Course/Programme:	MSc. Artificial Intelligence
Module Name and Code:	Computer Vision and Deep Learning (AIN7302)
Student ID:	2011184
Student Name:	Bogomil Iliev
Tutor:	Dr. Naveed Islam and Dr. Eugen Harinda
Assessment Number:	2 of 2
Assessment Type:	Portfolio
Assessment Title:	A Deep learning-based Computer Vision Solution.
Indicative Word Count:	3000 words.
Weighting:	60% of overall module grade.
Submission Deadline:	22.12.2024 (no later than 23:59)
Submission Date:	

Learning Outcomes assessed:

LO2: Reconstruct and critically evaluate existing computer vision and deep learning models to solve real-world problems, demonstrating originality, creativity and independence.

LO3: Apply computer vision techniques while critically evaluating and discussing competing 2D to 3D methods.

Student ID#: 2011184

Detecting and Diagnosing Brain Tumours through Computer Vision and Deep Learning Techniques.

Contents

Table of Figures

Abstract

Brain Tumours pose a critical health concern, bringing forward the need of early and accurate detection to optimise patient outcomes. Computer vision and deep learning methods, present a promising avenue for automating the process of tumour identification, classification and monitoring in MRI scans. This work explores the implementation of CNNs and their variants, such as U-Net and Multimodal approaches, for segmenting and classifying regions with malformations in radiological data. A critical evaluation of both 3D and 2D approaches is performed to assess their performance on the latter tasks and identification of their benefits and drawbacks for the aforementioned needs of the professionals. Our findings demonstrate that these techniques can achieve high accuracy, offer valuable insights into tumour boundaries, and assist clinicians in making timely and evidence-based decisions. This is proved by the recreation of one of the discussed models that backed up its findings and rectified the reproducibility of the original model.

1. Introduction

Cancer diseases have plagued humankind for centuries, but with the technological advancements made through the Industrial Revolutions and our rapidly changing way of living, the number of affected individuals worldwide is booming. One of its variations that is significantly hard to locate and diagnose in its nascent phases is the Brain Tumour. It can be described as a formation of somatic cells that expands frantically and without a pattern in the encephalon (NHS, 2023). Such gatherings of cells can emerge in two basic forms – benign or malignant. They may appear within the brain or proliferate from other somatic sections. These malformations are not consistent in their size, shape or number. They may form in different parts of the brain as a single entity or as multiple clusters.

Identifying cancerous cells is done through conventional approaches that employ medical imaging as a technology. CT, PET and MRI are amongst the most common methods. Thus, due to the intricate structure of the viscus, the localisation and therapy of BTs are strenuous tasks, that require high levels of expertise and experience by the assessing medical personnel. To properly distinguish the form, measurements and position of the cancer, the professionals have to go through different modalities and carefully segment them, which is laborious and time-consuming (Bauer *et al.*, 2013; Soomro *et al.*, 2023).

Respectively, the timely recognition of the disease is of utmost importance in order to treat it and prevent it from claiming lives. Contemporary advancements in AI and Computer Vision have expanded implementations in medical imaging through the release of open-source datasets like BraTS 2015 - 2023 (Kazerooni *et al.*, 2023).

Hence, ML and DL have proved to be applicable to the problem region and alleviated the diagnostic process greatly. These can be grouped into several sections depending on their goals - ranging from image classification of malignant and benign malformations, detection

of cancerous cells, segmentation, to characterisation of tumour types (Pereira *et al.*, 2016; Litjens *et al.*, 2017; Dong *et al.*, 2017).

This work is discussing in depth two more contemporary approaches to assist doctors in early carcinoma detection and is aiming at reconstructing one of them to confirm its findings. Namely, the identification of such cell gatherings through Multimodal Neural Networks and MRI image segmentation via U-net applications. MNNs have only recently become a point of interest in medical imaging research. Several authors have already employed the models for classification tasks in the 3D plane, however inconsistent and non-satisfactory outcomes were delivered. Hence, one of the model's discussed below, managed to achieve high scores and still pertain the main concept of extracting both visual intel from the picture information and concatenating it with qualitative one deriving from the empirical representations produced by medical staff. This allows the architecture to better asses the cases presented by "drawing from two wells" (Latif *et al.*, 2018; Ma and Jia, 2020; Curci and Esposito, 2024).

The other technique that is assessed is a NN architecture – a modified U-net model in order to segment samples and support physicians in their daily duties. Here, the focus is in the 2D plane, where images are pre-processed from the raw information (Al Nasim *et al.*, 2022).

