Brain tumor classification on MRI in light of molecular markers

Jun Liu[1*], Geng Yuan[2], Hongyi Huang[3], Wenbin Zhang[4], Xue Lin[2], XiaoLin Xu[2], and Yanzhi Wang[2]

[1] Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
[2] Department of Electrical & Computer Engineering, College of Engineering, Northeastern University, Boston, MA, USA
[3] Servicing intelligence Machine Learning Department, Capital One Bank, DC, USA
[4] University of Maryland, Baltimore County, MD, USA

Abstract. In research findings, co-deletion of the 1p/19q gene is associated with clinical outcomes in low-grade gliomas. The ability to predict 1p19q status is critical for treatment planning and patient follow-up. The aim of this study is to utilize a specially MRI-based convolutional neural network for brain cancers detection. Although public networks such as RestNet, AlexNet can effectively diagnose brain cancers using transfer learning, however, the model includes quite a few weights that have nothing to do with medical images. As a result, the diagnostic results are unreliable by transfer learning model. To deal with the problem of trustworthiness, we create the model from the ground up, rather than depending on a pre-trained model. To enable flexibility, we combined convolution stacking with a dropout and full connect operation, it improved performance by reducing overfitting. During model training, we also supplement the given dataset and inject Gaussian noise. We use three–fold cross-validation to train the best selection model. Comparing InceptionV3, VGG16, and MobileNetV2 fine-tuned with pre-trained models, our model produces better results. On an validation set of 125 codeletion vs. 31 not codeletion images, the proposed network achieves 96.37% percent F1-score, 97.46% percent precision, and 96.34% percent recall when classifying 1p/19q codeletion and not codeletion images.

Keywords: low-grade gliomas, 1p/19q, imbalance, reliability, transfer learning, CNNs

1 Introduction

Tumors of the brain are malignant cell growths or aggregates in or around the brain.Depending on where they exist in the brain, low-grade gliomas[1–3] can manifest themselves in a variety of ways. The patient may have weakness or numbness in the right leg if the brain cancer grows in the area of the brain that controls it[1].

The most efficient method to detect brain tumors is MRI. Scanning generates a vast amount of magnetic resonance images, which is examined by a radiologist.Biomarker detection helps give patients the best appropriate treatment for their particular condition. This study is notable for the uniqueness and promising findings of merging deep learning with radiogenomics. Deep learning was better at detecting 1p/19q co-deletion on T2 images than it does on post-T1 contrast images.

Akkus et al. [2] were the first to use a deep learning approach to predict 1p19q from low-grade glioma MRI images in 2017. Chelghoum et al., 2020[4] employs transfer learning to classify 1p19q using popular pre-trained models such as AlexNet, VGG19, GoogleNet, and others.[5–7] They point out that transfer learning can still provide correct results even with limited datasets. To get the best accuracy, Abiwinanda et al. [8] create the network by combining various CNN operations. Maithra Raghu et al. [9] discovered that transfer learning had little impact on imaging tasks in medicine, with models trained from scratch performing nearly as well as ImageNet-transferred models.

Our contributions are listed below.:

- Using convolution stack, we develop a CNN specifically for detecting brain cancer in MRI images..
- To avoid overfitting and improve performance, we utilize a tunable composition of dropout and Gaussian noise during training.
- Comprehensive evaluation of discriminant results using the confusion matrix, F1-score, precision, and recall methods to evaluate unbalanced data in order to avoid false positives.

2 Materials and Methods

To train our proposed network, we use the Kaggle public dataset[10]. Meanwhile, on this dataset, we compare the results of VGG16, InceptionResNetV2, and MobileNetV2 all fine-tuned pre-trained models applying transfer learning.

2.1 Experimental Data

For evaluation and research, Kaggle Public Datasets offers brain MRI datasets. There are 253 brain Magnetic Resonance images in the collection, divided into two folders: Yes and No. The folder Yes contains 155 scans of tumors in the brain, while the folder No has 98 non-tumor brain MRI scans.

A brain with a tumor is on the left, while a healthy brain is on the right in Fig. 1.

