The UNet Model for Multimodal Brain Tumor Classification and Segmentation

R. Prathiba
Department of Computer Science & Engineering, PES Institute of Technology and Management, Shimoga, Karnataka, India | Visvesveraya Technological University, Belagavi, Karnataka, India
prathibhar@sjcit.ac.in (corresponding author)

B. S. Sunitha
Department of Computer Science Engineering (Data Science), PES Institute of Technology and Management, Shimoga Karnataka, India | Visvesveraya Technological University, Belagavi, Karnataka, India
sunitha.bs@pestrust.edu.in

Received: 7 April 2025 | Revised: 20 May 2025, 26 May 2025, and 4 June 2025 | Accepted: 6 June 2025

 DOI: https://doi.org/10.48084/etasr.11349

ABSTRACT

MRI manual detection relies on the ability and experience of the physicians. This study presents a brain tumor segmentation and identification system to address this issue and experiments on the BraTS dataset which provides four types of scans, i.e. T1, T2, T1CE, and FAIR, allowing the detection of different class labels and establishing the ground truth for brain tumor segmentation. UNet was used to develop a fully automated approach for the segmentation of glioma on preoperative MRI scans. Images were preprocessed and then classified using the UNet CNN.

Keywords-brain tumor; BRaTS dataset; CNN; MRI; multimodal; Unet

I. INTRODUCTION

Malignant brain tumors are a major cause of death around the world, with the 5-year survival rate being 34% for men and 36% for women. The World Health Organization (WHO) has stated that brain tumors kill 120,000 people every year. CNNs have been widely used for image classification and facilitate diagnosis [1-3]. Deep CNNs extract significant features. Cascade CNN models can handle MRI data well to detect various diseases. However, current fully automated CNN-based segmentation methods still have limitations. The proposed system aims to help radiologists by reducing the number of false positives and speeding up diagnosis, especially for gliomas and other difficult tumors, by automating segmentation, finding tumor areas in MRI scans, and trying to reach the same level of accuracy as an expert.

In [4], the BraTS-10 dataset was used with a 3D CNN and UNet, which achieved dice scores of 0.750 (for expanding tumor), 0.906 (for the entire tumor), and 0.846 (for the core). CNNs are better than traditional methods for biomedical imaging, but there are still challenges with generalizability, ethics, and computation resources. In [5], filters and segmentation methods, such as thresholding, were used along with feature fusion. The results were very accurate, but this model did not work well on all machines and patients. A deep LSTM model [6] showed high similarity scores, but it was not compared to other models or a wide range of datasets. In [7], different ways to segment data were explored, highlighting the importance of clear, fair, and annotated data. AI has made it easier to diagnose gliomas but has issues with ethics, privacy, and obtaining regulatory approval [8]. In [9], GANs and ensemble 3D models were explored, and the need for real time and high accuracy was highlighted. Quantitative methods [10] show how hard it is to segment and ask for better DL methods. A review from 2019 to 2023 discussed the progress and clinical limitations of AI and hybrid segmentation [11]. In [12], a 2D CNN was successful in classification but did not detail the tumor's characteristics. The study in [13] focused on the role of AI in detection using open-source datasets, but it did not cover all the different types of clinical data.

DL models, such as CNNs [14], work well for splitting up strange tumors, but they need large and labeled datasets. The system in [15] was 98.89% accurate, but it was not tested in the real world. In [16], BraTs 2012 was used with bilateral and Gabor filters, achieving 88% Dice score, but this model requires more validation on different datasets. In [17], transfer learning (VGG16) was used, achieving 98.69% accuracy, but there may still be bias and problems with the dataset. In [18], ResNet12 and MobileNet were compared for their speed and accuracy, but this study lacked a comprehensive validation or robustness analysis. Brain-DeepNet [19] achieved 96.3% accuracy on MRI with data augmentation, but this study did not describe how it could be used in the real world or how explainable the model is. Deep CNNs [20] did better than other

models on public datasets, but there were concerns about how well they could be used in other situations or their ethics. VGG-16 did better than InceptionV3 at classifying tumors [21], but the dataset was small. ML-based MRI segmentation helps in finding problems early, but it does not work well in rare cases and is hard to understand [22]. DL models such as VGG16 and ResNet152 can help in diagnosing glaucoma but require long time and resources [23]. In [24], CNN-Net and U-Net to DenseNet were compared, finding that they both performed very well but did not examine the computational resources they needed. Finally, in [26], DL models such as AlexNet, VGG16, and ResNet-50 were up to 99.8% accurate but were not tested in real-life clinical settings.

