Kocatepe Tıp Dergisi
Kocatepe Medical Journal
26:168-173/Nisan 2025 Sayısı

ARAŞTIRMA YAZISI / RESEARCH ARTICLE

RADYOMİK ÖZELLİK TABANLI MAKİNE ÖĞRENİMİ İLE MENİNGİOMALARIN PREOPERATİF DERECELENDİRİLMESİ: BİR AUTOML ÇALIŞMASI

RADIOMICS FEATURE-BASED MACHINE LEARNING FOR PREOPERATIVE GRADING OF MENINGIOMAS: A STUDY USING AUTOML

Emin DEMİREL, Çiğdem ÖZER GÖKASLAN

Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Radyoloji Ana Bilim Dalı

ÖZET

AMAÇ: En yaygın primer intrakranial neoplazmlardan biri menenjiomlardır. Bu tümörlerin ameliyat öncesi doğru sınıflandırılması, hastaları uygun şekilde yönetmede ve tedaviye karar vermede çok önemlidir. Bu güncel çalışmada, açık kaynaklı yazılım kullanarak grade I ve grade II hastaları tahmin etmek için radyomik özellik temelli makine öğrenme modeli geliştirmeyi amaçladık.

GEREÇ VE YÖNTEM: Meningioma -SEG-CLASS açık kaynaklı veri seti, 2010 ve 2019 yılları arasında cerrahi rezeksiyon geçiren 96 tedavi edilmemiş hastadan toplanmıştır. Segmentasyon verisi açık kaynak olarak paylaşılan tümörlerin radyomik özellikleri çıkartıldı. Otomatik makine öğrenimi algoritmalarımızı geliştirmek için AutoGluon AutoML platformu kullanıldı.

BULGULAR: Gerekli özellik seçimi işlemleri sonrasında geliştirilen AutoGluon AutoML makine öğrenme modellerinde, ansambl L2 modeli göre en iyi performans gösterdi. Bu sonuçlar, test setinde 0,8205 AUC ve 0,8000 F1 skoru ile kabul edilebilir olup, modelin iyi bir genelleme yeteneğine işaret ediyor.

SONUÇ: Bu araştırma, çeşitli MRG dizilerinden radyomik özelliklerin çıkarılması, geleneksel radyolojik testlerden daha iyi bir şekilde meningiomların derecelendirilmesine yardımcı olabilir. Bu, invaziv olmayan preoperatif tümör tahminini kolaylaştırarak daha iyi cerrahi planlama ve yönetimi sağlar.

ANAHTAR KELİMELER: Menenjiom, Makine Öğrenmesi, MRG.

ABSTRACT

OBJECTIVE: One of the most common primary intracranial neoplasms is meningiomas. Correct preoperative classification of these tumors is crucial for appropriate management of patients and treatment decisions. In this current study, we aimed to develop a radiomic feature-based machine learning model to predict grade I and grade II patients using open source software.

MATERIAL AND METHODS: Meningioma-SEG-CLASS open source dataset was collected from 96 untreated patients who underwent surgical resection between 2010 and 2019. Radiomic features of tumors were extracted from segmentation data shared as open source. AutoGluon AutoML platform was used to develop our automated machine learning algorithms.

RESULTS: AutoGluon AutoML machine learning models developed after necessary feature selection processes showed the best performance compared to the ensemble L2 model. These results are acceptable with 0.8205 AUC and 0.8000 F1 score on the test set, indicating good generalization ability of the model.

CONCLUSIONS: This study suggests that extraction of radiomic features from various MR sequences may help grade meningiomas better than traditional radiologic tests. This facilitates noninvasive preoperative tumor prediction, enabling better surgical planning and management.

KEYWORDS: Meningioma, Machine Learning, MRI.

Geliş Tarihi / Received: 07.11.2024
Kabul Tarihi / Accepted: 24.02.2025
Yazışma Adresi / Correspondence: Dr. Öğr. Üyesi Emin DEMİREL
Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Radyoloji Ana Bilim Dalı
E-mail: dremindemirel@gmail.com
Orcid No (Sırasıyla): 0000-0002-0675-3893, 0000-0001-5345-1735
Etik Kurul/Ethical Committee: Afyonkarahisar Sağlık Bilimleri Üniversitesi Etik Kurulu (01.11.2024/9-338).

