Towards Accurate Diagnoses: An In-Depth Analysis of Brain Tumor Classification via Machine Learning and Deep Learning

1st Malika Muradi
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
malika.muradi@g.bracu.ac.bd

2nd Basit Hussain
Department of Computer Science and Engineering
Brac University
Dhaka, Bangaladesh
basit.hussain@g.bracu.ac.bd

***Abstract*—In the domain of cancer genomics, the precise and dependable categorization of tumors holds utmost importance in terms of timely detection and efficient strategizing for therapy. This study introduces a comprehensive methodology for the classification of brain tumors, employing machine learning and deep learning methodologies. The process commences with curating a broad dataset of brain tumor MRI scans, obtained from Kaggle. This dataset is carefully organized into four separate categories: Glioma, Meningioma, No Tumor, and Pituitary. In order to address the issue of class imbalance, we intentionally utilize oversampling as a method of data augmentation. The initial step in data preprocessing involves the standardization of image dimensions and the optimization of the dataset to enhance its suitability for later machine learning tasks. The scope of our study spans various machine learning models, which include both conventional methods such as Logistic Regression, Support Vector Machine, Naive Bayes, Random Forest, and Decision Tree, as well as more advanced ensemble techniques like Gradient Boosting and XGBoost. In addition, we provide Convolutional Neural Networks (CNN) as a deep learning framework for effectively capturing complex patterns inherent in genomics data. The results reveal the strengths of each model, with the CNN emerging as the top performer, achieving a precision and recall of 95% and an accuracy of 96% after augmentation.**

***Index Terms*—Cancer Genomics, Brain Tumor Classification, Machine Learning, Deep Learning, Explainable AI, LIME, Medical Imaging**

I. Introduction

Tumors represent a significant challenge in the field of oncology, with their complex genetic and molecular landscapes driving both their initiation and progression. A tumor, at its core, is an abnormal growth of cells that may be benign or malignant. The malignant form, commonly referred to as cancer, has the potential to invade or spread to other parts of the body. The genesis of tumors is often associated with specific genetic alterations that act as key oncogenic drivers, propelling the uncontrolled growth and division of cells [2].

Oncology advances in recent times have brought attention to the function of immune checkpoint blockade (ICB) in the management of different types of cancer. Promising outcomes have been observed with agents that target cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), programmed cell death protein 1 (PD-1), and programmed death-ligand 1 (PD-L1). This is particularly true in a subset of patients with a high tumor mutation burden (TMB). High TMB consistently selects for benefit with ICB therapy, suggesting that TMB is emerging as a possible biomarker [1].

Moreover, the discovery of human epidermal growth factor receptor 2 (HER2) in breast cancer has completely changed how this patient subgroup is treated. According to [3], HER2 testing has emerged as a crucial component of breast cancer management, providing focused treatment options, and directing therapeutic choices.

The significance of precision oncology is emphasized by the ongoing advancement in our comprehension of malignancies, encompassing its genetic foundations and molecular characteristics. The objective of this method is to customize treatments by considering the distinct genetic and molecular characteristics of an individual patient's tumor, hence providing a more focused and efficacious therapeutic approach. **Glioma:** Gliomas are a classification of neoplasms that arise from glial cells, which are the supporting cells within the central nervous system (CNS). Tumors of this nature have the potential to develop within the cerebral or spinal regions. Gliomas are classified according to the specific glial cell type implicated and their grade, which serves as an indicator of the tumor's level of aggressiveness. The prevailing classifications encompass astrocytoma's, oligodendrogliomas, and ependymomas. Glioblastoma multiforme (GBM) is an exceptionally aggressive variant of astrocytoma and represents the prevailing primary malignant neoplasm of the brain in the adult population [13].

Meningioma: Meningiomas are neoplasms that originate from the meninges, the membranous coverings enveloping the cerebral and spinal structures. These growths often have a benign nature and exhibit a sluggish rate of progression; nonetheless, there exists the potential for them to undergo malignant transformation. Although the majority of meningiomas do not exhibit symptoms, the presence of neurological symptoms can be attributed to their specific location and size. The main course of treatment for tumors is surgical removal;

however, in cases when total removal is not feasible or when the tumors are malignant, radiation therapy may be employed [14].

Pituitary: Pituitary tumors, also known as adenomas, originate from the pituitary gland, a diminutive gland situated near the cranial base that governs the synthesis of hormones important for regulating several physiological processes within the body. The majority of pituitary tumors are non-malignant in nature and are categorized according to their hormone-secreting capabilities. The manifestation of symptoms is contingent upon the specific hormone that is impacted. The available treatment modalities encompass surgical intervention, radiation therapy, and pharmacological interventions aimed at restoring hormone equilibrium [15].

No Tumor: In the context of medical imaging and diagnosis, the term "No tumor" refers to the absence of any discernible tumor in the examined region. It's crucial to accurately identify the absence of a tumor to avoid unnecessary treatments and interventions. Advanced imaging techniques, combined with machine learning and deep learning models, can assist radiologists in distinguishing between healthy tissues and pathological changes, ensuring accurate diagnosis [16].