2. Critical Evaluation of Existing Models

2.1. Segmentation of BTs by Employing a Deepened 2D U-net

The proposed architecture utilises a modified 2D U-net. **The data** that was used to train, validate and evaluate the results of the project were the BraTS 2017 – 2020 datasets. They accommodate Giloma screening samples in a 3D format, that are labelled by doctors from different organisations. The information comprises of four variations of MRI screening samples (T1, T2-weighted, post-contrast T1-weighted, FLAIR). Each one of them retains 155

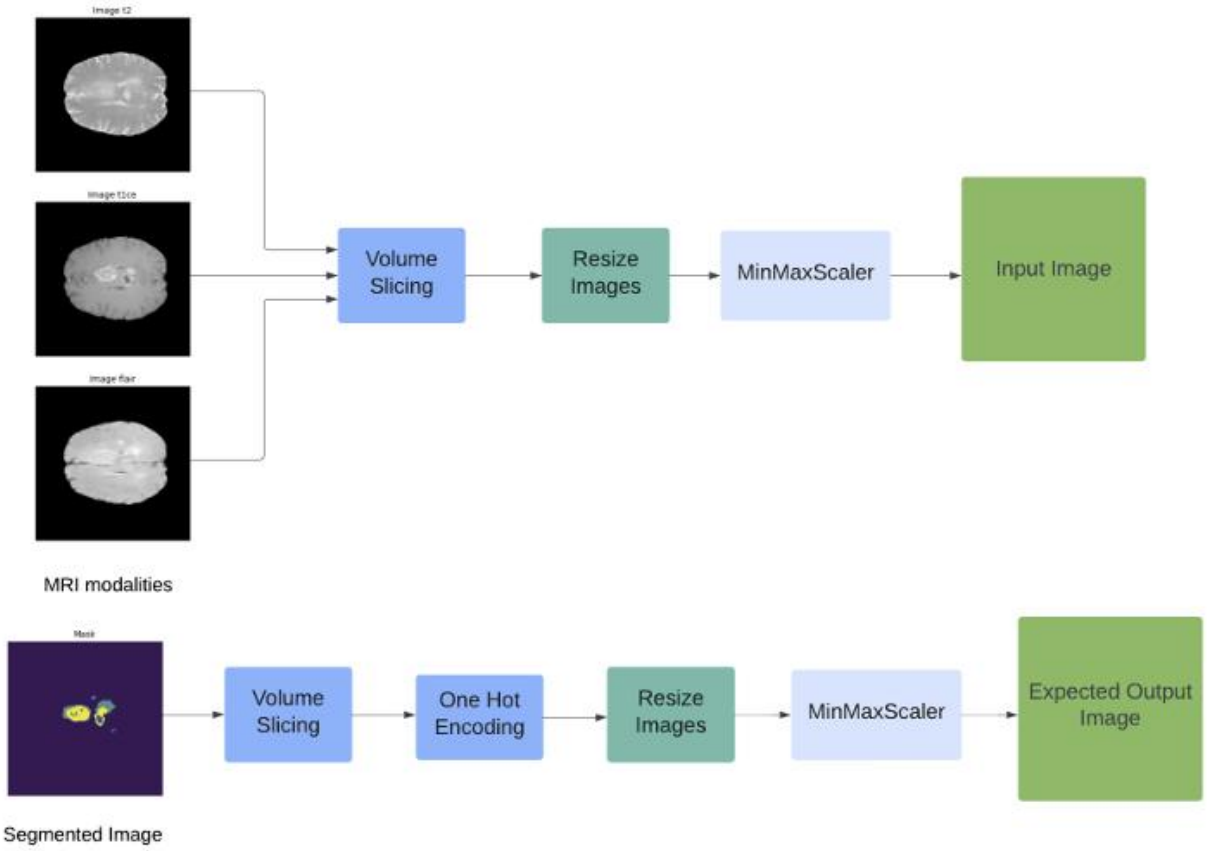
segments. The carcinoma areas that are labelled are – the necrotic (1) and the consistent carcinoma nucleus (1) - sharing a mutual label; the growing matter (4), the perilesional edema (2) and 0 for the rear setting of the sample (Kazerooni *et al.*, 2023; Al Nasim *et al.*, 2022).

