



OPEN

A weakly supervised deep learning-based method for glioma subtype classification using WSI and mpMRIs

Wei-Wen Hsu¹, Jing-Ming Guo¹, Linmin Pei^{2✉}, Ling-An Chiang¹, Yao-Feng Li³, Jui-Chien Hsiao¹, Rivka Colen^{4,5} & Peizhong Liu⁶

Accurate glioma subtype classification is critical for the treatment management of patients with brain tumors. Developing an automatically computer-aided algorithm for glioma subtype classification is challenging due to many factors. One of the difficulties is the label constraint. Specifically, each case is simply labeled the glioma subtype without precise annotations of lesion regions information. In this paper, we propose a novel hybrid fully convolutional neural network (CNN)-based method for glioma subtype classification using both whole slide imaging (WSI) and multiparametric magnetic resonance imagings (mpMRIs). It is comprised of two methods: a WSI-based method and a mpMRIs-based method. For the WSI-based method, we categorize the glioma subtype using a 2D CNN on WSIs. To overcome the label constraint issue, we extract the truly representative patches for the glioma subtype classification in a weakly supervised fashion. For the mpMRIs-based method, we develop a 3D CNN-based method by analyzing the mpMRIs. The mpMRIs-based method consists of brain tumor segmentation and classification. Finally, to enhance the robustness of the predictions, we fuse the WSI-based and mpMRIs-based results guided by a confidence index. The experimental results on the validation dataset in the competition of CPM-RadPath 2020 show the comprehensive judgments from both two modalities can achieve better performance than the ones by solely using WSI or mpMRIs. Furthermore, our result using the proposed method ranks the third place in the CPM-RadPath 2020 in the testing phase. The proposed method demonstrates a competitive performance, which is creditable to the success of weakly supervised approach and the strategy of label agreement from multi-modality data.

Brain tumors, originating in the glial cells, are cancerous masses in the central nervous system (CNS)¹. In 2011–2015, there are 23 out of 100,000 population diagnosed with brain tumors in the US². Prior to 2016, World Health Organization (WHO) categorizes the CNS gliomas into grades I–IV based on the histological features of a heterogeneous population of the tumor³. However, a new brain tumor classification criterion was released from the WHO in 2016. According to the new criterion, the tumor classification is determined based on both phenotypic and genotypic information⁴. There are many glioma subtypes: diffuse astrocytoma, isocitrate dehydrogenase (*IDH*)-mutant/-wildtype, anaplastic astrocytoma *IDH*-mutant/-wildtype, oligodendroglioma, *IDH*-mutant and *1p/19q-codeleted*, glioblastoma, *IDH*-mutant/-wildtype, etc. The prognosis of a patient with brain tumors is highly related to the tumor grade². In general, patients with higher-grade gliomas have a less survival period. Especially, for patients with glioblastoma (GBM), the median survival period still remains 12–16 months, even with treatment advancement⁵. Consequently, an accurate and robust glioma subtype prediction provides a valuable guide for diagnosis and treatment management. Conventionally, brain tumor diagnosis or grading is performed by pathologists, who examine tissue sections fixed on glass slides under a light microscope. Yet, the

¹Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC. ²Imaging and Visualization Group, ABCS, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA. ³Department of Pathology, Tri-Service General Hospital and National Defense Medical Center, Taipei 11490, Taiwan, ROC. ⁴Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15232, USA. ⁵Hillman Cancer Center, University of Pittsburgh Medical Center, Pittsburgh, PA 15260, USA. ⁶College of Engineering, Huaqiao University, Quanzhou, China. ✉email: linmin.pei@nih.gov

manual diagnosis/grading process is time-consuming and susceptible to human errors. Therefore, computer-aided brain tumor subtype classification is highly desired.

Prior to 2016, digital pathology images are the primary sources for the glioma subtype classification. Digital pathology is the digitized process of whole slide images (WSI) transforming into high-resolution images⁶. Digital pathology has become increasingly common because of the rich context information on the WSI. There are extensive studies on tumor subtype classification in the literature: Kothari et al. utilize a multi-class model for histological classification⁷. Chang et al. use the spatial pyramid matching framework (SPM) with a linear Support Vector Machine (SVM) classifier to classify glioblastoma multiforme (GBM)⁸. A hybrid machine learning method using SVM, random forest (RF), and neural network (NN) is proposed for glioma grading based on the WSI⁹. Barker et al. exploit an elastic net for brain tumor type classification¹⁰. However, a common limitation of these conventional machine learning methods is feature extraction, which requires professional clinical background and computer vision knowledge. In recent years, deep learning (DL)-based approaches have shown superior performance, and have been widely applied in many domains, e.g., computer vision¹¹, medical image analysis^{12,13}, and natural language processing (NLP)¹⁴. The deep learning-based methods is also adopted for glioma classification based on WSI¹⁵ and for glioma grading^{16,17}.

On the other hand, MRI is an alternative source for glioma grading because of the noninvasive property. The MRI-based approaches also provide promising results for glioma classification and grading. Zacharaki et al. apply a SVM-based method to classify tumor type on MRI¹⁸. In the paper, they first extract radiological features, e.g., tumor shape and intensity characteristics. They then apply feature selection using a SVM with recursive feature elimination. Finally, they perform the tumor classification using another SVM. In addition, a hybrid method using a SVM and k-nearest neighbour classifiers (named as SVM-KNN) is also utilized for brain cancer classification¹⁹. In Refs.^{20,21}, Random Forest (RF)-based methods are used for tumor classification as well. Recently, CNN-based methods have been becoming prevalent for such tasks^{22,23}. Sajjad et al. propose a deep learning-based method for multi-grade brain tumor classification²⁴. Liu et al. present a multi-task CNN algorithm for joint segmentation and genotype prediction of brainstem gliomas²⁵. Pei et al. utilize a 3D CNN-based method for brain tumor subtype classification, and achieve the state-of-the-art performance²⁶.

Unsurprisingly, a combination of pathology and radiology images provides more comprehensive context information than using a single modality alone. Ma et al. propose CNN-based methods for tumor classification on WSI and MRI^{27–30}. Kurc et al. investigate brain tumor classification using machine learning and deep learning on WSI and MRI³¹. The work of using both WSI and MRI in Refs.^{27–30} offers state-of-the-art performance. However, all the methods are fusion-based at the feature level. We argue that these methods undermine the priority of pathology in tumor classification, which conflicts with the criterion defined by the WHO. We should develop a computer-aided diagnosis system and should take pathological modality as the primary source for tumor classification.

Therefore, we propose a novel hybrid fully convolutional neural network (CNN)-based method for glioma subtype classification using both whole slide image (WSI) and multiparametric magnetic resonance image (mpMRIs). The proposed method primarily focuses on the WSI-based result while taking the mpMRIs-based result as the complementary reference to enhance the robustness.

Methodology

In this section, the proposed two approaches (the WSI-based approach and the mpMRIs-based approach) are elaborated. In addition, the label agreement strategy for fusing the prediction is also covered.