Cancer genomics has witnessed a transformative shift with the advent of machine learning (ML) and deep learning (DL) techniques, offering promising avenues for accurate tumor classification. The vast and intricate genomic data necessitates the use of sophisticated algorithms that can discern patterns and make predictions with high precision. Among the plethora of ML algorithms, several have gained prominence in the realm of tumor classification, including Logistic Regression (LogisticReg), Support Vector Machines (SVC), Decision Trees, Naive Bayes, Random Forest, Gradient Boosting, and XGBoost (XGBClassifier). Additionally, Convolutional Neural Networks (CNN), a deep learning model, have shown exceptional prowess in handling image-based tumor data.

Machine Learning Models:

- **Logistic Regression (LogisticReg):** Logistic regression is a statistical technique used to estimate the likelihood of a binary outcome by considering one or more predictor factors. The utilization of non-linear relationships between predictors and outcomes is particularly advantageous [4].
- **Support Vector Machines (SVC):** SVC works by finding the hyperplane that best divides a dataset into classes. It's particularly adept at handling high-dimensional data, making it suitable for genomic datasets [5].
- **Decision Trees:** The proposed structure has a flowchart-like configuration, in which individual nodes symbolize distinct features, branches signify decision rules, and leaves indicate corresponding outcomes. The approach is characterized by its intuitive nature and its capacity to facilitate visualization, hence enhancing interpretability [6].
- **Naive Bayes:** This algorithm makes the assumption of predictor independence based on Bayes' theorem. It's particularly effective for datasets with large dimensions, like genomic data [7].
- **Random Forest:** The ensemble approach generates numerous decision trees throughout the training process and produces the mode of the classes as the output for classification purposes. It offers high accuracy and can handle large data with higher dimensionality [8].
- **Gradient Boosting:** An iterative algorithm that adjusts for the errors of the previous tree. It's known for its efficiency and precision [9].
- **XGBoost (XGBClassifier):** An optimized gradient boosting algorithm, renowned for its speed and performance, and has been widely adopted in various ML challenges [10].

Deep Learning Model:

- **Convolutional Neural Networks (CNN):** Primarily used for image data, CNNs automatically and adaptively learn spatial hierarchies of features from input images. Their ability to learn from image data makes them indispensable for classifying tumors from medical images [11].

Importance in Tumor Classification

The aforementioned ML and DL models play a pivotal role in tumor classification. Their ability to handle vast genomic datasets, discern intricate patterns, and predict with high accuracy makes them invaluable tools in the fight against cancer. For instance, CNNs, with their image processing capabilities, can detect and classify tumors like Glioma, Meningioma, and Pituitary from medical images with remarkable accuracy [16]. Classification of Tumors: These models classify tumors like Glioma, Meningioma, No tumor, and Pituitary by learning from labeled genomic or image data. Features from this data are fed into the algorithms, which then make predictions based on the patterns they've learned. For instance, CNNs might identify specific shapes or textures in medical images indicative of a particular tumor type [16].

The brain tumor MRI dataset sourced from Kaggle will be employed for analysis. This dataset has a total of 2661 pictures, which have been classified into four distinct categories: Glioma, Meningioma, No Tumor, and Pituitary. The aim of this study is to construct and assess various machine learning and deep learning models, such as Logistic Regression, Support Vector Machines (SVM), Random Forest, Decision Tree, Naive Bayes, Gradient Boosting, XGBoost, and Convolutional Neural Networks (CNN), in order to achieve precise classification of brain tumors. We will assess model performance before and after data augmentation to address class imbalance. Additionally, we will integrate Explainable AI (XAI) using LIME with Logistic Regression to gain insights into complex model decision-making processes. Our research will aim to improve brain tumor classification in cancer genomics, enabling more precise diagnoses and treatment planning in the future.

II. Related work

The application of machine learning and deep learning methods in the field of tumor classification has received con-

siderable attention in recent years. the paper [7] underscored the significance of interpretability inside machine learning models, particularly in vital fields such as medical imaging. The authors put forth a machine learning framework that is designed to be easily understood and interpreted. Their focus is on the classification of images, specifically in the context of predicting glioma cancer using magnetic resonance images of brain tumors. Their approach not only achieved high prediction accuracy but also provided meaningful explanations for its predictions, addressing the ”black box” nature of many deep learning models.

the paper [18] expanded the scope of AI-based systems in breast cancer pathology. While many studies focused on binary classifications, such as benign versus malignant, Mi and colleagues developed a multi-class classification system. The methodology employed in their study utilized deep learning techniques to effectively categorize breast digital pathology images into distinct classifications, including normal tissue, benign lesions, ductal carcinoma in situ, and invasive carcinoma. Their results showcased the potential of AI in providing more nuanced and clinically relevant classifications in breast cancer pathology.

