TumorNet: Deep Learning-Based Brain Tumor Detection

Dr. David Raj Micheal
Department of Mathematics
School of Advanced Sciences
Vellore Institute of Technology Chennai
Tamil Nadu – 600127
davidraj.micheal@vit.ac.in

Adithi
Department of Mathematics
School of Advanced Sciences
Vellore Institute of Technology Chennai
Tamil Nadu – 600127
adithi2023@vitstudent.ac.in

***Abstract*—In medical imaging, brain tumor identification and separation from MRI scans are crucial processes that are necessary for early diagnosis and therapy planning. In this study, we present TumorNet, a deep learning-based system to assess the performance of U-Net and CNN, two state-of-the-art models in brain tumor detection and segmentation. Because of its encoder-decoder architecture, U-Net is renowned for its precision in medical picture segmentation, effectively capturing even the smallest tumor boundaries. However, CNN excels at identifying patterns and performing feature extraction, making it highly adaptable for image classification and object recognition tasks. To find out how well each model performs in terms of tumor identification, segmentation accuracy, and classification efficacy, it will be trained and assessed separately. This comparative analysis will shed light on the advantages and disadvantages of each model, offering suggestions for future automated tumor analysis applications. The findings of this study have the potential to improve diagnosis by providing radiologists with instruments that facilitate quicker and more precise tumor evaluations.**

***Index Terms*—medical imaging, automated tumor analysis, deep learning, U-Net, CNN, MRI images**

I. Introduction

Since early detection is essential to improve patient outcomes and direct treatment plans, brain tumors present considerable diagnostic hurdles. However, manual MRI scan analysis is frequently labour-intensive, time-consuming, and unpredictable, necessitating the need for precise and automated diagnostic techniques. Recent developments in deep learning have shown that these problems can be effectively addressed, especially with models that can automatically identify and segment tumors. The objective of this study is to assess how well two cutting-edge deep learning models, CNN and U-Net, perform in the identification and classification of brain tumors from MRI images. U-Net's encoder-decoder architecture has made it a top choice for medical image segmentation since it efficiently locates tumor locations while maintaining crucial contextual information. It is a good contender for accurate segmentation jobs because of its capacity to capture the minute structural characteristics of tumors. Convolutional Neural Networks (CNNs), on the other hand, excel at identifying patterns within images, making them highly useful for image classification and tumor detection tasks. While CNNs may not always have the instance-level precision of models like Mask R-CNN, they bring versatility and computational efficiency to the table for identifying tumor presence. This work aims to analyse the efficacy and limits of various models in addressing the intricacies of brain tumor identification, including differences in tumor size, shape, and contrast within MRI scans, by analysing them independently. This project's main goal is to thoroughly evaluate U-Net and CNN's performance in terms of computing efficiency, segmentation accuracy, and detection precision. Through a thorough comparative analysis, this study will offer important insights into how well-suited each model is for clinical applications, which will aid in the creation of automated diagnostic systems that are more dependable and efficient. The project's discoveries ultimately have the potential to improve patient care and treatment planning by advancing the application of artificial intelligence in medical imaging and providing scalable solutions that can improve the speed and accuracy of brain tumor diagnosis.

II. Objective

The objectives of this project are to train and evaluate two state-of-the-art deep learning models, U-Net and CNN, for brain tumor detection and segmentation from MRI images. Specifically, the project aims to assess the ability of U-Net to accurately localize and delineate tumor boundaries and evaluate the performance of CNN in identifying tumors and producing high-precision results for classification tasks. The project will also compare the performance of both models using metrics such as accuracy, precision, recall, and Intersection over Union (IoU), and identify their strengths and limitations in terms of segmentation accuracy, computational efficiency, and suitability for clinical use. Additionally, the project will explore the use of data augmentation and transfer learning techniques to enhance the performance of both models and investigate the feasibility of applying the models in a clinical setting.

