Brain Image Segmentation using Deep Learning Approach

Vikas Humbe[1], Srikant Somanna[1], Satish Bhalshankar[2]

[1,2]School of Technology, S. R. T. M. University, Sub-Campus, Latur

Dept. of Management Science, Dr. BAM University, Chh. Sambhajinagar

[1]vikashumbe@gmail.com,

Abstract: Brain MRI image segmentation plays a crucial role in diagnosing neurological disorders. This research paper explores the application of deep learning techniques, specifically U-Net, ResNet, and advanced architectures, for segmenting brain MRI images sourced from Kaggle. We analyze existing work utilizing these datasets and propose a comparative study evaluating the performance of these models on different segmentation tasks, highlighting their strengths and weaknesses. We investigate the impact of data augmentation, hyperparameter tuning, and architectural modifications on segmentation accuracy, focusing on metrics like Dice coefficient, IoU, and precision/recall.

Keywords: Medical Imaging, MRI images, Segmentation, Deep Learning

(Article history: Received: Oct 31,2025 and accepted Jan, 30,2026)

1 Introduction

Accurate and efficient segmentation of brain MRI images is essential for various clinical applications, including tumor detection, stroke analysis, and Alzheimer's disease diagnosis. Manual segmentation is time-consuming and prone to inter-rater variability. Deep learning offers a powerful alternative, achieving high accuracy and automation. Kaggle provides several publicly available datasets of brain MRI images, offering a valuable resource for researchers. This paper reviews existing research on these datasets and presents a comparative evaluation of different deep learning architectures for brain MRI segmentation.

Brain Tumor affects millions of people worldwide. Among its various manifestations, BT stands out as a significant source of pain, disability and reduced quality of life, particularly in the ageing population (see Fig. 1) [1, 19]. The motivation behind addressing BT arises from its substantial impact on people's lifestyles, particularly in recent years. Among various medical conditions, brain tumors represent a common and aggressive form of malignant disease, often associated with a short life expectancy when diagnosed at advanced stages. We have to consider the disease's profound impact on daily activities [2]. There is a growing recognition of the potential of computer-based tools, machine learning, and deep learning techniques to improve the diagnosis and classification of various medical conditions. By leveraging advanced technologies, we aim to provide more accurate and efficient methods for identifying and categorizing Brain Tumor, ultimately leading to better treatment decisions. The integration of these innovative approaches reflects a commitment to advancing medical diagnostics and addressing the unique challenges posed by deadly diseases like BT [3, 4].

Medical MRI brain image segmentation for tumor detection using deep learning has seen significant advancements, primarily driven by the development of convolutional neural networks (CNNs) and their derivatives. The U-Net architecture, a popular choice for biomedical image segmentation, employs a contracting path to capture context and a symmetric expanding path for precise localization, making it highly effective for segmenting complex structures like brain tumors from MRI images[1] [3]. Fully convolutional networks (FCNs) have also been adapted for semantic segmentation tasks, including brain tumor detection, by leveraging their ability to process images of arbitrary sizes and produce detailed segmentations efficiently[2]. The SegNet model, another fully convolutional approach, has been applied to multi-modal MRI datasets to segment brain tumors and their sub-regions, achieving high accuracy by integrating feature maps from different MRI modalities[15]. Additionally, deep learning frameworks like nnU-Net have automated the design of segmentation methods, adapting to various datasets without manual tuning, thus enhancing the accessibility and performance of segmentation tasks[12]. The use of VGG-16 and ensemble learning approaches has

further improved the detection and classification of brain tumors, achieving high precision and accuracy in identifying tumor regions[16]. Despite these advancements, challenges remain, such as the need for large annotated datasets and the refinement of automatic segmentation results for clinical use[4] [5]. Interactive frameworks like DeepIGeoS address these challenges by allowing user interactions to refine initial CNN-based segmentations, thereby improving accuracy and robustness[8]. Overall, deep learning has significantly enhanced the capability of MRI brain image segmentation for tumor detection, offering promising tools for early diagnosis and treatment planning.

