Pakistan Journal of Scientific Research, PJOSR
ISSN (p): 0552-9050
Volume: 3, Number: 2, Year: 2023

Classification of Isocitrate Dehydrogenase (IDH) Mutation Status in Gliomas Using Transfer Learning

Maria Fayyaz[1], Nauman Riaz Chaudhry[2], Reema Choudhary[3]

[1]Department of Computer Science, Virtual University of Pakistan, Lahore, 54000, Pakistan
[2]Department of Computer Science, University of Gujrat, Gujrat, 50700, Pakistan
[3]Department of Software Engineering, University of Gujrat, Gujrat, 50700, Pakistan

Corresponding author: Nauman Riaz Chaudhry (Email: nauman.riaz@uog.edu.pk).

***Abstract*: Isocitrate Dehydrogenase (IDH) mutation is a significant genetic alteration in brain tumors. Its diagnosis is vital for the prognosis of low-grade Glioma and secondary-grade Glioma patients. Physicians used invasive methods to diagnose the Gliomas, which was an unsafe method, but now advanced magnetic resonance imaging techniques are being used for tumor visualization and treatments. Some studies have used classical machine learning and deep learning methods for the problem of IDH mutation status detection using magnetic resonance images. Recent studies have used concatenation of deep and handcrafted features to achieve superior performance. This study used concatenations of the deep features extracted through pre-trained convolutional neural networks (CNNs) to detect IDH mutation status using magnetic resonance images. Five top accuracies on the ImageNet dataset were considered when selecting the pre-trained CNNs. Magnetic resonance images were acquired from The Cancer Genome Atlas Glioblastoma Multiforme and The Cancer Genome Atlas Low-Grade Glioma. All experiments (performed using features extracted from each CNN and their concatenation) were compared with each other and state-of-the-art. The proposed technique achieved 99% accuracy while efficiently using data and computational resources.**

***Index Terms*--Isocitrate Dehydrogenase, Convolution Neural Networks, IDH Wild-Type, IDH Mutated.**

I. INTRODUCTION

A brain tumor is the abnormal growth of cells in the brain. There are many types of brain tumors. Glioma is the most common type of brain tumor, which originates within the glial cells of the brain. Gliomas can be classified according to these glial cells and the hereditary features of the tumor. It supports predicting the tumor's present and future behavior for the prognosis and treatment. A Glioma impacts the function of the brain, and it can be dangerous for life depending upon tumor growth rate and its occurrence place. World Health Organization (WHO) has classified tumors into four grades (from I to IV) depending upon the severity and aggressiveness. Grade I is the least aggressive, while grade IV Glioma is the most aggressive and life-threatening. IDH mutations are found in 70% to 80% of lower grades. Tumors with normal IDH genes, IDH negative or IDHwt, tend to behave more aggressively. Patients withIDH-mutated Gliomas are expected to survive longer than those with IDHwt Gliomas. Detection of such biomarkers helps diagnose tumors and plan treatment more effectively. Hence, IDH mutation status detection is vital for doctors to treat their patients effectively and guide them in clinical decisions [1].

Invasive methods of biopsy are risky and sometimes impossible because of tumor location and the patient's health. Therefore, non-invasive methods are preferred instead of biopsy [2].

MRI of the brain is a harmless test in which magnetic fields and radio waves can generate detailed brain and spinal cord images. MRI can have three orientations: axial orientation is from top to down, and MRI's front-to-back plane is called coronal. Sagittal is side-to-side orientation.

A. ARTIFICIAL INTELLIGENCE-BASED ANALYSIS OF MEDICAL IMAGES

Brain tumor MRI can be examined manually, which is time-consuming and less accurate. However, many computer-aided systems can detect IDH mutation status to overcome this issue. Many AI models can play a significant role in medical and healthcare research. Classical machine learning techniques can be used to classify brain tumor types, but the issue with these techniques is that they rely on handcrafted features that are decided by human experts [3]. Handcrafted feature extraction requires a deep technical skill set and understanding of data engineering and machine learning algorithms [4]. These issues have been combated by using deep learning models, which automatically extract features and compute weights during model training. These extracted features are then inputted into the classifier. Deep neural networks need much data to avoid overfitting during image classification [5]. This data-hungry nature of deep learning models becomes problematic for data-

scarce domains like medical imaging. Transfer learning is a machine learning technique in which pre-trained models are used for a new problem. Any pre-trained neural network model is selected as an initial point for a related problem. This model is tuned to align with the latest issue. It avoids the requirement to train from scratch, speeding up the training and improving model performance. In biomedical image classification, training data acquisition may be expensive and complicated. Transfer learning solves this shortcoming by harnessing the training performed on a source domain for the target domain. Transfer learning can play a vital role in classifying IDH mutation status in Gliomas as it overcomes the challenges related to feature extraction, data scarcity, and generalization of diverse datasets. It reduces model training time and achieves high performance in medical image analysis. It is like a tool that uses the knowledge learned from large-scale datasets and applies that knowledge to specific tasks of medical imaging classification. It assists in efficient and accurate planning of diagnosis and treatment of glioma patients [6]. The current study aims to combine features extracted from ImageNet pre-trained CNNs to classify IDH mutation status in Glioma patients non-invasively and with acceptable reliability. This research proposed a transfer learning-based method for the task. Five pre-trained CNN models named Xception, ResNet152V2, InceptionV3, InceptionResNetv2, and NASNetLarge were selected based on their high accuracy on ImageNet pre-training. The study not only used these features individually for the IDH mutation status classification but also used their concatenation for this purpose. The datasets of TCGA-GBM (TCGA-GBM - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki, 2019) and TCGA LGG (TCGA- LGG - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki, 2019) were used to train these pre-trained CNN models. Brain tumor image features were extracted by using selected pre-trained CNN models. These features were inputted into the classifier. Afterward, these five feature sets were concatenated horizontally and passed to the classifier. Classical Machine learning and deep learning techniques in the medical field face challenges, such as limited and imbalanced datasets. Moreover, these techniques require computational resources management and addressing data quality and ethical data concerns. This study has advantages in classifying IDH mutation status in Gliomas. It makes pre-trained model usage more efficient and reduces the need for time and computation resources for model training. The proposed technique increases the model performance by leveraging features learned from different image datasets. It enables image feature extractions to act more effectively, adapting to relevant small datasets through fine tuning, adapting to small datasets through fine-tuning, thus improving predictive accuracy in clinical settings.