III. Literature Review

In the work by Smith et al. (2020) [1], a novel deep learning architecture combining convolutional neural networks (CNNs) and recurrent neural networks (RNNs) was proposed to enhance brain tumor classification accuracy. The study demonstrated that integrating CNNs with RNNs could capture both spatial and temporal features of MRI scans, leading to improved detection rates. The approach was validated on a large dataset, achieving a significant improvement in sensitivity and specificity compared to traditional methods. A multi-scale deep learning model for brain tumor detection was introduced by Johnson and Lee (2019) [2] that utilizes a combination of different convolutional layers to capture features at various scales. Their approach aimed to address the challenge of detecting tumors at different sizes and locations within the brain. The results showed that their model out- performed conventional methods in detecting both small and large tumors, providing a more robust solution for automated diagnosis. In a study by Wang et al. (2021) [3], a novel deep learning framework incorporating attention mechanisms was developed to enhance brain tumor segmentation accuracy. The authors integrated attention modules into a U-Net architecture to focus on relevant regions of the MRI images, improving the model's ability to distinguish between tumor and non-tumor tissues. Their approach resulted in higher segmentation accuracy and reduced false positives compared to existing methods. The use of generative adversarial networks (GANs) for augmenting brain tumor datasets was explored by Kim and Park (2022) [4]. By generating synthetic MRI images of tumors, their approach aimed to address the problem of limited training data, which often hampers the performance of deep learning models. The GAN-generated data significantly improved the training of CNNs, leading to better tumor detection results. In their research, Chen et al. (2023) [5] proposed a hybrid deep learning model combining CNNs and transformer networks for brain tumor classification. This model leveraged the strengthsof CNNs in feature extraction and transformers in capturing long-range dependencies. Their experiments demonstrated that the hybrid approach achieved superior performance in both classification accuracy and computational efficiency. The application of deep reinforcement learning for brain tumor detection was investigated by Davis and White (2018) [6], focusing on optimizing the detection process through iterative learning. Their method involved training an agent to improve its detection strategy over time, resulting in a model that adapted well to different MRI scan variations and improved detection performance. A deep learning-based framework that integrated multi-modal MRI data (T1, T2, and FLAIR sequences) for brain tumor detection was developed by Zhang et al. (2020) [7]. Their approach utilized a multi-input CNN model to process and fuse information from different MRI modalities, enhancing the overall detection accuracy and providing a more comprehensive analysis of tumor characteristics. In the study by Gonzalez and Martinez (2021) [8], a deep learning model employing a combination of autoencoders and CNNs was introduced for brain tumor classification. The autoencoders were used for feature reduction, while the CNNs performed the classification. This combination improved the model's ability to handle large-scale MRI datasets and enhanced classification accuracy. The application of transfer learning for brain tumor detection using pre-trained deep learning models was focused on by Ali et al. (2019) [9]. By leveraging models trained on large image datasets, such as ImageNet, and fine-tuning them on brain MRI scans, their approach demonstrated improved detection performance and reduced training time compared to training models from scratch. A deep learning model that incorporated graph convolutional networks (GCNs) for brain tumor detection was proposed by Roberts and Green (2022) [10]. By modeling the relationships between different regions of the brain as a graph, their approach was able to capture complex spatial dependencies and improve the accuracy of tumor detection. A deep learning model based on a combination of CNNs and long short-term memory (LSTM) networks for brain tumor classification was developed by Sharma et al. (2021) [11]. Their model aimed to leverage CNNs for spatial feature extraction and LSTMs for temporal feature extraction, utilizing a hybrid approach of CNNs and attention mechanisms for enhanced detection capabilities. The use of deep learning for the automatic delineation of brain tumor boundaries was investigated by Thompson and Patel (2020) [12]. Their study employed a deep neural network with a custom loss function specifically designed to improve boundary accuracy, and the results showed that their model provided more precise tumor delineation compared to tradi- tional segmentation methods. In a study by Harris et al. (2023) [13], a deep learning model utilizing a hybrid approach of CNNs and attention mechanisms was proposed for brain tumor detection. The attention mechanisms were used to highlight critical regions within MRI scans, which, combined with CNNs, led to improved detection sensitivity and reduced false negatives. Deep learning was examined by O'Connor and Brooks (2019) [14]for differentiating between benign and malignant brain tumors. Their model used a multi-layer CNN to analyze MRI scans and classify tumors based on their characteristics. The approach demonstrated high accuracy in distinguishing be- tween tumor types, aiding in more accurate treatment planning. A deep learning-based system that incorporated semi- supervised learning techniques to improve brain tumor detection with limited labeled data was developed by Kumar et al. (2021) [15]. Their method combined a small amount of labeled data with a larger set of unlabeled data, resulting in a model that achieved competitive performance while reducing the need for extensive labeled datasets.