This study aimed to address the problems of current methods, such as low accuracy, limited generalization, high computational cost, lack of interpretability, and not enough testing on real-world clinical datasets. A UNET-based CNN architecture was used to automatically segment brain tumors, achieving more than 90% accuracy and fewer false positives. This architecture is efficient in terms of computational resources, focuses on gliomas, and aims to help radiologists in real time by automating segmentation and reducing human errors. This study also suggests a clear workflow to meet clinical needs.

II. PROPOSED METHOD

Figure 1 describes a sequential method for the detection of brain tumors. Before the training stage, data collection, preprocessing, and data augmentation are applied. The second stage involves preprocessing the images and training the UNet and CNN models. In the final step, the models are tested for detecting and classifying tumors.

A. Data Collection

This study used 253 MRI scans from the publicly available BraTS dataset [27, 28], categorized into 155 tumorous (malignant) (61%) and 98 healthy (benign) images (39%). Figure 2 shows a sample input image. Tumors are split into multimodal categories: uh, t1, t1, and t2.

B. Data Augmentation

This stage aims to expand the data using reflection, gaussian blur, histogram equalization, rotation, linear transformation, and horizontal and vertical shifts. The augmented dataset includes 2065 images, with 1085 instances (53%) and 980 negative samples (47%).

C. Image Preprocessing

In this stage, the images were converted to grayscale to reduce the impact of noise. Each pixel in an image had an intensity level of 1 or 0. This technique provides accurate data and aids the segmentation procedure. Applying a filter after grayscale conversion can eliminate noise.

D. Data Split

The dataset was divided into 70% for training, 20% for validation, and 10% for testing.

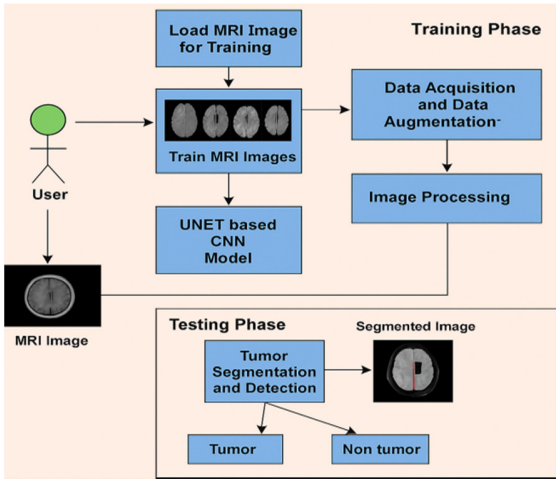

Fig. 1. System architecture.

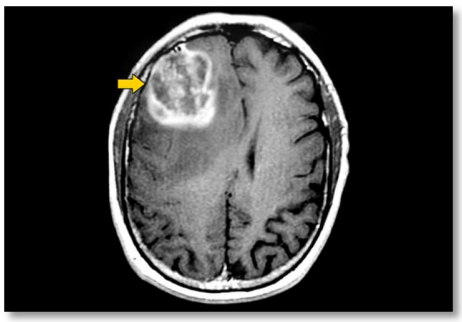

Fig. 2. Sample input image.

E. Convolutional Neural Network

Classification is crucial for tumor identification in medical imaging. CNNs are effective in identifying even the smallest details. ConvNets, prioritize distinct parts of an input image, require less preprocessing, and can learn filters and features with training. They can adapt to the image collection, utilizing fewer parameters and weights for reuse. Input images to the neural network have a shape of 240, 240, 3. The CNN consists of the following:

- A zero-padding layer with parameters set to (2, 2).
- The convolutional layer has a size of 7×7 and 32 filters. To expedite processing, a batch normalization layer is used to standardize pixel values.
- ReLU activation function is used.
- In the max-pooling layer, the f value is 4 and the s value is also 4.
- A flatten layer is used to convert the three-dimensional matrix into a one-dimensional vector.
- Finally, a densely packed, fully linked layer contains a single neuron activated with a sigmoid function.