INTRODUCTION

Meningiomas are one of the most common primary intracranial neoplasms of the CNS. These tumors, classified into three different grades according to the WHO classification, exhibit great variability both in clinical course and treatment strategies (1). Grade I meningiomas are generally benign and have a good prognosis following surgical resection, while grade II and III tumors are more aggressive with higher risks of recurrence and mortality (2). Accurate classification of meningioma pre-operatively is thus of significant importance in developing appropriate treatment plans and effectively managing patients.

Histopathological examination of tissue specimens obtained after surgical resection is the gold standard for the final grading of meningiomas (3). However, preoperative determination of the tumor grade, before an invasive treatment like surgery, would have considerable benefits for the clinician in designing therapy and for the patient in obtaining information about their disease. In this regard, the contribution of non-invasive imaging modalities has been long researched.

Because of its multiplanar imaging ability and soft-tissue contrast, magnetic resonance imaging has become the gold standard in the assessment of intracranial pathologies. Conventional MRI sequences are often used to diagnose and locate meningiomas, but they usually cannot accurately assess tumor grade (4). Recent magnetic resonance imaging (MRI) advances in techniques and the development of image analysis tools provide more comprehensive insights into tumor biology.

Radiomics is an advanced imaging analysis method that extracts high-dimensional quantitative features from medical images and correlates these features with clinical outcomes (5). This strategy offers vital information on tumor heterogeneity and microenvironment by characterizing quantitatively picture features that are not visible to the human eye (6). Radiomics analysis includes many features of tumor shape, intensity, texture, and wavelet features. Various combinations of these characteristics are referred to as "radiomics signatures, " which indicate tumor biology (7).

In recent years, radiomics technique has drawn great interest in predicting the grade of meningioma preoperatively. Park et al. (8), demonstrated that radiomics features extracted from contrast-enhanced T1-weighted MRI images maintained a high accuracy in distinguishing grade I from grade II/III meningiomas. Similarly, Coroller et al. (9) demonstrated that radiomics characteristics obtained from multiparametric MRI were superior to conventional radiologic assessment for meningioma grading. Our work is trying to establish a predictive model for Grade I and Grade II meningioma patients based on segmentation radiomics feature extraction, feature selection, and machine learning using open-source software**.**

MATERIAL AND METHOD

Patient Selection

"Segmentation and Classification of Grade I and II Meningiomas from Magnetic Resonance Imaging: An Open Annotated Dataset (Meningioma-SEG-CLASS) open-source dataset was used (10). The dataset included 96 consecutive untreated patients with intracranial meningiomas who were treated with surgical resection between 2010 and 2019. All patients had available preoperative T1, T1 postcontrast, and T2-FLAIR MR images and underwent partial or complete resection of pathologically confirmed grade I or grade II meningiomas. Meningioma grade was confirmed according to the classification guidelines defined in the 2016 WHO Health Organization criteria. The dataset also included clinical information, grade, subtype, type of surgery, tumor location, and atypical features. Meningioma labels on T1-CE and T2-FLAIR images and segmentation data were retrieved using the TCIA system (11). All patients in the TCIA system were anonymized, and the necessary ethics committee approval was obtained for the reference study. All patients in the dataset were included in the study.

Extraction of Radiomics Features

The Pyradiomics library is a powerful tool that is widely used for the extraction of radiomic features from medical imaging data. This library plays an important role in the analysis of image-based features, particularly in cancer research and diagnosis. In our work, we extracted several features using pyradiomics and prepared these features for further analysis. First, we extracted shape, original, logarithmic (log 3 and log 5) and wavelet features from the image data. Shape features describe the geometric structure and size of tumors, whereas the original features contain the basic intensity information of the image. Logarithmic and wavelet features allow us to analyze the fine details and different frequency components in the image. In particular, wavelet transforms help us better understand tumor heterogeneity by capturing information at different scales in the image (5).