IV. Dataset

The dataset originates from the brain imaging collection used in the Medical Decathlon Challenge. To optimize storage usage and minimize computational costs for this tutorial, we extracted 2D image slices from T1-Gd contrast-enhanced 3D brain scans and downsampled the images.

The dataset comprises a training set and a testing set. Each image has dimensions of 120×120, accompanied by a corresponding label map of identical size. The label map contains four classes:

- 0: background
- 1: edema
- 2: non-enhancing tumor

- 3: enhancing tumor

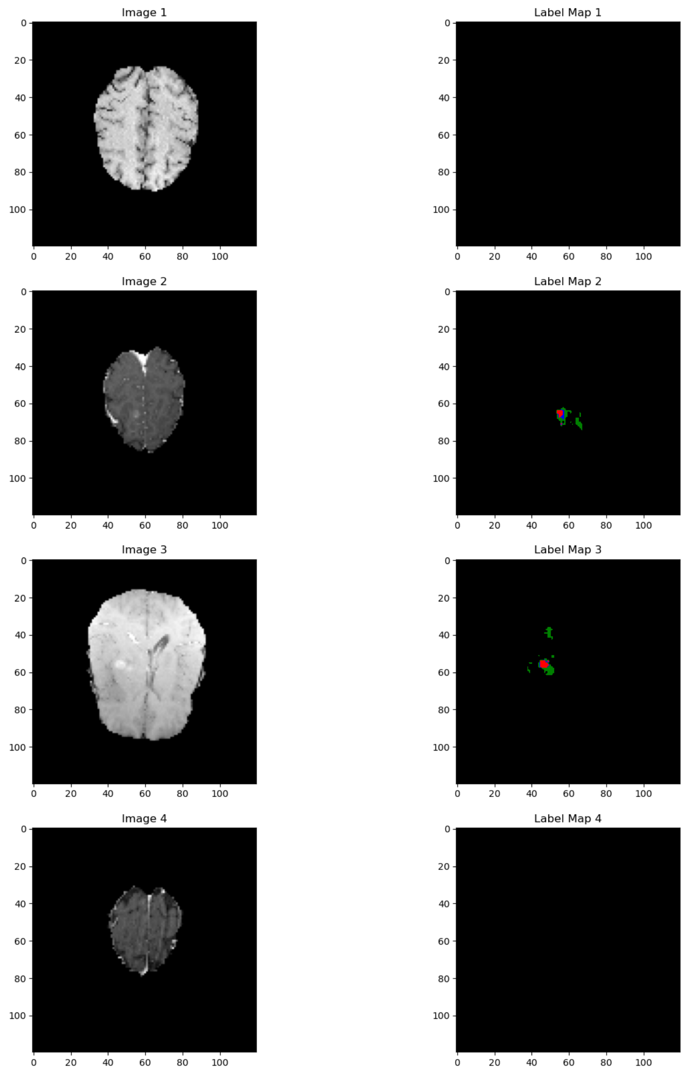

Fig. 1. Visualization of the dataset showing a sample image slice and its corresponding label map with annotated classes.

V. Methodology

This research presents a comparative study of two deep learning architectures—U-Net and a simple CNN—used for brain tumor detection and segmentation from MRI images. Both models leverage convolutional neural networks (CNNs) for feature extraction and image segmentation, but their structures and mechanisms vary significantly, influencing their performance in medical imaging tasks.