Figure 1 - BraTS Datasets Overview (Kazerooni et al.*, 2023; Al Nasim* et al.*, 2022).*

Dataset	Type	Patients(No of Cases)	HGG	LGG
BraTS 2017	Training	285	210	75
	Validation	46		
BraTS 2018	Training	285	210	75
	Validation	66		
BraTS 2019	Training	335	259	76
	Validation	125		
BraTS 2020	Training	369	293	76
	Validation	125		

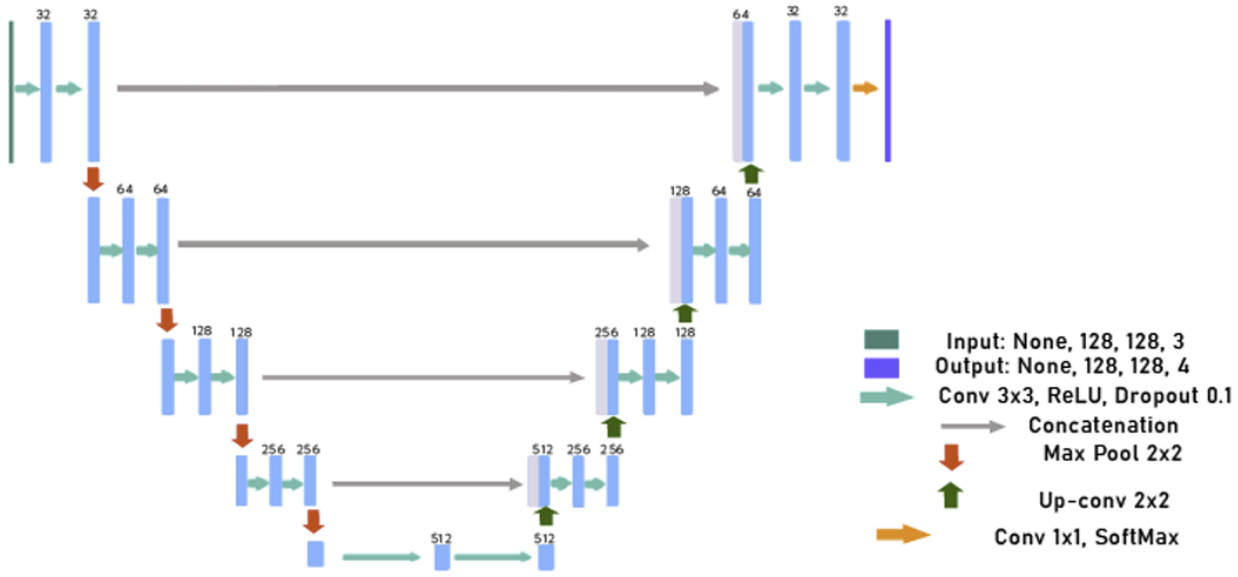
As it can be observed in **Figure 1**, the information in all four batches is spread is very alike. **Figure 2** explains what actions were taken to adjust the data to requirements. Since the data corpus consists of four distinct image variations for each slice, the authors utilised only T1ce and FLAIR shots in order to reduce hardware costs. Also, the dark regions of the volumes carry no information, hence they've been cut out and only the meaningful part of the data was employed, so time-efficiency can be improved. The data fed to the algorithm was reduced to the appropriate for the algorithm dimensions and thus were the resulting images. Normalisation was also performed through MinMaxScaler and then one hot encoding was administered. Then it was divided into proportions of 17:5:3 for training, validation and testing (Al Nasim *et al.*, 2022).

Figure 2 - Pre-processing (Al Nasim et al.*, 2022).*



In **Figure 3** can be observed the that the model consists of two portions – the extending one and the shrinking. Where the **architecture** shrinks - a CNN is being employed comprising of a double 3x3-layer system with a ReLU activation. The maxpool is set to 2x2, thus the count of attributes increases twofold with each iteration. The extensive segment involves 2x2 transposed convolutional layers to upscale the activation map and join it with the respective map of the previous portion of the network. At the end of it stand two ReLU-activated convolutional layers of 3x3. Every upscaling stride is followed by a 50% reduction in the count of attributes. The output one is a 1x1 with a SoftMax function (Al Nasim *et al.*, 2022).

Figure 3 - Architectural Design of 2D U-net (Al Nasim et al.*, 2022).*



The network was produced via the Keras and Tensorflow libraries and it was trained over 235 epochs with a batch size of 1. The optimisation algorithm utilised was Adam and the Loss function employed – Categorical cross-entropy. Techniques like model checkpoint, csv logging and early stopping were also used. To assess the architecture's deliverables **metrics** such as dice coefficient, Precision, Mean-IoU, Sensitivity, Accuracy, and Specificity were utilised (Al Nasim *et al.*, 2022).