Overall pipeline. Figure 1 shows the overall pipeline of the proposed method. It consists of a WSI-based approach at the top and a mpMRIs-based approach at the bottom. Each approach outputs a probability of each subtype for each case. The final prediction is primarily derived from the WSI-based result. However, the final prediction is corrected as the mpMRIs-based result when the confidence index of WSI-based result is less than a threshold value. The threshold value is obtained in the validation phase of the Challenge, which achieves the best performance by using the proposed method.

WSI-based approach. Even the training data with paired image and classification label is available in the CPM-RadPath 2020 challenge³², the classification task is still challenging because of the small number of cases and the label constrain issue. The label of the WSI is given, however, the precise lesion region information is missing. The issue becomes worse when considering the massive size of WSI. The inexact labeling in a weakly supervised learning task results in inaccurate samples in the training process. Extracting representative patches is of importance for the task. To overcome the label constrain issue, we extract multiple patches according to the intensity distribution, and assign the corresponding label, as shown in Fig. 2. These patch candidates are randomly selected from the areas with tissues on each WSI. In addition, to further screen out those samples without dense cell distribution, we apply two following criteria to the selecting patches: (1) The mean of all pixel intensities for each sampling patch should lie in between 50 and 150, and the standard deviation of pixel intensities in each channel of R, G, and B should be greater than 20. (2) The difference of maxima and minima of pixel intensity mean should be smaller than 100. The first condition ensures the selecting patches having rich context, rather than blank samples. The second condition screens out those patches that contain color markers. By following the rules, we extract 300 patches for each WSI in both the training and inference phases. Since the ratio of noisy samples is unknown and unpredictable, the prototype selection method can reduce the impact of noisy samples in the training process. Figure 3 shows the pipeline of the training phase in the WSI-based approach. First, all sampling patches for each category are collected to train using a ResNet50³³. Subsequently, the trained CNN model is used to extract the convolutional representations (deep features) for each sampling patch. To

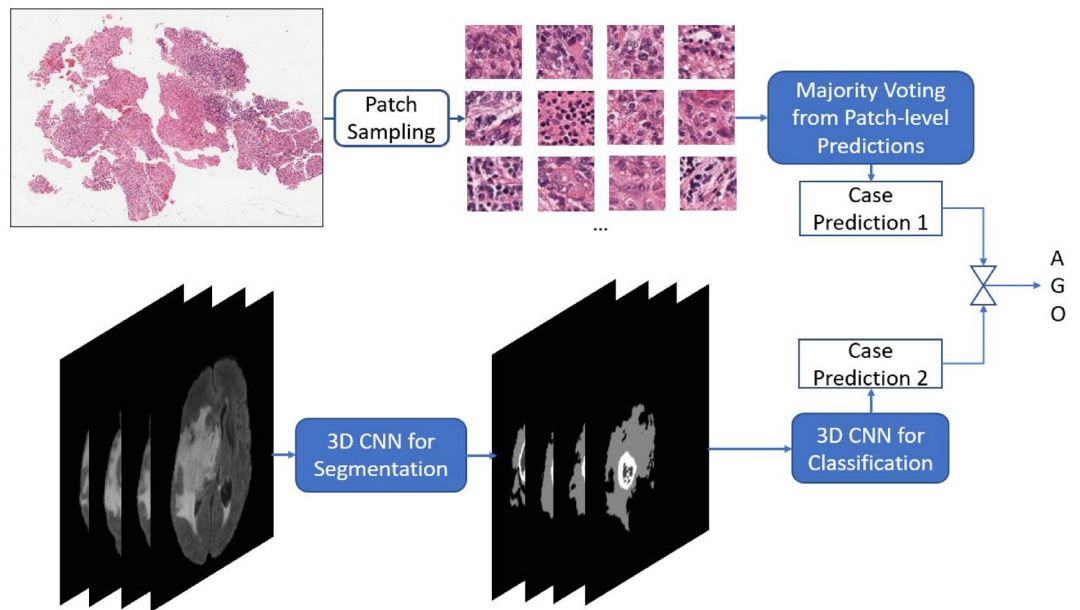


Figure 1. Overall pipeline of the proposed method in inference phase.

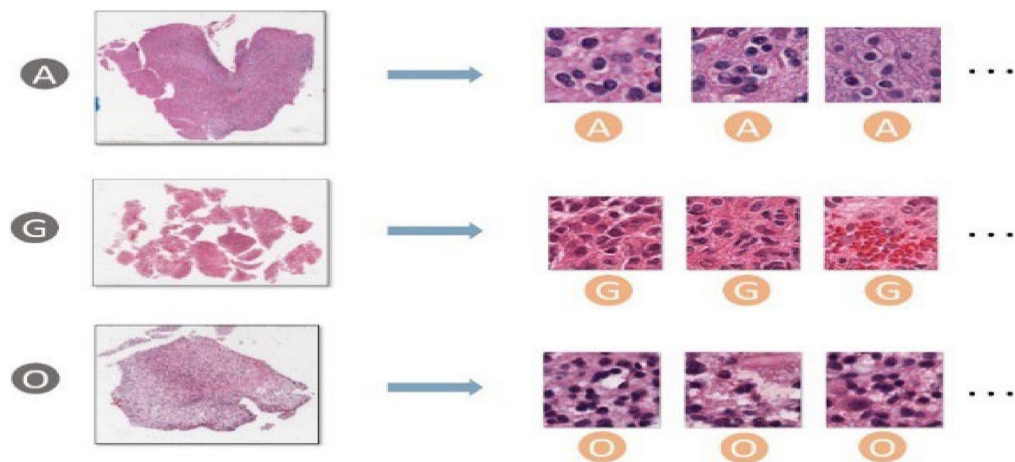


Figure 2. Patch sampling: Patches are randomly sampled from the densely cell-distributed areas of a WSI. All sampled patches have the same category as the label of the corresponding WSI.

reduce the impact of non-representative samples in each class, several prototypes for each category are selected by pair-wise similarity check. The extracted prototypes (e.g. 100 prototypes for each category in our experiment) are selected according to the similarity of intensity distribution. However, non-representative prototypes may also be collected because of the existence of irrelevant tissues such as lymphocytes, red blood cells, and mostly stroma, etc. In addition, astrocytoma and oligodendroglioma have similar morphological features and are commonly confounded diagnoses with large intraobserver variabilities^{34,35}. Last, an expert intervention step is to screen out the non-representative prototypes further or re-assign to another category.

