Triplanar-CNN for Automated Grading of Gliomas Using Preoperative Multi-Modal MR Images

[Abdela Ahmed Mossa, Ulus Çevik]

Abstract— Glioma has been one of the most common life-threatening brain tumor diseases all over the world with different levels of aggressiveness: Low Grade Glioma (LGG) and High Grade Glioma (HGG), and consequently automated glioma grade prediction methods based on multi-modal MRI images are of great interest. However, the development of effective automated methods, and in particular convolutional neural networks (CNN) for fast and accurate medical image analysis has relied on the availability of large annotated training datasets. The purpose of this study was to develop a 2D CNN model, Triplanar-CNN, to a fully automated and accurate glioma grade prediction, using a small training dataset of less than 300 glioma patients who underwent pre-operative volumetric MRI exams, which included FLAIR, T1Gd, T1, and T2 modalities. Our approach operates on all of the MRI modalities and plane slices (axial, coronal, and sagittal) based on reconstructing the volumetric MRI as a set of 2D stacked slices in the sagittal, coronal and axial planes, and allows to leverage pre-trained CNN models for feature extraction, which is essential given the inadequate amount of labeled training dataset. The proposed Triplanar-CNN architecture consists of three sub-networks, each based on leveraging CNN model pre-trained on natural images, and separately applied to axial, coronal and sagittal view of a 3D MRI, respectively, followed by a common fusion-layer to integrate the extracted features by each sub-networks, which is an input to a fully connected layer used for prediction. On the BraTS 2017 dataset, the Triplanar-CNN were trained separately for each modality, and each corresponding model yields more than 0.9 AUC for classifying glioma into two groups. Moreover, averaging the probability of glioma grading by all four MRI modalities boosts the classification performance compared to either of the four separately, achieved a patient-level grading result of 95.8% and 0.985 in accuracy and AUC, respectively, outperforming state-of-the-art results. Five-fold-cross validation was used to evaluate the models. To sum up, we developed a fully automated CNN-based model that can be translated into a clinical tool for non-invasive diagnosis of glioma using pre-operative MRI scans in a rapid and accurate way, leading to better outcomes for patients

Keywords— Deep Learning, Convolutional Neural Networks, Glioma grading, Multi-Modal MRI

I. Introduction

Gliomas that arise from glial cells are the most dangerous cancer in the world. It is also the most common primary brain tumor that represents about 80% of total malignant brain tumors [1]. Depending on the underlying histopathological and molecular parameters of the tumor along with the World Heal-th Organization (WHO) tumor stratification guidelines, gliomas can be categorized as high-grade glioma (HGG, grade III and IV) or low-grade glioma (LGG, grade I and II) [2]. Patients with LGG, which are slow-growing tumors, have a higher survival rate than patients with HGG tumors, which occur in the most vital and complex organ in human bodies [3]. The most common treatment of patients with glioma encompasses surgical resection, radiotherapy, chemotherapy, and follow-up plans, either separately or in combination. However, patients with LGG and HGG may follow different treatment options. Therefore, a rapid and accurate preoperative stratification of glioma grades is critical in predicting the disease prognosis, and in using a custom-made treatment strategy perfectly tailored to each individual patient.

For the past years, in practice, prediction of glioma grades has been based on either removing a section of the tumor during surgery or a tumor biopsy is used to take a sample of the tumor by inserting a needle through the skull into the brain, then a pathologist examines the tumor tissue under a microscope. However, these approaches are invasive, subjective task, prone to sampling and even time-consuming for experienced specialists due to the complex mechanical properties of the tissues, which vary from patient to patient [4]. Thus, it was indispensable to come up with a preoperative approach that can be used to noninvasively classify glioma grades rapidly and accurately via a comprehensive imaging of the tumor area.

Recently, technological advances have led to the acquisition of higher resolution images of a human brain with noninvasive medical imaging devices, such as magnetic resonance imaging (MRI). These powerful diagnostic tools adjunct the clinicians to visualize the smaller structures and abnormalities in the brain through visual inspection of images that provide valuable information concerning tumor type and grade [5]. Thus, glioma grading from imaging data avoids the invasive biopsy and the sampling error procedure. However, this comes at the cost of generating a bunch number of images per patient, which does not appear feasible that the clinician himself has to visually inspect morphological features of the tumor region in the whole brain MRI images, calculate a bunch of numbers, makes some advanced statistics and then combines the numbers into a meaningful diagnosis. Therefore, the development of a computer-assisted technique (CAD) to aid the radiologists for more accurate and fast glioma grading based on the multi-modal brain MRI volumetric data acquired prior to any invasive examination is of great current interest [6].

