

Ryohei Fukui¹, Masataka Onishi², Koshi Hasegawa³, Miyu Ohata³,
Katsuhiro Kida¹, Sachiko Goto¹

Received: 13.02.2024

Accepted: 23.05.2024

Published: 29.07.2024

Effect of segmentation dimension on radiomics analysis for *MGMT* promoter methylation status in gliomas


Wpływ rodzaju segmentacji na ocenę statusu metylacji promotora *MGMT* w glejakach podczas analizy radiomicznej

¹ Department of Radiological Technology, Faculty of Health Sciences, Okayama University, Okayama, Japan

² Department of Radiological Technology, Faculty of Health Sciences, Okayama University Medical School, Okayama, Japan

³ Department of Radiological Technology, Graduate School of Health Sciences, Okayama University, Okayama, Japan

Correspondence: Ryohei Fukui, Department of Radiological Technology, Faculty of Health Sciences, Okayama University, 2-5-1, Shikata-cho, Kitaku, 700-8558, Okayama, Japan, e-mail: rfukui@okayama-u.ac.jp

 <https://doi.org/10.15557/AN.2024.0002>

ORCID iDs

1. Ryohei Fukui <https://orcid.org/0000-0003-0594-3465>
2. Masataka Onishi <https://orcid.org/0009-0001-6816-1797>
3. Koshi Hasegawa <https://orcid.org/0009-0006-6141-7223>
4. Miyu Ohata <https://orcid.org/0009-0004-2606-750X>
5. Katsuhiro Kida <https://orcid.org/0000-0002-9230-003X>
6. Sachiko Goto <https://orcid.org/0000-0002-2809-6441>

Abstract

Introduction and objective: We investigated the impact of 2D (2D_seg) and 3D (3D_seg) segmentation on the accuracy of prediction models in the radiomics analysis to determine the presence or absence of methylation in the O⁶-methylguanine DNA methyltransferase (*MGMT*) gene promoter region of gliomas. **Materials and methods:** Magnetic resonance imaging images of gliomas were obtained from the Cancer Imaging Archive for 50 methylated and 50 unmethylated cases respectively. For each case, 2D_seg and 3D_seg were performed, and 788 radiomics features, including wavelet transform, were obtained. Ten features were selected by LASSO regression. The coefficients of determination (R^2) and root mean squared error (RMSE) were calculated by multiple regression analysis. Discriminant boundaries to discriminate methylation were created by linear discriminant analysis, and the sensitivity and specificity of each method were calculated. The discriminant accuracy of both methods was evaluated by receiver operating characteristics (ROC) analysis. **Results:** The R^2 value and RMSE were 0.72/0.28 and 0.73/0.33 for 2D_seg and 3D_seg, respectively. Similarly, sensitivity and specificity were 82.5/67.5% and 85/62.5%, respectively. The area under the curve determined by ROC analysis was 0.80 and 0.79, respectively, i.e. slightly larger for 2D_seg. The p -value by the DeLong method was 0.73. **Conclusions:** In the radiomics analysis using 2D_seg and 3D_seg, no difference in discriminant accuracy was observed between them. Therefore, 2D segmentation should be chosen because it is easier to segment.