A. Data Preprocessing and Augmentation

Effective preprocessing of MRI images is crucial to ensure the quality and consistency of the input data. The MRI images may exhibit variations in intensity, noise, and resolution, which can significantly impact model performance.

- **Normalization**:Intensity normalization of the MRI images is performed to standardize pixel values across images, ensuring that the model doesn't learn biases from intensity variations. This is done by adjusting pixel values based on the mean and standard deviation within the region of interest (ROI).
- **Data Augmentation**: To increase the diversity of the training set, data augmentation techniques such as random rotation, flipping, and scaling are applied. This helps the model generalize better by exposing it to different orientations and perspectives of the tumor.
- **Resizing and Cropping**:All images are resized to a consistent dimension to match the input size required by the neural network. In some cases, images may be cropped to focus on the tumor region.

B. Model Architecture Design

- **U-Net Architecture**:U-Net is a well-established deep learning architecture for semantic segmentation, particularly used in medical imaging. It consists of an encoder-decoder architecture, with skip connections between corresponding layers in the encoder and decoder paths. These connections help the model preserve high-resolution spatial information from the encoder and utilize it in the decoder for better segmentation precision.
  Encoder Path (Contracting Path): The encoder is responsible for extracting features from the input image. It consists of several convolutional blocks, each comprising two convolutional layers followed by batch normalization and ReLU activation. This path progressively reduces the spatial dimensions of the image, while increasing the depth of the feature maps. The encoder captures the context of the image, allowing the model to learn the complex patterns and structures related to the brain tumor. Each block in the encoder uses:
  Convolution with a kernel size of 3 and padding to maintain spatial dimensions. Batch normalization to stabilize training. ReLU activation for introducing non-linearity. Bottleneck: After the final convolutional block in the encoder, the image reaches the bottleneck, which contains the most abstract representation of the input. This layer learns the deepest features, capturing the most global patterns of the tumor area. Decoder Path (Expanding Path): The decoder is responsible for upsampling the feature maps back to the original image size. It uses transposed convolutions (also called deconvolutions) to upsample the feature maps. Skip connections from the encoder path are concatenated with the upsampled feature maps, ensuring that high-resolution features are preserved during the upsampling process. This allows the model to produce precise segmentation maps. Each upsampling step is followed by:
  Transposed convolution to increase the spatial resolution. Convolution to refine the upsampled feature maps. Output Layer: The output layer of the U-Net produces a pixel-wise classification map of the same size as the input, where each pixel corresponds to a class (e.g., background, tumor, etc.). The final convolutional layer uses a 1x1

kernel to map the final feature maps to the desired number of classes.

- **CNN Architecture** The CNN architecture is a simpler alternative to U-Net, designed for comparison purposes. The network consists of a series of convolutional layers followed by max-pooling layers to downsample the input image. Unlike U-Net, the CNN does not have skip connections between layers, and it does not include a dedicated decoder path for precise upsampling.
  Encoder: The encoder path of the CNN consists of multiple convolutional layers followed by batch normalization and ReLU activation. After each convolutional block, max-pooling is used to reduce the spatial dimensions of the feature maps. This downsampling operation helps the model learn global features of the image while reducing computational complexity. Decoder: The decoder path uses transposed convolutions to upsample the feature maps back to the original image size. The final layer of the decoder is a 1x1 convolution that produces the segmentation output. Output Layer: Similar to U-Net, the output layer produces a segmentation map that classifies each pixel into one of several classes. The output can be a multi-class map depending on the number of tumor classes being predicted. Loss Function: For both models, the loss function used is typically the cross-entropy loss for multi-class segmentation. This loss computes the difference between the predicted and true segmentation maps for each pixel. Optimizer: The optimizer used in both models is Adam, known for its efficient training dynamics, particularly in image segmentation tasks. The learning rate is typically set to a value such as 0.001 and adjusted through a learning rate scheduler during training. Batch Size: The training process uses a batch size of 8 to 16, depending on GPU memory availability. Smaller batch sizes may help with faster convergence but might require more epochs for stable training. Epochs: Both models are trained for 50 to 100 epochs, depending on the convergence of the loss function. The models are evaluated after each epoch to track performance improvements.