Numerous studies have utilized Kaggle's brain MRI datasets for deep learning-based segmentation. A review of relevant papers reveals:

- **U-Net based approaches:** Many studies have adopted the U-Net architecture, known for its effectiveness in medical image segmentation. Modifications like incorporating attention mechanisms, residual connections, or dense blocks have been explored to improve performance. Commonly used datasets include the "Brain Tumor Segmentation" dataset and the "Medical Image Segmentation" dataset (specify dataset names and links if available on Kaggle). These studies often report high Dice coefficients and IoU scores. However, limitations may include sensitivity to hyperparameter tuning and potential overfitting on smaller datasets.

- **ResNet based approaches:** ResNet's ability to handle deep architectures effectively has also been applied to brain MRI segmentation. Studies often utilize pre-trained ResNet models (e.g., ResNet50, ResNet101) as feature extractors, followed by a decoder network for segmentation. This approach leverages the pre-trained weights to improve performance, especially with limited training data. Challenges may involve adapting the pre-trained model to the specific characteristics of brain MRI images and potential computational cost.

- **Advanced Architectures:** More recent research explores advanced architectures like Attention U-Net, TransUNet, and other hybrid models that combine the strengths of different architectures. These models often incorporate attention mechanisms to focus on relevant features and improve segmentation accuracy, especially in complex regions. However, they may require more computational resources and careful hyperparameter tuning.

We take advantage of ResNet50, VGG-16 and Inception-V3 and employed ensemble techniques to enhance the overall performance of the models.

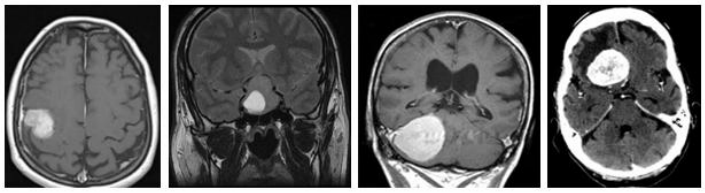


Fig. 1. Brain Tumor image.

ResNet50*:* ResNet50 consists of 50 convolutional layers, making it a Deep Convolutional Neural Network designed for various computer vision tasks (see Fig. 2) [20]. Pre-trained on large datasets such as ImageNet, ResNet50 can extract and leverage high level features, making it suitable for transfer learning across various visual recognition tasks [5].

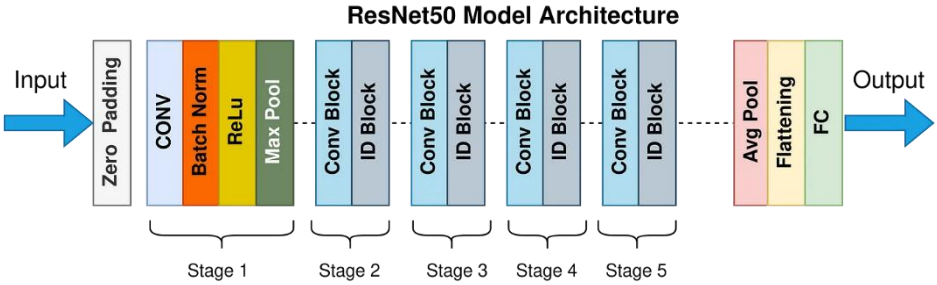


Fig. 2. ResNet50 architecture.

VGG-16: The input size for the first Convolutional Neural Network in VGG-16 is 224x224 pixels. VGG-16 employs a padding of 1 pixel for its 3x3 convolutional layers. The VGG-16 ensures that the spatial dimensions are preserved after convolution. Max-pooling is a down-sampling operation that helps in capturing the most important features while reducing the spatial dimensions using the equations (1, 2) [16].

- The soft function of VGG-16:

$$y_i = \sum_{j=1}^{n} e^{z_j} \,.\, e^{z_i} \tag{1}$$

- The loss function will be:

$$E = \frac{1}{3}\big(d(c_1, G_1) + d(c_2, G_2) + d(c_3, G_3)\big) \tag{2}$$

Inception-V3*:* One of the most important functions of deep neural networks is image classification. When it comes to identifying objects in images, one of the most widely used convolutional neural network models is Inception-V3. Deep learning-driven image identification is used by self-driving cars to detect traffic hazards, clinicians to identify malignant tissue in medical images, and Facebook to assist users with photo tagging [18].

2 Literature Survey

Early detection and treatment of BT is crucial for improving the quality of life. BT classification accuracy through deep learning and machine learning also applied transfer learning ResNet-34 and others. Ensemble models to improve performance. The model achieved high accuracy, precision, recall and F-score result [6].