The paper is organized as follows: Section 2 presents the literature review and analysis of classical machine learning and deep learning methods. Section 3 describes materials and techniques used in the study. Section 4 evaluates the performance of the process and discusses the contribution. Section 5 ends with a conclusion and future work.

II. RELATED WORK

A. CLASSICAL MACHINE LEARNING-BASED TECHNIQUES

In classical machine learning, features are extracted manually. Features are identified and described based on relevance to the given problem. Background understanding and domain knowledge are required to extract useful features. This study needs experts if classical machine learning is used for feature extraction from brain tumor images. The study [7] classified data into subcategories using a support vector machine (SVM). The model attained an accuracy of 84.62%. The study [8] focused on developing a technique that automatically predicts brain tumor type. Features were extracted by using pre-trained AlexNet. Brain tumors were classified by using SVM with an accuracy of 95.9%. The study [9] focused on different genes. Extreme Gradient Boosting (XGBoost) was used for the training. BraTS 2017 and 2018 were used for model testing. The test accuracy of the model was 83%. The study [10] selected features in three stages: Mann–Whitney U-test, Pearson test, and most minor absolute shrinkage and selection operator (LASSO). These selected features were used to train the SVM model. The best model was established without the Pearson test and attained an accuracy of 0.719 for IDH mutation status classification.

In the study [11], eight classical ML methods were evaluated based on the criteria of performance and stability. Results showed that Random Forest is a promising ML method for IDH genotype prediction. For the study[12], three-level models based on multimodal magnetic resonance radiomic Glioma subtype classification SVM were used. The model was proposed to reduce the complication of multi-class classification to multi-binary classification. The performance of the model on the test set was 83.4% accuracy. In the study [13], five MRI parameters were derived through brain scanning and inputted to a random forest algorithm, giving IDH mutation status classification accuracy of 88%. Table 1 summarizes the literature review of classical machine learning-based techniques.

TABLE I
SUMMARY OF LITERATURE REVIEW CLASSICAL MACHINE-LEARNING-BASED TECHNIQUES

Study	Methodology	Dataset(s)	Results
[7]	SVM	TCIA	82.7
[8]	*Pre-trained Alexnet,SVM*	Standard CE-MRI dataset	95.9
[9]	*XGBoost*	*BraTS 2017,2018*	83
[10]	*SVM,l-SVM,r-SVM,*	*TCIA*	71.9
[11]	*kNN,FDA, NB,Adaboost and RF*	*TCIA*	RF(89) NN(83)
[12]	SVM	*TCIA, Beamount*	83
[13]	*Random Forest*	Hospital Medical	88

B. DEEP LEARNING-BASED TECHNIQUES

On the other hand, feature extraction can be done automatically using deep learning-based techniques without human intervention. Deep learning also requires a large amount of data, normally unavailable in the medical imaging domain. Another

point is that deep learning needs high computational power and computer memory for valuable feature extraction and model training. The issues mentioned above of less medical data availability and high computational requirements to train the models can be solved by applying transfer learning. In transfer learning, pre-trained models can be reused for a new problem. This is popular in deep learning as it allows deep neural networks to be trained with a small amount of data. As in the current study, massively labeled medical information is not available. Therefore, transfer learning is the best option.

The study[14] proposed a three-dimensional CNN (3D CNN) named TDABNet that can determine IDH status in 3D Gliomas MRIs with 96.44% accuracy. In this study[15], Sparse graph convolutional network (SGCN) and ResNet-152 were used as pre-training to learn representations of mRNA expression data and histopathological images. In multi-model fusion, a regressor and a classifier are used for result output (classifier and survival analysis). The accuracy of MultiCo-Fusion is 0.756 ± 0.032. In the study [16], the proposed model contained three components. First is the image learner based on 3D CNN, a geometric learner based on GNN. The collaboration between the two networks maximizes prediction performance by up to 89.2%. The study [17] used the voxel-wise clustering method for IDH-mutation status classification for multiparametric MRI. A linear SVM was used to classify IDHmut status. Ten-class clustering showed an accuracy of 91%. In the study [18], a deep orthogonal fusion (DOF) model was used to predict Glioma patients' overall survival (OS) by using diverse multi-model data. Its performance was a median concordance index (C-index) of 0.718 ± 0.064. In this study [19], artificial intelligence (AI) techniques were applied to fluorescent images to examine Glioma during surgery speedily and with high accuracy. Auto contrast algorithms and denoising were implemented to normalize image data. CNN achieved 0.945 AUC without any time overhead.

The study [20] proposed a learnable group convolution-based segmentation method. In this method, learnable group convolution has replaced convolution in feature extraction. It has obtained 90.25%, 86.20%, and 80.36% precision for whole tumors, tumor cores, and enhancing tumors, respectively.