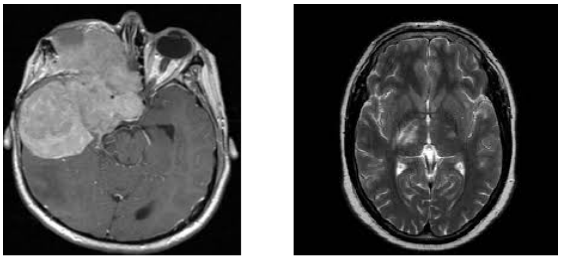

Fig. 1. Brain Magnetic Resonance Image

2.2 Network Composition

The model seen in Fig. 2 has 14 layers. Kernels with convolutions (3 x 3) had positive results, because the small convolutions catch some of the finer details of the edges. It starts with 16 kernels in first two CNN layers and gradually progresses to 32, 64, and lastly 128 kernels per layer.

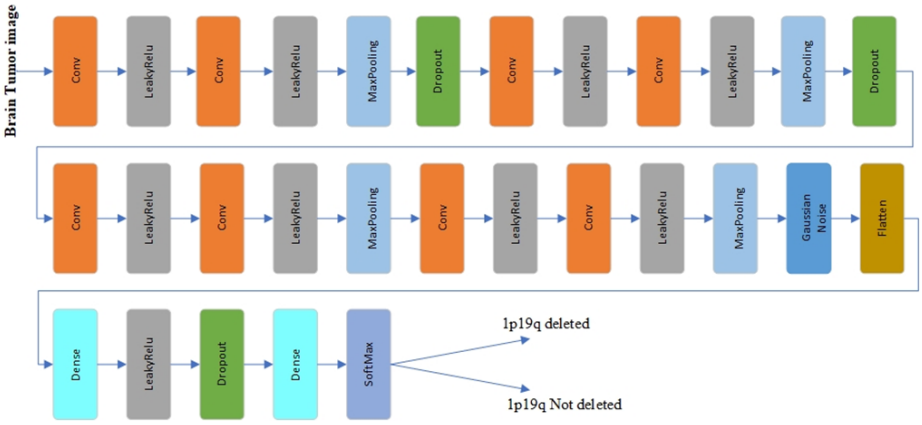

Fig. 2. Network Architectures

Convolution layers, pooling layers, LeakyRelu layer, Softmax layer, dropout layers, Dense layers[11], Flatten layer, make up the network represented in Fig.2 .

Table 1.

Layer	Kernels	Kernel Size	Stride	Feature Map Size	Activation
InputLayer	-	-	-	256 x 256	-
Convolution	16	3 x 3	1	254 x 254 x 16	LeakyReLU
Convolution	16	3 x 3	1	252 x 252 x 16	LeakyReLU
MaxPooling	-	2 x 2	1	126 x 126 x 16	-
Dropout	-	-	-	126 x 126 x 16	-
Convolution	32	3 x 3	1	124 x 124 x 32	LeakyReLU
Convolution	32	3 x 3	1	122 x 122 x 32	LeakyReLU
MaxPooling	-	2 x 2	1	61 x 61 x 32	-
Dropout	-	-	-	61 x 61 x 32	-
Convolution	64	3 x 3	1	59 x 59 x 64	LeakyReLU
Convolution	64	3 x 3	1	57 x 57 x 64	LeakyReLU
MaxPooling	-	2 x 2	1	28 x 28 x 64	-
Convolution	128	3 x 3	1	26 x 26 x 128	LeakyReLU
Convolution	128	3 x 3	1	24 x 24 x 128	LeakyReLU
MaxPooling	-	5 x 5	1	4 x 4 x 128	-
GaussianNoise	-	-	-	4 x 4 x 128	-
Flatten	-	-	-	2048	-
FullConnect	-	-	-	1024	-
Dropout	-	-	-	1024	-
FullConnect	-	-	-	2	-
Softmax	-	-	-	2	-

The composition of the network, including the Kernel Size,stride ,Feature Map Size etc, are listed in Table 1.

- All brain tumor images supplied into the network are resized to 256 by 256 pixels.
- For all convolutional layers, the network employs kernels of size 3 x 3, with 16, 32, 64, and 128 kernels in each layer in turn.
- LeakyReLU was used as the activation function since negative numbers are preserved together with concerns about saturation are eliminated when using tanh.
- This model starts with 2 x 2 maxpooling and subsequently progresses to 7 x 7 maxpooling.
- The Full connect layer is utilized twice to decrease the quantity of neurons.
- On purpose to reduce overfitting, Gaussian Noise was added in the training process and it can be thought of as a method of random data enrichment.
- Full Connected's purpose is to reduce multi-dimensional inputs to a single dimension.
- In the course of network training, some neural network units are dropped from the network with a certain probability.

2.3 Hyperparameters setting

Some hyperparameter settings were investigated during this research.

Learning rate

Since initial weight values are relatively random, starting with a higher learning rate usually works alright. As the training phase advances, the findings often get nearer either to global or local minima.

EarlyStopping

The model is prevented from overfitting the training data by using early stop approaches. The model will be stopped from training if there is no change of at least 0.001 in 8 epochs.

Batch Size

The batch size is limited by the amount of RAM available. Moreover, although a larger batch size helps for weights update less frequently, faster training, which may result in less effective results.

Count of Epochs

The number of epochs indicates how many times the entire training data is iterated over when training the model.

2.4 Experiments

We develop with Keras within the TensorFlow framework in this experiment.

Data preprocessing The data preprocessing steps are the removal of the third class, normalization of the data, and reshuffling of the training data. Since there is no enough image to train the model with this small dataset, so data augmentation aids in addressing the data imbalance issue.

Proposed training procedure In order to improve performance, we develop, evaluate, and train our model, which is represented in Fig. 3 using cross-validation in model training.

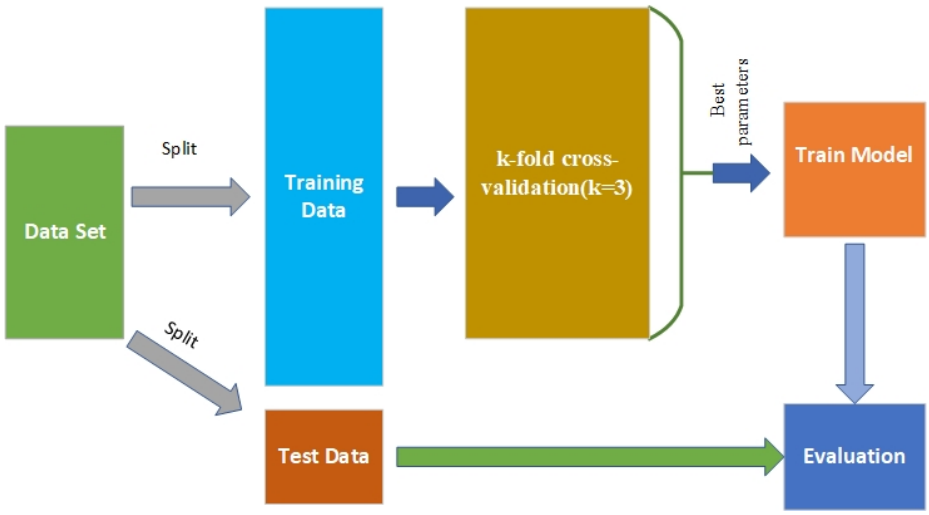

Fig. 3. Proposed training procedure

- The data set is separated into train and test sets at stochastic, with the model was build using the train set, while the test set was utilized to assess its accuracy.
- The program fine-tunes the model by cross-validation with k-fold [12] to obtain the finest quality model.
- Assess the model's anticipated accuracy on the test set.

Evaluation method To evaluate our designed model, we use the confusion matrix, accuracy, precision, recall, and F1.

Confusion Matrix

A method for analyzing the performance of classification algorithms is called the confusion matrix. If the dataset have an unbalanced amount of observations in each class or if the dataset has more than two classes, it would be incorrect to use classification accuracy alone as a measurement.

Table 2. Confusion Matrix

	1p/19q deleted	1p/19q not deleted
1p19q deleted	True Positive	False Positive
1p19q not deleted	False Negative	True Negative

Table 2 explicitly displays the proportion of accurate and wrong identifications for each category.

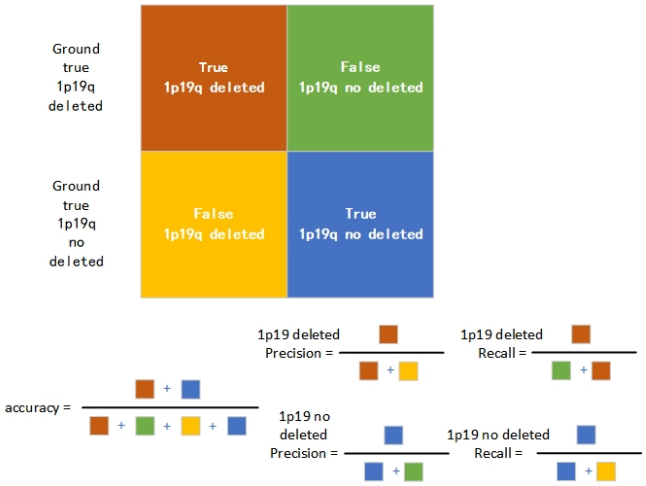

Fig. 4. Evaluation matrix

From Fig. 4 most metrics that can be derived from the confusion matrix.