Figure 4 – Performance of the 2D U-net over the different datasets (Al Nasim et al.*, 2022).*

Dataset	Loss	Accuracy	Mean IoU	Precision	Sensitivity	Specificity	Dice Score	Dice Score (necrotic/ core)	Dice Score (edema)	Dice Score (enhancing)
BraTS 2017	0.0056	0.9980	0.9637	0.9973	0.9970	0.9972	0.8453	0.8782	**0.9545**	**0.9490**
BraTS 2018	0.0057	0.9979	0.8927	0.9972	0.9970	0.9940	0.8160	0.8843	0.9368	0.8348
BraTS 2019	**0.0054**	**0.9981**	**0.9130**	**0.9974**	**0.9971**	**0.9991**	**0.8409**	0.8717	0.9506	0.9427
BraTS 2020	0.0056	0.9980	0.8935	0.9973	0.9970	0.9983	0.8300	**0.8846**	0.9478	0.8544

As it can be seen from **Figure 4**, the architecture's strongest **performance** was recorder on the BraTS 2019 data batch – an Accuracy of 99%, Precision – 99%, Mean IoU of 91%, Dice Score of 84% and Specificity – 99%, alongside a notably low loss. The variation over the different sets is minimal (Al Nasim *et al.*, 2022).

Through the conducted tests in this work, the authors have managed to prove that their 2D architecture is outmatching other known 3D CNN & ML implementations for segmentation purposes. The good side of this approach is that it is significantly more economical and time-effective than its 3D rivals for the aforementioned task, however it should be noted that the architecture's incapacity to utilise granular details leads to a reduction in the intel gained. Hence, a further field for development can be to employ a 3D variation of the U-net (Habib *et al.*, 2020; Hambarde *et al.*, 2020; Al Nasim *et al.*, 2022).

2.2. Classification of BTs by Utilising a 3D Multimodal Neural Network.

The goal of this architecture is to determine whether a patient's screening classifies them as cancer positive or negative and to aid physicians to diagnose oncological formations in the brain at an early stage. This was achieved through a DenseNet-based network (Curci and Esposito, 2024).

The **data** employed originated from the 2015 version of the BraTS dataset - 3762 MRI samples in 240x240 format within the RGB gamma. The information is binary labelled, where "0" stands for "healthy" and "1" for ill. The tabular information includes 13 empirical attributes in first and second degree. Since, the information batch is a bit uneven, some images were arbitrarily discarded. The empirical representations were also normalised to adhere to the binary requirement (Bohaju, 2020; Curci and Esposito, 2024).

Figure 5 - Architecture of the 3D Multimodal Model (Curci and Esposito, 2024).

