

Latent Diffusion Model を用いた脳 MR 画像のドメイン調和の評価

池上宙[†] 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

医療現場において、蓄積された脳 MR 画像の活用が強く望まれている。診断支援を目的とした機械学習タスクでは、拠点ごとの MR 画像撮像器や撮像プロトコルの違いが疾患に基づく特徴抽出の妨げになるドメインシフトとよばれる課題が存在する。従来の画像そのものを対象とした GAN をベースとしたドメイン調和手法は優れた効果が期待できる反面、ドメイン毎の学習が必要であり、また大きな違いのある画像同士の調和が困難である。近年注目されている生成技術は、様々なタスクと親和性が高くドメイン調和へも応用が期待できる。本研究ではテキストから画像生成が可能な潜在拡散モデルを活用し脳 MR 画像の調和のための画像変換について検討した。

1 はじめに

医療診断支援のために、機械学習モデルによる脳 MR 画像を使用した研究が行われている。近年の研究の大規模化に伴い、撮影された機器や撮像プロトコルといった撮像拠点（ドメイン）毎の非生物学的な条件の差がモデルの性能に悪影響を及ぼしていることが明らかになり [1]、こうしたドメイン差を取り除く調和を行うことが求められている。

調和の方法には大きく分けて二つのアプローチがある。一つは撮像された脳 MR 画像を特定のタスクを実行する過程で得られる低次元表現上でドメイン差を取り除く方法である。Dinsdale らは反復学習構造の敵対的ドメイン適応 [2] が、深層学習タスクの調和におけるデータセットバイアス除去に適応でき、データセットによる偏りを調和した低次元表現を獲得できることを示した。このアプローチでは低次元表現を学習するため学習コストが低いという利点があるが、目的に応じて毎回モデルを設計する必要があり、組み込みの柔軟性が低いことが課題として挙げられる。

もう一つは、生成モデルによるスタイル変換技術を使用して画像を直接処理する方法である。Arai らの CycleGAN を用いた調和手法 [3] では、二つの異なる機器で撮像された画像のうち一方にスタイル変換を行

い、もう片方の機器で撮像されたかのようにすることで、機器の違いによるドメイン差を取り除くことに成功した。しかしこのような CycleGAN をベースにする手法では、スタイル変換を行う際に調和するドメイン差を 1 対 1 で学習を行う必要があり、3 種類以上のドメインの調和に適応するためには、多くのモデルを用意する必要がある。また、画像の形状を大きく変えることができないため、ドメイン間にサイズや形状の違いがある場合に調和が困難であった。

一方、指示文からの画像生成 (text-to-image) や編集 (image-to-image) が可能な Latent Diffusion Model (LDM) [4] が近年注目されており、LDM を大規模データで学習させた Stable Diffusion は広く利用されている。本研究では、指示文を活用できる Stable Diffusion の image-to-image 変換を利用した新しい脳 MR 画像のドメイン調和のための新しい方法を試作し評価した。Stable Diffusion モデルに対し、画像の構造をより忠実に反映できる DreamBooth [5] を活用し脳 MR 画像の追加学習を行い、入力された MR 画像が特定の撮像器で撮像された画像であるように変換する。本手法は 1 つのモデルで無数のドメインに対応する調和技術の実現が期待できる。

2 拡散モデルを用いた脳 MR 画像の調和

2.1 データセットと前処理

本研究では、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 MR 画像のデータセットである Alzheimer's Disease Neuro imaging Initiative (ADNI) ¹2 と 3 を使用した。ADNI2 から 1000 枚、ADNI3 から 1861 枚を学習に使用し、前処理として、頭蓋骨除去、体積補正、角度補正を行った [6]。前処理を施した 160 × 224 × 160 の画像から冠状面をランダムに 10 スライス切り取り、解像度を 1mm 間隔にして追加学習に用いた。

2.2 手法の詳細

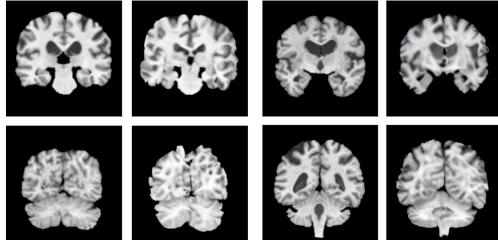
本研究では、Stable Diffusion を用いた image to image によるドメイン調和を試作した。imgae to image は、入力画像に対してノイズを付与し、ノイズを除去する過程で、指定したプロンプトに適応するように変換する手法である。この手法を利用して、1 枚の脳 MR 画像を入力し、プロンプトで指定したデータセットに変換することで、ドメイン変換が行える可能性がある。この手法の利点は、3 種類以上のドメイン差の調和を