Approaches of CAD-based grading of glioma can be coarsely divided into two categories: (1) the handcrafted-based approaches. In this approach, the tumor regions are usually segmented first either through the manual or automatic way,

Abdela Ahmed Mossa   Ulus Çevik
Faculty of Engineering / Çukurova University
Turkey

the features including multi-parametric histogram and image texture are extracted from the segmented regions, and machine learning methods are applied to give prediction after feature selection [7]. However, major challenges remain for methodological development to optimize feature extraction and selection. In addition, the preprocessing procedure is complex and time-consuming. (2) the deep learning (DL) based approaches- DL methods especially CNNs that learn the features relevant to the problem directly from datasets, without the intermediate step of radiologists or handcrafted radiomics feature extraction, is the most common emerging technology in the medical field and other domains [8]–[10], and have also recently been employed for glioma grading [11].

There have been attempts to accurately predict brain tumor patients' glioma grade from volumetric MRI images, using various CNN algorithms, and have reported good results. On one hand, the researchers modified the state-of-the-art CNN architectures, which have been proved to have high accuracy in natural image classification or leverage them via transfer learning. For instance, Khawaldeh et al. [12] reconstructed a 3D FLAIR modality in the axial plane, and selected slices that contain lesions to develop a 2D CNN model, which was based on AlexNet architecture [13], for glioma grading. Similarly, Yang et al. [14] used CNN models trained on natural images, GoogLeNet [15], and fine-tuned it using axial T1Gd slices and achieved a glioma grading prediction accuracy of 94.5%. On the other hand, many researchers developed their own CNNs. For instance, Pereira et al. developed a 3D CNN architecture to classify glioma grades from automatically defined tumor regions [16].

Although there exist several DL-based techniques for glioma grade prediction from pre-operative volumetric MRI images, they have certain limitations that need to be considered. The first limitation of most of these systems is, they exploit only one modality or only axial plane slices, despite the different type of plane slices and modalities could provide complementary information on grading glioma patients. Furthermore, many existing DL models are based on 3D CNN without considering the insufficient training dataset available to reach the full potential of the CNN architecture. Lack of available pre-trained 3D CNN model is also another challenge to use 3D CNN for small datasets. Besides, 3D CNNs are not computationally efficient. When 3D datasets sizes are small, multi-view CNN architectures have been designed and trained to recognize volumetric images from a collection of their views on 2D images with greater accuracy compared to 3D CNNs [17]. Hence, to overcome the above described limitations, we developed a novel 2D CNN architecture, Triplanar-CNN, which is an extension of the multi-view CNN, for predicting glioma grade from multi-slice 2D images of a volumetric MRI that enables us to leverage pre-trained 2D CNNs and achieves high classification performance for noninvasively classifying glioma into two groups, using MRI data obtained from BraTS'17 dataset.

In summary, we developed a deep learning model, Triplanar-CNN, to discriminate between LGG and HGG using pre-operative multimodal MRI images. The Triplanar-CNN architecture, which is a three-column CNN, takes as input a three-dimensional MRI scan and computes the output probability indicating the grade of glioma in that scan in five stages (Fig. 1). In the first stage, preprocessing, we reconstructed all the multimodal 3D MRI scans as a set of 2D stacked slices in the axial, coronal and sagittal planes. In the second stage, using the three columns, we leverage pre-trained 2D models for extracting features from each plane each modality 2D projected stacked slices. In the third stage, extracted features from the three columns are consolidated into a fixed size feature map. In the fourth stage, for each modality, classification is performed using a fully connected layer and a sigmoid activation function on top of it. Finally, the individual classification performance of each modality was compared and the probability predictions of the four modalities were averaged to obtain the final more accurate classification result at the patient level.