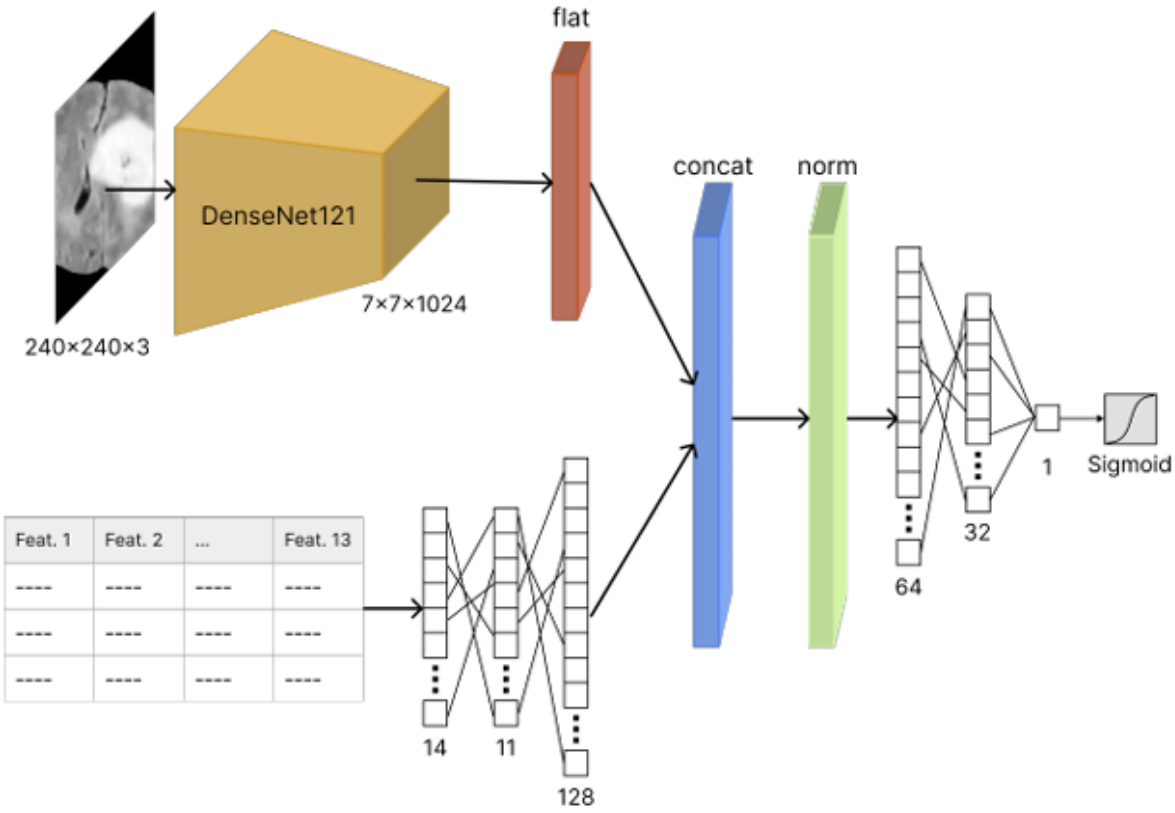


Figure 5 represents the employed **architecture**. The algorithm has dual input branches. One learns representations from the imagery part of the dataset through the DenseNet121 and accepts 240x240 samples in the RGB gamma, and the other processes the empirical information through a plain CNN with ReLU activation. The resulting images from the first algorithm are 7x7x1024. The delivered results by the algorithms are then joined and normalised. Following, the provided embedding is passed to a third ReLu-activated NN that delivers its decision to an output Layer with a SoftMax activation and two classes (Huang *et al.*, 2016; Curci and Esposito, 2024).

The model was drafted through the TensorFlow library. The approach was then assessed through a graded cross-validation in ten iterations, where each one retained approximately equivalent number of the “ill” and “healthy” labels. The loss function was binary cross-entropy with an Adam optimiser. The pace of learning was 0.001 with a size of batch of 32. The

epochs were capped at 100, where the early stopping technique was employed and taking into consideration the validation loss. This limit was 5 epochs of unsatisfactory results (Curci and Esposito, 2024).

To evaluate the model, the **evaluation metrics** F1, AUC, Precision, Accuracy and Recall were used.

Figure 6 - Results of the 3D Multimodal NN (Curci and Esposito, 2024).

CV Fold	Accuracy	AUC	Loss	Precision	Recall	F1-Score
1	0.99	0.99	0.18	0.99	0.98	0.99
2	0.97	0.97	1.5	0.99	0.95	0.97
3	0.99	0.99	5.6e-05	0.99	0.99	0.99
4	0.98	0.98	1.3	0.98	0.98	0.98
5	0.97	0.98	0.67	0.95	0.99	0.97
6	0.99	0.99	0.84	0.99	0.99	0.99
7	0.99	0.99	0.72	0.99	0.99	0.99
8	0.98	0.98	2.9	0.99	0.96	0.98
9	0.99	0.99	0.22	0.99	0.99	0.99
10	0.99	0.99	0	0.99	0.99	0.99
Avg.	0.99	0.99	0.83	0.99	0.98	0.98

From **Figure 6** is visible that the average Accuracy delivered is 99% with 99% precision and a Loss of 0.8.

The MNNs are quite novel and understudied, but this architecture proves they can be useful by better founding their decisions from two sources, which makes it more explainable. Although, it still managed to deliver results that can compete with robust CNN and ML implementations in the field, it has some downsides - the need of larger datasets to refrain from overfitting, also the cost-effectiveness and time consumption required to train the architecture are greater due to its comprehensive structure. It is also under question how well the model will generalise over novel information (Ullah *et al.*, 2024; Li *et al.*, 2019; Curci and Esposito, 2024).

2.3. Architecture Comparison

Although, the two models have slightly different tasks, it is safe to assume that both of these performed well and would benefit professionals in the early diagnose and detection of carcinomas. The first approach may uplift the ability of doctors to visualise affected tissues easier by removing the need for manual segmentation and administer a planned therapy. However, it should not be ignored that due to the comprehensive structure of the organ, it is quite possible that malicious formations may be obscured by healthy material, which has the potential to mislead the algorithm and throw the expert assessing its output into wrong conclusions. The 2D U-net implementation is significantly less costly and faster than the 3D networks, however it lacks the depth of the 3D CNN approaches. On the other hand, the second 3D detection method provides a more complex and easier to explain technique due to its two-way backed decision taking. Nonetheless, it still needs further research (Li et al., 2019; Ullah et al., 2024; Habib et al., 2020; Al Nasim et al., 2022; Dong et al., 2017; Curci and Esposito, 2024).