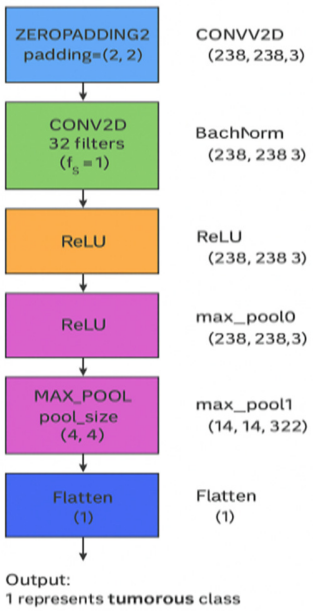

Fig. 3. CNN architecture.

F. UNet Architecture

Medical image segmentation is essential because it enables models to simultaneously perceive large images and small details. U-Net is widely used because of its exceptional performance across a range of tasks. To combine coarse contextual features with fine-grained local visual features, CNN increases the sampling ratio while decreasing feature sampling. Higher output resolution is achieved by adding more layers to a standard contracting network through the U-Net model. Combining upsampled output with high-resolution features allows for localization, and the convolutional layer learns to generate output with improved precision. The encoder (contracting route) and decoder (expanding path) are the two primary components of UNet. The decoder uses upsampling and combines features from the encoder to improve segmentation details, while the encoder uses convolutional layers and down-sampling to acquire context.

1) Encoder (Contracting Path)

As the feature depth increases, the encoder gradually decreases the input's spatial dimensions.

- Convolutional layers: To decrease the spatial dimensions, each encoder layer employs max-pooling, an activation function (usually ReLU), and a pair of convolution operations. The convolution operation is:

$$Input\_X_{(l+1)} = ReLU(Conv(Input_X)_l, W_l) + b_l \quad (1)$$

where $Input\_X_{(l)}$ is the input feature map, W_l is the weight matrix, also called filters, and b_l is the bias.

- Max pooling layer: By maximizing each pooling window, the feature map is down-sampled. Max-pooling is given by:

$$X_{Pool} = Max(Input_X) \quad (2)$$

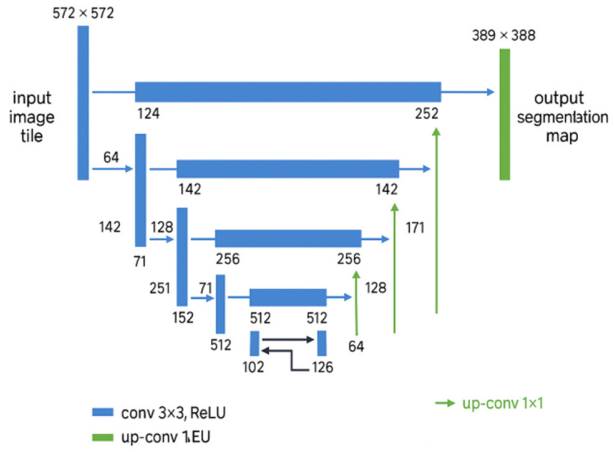

Fig. 4. Basic UNet architecture.

2) Bottleneck

Positioned between the encoder and decoder, the bottleneck layer employs convolutions to record the most condensed and detailed feature representation.

3) Decoder (Expanding Path)

To recreate the image segmentation mask, the decoder uses skip connections to combine coarse features from previous layers with coarser ones from the bottleneck and then up-sample the feature maps. The transposed convolution (up-sampling) expands the spatial resolution of the feature maps:

$$X_{upsample} = ReLU(Conv^U(Input_X, W_l) + b_l \quad (3)$$

where $Conv^U$ is the transposed or up-sampled convolution, also called deconvolution. To aid in the recovery of spatial information, skip connections merge encoder and decoder characteristics. The merging skip connections are given by:

$$X_{merge-skip} = ConCat(X_{upsample} + X_{skip-encode}) \quad (4)$$

where $X_{upsample}$ is the upsampled feature map and $X_{skip-encode}$ is the corresponding feature map from the encoder path.

4) Output Layer

The final layer uses a convolution to map the combined features into the desired number of segmentation classes:

$$X_{output\_layer} = SoftMax\left(Conv_{1x1}(X_{merge\_Skip})\right) \quad (5)$$

where $Conv_{1x1}$ is a convolution with a 1×1 kernel, and $SoftMax$ is an activation function used to output class probabilities.