This study [21] presented a GBM-tailored deep learning model in which CNN was applied to multiparametric MRIs. The model contained a 4-block 2D CNN applied to all MRI sequences. GBM-specific deep learning model obtained 83% accuracy on rCBV maps for IDH mutation prediction. The study [22] implemented a deep CNN to segment different brain tumor regions. It was found that gene signatures are highly correlated with patients' survival. Experiments obtained 70% segmentation accuracy. In the study[23], a novel 3D multi-task deep learning model was presented for the segmentation and IDH genotype prediction (SGPNet), which was based on CNN and obtained an accuracy of 82%. The study [24] proposed a deep learning-based multimodal feature fusion model. The VGG16 model was applied to brain Glioma classification using transfer learning. It attained 95% accuracy for IDH detection. The study [25] proposed a fully automated hybrid CNN model, and radiomics were used.This model contained two sub-models; one CNN part is used for tumor segmentation, whereas the 2nd part is a CNN-based classifier. CNN achieved 87.9% accuracy. In the research[26], CNN features were combined with a new framework, graph-based semi-supervised learning, that learns for labels of the unlabeled data. The model obtained an accuracy of 86.53%.

For the study [27], key features were extracted using principal component analysis (PCA), and the results showed that deep neural networks can learn key components of an image without manual feature selection. The model attained an accuracy of 94%. In the study [28], the residual CNN was trained for MR sequences. Training, validation, and test accuracies were 87.3%, 87.6%, and 89.1%, respectively. Table 2 summarizes the literature review based on deep learning techniques.

TABLE 2

SUMMARY OF LITERATURE REVIEW DEEPLEARNING-BASED TECHNIQUES

Study	Methodology	Dataset(s)	Results
[14]	*-*	*TDABNet*	*96.44*
[15]	*SGCN and ResNet-152*	*TCGA*	*75.6*
[16]	*GNN + CNN*	*TCIA*	*89.2*
[17]	*SVM*	*approved by the institutional review board*	*91*
[18]	*CNN*	*TCIA, Beijing Tiantan*	*C-index 0.718 ± 0.064*
[19]	*CNN*	*Hospital, Capital Medical University*	*94*
[20]	*DCNN*	*TCGA, BraTS, and TCGA*	*70*
[21]	*4-block 2D CNN+Softmax*	*Approved by the institution Review Board (IRB)*	*83*
[22]	*Learnable group convolution-based segmentation method*	*BraTS 2018*	*90.25*
[23]	*SGPNet*	*BraTS and TCGA Glioma databases*	*82*
[24]	*Pre-trained VGG16*	*Human Cancer Hospital, Severance Hospital*	*95*
[25]	*Fully automated hybrid CNN*	*Seoul National University Hospital,TCIA*	*87.9*
[26]	*Graph-based semi-supervised*	*TCGA*	*86.5*
[27]	*PCA*	*TCGA*	*94*
[28]	*Residual CNN*	*HUP. BWH,TCIA*	*89.1*

Machine-based and deep learning-based techniques for classifying IDH mutation status face challenges, including data diversity in clinical settings, interpretability, and data scarcity. Data imbalance, ethical considerations, and the requirement of robust validation and clinical integration workflows complicate their implementation. These issues need novel transfer learning techniques as they can overcome many issues faced by machine learning and deep learning in IDH mutation status in Gliomas. Transfer learning can mitigate the issue of data scarcity by reusing pre-trained models and knowledge gained from big-size databases. It also helps in feature extraction, generalizing the

medical image data, and enhancing interpretability. Transfer learning offers more trustworthy and clinically applicable solutions for diagnosing and treating Glioma patients.

III. MATERIALS AND METHODS

For this study experiments, medical image data of Gliomas was obtained from the TCIA repository then these images were labeled. Five pre-trained models were used to extract medical images features, these five image features sets were inputted to the classifiers individually. Afterwards, these five feature sets were concatenated horizontally to the classifier.

A. NEED OF TRANSFER LEARNING

This research used a transfer learning-based method to classify IDH mutation status using MRI scans. Like typical tasks involving medical images, limited data is available for the IDH mutation status. The transfer learning-based method has been chosen considering this data scarcity. ImageNet pre-trained CNNs, available as open-source through the Keras library, have been used. Five CNNs were selected out of all the available, and the criteria for this selection were the ImageNet accuracy and the model complexity. Brain tumor medical images were downloaded from TCIA. TCIA is a collection of medical imaging data of different cancer patients' imaging modalities. It offers imaging data diversity, such as MRI scans and CT scans. Various techniques are applied to glioma image data for preprocessing. Intensity normalization is used to reduce the variability of scan intensity value ranges. Non-brain tissues are removed by skull stripping to focus on the brain region of interest. In MRI scans, noise reduction techniques are used to enhance image quality. Data augmentation techniques are applied to TCIA data to enhance the dataset's quality, consistency, and diversity. Rotations, translations, scaling, and flipping to simulate variations in imaging orientation and improve model robustness. This dataset contained data labeled as 0 and 1 for IDHmut and IDHwt, respectively. Min-max normalization was applied to normalize data. Data normalization ensures that various features have similar value ranges. As well as increasing CNN model accuracy dramatically [29]. The dataset was divided into two subsets, one for training the model and the other for the model's evaluation. The top five pre-trained CNN models extracted features from Brain tumor nifty files. In this way, five feature sets were obtained. Each model was trained and evaluated individually, and after that, extracted features were concatenated horizontally and inputted into the model for training purposes. This model was evaluated, and the results showed that it had achieved excellent accuracy for IDH mutation classification.

B. DATA LABELING

The dataset contains two categories of brain tumors, one with IDHmut and the second with IDHwt, and it is labeled as 0 and 1 for IDHmut and IDHwt, respectively.

C. PRE-PROCESSING

For the TCIA website medical collection, medical images have been obtained from different medical centers using various MRI scanners. So, it was obvious that these images were diverse and needed to be normalized to overcome this heterogeneity. Min-max normalization was used for this research to normalize brain images.