The remaining of this article is organized as follows. In Section II, we present the dataset used for this paper followed by our proposed approach in detail including image pre-processing, and the 2D-CNN architecture. Section III illustrates the results. Finally, discussion, and conclusions are presented in sections IV and V, respectively.
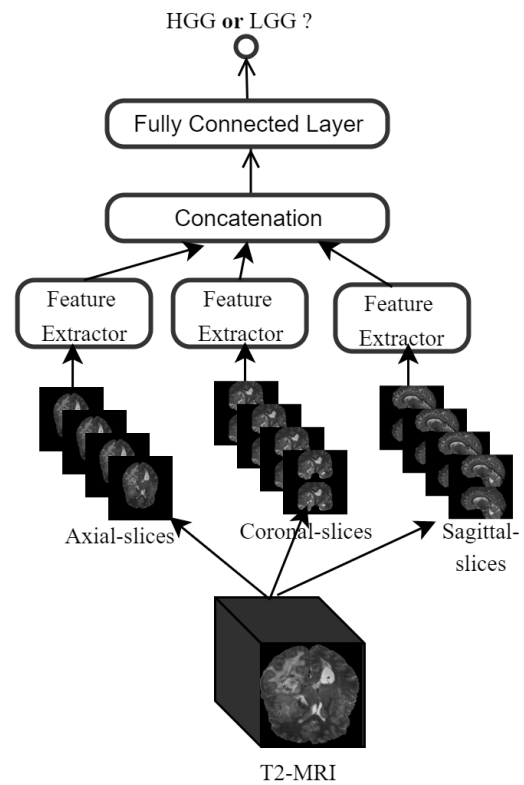

Figure 1. A schematic representation of the Triplanar-CNN architecture pipeline. Each 3D MRI modality is represented by a sequence of 2D slices in the axial, coronal and sagittal planes, giving rise to our Triplanar CNN architecture. Consecutive axial, coronal, and sagittal slices, respectively, are stacked as image input to the feature extractor. The feature extraction layers are employed to extract view-specific features from each modality. The concatenation layer integrates features obtained from each column, followed by a fully connected layer with a sigmoid activation function that maps the 3D MRI scan to probability prediction in the 0 to 1 range. The figure depicts only for T2 modality. Similar architectures were used for FLAIR, T1Gd, and T1 modalities.

II. Materials and Methods

A. Dataset and Image Preprocessing

The data used in the preparation of this article were obtained from the BRATS'17 training dataset [18], [19], which consists of multi-modal MRI exams of 285 glioma patients (75 LGG and 210 HGG). For each patient, four MR imaging modalities were provided: FLAIR, T1Gd, T1, and T2. The images in the dataset were skull-stripped, registered to a common space, resampled to a uniform isotropic resolution (1mm3), and size-adapted to 240×240×155.

For deep learning models, input sizes of all the images should be the same and should match the input size of the model architecture. As we planned to leverage 2D CNN models pre-trained on natural images of a fixed image resolution as well, we reconstructed each 3D MRI scan to a group of 2D stacked slices in the axial, coronal and sagittal planes. Some slices at the beginning and end of each plane, which did not contain any brain tissue, were discarded to avoid processing the background. We then crop each slice to a fixed size of 224×224 pixels for axial, and 112×224 pixels for coronal and sagittal views, respectively. To avoid the effect of image enlarging on the classification performance, two consecutive coronal slices were concatenated to form 224×224 pixels. Similarly, two consecutive sagittal slices were concatenated and reshaped to 224×224 pixel sizes.

We rescaled the intensity values of the slices to (0,255) range, convert them to PNG format, and normalized to have zero mean and unit variance. Then, all the reconstructed PNG images of an MRI in the axial, coronal and sagittal views were stacked together, and saved it in serialized form with pickle toolbox, respectively. Therefore, our input shape turned to be (n, 3,224,224). The values can be interpreted such that first value holds for the number of stacked images in the axial, coronal or sagittal views, respectively. The second value for the number of color channels, third and fourth values for width and height of PNG images. The first value varies across different views and modalities. All the preprocessing stages were done using the Python programming language and NiBabel package [20].