3. Reconstruction and Development Approach

This work reconstructed the aforementioned U-net approach in order to try to replicate and verify its results.

3.1. The Employed Data

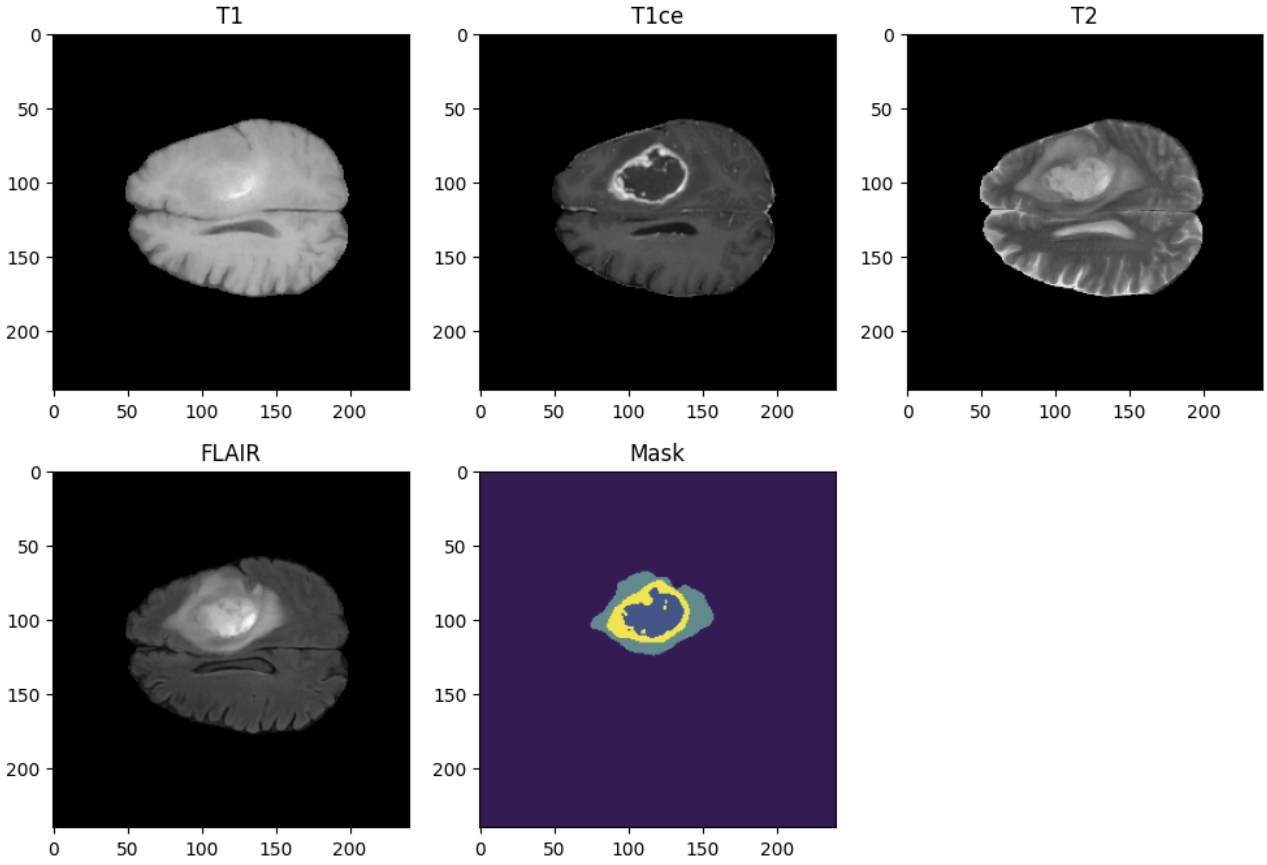
The dataset that was used to reproduce the algorithm was only the 2020 edition of the BraTS dataset. It is a gathering of MRI multimodal visual data with four variations - T1, T1ce, T2, FLAIR. They are annotated by doctors according to their sub-areas, as it was discussed earlier. The information batch entails 369 patients, whose samples have four modalities and a segmentation mask. Each picture sample is the 3D plane (224x224x155), hence there are three assessing angles that are depicted in 2D partitions (Kazerooni *et al.*, 2020). This was

the only dataset employed due to hardware and time constraints, when the reconstruction was being performed.