Keywords: segmentation, glioma, *MGMT*, radiomics

Streszczenie

Wprowadzenie i cel: W pracy zbadano wpływ rodzaju segmentacji – 2D (2D_seg) i 3D (3D_seg) – na dokładność modeli predykcyjnych statusu mutacji promotora genu metyloguaniny (*MGMT*) w glejakach podczas analizy radiomicznej. **Materiał i metody:** Z zasobów zgromadzonych w Cancer Imaging Archive pobrano obrazy glejaków uzyskane metodą rezonansu magnetycznego: 50 przypadków z metylacją promotora *MGMT* i 50 przypadków bez metylacji. Dla każdego przypadku przeprowadzono segmentację 2D_seg i 3D_seg oraz wyodrębniono 788 cech radiomicznych, m.in. transformatę falkową. Przy zastosowaniu modelu regresji LASSO wybrano 10 cech. Wykorzystując analizę regresji wielokrotnej, obliczono współczynniki determinacji (R^2) i średnią kwadratową błąd (*root mean squared error*, RMSE). Przy pomocy liniowej analizy dyskryminacyjnej określono graniczne wartości dyskryminujące umożliwiające różnicowanie statusu metylacji. Następnie obliczono czułość i swoistość każdej z metod. Dokładność dyskryminacyjną obu badanych metod oceniono przy wykorzystaniu analizy krzywych ROC. **Wyniki:** Wartości R^2 i RMSE wyniosły 0,72/0,28 i 0,73/0,33 odpowiednio dla metody 2D_seg i 3D_seg. Uzyskano czułość i swoistość odpowiednio 82,5/67,5% i 85/62,5%. Pole pod krzywą wyznaczone na podstawie analizy ROC wyniosło odpowiednio 0,80 i 0,79, a zatem było nieco większe dla metody

2D_seg. Wartość p obliczona metodą DeLonga wyniosła 0,73. **Wnioski:** W analizie radiomicznej z wykorzystaniem 2D_seg i 3D_seg nie zaobserwowano istotnych różnic pomiędzy metodami pod względem uzyskanej dokładności dyskryminacyjnej. Należy zatem uznać, że metodą zalecaną jest segmentacja 2D z uwagi na jej prostszy przebieg.

Słowa kluczowe: segmentacja, glejak, *MGMT*, radiomika

INTRODUCTION

Brain tumours are classified pathologically into a variety of categories. In particular, twenty-two new classifications have been added to the World Health Organization (WHO) Classification of Tumours, 5th edition (WHO CNS5) (Louis et al., 2021). In addition to nerve cells and nerve fibres, the brain and spinal cord contain glial cells that support them; tumours arising from these cells are collectively called gliomas. Gliomas constitute the most frequent primary intra-axial brain tumour in Japan, accounting for approximately 30% of all malignant tumours (The Committee of Brain Tumor Registry of Japan, 2017). As evident from the development of genetic diagnosis, specific genetic features impact treatment prognosis. For example, O⁶-methylguanine DNA methyltransferase (*MGMT*) is a DNA repair enzyme that removes alkyl groups from the O⁶ position of guanine in DNA and is considered to protect cells from alkylating agents. When the gene promoter region of *MGMT* expressed in glioma cells is methylated, the prognosis has been reported to be favourable for chemotherapy with the alkylating agent temozolomide (Hegi et al., 2005). However, such a genetic diagnosis requires brain tumour cells, which necessitates a highly invasive tissue histopathological examination.

In contrast, a new field of study – radiomics – has been focused on the identification of lesion histology and genotype by using the signal intensity of lesions depicted on computed tomography (CT) and magnetic resonance imaging (MRI) images. Various studies related to radiomics have been actively performed by using CT and MRI images (Feng et al., 2023; Sollini et al., 2021; Yang et al., 2019). Radiomics entails the measurement of relevant features based on the magnitude and distribution of the signal intensity in the lesion; on this basis, the prediction model is constructed. The signal intensity of a lesion is obtained by extracting it from the image; therefore, a segmentation process with regions of interest (ROIs) is necessary. Because CT and MRI images contain three-dimensional (3D) information, segmentation can also be performed three-dimensionally (3D_seg); however, the larger the number of slices, the greater the effort required to set the ROIs. If only one slice with a lesion was selected and a two-dimensional (2D) ROI was set (2D_seg), and if there was no difference between the accuracy of the predicted model and that of 3D_seg, the segmentation process could be considerably simplified. In this study, we compared the accuracy of 2D_seg and 3D_seg in predicting the methylation status of the *MGMT* gene promoter in gliomas.