B. The Triplanar-CNN Model Architecture and Training

CNN architecture, which was introduced by LeCun et al. for handwritten digit recognition [21], consists of convolutional and pooling layers occurring in an alternative fashion to extract high-level features from the input 2D images, followed by dense layers to perform classification. As depicted in Fig 2. we developed a 2D CNN architecture, Triplanar-CNN, for classification of gliomas from volumetric MRI exams, inspired by prior work of Nan Wu [22]. It takes a three-dimensional MRI scan (a sequence of slices in the axial, coronal and sagittal planes) as input and outputs a probability indicating the grade of glioma in that scan. The architecture consists of three core modules: (i) three view-specific columns, each based on the feature extraction layers of AlexNet architecture that consists of five convolutional layers and three pooling layers, and a customized global and average pooling layers to reduce the number of features and consolidate into a fixed size feature map. For each of the three columns, we apply transfer learning from AlexNet pre-trained on natural images and fine-tuned the columns with the preprocessed axial slices, coronal slices, and sagittal slices, respectively. The three columns output a fixed-dimension hidden representation for each of the three views (Fig. 3). Similar feature extraction architecture has been used in our prior work for tuberculosis report generation using chest CT scans [8]. Besides AlexNet, three different existing state-of-the-art deep learning models (ResNet [23], VGG [24], and GoogLeNet) as the feature extractor, pre-trained on natural image classification tasks, were also tested in our experiments, and discussed in detail in section III. (ii) a fusion layer to integrate the output of the three-column vectors into a single dimension vector, which is fed into the next layer, fully connected layer and (iii) a fully connected layer with a binary cross-entropy loss and sigmoid activation function to obtain the final prediction in the 0 to 1 range.

We trained four instances of the same Triplanar-CNN architecture, one for each modality, resulting in four different results for a patient, one from each modality. The probability of glioma grading by the four modalities on the validation set were also averaged to obtain the final more accurate single classification result at the patient level. Class labels for HGG patients were set to 1, and class labels for LGG patients were set to 0.

The dataset was split into five-folds for cross-validation experiments using stratified random sampling to ensure that each fold contains an approximately equal proportion of patients of both the grades. The models were trained using the Adam optimization algorithm [25], and each model optimized at different learning rates between 10^{-4} and 10^{-6}. To combat overfitting, we early-stopped the parameter optimization and training process when the validation loss did not decrease for 20 epochs. We then selected the version of the model with the lowest average loss on the validation dataset as our final model candidate. Our models were implemented and trained by the widely used deep learning framework Pytorch [26], and a desktop computer with NVIDIA GeForce RTX 2070 GPU.
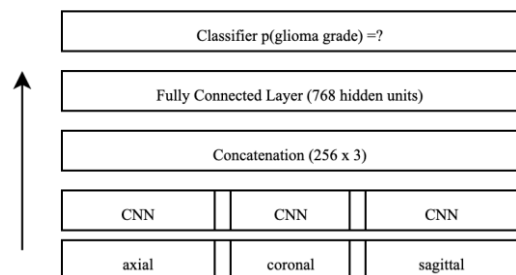

Figure 2. An overview of the proposed Triplanar-CNN network. CNN refers to the AlexNet-based convolutional neural network column from Fig 3. The arrow indicates the direction of the information flow.

Layer	Kernel size	Stride	#Features
max pooling			256
global average pooling			n x 256
max pooling	3 x 3	2 x 2	n x 256 x 6 x 6
convolutions	3 x 3	1 x 1	n x 256 x 13 x 13
convolutions	3 x 3	1 x 1	n x 256 x 13 x 13
convolutions	3 x 3	1 x 1	n x 384 x 13 x 13
max pooling	3 x 3	2 x 2	n x 192 x 13 x 13
convolutions	5 x 5	1 x 1	n x 192 x 27 x 27
max pooling	3 x 3	2 x 2	n x 64 x 27 x 27
convolution	11 x 11	4 x 4	n x 64 x 55 x 55
input (n x 3 x 224 x 224)			

Figure 3. Description of one CNN-based feature extractor column for a single view, and the dimension of feature maps after each layer. It transforms the input view of dimension $n \times 3 \times 224 \times 224$ into a 256-dimensional vector.

III. Results

We used five-fold cross-validation upon 285 patient subjects to reduce overfitting, wherefrom each class, 80% of all the image dataset is used as the training set and the rest as a validation set. That is, for each validation fold, the remaining other folds are used to train all the four models (for FLAIR, T1, T1Gd, and T2). Each fold had a similar ratio of HGG and LGG. Only 75 (26.3%) patients in fact had low-grade gliomas, whereas 210 (73.7%) had high-grade gliomas. Predictions from the FLAIR, T1, T1Gd, and T2 Triplanar-CNNs on the validation set were also ensembled using averaging the individual prediction results and generate a single output for each patient. The performance measures averaged over all the five folds are reported in Table I, including the area under the receiver operating characteristic curve (AUC), accuracy (ACC) (5), sensitivity (SEN) (6), specificity (SPE) (7), precision (PRE), and F1-score (FSC). The AUC curve summarizes the trade-off between the true positive rate and false positive rate for a binary classifier at various probability thresholds settings. Precision indicates the positive predictive value for a model whereas F1-score is the harmonic mean of precision and sensitivity.

