Deep Learning Approach for Radiogenomic Classification of Brain Tumor

Spoorthy K R
Dept. of ECE
PES University
Bengaluru, India
spoorthykr@pesu.pes.edu

Akash R Mahdev
Dept. of ECE
PES University
Bengaluru, India
akashrmhdev@pesu.pes.edu

Vaishnav B
Dept. of ECE
PES University
Bengaluru, India
vaishnavb@pesu.pes.edu

Shruthi M.L.J
Dept. of ECE
PES University
Bengaluru, India
shruthimlj@pes.edu

2022 IEEE 19th India Council International Conference (INDICON) | 978-1-6654-7350-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/INDICON56171.2022.10039760

***Abstract* – Medical imaging of brain tumors is critical for tumor pathology characterization and early tumor detection. Current genetic analysis techniques take a long time and necessitate the surgical removal of samples of brain tissue. Multimodal brain tumor images can be accurately classified to speed up identification and lessen patient anxiety. In this research a novel deep-learning-based classification model is proposed that checks for the genetic subtype of glioblastoma to detect MGMT (O6-methylguanine-DNA methyltransferase) promoter methylation, the presence of which has demonstrated to be a positive prognostic factor and a powerful predictor of responsiveness to chemotherapy. It allows for less invasive diagnosis and treatments. A scenario-specific medical images dataset, the RSNA-MICCAI dataset, is used to test our model which comprises structural multi-parametric MRI (mpMRI) scans, in DICOM format. This work aims to investigate the feasibility of applying deep learning techniques for predicting the methylation state of the MGMT promoter.**

Keywords - Convolutional Neural Network, Efficient Net, Genetic sequence, MGMT promoter, MRI Scans.

I. Introduction

The growth of abnormal cells in the brain is called brain tumor which can further be classified into two types: malignant tumors and benign (non-cancerous) tumors. The normal method to detect brain tumors is by performing MRI scans (Magnetic Resonance Imaging scans). Information about the growth of the cancer tissue in the brain is inferred from the MRI scans. A Glioblastoma is a common type of malignant brain tumor which is life threatening. It is usually very aggressive and spreads fast with the median life expectancy being around a year. O6-Methylguanine-DNA Methyltransferase (MGMT), a DNA repair enzyme, lessens the impact of alkylating chemotherapy on tumor cells and hence is a favorable factor for determining the reactiveness of the tumor towards chemotherapy.

The detection of brain tumor can be done using a Machine Learning or Deep Learning algorithm. When MRI scans of the brain are used as input to these algorithms, the prediction of the brain tumor can be done very quickly and with high accuracy. This helps radiologists make quick decisions. Chemotherapy and radiotherapy are frequently used as post-surgical treatments after the tumor has been removed. Because radiation can kill both normal and cancer cells, radiotherapy can have serious side effects. On the other hand, chemotherapy kills cancer cells by applying a chemical to the guanine DNA, which prevents the synthesis of new DNA. Chemotherapy can, however, be rendered ineffective due to an enzyme known as MGMT. The methylation status of MGMT's promoter determines its function. The transcription of the enzyme is impacted by the promoter region's methylation, which may result in an efficient chemotherapy treatment. As a result, the MGMT promoter's methylation state has been identified as a prognostic indicator and predictor of chemotherapy response.

Currently, iterative and prolonged procedures are necessary for genetic analysis of the tumor to forecast the methylation status, requiring surgical excision of brain tissue samples and weeks of genetic characterization. Additionally, the procedure itself could have negative effects. The number of surgeries performed will be drastically reduced if the methylation status of the enzyme can be accurately predicted by means of computer vision (i.e., Radio Genomics).

The Radiological Society of North America (RSNA) and the Medical Image Computing and Computer Assisted Intervention Society (the MICCAI Society) have partnered together to enhance glioblastoma diagnosis and treatment planning. The use of MRI (magnetic resonance imaging) images to train and evaluate the model to determine the MGMT promoter's methylation status in this research helps patients with brain cancer benefit from less invasive diagnostic procedures. The introduction of novel and tailored therapy options prior to surgery has the potential to improve brain cancer patients' care, survival, and results.