Feature Selection and Machine Learning

Feature extraction was followed by normalization of the data. More specifically, normalization is one crucial step in making the features comparable at different scales. This step is particularly necessary to improve the performance of machine learning models by preventing some features from being given higher weights than others in the process of learning (12). We split our dataset in a ratio of 8:2 for training and test sets, respectively, for model development and evaluation. It's one of the common ways to evaluate a model's generalization ability. The training set is used for learning, and the test set is used for evaluation of the model on an independent dataset.

We used a mutual-information-based approach to feature selection. Mutual information is a metric defining the amount of dependence existing between two variables; in this setting, it allows selecting features that share most information with the target variable. This may improve model performance while keeping its complexity low (13). In the last stage, we developed a number of machine-learning models using the AutoGluon AutoML platform. AutoGluon automates model selection, hyperparameter tuning, and model combination to speed up the process of developing high-performance models easily (14). The models developed using AutoGluon involve combinations of different algorithms. With the ensemble learning techniques applied, we developed a model set that would cover the weakness of individual models and enhance overall performance. Ensemble models generally have higher accuracy and better generalization ability. We found the most successful model and compared its performance metrics, including the area under the curve (AUC), F1 score, sensitivity, and specificity.

Ethical Committee

Ethical committee approval for the study was obtained from the Afyonkarahisar Health Sciences University Faculty of Medicine Non-Intervention Scientific Research Ethics Committee with the decision dated 01.11.2024 and numbered 2024/9–338.

Statistical Analysis

Data were analyzed using SPSS Statistics, version 25.0 (IBM Inc., Armonk, NY, USA). Descriptive statistics were expressed as mean ± standard deviation if the variables were continuous and normally distributed, and as median values if the variables were continuous and not normally distributed. Two level variables were compared with non-normally distributed continuous variables by using the Mann-Whitney U test. The chi-square analysis/Fisher's exact test was used for the analysis of categorical variables' associations. The area under the receiver operating characteristic curve (AUC) was used for the evaluation of the discrimination capability of models. Statistical significance was determined at $p < 0.05$.

RESULTS

We analyzed the demographic data of patients with Grade I and Grade II meningiomas included in our study. In terms of age, the mean age of Grade I meningioma patients was 52.76 ± 12.80 years, while the mean age of Grade II meningioma patients was 59.98 ± 14.66 years (p=0.0112). The proportion of females was 83.3% (n=45), and 14.8% (n=8) of Grade I patients were men. Among Grade II patients, the proportion of women (n=21, 48.8%) and men (n=22, 51.2%) was found (p=0.0005). When histological subtypes were analyzed, the meningothelial subtype (87.0%, n=47) was the most common subtype in patients with Grade I meningioma. In addition, secretory (9.3%, n=5), psammomatous (1.9%, n=1), and fibrous (1.9%, n=1) subtypes were detected. In Grade II meningiomas, atypical subtype was found in 90.7% (n=39), chordoid subtype in 7.0% (n=3), and mixed subtype in 2.3% (n=1). Significant differences were also observed in tumor localization. Grade I meningiomas were most commonly localized in the middle cranial fossa (48.1%, n=26), but also in other regions, such as the posterior cranial fossa (20.4%, n=11), falx and parasagittal regions (14.8%, n=8), anterior cranial fossa (7.4%, n=4), convexity (5.6%, n=3), and lateral ventricle (3.7%, n=2). The convexity was the most common site of localization in Grade II meningiomas (41.9%, n=18). This was followed by the middle cranial fossa (30.2%, n=13), falx and parasagittal regions (20.9%, n=9), anterior cranial fossa (4.7%, n=2), and posterior cranial fossa (2.3%, n=1) (p=0.0001). In the feature selection phase, four features were selected for the final model log-sigma-5-0-mm-3D_gldm_GrayLevelNonUniformity,original_shape_Maximum2DDiameterRow, wavelet-HLL_glrlm_ShortRunLowGrayLevlEmphasis,log-sigma-3-0-mm-3D_gldm_GrayLevelNonUniformity log-sigma-3-0-mm3D_glszm_ HighGray-level zone emphasis were the selected features. The most successful model is the ensembl L2 weighted model. AUC: 0.9923, F1 Score: 0.9296, Sensitivity:0.9429, Specificity: 0.9268. In the test group, the AUC was 0.8205, F1 Score was 0.8000, sensitivity was 1.0000, and specificity was 0.7692. The results for the training and test groups are presented in **Table 1** and **Figure 1**, respectively.