The loss function commonly used for brain tumor segmentation is the Dice loss or a combination of Dice loss and Cross-Entropy loss.

$$DiceLoss = 1 - \frac{2\sum_i PP_i GT_i}{\sum_i PP_i + \sum_i GT_i} \quad (6)$$

where PP_i is the predicted probability for pixel i and GT_i is the Ground Truth label for pixel i.

III. RESULTS AND DISCUSSION

The results in this section come from tests on the BraTS dataset, using it as a benchmark [4, 9, 14]. The proposed UNet-based segmentation method was compared with older methods, such as thresholding, watershed, and a baseline CNN model. Threshold and watershed segmentation are all examples of classical segmentation methods. These methods have problems, such as being sensitive to noise and changes in intensity, not being able to understand the context or generalize across different MRI modalities, and providing many false positives and false negatives. Although it requires much computing power, the proposed UNet-based CNN approach is a better and more accurate way to segment brain images.

The analysis of the CNN and UNet models for brain tumor segmentation focuses on accuracy and loss metrics. The CNN model showed a significant improvement in training accuracy, increasing from 36.2% to 90.05% by the 100th epoch, as shown in Table I and Figure 5. The training accuracy gap narrowed as the training progressed due to better generalization and less overfitting. The model learned effectively and decreased errors on the training set, with a significant drop in training loss from 1.38 in the first epoch to 0.22 by the 100th epoch, as shown in Figure 6.

TABLE I. PERFORMANCE OF CNN

Epochs	Accuracy	Loss
1	0.361991	1.384916
10	0.524887	1.078237
20	0.624434	0.950129
30	0.709821	0.73544
40	0.733032	0.696195
50	0.841629	0.471065
60	0.841629	0.400818
70	0.841629	0.381313
80	0.891403	0.34273
90	0.909502	0.273973
100	0.904977	0.22705

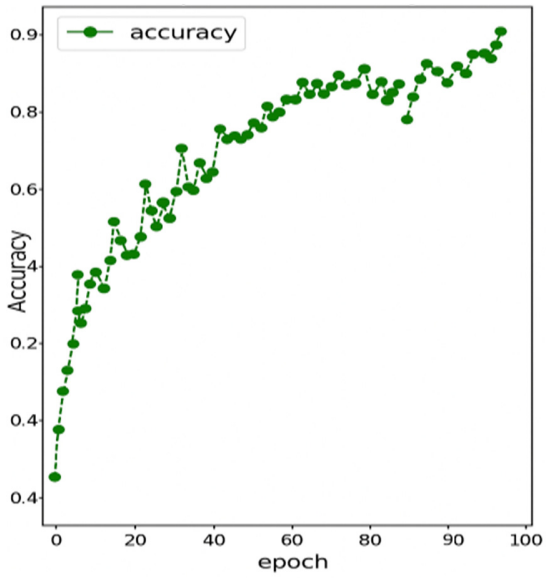

Fig. 5. Accuracy of the CNN model.

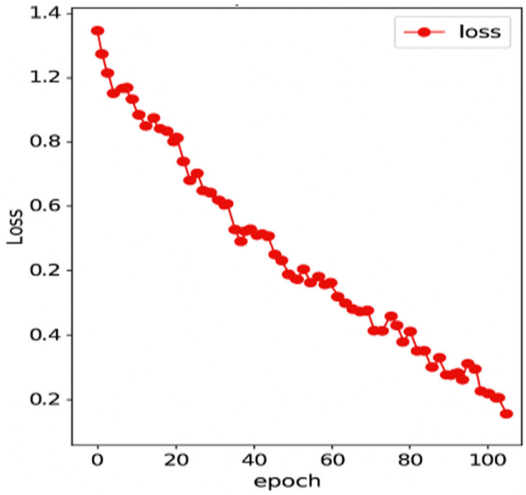

Fig. 6. Loss of the CNN model.

(a) (b) (c)

Fig. 7. (a) Input Image, (b) Threshold segmented image, (c) Watershed segmented image.