Recent advances in deep learning techniques have proven its effectiveness in identifying and learning patterns and hence have the potential of recognizing and extracting biomarkers that are indicative of the methylation status of MGMT promoter from MRI scans and provide a non-invasive, effective and precise alternative that would cause less suffering. The proposed work explores the use of multi-modal MRI data to detect MGMT methylation by training and evaluating a Deep Neural Network (DNN) framework developed using state-of-the-art convolutional neural network architectures. We pre-process the available 2D MRI images and investigate how transfer learning and fine-tuning can improve classification performance with the available data following which we present our implementation findings and discuss the efficacy of the approach in detecting the methylation status using MRI scans.

The rest of the paper is organized as follows. In Sec. II, we

provide a brief summary of relevant studies using CNNs for tumor classification and genetic analysis. In Sec. III, we talk about the dataset used and the image pre-processing steps applied along with model selection. Sec. IV talks about the proposed framework and implementation. In Sec. V, we present the findings and results of our implementation. We provide our concluding remarks and future work in Sec. VI

II. Literature Survey

The use of CNN for automatic detection of brain tumors is gathering significant attention [1], [2], [3]. The proposed system consisting of multiple layers including but not limited to the convolutional layer, pooling layer, non-linearity layer and fully-connected layer, trained on pre-processed MRI brain scans, classifies input images as tumorous or normal based on the training features. For visual learning and brain tumor recognition, convolutional neural networks (CNNs) are the most commonly used machine learning algorithm.

To classify T1-wce brain magnetic resonance images into four categories (glioma, meningioma, pituitary, and no tumor), the research in paper [4] suggests a CNN-based dense EfficientNet with min-max normalization. The developed network was an EfficientNet variant with dense and drop-out layers added.

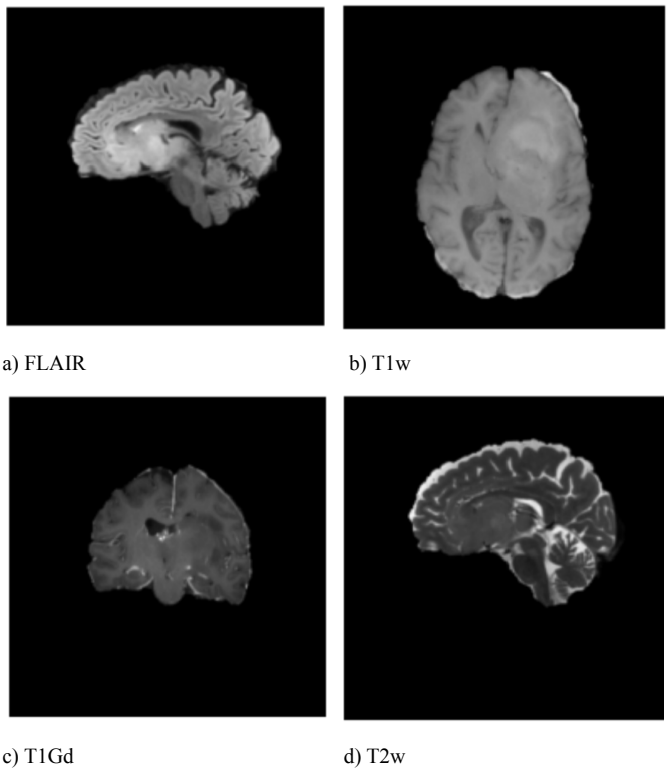
Considering the multi-modality characteristic of the MRI scans is drawing significant attention because of the abundant amount of information available for the model to learn [5]. A novel deep-learning-based classification model that explored lite attention mechanism, multi-modal feature aggregation and separable embedding is discussed that performs accurate classification of multi-modal brain tumor images.

A CNN model that is simpler than already-existing pre-trained networks for brain tumor classification, tested on T1-weighted contrast-enhanced magnetic resonance images is proposed in paper [6]. The performance of this network was evaluated using four approaches: combinations of two *10*-fold cross-validation methods and two databases. The generalization capability of this network was tested with one of the *10*-fold methods and the improvements were tested by using an augmented image database.

The work done in paper [7] proves that residual deep neural networks (ResNET) namely ResNET-50, ResNET-34 and ResNET-18 can be used to predict biomarkers from MRI scans without the need for extensive preprocessing like a distinct tumor segmentation step.

