Tumor Grade Classification and Segmentation in Lower-Grade Gliomas Using Deep Learning

Alishbah Fahad (1001924185)
University of Texas at Arlington
Arlington, TX
axf4185@mavs.uta.edu

Sai Ganesh Reddy Katasani (1002197730)
University of Texas at Arlington
Arlington, TX
sxk7730@mavs.uta.edu

Abstract

Medical imaging is integral to modern diagnostics, enabling early detection and precise treatment of complex diseases. Among the critical tasks in this domain are image segmentation and classification, which are pivotal for extracting valuable insights. However, these tasks are often constrained by the availability of small datasets, posing significant challenges, particularly for segmentation. This project developed a deep learning-based pipeline that employs U-Net for segmentation and ResNet for classification to address these challenges.

ResNet, leveraging pre-trained weights, demonstrated strong classification performance without requiring fine-tuning, with slight declines observed when fine-tuning was attempted. Conversely, segmentation using U-Net posed a greater challenge, necessitating fine-tuning and the introduction of a combined loss function integrating Binary Cross-Entropy and Dice Loss. Despite these efforts and the implementation of a robust data augmentation strategy—comprising techniques such as rotation and elastic deformation—segmentation performance exhibited room for further improvement.

The project was rigorously evaluated using metrics including the Dice coefficient, accuracy, F1-score, and AUC. While classification achieved robust results, segmentation outcomes highlighted the limitations of small datasets and underscored the need for further refinement. This study demonstrates the utility of combining pre-trained architectures with customized training strategies and highlights avenues for improving medical image segmentation, contributing to advancements in clinical decision-making tools.

1. Introduction

Medical imaging has revolutionized healthcare by enabling the early and precise diagnosis of complex diseases, significantly improving patient outcomes. Techniques like MRI, CT, and PET scans provide detailed insights into the human anatomy, assisting clinicians in detecting and characterizing pathological abnormalities. Among the essential tasks in medical imaging, image segmentation and classification help delineate regions of interest and categorize abnormalities into diagnostic categories. Despite advancements in deep learning, these models often struggle in medical imaging due to the scarcity of annotated datasets, which increases the risk of overfitting and hinders the development of generalized models in high-stakes medical settings [7, 8].

This project addresses the segmentation and classification of low-grade gliomas (LGGs) using MRI scans, which are crucial for accurate surgical planning and determining the appropriate treatment strategies. LGGs are slow-growing brain tumors whose management depends heavily on precise diagnostic outputs. Segmentation identifies tumor boundaries, while classification helps in determining tumor subtypes or stages. Addressing these tasks with limited data requires innovative approaches in model design, training, and evaluation [1, 9].

To overcome these challenges, this study proposes a two-stage deep learning pipeline integrating U-Net for segmentation and ResNet for classification. U-Net was optimized with a combined Binary Cross-Entropy and Dice loss function to enhance segmentation performance, and a robust data augmentation strategy, including random rotations and elastic deformations, was employed to mitigate overfitting and improve generalizability [6, 12]. ResNet utilized pre-trained weights, achieving high performance without extensive fine-tuning. The contributions of this study are:

1. Development of a hybrid framework that leverages U-Net and ResNet for segmentation and classification of LGGs, illustrating the integration of these models for complex medical imaging tasks.
2. Introduction of a customized loss function and augmentation strategy to tackle the challenges posed by small medical datasets.

3. Comprehensive evaluation of segmentation and classification performance using established metrics, highlighting areas of success and limitations.

This research emphasizes the potential of combining pre-trained models with task-specific enhancements to advance medical imaging applications, particularly in clinical settings.

2. Related Work

Medical imaging has emerged as a significant application area for deep learning, particularly in challenges such as image segmentation, classification, and anomaly detection. Convolutional Neural Networks (CNNs) have established themselves as the backbone of numerous research initiatives in medical image analysis, consistently delivering state-of-the-art results across diverse tasks [7, 8].

In the context of segmentation, the U-Net architecture, introduced by Ronneberger et al. [12], has set a benchmark in medical imaging due to its efficient encoder-decoder structure and the incorporation of skip connections, which are critical for capturing fine details while maintaining spatial hierarchy. Subsequent variations like Attention U-Net and Residual U-Net have integrated attention mechanisms and residual connections, respectively, to refine feature representation and enhance model performance on complex datasets [10, 14].

