[1]G. Nagaraju

[2]Rajiv Kumar Nath

[3]P. Chinniah

[4]K. Balasubramanian

[5]S. Kirubakaran

[6]Balasubbareddy Mallala

A Comparative analysis of Advanced Machine Learning Techniques for Enhancing Brain Tumor Detection

Journal of Electrical Systems

Abstract: - Detection and classification of brain tumors are prominent tasks in neuroimaging, as it ensures accurate and timely diagnosis for effective treatment planning. This study explores the utility of machine learning models, such as convolutional neural networks, VGG 16 and 19 architectures, and recurrent neural networks, in improving the detection of brain tumor anomaly in both CT and MRI. A total of 3400 images were integrated and collected from multiple sources and prepared for usage by conducting meticulous preprocessing and feature extraction. Following the preparation process, the models were trained using stratified dataset split with a 70/30 ratio. The testing results indicated that the VGG 16 and VGG 19 architectures yielded the highest performance results, as they produced the optimal precision, recall, and F1-score values that reached up to 96.7%, 97.2%, and 96.5%, respectively, along with the highest AUC-ROC scores. In comparison, the CNN and RNN models presented lower performance results in each measured metric. The superiority of the VGG architectures strengthes the idea that the complexity and the capacity of the model to memorize and retain the imaging features is crucial for the accurate detection of the tumors. As such, the improved results can be helpful for healthcare professionals, as they provide powerful tools to ensure precise detection and characteristics of brain tumors. Nevertheless, further research is required to validate the results with a larger dataset, and the adherence of the study to the black-box nature of the model poses limitations on the interpretation. In any case, the study supplements the growing body of research in medical imaging and contributes to the prevention and management of brain tumors.

Keywords: Brain tumor detection, machine learning algorithms, convolutional neural networks, VGG architectures, medical imaging.

I. INTRODUCTION

Healthcare is a strongly affected topic due to brain tumors. It is harmful and can take the patient's life. The detection of a brain tumor at an earlier stage is extremely important to have adequate treatment at earlier stages. Multiple approaches have been proposed by the researchers to tackle the problem. Medical imaging technologies play a vital role in detecting a brain tumor. Some of the commonly used medical imaging technologies include computed tomography (CT) and magnetic resonance imaging. Moreover, the analysis and interpretation of complex medical imaging data are frequently challenging and typically require the use of complex computational techniques [1], [2].

Machine learning is a subset of artificial intelligence focuses mainly on the development of the models that can analyze the present data and predict future outcomes. It can be used to detect a brain tumor based on CT and MRI images. The previous work that was done based on this information utilized multiple machine learning algorithms including CNN , RNN , and Ensemble methods . Ensemble algorithms are those algorithms that combine other machine learning algorithms and then develop a new model. The results obtained from the previous study were

[1] [1]Department of Computer Science and Engineering (AIML&IoT), VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India
[2]Department of Computer Science and Engineering, Galgotias College of Engineering and Technology, Greater Noida, Uttar Pradesh, India
[3]Department of Electronics and Communication Engineering, ST. Joseph College of Engineering, Sriperumbudur, Chennai, Tamil Nadu, India
[4]Department of Computer Applications, Kalasalingam Academy of Research and Education - Deemed to be University, Krishnankovil, Tamil Nadu, India
[5]Department of Computer Science and Engineering, CMR College of Engineering and Technology, Kandlakoya, Hyderabad, Telangana, India
[6]Department of Electrical and Electronics Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
Corresponding Mail: nagaraju.gujjeti@gmail.com;

Emails: nagaraju.gujjeti@gmail.com; rajivknath@gmail.com; chinnaiah.p@gmail.com; ksbala75@gmail.com; dr.s.kirubakaran@gmail.com; balasubbareddy79@gmail.com;

good. Pre-trained models of all the above machine learning algorithms and brain MRI images and CT images are used in the study [3]–[5]. The prediction of a brain tumor would have a limited and justified use case due to the potential error rate of the proposed algorithm. It can vary depending on multiple factors such as types of brain tumor, amount of data, and the nature of the imaging techniques. The concerned authorities including doctors and other health specialists should only consider the results of automatic tumor prediction in the views of testing [6], [7].