The research in paper [8] uses Convolutional recurrent neural networks (CRNN) to predict the methylation status of MGMT enzymes by using MRI scans as an input.

The correlation between preoperative MRI variables and MGMT methylation status in glioblastoma has recently attracted research attention, with factors like edema volume and tumor location among many others serving as the primary emphasis [9]. With a similar objective, we propose a deep-learning approach that utilizes the self learning capabilities of SOTA neural networks to identify key biomarkers that aid in predicting the methylation status of glioblastomas using MRI data.



a) FLAIR b) T1w

c) T1Gd d) T2w

Fig. 1. scans (a-d) represent the four modalities of imaging

III. Working Principle

A. Dataset

In this paper, we use the RSNA-MICCAI dataset [10], a multi-center brain tumor MRI dataset, which was made available as an outcome of a collaboration between the Medical Image Computing and Computer Assisted Intervention Society (the MICCAI Society) and the Radiological Society of North America (RSNA) to enhance the use of AI in the detection and treatment of brain tumors. The dataset was made available by these two institutions in an effort to better understand glioblastoma from a computer vision standpoint. We assess the effectiveness of CNNs in predicting the methylation state of the MGMT promoter using this dataset.

This dataset includes a large number of clinically collected multi-parametric MRI (mpMRI) scans of gliomas, as DICOM (.dcm) files, with a pathologically verified diagnosis and available MGMT promoter methylation status. These mpMRI scans are pre-processed, i.e., co-registered, interpolated, and skull-stripped, and describe four different imaging modalities T1-weighted pre-contrast (T1w), T1-weighted post-contrast (T1Gd), T2-weighted (T2w) and Fluid Attenuated Inversion Recovery (FLAIR) volumes as depicted in Figure 1. Each type of scan defines a specific focus during imaging and is obtained with various clinical protocols using different scanners from various institutions. For instance, FLAIR captures the effect following cerebrospinal fluid (CSF) suppression, when liquid signals like water are repressed to highlight other areas whereas T2W emphasizes the variation in lateral tissue

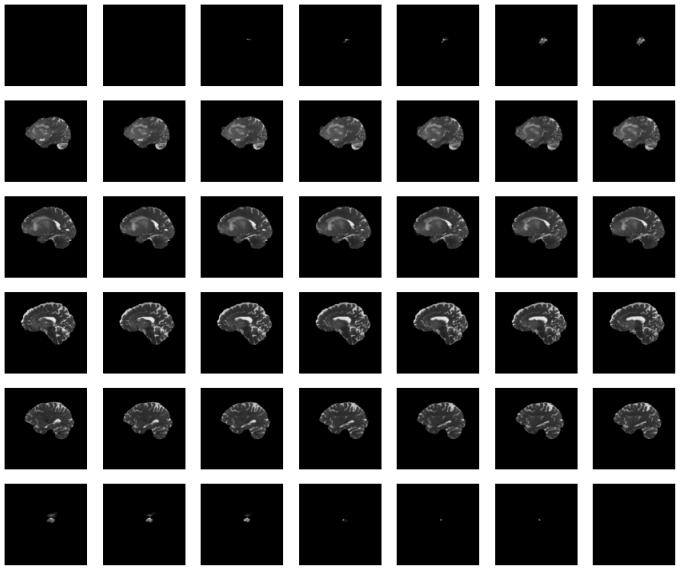


Fig. 2. A sequence of T2w scans for patient '00002' in the dataset

relaxation, and offers a thorough description of the lesion from various angles.

Each of the 585 annotated samples in the RSNA-MICCAI dataset describes a patient in terms of patient ID. Each case has a folder and a five-digit number (patient ID) used to identify it. Each of these "case" files has four sub-folders, one for each of the structural mpMRI (multi-parametric MRI) scans in DICOM format. Each of the four modalities have samples that range in number from a few tens to a few hundred. A single modality scan of a patient consists of a series of MRI scans. For patient "00002" in the dataset, Figure 2 shows an MRI sequence of T2w scans. There is a significant amount of data available for pattern learning given the average number of scans per patient, which ranges from 170 to 200.

B. Image Processing

Deep neural networks require a lot of training data to perform well and image augmentation is typically needed to improve the performance of the networks in order to create a robust classifier with very little training data. Through various processing techniques such as random rotation, shifts and flips, etc., image augmentation artificially generates training images from the 2D DICOM images of different modality available to us.

