Assignment 1

By Akanksha Verma (20MIP10002)

Submitted to Dr. Nikhil Pateria

TASK

https://www.kaggle.com/code/naghmehsh/image-segmentation-for-detecting-brain-tumor

Data / Input

The dataset used here comprises of magnetic resonance imaging (MRI) scans of the brain, specifically focusing on patients with suspected brain tumors. The dataset contains MRI scans of both Tumour and Healthy patient's brains, providing a comprehensive view of the brain structures. Each image indicates the presence and location of tumors, making it suitable for unsupervised learning tasks. The dataset is diverse, encompassing a range of tumor types, sizes, and locations, ensuring a robust evaluation of the segmentation model.

Literature Review

Traditional methods often face challenges in accurately delineating tumor boundaries, making deep learning models, specifically the U-Net architecture, a promising solution. The goal is to improve the precision and reliability of tumor segmentation, aiding medical professionals in diagnosis and treatment planning.

Problem Statement

The primary objective of this study is to develop an efficient and accurate method for detecting and segmenting brain tumors from MRI scans.

Data analysis (preprocessing)

Before feeding the data into the U-Net model, several preprocessing steps are implemented to enhance the quality and suitability of the input. This includes standardization of pixel intensities, normalization to correct for variations in scanner parameters, and resizing to ensure uniformity in image dimensions. Additionally, data augmentation techniques such as rotation, flipping, and zooming are applied to augment the dataset and improve the model's generalization capabilities.

The maths behind the transformations

1. **Standardization of Pixel Intensities:**

$$X \text{ standardized} = \frac{(\mathrm{X} - \mathrm{mean}(X))}{\mathrm{std}(X)}$$

2. **Normalization:**

$$X \text{ normalized} = \frac{\mathrm{X} - \min(X)}{\max(X) - \min(X)}$$

3. **Resizing**

X resized = resize (X , target_size)

4. **Data Augmentation**

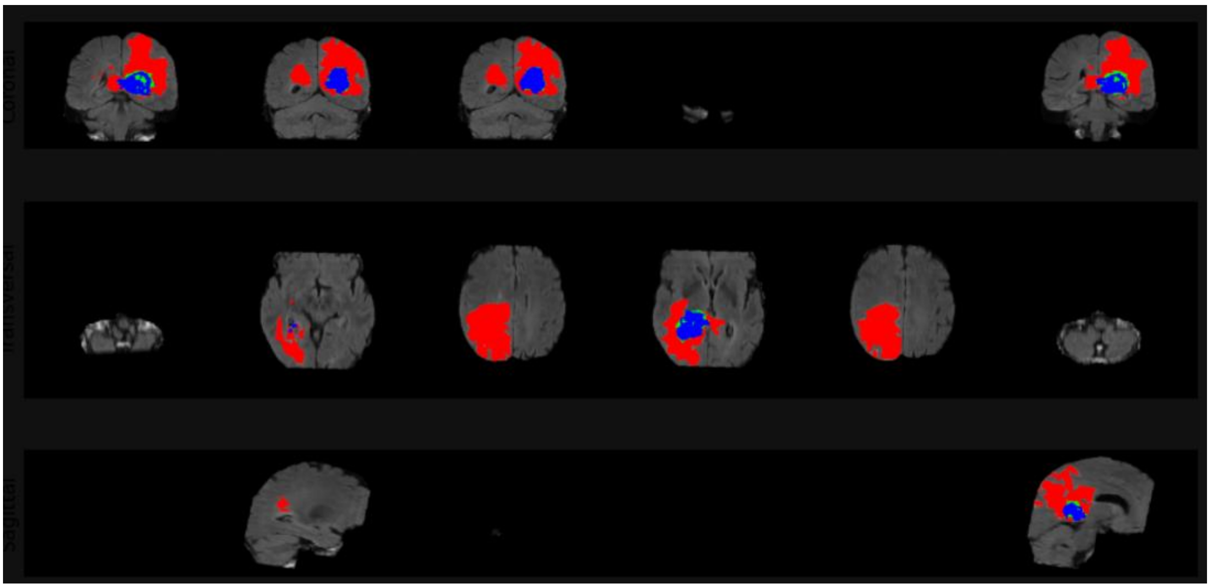
X augmented = augment(X)

