	Mandatory Summer Internship Report	Code : DO-PFE-01
		Revision Index : 00
		Edition : 07/2022

Mandatory Summer Internship Report

Field: Computer Networks and Telecommunications

Level: 4th Year

Subject :

Brain Metastases 3D Segmentation

Completed by: **Mariem Makni**

Host Company:

Laboratoire Ensi

Supervisor: Mme. Dorsaf Sébai	**Opinion of the Internship Committee**

Academic Year: 2023-2024

1. Introduction:

Many medical imaging modalities generate a large number of images per patient in each scan. For instance, a typical short-axis cardiac MRI sequence consists of more than 200 images, and manual segmentation of ventricles from all these images might take more than 20 minutes. Only a limited set of measurements associated with the scan is computed in regular clinical practice. Automating the analysis could lead to a comprehensive analysis of the underlying condition. However, automating the analysis using traditional image processing approaches poses a number of challenges, including accurately identifying the regions of interest. Recent studies have shown that deep learning approaches have the ability to accurately delineate the regions of interest provided that they are trained with sufficient data. This study aims to utilize deep learning to delineate Brain Metastases from 3D MRI images.

This internship was conducted within the framework of developing a robust and accurate segmentation model for brain metastases, utilizing state-of-the-art deep learning architectures. Throughout the process, various challenges were encountered and addressed. This report details the steps taken, the knowledge applied, and the outcomes achieved during this project, highlighting the intersection of theoretical learning and practical application in a cutting-edge research environment.

2. Presentation of the Host Laboratory

2.1. Presentation of ENSI

The laboratory where I conducted my research internship is part of **École Nationale des Sciences de l'Informatique (ENSI).** Founded in 1984, ENSI holds the distinction of being the first engineering school in the country dedicated specifically to computer science. ENSI's educational philosophy is centered on fostering lifelong learning, equipping its students with the essential skills and knowledge required to excel in both professional and personal spheres.

2.2. ENSI's Educational Offering

ENSI prepares students with a solid foundational knowledge, enabling them to specialize in various fields throughout their careers. Currently, ENSI offers six specialized tracks for its engineering students:

1. Software Engineering
2. Artificial Intelligence
3. Data Science for Computer Vision
4. Computer Engineering for Finance
5. Internet of Things (IoT) Services and Technologies
6. Embedded Systems and Software

2.3. Research Structures at ENSI

The school boasts four major research structures:

1. **CRISTAL Laboratory (LR99ES25)**: Established in 1999, the CRISTAL lab focuses on Networks, Image Systems, Architecture, and Multimedia.
2. **RIADI Laboratory (LR99ES26)**: Also founded in 1999, RIADI specializes in Arabized Informatics, Integrated Documentation, and Software Engineering.
3. **HANA Laboratory (LR14ES04)**: Launched in 2014, HANA is dedicated to Innovative Distributed Applications and Advanced Heterogeneous Networks.
4. **LARIA (UR 22ES01)**: Established in 2021, LARIA is the Research Unit for Artificial Intelligence, reflecting ENSI's commitment to staying at the forefront of AI research.

2.4. Collaborative Partnerships

ENSI’s success is also bolstered by its extensive network of partners, including industry leaders, academic institutions, and governmental bodies.