3.2. Data Preprocessing

The pixel values of the dataset were re-scaled and normalised, so they can be homogenous through all the different modalities in the data, only the doctor provided masks were not processed (Figure 7). This step is important as the model may develop biases and the segmentation can be affected(Brownlee, 2019).

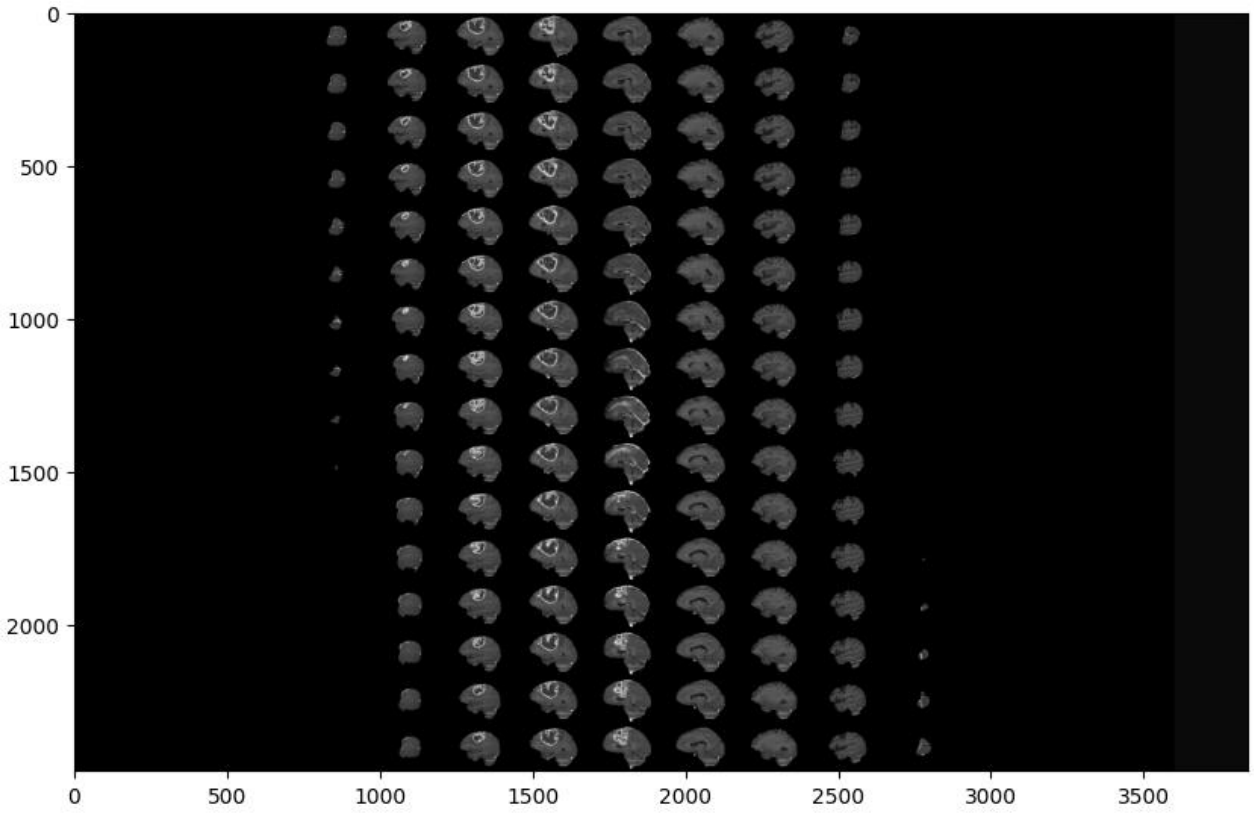
Figure 7 - Rescaled and Normalised Modality Examples.



From the four shot variations only two were utilised, like in the original work – FLAIR & T1ce. This is performed to reduce computational expenses and processing time (Al Nasim *et al.*, 2022).