Figure 7 shows the visual results of different segmentation techniques applied to brain tumor MRI scans. Figure 7(a) shows the original MRI scan, which shows the brain's structure with visible tumor regions. Figure 7(b) shows the segmentation result using simple intensity-based thresholding. Figure 7(c) shows the watershed segmented image. The watershed algorithm takes into account the topography and boundaries of the image.

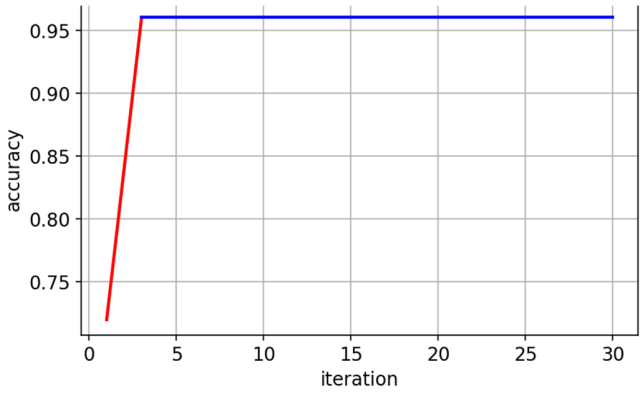

Fig. 8. Accuracy performance of the UNet model.

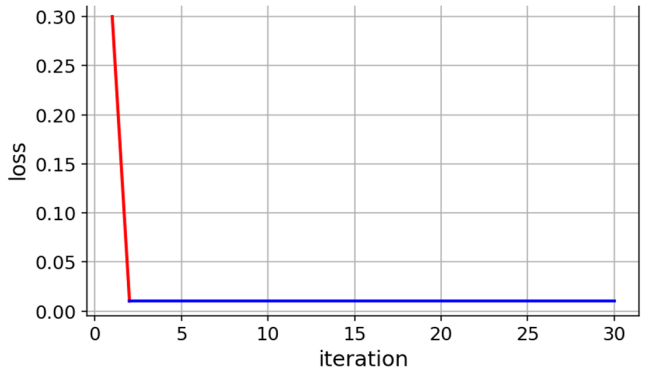

Fig. 9. Loss performance of the UNet model.

As shown in Table II, with 96% accuracy, the proposed UNET-based CNN strikes a balance between segmentation accuracy, training efficiency, and clinical applicability. This approach suits real-time diagnostic use because it provides a compromise between performance and resource requirements, minimizing reliance on large datasets or pre-trained models.

This steady and faster improvement compared to the CNN shows that the model extracted useful information from the training samples, as shown in Figure 8. Figure 9 displays the loss graph, which fluctuated less than the CNN. Figures 8 and 9 demonstrate the significant improvement in training accuracy and loss with UNet over CNN.

TABLE II. COMPARISON ANALYSIS OF BRAIN TUMOR SEGMENTATION METHODS

Method	Study	Model Type	Performance	Remarks
Proposed UNET-based CNN	This work	2D UNET with preprocessing and augmentation	Accuracy: 96% (10 epochs)	Fast convergence, lightweight, accurate segmentation and classification
Traditional ML Methods	[5], [16], [21], [22]	SVM, texture, rule-based	Accuracy: 85–89%	Low generalization, manual feature engineering
Basic CNN / LSTM Models	[1], [3], [6], [19]	CNN, LSTM	Accuracy: 90–94%	Slower training, less spatial precision
Transfer Learning Models	[15], [17], [18], [26]	VGG, ResNet, TL	Accuracy: 95–98%	High accuracy, but computationally intensive
3D / Ensemble UNETs	[4], [9], [14], [24]	3D UNET, Attention UNET	Dice: 0.89–0.92	Accurate but resource-heavy and slower to train

IV. CONCLUSION

Accurate and timely diagnosis of brain tumors remains one of the most critical challenges in neuro-oncology. Manual segmentation of MRI scans is time-consuming and prone to subjectivity, leading to potential diagnostic delays and variability. This study presents an automated, end-to-end brain tumor segmentation and classification framework based on a UNet-enhanced CNN architecture, designed specifically for multimodal MRI images sourced from the widely used BraTS dataset. The study on MRI inputs and tumor segmentation using a UNet-based CNN model demonstrated its effectiveness in reducing misdiagnosis and supporting radiologists. The model achieved 96% accuracy, surpassing traditional methods and complex models such as 3D UNet and ResNet. This study highlights the potential of this model in clinical settings, with future plans to add cross-slice correlation and 3D volumetric analysis.