In their recent study [19], proposed the utilization of a lightweight deep learning model coupled with a fine-tuning technique for the purpose of diagnosing brain cancers. For this goal, the researchers utilized the state-of-the-art object detection framework YOLO (You Only Look Once). The usefulness of the model in detecting malignant tumors, specifically glioblastoma, from MRI images was proved using the Brats 2021 dataset.

In their study [20], investigated the applicability of machine learning techniques for the prediction of tumor mutation burden (TMB) in lung adenocarcinoma, specifically through the analysis of histopathology pictures. Recognizing TMB as a crucial biomarker in lung cancer, they developed a system that combined histologic patterns and clinico-demographic data to predict TMB. Their results highlighted the value of incorporating histologic patterns in biomarker prediction, offering an informative and interpretable approach.

III. Methodology

In our methodology, we outline a comprehensive approach to handling and preparing the brain tumor image dataset for our study. Initially, we curated pertinent data from Kaggle, gathering a diverse set of brain tumor images. Subsequently, we embarked on data cleaning endeavors, which encompassed the removal of duplicates, correction of inconsistencies, and adept handling of missing values, culminating in the creation of a dependable and refined dataset. To address class imbalance, we strategically applied oversampling as a data augmentation technique. Our preprocessing phase involved standardizing image dimensions and optimizing the dataset for subsequent machine learning tasks. Beyond data preparation, our methodology extends to rigorous testing and validation procedures. We delve into the training of classifier models, encompassing both traditional machine learning approaches and deep learning techniques. Additionally, we introduce the implementation of Explainable AI, a pivotal component that enhances our understanding of the models' decision-making processes. This holistic methodology ensures the robustness and reliability of our study findings and analyses.

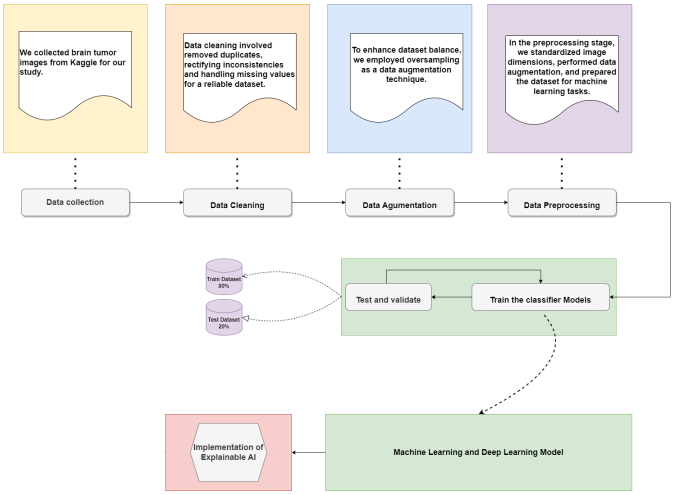

Fig. 1. A Visual Work Plan for Brain Tumor Classification in Cancer Genomics

The brain tumor MRI dataset available on Kaggle comprises a total of 2661 MRI images of human brains. These images have been meticulously categorized into four distinct classes, as illustrated in Figure 1. These classes are as follows: Gliomas originate from glial cells, which serve the essential function of nourishing and protecting neurons. Consequently, this type of brain tumor arises from abnormalities within these supportive cells. In contrast to gliomas, meningiomas are tumors that develop in the membranes surrounding the brain and spinal cord, known as the meninges. These tumors can exert pressure on the brain and spinal cord due to their location. Unlike the aforementioned tumor types, the “No tumor” category represents MRI scans that exhibit a negative result for the presence of any brain tumor. As a result, these scans provide a baseline for comparison and analysis. On a different note, pituitary tumors specifically originate in the pituitary gland, a small organ situated near the base of the brain. Consequently, these tumors can have various effects on hormone regulation and overall health.

Table 1 outlines the datasets' organization, which facilitates efficient data management and model development. The dataset is divided into two primary folders: the Training folder contains a total of 2227 MRI scans. These scans serve as the foundation for training the suggested model. By exposing the model to this substantial dataset, it acquires the necessary knowledge to make accurate predictions and classifications. In contrast, the “Testing” folder contains 434 MRI scans. These scans are reserved for the purpose of rigorously evaluating the performance of different model assumptions. They represent unseen data that the model has not encountered during training, ensuring a comprehensive assessment of its generalization

capabilities. The data is further categorized into four classes representing different types of brain tumors: Glioma, Meningioma, No Tumor, and Pituitary. Glioma, with 927 images in total, is the most prominent class, followed by Meningioma with 375 images, No Tumor with 432 images, and Pituitary with 927 images. These classes play a crucial role in the training of machine learning models to effectively differentiate between various tumor kinds in MRI data. The testing dataset comprises 434 MRI pictures and serves as an independent source of data for evaluating the models' performance. This ensures that the models' effectiveness is thoroughly examined on situations that have not been previously encountered. The structured and comprehensive dataset structure described here holds great value for academics and practitioners engaged in brain tumor detection and classification activities. It provides a well-organized and complete foundation for their investigations and model development endeavors.