MATERIALS AND METHODS

Creation of database

An outline of the radiomics analysis performed in this study is shown in Fig. 1. The samples used in this study were multiparametric MRI scans of de novo glioblastoma patients from the University of Pennsylvania Health System (UPENN-GBM) from the Cancer Imaging Archive (Frederick National Laboratory for Cancer Research, Frederick, MD, USA), a publicly available database. The database consists of 671 MRI-imaged glioma cases diagnosed according to the previous WHO criteria (4th edition) (Louis et al., 2016), each comprising 192 slices. We selected cases that fulfilled the following conditions: methylation or no methylation as confirmed by genetic diagnosis, a single lesion, and 3D-T1WI imaging after contrast enhancement. In addition, considering that the signal intensity of MRI images varies from case to case, we extracted cases that were imaged using equipment from the same manufacturer. We selected 50 methylated and 50 unmethylated cases from a total of 100 cases (80 cases were used for training and 20 cases for test data). The breakdown of the cases is as follows; 28 males and 22 females (63.8 ± 7.2 years old) were methylated, and 26 males and 24 females (61.2 ± 13.1 years old) were unmethylated. The MRI systems used were TrioTim or Verio, both with a magnetic field strength of 3.0 T, manufactured by Siemens Healthcare (Erlangen, Germany).

Segmentation

ROIs for segmentation were set on the extracted MRI images of the 40 cases by using 3D Slicer (v5.2.1). Segmentation was performed manually by the author (a radiological technologist with 14 years of experience) and confirmed by the co-authors (one student at the department of radiological technology, and four radiological technologists with from 1 to 25 years of experience). 3D_seg was used to segment all slices by setting the ROIs for every slices. 2D_seg was used to set the ROI in the slice with the largest tumour area in the transverse image. Although the images were obtained using equipment from the same manufacturer, the signal intensity differed across the cases because of the inherent characteristics of MRI imaging. Therefore, the maximum signal intensity in each case was normalised to 1.0 to minimise the influence of the differences in signal intensity. In addition, isotropic voxelisation was performed because the radiomics feature measurements were based on the assumption that the pixel size was the same in all directions.