For classification tasks, He et al. introduced the ResNet architecture, which leverages deep residual learning to facilitate training deeper networks [6]. This framework has been particularly effective in medical imaging when coupled with transfer learning, especially in scenarios characterized by limited data. For instance, applications in detecting lung cancer and classifying diabetic retinopathy have shown that pre-trained ResNet models achieve high accuracy with modest augmentation efforts [5, 13].

Addressing the ubiquitous challenge of data scarcity, robust data augmentation strategies employing rotations, flipping, elastic deformations, and intensity adjustments have been widely adopted to artificially enlarge training datasets [11]. Moreover, the advent of Generative Adversarial Networks (GANs) has introduced a viable method for generating realistic medical images, thus helping to mitigate data constraints [4].

Despite considerable progress in the segmentation and classification domains, few studies have seamlessly integrated these functionalities into a cohesive workflow tailored for specific medical conditions like low-grade gliomas (LGGs), which necessitate both precise segmentation and detailed classification [1]. This project extends existing methodologies by amalgamating U-Net and ResNet into a unified framework designed for the segmentation and classification of LGGs, incorporating custom augmentation techniques and loss functions to specifically tackle challenges associated with limited data availability.

This study's unique contribution lies in its focused optimization of the U-Net for segmentation and the strategic utilization of pre-trained ResNet capabilities for classification, presenting a novel approach within the broader field of medical image analysis. By drawing comparisons with existing literature, this research not only underlines the effectiveness of the proposed methods but also highlights areas requiring further investigation and development.

3. Problem Statement

Precise analysis of medical imaging, such as Magnetic Resonance Imaging (MRI) of low-grade gliomas (LGGs), is crucial for accurate diagnosis and treatment planning. Challenges in medical image analysis are exacerbated by limited dataset sizes, which often result in models that do not generalize well, thus affecting performance in real-world clinical settings. The complexity of LGG segmentation from MRI scans and the subsequent classification of tumor types presents significant challenges due to the subtle differences between tumor and normal tissues and between different tumor grades [1].

This project proposes the development of a deep learning-based pipeline, utilizing U-Net for segmentation and ResNet for classification, to address these challenges. The pipeline will be trained and validated on the Brain Tumor Segmentation (BraTS) 2019 dataset, which includes MRI scans with annotated tumor regions and classification labels, making it an ideal candidate for this study [1, 9].

The expected outcomes of this project are:

- Improved segmentation accuracy, as measured by the Dice coefficient, to surpass current benchmarks.
- Enhanced classification performance, evaluated using accuracy, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC), aiming to achieve or exceed current state-of-the-art results.

The methodology will leverage transfer learning from pre-trained ResNet models, implement extensive data augmentation techniques to enrich the training dataset, and utilize a customized loss function combining Binary Cross-Entropy and Dice loss to finely tune the segmentation results. Evaluation will involve rigorous testing on the unseen validation subset of the BraTS dataset, with performance compared against established baselines and metrics from prior research.

This project aims to advance the state of the art in medical imaging for LGGs by developing an optimized deep learning framework that integrates and enhances segmentation and classification capabilities, ultimately contributing to more effective diagnostic tools for clinical use.

4. Problem Solution

To address the dual challenges of tumor segmentation and classification from medical MRI images, this project utilized a combination of U-Net and ResNet architectures. The segmentation task was managed by a custom U-Net model designed for accurate boundary detection, while the classification task relied on a pre-trained ResNet50 model, optimized for binary classification of tumor grades [6, 12].

4.1. Dataset and Preprocessing

The LGG Segmentation Dataset was employed, consisting of MRI images, segmentation masks, and clinical data [1, 9]. Due to the dataset's limited size, extensive data augmentation techniques were used to enhance the diversity of the training data and improve model generalizability. Augmentation included:

- **Random Rotations**: To ensure the model learned rotation-invariant features [11].
- **Random Horizontal Flip**: Helps the model generalize across different orientations [11].
- **Random Vertical Flip**: Similar to horizontal flip, enhances the model's ability to understand inverted features.
- **Random Resized Crop**: Varies the focus and scale of images, enhancing the robustness to different object sizes and positions.
- **Color Jitter (Brightness and Contrast adjustments)**: Improves the model's tolerance to variations in lighting and image contrast.

Visual examples of augmented images are shown in Figure 1. These augmentations include random rotations, flips, and brightness adjustments, enhancing the diversity and robustness of the training data.