Table 1: Results of auto-gluon machine learning ensembl L2 weighted model

	AUC	Score	Sensitivity	Specificity
Train	0.9923	0.9296	0.9429	0.9268
Test	0.8205	0.8000	1.0000	0.7692

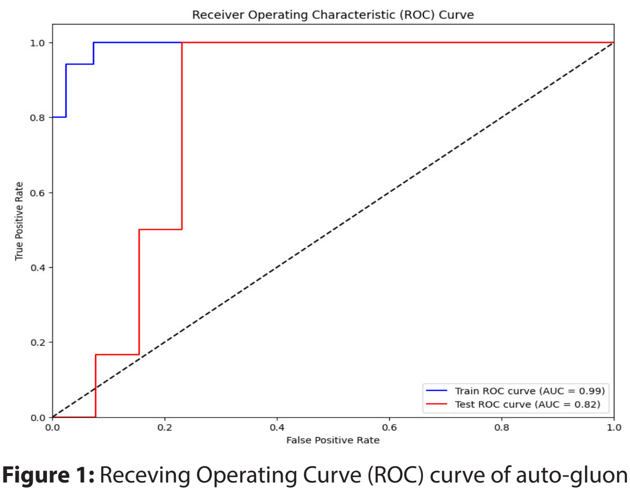

Figure 1: Receving Operating Curve (ROC) curve of auto-gluon machine learning ensembl L2 weighted model

DISCUSSION

Our study used the open-source Meningioma-SEG-CLASS dataset (10) to investigate the potential of radiomic features and automated machine learning in the preoperative differentiation of Grade I and Grade II meningiomas. Our findings show that the selected radiomic features, the AutoGluon platform (14), and the ansambl L2 weighted model have shown promising results in the classification of these tumors. This approach has the potential to optimize surgical planning and patient management by allowing preoperative prediction of tumor grade in a noninvasive manner.

One of the most important findings of our study was the high performance of our model on the training set (AUC: 0.9923, F1 Score: 0.9296). This is a combination of the selected radiomics features and AutoGluon's automatic hyperparameter optimization and model selection process (14) is effective in distinguishing between Grade I and Grade II meningiomas. In particular, our mutual information-based feature selection method (13) may have contributed to this high performance by selecting the most informative features while reducing model complexity. The performance on the test set (AUC: 0.8205, F1 Score: 0.8000) indicated that the generalization ability of the model was acceptable. In future studies, the generalization ability of the model to real-world data can be more accurately assessed using larger and more diverse datasets and different validation coefficients.

The selected radiomic features (log-sigma-5-0-mm-3D_gldm_GrayLevelNonUniformity, original_shape_Maximum2DDiameterRow, wavelet-HLL glrlm_ShortRun LowGrayLevel Emphasis,log-sigma-3-0-mm-3D_glszm_HighGrayLevelZoneEmphasis) reflect the textural and morphological differences between Grade I and Grade II meningiomas. The GrayLevelNonUniformity obtained from the gray-level distribution matrix (GLDM) measures tumor heterogeneity, with higher values indicating a more heterogeneous tumor (7). Since Grade II meningiomas usually have a more heterogeneous structure than Grade I (3), it is expected that this feature would be selected. Similarly, Maximum2DDiameterRow reflects tumor size and Grade II meningiomas can often reach larger sizes (15). Features derived from wavelet transforms help capture subtle textural differences by analyzing different frequency components of the image (16). Logarithmic transformations, however, allow features to be closer to a normal distribution. It can increase the performance of machine learning algorithms by providing (17).