We cycle through each image performing zooming, rotation, horizontal flip, etc., with random weights to bring about variation in our training, validation and testing data sets. We rescale and resize the whole image to a dimension of 300x300. We apply a zoom to the image by a value between 0.8 – 1.2. We rotate the image by a value between 0.8 – 1.2. The image is shifted along the vertical axis and horizontal axis by a random value. We either flip or choose not to flip the image based on a random value. We change the brightness of the image by a factor between 0.8 - 1.2. While all these augmentation are applied to the training data, the validation and testing scans are only rescaled to a value between [0,1] and resized to 300x300.

C. Model Selection

While several architectures of CNNs have been explored for similar works in the related fields as stated in Sec. II, we tried our implementation on a few models that we found to be computationally efficient in terms of the depth and size of the model, its performance in terms of accuracy and AUC and also considering the amount of data available to us for transfer learning. Taking the reported accuracy of the models into consideration we tested AlexNet and MobileNet architectures on the available MRI scans to take advantage of faster training, lightweight framework and easily re-trainable weights, but found that the performance of these models even with a small learning rate of 1e-4 did not cross 60%. Since the work in [9] extensively used ResNet architectures for a similar objective, we explored the performance of different ResNet frameworks on our dataset but found that their accuracy fluctuated around 56%. Table. 1 provides a performance review of the different architectures explored for prediction of the methylation status using the RSNA-MICCAI dataset.

As seen from the performance metrics in table.1, transfer learning and fine-tuning of the EfficientNetB0 model gave the best performance.

TABLE I. PERFORMANCE METRICS

Metrics	Model					
	Alex Net	***Mobile Net***	***ResNet10***	***ResNet50***	***Efficient NetB0***	***Efficient NetB3***
Learning Rate	1e-4	1e-4	1e-4	1e-4	1e-4	1e-4
Accuracy (in %)	55.68	52	57	56.81	70	66.92
Precision	0.55	0.52	0.50	0.56	0.71	0.58
Recall	0.55	0.52	0.77	0.56	0.70	0.56
F1 score	0.54	0.51	0.61	0.55	0.69	0.55

During ConvNet scaling it was stated that balancing all three dimensions—network depth, width, and resolution—equally is crucial for getting improved performance. In contrast to conventional practice, EfficientNet is a convolutional neural network architecture that equally scales all dimensions while also being lightweight and thus, faster with better accuracies compared to the previously explored models as reported by its performance on ImageNet and thus being the SOTA CNN for computer vision applications. Since EfficientNets also transfer well we compared the performance of two architectures namely EfficientNetB0 and EfficientNetB3 and found that even though EfficientNetB3 has higher number of sub-blocks and hence larger number of parameters learnt, it quickly overfits the data thereby limiting the achieved performance. It is also harder to tune the hyperparameters for a larger variant when the amount of data available is limited. Our experimentation revealed that EfficientNetB0 is the CNN model that best utilizes the benefits of EfficientNets while simultaneously minimizing overfitting by lowering the amount of parameters learned and henceforth all our implementation considers this architecture.

IV. Proposed Model

All the implementations were performed in python using the Keras library that runs on top of TensorFlow and utilized the pre-trained EfficientNetB0 model.

The dataset provided consisted of four modality scans of 585 patients which was split into two sets: 90% for training and validation, 10% for testing. The training and validation sets were then obtained using an 85% stratified split to ensure that both the classes were equally represented in both sets. In order to artificially increase the amount of data that was accessible and train the model to be robust, augmentation was carried out as previously mentioned, following which the CNN model is trained and fine-tuned. Figure 3 represents a flowchart of the learning framework employed during our implementation.

According to our implementations, the optimum method for improving classification accuracy on the available data was found to be transfer learning followed by fine-tuning, where features learnt on one problem are used for a different but related problem. It is typically applied for instances where the dataset is insufficient to fully train a model from scratch. We trained individual models for all the four modalities of scans that were available by only considering 20% of the slices around the midsection of the MRI sequence since they contained images of higher intensity and thus provided more expressive patterns for the model to train on.

The transfer learning workflow we used for the four models, each trained using a different modality, is as shown below :

1. Load the EfficientNetB0 model that has already been trained and initialize it with its weights.
2. Freeze the layers of this model to prevent losing any of the data they contain during subsequent training sessions.
3. On top of the frozen layers, add some new, trainable layers. On the new dataset, these new layers will discover how to perform classification using the previous features.
4. Train the new layers with the data available to us for the prediction task.

The transfer-learning phase consists of the following layers in their respective order:

1. Global Average Pooling : used for better representation of the vector as it uses a parser window which moves across the object and pools the data by averaging it. It accepts a 4D tensor and operates the mean on the height and width dimensionalities resulting in 2D tensor as output.
2. Batch Normalization : standardize the input to each layer and stabilize the learning process.
3. Dropout with a rate of 0.4 : regularization technique used to prevent overfitting in the neural network by ignoring certain instances during training.
4. Dense with 32 units and ReLU activation function
5. Batch Normalization and Dropout with a rate of 0.4
6. Dense with 1 unit and Sigmoid activation function as the output layer

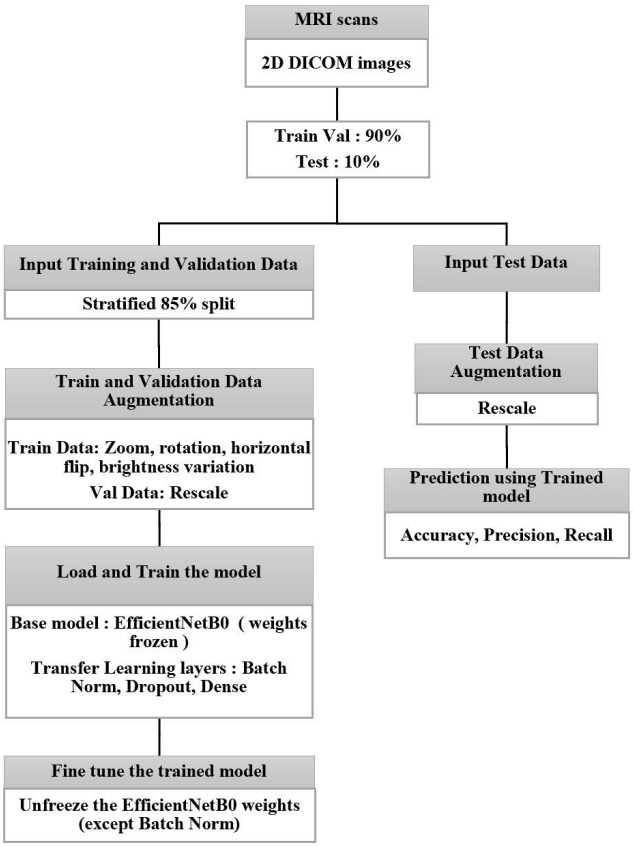
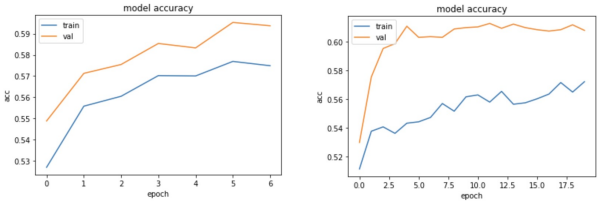


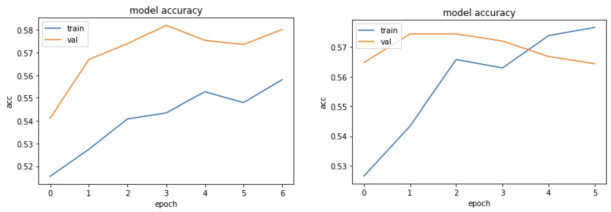
Fig.3. Learning framework

We employed the Adam optimizer, a computationally effective stochastic gradient descent method that works well for issues involving large amounts of data or parameters with a learning rate of 1e-4. Our implementation ran for 20 epochs and used the early stopping, reduce learning rate on plateau and model checkpoint as callbacks that monitored the validation loss. Since we framed it as a binary classification problem that predicts if the status of the MGMT promoter enzyme is methylated or unmethylated, we used the Binary Cross Entropy loss function that computes the cross-entropy loss between true labels and predicted labels and Binary Accuracy and AUC (Area under the curve) as metrics for monitoring the progress and measuring the performance of the model.