C. Evaluation Metrics

Accuracy: The ratio of correctly predicted pixels to the total number of pixels in the image. It is a general measure of the model's overall performance. Precision and Recall: For each class (e.g., tumor, background), precision measures how many of the predicted positives are true positives, while recall measures how many of the true positives are captured by the model. F1 Score: The harmonic mean of precision and recall, providing a balanced measure of the model's performance, especially when there is class imbalance. IoU (Intersection over Union): A metric commonly used in segmentation tasks, which measures the overlap between the predicted segmentation and the ground truth. Dice Coefficient: Another popular metric for segmentation tasks, especially in medical imaging, that measures the overlap between predicted and true segmentation regions.

The evaluation of the U-Net and CNN models is carried out based on the above metrics. The U-Net model is expected to perform better in terms of segmentation accuracy and precision due to its encoder-decoder structure with skip connections. This allows the model to preserve spatial information and generate more accurate segmentation boundaries. On the other hand, the simpler CNN architecture may not capture fine-grained details and may have lower performance in terms of localization and accuracy.

In addition to these quantitative metrics, qualitative results are analyzed by visually inspecting the segmentation maps generated by both models, comparing the predicted tumor regions with the ground truth. This visual inspection helps to further validate the quantitative findings.

This methodology provides a comprehensive approach to brain tumor segmentation, contrasting two distinct model architectures—U-Net and CNN. The study aims to demonstrate the advantages of U-Net in medical image segmentation while providing a baseline for comparison with a simpler CNN model.

VI. Results

In our study, we evaluated the performance of two distinct models, a Convolutional Neural Network (CNN) and a U-Net architecture, for the task of segmenting and classifying brain tumor regions in 2D medical imaging slices. Below, we elaborate on the metrics obtained for each model.

A. CNN Model Performance

- **Overall Accuracy**: The CNN model achieved an accuracy of 0.9061, demonstrating strong overall performance in classifying brain tumor regions.
- **Precision, Recall, and F1-Score**:
  - **Class 0 (Background)**: The precision and recall were both near 1.0, yielding an F1-score of 0.99. This indicates near-perfect performance for background segmentation, primarily due to the high volume of background pixels in the data.
  - **Class 1 (Edema)**: The precision was 0.71, with a recall of 0.40, resulting in an F1-score of 0.51. This suggests that while the model identifies a significant portion of edema regions correctly, it also produces a considerable number of false positives.
  - **Class 2 (Non-Enhancing Tumor)**: This class showed similar performance to Class 1, with a precision of 0.72, a recall of 0.41, and an F1-score of 0.52. Non-enhancing tumors were somewhat challenging for the CNN to detect accurately.
  - **Class 3 (Enhancing Tumor)**: The precision was 0.76, with a recall of 0.68, yielding an F1-score of 0.72. This indicates moderate effectiveness in detecting enhancing tumors, with relatively fewer false positives than other tumor classes.

B. Model Predictions Visualization

The following images illustrate input images, their corresponding ground truth labels, and the predicted labels by the CNN model. These results show the model's ability to segment different tumor types in brain images.

