

MRI モダリティを用いた 2D / 3D 深層学習モデルによる 脳腫瘍セグメンテーション性能の比較評価

Instructions to Authors

○荒川侑輝*, 曾昱雯**, 張曉勇***, 周雯童†, 杉田典大‡, 本間経康§

○Yuki Arakawa*, Zeng Yuwen**, Zhang Xiaoyong***,
Zhou Wentong†, Norihiro Sugita‡, Noriyasu Homma‡

*東北大学, **東北大学, ***仙台専門高等学校,
†東北大学, ‡東北大学, §東北大学

*Tohoku University, **Tohoku University, ***National Institute of Technology,
Sendai College.

†Tohoku University, ‡Tohoku University, §Tohoku University.

キーワード：畳み込み演算 (Convolution), 特徴抽出 (Feature Extraction), 転移学習 (Transfer Learning), データ拡張 (Data Augmentation), 意味的分割 (Semantic Segmentation)

連絡先：〒980 - 8579 仙台市青葉区荒巻字青葉 6-6-05 東北大学工学研究科・工学部 電子情報システム・応物系 1 号館, Tel.: (022)795-7130, E-mail: web-admin.cps-cc@grp.tohoku.ac.jp

1. 諸言

はじめに, 脳腫瘍は頭蓋骨の中でできる腫瘍の総称であり, 2024 年において年間約 2 万人が発症する病気である. 主な症状として, 頭痛や吐き気, まひ, 歩行障害, しびれ, ふらつきなどが挙げられる [1]. しかし, 脳は感覚, 運動機能など様々な機能を司る神経系の中枢機関であり, 発生箇所によって症状や腫瘍の種類が多岐にわたる. そのため, 治療もまた, これらの症状や発生箇所に合わせた処置を行わなければならない. こうした治療プロセスの第 1 段階として,

Magnetic Resonance Imaging (MRI) 画像や Computed Tomography (CT) 画像による画像診断が行われる.

しかし, MRI 画像や CT 画像は人体のスライス画像を複数枚組み合わせた 3D ボリュームの画像であり, 読影に必要な平均所要時間は, 1 件あたり 14.6 分である. これに検査適応決定や撮像方法の指示, 画像診断報告書の管理を含めると, 1 時間に 4 件以上の読影では質の担保が困難であり [2], 診断における医師の負担が大きい. また, CT や MRI は撮影時における患者の動きに非常に敏感であるため, 画像が不鮮明で腫瘍の境

界の診断が難しい場合が存在する [3]. このような診断上の負担を軽減し、腫瘍の抽出を適切に行うための一助として深層学習による脳腫瘍セグメンテーションが注目されている.

2. 先行研究

2.1 Zhou らによる研究

Zhou らの研究 [4] では 3D U-net [5] と Unetr [6] によるセグメンテーションタスクを行った. 本研究では医療画像における深層学習の基本モデルである U-net に着目する. 以下の Fig.1 に 3D U-net のモデルを示す. エンコーダ部分で画像を畳み込み処理, Max pooling 層で画像圧縮する. これらの処理により, 画像のグローバルな情報を抽出する. 一方, デコーダ部分ではスキップ接続で圧縮前の情報を補完しつつ, 転置畳み込みによって画像を拡大することで高解像度で特徴を掴みセグメンテーションを行う.



Fig. 1: 3D U-net モデル: 3D 画像を入力として, 画像の高さ, 幅, 深さ, チャンネルの 4 次元で学習を行う. エンコーダ部分で画像を畳み込み処理, Max pooling 層で画像圧縮する. これらの処理により画像のグローバルな情報を抽出する. 一方, デコーダ部分ではスキップ接続で圧縮前の情報を補完しつつ, 転置畳み込みによって画像を拡大することで高解像度で特徴を掴みセグメンテーションを行う.