D. TRAIN-TEST SPLIT

Data were split into two groups: training and testing. 80% of the data was used to train the CNN model, whereas 20% was used for model testing. In other words, 17205 slices (11160 of IDHwt and 6045 of IDHmut) were divided into two groups; 13764 were used for model training, whereas the remaining 3441 slices were used for model testing.

E. FEATURE EXTRACTION USING THE FIVE PRE-TRAINED MODELS

Many studies have used the concatenation of different handcrafted features [30]. In contrast, others have combined handcrafted features with deep features to improve accuracy [31-32]. Feature concatenation has been used for accuracy improvement [33]. This study has proposed a novel technique by concatenating deep features obtained from high-accuracy pre-trained CNNs. Five different, high-accuracy ImageNet pre-trained CNN models named Xception, ResNet152V2, InceptionV3, InceptionResNetv2, and NASNetLarge have been selected for this research. A detailed description of the selected pre-trained models is given in Table 3. The pre-trained networks were treated as automatic deep feature extractors during the feature extraction process. Their classifier layers were discarded, and only convolution bases were used. 2D slices were inputted to the convolution base to extract the deep features. The global average pooling layer was applied at the end to get one value for each feature map of the last layer from the convolution base.

TABLE 3
DESCRIPTION OF THE SELECTED PRE-TRAINED CNNS

Model	Size	Top-1 accuracy	Top-5 accuracy	Parameters	Depth
NASNetLarge	*343 MB*	*0.825*	*0.960*	*88, 949, 818*	-
InceptionResNetV2	*215 MB*	*0.803*	*0.953*	*55, 873, 736*	*572*
Xception	*88 MB*	*0.790*	*0.945*	*22, 910, 480*	*126*
ResNet152V2	*232 MB*	*0.780*	*0.942*	*60, 380, 648*	-
InceptionV3	*92 MB*	*0.779*	*0.937*	*23, 851, 784*	*159*

F. CLASSIFICATION

This study has two classes, IDHmut and IDHwt, so the final classification layer used Sigmoid activation to give the probability of input images belonging to these two classes. During the training of these classifiers, the loss was calculated from the predictions compared to the ground truth. Parameters of all layers were updated using the gradients of the loss function through backpropagation. Binary_crossentropy was used to compute the cross-entropy loss by comparing the predicted label

with the ground truth value. CNN model performance improved by decreasing loss. 'Adam' optimizer was used to modify the CNN attributes like weights and learning rate during training. As a result, it improves accuracy and reduces overall loss. The learning rate was set to 0.001(lr=0.001) and no of epochs to 100. This research used a transfer learning-based method to classify IDH mutation status using MRI scans. Like typical tasks involving medical images, limited data is available for the IDH mutation status. The transfer learning-based method has been chosen considering this data scarcity. ImageNet pre-trained CNNs, available as open-source through the Keras library, have been used. Five CNNs were selected out of all the available, and the criteria for this selection were the ImageNet accuracy and the model complexity. Brain tumor medical images were downloaded from TCIA. This dataset contained data labeled as 0 and 1 for IDHmut and IDHwt, respectively. Min-max normalization was applied to normalize data. Data normalization ensures that various features have similar value ranges. As well as it increases CNN model accuracy dramatically. The dataset was divided into two subsets, one for training the model and the other for the model's evaluation. The top five pre-trained CNN models extracted features from Brain tumor nifty files. In this way, five feature sets were obtained. Each model was trained and evaluated individually, and after being concatenated horizontally, it was inputted into the model for training purposes. This model was evaluated, and the results showed that it had achieved excellent accuracy for IDH mutation classification.

a) CLASSIFICATION USING INDIVIDUAL FEATURES

Five extracted feature sets from each pre-trained model were inputted to the classifier separately to classify for IDH mutation status for all the models individually.

b) CLASSIFICATION USING CONCATENATED FEATURES

The extracted feature sets of each pre-trained CNN were concatenated horizontally, and then the concatenated feature set was inputted to the classifier. As a result of the concatenation, the feature vector size for each example was 17204. Experiments were repeated with three different classifier configurations: classifier layers 52,256,128, classifier layers 256,256,128, and classifier layers 256,128. The block diagram of the methodology used in this research is given in Fig 1.

IV. RESULTS

Five top CNN models were used for feature extraction. These extracted feature sets from each pre-trained model were inputted into the classifier separately. Models' performances were evaluated by calculating accuracy, specificity, and sensitivity. Experiments were repeated three times for each model, and all these evaluation measures were carried out.

A. RESULTS USING INDIVIDUAL FEATURE SETS

a) ACCURACIES USING INDIVIDUAL FEATURE SETS

Table 4 presents the average accuracies of three experiments for the five CNN models. These average accuracies are represented in pictorial form to make them more understandable, as shown in Fig 2.

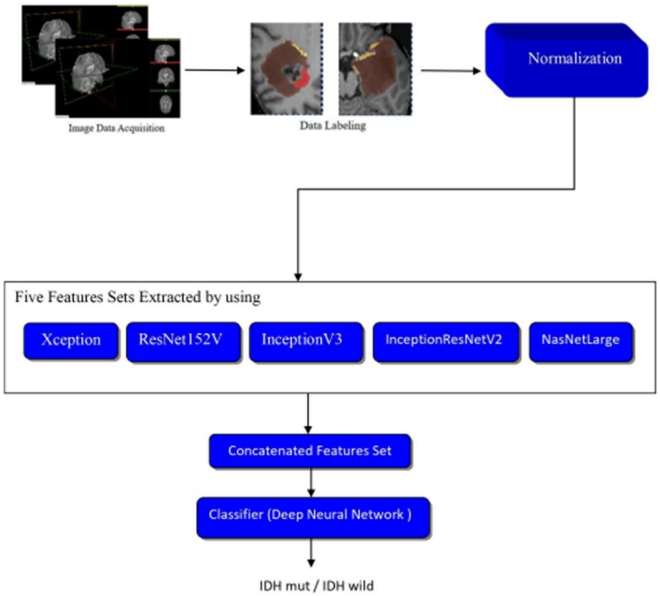

FIGURE 1. Complete Methodology

TABLE 4
AVERAGE TEST ACCURACIES OF THE SELECTED PRE-TRAINED CNNS

Models	Average Test Accuracy
Xception	*89.7%*
ResNet152V2	*92.1%*
InceptionV3	*92.1%*
InceptionResNetV2	*86.94%*
NASNetLarge	*91.1%*