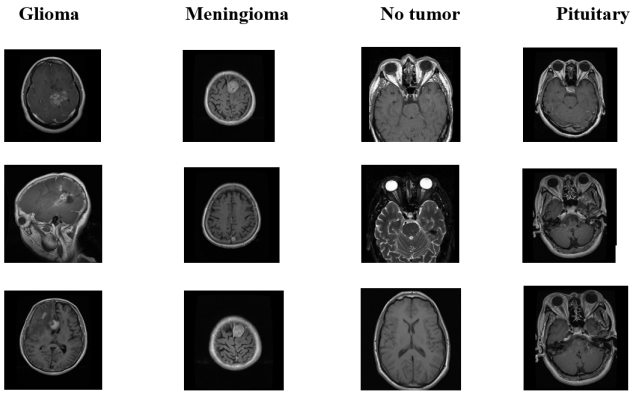

Fig. 2. Sample images of dataset

TABLE I
TUMOR DATASET SPLIT

Tumor	Total Dataset	Training Dataset	Testing Dataset
Glioma	927	826	101
Meningioma	375	247	128
No tumor	432	327	105
Pituitary	927	827	100
Total	**2661**	**2227**	**434**

A. Data Augmentation

Addressing class imbalance is crucial in machine learning, especially when working with datasets like the one you've mentioned, where some classes have significantly fewer samples than others. Oversampling is a common technique employed to tackle this issue.

The objective of oversampling is to achieve class distribution equilibrium by augmenting the number of instances in the minority class (namely, "Meningioma" and "No tumor") to align with the majority class (such as "Glioma" and "Pituitary"). This objective can be accomplished by replicating pre-existing samples from the underrepresented class or by producing artificial samples using methods such as SMOTE (Synthetic Minority Over-sampling Technique). By employing the technique of oversampling the minority classes, it is ensured that the machine learning and deep learning model is subjected to a more equitable and proportionate portrayal of the data. This approach mitigates the potential bias of the model towards the dominant class, hence improving the accuracy of predicting the underrepresented classes. In general, the utilization of oversampling techniques proves to be a viable approach for enhancing the efficacy and equity of machine learning models when dealing with imbalanced datasets, such as the one delineated in our discussion.

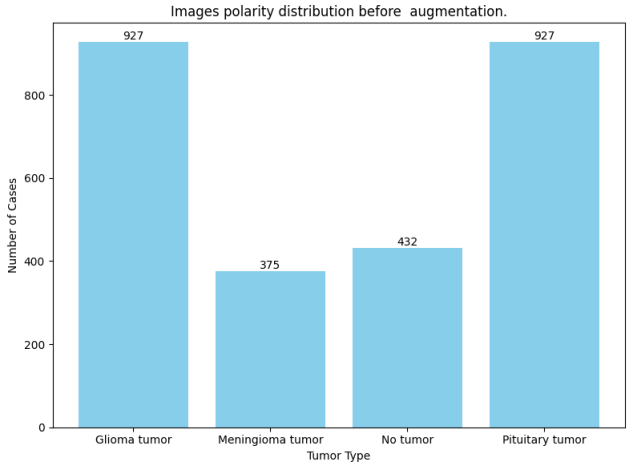

Fig. 3. Images Polarity distribution before augmentation

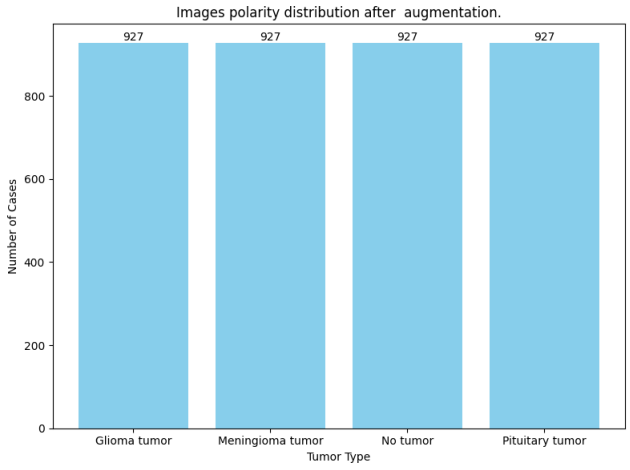

Fig. 4. Images Polarity distribution after augmentation

B. Data Pre-Processing

In our research, we conducted a series of crucial data preprocessing procedures. Initially, we procured a diverse collection of brain tumor images, meticulously sourced from the Kaggle dataset. These images were thoughtfully categorized to align with distinct tumor types, encompassing glioma, meningioma, the absence of tumors (no tumor), and the pituitary tumors class. A fundamental aspect of our preprocessing

pipeline involved image resizing. To establish uniformity in input dimensions, we uniformly resized all images to a standardized size of 224x224 pixels. We achieved this consistency by leveraging the 'target_size' parameter available within the Keras ImageDataGenerator. Furthermore, as part of the data preprocessing journey, we engaged in the process of flattening. This procedure involved transforming all images, across all categories, into a flattened array format. Consequently, our data repository was enriched with a collection of flattened image arrays and their corresponding category indices, meticulously stored in the labels list. This meticulous curation of data has rendered it ready for application in diverse machine learning tasks, prominently including image classification endeavors.