The purpose of the prototype approach is to measure the morphological similarities among all patches in each category for clustering and then select the representative patches from the major clusters. In this case, noisy samples or outliers will be excluded. The WSI-based method’s detail is as the following: first, all sampling patches in the same category are fed into the CNN model to derive deep features. Subsequently, the similarity matrix S is computed by pair-wise comparison of patches using cosine similarity, and each entity s_{ij} in the matrix can be derived by Eq. (1)

$$s_{ij} = \frac{G(x_i)^T G(x_j)}{\|G(x_i)\|_2 \|G(x_j)\|_2}, \tag{1}$$

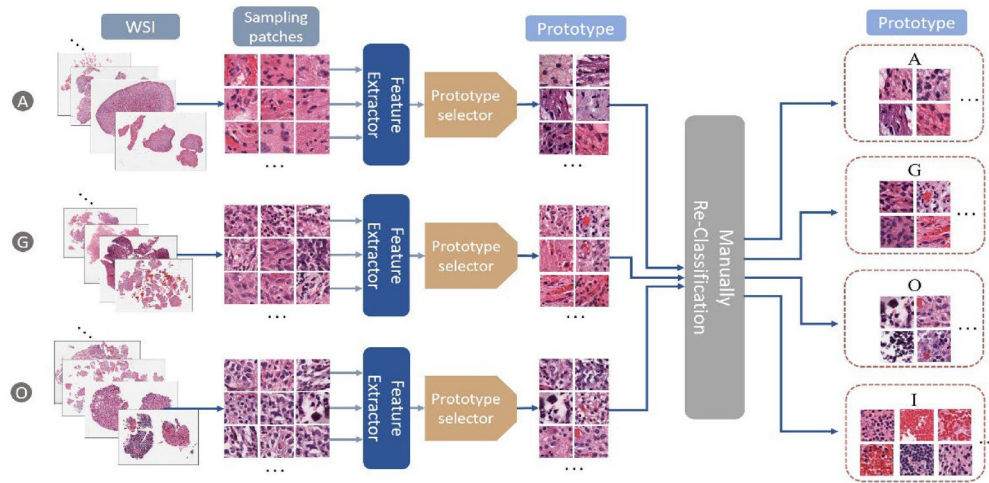


Figure 3. The pipeline of training phase in the WSI-based approach.

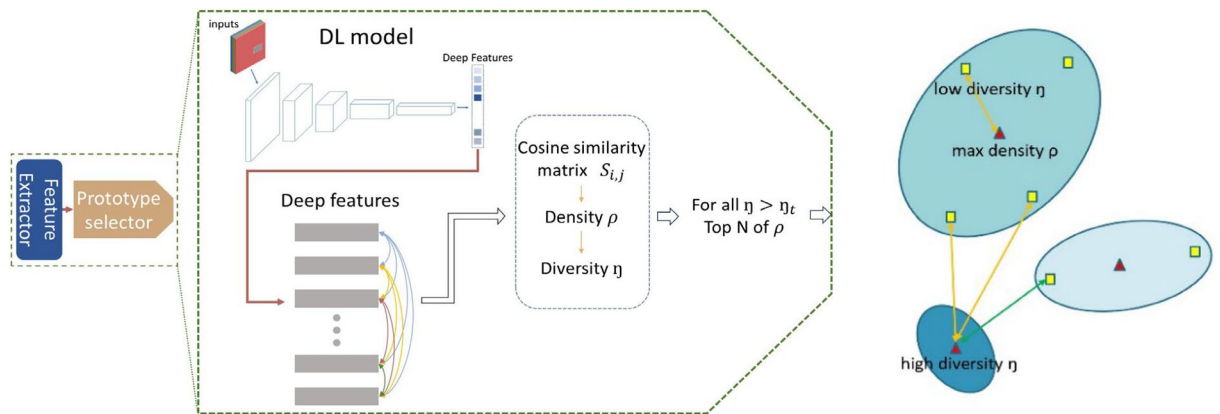


Figure 4. Prototype selection in each category. Left: The pipeline of prototype selection. Right: Measurements of density and diversity.

where $G(x_i)$ or $G(x_j)$ is the deep feature vector derived from the forward propagation of convolutional layers of the i th or j th sampling patch (x_i or x_j) in the CNN model. The entity s_{ij} reflects the similarity between two arbitrary sampling patches. A similarity threshold, s_t , is set to be the average value of all pair-wise similarity values in a category. The measurement of density for the i th sampling patch, ρ_i , is computed by Eq. (2).

It counts how many patches over the total m patches in a category.

$$\rho_i = \sum_{j=1}^m \text{sign}(s_{ij} - s_t). \tag{2}$$

For prototype selection in each category, the largest similarity value indicates the most representative index of the category. An ideal prototype set is able to distinguish all categories. To meet the requirements, another index of η_t is to measure the diversity of prototypes. The diversity index for each patch in a category is designed by Eq. (3)³⁶.

$$\eta_i = \begin{cases} 1 - \min_j s_{ij}, & \rho_i = \rho_{max} \\ 1 - \max_{j, \rho_j > \rho_i} s_{ij}, & \rho_i < \rho_{max} \end{cases} \tag{3}$$

Figure 4 shows the scheme of prototype selection for each category with the rules as follows: All sampling patches for a category are fed into the CNN model for feature extraction. Subsequently, the similarity matrix S is derived by performing pair-wise similarity comparison among patches in the category. Afterwards, the measurements of density (ρ_i) and diversity (η_i) of each patch are computed to determine the selection priority and selection condition. All patches in the category are ranked based on their corresponding density value ρ , which reflects the selection priority. In addition, the diversity threshold η_t is set to avoid the selected prototypes being too similar, resulting in redundancy during selection.

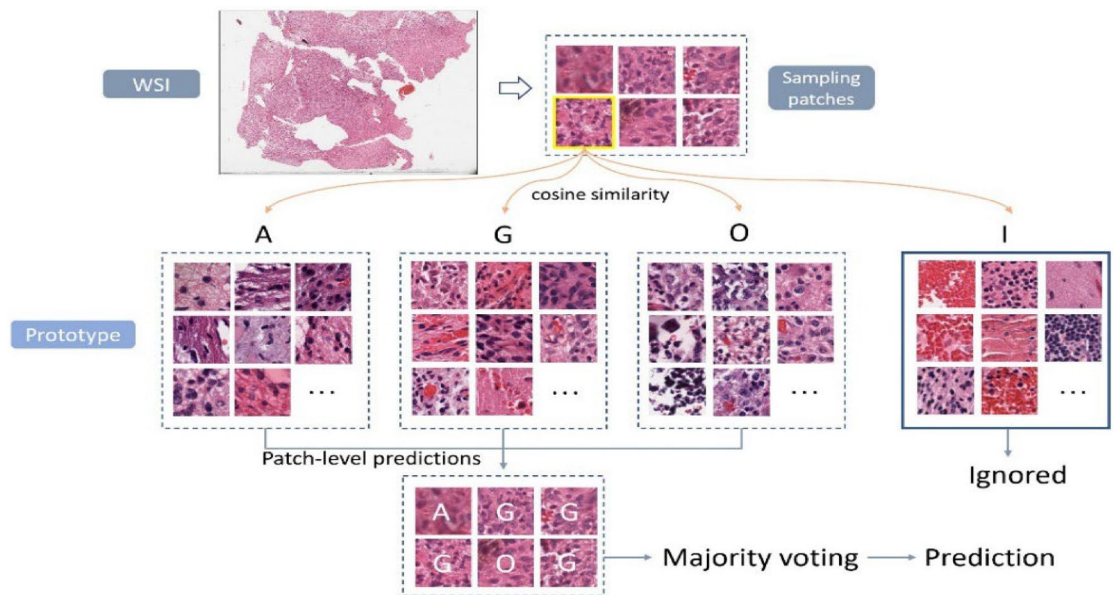


Figure 5. Example of case-level inference in the proposed WSI-approach.