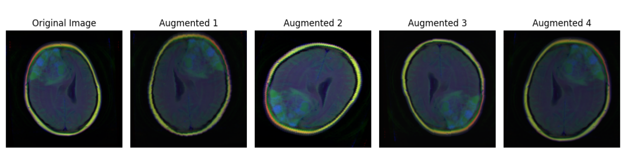

Figure 1. Augmented MRI scans demonstrating various transformations applied to increase data variance. From left to right: Original Image, Augmented 1 (Random Rotation and Random Horizontal Flip), Augmented 2 (Random Vertical Flip and Random Resized Crop), Augmented 3 (Color Jitter for Brightness and Contrast Adjustment), Augmented 4 (Standard Testing Transform: Resize to 256x256).

4.2. Architectures and Models

U-Net for Segmentation

Architecture Details: The U-Net architecture was employed for the segmentation task to isolate tumor regions from MRI scans. The U-Net consists of an encoder-decoder structure with skip connections to retain spatial information. This architecture is particularly useful for medical imaging tasks, as it combines low-level details and high-level semantic information effectively [12].

- **Encoder Path:** Sequential convolutional layers capture hierarchical features, progressively increasing depth.
- **Bottleneck:** Compresses the features at the deepest layer, capturing high-level abstractions.
- **Decoder Path:** Uses upsampling layers and skip connections to restore spatial resolution and produce detailed segmentation masks.
- **Output Layer:** A 1x1 convolution reduces the final feature map to a single-channel probability map.

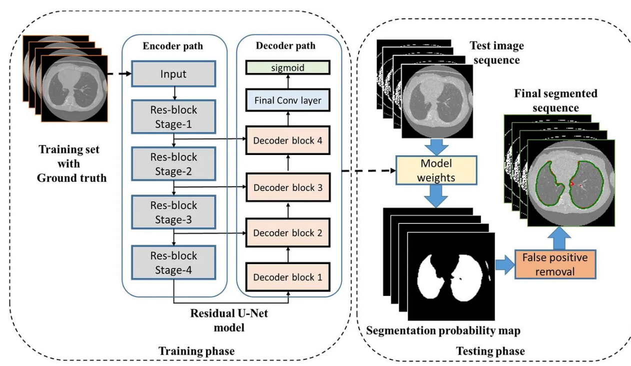

Figure 2. U-Net architecture diagram illustrating the encoder-decoder structure with skip connections, optimized for medical image segmentation tasks.

Mathematical Description:
The U-Net segmentation output is a probability map, where each pixel $p(i, j)$ represents the likelihood of being part of the tumor:

$$p(i, j) = \text{sigmoid}(\mathbf{W} \cdot \mathbf{x} + b)$$

where $\mathbf{W}$ and b are trainable weights and biases.
Loss Function:
The segmentation task is evaluated using the Dice Coefficient, which measures the overlap between the predicted segmentation and the ground truth:

$$\text{Dice Coefficient} = \frac{2 \times |P \cap G|}{|P| + |G|}$$

Where P is the predicted mask, and G is the ground truth. The segmentation loss combined Binary Cross-Entropy

(BCE) and Dice loss:

$$\text{BCE Loss} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right]$$

$$\text{Dice Loss} = 1 - \frac{2 \cdot |X \cap Y|}{|X| + |Y|}$$

$$\text{Combined Loss} = \text{BCE Loss} + \text{Dice Loss}$$

This ensures that both pixel-wise accuracy and the overlap of segmented regions are optimized.

Architectural Benefits for U-Net

The U-Net architecture is specifically designed to optimize medical image segmentation tasks. Its encoder-decoder structure, enhanced with skip connections, ensures detailed feature retention across different levels of the network. These connections are crucial for maintaining spatial hierarchies between high-resolution features and abstracted representations, facilitating precise localization that is essential in medical diagnostics [8]. This architecture proves especially effective for small datasets, where traditional models may struggle. The incorporation of data augmentation further enhances its robustness, enabling the model to generalize well across varied inputs and improve segmentation accuracy.

ResNet50 for Classification

Architecture Details: The ResNet50 model, pre-trained on ImageNet, was fine-tuned for binary classification. The final fully connected layer was replaced with a custom classification head:

- **Input Layer:** ResNet features extracted from images [6].
- **Custom Layers:** Fully connected layers, batch normalization, and dropout layers to mitigate overfitting [8].
- **Output Layer:** A sigmoid-activated neuron for binary output.