True positive (TP) = Number of HGG patients correctly predicted as HGG (1)

False positive (FP) = Number of LGG patients incorrectly predicted as HGG. (2)

True negative (TN) = Number of LGG patients correctly predicted as LGG. (3)

False negative (FN) = Number of HGG patients incorrectly predicted as LGG. (4)

$$ACC = \frac{TP + TN}{(TP + FP + TN + FN)} \times 100 \quad (5)$$

$$SEN = \frac{TP}{(TP + FN)} \quad (6)$$

$$SPE = \frac{TN}{(TN + FP)} \quad (7)$$

As shown in Table I, the proposed approach achieved higher performance on every modality. It is also observed that ensembling all the four models resulted in improved performance over the individual models, which is 95.8% accuracy, and 0.985 AUC. The developed model achieved an AUC value of 0.89, 0.957, 0.96, and 0.983 when trained solely on FLAIR, T2, T1Gd and T2 modalities, respectively. In addition, the individual models obtained an accuracy, SPE, and SEN values greater than 90%, except FLAIR modality, which is around 80% in ACC, SPE, and SEN, demonstrating that the Triplanar-CNN model trained solely on single modalities are also able to extract discriminative features for distinguishing between HGGs, and LGGs. In addition, the result indicated that the Triplanar-CNN model trained solely with T1 modality has the best classification performance, 94.7% accuracy, with more than 2.5% improvement over the other individual modalities classification performance. The results indicated that T1 modalities have more discriminative features for glioma grading compared to the other modalities. In addition, we know that the higher the specificity, the lower the chance of misclassifying the LGG patients, and the higher the sensitivity, the lower the chance of misclassifying the HGG patients. But, regardless of the class imbalance on the dataset, the ensemble model stratifies the LGG and HGG patients with 0.99 and 0.95, in SPE and SEN, which shows the robustness of the approach that we used. As shown in Table II, the performance of the proposed approach on the individual cross validation was also consistent and high. The minimum AUC was 0.93, and the maximum was 1.0. To sum up, the result indicated that our ensemble model, comprising the four modalities, can effectively separate patients with LGG from those with HGG.

To further investigate the effectiveness of our proposed method, we also evaluated the performance of our proposed approach by using only 60% of the dataset (Folder, folder 4, and Folder 5) for training, 20% for validation (Folder 2), and 20% for testing (Folder 3). As shown in Table III, our model achieved similar performance to that of the five-fold cross-

validation, which shows that Triplanar CNN architecture pre-trained on natural images may even effectively discriminate HGG subjects from LGG when trained on very small datasets.

TABLE I. AUC AND ACCURACY VALUES OF THE INDIVIDUAL MODALITIES AND THEIR ENSEMBLE ON FIVEFOLD CROSS-VALIDATION

Models	AUC	ACC	SPE	SEN	PRE	FSC
FLAIR	0.89	83.2	0.8	0.84	0.92	0.88
T2	0.957	90.2	0.95	0.89	0.98	0.93
T1Gd	0.96	92.3	0.93	0.92	0.97	0.94
T1	0.983	94.7	0.95	0.95	0.98	0.96
Fused	**0.985**	**95..8**	**0.99**	**0.95**	**0.99**	**0.97**

TABLE II. AUC AND ACCURACY OF THE FOUR MODALITIES AND FUSED MODELS ON THE INDIVIDUAL FOLDER OF THE FIVEFOLD CROSS-VALIDATION

Folders	Models (ACC / AUC)				
	FLAIR	*T1*	*T1Gd*	*T2*	*Fused*
Folder 1	77.2 / 0.857	94.7 / 0.98	96.5 / 1.0	98.2/0.99	**98.2 / 1.0**
Folder 2	91.2 / 0.963	98.2 / 1.0	98.2/0.99	94.7 /1.0	**98.2 / 1.0**
Folder 3	87.7 / 0.91	98.2 / 0.99	100 / 1.0	93 / 0.99	**100 / 1.0**
Folder 4	89.5 / 0.959	98.2 / 1.0	94.7/0.99	89.5/0.96	**96.5 / 1.0**
Folder 5	70.2 / 0.763	84.2 / 0.93	71.9/0.81	75.4/0.84	**86 / 0.93**
Average	83.2 / 0.89	94.7 / 0.98	92.3/0.96	90.2/0.96	**95.8/ 0.98**