In a recent study concentrating on BT, the methodology involved utilizing deep features, with a particular emphasis on the application of a CNN to gain these intricate features from images associated with BT. Specifically, the SVM, K-NN and Naive Bayes. The results of the experiment were noted, demonstrating an accuracy of 90%.

Classify the feature by support vector machine with an accuracy 95%. Restricted Boltzmann machine model for feature extraction [7].

BT is a significant health issue worldwide, causing decreasing overall quality of life and death threat. This common condition primarily affects overall life, and since there's no cure for BT, treatments mainly target symptom management. The main goals include relieving pain, enhancing brain function and maintaining stability [8].

Medical image segmentation using MRI is essential for accurate diagnosis, treatment planning, and monitoring of diseases, particularly brain tumors and bone abnormalities. The complexity of anatomical structures and imaging artifacts in MRI necessitates robust and automated segmentation techniques.

Fuzzy C-Means (FCM) and its variants have been widely used for MRI segmentation, addressing challenges like intensity inhomogeneities by incorporating weighted bias field information and kernel-induced distances. These methods have shown promising results in segmenting both synthetic and real MRI datasets, validated by metrics such as the Silhouette accuracy index [1]. Modified active contour methods have also been effective, especially for extracting bone structures from 3D MRI data, demonstrating improved accuracy over traditional approaches [2].

Multimodal optimization approaches, such as modified 3D histogram-based segmentation combined with particle swarm optimization, automate the selection of segment numbers and outperform traditional clustering methods in robustness and effectiveness [3]. Otsu's thresholding, enhanced with preprocessing and morphological operations, has been applied to brain tumor segmentation, with optimal class selection significantly improving accuracy and sensitivity [4]. Hybrid metaheuristic algorithms, like the combination of Coronavirus Optimization and Harris Hawks Optimization, further enhance multilevel thresholding for both 2D and 3D MRI images, outperforming other metaheuristics in segmentation quality [5].

Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized MRI segmentation. Enhanced CNNs, optimized with algorithms like BAT, automate brain tumor segmentation and outperform traditional methods in precision, recall, and accuracy [6]. Reviews highlight that combining Conditional Random Fields (CRF) with Fully Convolutional Neural Networks (FCNN) or ensemble models yields superior performance for brain tumor segmentation [7]. Overall, deep learning methods are consistently more effective than classical approaches for MRI brain tumor segmentation, enabling faster and more accurate diagnosis [8].

Stochastic multi-agent approaches, using autonomous agents and modified region growing algorithms, have been proposed for 3D MRI segmentation. These methods allow for efficient and refined segmentation through agent collaboration, showing promising results in both healthy and pathological brain images [9]. Foundation models like the

Segment Anything Model (SAM) offer zero-shot segmentation capabilities, but their performance varies across datasets and tasks, excelling with well-defined objects but struggling with ambiguous regions like brain tumors. SAM performs better with box prompts and outperforms similar methods in single-point prompt settings, though its performance is moderate to poor in some medical imaging scenarios [10].

Segmentation quality is commonly assessed using metrics such as the Silhouette accuracy index, Jaccard similarity index, sensitivity, specificity, accuracy, PSNR, SSIM, and normalized correlation coefficient, ensuring objective comparison across methods [1] [4] [5].

Recent advances in MRI image segmentation leverage a combination of classical, optimization-based, and deep learning methods. Deep learning, especially CNN-based approaches, currently offers the best performance for brain tumor segmentation. However, challenges remain in handling data variability, annotation limitations, and complex anatomical structures, highlighting the need for continued innovation and careful validation in clinical applications.

3 Proposed Model

This study will evaluate the performance of U-Net, ResNet-based, and an advanced architecture (e.g., TransUNet) for segmenting brain MRI images from selected Kaggle datasets. The specific datasets used will be clearly identified and justified based on their relevance and size. The methodology includes:

- **Data Preprocessing:** This involves image normalization, resizing, and data augmentation techniques like rotation, flipping, and elastic transformations to address data imbalance and improve model generalization.
- **Model Architecture:** We will implement and compare U-Net, a ResNet-based encoder-decoder architecture, and an advanced architecture (e.g., TransUNet). The choice of the advanced architecture will be motivated by the literature review and its potential to improve segmentation accuracy. Detailed architectural specifications will be provided.
- **Training and Evaluation:** The models will be trained using an appropriate optimizer (e.g., Adam) and loss function (e.g., Dice loss, cross-entropy). We will employ k-fold cross-validation to evaluate the models' robustness and generalize performance. Evaluation metrics will include Dice coefficient, Intersection over Union (IoU), precision, recall, and F1-score. We will analyze the impact of different hyperparameters, including learning rate, batch size, and number of epochs.
- **Comparative Analysis:** The performance of the different models will be compared using statistical tests (e.g., t-test) to determine if any significant differences exist. We will also conduct a qualitative analysis by visualizing the segmentation results to assess the models' strengths and weaknesses in different regions of the brain.

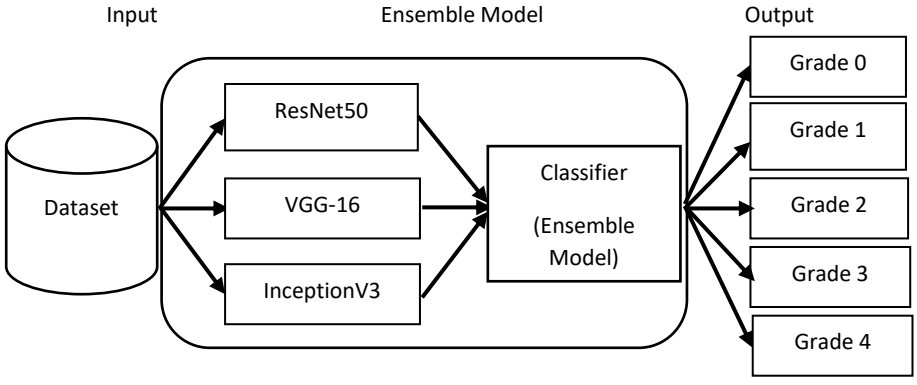


Fig. 3. Proposed ensemble model for knee osteoarthritis images classification.

The training dataset phases are data collection, data enhancement, region of interest, and image segmentation.

3.1 Data description

Non-pathological images were collected from the IXI Dataset. Among the available four types of MRIs, T1 contrast enhanced (T1ce) is the most commonly used contrast while performing single-contrast tumour classification, hence in this research, 259 T1 weighted volumes were chosen randomly from the IXI dataset as healthy samples.

Grades based on severity images are shown in Fig. 4.

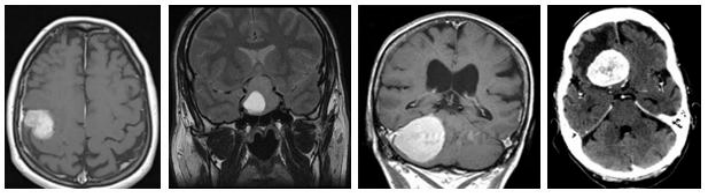


Fig. 4. Images of different grade based on severity level.

3.2 Data enhancement

Data enhancement is performed using histogram equalization. This is useful for images with a background and light or dark foreground. In particular, the approach can lead to better visualization of bone structure in radiographs and better detail in photographs that are either overexposed or underexposed (see Fig. 5). Enhanced MRI images may undergo processing to remove unwanted features that could hinder an accurate diagnosis using equations (3, 4). [10, 11]

- The probability of an occurrence of a pixel of level i in the image is

$$p_x(i) = p(x = i) = \frac{n_i}{n}, 0 \leq i \leq L \tag{3}$$

where n_i is the frequency of occurrence of i and n is the total number of observations.

- The cumulative distribution function $cdf_x(i)$can be expressed as:

$$cdf_x(i) = \sum_{j=0}^{i} p_x(x = j) \tag{4}$$

where p_x(x = j) is the probability distribution function for x.