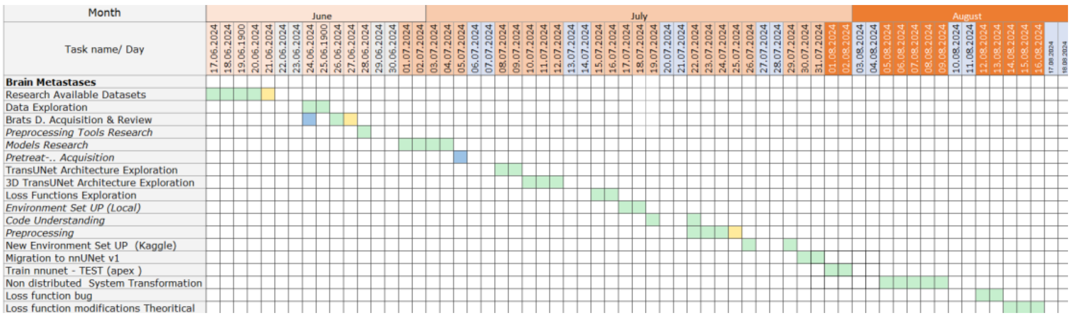
Figure 1 - ENSI's Partners

3. Goals

The primary objectives of my internship were centered around the implementation and customization of advanced deep learning models for brain metastasis segmentation. The tasks involved in this process are outlined below:

- **Choice of Dataset:** The first objective was to select a suitable dataset for training and testing the segmentation models.
- **Choice of Model:** The next step involved selecting an appropriate model for our research.
- **Loss Function Adjustment to Work with Hausdorff-Based Approach:** The aim is to evaluate the results of the model on the same dataset using a base loss function compared to one that includes the Hausdorff distance metric. This requires modifying the loss function, originally detailed in [1] for binary applications, to accommodate multilabel 3D segmentation tasks.

4. Internship’s journal :



Month	June														July																																						August															
Task name/ Day	17.06.2024	18.06.2024	19.06.1900	20.06.2024	21.06.2024	22.06.2024	23.06.2024	24.06.2024	25.06.1900	26.06.2024	27.06.2024	28.06.2024	29.06.2024	30.06.2024	01.07.2024	02.07.2024	03.07.2024	04.07.2024	05.07.2024	06.07.2024	07.07.2024	08.07.2024	09.07.2024	10.07.2024	11.07.2024	12.07.2024	13.07.2024	14.07.2024	15.07.2024	16.07.2024	17.07.2024	18.07.2024	19.07.2024	20.07.2024	21.07.2024	22.07.2024	23.07.2024	24.07.2024	25.07.2024	26.07.2024	27.07.2024	28.07.2024	29.07.2024	30.07.2024	31.07.2024	01.08.2024	02.08.2024	03.08.2024	04.08.2024	05.08.2024	06.08.2024	07.08.2024	08.08.2024	09.08.2024	10.08.2024	11.08.2024	12.08.2024	13.08.2024	14.08.2024	15.08.2024	16.08.2024	17.08.2024	18.08.2024
Brain Metastases																																																															
Research Available Datasets	🟩	🟩	🟩	🟩	🟨																																																										
Data Exploration								🟩	🟩																																																						
Brats D. Acquisition & Review								🟦		🟩	🟨																																																				
Preprocessing Tools Research												🟩																																																			
Models Research															🟩	🟩	🟩	🟩																																													
Pretreat-.. Acquisition																			🟦																																												
TransUNet Architecture Exploration																						🟩	🟩																																								
3D TransUNet Architecture Exploration																								🟩	🟩	🟩																																					
Loss Functions Exploration																													🟩	🟩																																	
Environment Set UP (Local)																															🟩	🟩																															
Code Understanding																																	🟩			🟩																											
Preprocessing																																				🟩	🟩	🟩	🟨																								
New Environment Set UP (Kaggle)																																								🟩			🟩																				
Migration to nnUNet v1																																												🟩	🟩																		
Train nnunet - TEST (apex)																																														🟩	🟩																
Non distributed System Transformation																																																		🟩	🟩	🟩	🟩	🟩									
Loss function bug																																																									🟩	🟩					
Loss function modifications Theoritical																																																											🟩	🟩	🟩		

5. Work Done:

5.1. The BraTS METS 2023 Dataset

The BraTS METS 2023 dataset [2] was chosen as the primary dataset for this project due to its comprehensive and well-annotated nature. It is one of the most robust datasets available, featuring a wide selection of cases from prestigious institutions.