2.2 Multi Modality

脳 MRI 画像には, T1n, T1c, T2, FLAIR の 4 種のモダリティが存在する. Fig2 に示すように, これらはそれぞれ異なる特徴を有する. T1n, T1c では水分が多い部分 (浮腫や脳脊髄液など) は暗く映すが脂肪などの脳組織は明るく映し, T1c は T1n の撮影においてガドリニウムと呼ばれる造影剤を投与した後の画像である. そのため, T1c 画像は T1n 画像に比べて腫瘍の境界がより鮮明に映される. 対して, T2, FLAIR では水分が多い部分を明るく映す. また, FLAIR は T2 のうち脳脊髄液の信号を抑制し, 腫瘍の境界が明瞭となる. これら画像が持つ特徴によって, 腫瘍位置はシーケンス画像ごとに異なる. そのため, 1 つのみのシーケンス画像から腫瘍位置, 大きさを特定することは正確性に欠ける. これを改善するため, 4 つのモダリティを用いてセグメンテーションを行う.

2.3 Dataset

データセットは脳腫瘍の MRI 画像と専門医によるアノテーション画像を含む公開データセットである BraTS2023 [9] を使用した. また, Training 1000 cases, Validation 200 cases, Test 50 cases で学習を行った. さらに, セグメンテーション領域は造影増強腫瘍領域 (Enhancing Tumor:ET), ET と壊死・非影像領域 (Necrotic Tumor Core:NCR) を合わせた Tumor core (TC) 領域, TC 領域に浮腫 (Edema:ED) を合わせた Whole Tumor (WT) 領域の 3 つに分割して評価を行い, 3 つの領域を Fig.3 に示す.

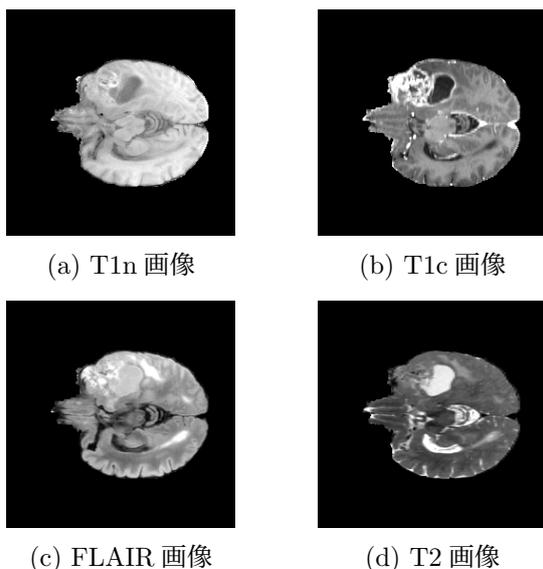


Fig. 2: 4 種のモダリティ MRI 画像：(a) T1n 画像は浮腫や脳脊髄液などは暗く映し、脂肪などの脳組織は明るく映す (b) T1c 画像は T1n 画像に対して造影剤であるガドリニウム投与後の画像 (c) FLAIR 画像は T2 のうち脳脊髄液の信号を抑制し、腫瘍の境界が明瞭 (d) T2 画像

2.4 Loss

深層学習における Loss は Dice Loss [10] および Cross-Entropy Loss [8] を組み合わせたものを使用した。Dice Loss を式 1, Cross-Entropy Loss を式 2, それらを組み合わせたものを式 3 に示す。ここで、式 1 で使用している p は予測領域, g は正解領域, N は総ピクセル数を表す。

$$dice\ loss = 1 - \frac{\sum_{i=1}^N (p_i * g_i)}{\sum_{i=1}^N p_i + \sum_{i=1}^N g_i} \quad (1)$$

また、式 2 で使用している C はクラス, w は重みであり、それぞれの重みは (0.2, 0.3, 0.25, 0.25) で (背景, ET, TC, WT) に対応している。 g_i はピクセル i の正解クラスを one hot 形式で示したものであり, p_i はピクセル i が各クラスに属する予測確立である。