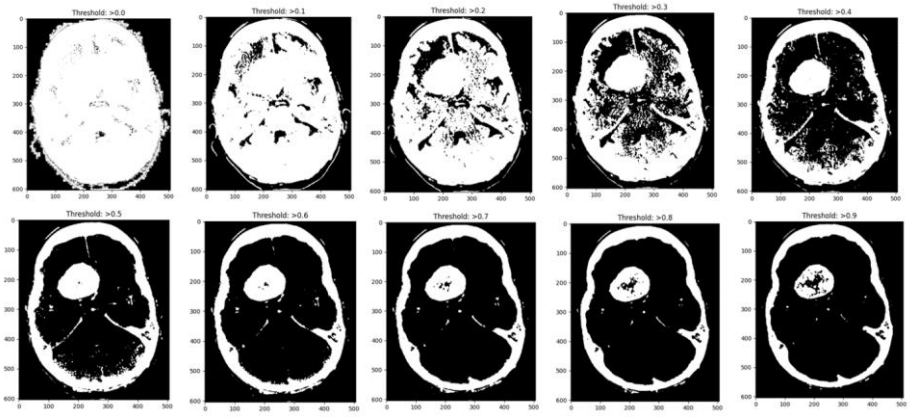


Fig. 5. Normal image and Enhanced images.

3.3 Region of Interest (RoI)

To classify the severity of BT, the area that helps us to determine the severity of the injury is the area of the meniscus. RoI imaging is taken from the MRI images which have been cropped. A center pinhole filter is used to substantially reduce patient dose outside the region of interest while preserving or enhancing the image as shown in Fig. 6. [12]

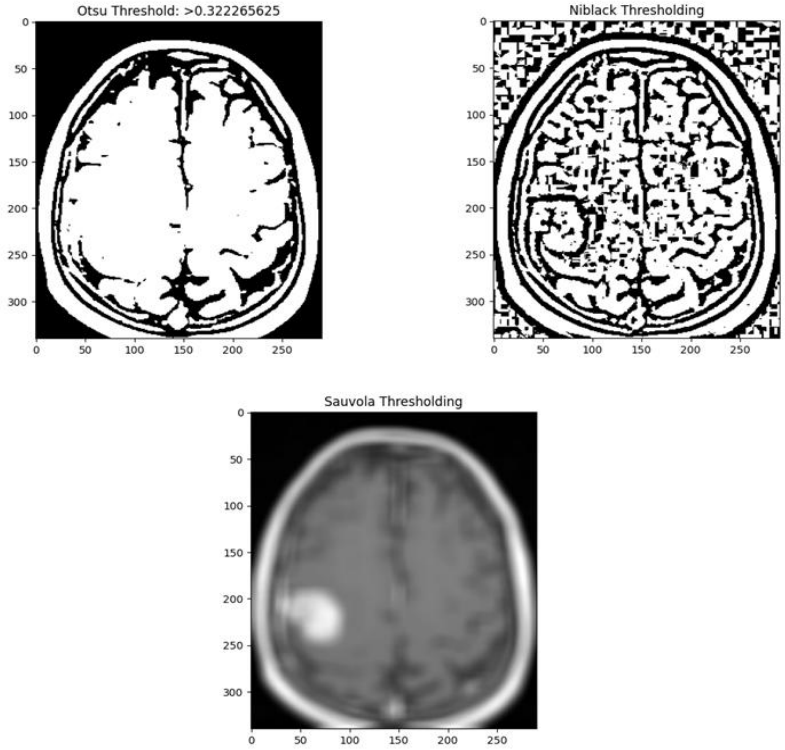


Fig. 6. Image shows the Region of Interest.

3.4 Image segmentation

Image segmentation is a process of dividing an image into several parts. The purpose of segmenting an image is to get more meaningful and easier to analyze. It is usually used to locate objects and create boundaries. K-means is the method of choice for MRI image segmentation. After K-means segmentation with k=3, distinct regions corresponding to clusters with different colors can be seen shown in Fig. 7. Each cluster represents a different group of pixel intensities [13].

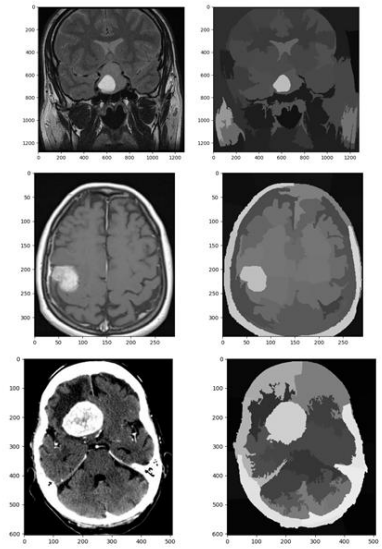


Fig. 7. Segmented MRI images.