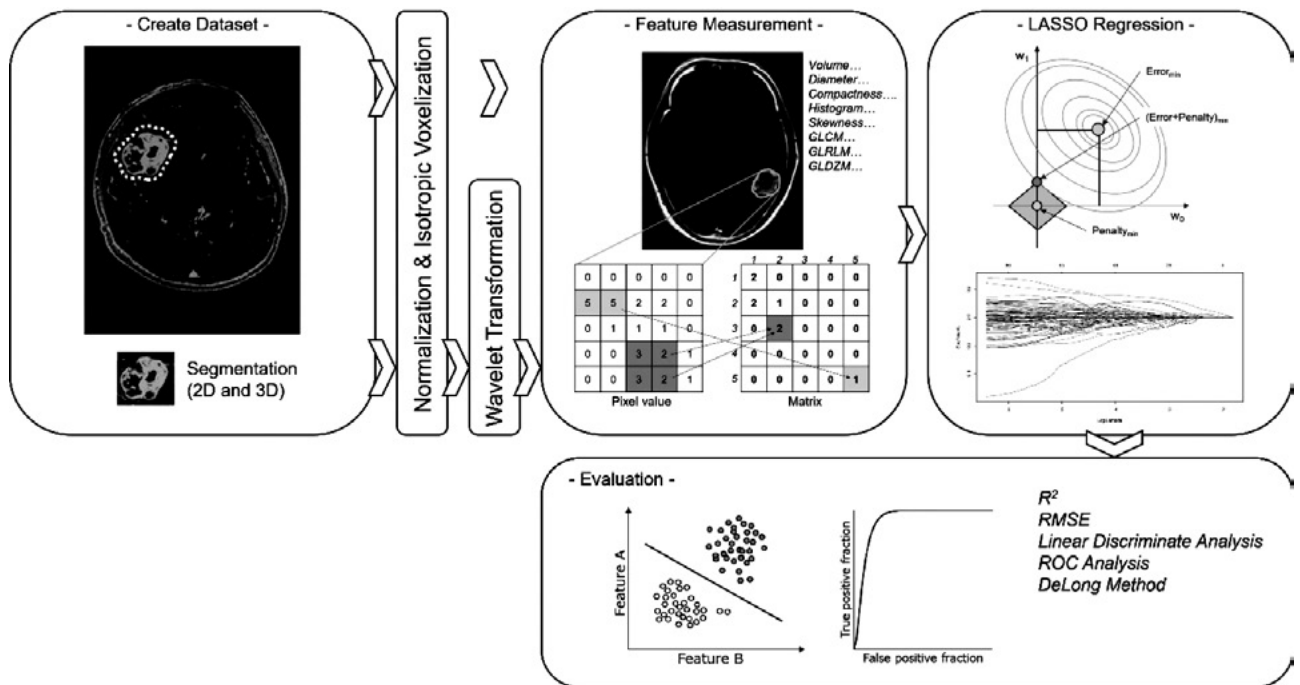


Fig. 1. Outline of the radiomics analysis

Radiomic feature measurement

The radiomic features of the segmented tumour regions were quantified using PyRadiomics, a Python package (van Griethuysen et al., 2017). We measured 14 shape, 18 first-order, and 68 textural features. Additionally, a wavelet transform was added as a preprocessing step before the measurement to increase the number of feature types. In total, 788 features were measured.

LASSO regression

Overlearning occurs when the explanatory variable (radiomic feature) is excessively large compared to the objective variable (whether methylated or unmethylated). Therefore, the obtained features were reduced by means of least absolute shrinkage and selection operator (LASSO) regression using RStudio (v2023.03.1+446) (Hastie et al., 2015). LASSO regression is a regularised linear regression method in which the sum of the weights (L1-norm term) is added to the least-squares cost function, as shown in the following equation:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (1)$$

where $\hat{\beta}$ is the LASSO solution, N is the number of cases, p is the number of features, y_i is the objective variable, x_{ij} is the feature, β_0 is the constant term, β_j is the regression coefficient, and λ is the normalisation parameter. The leave-one-out cross validation (LOOCV) method was used to select the optimal λ

(Shao, 1993). Ten features were extracted in 2D_seg and 3D_seg after determining the optimal λ by using the LOOCV method.

Evaluation and statistical analysis

Multiple regression analysis was performed using the features obtained by 2D_seg and 3D_seg. The coefficient of determination (R^2) and the root mean square error (RMSE) were calculated as evaluation indices. These values were calculated using the following equations:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2} \quad (3)$$

where y_i is the correct value, \hat{y}_i is the predicted value, \bar{y} is the average of the correct values, and n is the number of radiomic features.

Fisher's linear discriminant analysis (LDA) was also performed using the two features that were most highly correlated with the presence or absence of methylation (Fisher, 1936). Sensitivity and specificity were determined by establishing an optimal linear decision boundary that effectively separates the objective variables. In addition, a receiver operating characteristic (ROC) analysis was performed. The area under the ROC curve (AUC) was analysed for statistical significance using the DeLong method (DeLong et al., 1988).