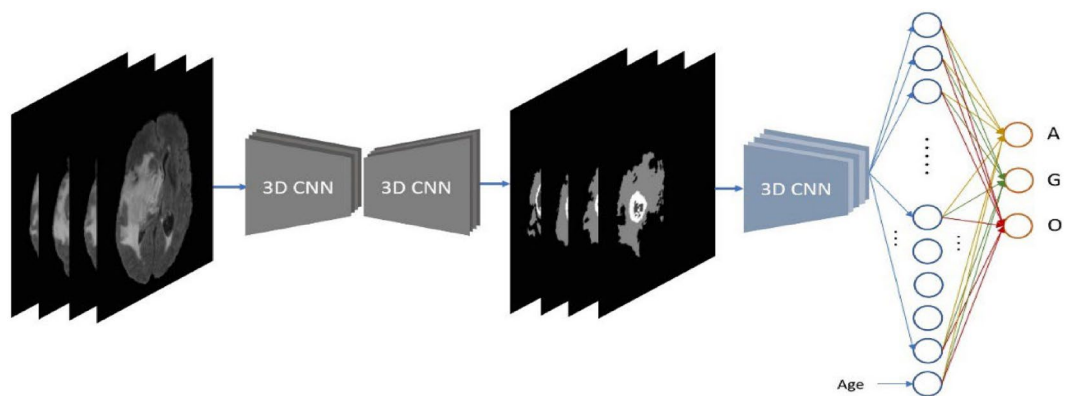


Figure 6. Pipeline of the proposed MRI-approach.

In Eq. (3), the sampling patch with the highest density value, i.e., ρ_{max} , among all sampling patches in a category will be assigned the top selecting priority, and it will also be assigned a large diversity index to surpasses the diversity threshold η . For other sampling patches, the diversity index η measures the feature distance to the most similar patch with a higher density value ρ .

Accordingly, if the i th sampling patch is the prototype candidate, we check the η_i to ensure it is greater than the diversity threshold to retain high diversity among the selected prototypes within a category.

In the inference phase, several patches are extracted from each WSI, and then fed into the CNN model for feature extraction. For each sampling patch, the cosine similarity is computed between the extracted feature and the deep features of each prototype in each category. Each patch is classified into the category with the highest average of cosine similarities among prototypes. Finally, a majority voting from all predictions of samples is performed to determine the final label prediction for the case. Figure 5 illustrates an example of case-level inference using the proposed WSI-based approach. Notably, we employ an additional category (I) for irrelevant classification of gliomas subtype. Sampling patches of the category (I) will be ignored in the voting process.

MRI-based approach. For the mpMRIs-based approach, we employ a cascade deep learning-based method. Brain tumors are firstly segmented using a 3D CNN model^{37,38}, and then the segmentations are fed into another 3D CNN model for glioma subtype classification. The pipeline of the proposed MRI-based approach is shown in Fig. 6. Accurate segmentation of brain lesions leads to an outstanding performance on brain tumor classification. Since the intensity of MRI varies across all cases, intensity normalization is desirable to reduce the bias. In our experiments, a z-score normalization is applied for all MRIs. In doing so, all voxel values are subtracted by the mean and divided by the standard deviation of the brain region. In addition, several data augmentations are applied in the training phase of both segmentation and classification, such as rotation, random

flipping, and affine translation. According to a report of 2008–2011, age is associated with brain tumor subtype². Therefore, the age of the patient is also taken into consideration for glioma subtype classification.

Label agreement. For the WSI-based approach, the prediction is derived by choosing the glioma subtype with the highest votes. In addition, the distributions of votes among all subtypes, regardless of the category (I), turn into probabilities indicating the confidence score for each glioma subtype. While for the MRI-based approach, the prediction probability for each class can be retrieved directly from the CNN classification network. According to the guideline of CNS glioma classification defined by WHO, histopathology information is the primary image source for the glioma subtype classification. Thus, we primarily consider the WSI-based result as the final prediction, but also take the mpMRIs-based result as a complementary reference when the confidence index of the WSI-based result is less than the pre-defined threshold, C_c . In such a strategy, the final consensus of glioma subtype classification is decided by the fusion of WSI and mpMRIs information. We believe that integrating WSI and mpMRIs offers a more robust and reliable result than using a single image type.

Datasets and experiments

Ethics approval and guidelines. In this study, the MRI images and pathology images sourced from two public dataset: multimodal Brain Tumor Segmentation Challenge (BraTS) 2020^{39–45} and Computational Precision Medicine: Radiology-Pathology Challenge on brain tumor classification (CPM-RadPath) 2020³¹. Approval was granted on the grounds of existing datasets. Informed consent was obtained from all of the patients in this study. All methods were carried out in accordance with relevant guidelines and regulations. Ethical approval for the use of these data was obtained from the ethics committee of University of Pittsburgh.

Dataset. The BraTS 2020 training data contains 369 cases including 76 low-glioma grade (LGG) patients and 293 high-glioma grade (HGG) patients. For each case, there are multiparametric MRIs (mpMRIs) and the corresponding ground truths of brain tumors. The mpMRIs include T1-weighted MRI (T1), T1-weighted MRI with contrast enhancement (T1ce), T2-weighted MRI (T2), and T2-weighted MRI with fluid-attenuated inversion recovery (T2-FLAIR). Each modality of each case has a size of $240 \times 240 \times 155$. The ground truth of a tumor segmentation contains multiple tumor subtypes, such as tumor tissues of peritumoral edema (ED), enhancing tumor (ET), and necrosis/non-enhancing tumor (NCR/NET).

For the data from CPM-RadPath 2020, the training dataset is comprised of 221 cases with paired radiology and digital pathology images. Within the 221 cases, there are 54, 34, and 133 cases for lower grade astrocytoma, *IDH-mutant* (A), oligodendroblastoma, *1p/19q codletion* (O), and glioblastoma and diffuse astrocytic glioma with molecular features of glioblastoma, *IDH-wildtype* (G), respectively. In addition, there are 35 and 73 cases for the validation and testing sets in the CPM-RadPath 2020 challenge, respectively. It notices that the challenge organizer privately owns the ground truth of the glioma subtype of the validation and testing data. In the validation phase, participants submit the prediction results to the challenge for online evaluating the algorithm. However, participants are only allowed to submit the algorithm wrapped in a Docker container in the testing phase, and the organizer executes the algorithm for final ranking.

Evaluation metrics. Three metrics are utilized for performance evaluation in the challenge of CPM-RadPath 2020 for glioma subtype classification, which are micro-F1, Cohen's Kappa Coefficient, and balanced accuracy. In the tasks of multi-class classification, the micro-F1 is equivalent to the overall accuracy, as Eq. (4).

$$\text{Micro-F1} = \frac{\text{total TruePositives}}{\text{total \# cases}} \text{ (for multi-class)}. \quad (4)$$

The Cohen's Kappa Coefficient (κ) is to measure inter-rater and intra-rater reliability for categorical items, and the definition is as Eq. (5).

$$\kappa = \frac{p_o - p_e}{1 - p_e}, \quad (5)$$

where p_o is the relative observed agreement among raters, and p_e is the hypothetical probability of chance. Lastly, the balanced accuracy computes the average of the proportion corrects of each class individually, which is formulated in Eq. (6).

$$\text{Balanced accuracy} = \frac{\sum_{i=1}^{i=n} \text{Recall}_i}{n}, \quad (6)$$

where n indicates the number of classes in the task.