Following the data collection, labeling and resizing phase, we embarked on data splitting, a pivotal step in our study. We accomplished this by employing the 'train_test_split' function generously provided by ScikiLearn. This strategic division of the dataset into training, validation, and testing subsets enabled us to comprehensively evaluate our model's performance across different data partitions, thereby ensuring its robustness.

IV. Results and Discussion

In the following discussion, we will thoroughly assess the performance of our machine learning (ML) and deep learning (DL) models. The accompanying figure illustrates the accuracy of each model, both before and after the implementation of data augmentation techniques. The research will begin by assessing the levels of accuracy prior to augmentation, in order to establish a fundamental understanding of the models' basic performance. This allows us to evaluate their abilities before making any alterations. In the subsequent part, an analysis will be conducted to evaluate the attained accuracies subsequent to augmentation. This analysis will primarily focus on analyzing the impacts of different techniques on each unique model. The aim of this study is to examine the influence of data augmentation on the prediction abilities and performance improvements of the models. The objective of this presentation is to highlight the notable attributes and advancements achieved in our diverse array of machine learning (ML) and deep learning (DL) models.

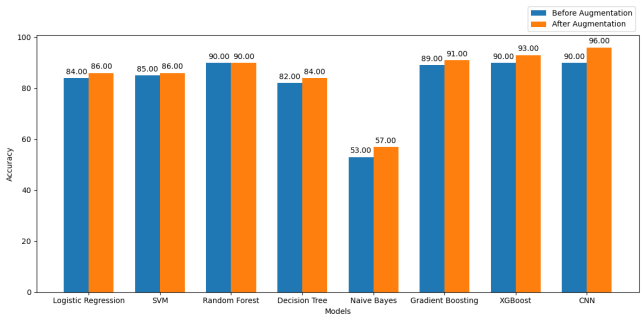

Fig. 5. A Comprehensive Accuracy Comparison Across Multiple Models before and after augmentation

The table provided illustrates the performance metrics of seven machine learning (ML) models and one deep learning (DL) model in the classification of brain tumors using cancer genomics data. These metrics are essential indicators of the performance of each model before the use of any data augmentation approaches. Initially, it is important to highlight that both Logistic Regression and Support Vector Machine (SVM) exhibit comparable performance measures, encompassing precision, recall, F1-Score, and accuracy, with average scores of approximately 84-85%. Hence, it is evident that traditional machine learning (ML) models regularly and proficiently exhibit the capacity to discern brain tumours in genomics data, while upholding a well-rounded performance. Although they may not be the most high-performing choices, these solutions offer a reliable basis for categorization jobs within this particular field. In contrast, the Random Forest algorithm has notable performance in terms of precision, recall, F1-Score, and accuracy scores, all of which consistently approximate 90%. The aforementioned model demonstrates a notable proficiency in effectively capturing the intricacies inherent in genomics data, hence achieving a harmonious equilibrium between precision and recall. Therefore, these results imply that Random Forest may be a compelling choice for brain tumor classification in cancer genomics. Moving on, the Decision Tree model exhibits balanced performance with precision, recall, F1-Score, and accuracy all at 82%. While not the most accurate model, its stability across metrics suggests its suitability for this classification task. In contrast, Naive Bayes lags behind other models with lower precision, recall, F1-Score, and accuracy, all around 53%. These scores indicate that Naive Bayes may not be the most appropriate choice for accurately classifying brain tumors based on genomics data in your project. Transitioning to the ensemble learning techniques, Gradient Boosting and XGBoost deliver strong performance. They achieve high precision, recall, F1-Score, and accuracy scores around 89-90%. As a result, these results suggest that these models have the potential to effectively classify brain tumors in the context of cancer genomics. However, the CNN (Convolutional Neural Network), being a deep learning model, stands out with the highest precision at 91%. This indicates its ability to minimize false positives, which is particularly crucial in medical applications like brain tumor classification. Furthermore, it maintains a recall of 90%, an F1-Score of 90%, and an accuracy of 90%. Consequently, these strong results underscore the suitability of CNNs for your project. They excel at capturing intricate patterns within the genomics data, making them a compelling choice when precision is of utmost importance in cancer genomics research.