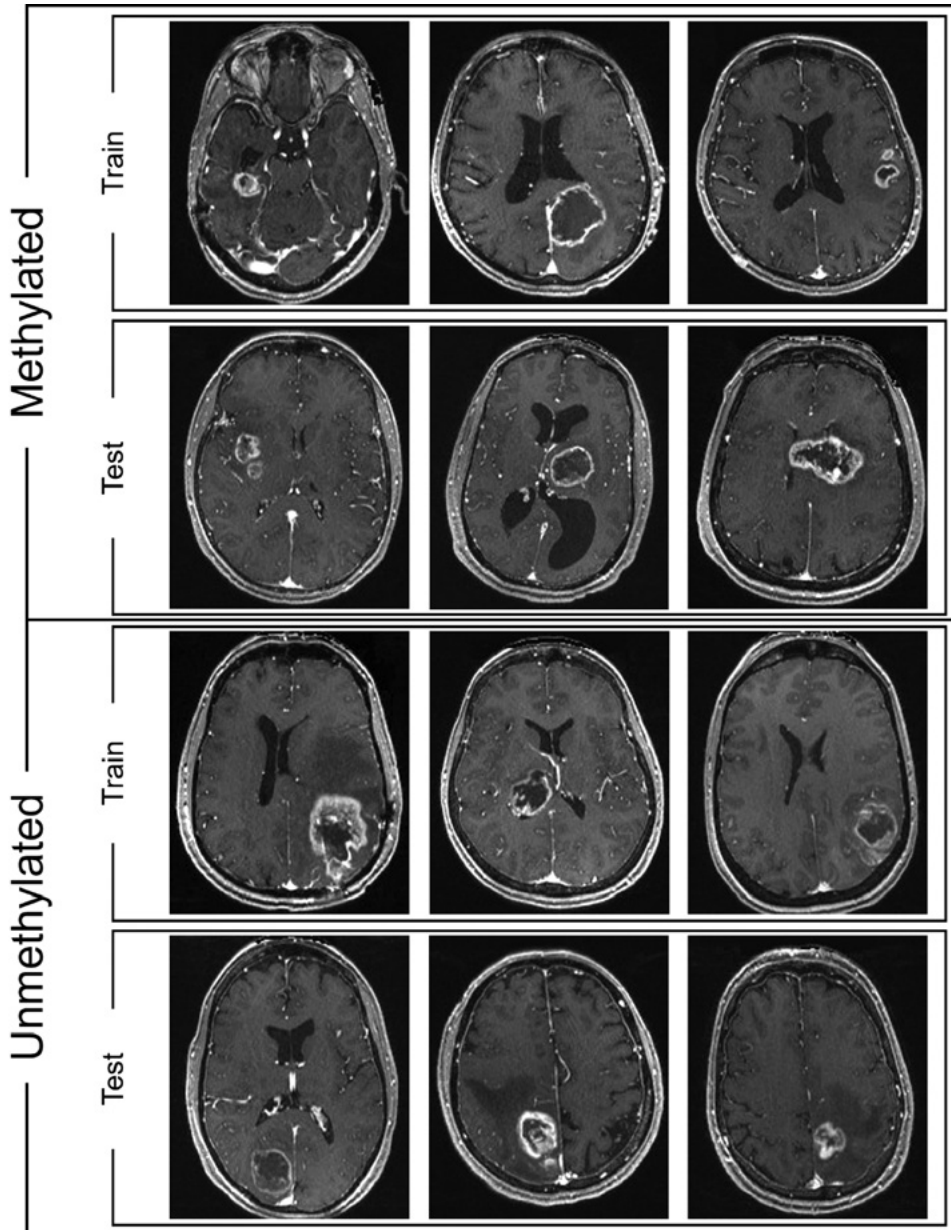


Fig. 2. Glioma cases with or without methylation used for this study

RESULTS

Fig. 2 shows six cases for each sample. It is reported that it is possible to distinguish *MGMT* methylation from MRI imaging findings, however it was difficult task (see Fig. 2) (Drabycz et al., 2010). The radiomics features selected by LASSO regression are shown in Tabs. 1 and 2. The categories in the table indicate the types of features. Overall, many texture-related features were selected. The features with “wavelet” in the feature names are those that were measured after the wavelet transform was applied as a pre-processing step. Conversely, “original” is a feature without wavelet transform. Moreover, as regards the designations “HLH” and “LHH”, “H” denotes a high-pass filter and “L” indicates a low-pass filter. Furthermore, “gldm” and “glrlm”

are texture features, which are calculated by transforming the signal intensity into a matrix and then calculating variance and similarity.

Tab. 3 shows the R^2 value and RMSE obtained by the multiple regression analysis. Neither the R^2 value nor RMSE differed significantly among the segmentation dimensions in the train and test data.