Glioma subtype classification. For the WSI-based approach, 300 patches are sampled from each case of WSI and are fed into the proposed pipeline for patch-level classification in our experiments. Subsequently, the glioma subtype for each case can be determined by choosing the subtype with the highest votes, and the vote distributions are normalized into the probabilities as the confidence scores for each glioma subtype.

On the other hand, for the MRI-based approach, a 3D CNN model of ResUNet^{37,38} is trained on the dataset from the challenge of BraTS 2020 to perform tumor subregion segmentation. The task targets three tumor subregions, including peritumoral edema (ED), enhancing tumor (ET), and necrosis or non-enhancing tumor (NCR/NET). The segmentation results of three targeted tumor subregions are shown in Fig. 7. Subsequently,

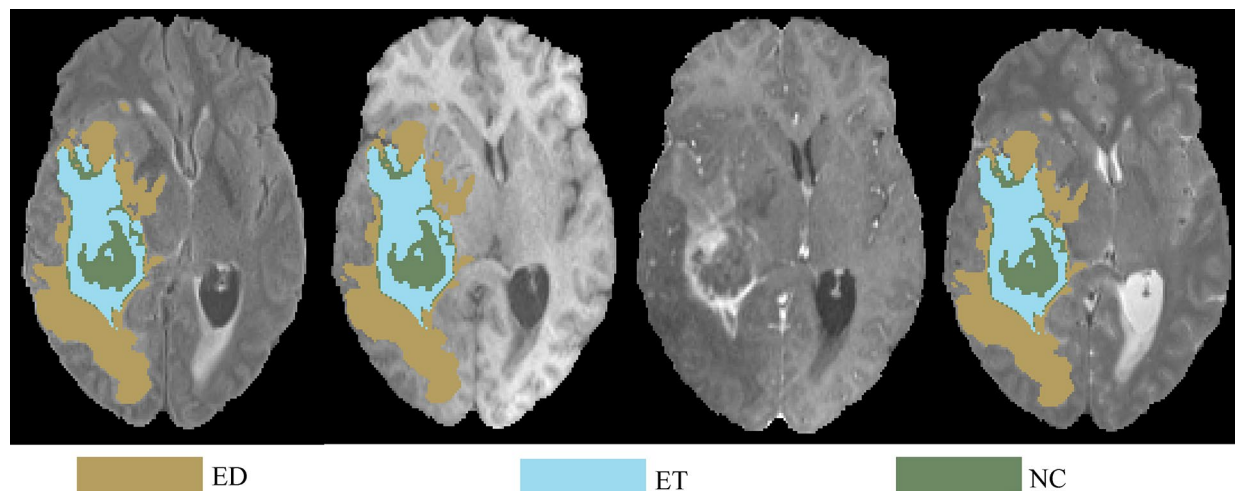


Figure 7. An example of segmentation results of three subregions of tumor, which are peritumoral edema (ED) in green, enhancing tumor (ET) in red, and necrosis or non-enhancing tumor (NCR/NET) in blue.

Phase	Method	Case	F1-score	Kappa	Balance_acc	Average
Validation	MRI-based	35	0.771	0.627	0.698	0.699
	WSI-based	35	0.886	0.798	0.777	0.821
	Fusion	35	0.886	0.801	0.8	0.829
Test	Fusion	73	0.712	0.505	0.654	NA

Table 1. Performance of two approaches and the scheme of label agreement.

Phase	Method	F1-score	Kappa	Balance_acc	Average	Ranking
Valid	Ma et al. ²⁷	0.943	0.903	0.833	NA	NA
	Pei et al. ²⁶	0.829	0.715	0.749	0.764	NA
	Chan et al. ²⁹	0.72	NA	NA	NA	NA
	Xue et al. ²⁸	0.849	NA	NA	NA	NA
	Our method	0.886	0.801	0.8	0.829	NA
Test	Pei et al. ²⁶	0.603	0.39	0.596	NA	2nd (2019)
	Our method	0.712	0.505	0.654	0.654	3rd (2020)

Table 2. Performance comparison among state-of-the-art approaches in CPM-RadPath 2019/2020. All references are coming from the top teams in the CPM-RadPath 2019/2020.

the segmentation model is applied on all the MRI cases in the training dataset from CPM-RadPath 2020, and the segmentation results are fed into another 3D CNN model of ResNet^{31,33} to train the relations between tumor morphology and glioma subtypes.

It notes that the CPM-RadPath 2020 has identical datasets as CPM-RadPath 2019. Since the glioma subtype of ground truth for each case in the testing set from the challenge of CPM-RadPath 2020 is not publicly available, in our experiments, the performance on the validation set from the challenge is focused for evaluation and comparison.

Result

The performance of the two approaches and the scheme of label agreement is listed in Table 1. The confidence threshold in the scheme of label agreement, C_p , is empirically set to be 0.6 in our experiment. The comparison shows a promising result that a label fusion-based predictions using the proposed method outperform the ones from single approach.

In addition, we also compare our result to the other top-ranked teams in the validation phase, as shown in Table 2. The results show that our proposed method achieves a competitive performance among all on the validation set. Moreover, we participate the CPM-RadPath testing phase. It notices that all participants are required to submit the algorithm wrapped with Docker in the testing phase of CPM-RadPath 2020. The challenge organer

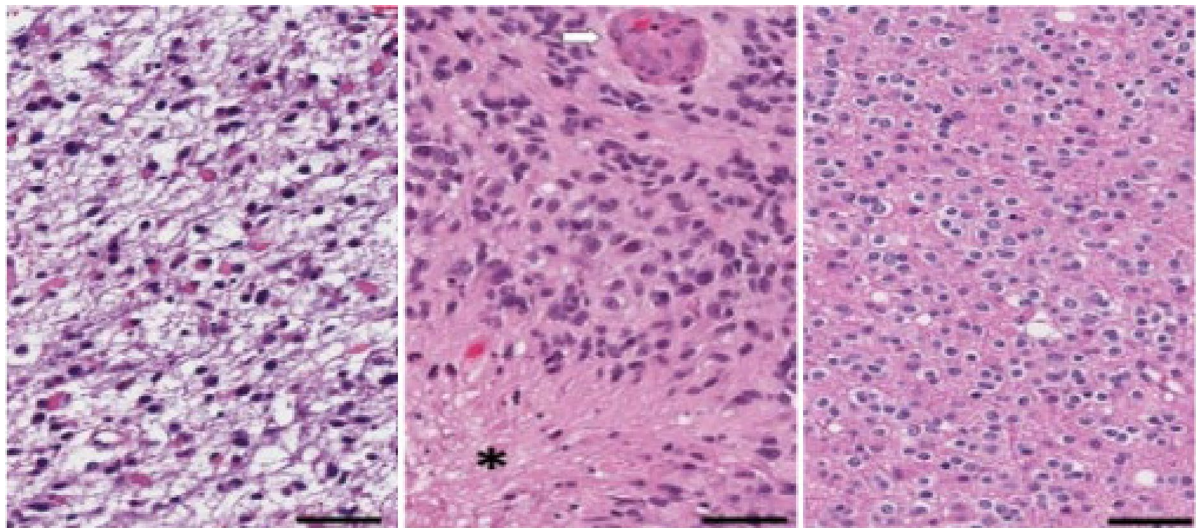


Figure 8. Pathological morphologies corresponding to each subtype of glioma. The scale bar at the bottom-right corner indicates the actual size of 50 μm at the magnification. Left: astrocytoma (A). Middle: glioblastoma (G). Right: oligodendroglioma (O).