In the literature, there is growing interest in the use of radiomic features in the preoperative grading of meningiomas (8, 9). These studies support our findings. These studies have shown that radiomic features extracted from different MRI sequences may perform better than conventional radiologic assessments for meningioma grading. However, most of these studies used different imaging protocols, feature extraction methods, and machine learning algorithms. Therefore, it is difficult to directly compare the results of the different studies. Moreover, conducting larger multicenter studies would increase the usability of the radiomic approach in clinical practice. This study utilized open-source software and automated machine-learning methods. This approach increases the reproducibility of the study and allows other researchers to test similar models using their own datasets. Furthermore, advanced AutoML platforms such as AutoGluon have automated model selection and optimization processes, enabling a more efficient and comprehensive analysis (14).

Our study had some limitations. First, the dataset was relatively small. This may limit the generalization ability of the model. Second, we analyzed only grade I and II meningiomas. Studies including Grade III meningiomas are important to evaluate the validity of the radiomic approach for all grades. Third, only MRI was used in our study. Integrating other imaging modalities (e.g., CT and PET) and clinical data (e.g., age, sex, and symptoms) may further improve the performance of the model. Finally, radiomic features may be sensitive to imaging parameters (e.g. as magnetic field strength and slice thickness).

In conclusion, our study demonstrated that radiomic features and automated machine learning are potential tools for the preoperative discrimination of Grade I and Grade II meningiomas. Future studies should address the aforementioned limitations and further improve the performance of the model to increase its usability in clinical practice. In this way, the preoperative grading of meningiomas can be performed more accurately, surgical planning can be optimized, and more appropriate treatment strategies can be developed for patients.

REFERENCES

1. 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(6):803-20.

2. Goldbrunner R, Minniti G, Preusser M, et al. EANO guidelines for the diagnosis and treatment of meningiomas. Lancet Oncol. 2016;17(9):e383-91.

3. Perry A, Stafford SL, Scheithauer BW, Suman VJ, Lohse CM. Meningioma grading: an analysis of histologic parameters. Am J Surg Pathol. 1997;21(12):1455-65.

4. Watts J, Box G, Galvin A, et al. Magnetic resonance imaging of meningiomas: a pictorial review. Insights Imaging. 2014;5(1):113-22.

5. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441-6.

6. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.

7. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016;278(2):563-77.

8. Park YW, Oh J, You SC, et al. Radiomics and machine learning may accurately predict the grade and histological subtype in meningiomas using conventional and diffusion tensor imaging. Eur Radiol. 2019;29(8):4068-76.

9. Coroller TP, Bi WL, Huynh E, et al. Radiographic prediction of meningioma grade by semantic and radiomic features. PLoS One. 2017;12(11):e0187908.

10. Vassantachart A, Cao Y, Shen Z, et al, W. Segmentation and Classification of Grade I and II Meningiomas from Magnetic Resonance Imaging: An Open Annotated Dataset (Meningioma-SEG-CLASS) (Version 1) The Cancer Imaging Archive. 2024;51(3):2334-44.

11. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging. 2013;26(6):1045-57.

12. Demircioğlu A. The effect of feature normalization methods in radiomics. Insights Imaging. 2024;15(1):2.

13. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence. 2005;27(8):1226-38.

14. Erickson N, Mueller J, Shirkov A, et al. Autogluon-tabular: Robust and accurate automl for structured data. arXiv preprint arXiv:200306505. access date: 13.03.2020.

15. Oya S, Kim SH, Sade B, Lee JH. The natural history of intracranial meningiomas. J Neurosurg. 2011;114(5):1250-6.

16. Unser M. Texture classification and segmentation using wavelet frames. IEEE Transactions on Image Processing. 1995;4(11):1549-60.

17. Feng C, Wang H, Lu N, et al. Log-transformation and its implications for data analysis. Shanghai Arch Psychiatry. 2014;26(2):105-9.