In the context of brain tumor classification using cancer genomics data, the table below presents the performance metrics of seven machine learning (ML) and one deep learning (DL) models. This assessment is conducted after the application of augmentation techniques, which aim to enhance the models' effectiveness in accurately classifying brain tumors. Logistic Regression and SVM (Support Vector Machine) exhibit incremental improvements in their performance metrics following

TABLE II
MODEL PERFORMANCE BEFORE AUGMENTATION

Model	Precision	Recall	F1-Score	Accuracy
Logistic Regression	84	84	84	84
SVM	84	84	84	85
Random Forest	90	89	89	90
Decision Tree	82	82	82	82
Naive Bayes	53	53	50	53
Gradient Boosting	90	90	89	89
XGBoost	90	90	90	90
CNN	91	90	90	90

augmentation. Specifically, both models achieve precision and recall scores of 85%, indicating a slight but positive impact on their ability to correctly identify brain tumors. Moreover, their F1-Score and accuracy have seen marginal improvements, with Logistic Regression reaching an accuracy of 86%. These subtle enhancements imply that the augmentation techniques have contributed to a modest increase in the overall effectiveness of these traditional ML models. Meanwhile, Random Forest maintains its robust performance with precision, recall, F1-Score, and accuracy scores all remaining steady at 90%. This consistency demonstrates the model's capacity to encompass the complexity inherent in genomics data, even after being supplemented. The Decision Tree model exhibits moderate advancement, specifically in its F1-Score, which currently rests at 83%, and in its accuracy, attaining 84%. The observed positive changes indicate that the utilization of augmentation approaches has played a role in enhancing the model's capacity to achieve a more desirable trade-off between precision and recall. Consequently, this renders the model a marginally more dependable option. On the other hand, Naive Bayes demonstrates a significant enhancement in performance after augmentation. The precision and recall metrics demonstrate an increase to 55% and 56% correspondingly. Furthermore, the F1-Score and accuracy metrics exhibit more notable enhancements, reaching 52% and 57% respectively. Although not among the top-performing models, these enhancements illustrate the potential of augmentation techniques to significantly boost the performance of less complex models, making them more viable options. Gradient Boosting and XGBoost continue to exhibit strong performance, with some models even achieving better results than before. Gradient Boosting, for instance, attains a recall of 91%, while XGBoost achieves an impressive 93% accuracy. These outcomes highlight the resilience of ensemble learning techniques, particularly when applied to augmented data. Lastly, the CNN (Convolutional Neural Network) stands out as the leader in performance after augmentation. It now boasts a remarkable 95% precision and recall, further minimizing false positives. Additionally, the F1-Score and accuracy rise to 95% and 96%, respectively. These results emphasize that the CNN remains the top choice for highly accurate and precise brain tumor classification in cancer genomics data, even after augmentation techniques have been applied. In summary, augmentation techniques have generally had a favorable impact on the performance of these models in the realm of brain tumor classification using cancer genomics data. While several models exhibit minor improvements, others, such as Naive Bayes, demonstrate significant enhancements. CNN remains in the forefront in terms of precision and accuracy, rendering it a compelling selection for the aims of your project. It is advisable to conduct additional investigation, experimentation, and verification using augmented data in order to identify the best appropriate model that is customized to meet your individual requirements and objectives.

TABLE III
MODEL PERFORMANCE AFTER AUGMENTATION

Model	Precision	Recall	F1-Score	Accuracy
Logistic Regression	85	85	85	86
SVM	85	86	85	86
Random Forest	90	90	90	90
Decision tree	82	82	83	84
Naive Bayes	55	56	52	57
Gradient Boosting	90	91	91	91
XGBoost	92	91	92	93
CNN	95	95	95	96

The generation of a confusion matrix is a basic procedure in evaluating the efficacy of various machine learning (ML) and deep learning (DL) models, as well as comprehending their ability to accurately classify data. The application of a confusion matrix is a valuable methodology for assessing the effectiveness of a model by providing a comprehensive breakdown of its prediction results. The classification process entails the distinction between true positives, which are instances that are correctly classified as positive, true negatives, which are instances that are correctly classified as negative, false positives, which are negative instances that are incorrectly classified as positive, and false negatives, which are positive instances that are incorrectly classified as negative. The performance of each model can be properly evaluated by creating confusion matrices and analyzing metrics like as precision, recall, accuracy, and F1-score. This evaluation aids in ascertaining the optimal model for a certain job. The assessment of machine learning (ML) models, such as Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Decision Trees (DT), Naive Bayes (NB), Gradient Boosting (GB), and XGBoost, can be conducted by employing confusion matrices to evaluate their effectiveness and constraints in handling class imbalance within the dataset. Within the domain of deep learning models, namely Convolutional Neural Networks (CNN), the confusion matrix holds significant utility in evaluating the neural network's proficiency in accurately distinguishing between different classes. The comprehensive evaluation of models through the utilization of confusion matrices is a crucial procedure in ascertaining the best suitable approach for a specific categorization undertaking.

