman. Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 22500–22510, 2022.
- [4] Kei Nishimaki, Kumpei Ikuta, Shingo Fujiyama, Kenichi Oishi, and Hitoshi Iyatomi. Pcss: Skull stripping with posture correction from 3d brain mri for diverse imaging environment. *IEEE Access*, Vol. 11, pp. 116903–116918, 2023.