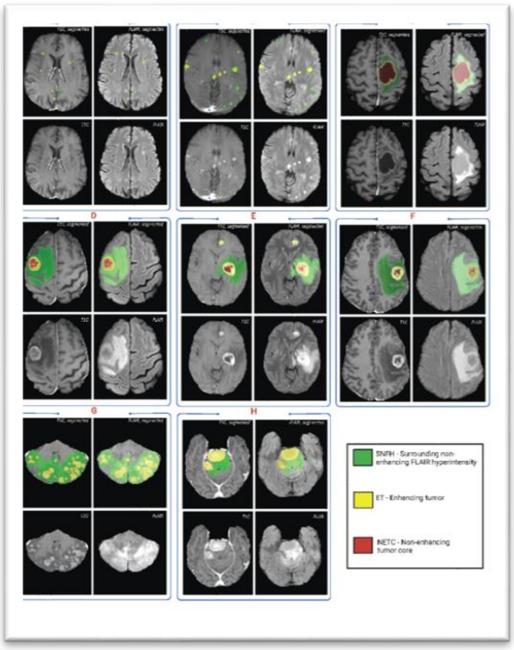

D
E
F
G
H
SNFH - Surrounding non-enhancing FLAIR hyperintensity
ET - Enhancing tumor
NETC - Non-enhancing tumor core

Figure 2 - BraTS METS Data samples

The dataset includes four distinct channels of information, each representing a different volume of the same brain region:

- **T1 Native (T1)**: Provides a baseline image of the brain.
- **T1 Post-contrast (T1CE)**: Highlights disruptions in the blood-brain barrier, making it crucial for identifying and delineating tumor boundaries.

- **T2-weighted (T2)**: Visualizes fluid content and edema, offering high-contrast images of brain tissues, which is vital for visualizing the tumor environment.
- **T2 Fluid Attenuated Inversion Recovery (FLAIR)**: Suppresses fluid signals, making it essential for identifying and segmenting peritumoral edema.

The annotations provided in this dataset are particularly valuable for 3D segmentation tasks. These include:

- **Label 1: Non-enhancing Tumor Core (NETC)**: Represents the central, non-enhancing part of the tumor, often necrotic or cystic, important for assessing the internal structure of the tumor.
- **Label 2: Peritumoral Edematous/Infiltrated Tissue (SNFH)**: Surrounding non-enhancing FLAIR hyperintensity, crucial for differentiating infiltrative tumor tissue from other brain tissues.
- **Label 3: Enhancing Tumor (ET)**: Indicates active tumor regions, critical for detailed tumor characterization.

The dataset's structure and comprehensive annotation make it an excellent choice for developing and evaluating advanced segmentation models.

5.2. Choice of Model:

Recent advancements in machine learning, particularly the introduction of transformer-based models, have opened new avenues in multiple fields. These models offer substantial improvements in capturing global contextual information compared to traditional solutions. With the rise of Vision Transformers (ViTs) in 2016, attempts to apply these architectures to medical 3D segmentation tasks raised. Among the models explored, 3D TransUNet [3] stood out due to its hybrid architecture that integrates transformers with a U-Net backbone, leveraging the strengths of both convolutional and transformer-based models.

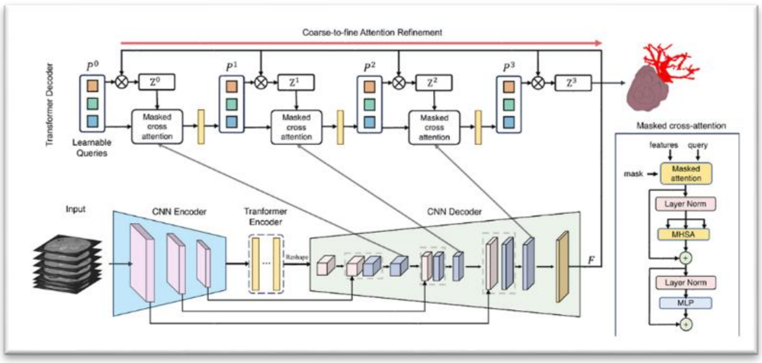


Figure 3 - 3D TransUNet Architecture

3D TransUNet offers three configurations:

- **Transformer in Encoder Only**: Suitable for large-scale segmentations like organs.