REFERENCES

[1] M I. Sharif and K. Kaushik, "Recent Advancements in Brain Tumor Segmentation and Classification using Deep Learning: A Review," *International Journal of Engineering Research and*, vol. 8, no. 12, Dec. 2019, https://doi.org/10.17577/IJERTV8IS120190.

[2] Md. F. Ahamed *et al.*, "A review on brain tumor segmentation based on deep learning methods with federated learning techniques," *Computerized Medical Imaging and Graphics*, vol. 110, Dec. 2023, Art. no. 102313, https://doi.org/10.1016/j.compmedimag.2023.102313.

[3] M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, "A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned," *Magnetic Resonance Imaging*, vol. 61, pp. 300–318, Sep. 2019, https://doi.org/10.1016/j.mri.2019.05.028.

[4] M. Ali, S. O. Gilani, A. Waris, K. Zafar, and M. Jamil, "Brain Tumour Image Segmentation Using Deep Networks," *IEEE Access*, vol. 8, pp. 153589–153598, 2020, https://doi.org/10.1109/ACCESS.2020.3018160.

[5] J. Amin, M. Sharif, M. Raza, T. Saba, and M. A. Anjum, "Brain tumor detection using statistical and machine learning method," *Computer Methods and Programs in Biomedicine*, vol. 177, pp. 69–79, Aug. 2019, https://doi.org/10.1016/j.cmpb.2019.05.015.

[6] J. Amin, M. Sharif, M. Raza, T. Saba, R. Sial, and S. A. Shad, "Brain tumor detection: a long short-term memory (LSTM)-based learning model," *Neural Computing and Applications*, vol. 32, no. 20, pp. 15965–15973, Oct. 2020, https://doi.org/10.1007/s00521-019-04650-7.

[7] Z. Liu *et al.*, "Deep learning based brain tumor segmentation: a survey," *Complex & Intelligent Systems*, vol. 9, no. 1, pp. 1001–1026, Feb. 2023, https://doi.org/10.1007/s40747-022-00815-5.

[8] S. Khalighi, K. Reddy, A. Midya, K. B. Pandav, A. Madabhushi, and M. Abedalthagafi, "Artificial intelligence in neuro-oncology: advances and challenges in brain tumor diagnosis, prognosis, and precision treatment," *npj Precision Oncology*, vol. 8, no. 1, Mar. 2024, Art. no. 80, https://doi.org/10.1038/s41698-024-00575-0.

[9] S. G. Domadia, F. N. Thakkar, and M. A. Ardeshana, "Recent advancement in learning methodology for segmenting brain tumor from magnetic resonance imaging -a review," *Multimedia Tools and Applications*, vol. 82, no. 22, pp. 34809–34845, Sep. 2023, https://doi.org/10.1007/s11042-023-14857-5.

[10] T. Magadza and S. Viriri, "Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art," *Journal of Imaging*, vol. 7, no. 2, Jan. 2021, Art. no. 19, https://doi.org/10.3390/jimaging7020019.

[11] Y. Zakeri, B. Karasfi, and A. Jalalian, "A Review of Brain Tumor Segmentation Using MRIs from 2019 to 2023 (Statistical Information, Key Achievements, and Limitations)," *Journal of Medical and Biological Engineering*, vol. 44, no. 2, pp. 155–180, Apr. 2024, https://doi.org/10.1007/s40846-024-00860-0.

[12] S. Saeedi, S. Rezayi, H. Keshavarz, and S. R. Niakan Kalhori, "MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, Jan. 2023, Art. no. 16, https://doi.org/10.1186/s12911-023-02114-6.

[13] A. Batool and Y. C. Byun, "Brain tumor detection with integrating traditional and computational intelligence approaches across diverse imaging modalities - Challenges and future directions," *Computers in Biology and Medicine*, vol. 175, Jun. 2024, Art. no. 108412, https://doi.org/10.1016/j.compbiomed.2024.108412.

[14] V. N. V. L. S. Swathi, K. Sinduja, V. Ravi Kumar, A. Mahendar, G. V. Prasad, and B. Samya, "Deep learning-based brain tumor detection: An MRI segmentation approach," *MATEC Web of Conferences*, vol. 392, 2024, Art. no. 01157, https://doi.org/10.1051/matecconf/202439201157.