The trained models, once they reached convergence on the MRI data we had, were then fine-tuned to increase classification accuracy for 10 epochs with Binary Cross entropy as the loss function and accuracy as the metrics. This involved unfreezing all of the model weights generated above (except for Batch Normalization Layer) and retraining the model on the available data at a very slow learning rate with Adam optimizer.

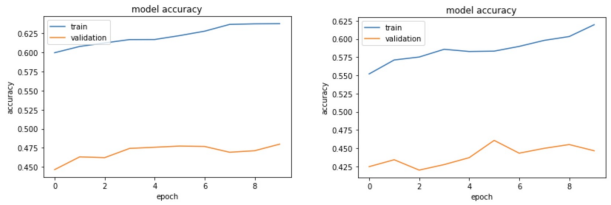


a) Flair b) T1w

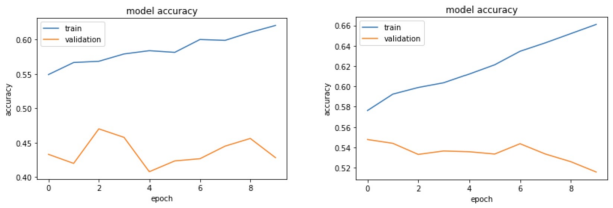


c) T1wce d) T2w

Fig.4. Accuracy of the 4 types of MRI scans before fine tuning



a) Flair b) T1w



c) T1wce d) T2w

Fig.5. Accuracy of the 4 types of MRI scans after fine tuning

V. RESULT

The training curves of the models monitored on accuracy of classification, before fine tuning, evaluated on the 4 modalities of MRI scans is demonstrated in Figure 4 (a-d). The performance metrics used for evaluating transfer-learning were classification accuracy and AUC. To determine the effectiveness of chemotherapy, we must accurately determine the status of MGMT promoter methylation and hence accuracy of the classification is the key metric of RSNA-MICCAI. In the current situation, the model should enhance the categorization of positive samples while simultaneously lowering the risk of unwanted surgery and post-operative discomfort due to misclassification.

It can be observed from Fig. 4 (a, c, d) that early stopping occurs as the model failed to improve in its performance over the validation dataset due to overfitting which is mostly caused by the nature of the dataset as several other studies in Sec. II too obtained a similar result. The results obtained based on the implementation proposed in Sec. 1V is mentioned in the table. II for all the four modalities.

Since we used pre-trained weights from a completely different dataset to initialize our EfficientNet model, fine-tuning becomes essential as the entire model is now trained to extract features from the MRI images that are available to us. As seen from the table. II , performance of the proposed architecture significantly improved after fine-tuning the model for 10 epochs with the batch size reduced from 128 to 64.

TABLE II. MODEL PERFORMANCE

Modality	Transfer-learning alone		Transfer-learning + Fine-tuning
	Accuracy (%)	AUC	Accuracy (%)
FLAIR	52.40	0.5077	64.20
T1-weighted	48.61	0.4443	71.92
T1wCE	46.73	0.4945	53.6
T2-weighted	53.97	49.45	61.51

The performance of the four models, trained on the different modalities of MRI scans while fine-tuning it is demonstrated in Figure 5 (a-d).

By gradually adjusting the pretrained features to the available data, fine-tuning generated significant improvements and contributed to the 70% accuracy of T1w modality scans while the accuracy of the FLAIR and T2w fluctuated around 65%. Fine-tuning significantly improved the separability of the two classes as can be seen from Figure 6 that provides a distribution plot of the model's prediction before and after fine-tuning for the FLAIR imaging modality. A similar trend was observed for models trained on T1w, T1wCE and T2w MRI scans as well.