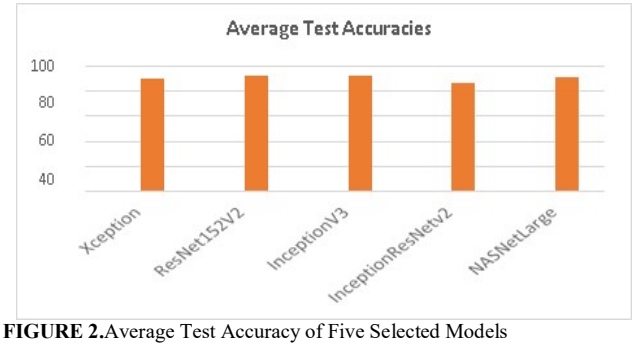

FIGURE 2. Average Test Accuracy of Five Selected Models

Features extracted by using the top five models were concatenated horizontally. This concatenation obtained eleven thousand seven hundred twelve features against one example. When the model was trained with a concatenated features set, it attained higher accuracy than the individual models.

b) ACCURACIES FOR CONCATENATED FEATURES

Since the results for the concatenated features were excellent, we

repeated the experiments with classifiers for different complexity. Three classifiers were used for this purpose. The first one had 512, 256, and 128 units (and the final sigmoid layer) in its three layers, the second 256, 256, and 128 units, while the third one had 256 and 128 units in its two layers. Each experiment was repeated three times for each classifier using concatenated features like individual feature sets. Accuracy values showed that the classifier with the highest complexity attained the best accuracy. Table5 shows the average test accuracies of different configurations of the proposed model.

TABLE 5
AVERAGE TEST ACCURACIES OF THE DIFFERENT CONFIGURATIONS OF THE PROPOSED MODEL

Models	Average Test Accuracy
512,256,128	*99.5%*
256,256,128	*99.2%*
256,128	*99.1%*

c) COMPARISONS OF ALL MODELS AS WELL AS CONCATENATED FEATURES

Fig 3 shows the comparisons of all models and compares all model's performance with the model performing concatenated features. Concatenated features attained the highest accuracy.

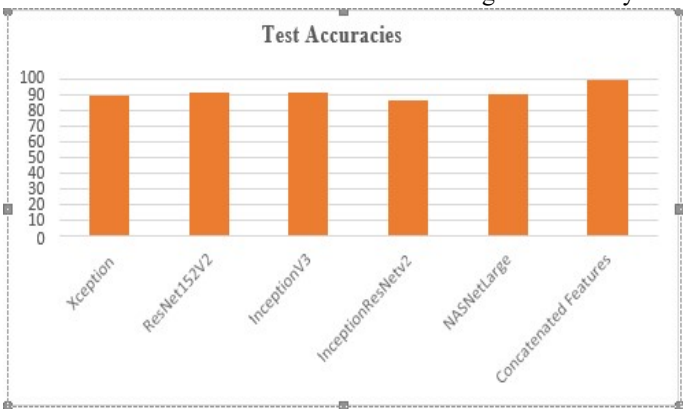

FIGURE 3. Bar Chart for Accuracies of All Models and Concatenated Features

B. SENSITIVITY AND SPECIFICITY

Sensitivity and specificity are the measures that give us the per-class performance of the proposed technique. Table 6 represents the sensitivity score for three experiments repeated for individual CNN models. Table 7 represents the specificity score for three experiments repeated for individual CNN models.

TABLE 6
AVERAGE SENSITIVITY OF INDIVIDUAL MODELS

Models	Average Test Accuracy
Xception	*90.8%*
ResNet152V2	*97.2%*
InceptionV3	*95.7%*
InceptionResNetV2	*90.8%*
NASNetLarge	*89%*

TABLE 7
AVERAGE SPECIFICITY OF INDIVIDUAL MODELS

Models	Average Test Accuracy
Xception	*89.73%*
ResNet152V2	*92.2%*
InceptionV3	*90.9%*
InceptionResNetV2	*87.1%*
NASNetLarge	*90%*

Table 8 and 9 represent the average sensitivity and specificity scores For the three Configurations, respectively.

TABLE 8
AVERAGE SENSITIVITY SCORES OF DIFFERENT CONFIGURATIONS

Models	Average Sensitivity
512,256,128	*99.1%*
256,256,128	*98.82%*
256,128	*99.0%*

TABLE 9
AVERAGE SPECIFICITY SCORES OF DIFFERENT CONFIGURATIONS

Models	Average Specificity
512,256,128	*99.3%*
256,256,128	*99.5%*
256,128	*97.8%*

C. AUC-ROC USING INDIVIDUAL FEATURE SETS

Table 10 presents the average AUC-ROC of three experiments for the five CNN models, whereas for the model when given concatenated features, has AUC-ROC 0.99.

TABLE 10
AUC-ROC CURVE OF THE SELECTED PRE-TRAINED CNNS

17with IDHwt, labeled as 0 and 1 for IDHmut and IDHwt, respectively.