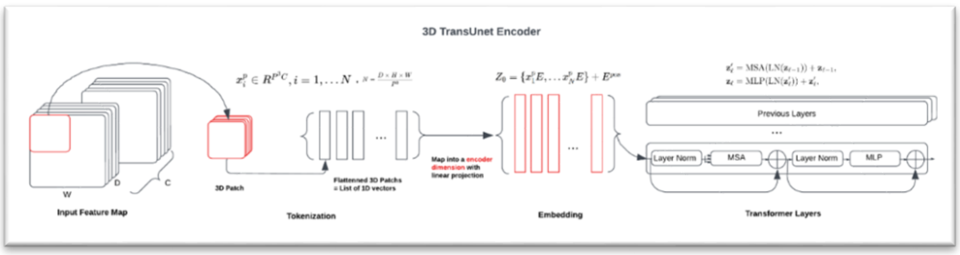


Figure 4 - 3D TransUNet Encoder

- **Transformer in Decoder Only**: Optimal for small parts segmentation, which is particularly relevant to our case.

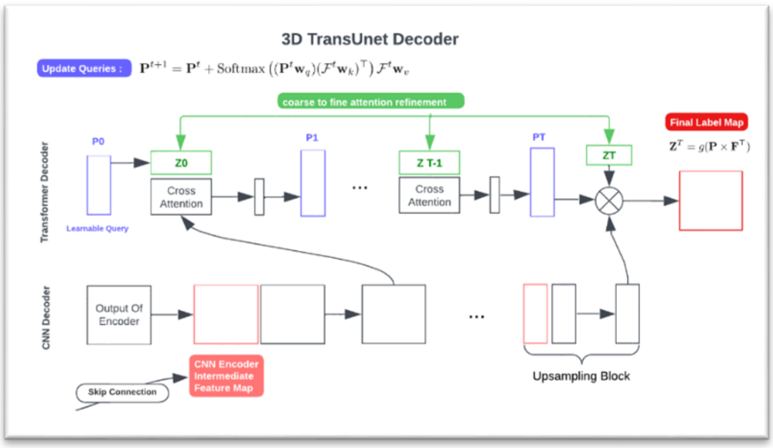


Figure 5 - 3D TransUNet Decoder

- **Transformer in Both Encoder and Decoder**: Found to be less effective, likely due to interference between the encoder and decoder mechanisms.

Configuration Choice:

The integration of transformers in the decoder allows for refined attention mechanisms, enhancing the model's ability to accurately segment small lesions. Which gives it superior performance in segmenting small, intricate structures like brain metastases.

5.3. Introduction to nnUNet

nnUNet [4] is a highly versatile and automated deep learning framework designed to simplify the development of medical image segmentation models. It automatically configures itself based on the characteristics of the dataset, reducing the manual effort typically required for model setup and data processing.

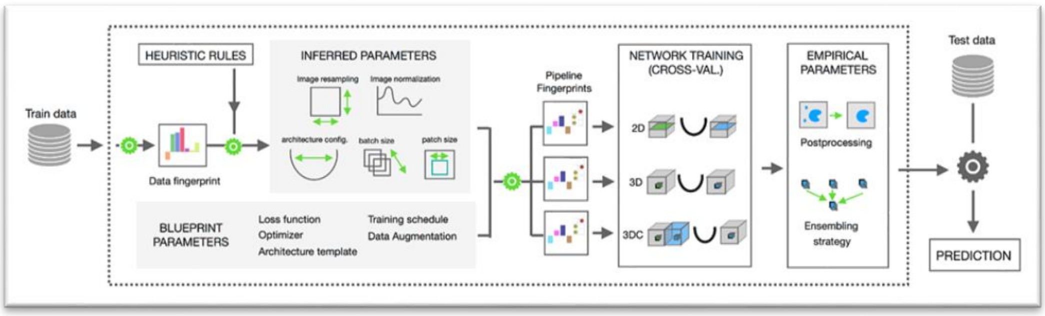


Figure 6 - nnUNet Architecture

Versions Overview:

- **nnUNet v1**: The initial version laid the groundwork for automated configuration and quickly gained popularity across various medical imaging challenges. However, it presents stability issues on Windows OS, and lacks some of the user-friendly features introduced in later versions.
- **nnUNet v2**: This version builds on the original, offering enhanced automation in handling datasets, broader support for various image file formats, and a more user-friendly experience overall. It provides greater customization options, making it adaptable to a wider range of applications.

Given that 3D TransUNet is built on the **nnUNet v1 framework**, this version was utilized.

5.4. Data Format Conversion and Preprocessing

To effectively utilize the nnUNet v1 framework, it is crucial to organize the dataset according to its specific directory structure [5] and file format requirements. The nnUNet framework expects the data to be structured in a particular way to facilitate automated pre-processing, training, and evaluation.

Directory Setup:

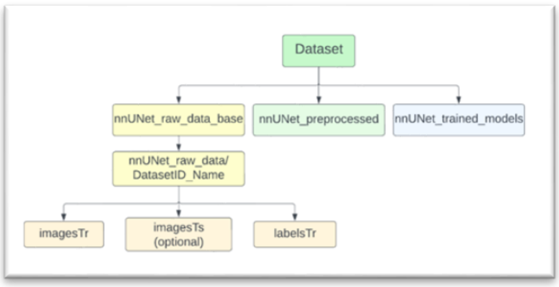


Figure 7 - nnUNet's Data Setup Requirements

1. **Dataset Preparation and Exploration**:

The training data was initially organized into separate folders for each patient. Each folder contained five files: four imaging modalities (T1, T1CE, T2, and FLAIR) and one segmentation mask file. Consequently, I implemented a script to restructure the data as required.

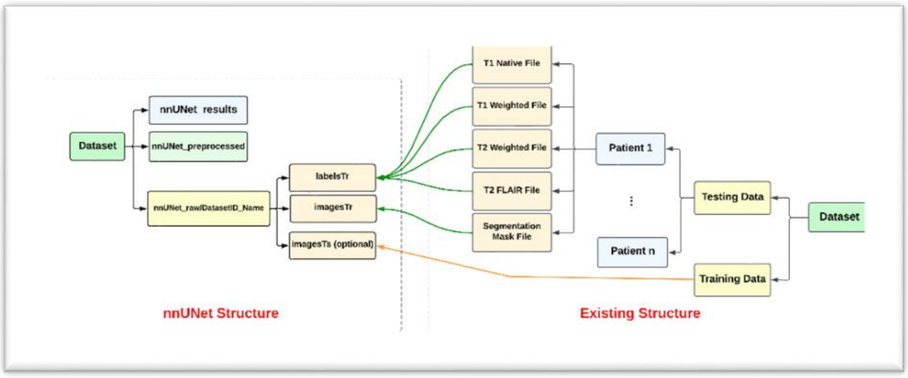


Figure 8 - Data Structure Adjustments

Additionally, another script was implemented to create the `dataset.json` needed by nnUNet. This file includes information about the modalities, labels, and any specific preprocessing instructions.

I conducted thorough data exploration to understand the nuances of handling 3D MRI data and presented the findings.

Pre-processing:

Once the data was organized, I proceeded with pre-processing using the nnUNet framework. This step involves a series of transformations that prepare the data for training.