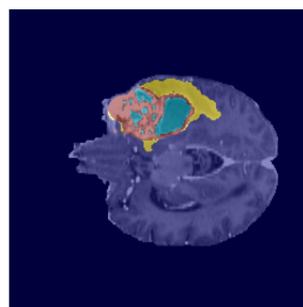


Fig. 3: セグメンテーション領域：ET が青色, NCR が赤色, ED が黄色の領域

$$ce\ loss = \sum_{i=1}^N \sum_{c=1}^C w_c \cdot g_i^c \cdot \log(p_i^c) \quad (2)$$

$$Loss = dice\ loss + ce\ loss \quad (3)$$

テスト時の性能評価は Dice Similarity Coefficient (DSC) を用いた。その式を式 4 に示し、結果を以下の Table1 に示す。

$$DSC = \frac{\sum_{i=1}^N (p_i * g_i)}{\sum_{i=1}^N p_i + \sum_{i=1}^N g_i} \quad (4)$$

スコアは百分率で表されており、数値が高いほどセグメンテーションが上手く行われていることを示す。各領域、モデル毎のスコアの最大値は太字で表記されている。本研究では、3D U-net モデルのスコアに着目し、これらのスコアと 2D U-net モデルによるスコアとを Dataset や Loss 等の条件を揃えて比較する。

2.5 3D セグメンテーションにおける課題

以上で述べた 3D セグメンテーションには 2 つの課題が挙げられる。1 つめの課題として計算量の問題が挙げられる。2.1 節で

Table 1: Zhou らのセグメンテーションスコア [4]

Modalities				Segmentation Regions						Average	
				ET		TC		WT			
T1c	T1n	FLAIR	T2w	3D U-net	UNETR	3D U-net	UNETR	3D U-net	UNETR	3D U-net	UNETR
•	○	○	○	82.21	68.31	86.50	67.58	75.84	62.09	81.52	65.99
○	•	○	○	46.11	2.38	66.48	16.59	78.05	78.05	63.55	32.34
○	○	•	○	52.06	29.59	70.72	37.18	88.64	78.54	70.47	48.44
○	○	○	•	53.11	31.14	67.62	44.46	84.37	73.60	68.37	49.73
•	•	○	○	81.40	70.76	85.92	69.98	78.00	65.69	81.77	68.81
•	○	•	○	80.72	75.57	86.11	76.44	88.46	82.84	85.10	78.28
•	○	○	•	83.29	76.51	87.02	76.37	86.38	79.18	85.56	77.35
○	•	•	○	56.41	29.44	73.89	46.34	88.77	81.85	73.02	52.54
○	•	○	•	53.31	30.46	68.18	43.87	84.72	75.01	68.74	49.78
○	○	•	•	55.99	35.84	70.64	50.19	88.46	83.46	71.70	56.50
•	•	•	○	81.14	80.14	84.12	82.03	88.04	85.03	84.43	82.40
•	•	○	•	80.44	80.44	82.70	77.82	81.01	79.17	81.38	79.14
•	○	•	•	81.51	80.01	85.04	81.09	89.34	85.79	85.30	82.30
○	•	•	•	54.85	35.38	68.52	50.26	88.41	83.32	70.59	56.32
•	•	•	•	81.32	80.36	85.19	82.12	88.39	86.51	84.97	82.99

述べたように、3D セグメンテーションにおける画像データは4次元の情報を持つため、高い計算機性能が要求される。また、1節で述べたように、MRI 画像はシークエンス画像であることから、すべての方向において均一に鮮明な画像が得られるとは限らない。そこで本研究では、こうした課題を軽減するため、2D 画像による U-net モデル [8] を用いたセグメンテーションを実施し、3D 手法との性能比較を行う。

3. 比較実験

3.1 実験概要

本実験では、Zhou らの先行研究 [4] で使用した 3D U-net モデルに対して 2D U-net モデルを設定して実験を行う。2D U-net は 2D 画像を入力として、高さ、幅、チャンネルの 3次元で学習を行い、中で行う処理に関しては 3D U-net のものと変わらない。また、本実験では、一連の MRI 画像から腫瘍領域を含むスライス画像のみを抽出した。具体的には、腫瘍の最大断面積を有するスライ