3.5 Ensemble model

The ensemble learning system's strength and effectiveness allow it to enhance models' prediction performance. In recent years, ensemble learning has gained prominence as a research trend, which has increased the number of studies utilizing ensemble learning across several application domains.

The idea of ensemble learning is to blend different models' outputs to get a more accurate forecast. Ensemble learning has shown to be a powerful technique in several disciplines, providing more robust and reliable forecasts by efficiently combining predictions from multiple models [17]. The proposed ensemble model combines the strengths of three distinct models: VGG-16, InceptionV3 and ResNet50 to enhance the overall performance. An ensemble model has been used to classify the grade of Brain Tumor (BT) based on the severity of the disease.

We employed a technique for ensemble learning called hard voting or majority voting. It is as easy as tallying up each base model's predictions and choosing the class with the highest number of votes to be the final prediction. It works well in classification work where the classes are distinct and incompatible with one another. The grade predicted by the highest number of models is chosen as the final grade for a particular image using the hard voting approach. For instance, if two of the three models predict a grade of 0 while the other model predicts a grade of 1, the grade with the highest prediction is the final grade. The hard voting ensemble approach is shown in Fig. 3.

4 Results and Discussion

This section will present the quantitative and qualitative results of the experiments. Tables and figures will summarize the performance of each model across different evaluation metrics. We will analyze the impact of data augmentation, hyperparameter tuning, and architectural choices on the segmentation accuracy. A detailed discussion will interpret the results and explain the observed performance differences between the models. We will highlight the strengths and limitations of each architecture and discuss potential areas for future improvement.

The objective of the ensemble model is to classify images as Healthy, Doubtful, Minimal, Moderate, and Severe [14, 15]. Each model in the ensemble brings its unique architectural characteristics and learned features to the diverse and complementary strengths of each model. This approach enhances the overall robustness and accuracy of the classification system [16].

We used the 7,434 BT image dataset for the final training and testing of the model. The 1,656 images were randomly chosen from the BT image dataset for the final model evaluation. The accuracy archived of our suggested ensemble model is 90%. While ResNet, with an accuracy of 68%, VGG-16, with an accuracy of 77% and Inception-V3, with an accuracy of 64%, the ensemble model surpasses them with an accuracy of 90% and is listed in Table 1. Furthermore, in Fig. 9, we show the BT training loss ROC curve.

Table 1. Comparison of accuracy of other models with ensemble model

Model	ResNet50	VGG-16	Inception-V3	Ensemble Model
Accuracy	68%	77%	64%	**90%**

The primary work objective prioritizes higher accuracy, the ensemble model emerges as a promising choice due to its superior performance. However, it is crucial to weigh this decision against other factors, such as computational complexity, interpretability and real-time requirements.

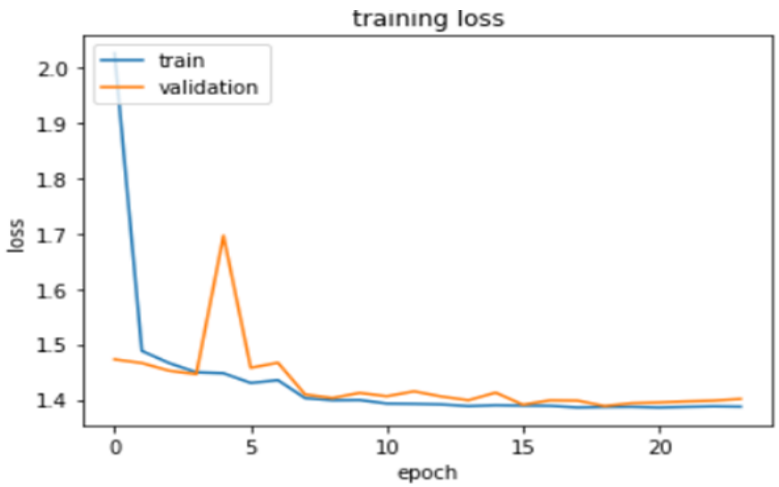


Fig. 9. ROC curve for BT training loss

5 Conclusion and Future Scope

The Brain Tumor (BT) work is dedicated to the development of an ensemble model with the primary objective of classifying the severity of BT based on medical images. This ensemble model, a collaborative effort of VGG-16, Inception-V3 and ResNet50, harnesses the strengths of these advanced architectures to deliver predictions that are accurate and reliable. By leveraging each model's distinctive features and capabilities, the ensemble ensures a comprehensive approach to severity classification.