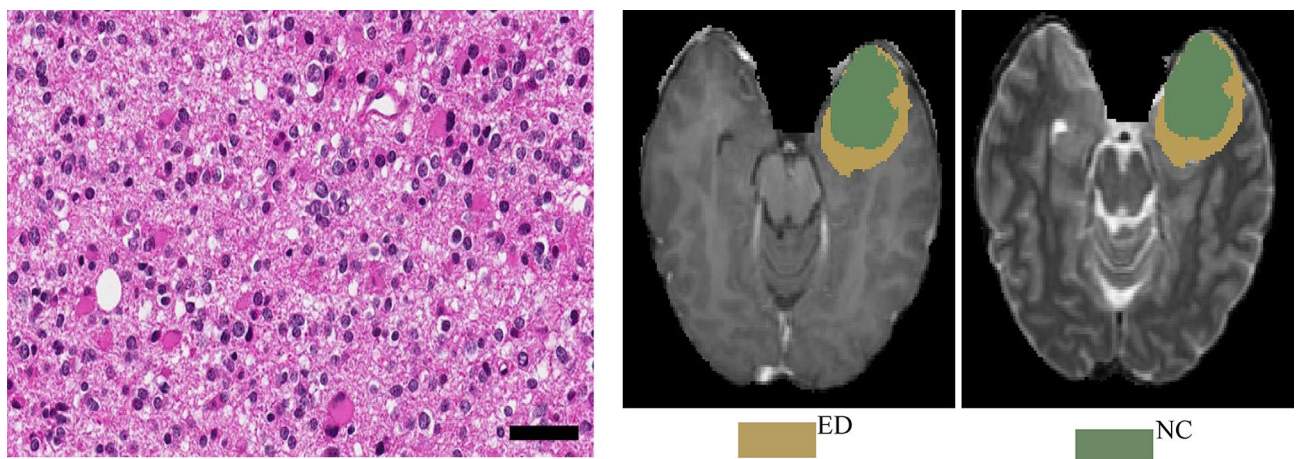


Figure 9. One of the cases that was overdiagnosed by the WSI-based approach but successfully corrected by the MRI-based approach after the proposed label agreement scheme. Left: Pathological observation. Right: The brain tumor segmentation on MRI.

executes the algorithm, and ranks the performance. Our result ranks the third place in the testing phase. It indicates that the proposed method offers a competitive performance on glioma subtype classification.

Discussion

From the pathological perspective, the morphologies of glioma subtypes are distinct. As it can be observed in Fig. 8-Left, astrocytoma revealed hypercellularity with irregular and hyperchromatic nuclei in the fibrillary background. The pink and abundant cytoplasmic with eccentric nuclei (gemistocytic differentiation) occasionally showed in this group. While glioblastoma, as shown in Fig. 8-Middle, is derived from astrocytoma; hence, many morphologic findings are shared. However, few discriminatively histopathologic features can distinguish them, such as glomerular endothelial proliferation (pointed by a white arrow) and tumor necrosis (the area marked with a star-shaped sign) only show in glioblastoma. Last, the oligodendroglioma appears round shape nuclei with open chromatin and artifactual cytoplasmic retraction, leading to the “fried egg” appearance, as shown in Fig. 8-Right.

Figure 9 shows an example that the prediction from the MRI-based approach successfully corrected the final diagnostic result using the proposed label agreement scheme. In this case, the prediction from the WSI-based approach is overdiagnosed astrocytoma as glioblastoma, which might occur because both of them have shared many morphologic features. The weight of decisive features, including necrosis or vascular proliferation, should be recruited more for model training to increase the chance of separation. Fortunately, this misinterpretation is corrected by the mpMRIs-based approach. MRI is a powerful source to detect necrosis and vascular abnormality by evaluating the amount of enhancement, degree of heterogeneity, and liquid components. As it can be observed in Fig. 9, there is no enhancement or necrosis identified in MRI. It can be a clue for the computer-aided system

Method	Astrocytoma	GBM	Oligodendroglioma	Assigned label
WSI-based	0.3844	0.5069	0.1088	G
mpMRI-based	0.9413	0.0145	0.0441	A
Final agreement	–	–	–	A

Table 3. Voting distribution and predicting probabilities by using WSI-based and mpMRI-based methods, respectively. The maximum probability is in bold.