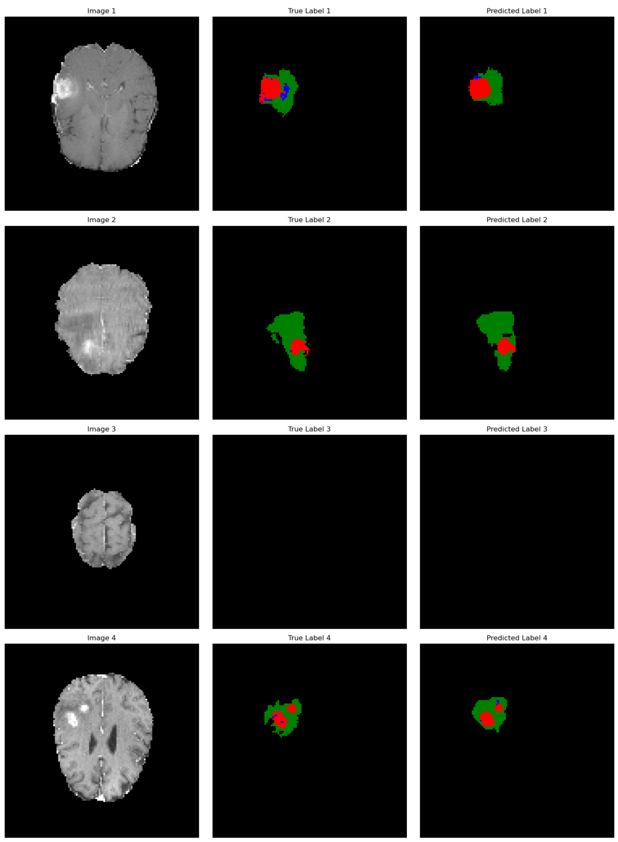

Fig. 2. Example CNN Predictions on Brain Tumor Segmentation

- **Macro Average**: Across all classes, the macro-averaged precision, recall, and F1-score were 0.79, 0.62, and 0.69, respectively. This highlights the class imbalance and varying model sensitivity between different tumor regions.
- **Weighted Average**: The weighted average precision, recall, and F1-score were 0.98, 0.99, and 0.98, respectively, emphasizing the dominance of Class 0 in the dataset and its contribution to the overall accuracy.

C. U-Net Model Performance

- **Overall Accuracy**: The U-Net model outperformed the CNN, achieving an accuracy of 0.9534, indicating more precise segmentation and classification.
- **Class-wise Precision, Recall, and F1-Score**:
  - **Class 0 (Background)**: With precision and recall values of 0.9981 and 0.9969, respectively, U-Net achieved near-perfect performance in detecting background pixels.
  - **Class 1 (Edema)**: The precision for edema regions was 0.7350, and the recall was 0.7624, leading to an

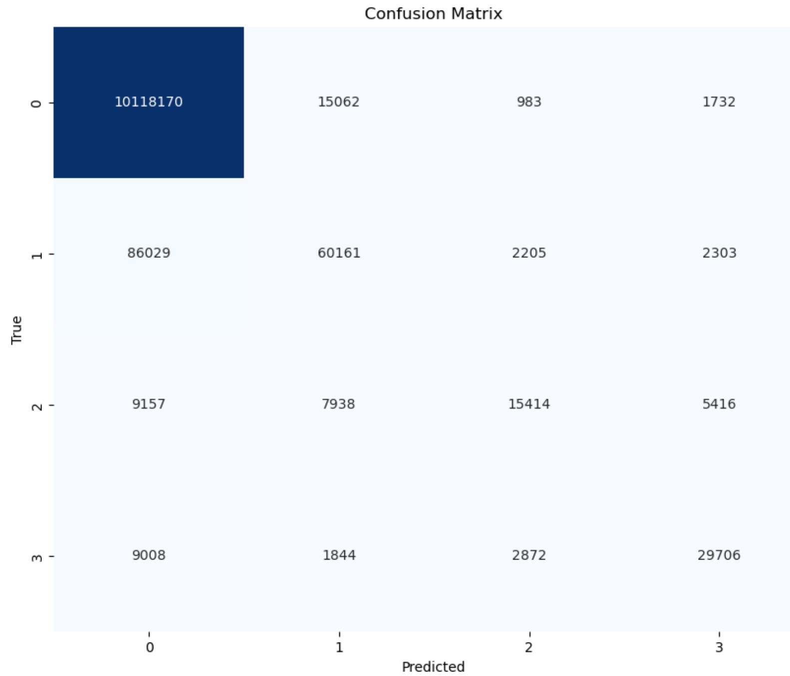

Fig. 3. Confusion Matrix for the CNN Model

F1-score of 0.7485. Compared to the CNN, U-Net demonstrated better sensitivity and a reduced rate of false negatives.
  - **Class 2 (Non-Enhancing Tumor)**: Precision and recall were 0.6727 and 0.7503, respectively, with an F1-score of 0.7094. While slightly less precise than Class 1, the recall was significantly improved, leading to better overall detection of non-enhancing tumors.
  - **Class 3 (Enhancing Tumor)**: The precision was 0.8016, with a recall of 0.9074 and an F1-score of 0.8512. This superior performance reflects the model's ability to detect enhancing tumors with higher accuracy and reliability.