The work has great potential to improve medical practitioners' efficiency by facilitating quick decisions and prompt actions based on real-time model predictions. The future evolution of the Brain Tumor work is poised to advance significantly with a strategic focus on real-time implementation. The integration of real-time capabilities will empower medical professionals to obtain immediate results during patient examinations, revolutionizing the diagnostic process for Brain Tumor.

This paper will conclude by summarizing the findings of the comparative study. We will identify the best-performing model for brain MRI segmentation based on the evaluation metrics and discuss its implications for clinical applications. We will also identify future research directions, such as exploring more advanced architectures, developing more robust data augmentation techniques, and addressing challenges related to data scarcity and variability.

References

1. T. P. Latchoumi S. Ranjeeth and P. V. Paul: A Survey on Predictive Models of Learning Analytics. In Procedia Computer Science 167(5), pp. 37-46 (2020). https://doi.org/10.1016/j.procs.2020.03.180
2. Sun-Woo Pi, Byoung-Dai Lee, Mu Sook Lee & Hae Jeong Lee: Ensemble deep-learning networks for automated osteoarthritis grading in knee X-ray images.Scientific Reports 13(22887), 1-17 (2023).https://doi.org/10.1038/s41598-023-50210-4
3. https://www.snaplogic.com/machine-learning-showcase/image-recognition-inception-v3
4. https://www.ahsmp.com/alleviating-knee-osteoarthritis-how-physiotherapy-can-make-a-difference/
5. https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758
6. Kannan, S., Sathya, A., Ramathilagam, S., & Devi, R. Novel segmentation algorithm in segmenting medical images. J. Syst. Softw.. 2010; 83. https://doi.org/10.1016/j.jss.2010.07.036
7. Voronin, V., Marchuk, V., Semenishchev, E., Cen, Y., & Agaian, S. Medical image segmentation using 3D MRI data. . 2017; 10221. https://doi.org/10.1117/12.2262857
8. Akan, T., Oskouei, A., Alp, S., & Bhuiyan, M. Brain magnetic resonance image (MRI) segmentation using multimodal optimization. Multimedia Tools and Applications. 2024 https://doi.org/10.1007/s11042-024-19725-4
9. Nyo, M., Mebarek-Oudina, F., Hlaing, S., & Khan, N. Otsu's thresholding technique for MRI image brain tumor segmentation. Multimedia Tools and Applications. 2022; 81. https://doi.org/10.1007/s11042-022-13215-1
10. Hosny, K., Khalid, A., Hamza, H., & Mirjalili, S. Multilevel segmentation of 2D and volumetric medical images using hybrid Coronavirus Optimization Algorithm. Computers in Biology and Medicine. 2022; 150. https://doi.org/10.1016/j.compbiomed.2022.106003
11. Thaha, M., Pradeep, K., Kumar, M., Murugan, B., Dhanasekeran, S., Vijayakarthick, P., & Selvi, S. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. Journal of Medical Systems. 2019; 43. https://doi.org/10.1007/s10916-019-1416-0
12. Wadhwa, A., Bhardwaj, A., & Verma, V. A review on brain tumor segmentation of MRI images.. Magnetic resonance imaging. 2019; 61. https://doi.org/10.1016/j.mri.2019.05.043
13. Soomro, T., Zheng, L., Afifi, A., Ali, A., Soomro, S., Yin, M., & Gao, J. Image Segmentation for MR Brain Tumor Detection Using Machine Learning: A Review. IEEE Reviews in Biomedical Engineering. 2022; 16. https://doi.org/10.1109/RBME.2022.3185292
14. Bennai, M., Guessoum, Z., Mazouzi, S., Cormier, S., & Mezghiche, M. A stochastic multi-agent approach for medical-image segmentation: Application to tumor segmentation in brain MR images. Artificial intelligence in medicine. 2020; 110. https://doi.org/10.1016/j.artmed.2020.101980
15. Mazurowski, M., Dong, H., Gu, H., Yang, J., Konz, N., & Zhang, Y. Segment Anything Model for Medical Image Analysis: an Experimental Study. Medical image analysis. 2023; 89. https://doi.org/10.1016/j.media.2023.102918