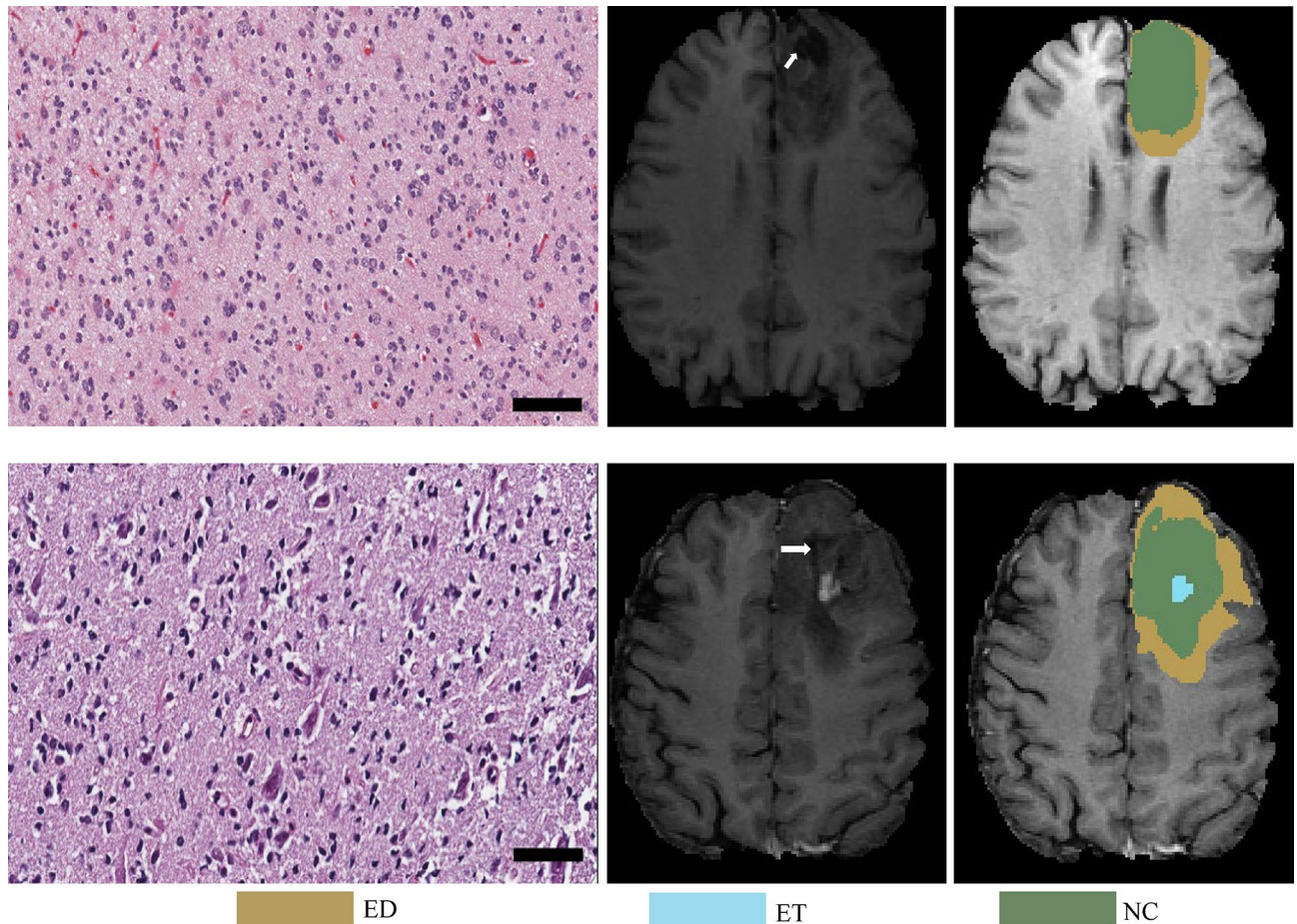


Figure 10. Two cases of misinterpretation from Oligodendroglioma to Astrocytoma. From left to right: WSI, corresponding T1ce and its segmentation. (Top): Case of CPM20_TCIA10_239_1, (Bottom): Case of CPM20_TCIA1_387_1. The arrow on T1ce points to cavity.

to classify the case to the category of astrocytoma, instead of glioblastoma. The voting distribution of the case by the WSI-based approach and the probabilities of subtype predictions by the mpMRI-based approach are shown in Table 3. The assigned labels of WSI-based and mpMRI-based method are “G” and “A”, respectively. However, since the confidence index (0.5069) by the WSI-based method is less than the threshold value (0.6), the final label agreement assigns “A” for the correction.

However, misclassification also occurs. In the experiment, there are another two cases misinterpreted, as shown in Fig. 10. For the first case, in Fig. 10-(Top), it is mis-classified oligodendroglioma as an astrocytoma by the WSI-based approach. There are two explanations for the misinterpretation: First, this case belongs to a higher grade oligodendroglioma (WHO grade III) and shows a more severe degree of nuclear atypia, mimicking astrocytoma. Second, the slide is mainly located in the infiltrating part mixing with tumor and adjacent brain pathologic images. Unfortunately, the mpMRIs-based approach does not correct the misclassification by following the strategy of label agreement. The prediction of the WSI-based approach is Astrocytoma with confidence of 0.93, while the prediction of the mpMRI-based approach is GBM. However, the ground truth label of this case is Oligodendroglioma. Figure 10-(Bottom) shows another case of misinterpretation. It is misdetermined oligodendroglioma to astrocytoma by the WSI-based approach due to poor fixation and staining procedure. In addition, the tumor cells have revealed marked pyknosis and dark nuclei without nuclear details, leading to

misinterpretation. Though, some features of oligodendroglioma can be identified, such as artifactual cytoplasmic retraction. It may be a dilemma for the model to make decisions with such two contradictory features observed. The prediction of the mpMRI-based approach is same as that of the WSI-based approach, but both are misclassified. Unfortunately, the ground truths of brain tumor segmentation are not available for public. We also notice that both cases contain a cavity after post-surgery, which may result in the misclassification.

Though the proposed method produces a competitive result, there are some limitations. First, we use both WSI and mpMRIs for the classification. To achieve the best result, we assume that an accurate brain tumor subregion results in a good classification. In doing so, we use an extra dataset from BraTS 2020, which limits its application. Second, the proposed method is sensitive to some empirical parameters, such as the threshold of confidence index, the number of extracting patches, etc. Third, the relatively small size of experimental data is a drawback, which widely exists in deep learning-based methods. Forth, the proposed method requires a qualified professional or expert intervention to screen out the non-representative prototypes.

Conclusion

In this study, we propose a novel hybrid fully convolutional neural network (CNN)-based method for glioma subtype classification using both whole slide image (WSI) and multiparametric magnetic resonance images (mpMRIs). It is comprised of two methods: a WSI-based method and a mpMRIs-based method. For the WSI-based method, we categorize the glioma subtype using a 2D CNN on WSIs. For the mpMRI-based method, we also develop a 3D CNN-based method by analyzing the mpMRI. The mpMRIs-based method consists of brain tumor segmentation and classification. We classify the glioma subtype primarily on WSI-based results with the guidance of the mpMRIs-based prediction when the confidence index of the WSI-based result is less than the pre-defined threshold. The experimental results show that the final label fusion-based predictions achieve a superior result and offer a competitive performance.

Data availability

In this study, Our data are coming from both <https://www.med.upenn.edu/cbica/brats2020/data.html> and <https://www.med.upenn.edu/cbica/cpm2020.html>. The data can be found: <https://www.med.upenn.edu/cbica/brats2020/data.html> (BraTS 2020), and <https://miccai.westus2.cloudapp.azure.com> (CPM-RadPath 2020).