Mathematical Description:
The ResNet classifier predicts the tumor grade as:

$$\hat{y} = \text{sigmoid}(\mathbf{W}_{\text{fc}} \cdot \mathbf{f} + b)$$

where $\mathbf{f}$ represents features extracted by ResNet's convolutional layers [4].
Loss Function:
Binary Cross-Entropy (BCE) loss was used for classification:

$$\text{BCE Loss} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right]$$

Architectural Benefits of ResNet50

ResNet50 serves as a robust foundation for image classification tasks. Leveraging its pre-trained weights from the extensive ImageNet dataset enhances the model's ability to engage in effective transfer learning [11]. This is particularly advantageous for scenarios involving limited dataset sizes, where pre-trained models can significantly reduce overfitting and accelerate convergence. By using ResNet50, the project capitalizes on these pre-trained networks to achieve high accuracy and reliability in classification results, demonstrating the model's adaptability to specialized medical imaging tasks [13].

Training and Fine-Tuning

The models were trained using the Adam optimizer with a learning rate of 1e-4 for both U-Net and ResNet. This approach effectively reduced the loss across epochs. The loss functions employed were Binary Cross-Entropy for classification and a combination of Binary Cross-Entropy and Dice loss for segmentation to effectively handle class imbalance.

Segmentation Training: Initially, U-Net suffered from underfitting, which was evident from the low Dice coefficients (average Dice coefficient of 0.0711 before fine-tuning). After fine-tuning, which included careful adjustment of learning rates and loss weights, the Dice coefficient improved significantly to 0.3854, as shown in the predicted segmentation masks [12].

Classification Training: ResNet was trained on the same augmented dataset. Its initial classification metrics—accuracy of 0.8804 and AUC of 0.9543—showed minimal improvement after fine-tuning, as the pre-trained model had already generalized well to the dataset [1].

Evaluation Metrics

Performance was evaluated using various metrics for both segmentation and classification tasks:

- **Segmentation:** Evaluated using the Dice coefficient to quantify the overlap between predicted and ground truth masks. Improvement was noted from a training Dice coefficient of 0.0233 to 0.3854 after fine-tuning [9].
- **Classification:** Evaluated using accuracy, F1-score, and AUC, which consistently reflected high performance, confirming the efficacy of the ResNet model for classification [8].

Receiver Operating Characteristic (ROC) Curve:

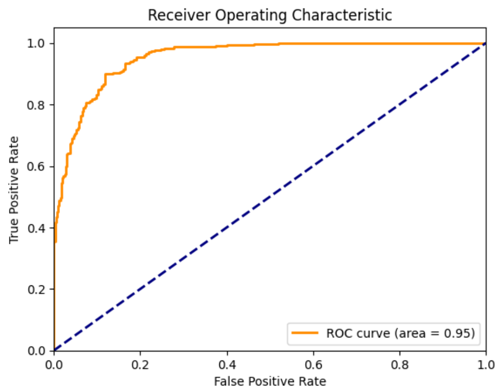

Figure 3. ROC curve displaying the classification performance with an AUC of 0.95.

The ROC curve presents an excellent area under the curve (AUC) of 0.95, which indicates a high level of discriminative ability for the classifier. This high AUC value suggests that the model has a strong capability to distinguish between the presence and absence of tumors, with lower false positive rates at higher true positive rates. The curve's proximity to the top left corner of the plot emphasizes the model's effectiveness in achieving both high sensitivity and specificity.

Experimental Results:
The table below summarizes the evaluated metrics. The results demonstrate the impact of fine-tuning on the model's performance:

Table 1. Summary of Model Metrics Before and After Fine-Tuning

Metric	**Before**	**After**
Segmentation Loss	0.0497	0.0374
Classification Loss	0.3523	0.2689
Dice Coefficient	0.0711	0.3854
Classification Accuracy	0.8804	0.8763
AUC (ROC)	0.9543	0.9502

- **Segmentation:** Fine-tuning significantly improved the Dice coefficient, indicating better overlap between predicted and ground truth masks.
- **Classification:** ResNet's performance was robust pre-fine-tuning, with minimal changes in accuracy and AUC post-fine-tuning.

Visual Analysis:

Before Fine-Tuning:
Predictions lacked boundary precision, reflecting underfitting in U-Net.