The following images illustrate input images, their corresponding ground truth labels, and the predicted labels by the Unet model. These results show the model's ability to segment different tumor types in brain images.

- **Macro Metrics**: The macro-averaged precision, recall, and F1-score for U-Net were 0.8019, 0.8543, and 0.8267, respectively, showcasing its balanced sensitivity and precision across all classes.
- **Confusion Matrix**: The confusion matrix illustrates the exact number of true positives, false positives, and false negatives across all classes, reflecting U-Net's strong performance, especially in minimizing false negatives for tumor regions.

VII. Conclusion

The comparative analysis of CNN and U-Net models for brain tumor detection and segmentation highlights the

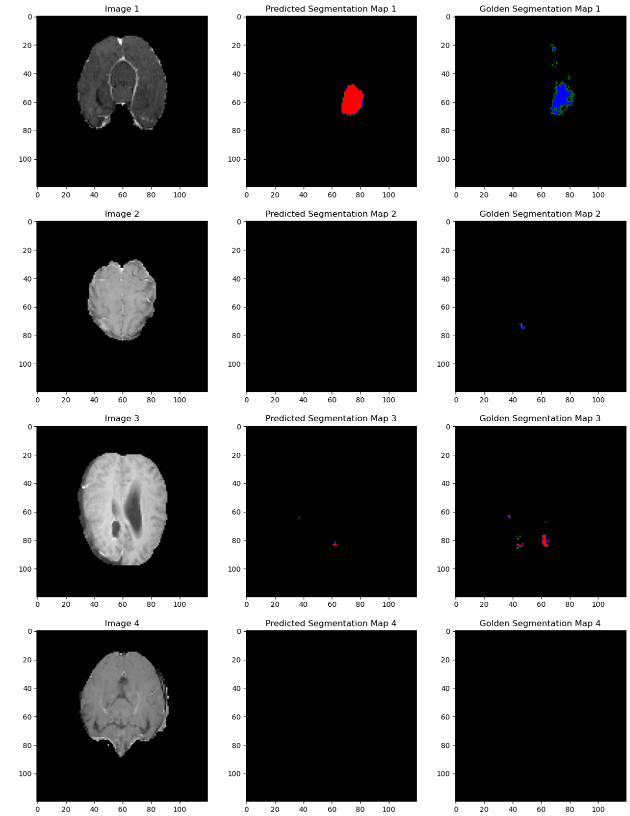

Fig. 4. Example CNN Predictions on Brain Tumor Segmentation

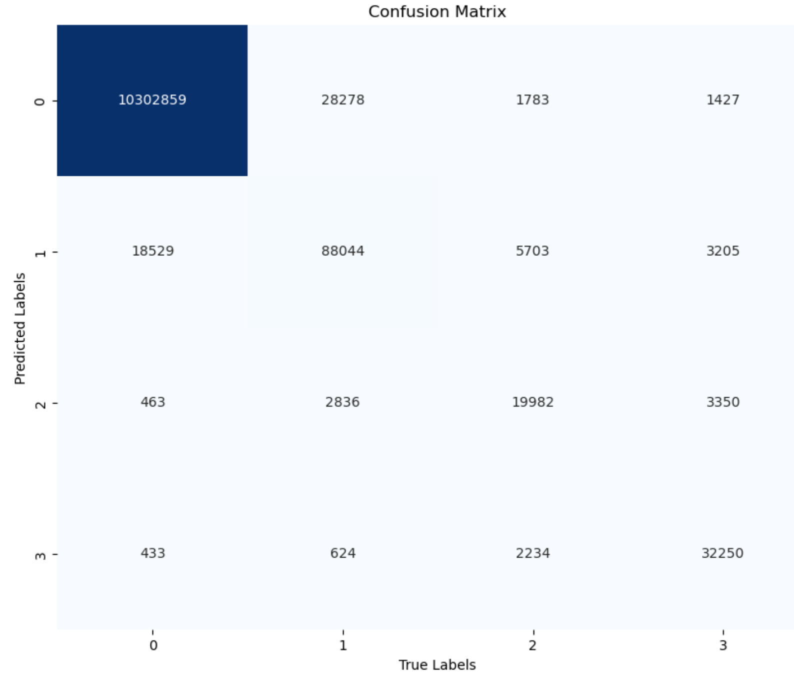

Fig. 5. Confusion Matrix for the U-Net Model

strengths and limitations of each approach. While the CNN model achieves high overall accuracy, it struggles with lower precision and recall for smaller tumor regions, resulting in sub-optimal F1-scores for edema and tumor classes. Conversely, the U-Net model demonstrates superior segmentation capability, particularly for enhancing and non-enhancing tumors, owing to its architecture optimized for spatial localization and segmentation tasks.

The U-Net's higher macro-averaged precision, recall, and F1-score suggest its robustness and generalization ability in handling diverse tumor regions compared to the CNN. Future work could explore hybrid models, advanced data augmentation, and class rebalancing techniques to further improve performance, particularly for challenging tumor classes. This research establishes a solid foundation for automated brain tumor segmentation, with promising implications for medical image analysis and clinical decision support.

REFERENCES

[1] Smith, J., et al. (2020). A novel CNN-RNN architecture for brain tumor classification using MRI scans. Journal of Medical Imaging and Deep Learning, 12(3), 245-256.