[15] M. Agarwal, G. Rani, A. Kumar, P. K. K, R. Manikandan, and A. H. Gandomi, "Deep learning for enhanced brain Tumor Detection and classification," *Results in Engineering*, vol. 22, Jun. 2024, Art. no. 102117, https://doi.org/10.1016/j.rineng.2024.102117.

[16] Z. U. Rehman, M. S. Zia, G. R. Bojja, M. Yaqub, F. Jinchao, and K. Arshid, "Texture based localization of a brain tumor from MR-images by using a machine learning approach," *Medical Hypotheses*, vol. 141, Aug. 2020, Art. no. 109705, https://doi.org/10.1016/j.mehy.2020.109705.

[17] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning," *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 757–775, Feb. 2020, https://doi.org/10.1007/s00034-019-01246-3.

[18] A. Al-Shahrani, W. Al-Amoudi, R. Bazaraah, A. Al-Sharief, and H. Farouquee, "An Image Processing-based and Deep Learning Model to Classify Brain Cancer," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15433–15438, Aug. 2024, https://doi.org/10.48084/etasr.7803.

[19] S. U. Habiba, Md. K. Islam, L. Nahar, F. Tasnim, M. S. Hossain, and K. Andersson, "Brain-DeepNet: A Deep Learning Based Classifier for Brain Tumor Detection and Classification," in *Intelligent Computing & Optimization*, vol. 569, P. Vasant, G. W. Weber, J. A. Marmolejo-Saucedo, E. Munapo, and J. J. Thomas, Eds. Springer International Publishing, 2023, pp. 550–560.

[20] K. Bathe, V. Rana, S. Singh, and V. Singh, "Brain Tumor Detection Using Deep Learning Techniques," in *Proceedings of the 4th International Conference on Advances in Science & Technology (ICAST2021)*, 2021, https://doi.org/10.2139/ssrn.3867216.

[21] J. Shetty, M. Shenoy, V. R. Das, M. Mishra, R. Prasad, and S. Seth, "Classification Of Brain Images For Identification Of Tumors," in *2022 IEEE Bombay Section Signature Conference (IBSSC)*, Mumbai, India, Dec. 2022, pp. 1–5, https://doi.org/10.1109/IBSSC56953.2022.10037548.

[22] G. Hemanth, M. Janardhan, and L. Sujihelen, "Design and Implementing Brain Tumor Detection Using Machine Learning Approach," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, Apr. 2019, pp. 1289–1294, https://doi.org/10.1109/ICOEI.2019.8862553.

[23] L. Nahar, M. S. Hossain, P. Das, T. Alam, and K. Andersson, "A Deep Learning-Based Ophthalmologic Approach for Retinal Fundus Image Analysis to Detect Glaucoma," in *Proceedings of the Third International Conference on Trends in Computational and Cognitive Engineering*, vol. 348, M. S. Kaiser, K. Ray, A. Bandyopadhyay, K. Jacob, and K. S. Long, Eds. Springer Nature Singapore, 2022, pp. 519–532.

[24] S. Bagyaraj, R. Tamilselvi, P. B. Mohamed Gani, and D. Sabarinathan, "Brain tumour cell segmentation and detection using deep learning networks," *IET Image Processing*, vol. 15, no. 10, pp. 2363–2371, Aug. 2021, https://doi.org/10.1049/ipr2.12219.

[25] M. Nazir, S. Shakil, and K. Khurshid, "Role of deep learning in brain tumor detection and classification (2015 to 2020): A review," *Computerized Medical Imaging and Graphics*, vol. 91, Jul. 2021, Art. no. 101940, https://doi.org/10.1016/j.compmedimag.2021.101940.

[26] V. K. Dhakshnamurthy, M. Govindan, K. Sreerangan, M. D. Nagarajan, and A. Thomas, "Brain Tumor Detection and Classification Using Transfer Learning Models," in *Engineering Proceedings*, Feb. 2024, https://doi.org/10.3390/engproc2024062001.

[27] B. H. Menze *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, Jul. 2015, https://doi.org/10.1109/TMI.2014.2377694.

[28] "Brain MRI Images for Brain Tumor Detection." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection.