2022

Uncertainty Estimation in Classification of MGNT Using Radiogenomics for Glioblastoma Patients

W. Farzana
Old Dominion University, wfarz001@odu.edu

Z. A. Shboul
Old Dominion University

A. Temtam
Old Dominion University, atemtam@odu.edu

K. M. Iftekharuddin
Old Dominion University, kiftekha@odu.edu

Original Publication Citation

Farzana, W., Shboul, Z. A., Temtam, A., & Iftekharuddin, K. M. (2022). Uncertainty estimation in classification of MGMT using radiogenomics for glioblastoma patients. In K. Drukker, K.M. Iftekharuddin, H. Lu, M.A. Mazurowski, C. Muramatsu, R.K. Samala (Eds.), *Proceedings of SPIE, Medical Imaging 2022: Computer-Aided Diagnosis* 120331E (1-7). Society of Photo-Optical Instrumentation Engineers. https://doi.org/10.1117/12.2612621

Uncertainty Estimation in Classification of MGMT Using Radiogenomics for Glioblastoma Patients

W. Farzana, Z. A. Shboul, A. Temtam, and K. M. Iftekharuddin

Vision Lab in Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, VA23529

ABSTRACT

Glioblastoma Multiforme (GBM) is one of the most malignant brain tumors among all high-grade brain cancers. Temozolomide (TMZ) is the first-line chemotherapeutic regimen for glioblastoma patients. The methylation status of the O6-methylguanine-DNA-methyltransferase (MGMT) gene is a prognostic biomarker for tumor sensitivity to TMZ chemotherapy. However, the standardized procedure for assessing the methylation status of MGMT is an invasive surgical biopsy, and accuracy is susceptible to resection sample and heterogeneity of the tumor. Recently, radio-genomics which associates radiological image phenotype with genetic or molecular mutations has shown promise in the non-invasive assessment of radiotherapeutic treatment. This study proposes a machine-learning framework for MGMT classification with uncertainty analysis utilizing imaging features extracted from multimodal magnetic resonance imaging (mMRI). The imaging features include conventional texture, volumetric, and sophisticated fractal, and multi-resolution fractal texture features. The proposed method is evaluated with publicly available BraTS-TCIA-GBM pre-operative scans and TCGA datasets with 114 patients. The experiment with 10-fold cross-validation suggests that the fractal and multi-resolution fractal texture features offer an improved prediction of MGMT status. The uncertainty analysis using an ensemble of Stochastic Gradient Langevin Boosting models along with multi-resolution fractal features offers an accuracy of 71.74% and area under the curve of 0.76. Finally, analysis shows that our proposed method with uncertainty analysis offers improved predictive performance when compared with different well-known methods in the literature.

Keywords: Radiogenomics, Uncertainity, Classification, MGMT, Glioblastoma.

1. INTRODUCTION

Glioblastoma (GBM) is an aggressive grade IV astrocytoma known to be the most malignant brain tumor, which involves a complex treatment regimen.[1] Temozolomide (TMZ) is the first-line chemotherapeutic treatment for GBM patients. The chemotherapeutic treatment produces cytotoxicity and apoptotic effect in cancer cells by inducing methylation at the O6-methylguanine methyltransferase (MGMT) site of DNA.[2] Therefore, the methylation status of MGMT has substantial prognostic significance and can impact GBM treatment. The standardized approach of assessing genomic information in GBM is tissue sampling. Tissue sampling requires an invasive surgical resection that could be imprecise due to the heterogeneity of the tumor.[3] However, several studies[4–6] have shown that genetic alterations may be associated with phenotypic changes and can be detected by mMRI texture features. The authors[5,7] have discussed an association between radiomics features such as gray-level co-occurrence matrix (GLCM), shape, area, volume, and intensity-relevant features for the prediction of MGMT status. GLCM attributes might be useful in identifying grey-level spatial variance within the image but may not be beneficial in evaluating the randomized surface structure variability of aberrant tumor tissues in MRI.[6] Multiresolution wavelet features, on the other hand, represent texture variation in tumor tissue across independent image resolution.[8] On the other hand, multi-resolution fractal features mathematically integrate conventional textural analysis and multi-resolution texture analysis. Multi-resolution fractal texture features reflect the randomized and complicated pattern of the tumor tissue at multiple scales. The work[9–11] has demonstrated the efficacy of multi-resolution fractal features for characterization, segmentation, and categorization of

Further author information: (Send correspondence to K. M. Iftekharuddin)
K. M. Iftekharuddin.: E-mail: kiftekha@odu.edu
W. Farzana.: E-mail: wfarz001@odu.edu

Medical Imaging 2022: Computer-Aided Diagnosis, edited by Karen Drukker, Khan M. Iftekharuddin, Hongbing Lu, Maciej A. Mazurowski, Chisako Muramatsu, Ravi K. Samala, Proc. of SPIE
Vol. 12033, 120331E · © 2022 SPIE · 1605-7422 · doi: 10.1117/12.2612621

tumor tissues in MRI. Moreover, uncertainty estimation is crucial in medical decision making. For example, the Bayesian approach has been applied in uncertainty estimation in neural network models.[12] This study hypothesizes that multi-resolution fractal texture features along with conventional features will offer improved MGMT methylation classification performance. In addition, we also incorporate uncertainty estimation with an ensemble of Stochastic Gradient Langevin Boosting (SGLB) models to further improve the prediction of MGMT status in glioblastoma patients.

2. METHODS

2.1 Dataset

In this study, we use the pre-operative scans of glioblastoma patients[13–15] to include four modalities of MRI scans: pre-contrast T1-weighted (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2) and T2 Fluid Attenuated Inversion Recovery (FLAIR). All scans are skull-stripped and co-registered. The dataset also includes segmented tumor sub-regions of GBM: Gadolinium enhancing tumor (ET), the peritumoral edema (ED), and the necrosis (NC) including non-enhancing tumor (NCR/NET). The status of the MGMT methylation of the patients is collected from the Genomic Data Commons Data Portal (TCGA). In this work, a total of 114 patients with a known MGMT methylation status are utilized. The distribution of the MGMT status is as follows: 71 patients are unmethylated and 43 patients are methylated.

2.2 Methodology

In this study, we obtain MGMT status predictive model based on multi-resolution fractal texture features along with other conventional mMRI image features. The extracted features depict the volumetric, area, histogram-based statistics, and texture characteristics of tumor and sub-regions (ET, ED, NC). The conventional texture features are extracted from grey-tone spatial dependence matrices (GTSDM), neighborhood grey-tone difference matrix (NGTDM), and gray level size zone matrix (GLZSM). The fractal texture features are PTPSA for fractal characterization,[16] multi-resolution Brownian motion mBm analysis,[17] and tumor region characterization with Holder Exponent (HE) modeling.[18] First, we divide the dataset randomly into 75% (91 cases) training and 25% (23 cases) testing set with a stratified distribution of MGMT. Around 1300 radiomics features are extracted from mMRI sequences of tumors in the training set. Then, K-best features are selected based on ANOVA F-value between target and features. Afterward, a recursive feature elimination with 5-fold cross-validation is performed by fitting Random Forest (RF) model as an estimator on feature sets where the number of features is 7, 9, 11, 13, 15. Based on the importance score which maximizes the accuracy, a feature set consists of 9 features selected from training samples. In our study, we have utilized ensembles of Stochastic Gradient Langevin Boosting (SGLB) models where SGLB itself is a cluster of trees. The selected features are utilized as an input, and the MGMT methylation status is utilized as a targeted output. The overall pipeline for MGMT prediction in GBM is presented in Figure 1.

An ensemble-based Bayesian approach is considered for estimation of uncertainty in prediction of MGMT status. Considering model parameters θ as a random variable and prior probability distribution as $p(\theta)$, the posterior probability distribution $p(\theta|D)$ according to Bayes' rule is

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} \tag{1}$$

where, $D = \{x^{(i)}, y^{(i)}\}$, defines the training dataset and $i = 1.....N$ are the number of the sample size in equation (1). The probabilistic ensemble model is represented as $\{P(y|x;\theta^{(m)}\}$,where $m = 1.....M$, are the number of models as shown in Figure 1.Stochastic Gradient Langevin Boosting (SGLB) is the proposed algorithm[19]for sampling of Gradient Boosting Decision Tree (GBDT) models from posterior distribution $p(\theta|D)$. SGLB integrates GBDT with stochastic gradient Langevin dynamics to obtain global optimum for loss function.[20]In SGLB gaussian noise is added into gradients throughout the learning process to achieve a solution space for global optimum. In this study, we utilize the Stochastic Gradient Langevin Boosting (SGLB) algorithm with different random seeds[12] to obtain each independent model (Fig.1, block 4) within the ensemble model. Each

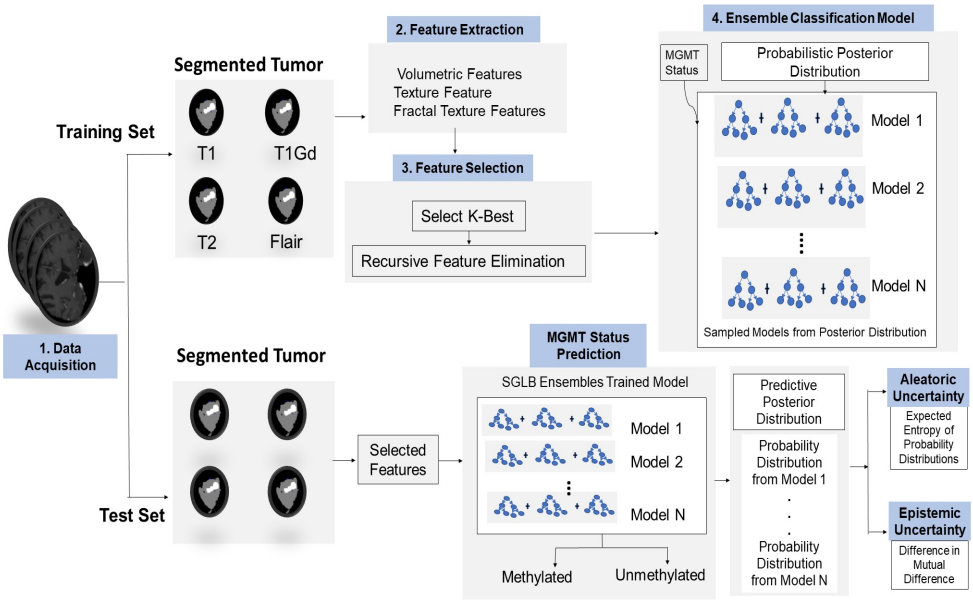


Figure 1: The overall pipeline for MGMT status prediction.

distinct set of parameters in GBDT model yields diversity in prediction. The ensemble of SGLB is performed with open-source Cat Boost Library.

The overall uncertainty in ensemble method can be divided into aleatoric (data) and epistemic (knowledge) uncertainty in a Bayesian framework which improves predictive performance[12] . Aleatoric (data) uncertainty is estimated as the entropy of the predictive distribution of each model. Epistemic (knowledge) uncertainty is estimated as dispersion or "disagreement" level of the model within the ensemble.[21]

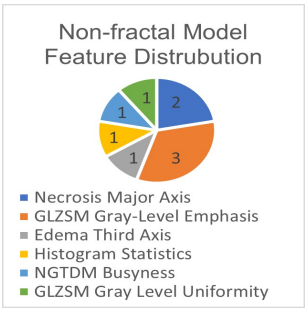
3. RESULTS

In this work, we evaluate the efficacy of multi-resolution fractal texture features in MGMT methylation classification. For evaluation of texture features, we first consider a single SGLB (Model 1 in Fig.1) classification model (fractal and non-fractal) followed by ensemble model consisting of twenty independent SGLB models for uncertainty analysis. For the fractal model, we consider fractal features with other conventional features as input to the feature selection step in Figure 1. The distribution of selected features for each model are shown in Figure 2. To avoid overfitting in model performance, 10-fold cross-validation is performed within the training dataset. The model performance is evaluated on the test dataset.

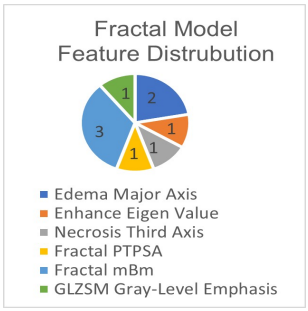
The predictive performance is improved with the inclusion of multi-resolution fractal features. In Table 1, Non-fractal model represents performance without inclusion of fractal features and Fractal model depicts inclusion of fractal features.

Table 1: Predictive Performance of Single Model on Test Dataset

Models	Training Accuracy	Test Accuracy	AUC	Precision	Recall	F1-score
Non-fractal Model	70.41%	65.22%	0.62	0.66	0.65	0.60
Fractal Model	74.80%	69.57%	0.73	0.80	0.70	0.63



(a) Non-fractal model



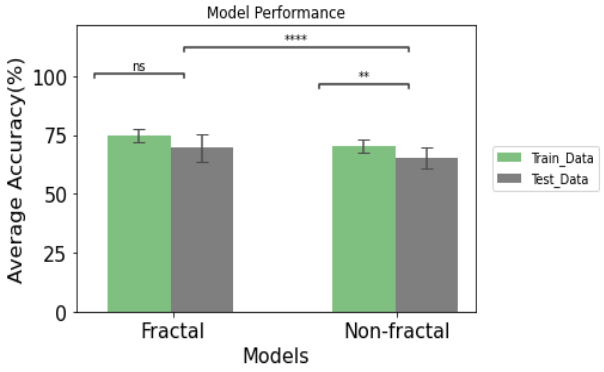
(b) Fractal model

Figure 2: Feature description for (a) Non-fractal model and (b) Fractal models. The number of selected features is denoted under each sub-category.

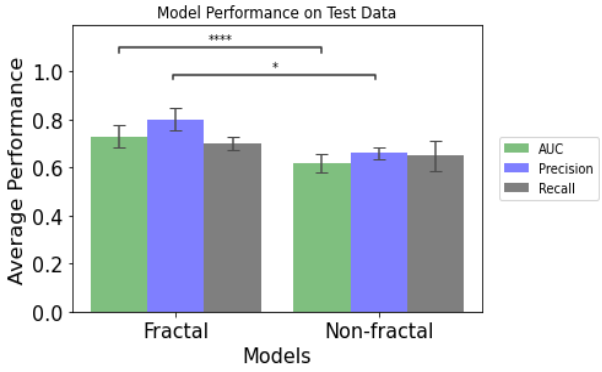
We apply an ensemble of Stochastic Gradient Langevin Boosting (SGLB) to further improve the performance of the Fractal (single) Model. The ensemble of SGLB consists of twenty independent models (different random seeds) where each model consists of 1000 trees. We compare the performance of the Fractal model, and the ensemble of models based on accuracy and area under the curve in the test data as shown in Table 2.

Table 2: Performance comparison of fractal model (single) and ensemble model (of twenty single model) on Test Dataset.

Classification Model	Accuracy (%)	Area Under Curve (AUC)
Fractal Model (Without Uncertainty)	69.57%	0.73
Ensemble of Models (With Uncertainty)	71.74%	0.76



(a) Models Performance on Train-Test Data

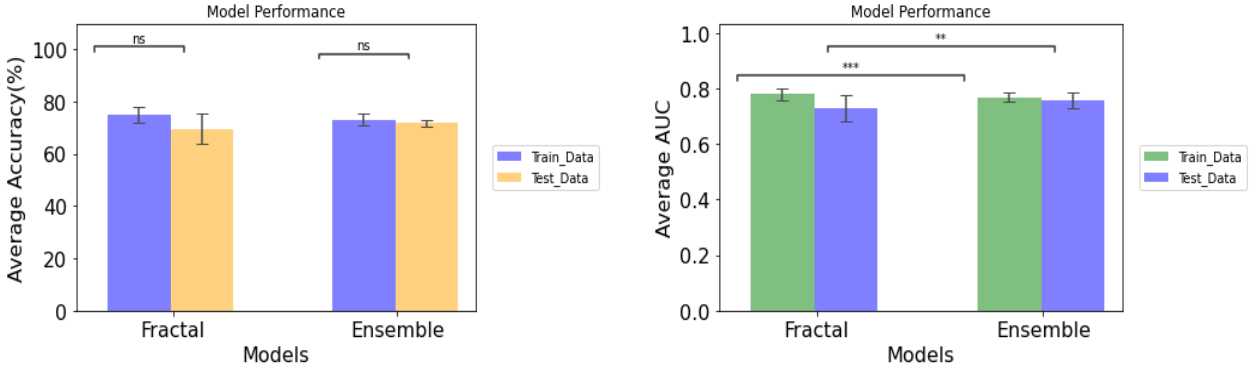


(b) Models Performance on Test Data

Figure 3: (a) Accuracy of Fractal and Non-fractal model on training and test Data. (b) Fractal and Non-fractal model performance on test data. Error bar represents one standard deviation and $*$ depicts significant difference between two group and 'ns' denotes non-significant difference between groups.

The cross-validated training accuracy and predictive accuracy improved to 74.80% and 69.57% with the inclusion of fractal features. Moreover, there is no significant difference (ANOVA test, p-value= 0.441) between the training and testing accuracy in fractal model which indicates exclusion from over-fitting. In Non-fractal model, there is notable difference (p-value=0.002) between training and testing accuracy. Moreover, model accuracy on test data, there is substantial difference (p-value=0.00005) between fractal and non-fractal model. The significant difference in AUC and Precision (p-value= 0.0000000003, 0.013 respectively) on test data is also observed

in fractal and non-fractal model.



(a) Models Accuracy on Train-Test Data.

(b) Model AUC on Train-Test Data.

Figure 4: (a) Fractal and Ensemble model accuracy on training and testing data. (b) Fractal and Ensemble model AUC on training and testing data. The error bar represents one standard deviation and $*$ depicts the significance level and 'ns' denotes non-significant.

In fractal and ensemble model, there is no significant difference (ANOVA test, p-value= 0.441 and 0.511 respectively) between cross-validated training and testing accuracy. However, there is notable difference in training AUC (p-value= 0.0002) and testing AUC (p-value=0.0016) between fractal and ensemble model. The training and testing AUC for fractal model is 0.78 and 0.73 respectively and in ensemble model the training and testing AUC is 0.77 and 0.76 respectively.

Total uncertainty is the summation of data and knowledge uncertainty. The measurement of total uncertainty and knowledge uncertainty from ensemble of models can be applied to measure misclassification or test error.[12] The error can be evaluated by Prediction-Rejection-Ratio (PRR) which ranks the uncertainty estimation values with error margins.[21] A higher value of PRR suggests that the model can recognize and reject inaccurate predictions based upon uncertainty measures.[21]

Table 3: Prediction Rejection Ratio (PRR) of single and ensemble model.

Type of Uncertainty	Fractal Model (Prediction Rejection Ratio)	Ensemble of Models (Prediction Rejection Ratio)
Total Uncertainty (TU)	8.668%	29.23%
Knowledge Uncertainty (KU)	N/A	15.47 %

Test error occurs because of noise and knowledge limitations. This explains the reason of higher PRR value when ranking with total uncertainty compared to knowledge uncertainty in Table 3. In addition, fractal model has no knowledge uncertainty. It is observed that in case of ensemble models PRR due to total uncertainty is higher than knowledge uncertainty which is consistent with the observations in other studies.[12,21]

The performance comparison of ensemble method with deep learning-based study in Table 4 shows that the ensemble-based machine learning model performs better. However, a direct comparison between these studies and our study may not be applicable because of the differences in MRI modalities, methodologies, and datasets.

The evaluation results show that the proposed Fractal model with multi-resolution fractal features improves the performance of MGMT methylation prediction. Our work further shows that incorporating ensemble modeling with uncertainty estimation offers improved classification performance.

Table 4: Performance Comparison between our method and other MGMT predictive method

Method	**Evaluation Method**	**Accuracy**	**Area Under Curve (AUC)**	**Total Number of Patients**	**MRI modality**
V. G. Kanas et al.[22]	10- fold CV	70.2 %	0.73	82	T1, T1Gd, T2
L. Han et al.[23]	15% Validation	63%	0.73	159	T1, T2, T2 FLAIR
E. Calabrese et al.[24]	10-fold CV	–	0.55	199	T1, T1Gd, T2, T2 FLAIR
Our Proposed Method **without Uncertainty**	10-fold CV	69.57%	0.73	114	T1, T1Gd, T2, T2 FLAIR
Our Proposed Method **with Uncertainty**	10-fold CV	71.74%	0.76	114	T1, T1Gd, T2, T2 FLAIR

4. CONCLUSION

The contributions of this study are in two folds. First, inclusion of multi-resolution fractal features with conventional radiomics features to augment predictive performance of MGMT methylation. Secondly, an uncertainty estimation analysis via ensemble of models is performed that further improves the model classification performance.In this paper, we show that the inclusion of multi-resolution fractal features improves the predictive performance of MGMT status in glioblastoma. The goal is to obviate invasive tissue-sampling and providing complimentary perception on genetic biomarker relevant to treatment response of GBM patients. Moreover, as uncertainty estimation is critical for machine learning models for medical decision making. We perform uncertainty estimation via ensembles of gradient boosting models to further improve the performance of the MGMT predictive model. The results suggest that the fractal features and uncertainty estimation via ensemble modeling offers improved predictive performance when compared to deep learning methods. In the future, we plan to develop the model with multi-omics data for a better prognosis of treatment response.

5. ACKNOWLEDGEMENTS

We acknowledge partial support for this work from National Institutes of Health grant #R01 EB020683.

REFERENCES

[1] Nam, J. Y. and Groot, J. F. D., "Treatment of glioblastoma," *J. Oncol. Pract.* **13**, 629 – 638 (2017).

[2] Margison, G. P., Povey, A. C., Kaina, B., and Santibáñez Koref, M. F., "Variability and regulation of O6-alkylguanine-DNA alkyltransferase," *Carcinogenesis* **24**, 625 – 635 (2003).

[3] Liu, D., Chen, J., Hu, X., Yang, K., Liu, Y., Hu, G., Ge, H., Zhang, W., and Liu, H., "Imaging-genomics in glioblastoma: Combining molecular and imaging signatures," *Frontiers in Oncology* **11**, 2666 (2021).

[4] Hajianfar, G., Shiri, I., Maleki, H., Oveisi, N., Haghparast, A., Abdollahi, H., and Oveisi, M., "Noninvasive O6 methylguanine-DNA methyltransferase status prediction in glioblastoma multiforme cancer using magnetic resonance imaging radiomics features: Univariate and multivariate radiogenomics analysis," *World Neurosurg.* **132**, 140 – 161 (2019).

[5] Korfiatis, P., Kline, T. L., Coufalova, L., Lachance, D. H., Parney, I. F., Carter, R. E., Buckner, J. C., and Erickson, B. J., "MRI texture features as biomarkers to predict MGMT methylation status in glioblastomas," *Med. Phys.* **43**, 2835 – 2844 (2016).

[6] Shboul, Z. A., Chen, J., and Iftekharuddin, K. M., "Brain tumor classification using wavelet and texture based neural network," *Sci. Rep.* **10**, 1 – 13 (2020).