Received: 12 January 2022; Accepted: 30 March 2022

Published online: 12 April 2022

References

1. Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127**(12), 2893–2917 (2010).
2. Ostrom, Q. T. *et al.* CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2011–2015. *Neuro Oncol.* **20**, 1–86 (2018).
3. Vignesswaran, K., Neill, S. & Hadjipanayis, C. G. Beyond the World Health Organization grading of infiltrating gliomas: Advances in the molecular genetics of glioma classification. *Ann. Transl. Med.* **3**, 7 (2015).
4. Louis, D. N. *et al.* The 2016 World Health Organization classification of tumors of the central nervous system: A summary. *Acta Neuropathol.* **131**(6), 803–820 (2016).
5. Chen, J., McKay, R. M. & Parada, L. F. Malignant glioma: Lessons from genomics, mouse models, and stem cells. *Cell* **149**(1), 36–47 (2012).
6. Janowczyk, A. & Madabhushi, A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J. Pathol. Inform.* **7**, 29 (2016).
7. Kothari, S., Phan, J. H., Young, A. N. & Wang, M. D. Histological image classification using biologically interpretable shape-based features. *BMC Med. Imaging* **13**(1), 9 (2013).
8. Chang, H., Zhou, Y., Spellman, P. & Parvin, B. Stacked predictive sparse coding for classification of distinct regions in tumor histopathology. In *Proc. IEEE International Conference on Computer Vision*, 169–176 (2013).
9. Wang, X. *et al.* Machine learning models for multiparametric glioma grading with quantitative result interpretations. *Front. Neurosci.* **12**, 1046 (2019).
10. Barker, J., Hoogi, A., Depeursinge, A. & Rubin, D. L. Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles. *Med. Image Anal.* **30**, 60–71. <https://doi.org/10.1016/j.media.2015.12.002> (2016).
11. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436. <https://doi.org/10.1038/nature14539> (2015).
12. Pei, L., Vidyaratne, L., Rahman, M. M. & Iftekharuddin, K. M. Deep learning with context encoding for semantic brain tumor segmentation and patient survival prediction. In *Medical Imaging 2020: Computer-Aided Diagnosis*, Vol. 11314, 113140H (International Society for Optics and Photonics, 2020).
13. Pei, L., Vidyaratne, L., Rahman, M. M. & Iftekharuddin, K. M. Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images. *Sci. Rep.* **10**(1), 1–11 (2020).
14. GaneGEDara, T. *Natural Language Processing with TensorFlow: Teach Language to Machines Using Python's Deep Learning Library* (Packt Publishing Ltd, 2018).
15. Hou, L. *et al.* Patch-based convolutional neural network for whole slide tissue image classification. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2424–2433 (2016).
16. Ertosun, M. G. & Rubin, D. L. Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks. In *AMIA Annual Symposium Proceedings*, Vol. 2015, 1899 (American Medical Informatics Association, 2015).
17. Pei, L., Jones, K. A., Shboul, Z. A., Chen, J. Y. & Iftekharuddin, K. M. Deep neural network analysis of pathology images with integrated molecular data for enhanced glioma classification and grading. *Front. Oncol.* **11**, 2572 (2021).
18. Zacharaki, E. I. *et al.* Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn. Reson. Med.* **62**(6), 1609–1618 (2009).
19. Machhale, K., Nandpuru, H. B., Kapur, V. & Kosta, L. MRI brain cancer classification using hybrid classifier (SVM-KNN). In *2015 International Conference on Industrial Instrumentation and Control (ICIC)*, 60–65 (IEEE, 2015).
20. Usman, K. & Rajpoot, K. Brain tumor classification from multi-modality MRI using wavelets and machine learning. *Pattern Anal. Appl.* **20**(3), 871–881 (2017).

21. Reza, S. M., Samad, M. D., Shboul, Z. A., Jones, K. A. & Iftekharuddin, K. M. Glioma grading using structural magnetic resonance imaging and molecular data. *J. Med. Imaging* **6**(2), 024501 (2019).
22. AlBadawy, E. A., Saha, A. & Mazurowski, M. A. Deep learning for segmentation of brain tumors: Impact of cross-institutional training and testing. *Med. Phys.* **45**(3), 1150–1158 (2018).
23. Zhuge, Y. *et al.* Automated glioma grading on conventional MRI images using deep convolutional neural networks. *Med. Phys.* **47**(7), 3044–3053 (2020).
24. Sajjad, M. *et al.* Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *J. Comput. Sci.* **30**, 174–182 (2019).
25. Liu, J. *et al.* A cascaded deep convolutional neural network for joint segmentation and genotype prediction of brainstem gliomas. *IEEE Trans. Biomed. Eng.* **65**(9), 1943–1952 (2018).
26. Pei, L., Vidyaratne, L., Hsu, W.-W., Rahman, M. M. & Iftekharuddin, K. M. Brain tumor classification using 3D convolutional neural network. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (eds Crimi, A. *et al.*) 335–342 (Springer, 2020).
27. Ma, X. & Jia, F. Brain tumor classification with multimodal MR and pathology images. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (eds Crimi, A. *et al.*) 343–352 (Springer, 2020).
28. Xue, Y. *et al.* Brain tumor classification with tumor segmentations and a dual path residual convolutional neural network from MRI and pathology images. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (eds Crimi, A. *et al.*) 360–367 (Springer, 2020).
29. Chan, H.-W., Weng, Y.-T. & Huang, T.-Y. Automatic classification of brain tumor types with the MRI scans and histopathology images. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (eds Crimi, A. *et al.*) 353–359 (Springer, 2020).
30. Pei, L. *et al.* A hybrid convolutional neural network based-method for brain tumor classification using mMRI and WSI. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (eds Crimi, A. *et al.*) 487–496 (Springer, 2020).
31. Kurc, T. *et al.* Segmentation and classification in digital pathology for glioma research: Challenges and deep learning approaches. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2020.00027> (2020).
32. Farahani, K. *et al.* *Computational Precision Medicine Radiology-Pathology challenge on Brain Tumor Classification 2020.* <https://doi.org/10.5281/zenodo.3718894> (2020).
33. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 770–778 (2016).
34. Coons, S. W., Johnson, P. C., Scheithauer, B. W., Yates, A. J. & Pearl, D. K. Improving diagnostic accuracy and interobserver concordance in the classification and grading of primary gliomas. *Cancer Interdiscipl. Int. J. Am. Cancer Soc.* **79**(7), 1381–1393 (1997).
35. Cooper, L. A. *et al.* An integrative approach for in silico glioma research. *IEEE Trans. Biomed. Eng.* **57**(10), 2617–2621 (2010).
36. Han, J., Luo, P. & Wang, X. Deep self-learning from noisy labels. In *Proc. IEEE/CVF International Conference on Computer Vision*, 5138–5147 (2019).
37. Zhao, Y.-X., Zhang, Y.-M. & Liu, C.-L. Bag of tricks for 3D MRI brain tumor segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (eds Crimi, A. *et al.*) 210–220 (Springer, 2020).
38. Yu, L., Yang, X., Chen, H., Qin, J. & Heng, P. A. Volumetric ConvNets with mixed residual connections for automated prostate segmentation from 3D MR images. In *Proc. AAAI Conference on Artificial Intelligence*, Vol. 31 (2017).
39. Menze, B. H. *et al.* The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* **34**(10), 1993–2024 (2014).
40. Bakas, S. *et al.* Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. Preprint at <http://arXiv.org/811.02629> (2018).
41. Bakas, S. *et al.* Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci. Data* **4**, 170117 (2017).
42. Bakas, S. *et al.* Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection. In *The Cancer Imaging Archive* (2017).
43. Bakas, S. *et al.* Segmentation labels and radiomic features for the pre-operative scans of the TCGA-GBM collection. In *The Cancer Imaging Archive* (2017).
44. Cordier, N., Delingette, H. & Ayache, N. A patch-based approach for the segmentation of pathologies: Application to glioma labelling. *IEEE Trans. Med. Imaging* **35**(4), 1066–1076 (2015).
45. Pereira, S., Pinto, A., Alves, V. & Silva, C. A. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans. Med. Imaging* **35**(5), 1240–1251 (2016).

Acknowledgements

The experiments in this study work on the datasets from the challenges of <https://www.med.upenn.edu/cbica/brats2020/data.html> and <https://www.med.upenn.edu/cbica/cpm2020.html> held by the BrainLes workshop of the MICCAI.

Author contributions

W.H. and L.P. conducted the experiment, and wrote the main manuscript text. J.G., L.C., J.H., Y.L., R.C., and P.L. reviewed and revised the manuscript.

Funding

Open Access funding provided by the National Institutes of Health (NIH).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022