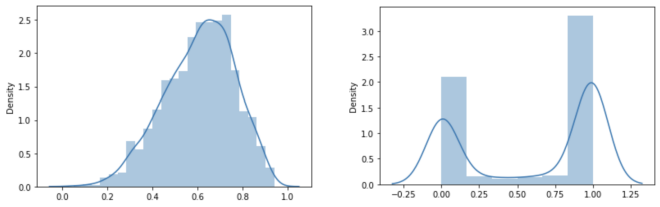


Fig.6. Distribution plot of the model's prediction before and after fine-tuning

Fig. 7 shows the classification report obtained for the model that has been fine-tuned for the T1w modality which has the best accuracy over all the other types examined.

	precision	recall	f1-score	support
0.0	0.67	0.81	0.73	27
1.0	0.75	0.58	0.65	26
accuracy			0.70	53
macro avg	0.71	0.70	0.69	53
weighted avg	0.71	0.70	0.69	53

Fig.7. Accuracy, Precision, Recall and F1-Score for the best model obtained on T1w modality

VI. Conclusion

In an attempt at trying to determine if there exists a relationship between the methylation status of MGMT promoter enzyme and MRI scans, we developed a novel classification architecture in this work that raised prediction accuracy on the RSNA-MICCAI dataset to around 70 percent by considering each modality of the scan separately. We compared several CNN architectures with accuracy as the metric and found EfficientNet to be the best model in terms of model complexity and weights retraining capability. We demonstrated how transfer learning followed by fine-tuning significantly improves the accuracy of classification by considering EfficientNetB0 as the base model followed by a few layers of Batch Normalisation, Dense and Dropout with Adam optimizer..

Although the suggested architecture yields an encouraging result, the majority of studies using the RSNA-MICCAI dataset were unable to achieve more than 80% accuracy, so we cannot rely on the current deep learning techniques that solely use MRI scans for accurate prediction of the methylation status.

This study is subject to the following constraints, which also provide future developments. To begin with, our suggested strategy ignores the relationship between data from various modality and only takes into account the temporal correlation between data from the same modality. While individual models trained on each modality separately, achieve around 65-70% accuracy, this can further be improved by multi-modal feature aggregation which looks into the intermodality correlation more thoroughly.

Use of more complex models that can further investigate the patterns from multi-modal image features can also be explored but this requires a large amount of data to prevent overfitting of the bigger models.

VII. References

[1] G. Raut, A. Raut, J. Bhagade, J. Bhagade and S. Gavhane, "Deep Learning Approach for Brain Tumor Detection and Segmentation," *2020 International Conference on Convergence to Digital World - Quo Vadis (ICCDW)*, 2020, pp. 1-5, doi: 10.1109/ICCDW45521.2020.9318681.

[2] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network", *2017 International Conference on Engineering and Technology (ICET)*, pp. 1-6, 2017.

[3] Korfiatis, P., Kline, T.L., Lachance, D.H. et al. "Residual Deep Convolutional Neural Network Predicts MGMT Methylation Status". *J Digit Imaging 30*, 622–628 (2017).

[4] R. Qu and Z. Xiao, "An Attentive Multi-Modal CNN for Brain Tumor Radiogenomic Classification," *Information*, vol. 13, no. 3, p. 124, Mar. 2022, doi: 10.3390/info13030124.

[5] D. R. Nayak, N. Padhy, P. K. Mallick, M. Zymbler, and S. Kumar, "Brain Tumor Classification Using Dense Efficient-Net," *Axioms*, vol. 11, no. 1, p. 34, Jan. 2022, doi: 10.3390/axioms11010034.

[6] M. M. Badža and M. Č. Barjaktarović, "Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network," *Applied Sciences*, vol. 10, no. 6, p. 1999, Mar. 2020, doi: 10.3390/app10061999.

[7] Korfiatis, P., Kline, T.L., Lachance, D.H. *et al.* "Residual Deep Convolutional Neural Network Predicts MGMT Methylation Status". *J Digit Imaging* 30, 622–628 (2017).

[8] Han, L. & Kamdar, M. R. "MRI to MGMT: Predicting methylation status in glioblastoma patients using convolutional recurrent neural networks". in *Pacific Symposium on Biocomputing 2018: Proceedings of the Pacific Symposium*, 331–342 (2018).

[9] Kanas, V.G.; Zacharaki, E.I.; Thomas, G.A.; Zinn, P.O.; Megalooikonomou, V.; Colen, "R.R. Learning MRI-based classification models for MGMT methylation status prediction in glioblastoma". *Comput. Methods Programs Biomed.* 2017, *140*, 249–257

[10] U.Baid, et al., "The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification", *arXiv*:2107.02314, 2021A.