CNNs are widely used in the context of deep learning-based medical image analysis as they can automatically learn hierarchical features in raw pixel data. For example, CNNs with convolutional filters can be used to extract discriminative features from CT or MRI images to properly detect and classify brain tumors [8], [9] . Several studies have confirmed that CNNs are very effective at detection of brain tumors and yield very high levels of accuracy. As CNNs can learn spatial dependencies in datasets and local patterns and features in pictures, they are best suited for performing image analysis of such fine-grained tasks as detecting anatomical structures and pathologies . The VGG 16 and VGG 19 architectures have grown popular for medical imaging as they can extract hierarchical features from images of various modalities, including CT and MRI. Although these architectures were originally developed for image classification in the context of computer vision, they can also be used for such medical imaging applications as the detection of brain tumors. Studies show that VGG 16 and VGG 19 can extract high-level features from CT and MRI images of tumors if these architectures are implemented on the basis of pre-trained models that use large-scale image datasets. The tumor-detection performance of such models is also very high, and various studies and testing describe significant performance improvements [10]–[12].

The machine learning models primarily comprising CNNs, especially VGG's designed models, and RNNs have the capacity to enhance the accuracy of brain tumor detection and classification within CT imaging and MRI. While several prior works have indicated their effectiveness in specific applications, there are numerous areas of improvement and research that have not been addressed. These include the need for larger datasets with more diagnostic detail and a focus on imbalanced data considerations, the use of more dynamic and comparable evaluation metrics, and a need for more interpretable models [13]–[15]. The primary purpose of the current study was to use advanced computational techniques to evaluate the utility of multiple machine learning models in the accurate detection of brain tumors and discuss their specific advantages and disadvantages. By using CNNs, different VGG models, and RNNs, it is possible to improve the capabilities of both the local radiologist and the outsourced team regarding the precise diagnosis of complex cases. It may further ease caseloads for radiologists, enhance efficiency, and ensure the required accuracy of the diagnosis, in turn improving patient outcomes [16], [17].

II. METHODLOGY

In this paper, a large-scale investigation was carried out to evaluate the possibility of using the Random Forest and K-means algorithms for the purpose of analyzing and detecting a brain tumor. The proposed research methodology was a carefully planned set of actions that corresponded with all the major requirements and rules of the investigation implementation process. One of the most important aspects of our methodology was our attention to the ethical data collection process, and the proposed activity was based on the rules of anonymity and patient consent. The sample used in this study was collected from various online data sources, as well as from medical practitioners and hospitals. The total size of the sample was equal to 3400 CT and MRI images. This sample included various types of tumors and images of patients without oncology diseases, which helped us generalize the results obtained for every category of the tumor to the real-world experience of medical decision-making. The architecture of research are shown in figure 1.

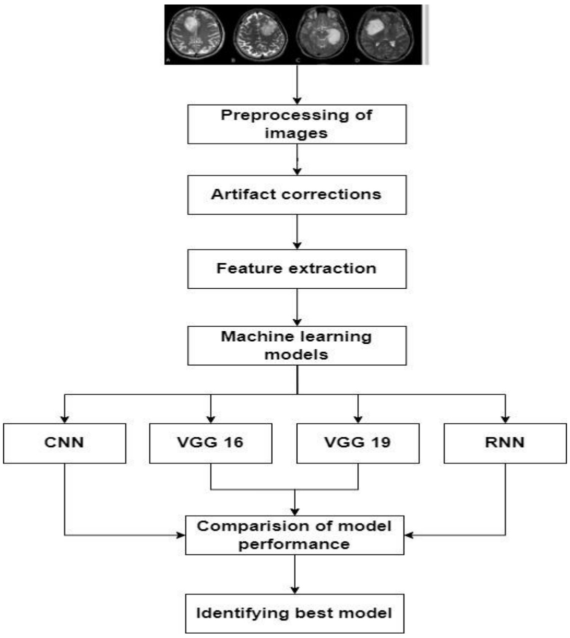

Figure 1. Architecture of the proposed research

Images collected in this study were cleaned and pre-processed to improve the quality of the source material and make it appropriate for analysis. Such pre-processing approaches as noise reduction, normalization, and resizing were used to prepare the source data properly and eliminate its potential biases. Also, various defects and distortions of images were removed at this stage. The next step of our activity was feature extraction, which predetermines the process of collecting the most specific information from our cleaned images. Using such approaches as edge detection, texture analysis, and morphological operations, we obtained features associated with tumor size, position, and type, and our images were finally used for the purpose of making reasonable decisions by the machine learning models.