[7] Xi, Y., Guo, F., Xu, Z., Li, C., Wei, W., Tian, P., Liu, T., Liu, L., Chen, G., Ye, J., Cheng, G., Cui, L., Zhang, H., Qin, W., and Yin, H., "Radiomics signature: A potential biomarker for the prediction of MGMT promoter methylation in glioblastoma," *J. Magn. Reson. Imaging.* **47**, 1380 – 1387 (2018).

[8] John, P., "Prediction of molecular mutations in diffuse low-grade gliomas using MR imaging features," *Int. J. Sci. Eng. Res.* **3** (2012).

[9] Islam, A., Reza, S. M. S., and Iftekharuddin, K. M., "Multifractal texture estimation for detection and segmentation of brain tumors," *IEEE Trans. Biomed. Eng.* **60** (2013).

[10] Reza, S. and Iftekharuddin, K. M., "Multi-fractal texture features for brain tumor and edema segmentation," *Medical Imaging 2014: Computer-Aided Diagnosis* **9035** (2014).

[11] Reza, S. M. S., Mays, R., and Iftekharuddin, K. M., "Multi-fractal detrended texture feature for brain tumor classification," *Medical Imaging 2015: Computer-Aided Diagnosis* **9414** (2015).

[12] Malinin, A., Prokhorenkova, L., and Ustimenko, A., "Uncertainty in gradient boosting via ensembles." Conference paper at ICLR 2021, http://arxiv.org/abs/2006.10562 (2020).

[13] Bakas, S., Akbari, H., and Sotiras, A., "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data* **4**, 170 – 117 (2017).

[14] Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., and Prior, F., "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," *J. Digit. Imaging* **36**, 1045 – 1057 (2013).

[15] Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., and Davatzikos, C., "Segmentation labels and radiomic features for the pre-operative scans of the TCGA-GBM collection," *Nat Sci Data* **4**, 170 – 117 (2017).

[16] Iftekharuddin, K. M., Jia, W., and Marsh, R., "Fractal analysis of tumor in brain MR images," *Mach. Vis. Appl.* **13**, 352 – 362 (2008).

[17] Islam, A., Iftekharuddin, K. M., Ogg, R. J., M.D., F. H. L., and Sivakumar, B., "Multifractal modeling, segmentation, prediction, and statistical validation of posterior fossa tumors," in [*Medical Imaging 2008: Computer-Aided Diagnosis*], Giger, M. L. and Karssemeijer, N., eds., **6915**, 1036 – 1047, International Society for Optics and Photonics, SPIE (2008).

[18] Ayache, A. and Véhel, J. L., "On the identification of the pointwise hölder exponent of the generalized multifractional brownian motion," *Stoch. Process. their Appl* **111** (2004).

[19] Ustimenko, A. and Prokhorenkova, L., "SGLB: Stochastic gradient langevin boosting," *Proceedings of the 38 th International Conference on Machine Learning, PMLR* **139** (2021).

[20] Raginsky, M., Rakhlin, A., and Telgarsky, M., "Non-convex learning via stochastic gradient langevin dynamics: a nonasymptotic analysis," *Proceedings of Machine Learning Research* **65**, 1 – 30 (2017).

[21] Malinin, A., Mlodozeniec, B., and Gales, M., "Ensemble distribution distillation." Conference paper at ICLR 2021, https://openreview.net/pdf?id=BygSP6Vtvr (2019).

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

[23] Han, L. and Kamdar, M. R., "MRI to MGMT: Predicting methylation status in glioblastoma patients using convolutional recurrent neural networks," *Pacific Symp. Biocomput.* **23**, 331 – 342 (2018).

[24] Calabrese, E., Villanueva-Meyer, J. E., and Cha, S., "A fully automated artificial intelligence method for non-invasive, imaging-based identification of genetic alterations in glioblastomas," *Sci. Rep.* **10**, 1 – 11 (2020).