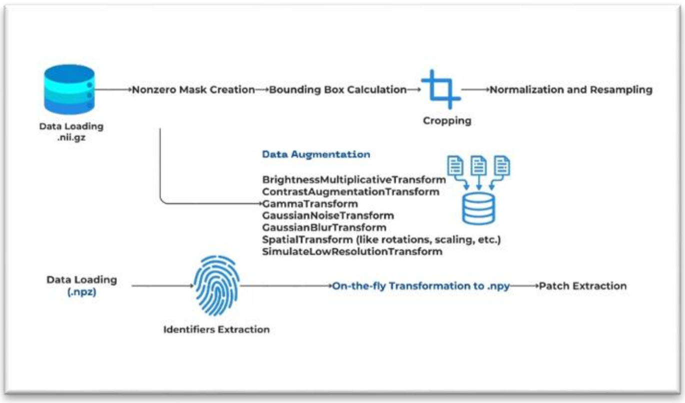


Figure 9 - Preprocessing Step

5.5. Training

During the training phase, I encountered several challenges:

1. **APEX Library Installation**:

The installation of the APEX library [6], which is essential for mixed precision training, presented some issues. To overcome this, I stopped it from isolating the build during the installation process to bypass certain steps that were causing failures.

Figure 10 -- NVIDIA Apex: Tools for Easy Mixed-Precision Training in PyTorch

2. **Adapting 3D TransUNet to Non-Distributed Training**:

The 3D TransUNet model is originally designed to operate in a distributed training environment, utilizing multiple GPUs. It relies on the Distributed Data Parallel (DDP) [7] implementation, which is more efficient than DataParallel (DP) [8] for large-scale training tasks.

However, due to resource limitations, I needed to adapt the model to work in a non-distributed, single-GPU setup. This required several modifications to the code. The changes focused on removing distributed training components and improving error handling through additional debugging and dynamic fallback logic. Furthermore, the network trainer and loss computation were streamlined to accommodate model-specific outputs effectively in a non-distributed environment.

Component	**Details**	**Modifications**
Initialization and Configuration	Originally, configurations included basic settings.	-Added parameters like layer_decay, lr_scheduler_name, and model-specific flags such as disable_ds, is_spatial_aug_only. -Updated initialize_network to set up a transformer-based model and specific convolution and normalization settings.
Data Augmentation	Standard methods were used for data augmentation.	Introduced setup_DA_params_BraTSRegions, a method tailored for brain tumor segmentation.
Distributed Training	Basic settings for distributed training across GPUs.	Configured batch size management and specific training strategies and optimizations for distributed environments.
Loss Computation and Backpropagation	Standard loss computation and backpropagation methods were used.	Extended the compute_loss method to handle transformer outputs and multiple outputs for deep supervision. Included NaN value checks and additional debug outputs in backpropagation.
Optimizer and Scheduler	Basic optimizer and scheduler configurations were present.	Extended configurations to include warmup phases and custom learning rate schedules.

Table 1 – Summary of implemented modifications

The issue remained and further modifications needs to be done.

5.6. Loss Function Adaptation

Another major task involved adapting the loss function to work with the specific requirements of the 3D TransUNet model. The original loss function proposed in [1] was designed for binary segmentation tasks. My goal was to adapt this loss function for multi-label segmentation, which is necessary for accurately segmenting different tumor regions.

To achieve this:

- **Target and Prediction Tensors:** Each label should be treated as a separate binary segmentation task within the overall calculation. For each label, we must create a binary mask from the multi-label segmentation maps of both predictions and targets.
- **Compute Loss for Each Label:** Loop over each label, apply the binary loss function, and store the results.
- **Aggregate the Losses:**
  - **Sum:** Add up the losses across all labels to get a total loss for the batch.
  - **Mean:** Compute the mean of these losses to normalize the loss across labels.

This cannot be tested until the previous issue is resolved. Thus, it is still in process.

6. Consolidation of Acquired Knowledge:

During this internship, I had the opportunity to apply and expand upon the essential knowledge and skills I have acquired throughout my academic journey. The theoretical foundations and practical insights from various courses were instrumental in successfully navigating the challenges of this project.