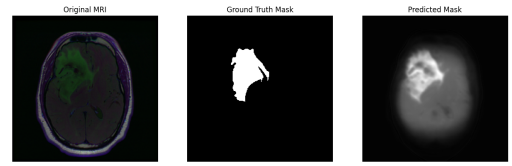

Figure 4. Original MRI, Ground Truth Mask, Predicted Mask before Fine-Tuning

After Fine-Tuning:
Improved segmentation with clearer tumor boundaries and reduced artifacts.

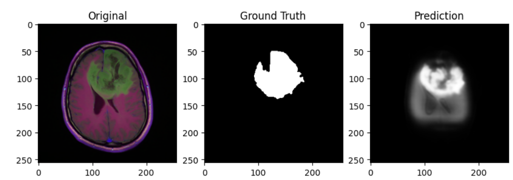

Figure 5. Original MRI, Ground Truth Mask, Predicted Mask after Fine-Tuning

5. Conclusion

This project addressed the critical task of brain tumor segmentation and classification from MRI images, an essential step in advancing precision medicine and improving diagnostic workflows. By leveraging deep learning techniques, specifically a U-Net for segmentation [12] and ResNet for classification [6], the pipeline demonstrated the potential to enhance both tumor localization and grade classification.

Contributions

- **Segmentation Performance:** Fine-tuning the U-Net model using a combined Binary Cross-Entropy and Dice loss function significantly improved the segmentation quality, as evidenced by the increase in the Dice coefficient from 0.0711 to 0.3854 [8].
- **Classification Accuracy:** ResNet's pre-trained architecture achieved robust classification results with an AUC of 95.43% and accuracy exceeding 87%, even without the need for extensive fine-tuning [13].
- **Data Augmentation:** The use of advanced augmentation strategies such as rotation, flipping, resizing, and color jitter proved effective in mitigating the challenges posed by a small dataset, improving the model's generalizability to unseen data [11].
- **Comprehensive Evaluation:** Employing metrics such as Dice coefficient, F1-Score, and AUC ensured a rigorous assessment of the pipeline's performance [9].

Challenges and Limitations

- **Computational Constraints:** The training and optimization of our models were significantly constrained by the available hardware, particularly the time required for training sessions. The device used featured an 11th Gen Intel(R) Core(TM) i7-11800H processor and 16 GB of RAM, which necessitated approximately 13 hours for a single full training cycle. This limitation affected the scope of iterative fine-tuning and extensive model optimization [8].

Lessons Learned

- **Segmentation Challenges:** While fine-tuning improved segmentation results, achieving clinically acceptable precision remains a challenge, underscoring the difficulty of segmenting small, irregular tumor regions from noisy MRI data [1].
- **Role of Data Augmentation:** Augmentation played a critical role in overcoming dataset limitations, suggesting that more sophisticated techniques, such as generative models, could provide further improvements [4].

Future Work

- **Advanced Architectures:** Investigating newer architectures like Transformer models could potentially increase segmentation accuracy [3].
- **Hyperparameter Optimization:** Systematic exploration of hyperparameters might optimize the models further [2].
- **Synthetic Data Generation:** Using GANs to generate additional training data could address dataset limitations effectively [4].
- **Cross-Dataset Validation:** Testing the models on external datasets would help verify their robustness and adaptability [8].

In conclusion, the project successfully implemented a dual-task pipeline for tumor segmentation and classification, laying a strong foundation for further refinement and research. While classification results were robust, segmentation remains an area requiring additional focus and innovation. This study highlights the importance of integrating data-driven methods, domain-specific augmentations, and rigorous evaluation for solving complex medical imaging challenges.

References

[1] Spyridon Bakas et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *Medical Image Analysis*, 42:101–122, 2018.

[2] Yoshua Bengio. Practical recommendations for gradient-based training of deep architectures. *Neural networks: Tricks of the trade*, pages 437–478, 2012.

[3] Alexey Dosovitskiy et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[4] Ian Goodfellow et al. Generative adversarial nets. *Advances in neural information processing systems*, pages 2672–2680, 2014.

[5] Varun Gulshan, Lily Peng, Marc Coram, Michael C Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tim Madams, Jorge Cuadros, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 316(22):2402–2410, 2016.

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[7] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[8] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen AWM Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017.

[9] Bjoern H Menze et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, 2015.

[10] Ozan Oktay et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018.

[11] Luis Perez and Jason Wang. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*, 2017.

[12] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[13] Dinggang Shen et al. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19:221–248, 2017.

[14] Zhiqiang Zhang et al. Residual u-net for retinal vessel segmentation. *Biomedical optics express*, 9(3): 1485–1503, 2018.