TABLE III. CLASSIFICATION ACCURACIES ACHIEVED FOR TRAINING, VALIDATION AND TEST DATASETS

Datasets	Models (AUC ± ACC)				
	FLAIR	*T1*	*T1CE*	*T2*	*Fused*
train	1.0 ± 99.4	1.0 ± 98.8	1.0 ± 100	1.0 ± 94.2	**1.0± 100**
valid	0.9 ± 89.3	1.0 ± 98.2	0.99± 96.5	0.99 ± 94.7	**1.0±96.5**
test	0.83±77.2	0.983±94.7	1.0 ± 100	1.0 ± 96.5	**1.0±100**

IV. Discussions

Pre-operative grading of gliomas is a critical step in brain tumor diagnosis as it can help to choose the appropriate treatment strategy, and MRI is usually the preferred noninvasive imaging technique in the pre-surgical assessment of patients with glioma. Hence, studies have been conducted that can predict the grading of gliomas solely based on MRI features. However, conventional grading of gliomas using traditional machine learning methods on MRI scans that rely on the extraction of handcrafted features, which may not fully capture the useful information from every cross-sectional MRI slices, is time-consuming, subjective and sometimes could not be accurate. In this work, we substantially advanced this body of knowledge by using deep learning techniques that discover automatic features and does prediction directly from the image. We did this through a novel 2D convolutional neural network using pre-operative multi-modal volumetric MRI scans of the brain, which reduce the computational requirement of exploring volumetric MRI images. We found that the proposed approach achieved high classification performance (Accuracy = 95.8%, AUC = 0.985, SEN = 0.95, SPE = 0.99) using a small training dataset of less than 285 patients MRI exams.

Based on the obtained findings, our results are consistent with some previous reports that classify glioma using solely on single and multiple MRI modality, in that using multiple modalities improves classification performance which is consistent with the visual examination from medical doctors. However, in contrary to previous studies where FLAIR modalities were reported to be superior to T1Gd images [27], and T2 to be more superior to FLAIR [28] in providing a more complete characterization of glioma, our findings show that T1 modality is more predictive than the other three modalities, and the individual classification performance of FLAIR modality was the lowest. Though in our experiment, the individual classification performance of the four modalities, in multiple performance measurements, were consistent across different 2D architectures and sub-datasets, more investigation has to be done on comparison of the individual classification performance of various modalities for glioma grading.

A. *Choice of Feature Extractor*

Feature extractor is the most important component of any deep learning architecture, and hence in our proposed architecture too. As it is shown in the methods section, we used pre-trained models as a feature extractor. Hence, in order to choose the right state-of-the art CNN model as part of our Triplar-CNN architecture, we experimented with four state-of-the-art deep learning models, each pre-trained on natural images, and found AlexNet to perform best (Table IV). As depicted in Fig.3, AlexNet outperforms the other architectures, though its difference is not significant with VGG19. For this reason, we chose AlexNet as a feature extractor in our proposed Triplanar-CNN architecture.
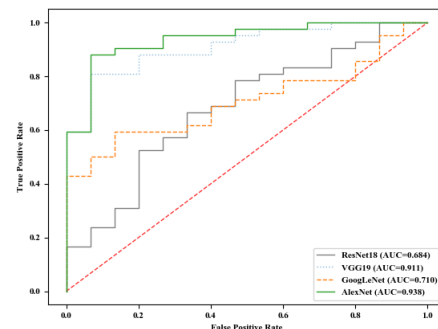

Figure 4. Comparison among ReseNet18, VGG19, GoogLeNet, and AlexNet. A Scatter plot of the false positive rate (1-specificity) versus true positive rate (sensitivity) for the models shows that AlexNet achieves the highest AUC.

