SEGMANTIC SEGMENTATION OF BRAINSCANS

ADVANCING AUTOMATED BRAIN TUMOR SEGMENTATION: PROGRESS AND REFLECTIONS ON THE WORKFLOW DEVELOPMENT

AUTHORS

Isabella Hofstede
398319, BFV3
https://github.com/ishofstede/modelleren_van_kanker_brats/tree/tensorflow

ACKNOWLEDGEMENTS

This project is based on a tutorial by Sreenivas Bhattiprolu, where a 2D U-Net model is adapted for 3D medical image segmentation. The implementation here is built upon concepts discussed in his tutorial series.

INTRODUCTION

THIS PROJECT EXPLORES THE APPLICATION OF SEMANTIC SEGMENTATION TO THREE-DIMENSIONAL MRI SCANS OF THE BRAIN, UTILIZING THE BRATS DATASET. THE PROJECT FOCUSED ON BUILDING A USABLE WORKFLOW AND GENERATING PREDICTIONS USING A 3D-UNET MODEL. THE PROJECT AIMS TO HIGHLIGHT THE CHALLENGES AND ACHIEVEMENTS IN AUTOMATING BRAIN TUMOR SEGMENTATION, WHICH HAS SIGNIFICANT IMPLICATIONS FOR MEDICAL DIAGNOSTICS.

OBJECTIVE

The primary objective of this project was to establish an efficient and scalable workflow for segmenting brain tumors in volumetric MRI data. This was achieved by leveraging a pre-trained 3D-Unet model, with the longterm goal of eventually training a custom model to enhance performance and accuracy.