Models	AUC
Xception	*0.94*
ResNet152V2	*0.97*
InceptionV3	*0.98*
InceptionResNetV2	*0.89*
NASNetLarge	*0.97*

Table 11 shows the results of IDH mutation status prediction results of individual feature sets as well as concatenated features set. From all evaluation measures, it can be seen that concatenated features performed the highest among all models' feature sets. Average accuracy of concatenated features is highest among all individual feature sets extracted by five different pre-trained models. Likewise, average sensitivity and specificity of concatenated features are highest as compared to their relative values respectively 98.9% and 98.8%.

TABLE 11
COMPARISON OF MODELS WITH INDIVIDUALS FEATURES SETS AND CONCATENATED FEATURES SET

Models	Average Accuracy	Average Sensitivity	Average Specificity	AUC
Xception	*89.7%*	*90.8%*	*89.73%*	*0.94*
ResNet152V2	*92.1%*	*97.2%*	*92.2%*	*0.97*
InceptionV3	*92.1%*	*95.7%*	*90.9%*	*0.98*
InceptionResNetV2	*86.9%*	*90.8%*	*87.1%*	*0.89*
NASNetLarge	*91.1%*	*89%*	*90%*	*0.97*
Concatenated Features	***99.2%***	***98.9%***	***98.8%***	***0.99***

V. COMPARISONS OF THIS STUDY WITH PREVIOUS STUDIES

As this study is conducted on the TCIA TCGA dataset, here is a comparison of this study's results with the previous studies using different techniques in which the same dataset was used for their experiments. The results of these studies and the current study are given in Table 12, showing that this study showed the best results among all previous studies. A recent study gave 99% accuracy in the classification of IDH mutation status in Gliomas.

TABLE 12

COMPARISON OF THE STUDY WITH PREVIOUS STUDIES

Study	Methodology	Dataset(s)	Results
[15]	*SGCN and ResNet-152*	*TCGA*	*75.6*
[16]	*GNN+CNN*	*TCIA*	*89.2*
[18]	*CNN*	*TCIA*	*C-index 0.718±0 .064*
[20]	*DCNN*	*TCGA, Severance Hospital*	*70*
[25]	*Fully automated hybrid CNN*	*Seoul National University Hospital, TCIA*	*87.9*
[26]	*Graph-based semi-supervised*	*TCGA*	*86.5*
[28]	*Residual CNN*	*HUP. BWH, TCIA*	*89.1*

VI. DISCUSSION

Accurate oncology helps doctors to make an effective prognosis plan for individual patients. Features extraction from medical images can make an investigation of brain tumors more accurate. In this paper, the prediction of IDH mutation status has been performed. The method has proposed a deep learning transfer learning technique and concatenation of feature sets to enhance the prediction accuracy. As deep learning techniques were used, feature engineering was not required, and raw inputs (image slices) were inputted into the models. Deep learning is beneficial for dataset classification. Even when data is complicated and infeasible to recognize by a human, such as genomic expression profiles, deep learning can extract features by learning them. TCIA medical datasets have inconsistent images, so normalization is required to attain unified patterns of all brain tumor image slices. High overall performance was achieved on TCIA for the IDH mutation classification task. The accuracy of the model trained using a concatenated feature set was compared with the models' accuracies using individual features. Results showed that the proposed technique of applying transfer learning for feature extraction using five high-accuracy models and then concatenating these feature sets horizontally has boosted the model. Transfer learning has increased the accuracy of classification of IDH mutation status in Glioma patients. This is helpful for doctors in taking more precise and valid diagnostic decisions. Transfer learning is time saving as it allows physicians to get IDH mutation status classification results more quickly. By applying transfer learning techniques, it becomes very easy for doctors to interpret medical images data and make diagnostic assessments. Consequently, it leads to improved Glioma patient's outcomes by facilitating in treatment planning personalization. The collaboration of machine learning experts and physicians has advanced the understanding of IDH mutation status in Gliomas.

VII. CONCLUSION

IDH mutation is a genetic alteration found in brain tumors. Its diagnosis can play an essential role in the prognosis of LGG and secondary-grade Glioma patients. Physicians were using invasive methods for IDH mutation investigation, which were unsafe. But now, MRI are being used for tumor visualization. There are machine-learning techniques that can be used for brain tumor classification. In classical machine learning methods, handcrafted features train the models. Meanwhile, in deep learning methods, CNN has been used to analyze visual imagery. In some studies, deep learning was used to detect IDH mutation status using MRIs. For this study, MRIs of brain tumors were obtained from TCGA-LGG and TCGA-GBM. Data was obtained from TCIA, an extensive publicly available cancer image archive. One hundred eleven Glioma patients.'

Data were used for the study, 39 cases were IDH-wild type, and 72 were IDH mut. Data were normalized by using the min-max normalization method. Image slices were labeled 0 and 1. All image' slices of the IDHwt class were labeled 0, and image slices of the IDH mut class were labeled 1. 2D tumorous slices were selected from 3D MR volumes and then inputted to the selected pre-trained CNN models for deep features extraction.

Brain tumor image slices' features were extracted using five high-accuracy Image Net pre-trained CNN models. These features were used to train the models individually, and their performance was evaluated. Afterward, five feature sets were concatenated horizontally. The sequential model was trained and evaluated on the test set. The final classifier used the sigmoid function to predict belonging to one of the two classes. The research has shown that the performance of pre-trained CNNs for IDH mutation status prediction about their performance in Image Net classification was excellent. Moreover, the performance of features extracted from individual pre-trained CNNs was compared to when these features were combined for IDH mutation status prediction. It can be concluded that the model that used concatenated features attained higher accuracy. The current

study combined features extracted from ImageNet pre-trained CNNs to classify IDH mutation status in Glioma patients non-invasively and with acceptable reliability. The proposed method has used pre-trained CNNs, producing reliable predictions while being computationally cheap. The contribution of this study is that deep features extracted from different pre-trained CNNs achieve better results than these features used individually.