Precision

Precision refers to the percentage of 1p/19q deleted sufferers correctly predicted by the model terms of number of sufferers with 1p/19q deleted.

Precision = True 1p/19q deleted / (True 1p/19q deleted + False 1p/19q deleted)

Recall

The percentage of 1p/19q deleted sufferers detected by the model to all 1p/19q deleted sufferers is used to calculate specificity.

Recall = True 1p/19q deleted / (True 1p/19q deleted + False 1p/19q non-deleted)

F1 score

The F1 score was established to work successfully with unbalanced data because of a disparity in the percentage of brain and non-brain malignancies. in this dataset. Its formula is as follows:

F1score = 2 * (recall * precision) / (recall + precision)

3 Results

The suggested network achieves a f1-score of 0.9637 in Table 3 using 164 images, including 125 1p19q deleted and 39 1p19q undeleted.

Table 3. Result

type	Precision	Recall	F1-score	Images
1p/19q deleted	0.9881	0.9635	0.9742	125
1p/19q not deleted	1	0.9644	0.9375	31
avg / total	0.9746	0.9634	0.9637	156

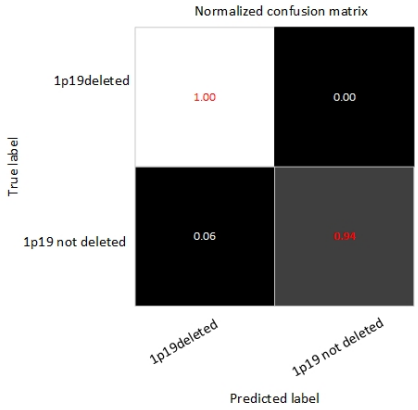

Fig. 5. Confusion matrix

On the test set, Fig. 5 illustrates the 1p19q values of the confusion matrix. We can observe that all 125 1p19q deleted images have been correctly identified.

To compare with our model, we used transfer learning based models such InceptionResNetV2, MobileNetV2, and VGG16.

From Table 4 We discovered that our model is very well balanced regarding Precision, Recall and F1 score, these values are very close. As the comprehensive indicator of Precision and Recall, the F1 score can see the model's performance in recognition more clearly. In terms of performance,our model exceeds all other models in terms of F1 score.

4 Discussion

We present a viable CNN-based technique for predicting deletions of the 1p/19q chromosomal arm. In the procedure of applying deep learning algorithms to medical imaging, having a sufficient quantity of datasets is a big challenge. For medical diagnosis, current transfer learning methods leverage a variety of publically available models trained on big public ImageNet datasets. These models, on the other hand, generate a significant quantity of medical images, endangering the accuracy of clinical diagnosis[16].

Table 4. Baseline Comparisons

Model	Precision	Recall	F1 score
We proposed	0.9746	0.9634	0.9637
InceptionV3[13]	0.9230	1	0.9600
MobileNetV2[14]	1.0	0.8709	0.9200
InceptionResNetV2[14]	0.9062	0.9354	0.9153
VGG16[15]	0.8960	0.9286	0.9123

Transfer learning is not used in our brain tumor detection model, and the parameters it creates are all based on medical imaging datasets, ensuring complete accuracy in the brain tumor detection.

Small dataset can result in substantial training errors.In order to solve this problem, the training phase, Gaussian noise is included to improve the model's normalizing ability and fault - tolerant. Adding noise means that the network cannot remember training samples because they are changing all the time, as a result of which network weights are reduced and the network becomes more robust with lower generalization error. Because fresh samples are drawn from the field proximity to existing samples, The input space's shape has been smoothed out. It make easier for the network to learn the mapping function, resulting in better and faster learning.

5 Conclusion

In this paper, we use convolution stacking in conjunction with Gaussian noise and random drop in training to create a model for identifying small dataset of brain tumors that outperforms the transfer learning method while having the advantages of small model size and high reliability.Cross-validation in training is an excellent way to overcome the problem of data imbalance.