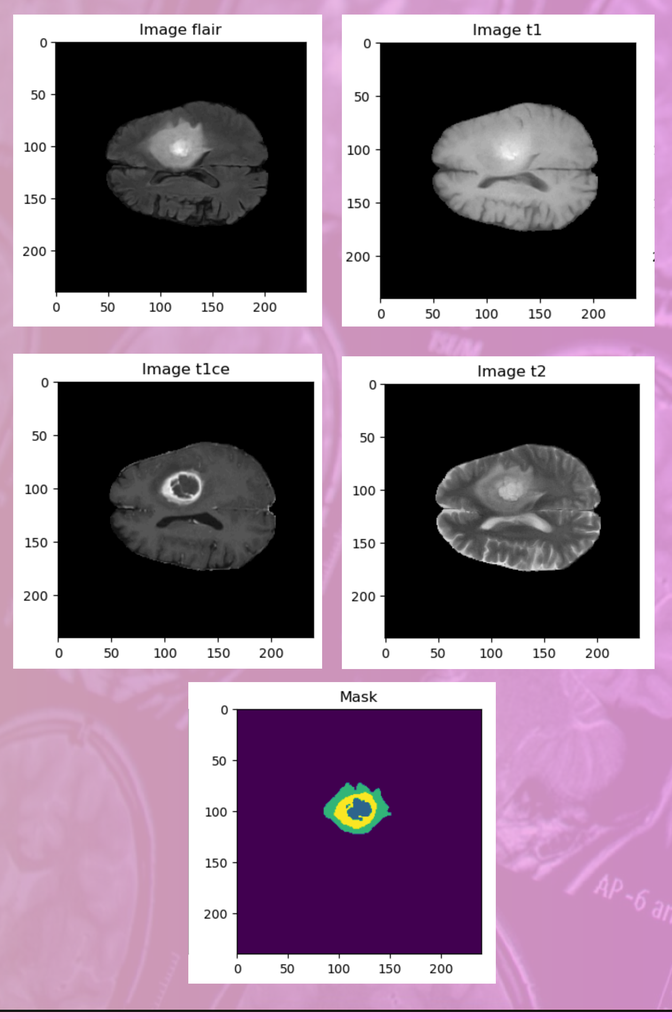
METHODOLOGY

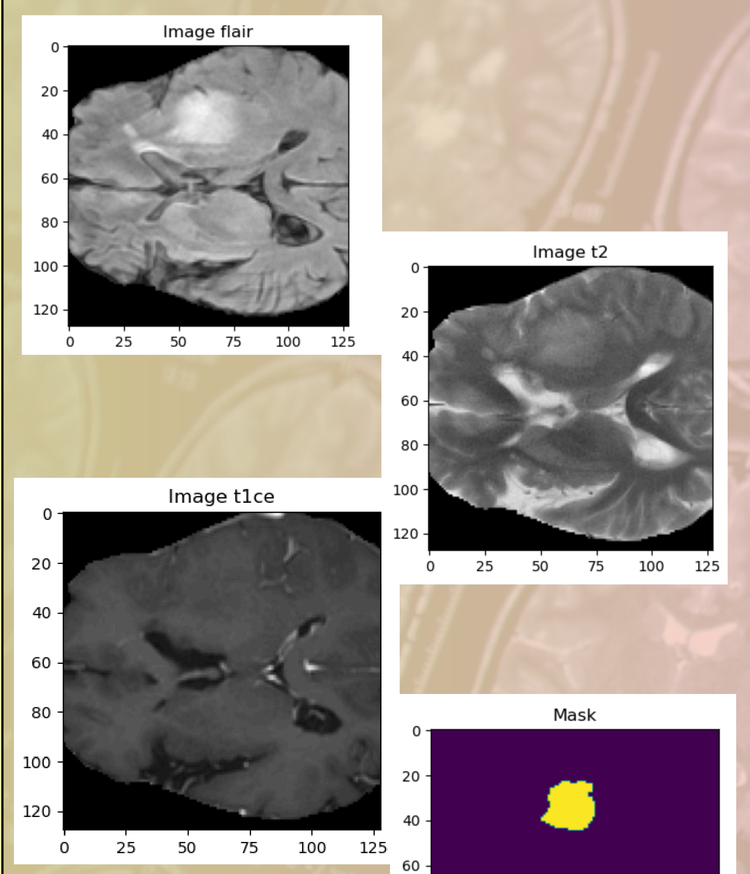
- Data Preprocessing: A comprehensive exploratory data analysis (EDA) was conducted to understand the structure and challenges of the dataset. Preprocessing steps were then automated using custom scripts, ensuring a streamlined pipeline for handling large volumes of MRI data.
- Prediction Execution: Instead of training the model due to time limitations, predictions were generated using a pre-existing, pre-trained model, which allowed for immediate analysis and validation of the segmentation approach.
- Reflection on Methodology: The decision to use TensorFlow over PyTorch was driven by time constraints and familiarity, though migration to PyTorch was explored. Several technical hurdles with Keras were encountered and successfully mitigated, contributing to the overall learning process.

RESULTS/FINDINGS

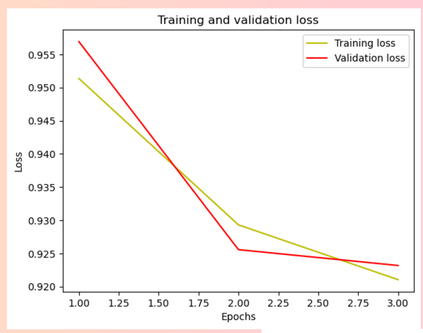
While the generated segmentations were not perfect, they demonstrated promising potential as a starting point for further refinement. The tumor segmentation predictions were mostly accurate but contained some inconsistencies, which highlighted areas for future optimization.

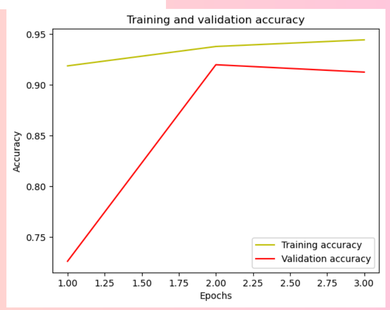
[RAW DATA] Raw data used before preprocessing step.

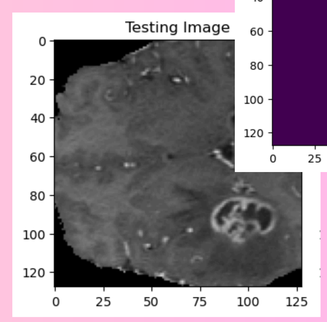


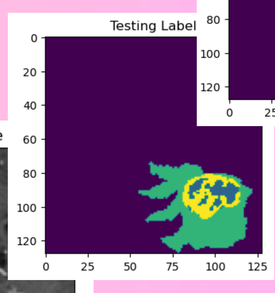


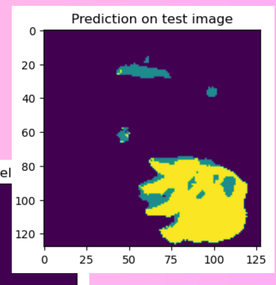
[PROCESSED DATA] This data has been adjusted so that dimensions are cropped for better performance.











[RESULTS OF TESTING 3 EPOCHS] LOSS FUNCTIONS CONTINUES TO GO DOWN, MEANING MODEL STILL HAS ROOM FOR IMPROVEMENT.

ANALYISIS

- Strengths: The automated preprocessing pipeline was highly effective in handling large datasets efficiently. The use of a custom data generator facilitated great integration with the model, even when dealing with complex volumetric data. Furthermore, the generated segmentations provided an insightful basis for further refinement.
- Challenges: The inability to train the model due to time limitations prevented the optimization of hyperparameters and fine-tuning. Additionally, the migration from Keras to PyTorch was deemed too resource-intensive for the scope of the project, and certain legacy issues with Keras functions introduced unforeseen challenges.
- Comparison to Expectations: Although the goal of training a custom model was not achieved, the results still showed a viable path forward for further research, with substantial potential to refine the model's accuracy and robustness.

CONCLUSION

This project demonstrates the feasibility of implementing an automated pipeline for 3D medical image segmentation, which is useful for future advancements in brain tumor detection. Despite time limitations, the successful execution of preprocessing and prediction steps provides a strong foundation for further model training and optimization. Future work will focus on training a custom model, selecting the optimal framework (TensorFlow vs. PyTorch), and enhancing the segmentation accuracy through advanced optimization strategies.

https://www.kaggle.com/datasets/awsaf49/brats20-dataset-training-validation
https://www.youtube.com/playlist?list=PLZsOBAyNTZwYgF8O1bTdV-lBdN55wLHDr

Hanze