V. DISCUSSION

In our study, we embarked on a comprehensive exploration of machine learning and deep learning techniques applied to the critical task of brain tumor classification within the realm of cancer genomics. Our research yielded several noteworthy

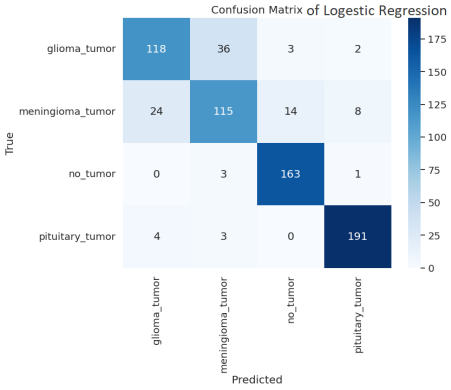

Fig. 6. Confusion Matric

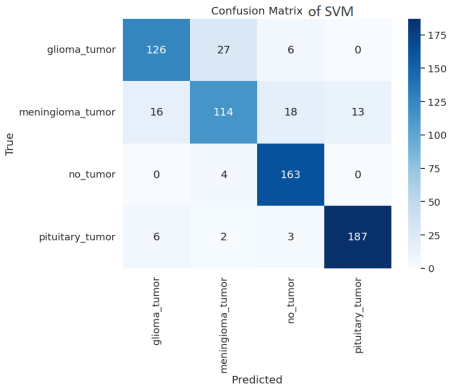

Fig. 7. Confusion Matric

findings and considerations. Firstly, our investigation revealed the potential of these advanced computational methods in enhancing the accuracy of brain tumor classification. The Convolutional Neural Network (CNN) emerged as the standout performer, achieving impressive precision, recall, and accuracy metrics. This underscores the promise of deep learning models in capturing intricate patterns within medical images for precise diagnoses.

A crucial aspect of our research was the strategic han-

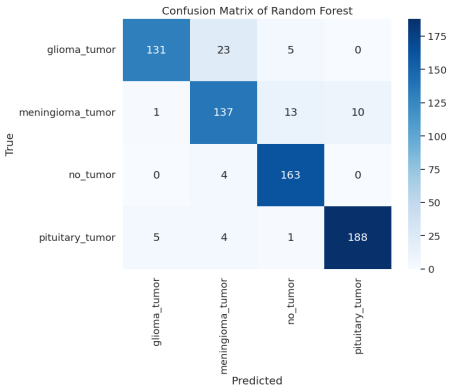

Fig. 8. Confusion Matric

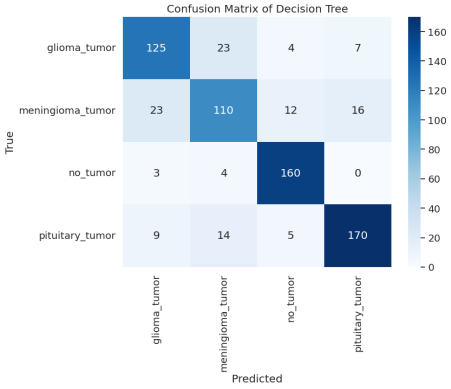

Fig. 9. Confusion Matric

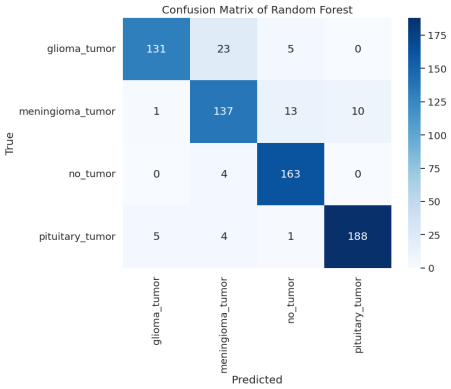

Fig. 10. Confusion Matric

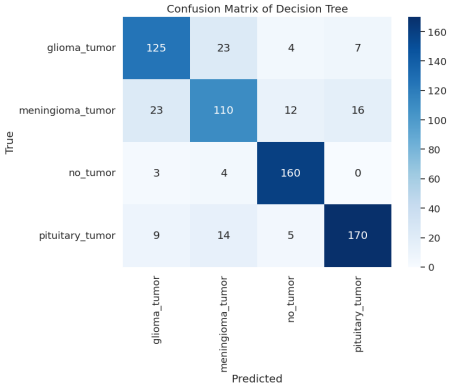

Fig. 11. Confusion Matric

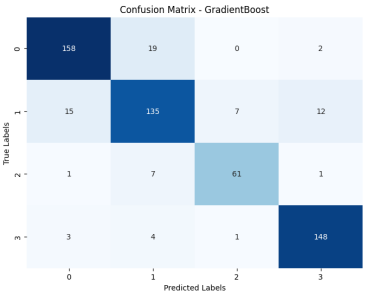

Fig. 12. Confusion Matric

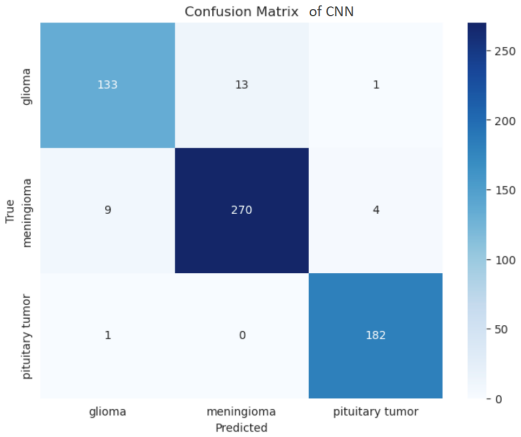

Fig. 13. Confusion Matric

dling of class imbalance, a common challenge in medical image classification tasks. Through the skillful application of data augmentation techniques, specifically oversampling, we successfully addressed this issue, leading to more balanced model training. The significant improvements observed across all models following data augmentation underscore the importance of careful dataset curation and preprocessing in the context of medical image analysis.