Evaluation of domain harmonization in Brain MR Images using a latent diffusion model
SORA IKEGAMI[†], KEI NHISHIMAKI[‡], SHUYA TOBARI[†] and HITOSHI IYATOMI[†]
Applied Informatics, Faculty of Science and Engineering, Hosei University 184-8584, Tokyo, Japan
{sora.ikegami.6u@stu., kei.nishimaki.6b@stu., shuya.tobari.7b@stu., iyatomi@}hosei.ac.jp

¹<https://adni.loni.usc.edu/>

表 1: 生成時のパラメータ

パラメータ	値
seed	2500
num_inference_step	1000
guidance_scale	100
num_images_per_prompt	100



(a) ADNI2からADNI3 (b) ADNI3からADNI2
図 1: image to image の結果 (左: 入力, 右: 出力)

行う際に必要なモデルは、使用する画像のドメインを学習したモデル1つとなり、先行研究の課題であった学習コストを抑制することが期待できる。全てのデータセットの脳MR画像を入力画像の対象としているため、1つのモデルで様々なドメインに変換することが可能である。脳MR画像は3次元画像であるが、本実験では初期検討として、2次元スライスの画像に対して変換した。

Stable Diffusion v-1.5 モデルに、DreamBoothで脳MR画像の追加学習したモデルを変換に用いた。モデルの学習の際に、instance prompt にデータセット名をclass prompt に"a brain MR image"を使用した。

追加学習にて構築したモデルを用いて、脳MR画像から任意の冠状面スライスを入力し、image to imageにて、変換先データセットの名前("ADNI3"または"ADNI2")をプロンプトにしてスタイル変換した。画像を変換する際のパラメータの設定を表1に示す。guidance scale は入力画像とプロンプトの影響を決めるパラメータであり、値が大きいほど影響力が強まる。デフォルト値は7.5であるが、予備実験の結果、本実験では100に設定した。

2.3 評価

変換した画像が、脳の構造を失っていないかを確認するため、ADNI3からADNI3の同一ドメインによる変換を行い、SSIM (Structural Similarity Index Measure) とピクセル間の平均的な差の二乗を計算したMSE (Mean Squared Error) の2つの指標で評価した。

3 実験と結果

図1aにADNI2の画像をADNI3に変換した結果を、図1bにADNI3の画像をADNI2に変換した結果を示す。視覚的には形が崩れずに変換画像が生成された。

ADNI2からADNI3及びADNI3からADNI3に変換した画像を図2に示す。また、100枚の変換した画像と元画像との評価スコアの平均±標準偏差と最も良

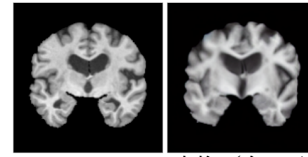


図 2: ADNI3からADNI3の変換 (左: 入力, 右: 出力)

表 2: スコア

	ADNI2 to ADNI3		ADNI3 to ADNI3	
	平均 (± 標準偏差)	最良	平均 (± 標準偏差)	最良
SSIM	0.581(±0.046)	0.652	0.601(±0.054)	0.697
MSE	0.038(±0.023)	0.023	0.031(±0.012)	0.018

かったスコアを表2に示す。SSIMの値は0.7に達せず、MSEでは0.03を上回る思わしくない結果が得られた。

4 考察とまとめ

図1と図2より、変換によって、定性的には脳の大きな構造を失っていないように考えられるが、表2より、SSIMの値が低かった。これは、指定したプロンプトの影響を強く受けて、学習画像の形状に変えてしまっているからだと考えられる。今回はguidance scaleを100まで高めている。しかし、これによってプロンプトの影響が高くなりすぎてしまい、入力画像よりも構造が大きく変わってしまった。今後の展望として、入力画像とプロンプトの重みを変えて生成できるような生成方法を確認し、脳の形状を変えないスタイル変換を実現していく。

5 謝辞

本研究はJSPS 科研費21K12656の助成を受けた。

参考文献

- [1] E. Tzeng, J. Hoffman, T. Darrell, and K. Saenko, "Simultaneous deep transfer across domains and tasks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015.
- [2] X. Han, J. Jovicich, D. Salat, A. van der Kouwe, B. Quinn, S. Czanner, E. Busa, J. Pacheco, M. Albert, R. Killiany *et al.*, "Reliability of mri-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer," *Neuroimage*, vol. 32, no. 1, pp. 180–194, 2006.
- [3] H. Arai, Y. Onga, K. Ikuta, Y. Chayama, H. Iyatomi, and K. Oishi, "Disease-oriented image embedding with pseudo-scanner standardization for content-based image retrieval on 3d brain mri," *IEEE Access*, vol. 9, pp. 165 326–165 340, 2021.
- [4] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 10 684–10 695.
- [5] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, "Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 22 500–22 510.
- [6] K. Nishimaki, K. Ikuta, S. Fujiyama, K. Oishi, and H. Iyatomi, "Pcss: Skull stripping with posture correction from 3d brain mri for diverse imaging environment," *IEEE Access*, 2023.