TABLE IV. PRE-TRAINED MODELS COMPARISON

Models	AUC	ACC	SEN	SPE	PRE	FSC
ResNet	0.684	71.9	0.905	0.2	0.76	0.826
VGG19	0.911	86.0	0.881	0.8	0.925	0.902
GoogLeNet	0.71	71.9	**0.952**	0.067	0.741	0.833
AlexNet	**0.938**	**87.7**	0.857	**0.933**	**0.973**	**0.911**

B. ***Results Comparison with Some Existing Works***

A comparison of our results with some of the state-of-the-art methods recently published works for glioma grading evaluated on the same public dataset and other private or public datasets are presented in Table V. Our method is superior to other existing algorithms in multiple performance metrices including AUC and accuracy. However, direct comparisons between our study and some previous studies based on private data are not possible without having access to MRI data of the patient cohort, as reported in these studies. Nevertheless, we show the state-of-the-art studies on glioma grading that have used pre-operative MRI modalities. Overall, the results reported in this paper are high and innovative, and can be considered as the state-of-the-art for stratifying glioma grades using deep learning models.

TABLE V.   COMPARISON WITH THE EXISTING CAD-BASED STUDIES

Authors	Approaches	Year	Dataset size	AUC	Accuracy (%)
Hsieh [29]	Handcrafted	2017	107	0.89	88
Cho [30]	Handcrafted	2017	274	0.89	89.8
Zhang [7]	Handcrafted	2017	120	0.945	94.5
Chen [31]	Handcrafted	2018	274	0.96	91.3
Cho [32]	Handcrafted	2018	*285	0.921	88.8
Ge [28]	DL	2018	*285	-	93.3
Khawaldeh [12]	DL	2018	*285	-	91.2
Pereira [16]	DL	2018	*285	0.984	92.98
Bi [33]	Handcrafted	2019	60	0.86	91
Zhu [34]	DL	2019	181	0.82	-
Reza [35]	Handcrafted	2019	*285	0.88	88
Proposed	DL	2020	*285	**0.985**	**95.78**

*Indicates authors used the same cohort dataset used in this paper

V.   **Conclusion**

We developed a deep learning model based on a novel 2D-CNN architecture for predicting glioma grade fast and accurately, using only small multi-modal MRI images, and this could potentially be of great help to radiologists to employ better decisions in further management of patients with more appropriate treatments. Future studies will need to extend the results presented herein larger datasets, increase the number of classes that can be classified by including more sub-grades of glioma tumors.

Acknowledgment

References

[1] Z. He, R. A. Mitteer, Y. Mou, and Y. Fan, "Multimodality Targeting of Glioma Cells," in *Glioblastoma*, Elsevier, 2016, pp. 55–72.

[2] D. N. Louis *et al.*, "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary," *Acta Neuropathol.*, vol. 131, no. 6, pp. 803–820, Jun. 2016.

[3] Q. T. Ostrom, H. Gittleman, G. Truitt, A. Boscia, C. Kruchko, and J. S. Barnholtz-Sloan, "CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011–2015," *Neuro. Oncol.*, vol. 20, no. suppl_4, pp. iv1–iv86, Oct. 2018.

[4] R. J. Jackson *et al.*, "Limitations of stereotactic biopsy in the initial management of gliomas," *Neuro. Oncol.*, vol. 3, no. 3, pp. 193–200, 2004.

[5] M. Smits and M. J. Van Den Bent, "Imaging correlates of adult glioma genotypes," *Radiology*, vol. 284, no. 2, pp. 316–331, Aug. 2017.

[6] J. Kong *et al.*, "Machine-based morphologic analysis of glioblastoma using whole-slide pathology images uncovers clinically relevant molecular correlates," *PLoS One*, vol. 8, no. 11, p. e81049, Nov. 2013.

[7] X. Zhang *et al.*, "Optimizing a machine learning based glioma grading system using multi-parametric MRI histogram and texture features," *Oncotarget*, vol. 8, no. 29, pp. 47816–47830, Jul. 2017.

[8] A. A. Mossa, A. M. Yibre, and U. Çevik, "Multi-view CNN with MLP for diagnosing tuberculosis patients using CT scans and clinically relevant metadata," in *CEUR Workshop Proceedings*, 2019, vol. 2380.

[9] H. Eriş and U. Çevik, "Implementation of Target Tracking Methods on Images Taken from Unmanned Aerial Vehicles," in *SAMI 2019 - IEEE 17th World Symposium on Applied Machine Intelligence and Informatics, Proceedings*, 2019, pp. 311–316.

[10] N. Bien *et al.*, "Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet," *PLoS Med.*, vol. 15, no. 11, Nov. 2018.