My **Python course** played a pivotal role in the development process, as Python is the primary programming language used in machine learning and deep learning frameworks. The familiarity with Python allowed me to efficiently script and automate tasks, manipulate large datasets, and integrate various tools and libraries necessary for the project.

The **Deep Learning course** was particularly valuable, as it provided me with a comprehensive understanding of model architectures, including CNNs, and the critical components like hyper parameters, loss functions, and up-sampling techniques. This course equipped me with the knowledge needed to understand and implement the 3D TransUNet model, adapt it to the specific requirements of the project, and fine-tune it for optimal performance.

In the **Image Processing course**, I learned about various pre-processing techniques, which I applied during the data preparation stages of this project. The course also covered essential image enhancement and transformation methods that were instrumental in preparing the medical images for analysis and segmentation.

The **Mixed and Augmented Reality course** provided a solid grounding in 3D data structures and the concept of slices (frames). This knowledge was directly applicable when working with medical imaging data, where I frequently dealt with multi-dimensional MRI scans.

Finally, the **PPP (Professional Practice Project)** was an invaluable experience that bridged the gap between academic learning and real-world application. Through this project, I gained hands-on experience in a research-oriented environment, which greatly enhanced my understanding of the complexities and challenges involved in developing and deploying machine learning models in practical settings

Collectively, these courses provided me with the necessary theoretical and practical knowledge to understand and implement these advancements.

7. Conclusion:

This internship provided a valuable opportunity to delve into the complexities of medical image segmentation and apply advanced deep learning techniques to a real-world problem. The research work is still in progress, and as we know, computer vision, especially in this field, is computationally expensive and resource-intensive.

Through this project, I not only reinforced my understanding of key concepts in machine learning, deep learning, and image processing but also gained practical experience in adapting models to specific resource constraints and enhancing model performance through customized loss functions.

This experience has further solidified my interest in the intersection of technology and healthcare, inspiring me to continue exploring and contributing to advancements in this critical area.

Bibliographies

[1] Karimi D, Salcudean SE. Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks. IEEE Trans Med Imaging. 2020 Feb;39(2):499-513. doi: 10.1109/TMI.2019.2930068. Epub 2019 Jul 19. PMID: 31329113.

[2] Moawad, A. W., Janas, A., Baid, U., Ramakrishnan, D., Saluja, R., Ashraf, N., Jekel, L., Amiruddin, R., Adewole, M., Albrecht, J., … et al. (2023). The Brain Tumor Segmentation (BraTS-METS) Challenge 2023: Brain Metastasis Segmentation on Pre-treatment MRI. *arXiv*. https://doi.org/10.48550/arXiv.2306.00838

[3] Chen, J., Mei, J., Li, X., Lu, Y., Yu, Q., Wei, Q., Luo, X., Xie, Y., Adeli, E., Wang, Y., Lungren, M., Xing, L., Lu, L., Yuille, A., & Zhou, Y. (2023). *3D TransUNet: Advancing Medical Image Segmentation through Vision Transformers*. https://doi.org/10.48550/arXiv.2310.07781

[4] Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S., Wasserthal, J., Koehler, G., Norajitra, T., Wirkert, S., & Maier-Hein, K. H. (2018). *nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation*. https://doi.org/10.48550/arXiv.1809.10486

[5] nnUNet v1 - Dataset conversion instructions. https://github.com/MIC-DKFZ/nnUNet/blob/nnunetv1/documentation/dataset_conversion.md

[6] NVIDIA- Apex (A PyTorch Extension). https://nvidia.github.io/apex/

[7] Network Trainer DDP. https://github.com/MIC-DKFZ/nnUNet/blob/nnunetv1/nnunet/training/network_training/nnUNetTrainerV2_DDP.py

[8] Network Trainer DP. https://github.com/MIC-DKFZ/nnUNet/blob/nnunetv1/nnunet/training/network_training/nnUNetTrainerV2_DP.py