スを基準として、その前後 5 枚ずつを加えた計 11 枚を抽出し、2D U-net への入力とした。これは、予備実験において全スライスを用いてセグメンテーションを行った際に十分なスコアが得られなかったためである。腫瘍領域の端部を含むスライスが学習に悪影響を与えると考えられることから、腫瘍の中心付近のスライスに限定することで、学習データの質の向上を図った。

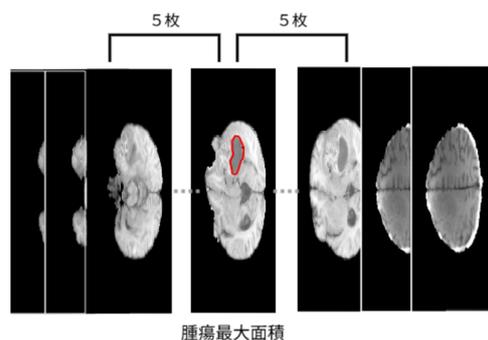


Fig. 4: 2D 変換：腫瘍が存在する画像の中から、腫瘍の最大面積を含む前後 5 枚の計 11 枚を抽出

以上の処理を行い、Zhou らの先行研究 [4] で使用した Dataset, Loss 等の条件を一致

させ、セグメンテーションを行う。モデルは2D U-netを用い、汎化性能を向上させるためのData Augmentation手法も同様に先行研究で用いたものと一致させて実験を行う（使用した処理・・・Random crop, Random Rot Flip, Elastic Transform, Gaussian noise, Contrast, Brightness Transform）。テスト時の評価領域、スコアも先行研究と同様である。

4. 結果

4.1 性能比較

2D U-netによるセグメンテーションスコアと3D U-netによるものを比較した表を以下のTable2に示す。3D U-netのスコアは第2章で示したスコアを参照した。

Table2より、セグメンテーション性能は全体的に3D U-netの方が高いことが確認できる。モダリティ単体での性能を比較すると、T1n, FLAIR, T2wは非常に低いスコアを示した。一方、T1cは比較的高い性能を維持しており、単一モダリティの中では最も有効であることが示された。さらに、モダリティの組み合わせにおいて、2D U-netのアベレージスコアは約77で頭打ちとなるのに対し、3D U-netは最大85.56に達し、両者の間には顕著な性能差が認められた。

また、セグメンテーション結果を以下のFig.5に示す。図中のa, b, c, dはそれぞれT1c, T1n, FLAIR, T2のMRI画像を表し、eは正解ラベル（Ground Truth:GT）、fは2D U-netによるセグメンテーション結果、gは3D U-netによるセグメンテーション結果を示す。緑色がWT 青色TCが赤色がETである。

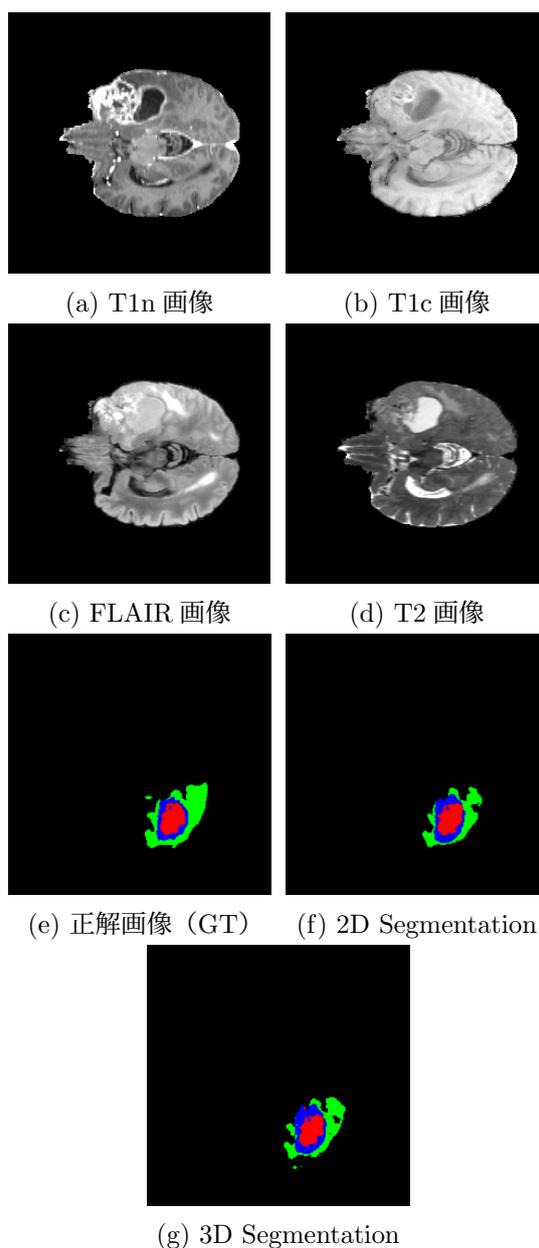


Fig. 5: 2D セグメンテーションの結果および3D セグメンテーションとの比較：(a)T1c 画像は浮腫や脳脊髄液などは暗く映すが脂肪などの脳組織は明るく映し、T1n よりも腫瘍の境界が鮮明 (b)T1 画像は 浮腫や脳脊髄液などは暗く映し、脂肪などの脳組織は明るく映す (c)FLAIR 画像は T2 のうち脳脊髄液の信号を抑制し、腫瘍の境界が明瞭 (d)T2 画像 (e)GT 画像は BraTS2023 に含まれる正解ラベル, (f)2D セグメンテーションの結果 (g)3D セグメンテーションの結果

Table 2: セグメンテーションスコア比較表

Modalities				Segmentation Regions						Average	
				ET		TC		WT			
T1c	T1n	FLAIR	T2w	3D U-net	2D U-net	3D U-net	2D U-net	3D U-net	2D U-net	3D U-net	2D U-net
•	○	○	○	82.21	73.60	86.50	60.62	75.84	79.11	81.52	71.11
○	•	○	○	46.11	40.59	66.48	52.50	78.05	38.54	63.55	43.88
○	○	•	○	52.06	34.00	70.72	67.14	88.64	42.63	70.47	47.92
○	○	○	•	53.11	40.83	67.62	62.61	84.37	46.56	68.37	50.00
•	•	○	○	81.40	72.08	85.92	62.37	78.00	79.45	81.77	71.30
•	○	•	○	80.72	71.01	86.11	77.57	88.46	78.70	85.10	75.76
•	○	○	•	83.29	73.90	87.02	72.94	86.38	79.66	85.56	75.50
○	•	•	○	56.41	46.90	73.89	69.27	88.77	48.35	73.02	54.84
○	•	○	•	53.31	46.47	68.18	66.85	84.72	47.89	68.74	53.74
○	○	•	•	55.99	46.51	70.64	73.13	88.46	51.85	71.70	57.16
•	•	•	○	81.14	72.32	84.12	79.52	88.04	80.32	84.43	77.39
•	•	○	•	80.44	72.44	82.70	74.46	81.01	79.86	81.38	75.59
•	○	•	•	81.51	43.65	85.04	73.61	89.34	49.36	85.30	55.54
○	•	•	•	54.85	71.95	68.52	80.26	88.41	79.27	70.59	77.16
•	•	•	•	81.32	72.59	85.19	79.46	88.39	79.41	84.97	77.15

5. 考察

2D U-net が 3D U-net よりも性能が低くなったことについて考察する。本実験では、トレーニングデータとして脳腫瘍が存在するスライスデータの内、最大面積から前後5枚の合計11枚を用いた。これは3D U-net で用いたデータセットに対して情報量、データ数ともに非常に少ない。このことが影響して、2D U-net の性能が低くなったと考えられる。また、2D モデルはスライス間の空間情報が使えないため、セグメンテーションの精度に悪影響を与えていると考えられる。次に、T1c のセグメンテーションスコアが他の単一モダリティ T1n, FLAIR, T2 よりも高かったことについて考察する。これは、T1c 画像が他のモダリティに比べて腫瘍が鮮明に区別できる画像であり、2D 画像でのセグメンテーションは、境界が不鮮明な画像に対しては不向きであると考えられる。

6. 結言

本研究では先行研究である 3D U-net モデルを用いたセグメンテーションと 2D U-net によるセグメンテーションの性能比較を行った。その結果、T1n, FLAIR, T2 の組み合わせ以外のすべてのモダリティの組み合わせで、3D U-net の方がスコアが優れていることが分かった。しかし、データセットの画像枚数や情報量を減らすことができたため、計算時間やメモリの削減につながり、2D U-net モデルのセグメンテーション精度を向上させることで現場での活用もできる可能性があると考えられる。今後の課題として、やはり 2D セグメンテーションの精度が低いことが挙げられる。精度を向上させる方法として、モデルの改良や Unetr や Attention Unet などの他のモデルの適用が考えられる。特に、Attention Unet はスキップ接続部分に Attention gate と呼ばれる注目領域を関数的に特徴付ける機構があるため、細かい構造や小さな病変部分におけるセグメンテーションの精度向上が期待できる。

参考文献

- 1) 脳腫瘍〈成人〉について：[国立がん研究センター がん情報サービス 一般の方へ]（参照 2026.2.11）
https://ganjoho.jp/public/cancer/brain_adult/about.html
- 2) 【提言】医療の質と安全を担保するための読影量について 2022.02.17（参照 2026.2.17）
<https://jcr.or.jp/news/recommendations20220216/>
- 3) 北里大学病院 放射線部 水上 慎也, 神奈川 MRI 技術研究会「アーチファクト対策」（参照 2026.2.17）
https://kanamri.umin.ne.jp/26-1-1_PDF.pdf
- 4) W. Zhou, Y. Zeng, X. Zhang, R. Li, A. Nagai, M. Kanamori, H. Endo, K. Niizuma, and N. Homma, "Modality-Guided Edge Fusion and Semantic Enhancement for Multi-Modal Brain Tumor Segmentation," in 2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) Workshop on AI-Driven Multimodal Approaches for Brain Disease Diagnosis and Treatment, 2025.
- 5) 'zg'n 'i'ek and Ahmed Abdulkadir and Soeren S. Lienkamp and Thomas Brox and Olaf Ronneberger: 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, MIC-CAI,424"432 (2016).
- 6) Ali Hatamizadeh and Yucheng Tang and Vishwesh Nath and Dong Yang and Andriy Myronenko and Bennett Landman and Holger Roth and Daguang Xu: UNETR: Transformers for 3D Medical Image Segmentation,IEEE (2022)
- 7) W. Zhou, Y. Zeng, X. Zhang, H. Liu, K. Ichiji, and N. Homma, "Comparative Evaluation of MRI Modalities for Deep Learning-Based Brain Tumor Segmentation," in Proc. 60th Anniversary Meeting of the SICE Tohoku Chapter, Sendai, Japan, Dec. 2024, Paper ID: S60-3.
- 8) Olaf Ronneberger and Philipp Fischer and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", Medical Image Computing and Computer-Assisted Intervention " MIC-CAI,234"241 (2015).
- 9) 'BraTS 2023: Brain Tumor Segmentation Challenge", (参照 2026-2-18)
- 10) Fausto Milletari and Nassir Navab and Seyed-Ahmad Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation",International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission-IEEE(2016)
- 11) Adrian Galdran and Gustavo Carneiro and Miguel 'ngel Gonz'lez Ballester, "On the Optimal Combination of Cross-Entropy and Soft Dice Losses for Lesion Segmentation with Out-of-Distribution Robustness",Third Challenge, DFUC, 40"51(2022)