References

1. Amin, J., Sharif, M., Haldorai, A. et al. Brain tumor detection and classification using machine learning: a comprehensive survey. Complex Intell. Syst. (2021). https://doi.org/10.1007/s40747-021-00563-y
2. Akkus, Z., Ali, I., Sedlář, J. et al. Predicting Deletion of Chromosomal Arms 1p/19q in Low-Grade Gliomas from MR Images Using Machine Intelligence. J Digit Imaging 30, 469-476 (2017). https://doi.org/10.1007/s10278-017-9984-3
3. Bhattacharya, D., Sinha, N. & Saini, J. Determining chromosomal arms 1p/19q co-deletion status in low graded glioma by cross correlation-periodogram pattern analysis. Sci Rep 11, 23866 (2021). https://doi.org/10.1038/s41598-021-03078-1
4. Lombardi G, Barresi V, Castellano A, Tabouret E, Pasqualetti F, Salvalaggio A, Cerretti G, Caccese M, Padovan M, Zagonel V, Ius T. Clinical Management of

Diffuse Low-Grade Gliomas. Cancers (Basel). (2020) Oct 16;12(10):3008. doi:10.3390/cancers12103008.
5. Alanazi MF, Ali MU, Hussain Chelghoum R., Ikhlef A., Hameurlaine A., Jacquir S.Transfer learning using convolutional neural network architectures for brain tumor classification from MRI images (2020), dio:10.1007/978-3-030-49161-1_17
6. SJ, Zafar A, Mohatram M, Irfan M, AlRuwaili R, Alruwaili M, Ali NH, Albarrak AM. Brain Tumor/Mass Classification Framework Using Magnetic-Resonance-Imaging-Based Isolated and Developed Transfer Deep-Learning Model. Sensors (Basel). 2022 Jan 4;22(1):372. doi:10.3390/s22010372.PMID:35009911;PMCID:PMC8749789.
7. Swati ZNK, Zhao Q, Kabir M, Ali F, Ali Z, Ahmed S, Lu J. Brain tumor classification for MR images using transfer learning and fine-tuning. Comput Med Imaging Graph. 2019 Jul;75:34-46. doi:10.1016/j.compmedimag.2019.05.001.
8. Abiwinanda N., Hanif M., Hesaputra S.T., Handayani A., Mengko T.R.Brain tumor classification using convolutional neural network World congress on medical physics and biomedical engineering 2018, Springer, Singapore (2019), pp. 183-189
9. Raghu, M.; Zhang, C.; Kleinberg, J.; Bengio, S. Transfusion: Understanding Transfer Learning for Medical Imaging, NeurIPS 2019. Available online: https://arxiv.org/abs/1902.07208 Accessed on 19 Feb 2022.
10. Brain MRI Images for Brain Tumor Detection, https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection. Accessed on 26 March 2022
11. Keras API reference / Layers API / Core layers / Dense layer, https://keras.io/api/layers/core_layers/dense/.Accessed on 26 Feb 2022
12. Ashalata Panigrahi, Manas R. Patra, Chapter 6 - Network Intrusion Detection Model Based on Fuzzy-Rough Classifiers,Editor(s): Pijush Samui, Sanjiban Sekhar, Valentina E. Balas,Handbook of Neural Computation,Academic Press,2017,Pages 109-125,ISBN 9780128113189, https://doi.org/10.1016/B978-0-12-811318-9.00006-5.
13. Mohd. Farhan Israk Soumik, Brain Tumor Detection InceptionV3. https://www.kaggle.com/code/mdfarhanisraksoumik/brain-tumor-detection-inceptionv3-auc-99-84.Accessed on 26 Feb 2022
14. Farid Taghiyev Brain Tumor Detection using TensorFlow 2.x. https://www.kaggle.com/code/faridtaghiyev/brain-tumor-detection-using-tensorflow-2-x/notebook. Accessed on 26 Feb 2022
15. Kuraparthi, S., Reddy, M.K., Sujatha, C.N., Valiveti, H., Duggineni, C., Kollati, M., Kora, P., V, S. (2021). Brain tumor classification of MRI images using deep convolutional neural network. Traitement du Signal, Vol. 38, No. 4, pp. 1171-1179. https://doi.org/10.18280/ts.380428
16. Liu, J.; Deng, F.; Yuan, G.; Lin, X.; Song, H.; Wang, Y. An Explainable Convolutional Neural Networks for Automatic Segmentation of the Left Ventricle in Cardiac MRI. In Proceedings of the CECNet 2021, Beijing, China, 18-21 November 2021. https://doi.org/10.3233/FAIA210416.