VI. Future Directions and Conclusion

Building upon the foundations laid by our study, several promising avenues for future research emerge. Firstly, expanding the dataset to include multi-modal data, such as genetic markers or clinical information, holds potential for enhancing model robustness. This holistic approach could provide a more comprehensive understanding of the factors influencing brain tumors, leading to more accurate diagnoses.

Ensemble methods, which harness the strengths of both machine learning and deep learning models, deserve further exploration. Techniques like stacking or boosting could be investigated to create hybrid models that leverage the advantages of each approach, potentially achieving even higher accuracy.

Additionally, the concept of transfer learning, particularly leveraging pre-trained models on large medical imaging datasets, should be considered. This approach has the potential to accelerate model development by extracting relevant features and patterns from brain tumor images.

Clinical validation represents a crucial next step. Collaborating with healthcare institutions to assess model performance on real patient data is essential for ensuring the clinical relevance and reliability of our findings.

Real-time implementation of our models in clinical settings is a significant advancement. Developing applications that can be seamlessly integrated into radiology departments or medical facilities can expedite the diagnostic process, potentially leading to quicker treatment decisions and improved patient outcomes.

Finally, ethical considerations must not be overlooked. As AI continues to advance in healthcare, addressing concerns related to patient data privacy, model biases, and ethical use of AI in clinical decision support is paramount. Future research should incorporate ethical considerations and guidelines to ensure responsible AI deployment in medical applications.

In conclusion, our study serves as a stepping stone toward more accurate and interpretable brain tumor classification in the context of cancer genomics. Future research endeavors should focus on expanding the scope, improving model performance, ensuring seamless clinical integration, and upholding ethical standards, ultimately benefiting healthcare professionals and patients alike.

References

[1] Chan, T., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S., Stenzinger, A., & Peters, S. (2018). Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic.

[2] 2. Cheng, D. T., Mitchell, T. N., Zehir, A., Shah, R. H., Benayed, R., Syed, A., ... & Berger, M. (2015). Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology.

[3] Wolff, A., Hammond, M., Allison, K., Harvey, B. E., Mangu, P., Bartlett, J., ... & Dowsett, M. (2018). Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update.

[4] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression (Vol. 398). John Wiley & Sons.

[5] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

[6] Quinlan, J. R. (1986). Induction of decision trees. Machine learning, 1(1), 81-106.

[7] Rish, I. (2001). An empirical study of the naive Bayes classifier. In IJCAI 2001 workshop on empirical methods in artificial intelligence (Vol. 3, No. 22, pp. 41-46). IBM New York.

[8] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

[9] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.

[10] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

[11] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.

[12] Akselrod-Ballin, A., Galun, M., Gomori, J. M., Brandt, A., & Basri, R. (2016). An integrated segmentation and classification approach applied to multiple sclerosis analysis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4424-4432).

[13] Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W. K., ... & Ellison, D. W. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta neuropathologica, 131(6), 803-820.

[14] Rogers, L., Barani, I., Chamberlain, M., Kaley, T. J., McDermott, M., Raizer, J., ... & DeAngelis, L. M. (2015). Meningiomas: knowledge base, treatment outcomes, and uncertainties. A RANO review. Journal of neurosurgery, 122(1), 4-23.

[15] Melmed, S. (2011). Pathogenesis of pituitary tumors. Nature Reviews Endocrinology, 7(5), 257-266.

[16] Giger, M. L. (2018). Machine learning in medical imaging. Journal of the American College of Radiology, 15(3), 512-520.

[17] Pintelas, E. G., Liaskos, M., Livieris, I., Kotsiantis, S., & Pintelas, P. (2020). Explainable Machine Learning Framework for Image Classification Problems: Case Study on Glioma Cancer Prediction.

[18] Mi, W., Li, J., Guo, Y., Ren, X., Liang, Z., Zhang, T., & Zou, H. (2021). Deep Learning-Based Multi-Class Classification of Breast Digital Pathology Images.

[19] Shelatkar, T., Urvashi, D., Shorfuzzaman, M., Alsufyani, A., & Lakshmanna, K. (2022). Diagnosis of Brain Tumor Using Light Weight Deep Learning Model with Fine-Tuning Approach.

[20] Sadhwani, A., Chang, H. W., Behrooz, A., Brown, T., Auvigne-Flament, I., Patel, H. L., ... & Steiner, D. F. (2021). Comparative analysis of machine learning approaches to classify tumor mutation burden in lung adenocarcinoma using histopathology images.