[11] Z. Li, Y. Wang, J. Yu, Y. Guo, and W. Cao, "Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma," *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, Dec. 2017.

[12] S. Khawaldeh, U. Pervaiz, A. Rafiq, and R. S. Alkhawaldeh, "Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks," *Appl. Sci.*, vol. 8, no. 1, p. 27, Dec. 2017.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[14] Y. Yang *et al.*, "Glioma grading on conventional MR images: A deep learning study with transfer learning," *Front. Neurosci.*, vol. 12, no. NOV, Nov. 2018.

[15] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015, vol. 07-12-June, pp. 1–9.

[16] S. Pereira, R. Meier, V. Alves, M. Reyes, and C. A. Silva, "Automatic brain tumor grading from MRI data using convolutional neural networks and quality assessment," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11038 LNCS, pp. 106–114.

[17] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller, "Multi-view convolutional neural networks for 3D shape recognition," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, vol. 2015 Inter, pp. 945–953.

[18] S. Bakas *et al.*, "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data*, vol. 4, p. 170117, 2017.

[19] B. H. Menze *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.

[20] B. C. Matthew Brett, Michael Hanke, MARC-ALEXANDRE CÔTÉ, Paul McCarthy, Chris Cheng, "NiBabel- Access a cacophony of neuro-imaging file formats." .

[21] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, 1998, vol. 86, no. 11, pp. 2278–2323.

[22] N. Wu *et al.*, "Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening," Mar. 2019.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016-Decem, pp. 770–778.

[24] [K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[25] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[26] A. Paszke *et al.*, "Automatic differentiation in PyTorch," in *Advances in Neural Information Processing Systems 32*, 2019, pp. 8024–8035.

[27] S. Gurwara, A. Azzawe, S. Jacobs, A. Ravi, M. Hardjasudarma, and E. G. Toledo, "Comparing the volume of gliomas in the brain in FLAIR and post-contrast T1-weighted MRI sequences.," *J. La. State Med. Soc.*, vol. 162, no. 5, pp. 265–6.

[28] C. Ge, I. Y. H. Gu, A. S. Jakola, and J. Yang, "Deep Learning and Multi-Sensor Fusion for Glioma Classification Using Multistream 2D Convolutional Networks," *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2018, pp. 5894–5897, Jul. 2018.

[29] K. L. C. Hsieh, C. M. Lo, and C. J. Hsiao, "Computer-aided grading of gliomas based on local and global MRI features," *Comput. Methods Programs Biomed.*, vol. 139, pp. 31–38, Feb. 2017.

[30] H. H. Cho and H. Park, "Classification of low-grade and high-grade glioma using multi-modal image radiomics features," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2017, pp. 3081–3084.

[31] W. Chen, B. Liu, S. Peng, J. Sun, and X. Qiao, "Computer-Aided Grading of Gliomas Combining Automatic Segmentation and Radiomics," *Int. J. Biomed. Imaging*, vol. 2018, 2018.

[32] H. ho Cho, S. hak Lee, J. Kim, and H. Park, "Classification of the glioma grading using radiomics analysis," *PeerJ*, vol. 2018, no. 11, 2018.

[33] X. Bi, J. G. Liu, and Y. S. Cao, "Classification of low-grade and high-grade glioma using multiparametric radiomics model," in *Proceedings of 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, ITNEC 2019*, 2019, pp. 574–577.

[34] Z. Y. *et al.*, "A deep learning radiomics model for preoperative grading in meningioma," *Eur. J. Radiol.*, vol. 116, pp. 128–134, 2019.

[35] S. M. S. Reza, M. D. Samad, Z. A. Shboul, K. A. Jones, and K. M. Iftekharuddin, "Glioma grading using structural magnetic resonance imaging and molecular data," *J. Med. Imaging*, vol. 6, no. 2, 2019.

About author (s) :

Abdela Ahmed Mossa received his B.Sc. and M.Sc. degree in Computer science from BahirDar University, Ethiopia in 2008 and Osmania University, India, in 2013, respectively. He is currently pursuing PhD in Computer Engineering at Çukurova University, Turkey. His current research include medical image analysis, machine learning, and deep learning

Ulus Çevik received his Ph.D. degree in Electrical and Computer Engineering from University of Sussex, in 1996. He is currently a Professor, and head of Electrical and Electronics department at Çukurova University, Turkey. His current research area include Computer vision, Medical image analysis, and volume rendering