ML or DL model used (description)

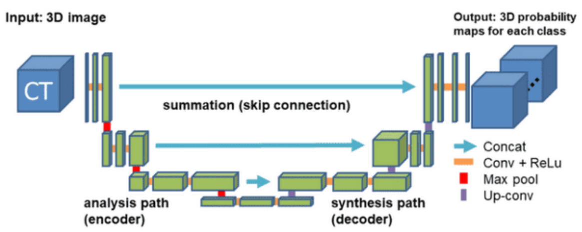
The U-Net architecture is employed for brain tumor segmentation due to its proven effectiveness in medical image analysis tasks. U-Net is a convolutional neural network (CNN) designed for semantic segmentation, featuring a contracting path for capturing context and a symmetric expanding path for precise localization. The model is trained using a combination of binary cross-entropy loss and dice coefficient, optimizing for both pixel-wise accuracy and segmentation overlap. Transfer learning may also be considered, using pre-trained weights on a large dataset to boost performance.

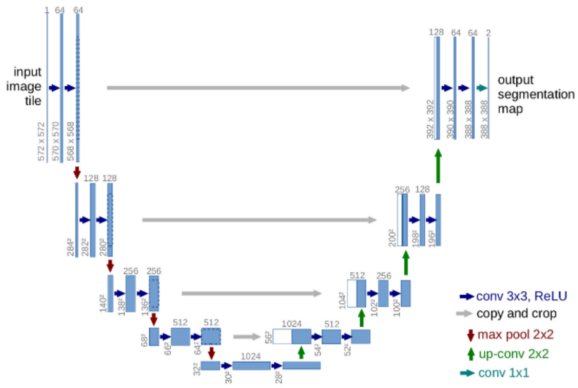
Results and discussion

The evaluation of the U-Net model is conducted using standard metrics such as precision, recall, F1-score, and Dice coefficient. The results showcase the model's ability to accurately segment brain tumors from MRI scans, highlighting its superiority over traditional methods. Furthermore, visualizations of the segmentation maps and a qualitative assessment of the model's performance on various tumor types are presented. Potential challenges and limitations, such as sensitivity to dataset variations, are discussed, along with avenues for future improvements and research directions in brain tumor segmentation using deep learning techniques.



Note on the use of U-Net here





<u>U-Net Definition</u>

U-Net is a deep learning architecture designed for biomedical image segmentation, especially in scenarios with limited annotated data.

Architecture Overview:

- U-Net consists of two main paths:
    - **Contracting Path (Encoder)**: Captures contextual information and reduces spatial resolution.
    - **Expansive Path (Decoder)**: Decodes the encoded data and generates a segmentation map.
- Skip connections connect the contracting and expansive paths to preserve spatial information.

<u>Working Process</u>

Input

The input to the U-Net model is a brain MRI scan, represented as a 2D or 3D image

Contracting Path (encode)

- Encoder layers perform convolutional operations:
    - Reduce spatial resolution while increasing depth.
    - Capture increasingly abstract representations.
- The input image progressively shrinks in height and width but increases in channel depth.
- At the bottleneck, a final convolution operation generates a feature map.

Bottleneck

- A bottleneck layer captures high-level contextual information using multiple convolutional blocks.

Expansive Path (decod)

  - Decoder layers upsample the feature map:
    - Increase spatial resolution while reducing channel depth.
  - Skip connections from the contracting path help refine features.
  - The output image represents a binary segmentation map (foreground vs. background).

Output Layer

- The final layer uses a 1x1 convolution with a sigmoid activation function to produce pixel-wise predictions, representing the probability of tumor presence.

Loss Function

- Binary Cross-Entropy Loss measures the dissimilarity between predicted and true pixel-wise binary labels.

Evaluation Metrics

- Dice Coefficient, Precision, Recall, and F1-Score are commonly used metrics for assessing segmentation performance.

Training

- Adam or SGD with momentum is employed as the optimization algorithm to adjust model parameters.
- Learning rate scheduling may be used to facilitate convergence.

Regularization

- Dropout is applied to prevent overfitting during training.