REFERENCES

[1] Al-Khallaf, H. (2017). Isocitrate dehydrogenases in physiology and cancer: Biochemical and molecular insight. Cell & Bioscience, 7(1), 37. https://doi.org/10.1186/s13578-017-0165-3

[2] Sciortino, T., Secoli, R., d'Amico, E., Moccia, S., Conti Nibali, M., Gay, L., Rossi, M., Pecco, N., Castellano, A., De Momi, E., Fernandes, B., Riva, M., & Bello, L. (2021). Raman Spectroscopy and Machine Learning for IDH Genotyping of Unprocessed Glioma Biopsies. Cancers, 13(16), 4196. https://doi.org/10.3390/cancers13164196

[3] Van der Velden, B. H. M., Kuijf, H. J., Gilhuijs, K. G. A., &Viergever, M. A. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Medical Image Analysis, 79, 102470. https://doi.org/10.1016/j.media.2022.102470

[4] Zaharia, G.-E., Cercel, D.-C., &Dascalu, M. (2021). UPB at SemEval-2021 Task 1: Combining Deep Learning and Hand-Crafted Features for Lexical Complexity Prediction (arXiv:2104.06983). arXiv. https://doi.org/10.48550/arXiv.2104.06983

[5] Yadav, S. S., &Jadhav, S. M. (2019). Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data, 6(1), 113. https://doi.org/10.1186/s40537-019-0276-2

[6] Kim, H. E., Cosa-Linan, A., Santhanam, N., Jannesari, M., Maros, M. E., &Ganslandt, T. (2022). Transfer learning for medical image classification: A literature review. BMC Medical Imaging, 22(1), 69. https://doi.org/10.1186/s12880-022-00793-7

[7] Islam, M., Wijethilake, N., & Ren, H. (2021). Glioblastoma multiforme prognosis: MRI missing modality generation, segmentation and radiogenomic survival prediction. Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society, 91, 101906. https://doi.org/10.1016/j.compmedimag.2021.101906

[8] Taha, B., Li, T., Boley, D., Chen, C. C., & Sun, J. (2021). Detection of Isocitrate Dehydrogenase Mutated Glioblastomas Through Anomaly Detection Analytics. Neurosurgery, 89(2), 323–328. https://doi.org/10.1093/neuros/nyab130

[9] Peng, H., Huo, J., Li, B., Cui, Y., Zhang, H., Zhang, L., & Ma, L. (2021). Predicting Isocitrate Dehydrogenase (IDH) Mutation Status in Gliomas Using Multiparameter MRI Radiomics Features. Journal of Magnetic Resonance Imaging: JMRI, 53(5), 1399–1407. https://doi.org/10.1002/jmri.27434

[10] Liu, S., Shah, Z., Sav, A., Russo, C., Berkovsky, S., Qian, Y., Coiera, E., & Di Ieva, A. (2020). Isocitrate dehydrogenase (IDH) status prediction in histopathology images of Gliomas using deep learning. Scientific Reports, 10(1), 7733. https://doi.org/10.1038/s41598-020-64588-y

[11] Wu, S., Meng, J., Yu, Q., Li, P., & Fu, S. (2019). Radiomics-based machine learning methods for isocitrate dehydrogenase genotype prediction of diffuse Gliomas.

[12] Lu, C.-F., Hsu, F.-T., Hsieh, K. L.-C., Kao, Y.-C. J., Cheng, S.-J., Hsu, J. B.-K., Tsai, P.-H., Chen, R.-J., Huang, C.-C., Yen, Y., & Chen, C.-Y. (2018). Machine Learning-Based Radiomics for Molecular Subtyping of Gliomas. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, 24(18), 4429–4436. https://doi.org/10.1158/1078-0432.CCR-17-3445

[13] De Looze, C., Beausang, A., Cryan, J., Loftus, T., Buckley, P. G., Farrell, M., Looby, S., Reilly, R., Brett, F., & Kearney, H. (2018). Machine learning: A useful radiological adjunct in determination of a newly diagnosed glioma's grade and IDH status. Journal of Neuro-Oncology, 139(2), 491–499. https://doi.org/10.1007/s11060-018-2895-4

[14] Ai, L., Bai, W., & Li, M. (2022). TDABNet: Three-directional attention block network for the determination of IDH status in low- and high-grade Gliomas from MRI. Biomedical Signal Processing and Control, 75, 103574. https://doi.org/10.1016/j.bspc.2022.103574

[15] Tan, K., Huang, W., Liu, X., Hu, J., & Dong, S. (2022). A multi-modal fusion framework based on multi-task correlation learning for cancer prognosis prediction. Artificial Intelligence in Medicine, 126, 102260. https://doi.org/10.1016/j.artmed.2022.102260

[16] Wei, Y., Li, C., Chen, X., Schönlieb, C.-B., & Price, S. J. (2022). Collaborative learning of images and geometrics for predictingisocitrate dehydrogenase status of glioma (arXiv:2201.05530). arXiv. https://doi.org/10.48550/arXiv.2201.05530

[17] Hagiwara, A., Tatekawa, H., Yao, J., Raymond, C., Everson, R., Patel, K., Mareninov, S., Yong, W. H., Salamon, N., Pope, W. B., Nghiemphu, P. L., Liau, L. M., Cloughesy, T. F., &Ellingson, B. M. (2022). Visualization of tumor heterogeneity and prediction of isocitrate dehydrogenase mutation status for human Gliomas using multiparametric physiologic and metabolic MRI. Scientific Reports, 12(1), Article 1. https://doi.org/10.1038/s41598-022-05077-2

[18] Braman, N., Gordon, J. W. H., Goossens, E. T., Willis, C., Stumpe, M. C., &Venkataraman, J. (2021). Deep Orthogonal Fusion: Multimodal Prognostic Biomarker Discovery Integrating Radiology, Pathology, Genomic, and Clinical Data (arXiv:2107.00648). arXiv.https://doi.org/10.48550/arXiv.2107.00648

[19] Shen, B., Zhang, Z., Shi, X., Cao, C., Zhang, Z., Hu, Z., Ji, N., & Tian, J. (2021). Real-time intraoperative glioma diagnosis using fluorescence imaging and deep convolutional neural networks. European Journal of Nuclear Medicine and Molecular Imaging, 48(11), 3482–3492. https://doi.org/10.1007/s00259-021-05326-y

[20] Liu, H., Li, Q., & Wang, I.-C. (2021). A Deep-Learning Model with Learnable Group Convolution and Deep Supervision for Brain Tumor Segmentation. Mathematical Problems in Engineering, 2021, e6661083. https://doi.org/10.1155/2021/6661083

[21] Pasquini, L., Napolitano, A., Tagliente, E., Dellepiane, F., Lucignani, M., Vidiri, A., Ranazzi, G., Stoppacciaro, A., Moltoni, G., Nicolai, M., Romano, A., Di Napoli, A., &Bozzao, A. (2021). Deep learning can differentiate IDH-mutant from IDH-wild type GBM.

[22] ZadehShirazi, A., McDonnell, M. D., Fornaciari, E., Bagherian, N. S., Scheer, K. G., Samuel, M. S., Yaghoobi, M., Ormsby, R. J., Poonnoose, S., Tumes, D. J., & Gomez, G. A. (2021). A deep convolutional neural network for segmentation of whole-slide pathology images identifies novel tumour cell-perivascular niche interactions that are associated with poor survival in glioblastoma. British Journal of Cancer, 125(3), 337–350. https://doi.org/10.1038/s41416-021-01394-x

[23] Wang, Y., Wang, Y., Guo, C., Zhang, S., & Yang, L. (2021). SGPNet: A Three-Dimensional Multitask Residual Framework for Segmentation and IDH Genotype Prediction of Gliomas. Computational Intelligence and Neuroscience, 2021, e5520281. https://doi.org/10.1155/2021/5520281

[24] Yao, W., & Thomas, S. (2021). Deep Learning-Based Magnetic Resonance Imaging Image Feature Analysis for Pathological Classification of Brain Glioma. Scientific Programming, 2021, e6778009. https://doi.org/10.1155/2021/6778009

[25] Choi, Y. S., Bae, S., Chang, J. H., Kang, S.-G., Kim, S. H., Kim, J., Rim, T. H., Choi, S. H., Jain, R., & Lee, S.-K. (2021). Fully automated hybrid approach to predict the IDH mutation status of Gliomas via deep learning and radiomics. Neuro-Oncology, 23(2), 304–313. https://doi.org/10.1093/neuonc/noaa177

[26] Ge, C., Gu, I. Y.-H., Jakola, A. S., & Yang, J. (2020). Deep semi-supervised learning for brain tumor classification. BMC Medical Imaging, 20(1), 87. https://doi.org/10.1186/s12880-020-00485-0

[27] Chang, P., Grinband, J., Weinberg, B. D., Bardis, M., Khy, M., Cadena, G., Su, M.-Y., Cha, S., Filippi, C. G., Bota, D., Baldi, P., Poisson, L. M., Jain, R., & Chow, D. (2018). Deep-Learning Convolutional Neural Networks Accurately Classify Genetic Mutations in Gliomas. American Journal of Neuroradiology, 39(7), 1201–1207. https://doi.org/10.3174/ajnr.A5667

[28] Chang, K., Bai, H. X., Zhou, H., Su, C., Bi, W. L., Agbodza, E., Kavouridis, V. K., Senders, J. T.,Boaro, A., Beers, A., Zhang, B., Capellini, A., Liao, W., Shen, Q., Li, X., Xiao, B., Cryan, J., Ramkissoon, S., Ramkissoon, L., … Kalpathy-Cramer, J. (2018). Residual Convolutional Neural Network for the Determination of IDH Status in Low- and High-Grade Gliomas from MR Imaging. Clinical Cancer Research, 24(5), 1073–1081. https://doi.org/10.1158/1078-0432.CCR-17-2236

[29] Why is data normalization necessary for machine learning models? (n.d.). Deepchecks. Retrieved January 23, 2024, from https://deepchecks.com/question/why-is-data-normalization-necessary-for-machine-learning-models/

[30] Roy, S., Nag, S., Maitra, I. K., &Bandyopadhyay, S. K. (2013). A Review on Automated Brain Tumor Detection and Segmentation from MRI of Brain (arXiv:1312.6150). arXiv. https://doi.org/10.48550/arXiv.1312.6150

[31] Hasan, A. M., Jalab, H. A., Meziane, F., Kahtan, H., & Al-Ahmad, A. S. (2019). Combining Deep and Handcrafted Image Features for MRI Brain Scan Classification. IEEE Access, 7, 79959–79967. https://doi.org/10.1109/ACCESS.2019.2922691

[32] Saba, T., Sameh Mohamed, A., El-Affendi, M., Amin, J., & Sharif, M. (2020). Brain tumor detection using fusion of hand crafted and deep learning features. Cognitive Systems Research, 59, 221–230. https://doi.org/10.1016/j.cogsys.2019.09.007

[33] Al-Qazzaz, S. (2020). Deep learning-based brain tumor images segmentation and its extension to stroke lesion segmentation [PhD Thesis]. Cardiff University]