The LDA results used for the train and test data are shown in Fig. 3. 2D_seg had sensitivities of 82.5% (train) and 80% (test), and specificities of 67.5% (train) and 70% (test) for discriminating unmethylated cases. Similarly, the sensitivity and specificity of 3D_seg were 85% (train) and 80% (test), and 62.5% (train) and 60% (test), respectively. The results of the ROC analysis are shown in Fig. 4. The AUC values of 2D_seg and 3D_seg were 0.80 and 0.79, respectively,

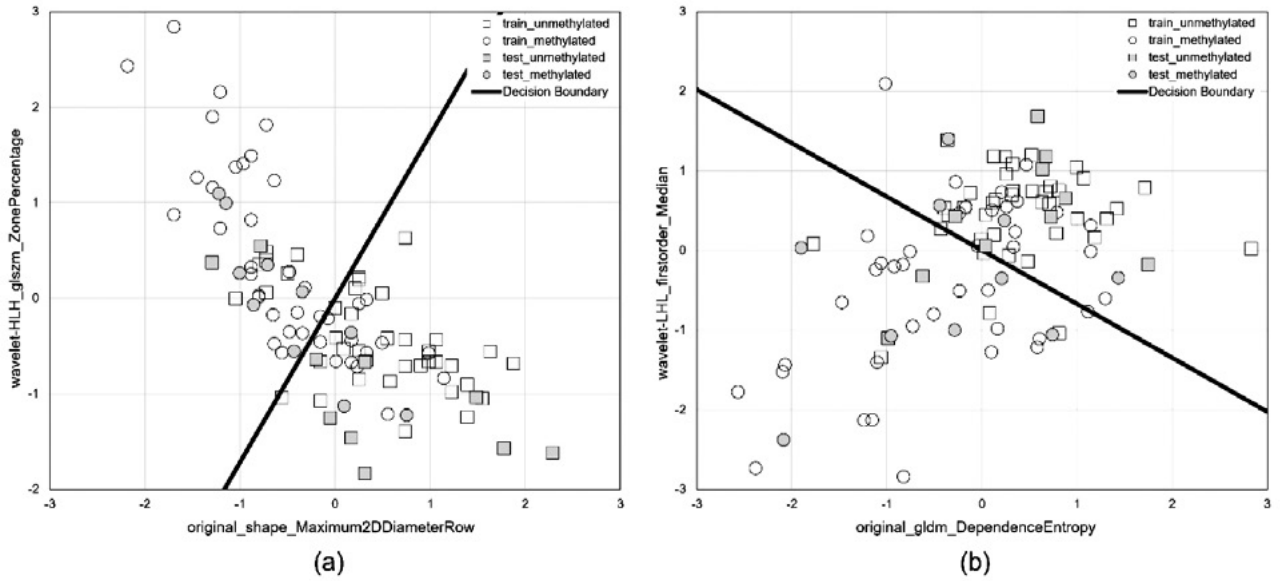


Fig. 3. Results of LDA and decision boundary for (a) 2D_seg and (b) 3D_seg

with 2D_seg being slightly larger than 3D_seg. There were no differences between these models (DeLong method, $p = 0.73$, 95% confidence interval -0.095 to 0.14).

DISCUSSION

The features selected by LASSO regression differed between 2D_seg and 3D_seg; however, they all had the same number of texture features. As the images used were MRI images,

the signal intensities were not absolute values. Certain image properties, such as signal intensity, affect radiomic feature measurements (Scalco et al., 2020). Therefore, the signal intensity was normalised before feature measurement;

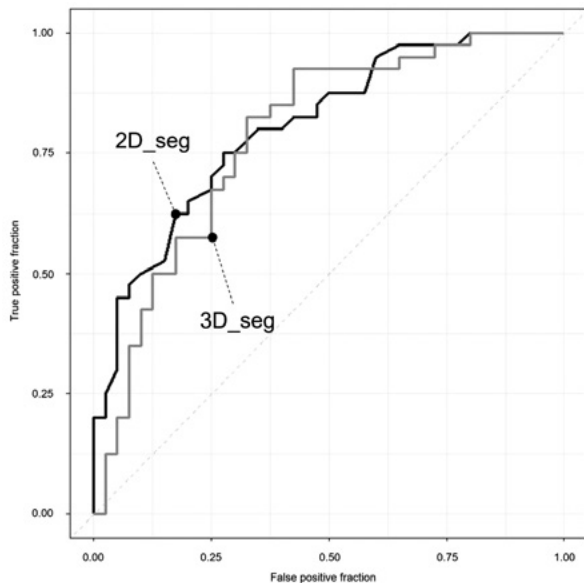


Fig. 4. ROC curve of the two types of segmentation dimension discriminated the methylation. AUCs of 2D_seg and 3D_seg are 0.80 and 0.79 ($p = 0.73$, 95% confidence interval -0.095 to 0.14)