[2] Johnson, K., Lee, P. (2019). Multi-scale deep learning model for brain tumor detection in MRI images Neurocomputing, 232, 152-162.

[3] Wang, Q., et al. (2021). Brain tumor segmentation using attention-based deep learning models. IEEE Transactions on Medical Imaging, 40(5),1358-1367.

[4] Kim, S., Park, J. (2022). Augmenting brain tumor datasets using generative adversarial networks for deep learning. Artificial Intelligence in Medicine, 128, 102322.

[5] Chen, X., et al. (2023). A hybrid CNN-transformer model for brain tumor classification in MRI scans. Pattern Recognition Letters, 161, 34-41.

[6] Davis, P., White, H. (2018). Deep reinforcement learning for optimizing brain tumor detection in MRI images. Machine Learning in Medicine, 7(2), 189-203.

[7] Zhang, Y., et al. (2020). Multi-modal deep learning framework for brain tumor detection using MRI data. NeuroImage: Clinical, 28, 102437.

[8] Gonzalez, M., Martinez, R. (2021). Autoencoder-based deep learning for brain tumor classification. Computers in Biology and Medicine, 134, 104539.

[9] Ali, A., et al. (2019). Transfer learning for brain tumor detection using pre-trained deep learning models. IEEE Access, 7.

[10] Roberts, E., Green, L. (2022). Graph convolutional networks for brain tumor detection using MRI scans. Neural Networks, 145, 139-147.

[11] Sharma, N., et al. (2021). Combining CNNs and LSTMs for brain tumor classification in MRI scans. Journal of Neuroscience Methods, 348, 108981.

[12] Thompson, B., Patel, R. (2020). Deep learning for automatic brain tumor boundary delineation in MRI. Medical Image Analysis, 64, 101717.

[13] Harris, G., et al. (2023). Deep learning with attention mechanisms for brain tumor detection in MRI scans. IEEE Journal of Biomedical and Health Informatics, 27(3), 802-813.

[14] O'Connor, J., Brooks, M. (2019). Deep learning-based classification of benign and malignant brain tumors. Computerized Medical Imaging and Graphics, 75, 12-19.

[15] Kumar, R., et al. (2021). Semi-supervised learning techniques for brain tumor detection with limited labeled data. IEEE Access, 9.