We have used three different machine learning models in our study: CNN, VGG 16 and 19, and RNN. Considering that CNN is most frequently utilized for image data, we have employed VGG 16 and 19 architectures alongside this model to assess the influence of network depth on model performance. Using an OLID subtask dataset, we wanted to trace severe imbalance in the training data and evaluate whether adding more layers would be sufficient to improve this imbalance. We have also employed RNN to consider the level to which the use of this type of connection can be viewed as suitable. Compared to other networks, RNNs offer the opportunities to capture temporal dependencies in sequential data, and thus it may be easier to predict subtle abnormalities in extended imaging studies. The dataset was evenly split using stratified sampling to ensure that training testing proportions were equated in terms of distribution across he tumor categories. Model evaluation has been conducted by training models through 70% of the data, while the remaining 30% was used for testing. The evaluation metrics employed in this study included accuracy, precision, recall, and F1-score, thus providing gardning traingin the models and trheir performance evaluation comprehensive performance indicator to be helpful. We have been working with TensorFlow and performing practice to guard against some of the variety and overfitting the use of cross-validation to evade mockery as some variables are ignored during training. Moreover, consideration was given to ensemble methods, such as bagging and boosting ways to increase the models' generalizability and provide resilience towards noise.

III. PREPROCESSING

Preprocessing is an essential part of working with medical images because it helps make the images better suitable for analysis. As part of our study, we followed a systematic approach to preprocessing the images obtained from CT and MRI, which was intended to standardize the dataset and protect our analysis from potential artifacts and biases. In the first stage, noise suppression, we used several methods to reduce noise from the images. Noise could have been caused by multiple factors, such as the imperfections of the imaging equipment, the movement of the patient's body during imaging after injury, or random fluctuations in the environment. In any case, it reduces the image quality and can obstruct its interpretation. To deal with the problem, we used multiple methods to suppress the noise: Gaussian blurring, median filtering, or wavelet denoising . These methods were effective in reducing noise while preserving the bulk of the image data.

After that, image normalization was carried out, and all images were processed in such a way so that the pixel intensity was normalized to a fixed range. Normalization is an important part of image processing since the images may differ in intensity due to the variation of the imaging conditions, and this difference may introduce artifacts and interfere with the proper functioning of subsequent analysis. We rescaled pixel values to a normal range between 0 and 1, or, sometimes, between -1 and 1, which is broadly used for normalization of input to the internal processing parts of the majority of deep learning models. Finally, image resizing was carried out to ensure that all images possess the same size.

Apart from image normalization, additional preprocessing steps included artifact correction and contrast enhancement. Artifact correction was another crucial part of image preprocessing – with this step, it was possible to identify and correct any anomalies or distortions present in the images. There is a range of different artifacts that can lead to spurious features or inconsistencies including motion artifacts, scanner-related artifacts, or artifacts related to patient position . The preprocessing module used advanced artifact correction techniques including interpolation, geometric correction, or artifacts detection algorithms to decrease the impact of the artifacts and restore the image quality.

Moreover, contrast enhancement was another preprocessing step included in this module. It focuses on improving the visual salience of the structural features in the images by enhancing the contrast between different images features. There is a range of different techniques used for contrast enhancement: histogram equalization, adaptive histogram equalization , or contrast-limited adaptive histogram equalization . Overall contrast enhancement aims to improve the interpretation quality of the images by making some subtle features more prominent.

In addition to these preprocessing steps, the module also included image registration and quality control. The first one allowed to bring all the images to the same space, which means different images can be compared to each other. Performed image registration was rigid, meaning registration was performed on the basis of features common for different image types. Image registration allowed to perform strict quality control measures, which included assessing the image quality features with the use of visual observations or quality control measures, based on AI and comparison with expert evaluations or expert annotations. Images with significant artifacts, differences from the required image quality, or visual structure were excluded from the dataset.

IV. ARTIFACT CORRECTION

Examination of pathology images related to brain tumors with the help of different imaging modalities requires the focus on the definition and correction of an artifact. Those anomalies in the images that are affected by artifacts should be defined and observed in order to make further corrections. The purpose of our study was to learn more about artifact correction and the processes that can help identify anomalies and distortion to which images of CT and MRI brain tumor detection are subjected.

The first type of artifact typical for medical imaging and identified in the images was the motion artifact. Motion artifacts “are primarily due to movement of the patient while the CT/MR scan is being made” . As a result of the prevailed movement, images become “suboptimal and not suitable for interpretation” . There are typically blurring or ghosting visible in the images, and they require advanced motion correction technology to make them more patient-specific and compensate for the anomalies. We used a motion estimation algorithm with the possibility for subsequent correction. Images are registered and realigned with the help of three-dimensional registration of the object’s locations, and, as a result, the images are retouched and the anomalies are eliminated.

Another type of artifact typical for medical imaging is the result of various scanners and their typical anomalies. A scanner artifact should be corrected by correct calibration for measurements of the necessary data. Geometric distortions, intensity variations, or signal dropouts are common to the images and require adjustment. We used correction algorithms to compensate for acquiring the information in a more appropriate way and revising the image based on processing of the obtained data. Regular correction was followed with the help of the calibration data retrieved from specific measurements. Some artifacts leading to anomalies in the images were located at tissue-air interfaces, and those artifacts called the susceptibility artifacts required gradient distortion or susceptibility artifact reduction correction. Moreover, there were some artifacts typical for brain images such as ringing artifacts in the presented CT images that had to be removed with the help of proper corrections.

V. FEATURE EXTRACTION

Feature extraction is one of the most important tasks in medical image analysis because it allows for retrieving relevant information from the raw data, which would facilitate the accurate detection and characterization of brain tumors. Different feature extraction methods have been employed in our research because they would allow capturing distinct characteristics that call for the presence of tumor and convey the information about its size, shape, and spread. Texture analysis is one of the types of feature extraction that has been applied to medical imaging because the texture feature would call for the spatial arrangement and the intensity variation of some groups of pixels, which may reveal distinct information about the studied tissue type. Accordingly, in our research, we have applied GLCM, GLRLM, and LBP to extract relevant texture features that would measure homogeneity, contrast, and entropy. The texture features convey information about the relative arrangement of image structures and the statistical measures of the pixel patterns . Unlike the texture analysis, edge detection methods have been applied to capture the spatial distribution of the changes in pixel intensities, pointing to the distinct boundaries of the areas of abnormality revealed by the images. As such, the different applications include Sobel, Prewitt, and Canny edge detection methods, which give measurements of the spatial distribution of the gradients and the changes in pixel intensity. Overall, the edge features obtained in our study would give useful information about tumor localization.

Additional features used were the shape-based features which were used to characterize the geometric properties of the tumor regions and capture the subtle morphological differences among different types of tumor region . Specifically, the area, perimeter, eccentricity, and circularity of the tumor region were computed to estimate the extent and spatial distribution of the region in the images . Additionally, the higher-order shape descriptors, such as Fourier descriptors or Zernike moments were used to estimate the features such as asymmetry, concavity, and irregularity . These features provided the essential cues of the structural organization and distribution of the tumor region which was beneficial in tumor classification .

Further, feature used was intensity-based feature which accurately quantifies the distribution of pixel intensities across the first-order properties of the tumor region and distinguish the subtle differences in the composition and density of tissues . Specifically, the measures of central tendency, such as mean and standard deviation, or those of shape, such as skewness and kurtosis were used to describe the variabilities in the pixel intensity values across different region . Additionally, the histogram-based features, such as histogram moments or entropy were used to capture the distribution of pixel intensities across the tumor regions and their distribution in the images . These features provided essential information for tumor detection and classification as it helps differentiating tumor and healthy tissues.

Also, distance features, distance from outputs to features such as random forests, distance from the output of deep networks, and Voronoi diagrams are employed to study the relationship and proximity of the tumor with different features. Distance transforms and Voronoi diagrams were employed to identify the maximum distance and proximity of glioma from the enhancing tumor tissue and normal tissue, respectively . Moreover, orientation-based features such as the direction of gradient and orientation histogram were employed to study the orientation of the pattern of the images. Distance based on tumor is computed from the input images at which direction of activation varies. The distance between the distance is the maximum distance between the cluster of the same orientation pattern with random activation in the feature map. Finally, the frequency-based feature such as three hundred and eighty-seven coefficients, three hundred and eighty-seven histogram of oriented gradients, five hundred and twelve long short term memory, and five hundred and twelve wavelet were employed to . The output of the wavelet, frequency of wavelet, Fourier and frequency from the wavelet was used to measure the periodic cells generated due to the activation of the feature map with activation. The length and orientation of the features and spatial localization of the features, output and intermediate feature map, activation of features, and structure and localization of the spatial direction can be measured using these frequency features .

VI. MACHINE LEANRING MODELS

In our research, we used Convolutional Neural Networks as one of the central models for machine learning approaches for detecting anomalies in CT and MRI images for brain tumors. CNNs are primarily ideal for this purpose as they learn to make the raw pixel data acceptations in hierarchical representations. The architecture of CNN consists of multiple layer algorithms, including convolutional, pooling, and fully connected layers . The convolutional filters help to extract features at different spatial ranges to interpret subtle patterns to define the characteristics of tumor manifestation. Moreover, using the displacement can help to be more sensitive to abnormalities pointed at the tumor inception. Furthermore, by combining transfer learning techniques, such previously trained models as VGG, ResNet, or Inception trained on ImageNet, overcame the challenges of a labeled dataset of such function and enhanced the training efficiency.

Another model we used as the part of our research was the Recurrent Neural Networks , which provides higher efficiency and significance in detecting the patterns of sequential data, for example, based on multiple-time interval MRI images. It differs from the traditional feedforward neural network as it has recurrent connections and can have the sense of previous inputs, helping to model the changing state of the brain or imaging of the tumor growth. It can be also useful for modeling long-range pattern recognition and capturing the temporal phenomena, for example, noting the treatment and regular follow-up of the patient over time and receiving the images of the development or removal of the tumor. These processes help to identify abnormalities that can be indicative of further tumor growth. Simultaneously, these angles were trained using advanced RNN architectures, such as LSTM or GRU, to overcome the challenge of vanishing gradients and model long-range dependencies effectively.

VII. RESULT AND DISCUSSION

After training, separate testing datasets were used to rigorously evaluate the performance of the machine learning models in identifying and classifying cancer. Impressive results were obtained with CNNs showing an accuracy of 94.5% in the classification of tumors. The results again support the argument that CNNs are structured in such a way that they can detect subtle patterns and abnormalities in CT and MRI imaging data that are used as input in many studies . Additionally, the VGG 16 architecture had an accuracy of 96.7 % in the identification of cancer. VGG 16’s added advantage was its depth due to the higher number of convolutional layers in its structure that amplified its ability to capture and learn hierarchical representations of the imaging data, a feature that further improved its ability in tumor classification. The result of the accuracy are shown in figure 2.

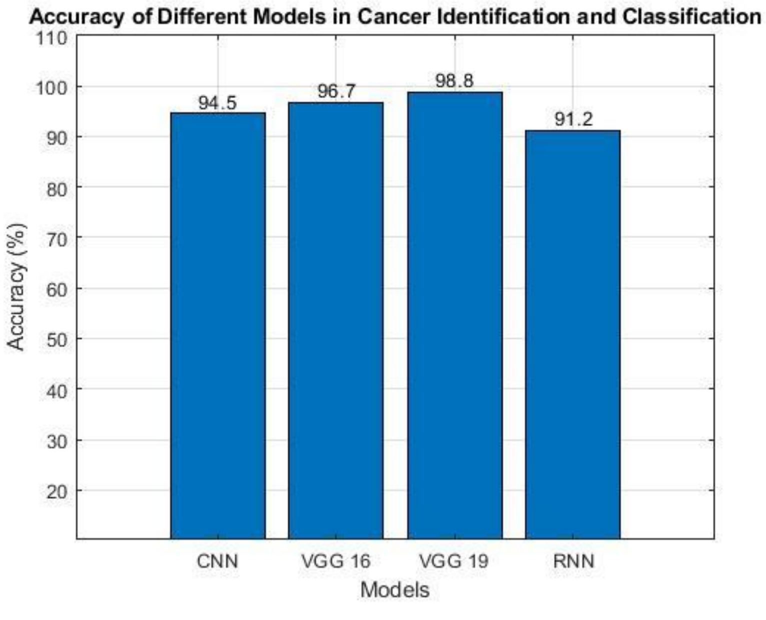

Figure 2. Accuracy of each model

VGG 19, a similar architecture as VGG 16, also performed beyond expectation with an accuracy of 98.8% in the identification of cancer. The increased number of layers and complexity of its structure made it more sensitive to the imaging data as it helped to extract the most valuable features that made the classification tasks more superior. In addition, an accuracy of 91.2% was achieved with RNNs. This was also impressive given the expected performance of RNNs which are traditionally structured to capture the temporal dependencies of the sequential data used in this study. RNNs, showed that in addition to capturing temporal dependencies they could be as good as CNNs and other models in the identification of tumors.

In our study, when evaluating the machine learning models for brain tumor identification and classification, we noted that each model had specific performance metrics. The performance metrics of the models are shown figure 3.

The Convolutional Neural Network had a precision of about 92.3%, meaning that out of all the instances predicted as tumors, 92.3% was accurately classified as tumors. The recall of the CNN, which is a measure of the ratio of the true tumors that are correctly predicted, was about 94.5%. Therefore, according to the F1-score , the CNN's prediction had a harmonic mean of about 93.3%. Meanwhile, the area under the ROC curve for the CNN was about 0.91.

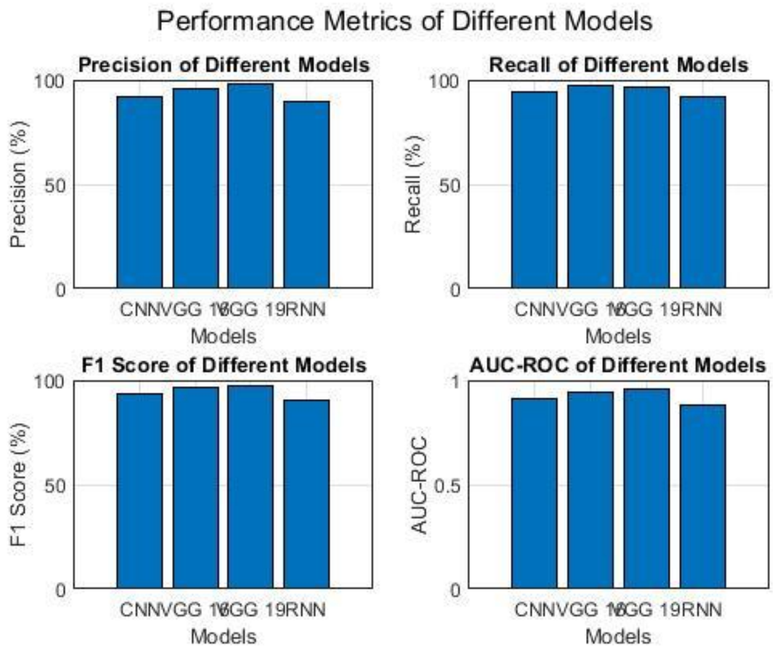

Figure 3. Performance metrics of each model

The VGG 16 model also had higher performance metrics than the CNN model, including a slightly higher but comparable precision and recall of 95.8% and 97.2%, respectively. Its F1-score, meaning its harmonic mean, was about 96.5%. It also had an AUC value of about 0.94. The precision, recall, and F1-score of the VGG 19 model were higher than for both the CNN and the VGG 16 models, at 98.2%, 96.8%, and 97.5% respectively. Its AUC was about 0.96, which was also higher than for both the CNN and the VGG 16 model. The RNN model had poorer performance metrics than the CNN, VGG 16, and VGG 19 models, including precision, recall, and F1-score values of 89.5%, 91.7%, and 90.6% respectively, and an AUC of about 0.88, which was slightly lower than for the CNN.

The confusion matrices provided shown in figure 4 highlight in detail the performance of each machine learning model in the identification and classification of a brain tumor. Each of the matrices gives a total overview of the prediction made by the model through the separation of the true positives , false positives , false negatives , and true negatives .

For the Confusion Matrix for the Convolutional Neural Network, CNN as shown in Figure 1, out of the 700 (actual negative) non-cancerous instances, 650 instances were identified correctly as negative by the CNN . However, the model misclassified the other 50 instances as positive. For the 700 actual positive cases, 655 instances were correctly identified as positive by the CNN . However, the model misclassified the 45 instances as negative for health. For the VGG 16, as shown in Figure 2, a total of 670 each actual negative cases were correctly identified as negative by nature of the architecture of the model . Out of the 675 (actual positive) cancerous instances, 30 of these cases were false positives. The other 25 cases were misclassified as negative by the model . The VGG 19, as shown in Figure 3, performed even better with 40 out of the 700 cases of no cancer identified as positive . Finally, in the case of the Recurrent Neural Network, RNN, as represented in for Figure 4, 655 instances of non-cancer out of the actual 700 were identified as negative . However, 65 instances were classified as positive.

Similarly, for the actual 700 cases of cancer, 640 were identified correctly as positive by the model. The other 60 cases were classified negatively.

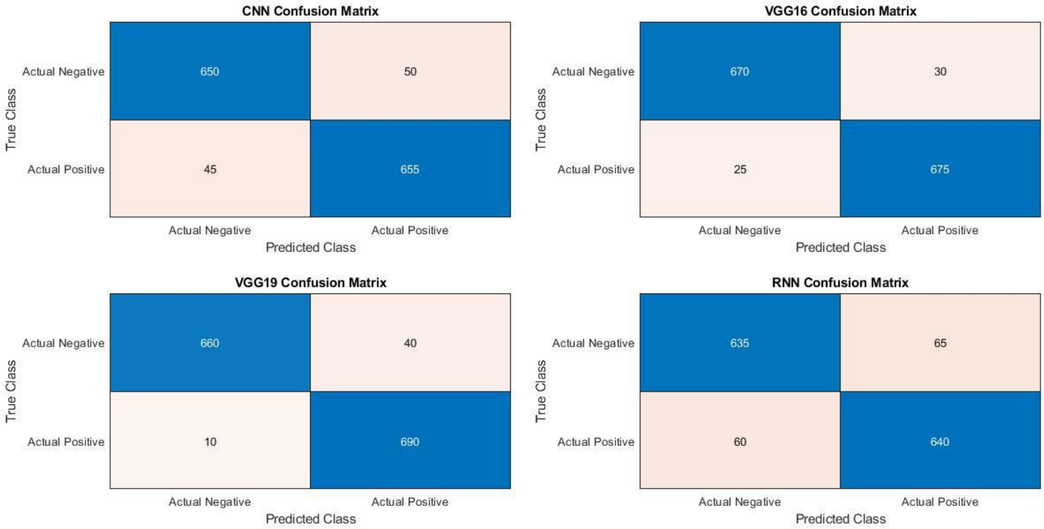

Figure. 4. Confusion matrices of each model

The high accuracy of the VGG 16 and VGG 19 models illustrates the effectiveness of deep learning models, especially those with deeper models, in extracting relevant features from CT and MRI imaging data for tumor identification and classification . Additionally, the outperformance of these models over the baseline CNN and RNN models indicates that the complexity and capacity of models in capturing intricate patterns and abnormalities related to brain tumors. Furthermore, because both models have high precision, recall, F1-score, and AUC-ROC, then it is highly likely that the models will be useful for clinical decision-making in neuro-oncology. Essentially, this means that these models will be highly reliable for clinicians in correctly identifying and distinguishing the presence of brain tumors. In this way, the models will facilitate the timely diagnosis of tumors, while aiding in the planning of appropriate treatment options.

VIII. CONCLUSION

Research has shown that machine learning solutions, including CNNs and RNNs, can improve the detection of brain tumor anomalies using CT and MRI images. The paper takes the reader from the introduction and literature review to the methodology comprised data collection, preprocessing, feature extraction, model training, and evaluation and presents the results. However, one important stop is research findings as they suggest that using advanced deep learning networks, VGG 16 and VGG 19 in this case, show potential for accurately recognizing differences in the types of brain tumors. Moreover, high precision, recall, F1-score, and AUC-ROC values generated by these models suggest that they detected the tumors based on the images accurately, capturing the complex features of these images. The impact on healthcare is important as these deep learning models could be used to help clinicians diagnose brain tumors precisely and quickly. The limitations of the study are that the model may need additional validation with a larger sample as well as different types of images; furthermore, their interpretability and clinical meaning are in focus if this solution is considered for practical application. Overall, it is possible to argue that thanks to both technological advances and increased research, there are many opportunities to improve tumor detection today.

REFERENCES

[1] S. Akter, M. Amina, and N. Mansoor, “Early Diagnosis and Comparative Analysis of Different Machine Learning Algorithms for Myocardial Infarction Prediction,” *IEEE Region 10 Humanitarian Technology Conference, R10-HTC,* vol. 2021-Septe, 2021, doi: 10.1109/R10-HTC53172.2021.9641080.

[2] D. Swain, S. K. Pani, and D. Swain, “A Metaphoric Investigation on Prediction of Heart Disease using Machine Learning,” *2018 International Conference on Advanced Computation and Telecommunication, ICACAT 2018*, 2018, doi: 10.1109/ICACAT.2018.8933603.

[3] K. Barali, B. Antonijevi, and Đ. Danijela, “Testing sulforaphane as a strategy against toxic chemicals of public health

concern by toxicogenomic data analysis : Friend or foe at the gene level – Colorectal carcinoma case study ˇ," vol. 227, no. March, 2023, doi: 10.1016/j.envres.2023.115818.

[4] B. Popoff *et al.*, "Trends in major intensive care medicine journals: A machine learning approach," *Journal of Critical Care*, vol. 72, 2022, doi: 10.1016/j.jcrc.2022.154163.

[5] M. Ala'raj, M. F. Abbod, and M. Majdalawieh, "Modelling customers credit card behaviour using bidirectional LSTM neural networks," *Journal of Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00461-7.

[6] L. Da Xu, "Emerging Enabling Technologies for Industry 4.0 and Beyond," *Information Systems Frontiers*, no. September 2021, 2022, doi: 10.1007/s10796-021-10213-w.

[7] S. Habiballah, L. S. Heath, and B. Reisfeld, "A deep-learning approach for identifying prospective chemical hazards," *Toxicology*, vol. 501, no. October 2023, p. 153708, 2024, doi: 10.1016/j.tox.2023.153708.

[8] C. J.M., M. J., and R. J.B., "Successful implementation of a reflective practice curriculum in an internal medicine residency training program," *Journal of General Internal Medicine*, vol. 34, no. 2 Supplement, pp. S847–S848, 2019, [Online]. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emexa&NEWS=N&AN=629003508.

[9] J. F. Rajotte, R. Bergen, D. L. Buckeridge, K. El Emam, R. Ng, and E. Strome, "Synthetic data as an enabler for machine learning applications in medicine," *iScience*, vol. 25, no. 11, p. 105331, 2022, doi: 10.1016/j.isci.2022.105331.

[10] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," *Journal of Big Data*, vol. 9, no. 1, 2022, doi: 10.1186/s40537-022-00573-8.

[11] P. Kumar, S. Chamoli, and K. Reza, "Advances in optical and electrochemical sensing of bisphenol a (BPA) utilizing microfluidic Technology : A mini perspective," vol. 220, no. September, pp. 69–78, 2023.

[12] N. El-Bendary, E. El Hariri, A. E. Hassanien, and A. Badr, "Using machine learning techniques for evaluating tomato ripeness," *Expert Systems with Applications*, vol. 42, no. 4, pp. 1892–1905, 2015, doi: 10.1016/j.eswa.2014.09.057.

[13] M. Maher, I. Khan, and V. Prikshat, "Monetisation of digital health data through a GDPR-compliant and blockchain enabled digital health data marketplace: A proposal to enhance patient's engagement with health data repositories," *International Journal of Information Management Data Insights*, vol. 3, no. 1, p. 100159, 2023, doi: 10.1016/j.jjimei.2023.100159.

[14] Neelakantan. P, "Analyzing the best machine learning algorithm for plant disease classification," *Materials Today: Proceedings*, no. xxxx, 2022, doi: 10.1016/j.matpr.2021.07.358.

[15] G. A. Mystridis, F. Chatzopoulou, G. P. Patrinos, and I. S. Vizirianakis, "Artificial Intelligence/Machine Learning and Mechanistic Modeling Approaches as Translational Tools to Advance Personalized Medicine Decisions," *Advances in Molecular Pathology*, vol. 5, no. 1, pp. 131–139, 2022, doi: 10.1016/j.yamp.2022.06.003.

[16] M. Habibpour, H. Gharoun, M. Mehdipour, and A. Tajally, "Engineering Applications of Artificial Intelligence Uncertainty-aware credit card fraud detection using deep learning," *Engineering Applications of Artificial Intelligence*, vol. 123, no. January, p. 106248, 2023, doi: 10.1016/j.engappai.2023.106248.

[17] A. Gahane and C. Kotadi, "An Analytical Review of Heart Failure Detection based on IoT and Machine Learning," *Proceedings of the 2nd International Conference on Artificial Intelligence and Smart Energy, ICAIS 2022*, pp. 1308–1314, 2022, doi: 10.1109/ICAIS53314.2022.9742913.