	Radiomics features	Category
1	original_shape_Maximum2DDiameterRow	Shape
2	wavelet-LLH_glrjm_HighGrayLevelRunEmphasis	Texture
3	wavelet-LLH_glrjm_LowGrayLevelRunEmphasis	Texture
4	wavelet-LHH_glcm_Autocorrelation	Texture
5	wavelet-LHH_glszm_SmallAreaLowGrayLevelEmphasis	Texture
6	wavelet-LHH_gldm_LargeDependenceHighGrayLevelEmphasis	Texture
7	wavelet-HLH_firorder_Mean	First-order
8	wavelet-HLH_glszm_ZonePercentage	Texture
9	wavelet-HHL_gldm_SmallDependenceLowGrayLevelEmphasis	Texture
10	wavelet-LLL_glcm_ClusterShade	Texture

Tab. 1. Selected radiomic features with use of 2D_seg

	Radiomics features	Category
1	original_shape_Elongation	Shape
2	original_gldm_DependenceEntropy	Texture
3	original_glrjm_RunLengthNonUniformityNormalized	Texture
4	original_glszm_LargeAreaLowGrayLevelEmphasis	Texture
5	wavelet-LLH_glcm_Correlation	Texture
6	wavelet-LLH_glszm_ZoneVariance	Texture
7	wavelet-LHL_firorder_Median	First-order
8	wavelet-HHL_gldm_DependenceNonUniformityNormalized	Texture
9	wavelet-LLL_firorder_Skewness	First-order
10	wavelet-LLL_glcm_ClusterShade	Texture

Tab. 2. Selected radiomic features with the use of 3D_seg

	Train		Test	
	2D_seg	3D_seg	2D_seg	3D_seg
R^2	0.72	0.73	0.71	0.66
RMSE	0.28	0.33	0.3	0.34

R^2 – coefficients of determination value; RMSE – root mean squared error.

Tab. 3. Results of multi regression analysis

however, the effects of frequency characteristics and other image corrections could not be removed. Therefore, if statistical features such as first-order features are calculated using signal intensity, the discrimination accuracy may be reduced. However, the texture feature, which converts the signal intensity into a matrix and measures dispersion as a scalar quantity, is largely unaffected by the differences in signal intensity among samples. This is likely because the features selected in LASSO regression contained more texture features. There was also a difference in the number of features to which the wavelet transform was applied between 2D_seg and 3D_seg. In 2D_seg, there was only one “original” feature for which the wavelet transform was not applied. In contrast, 3D_seg had four features. The advantage of using wavelet transforms for preprocessing is that they broaden the range of radiomic features and enable more features to be acquired. However, the amount of information in 3D_seg is significantly different from that in 2D_seg because it contains 3D information. Therefore, the original image without wavelet transform was correlated with the objective variable. However, the application of wavelet transforms for preprocessing is performed automatically before measurement in the PyRadiomics procedure. Therefore, the measurement time is not extended, and this is not a reason to actively adopt 3D_seg.

There was no significant difference in the discrimination accuracy between 2D_seg and 3D_seg for any of the indices. Therefore, the segmentation dimensions do not affect the discrimination accuracy. One of the reasons 2D_seg did not degrade in accuracy is that, as mentioned above, it utilises texture information and wavelet transforms to generate a variety of information from a small amount of signal intensity information. We also considered the effect of errors on segmentation performance. The setting of ROIs during segmentation affects the measurement of radiomics features (Eertink et al., 2022). This is because areas other than the lesion of interest or signals that are extremely high (low) in intensity compared to the signal intensity of the lesion are included. In this study, the ROI setting was completed manually, and 3D_seg set the ROIs for each slice, which may have accumulated segmentation errors. These factors may also explain the lack of differences between 2D_seg and 3D_seg. Furthermore, the AUC in the ROC analysis and statistical analysis showed no significant differences. However, there was a difference in the shapes of the ROC curves. 2D_seg is the ROC curve with the TPF rising first. This indicates that the number of false positives was small. 3D_seg is the curve where the FPF first increases, which indicates a high false-positive rate. This finding is consistent with the LDA results. However, when considering the AUC, there was no difference between the segmentation dimensions.

Finally, the limitation of this study is described. Gliomas are classified into subtypes, such as astrocytoma, oligodendroglioma, and paediatric-type diffuse glioma. However, in this study, we did not classify these subtypes, but only by

the presence or absence of *MGMT* methylation. The classification should be validated using only one category in the WHO CNS5 classification.

CONCLUSION

In this study, we investigated whether there was a difference in discrimination accuracy between 2D and 3D segmentation (2D_seg and 3D_seg) in radiomics analysis of the presence or absence of methylation in the promoter region of the *MGMT* gene in gliomas. Various evaluations demonstrated no difference in discrimination accuracy between 2D_seg and 3D_seg. Therefore, radiomics analysis to discriminate the presence or absence of methylation in *MGMT* may be reliably performed using 2D_seg.

Conflict of interest

The authors declare that they have no conflict of interest related to the publication of this article.

Author contribution

Original concept of study: RF. Collection, recording and/or compilation of data: RF, MOn, KH, MOh. Analysis and interpretation of data: RF, MOn, KH, MOh, KK. Writing of manuscript: RF. Critical review of manuscript: RF, MOn, KK, SG. Final approval of manuscript: RF, MOn, KH, MOh, KK, SG.

References

- DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44: 837–845.
- Drabycz S, Roldán G, de Robles P et al.: An analysis of image texture, tumor location, and *MGMT* promoter methylation in glioblastoma using magnetic resonance imaging. *Neuroimage* 2010; 49: 1398–1405.
- Eertink JJ, Pfaehler EAG, Wieggers SE et al.: Quantitative radiomics features in diffuse large B-cell lymphoma: does segmentation method matter? *J Nucl Med* 2022; 63: 389–395.
- Feng Z, Li H, Liu Q et al.: CT Radiomics to predict macrotrabecular-massive subtype and immune status in hepatocellular carcinoma. *Radiology* 2023; 307: e221291.
- Fisher RA: The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 1936; 7: 179–188.
- van Griethuysen JJM, Fedorov A, Parmar C et al.: Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017; 77: e104–e107.
- Hastie T, Tibshirani R, Wainwright M: *Statistical Learning with Sparsity: The Lasso and Generalizations*. 1st ed., CRC Press, New York 2015.
- Hegi ME, Diserens AC, Gorlia T et al.: *MGMT* gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 2005; 352: 997–1003.
- Louis DN, Perry A, Reifenberger G et al.: The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol* 2016; 131: 803–820.
- Louis DN, Perry A, Wesseling P et al.: The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro Oncol* 2021; 23: 1231–1251.
- Scalco E, Belfatto A, Mastropietro A et al.: T2w-MRI signal normalization affects radiomics features reproducibility. *Med Phys* 2020; 47: 1680–1691.
- Shao J: Linear model selection by cross-validation. *J Am Stat Assoc* 1993; 88: 486–494.
- Sollini M, Cozzi L, Ninatti G et al.: PET/CT radiomics in breast cancer: mind the step. *Methods* 2021; 188: 122–132.
- The Committee of Brain Tumor Registry of Japan: Brain tumor registry of Japan (2005–2008) *Neurol Med Chir* 2017; 57 (Suppl 1): 9–102.
- Yang L, Gu D, Wei J et al.: A radiomics nomogram for preoperative prediction of microvascular invasion in hepatocellular carcinoma. *Liver Cancer* 2019; 8: 373–386.