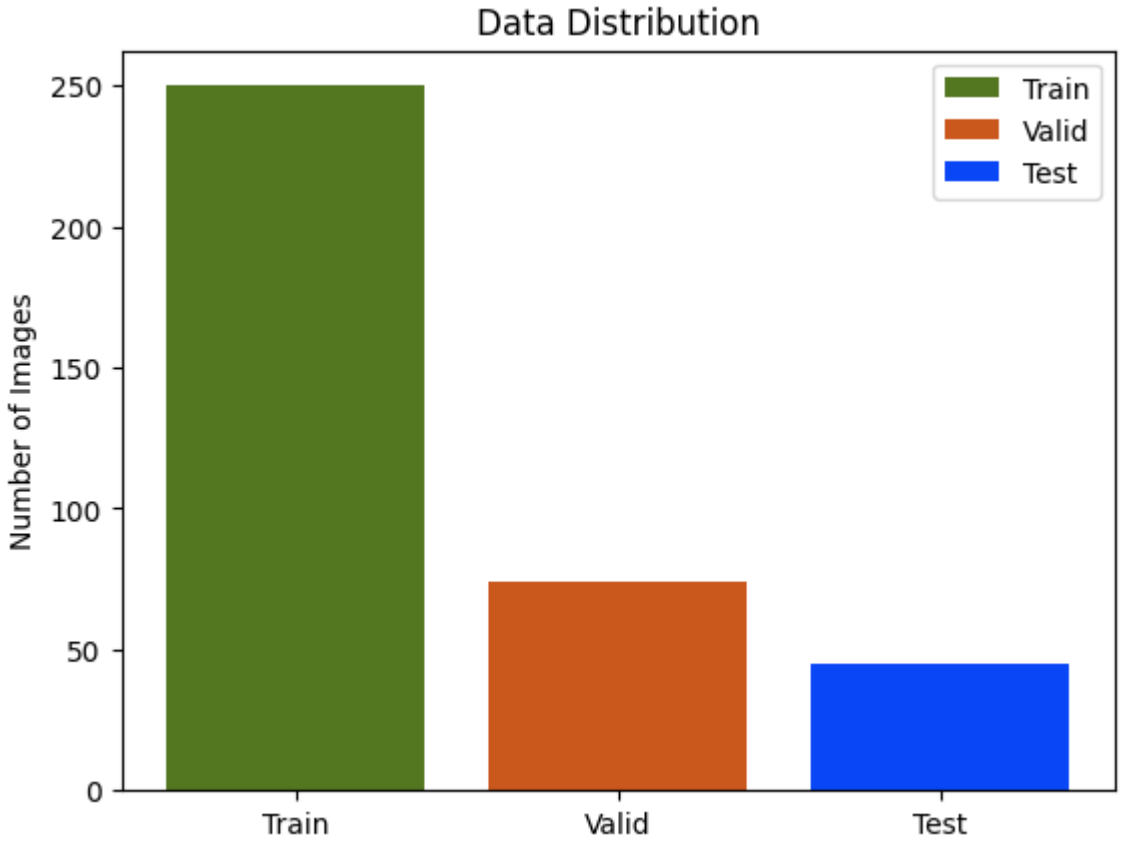
As it can be seen from Figure 8, some of the slices contain no information (black field), thus these were deprecated and a range of (50:-50) was used (Al Nasim *et al.*, 2022).

Figure 8 - All Slices of a Patient.



The data was then further split into Training – 250, Validation – 74, and Test Batches – 45 samples (Figure 9).

Figure 9 - Data Distribution.



As mentioned earlier only the segments from two modalities were included, where data-containing ones were segmented from 60 to 135 for each patient. One Hot Encoding was also applied to process the classes into empirical values. The pictures were further refined to 128x128 to reduce training time and resource cost. Every specimen comprises of 150 slices (100 partitions each deriving from the two modalities) (Al Nasim *et al.*, 2022).

3.3. Evaluation Metrics and Loss Function Employed

The correctness of the **loss function** being utilised during the network's training is pivotal as it assesses its deliverables. It forces the model to readjust its weights and decrease the loss,

thus improving the preciseness of outcomes. Hence, most employed one for tasks with multiple classes is **categorical cross-entropy**. The **Dice Loss** will also be applied, because it is standard for segmentation tasks, due to its ability to check for soundness in the overlaying of ground-truth and predicted output. (Brownlee, 2019; Al Nasim *et al.*, 2022).

Hence, a **Dice Coefficient** is presented as one of the evaluation metrics for every single class. The other evaluation metrics that are used are (Brownlee, 2019; Al Nasim *et al.*, 2022):

- **IoU** – Assesses the alignment among the predicted and ground truth values.
- **Accuracy**
- **Sensitivity** (Recall or True Positive Rate) - Evaluates how many positive ground truth pixels are precisely identified as positive.
- **Precision** (Positive Predictive Value) – Checks if predicted positive values are indeed positive.
- **Specificity** (True Negative Rate) - Checks if the negative ground truth values are precisely predicted as negative.

3.4. Model Architecture

Figure 3 provides a visual overview of the architecture used. **No improvements** were made to the original form of the model. **The reproduction was re-created** as it was proposed as a **2D U-net**, comprising of two sections consisting of multi-layer constructs – the extending one and the shrinking (Al Nasim *et al.*, 2022).

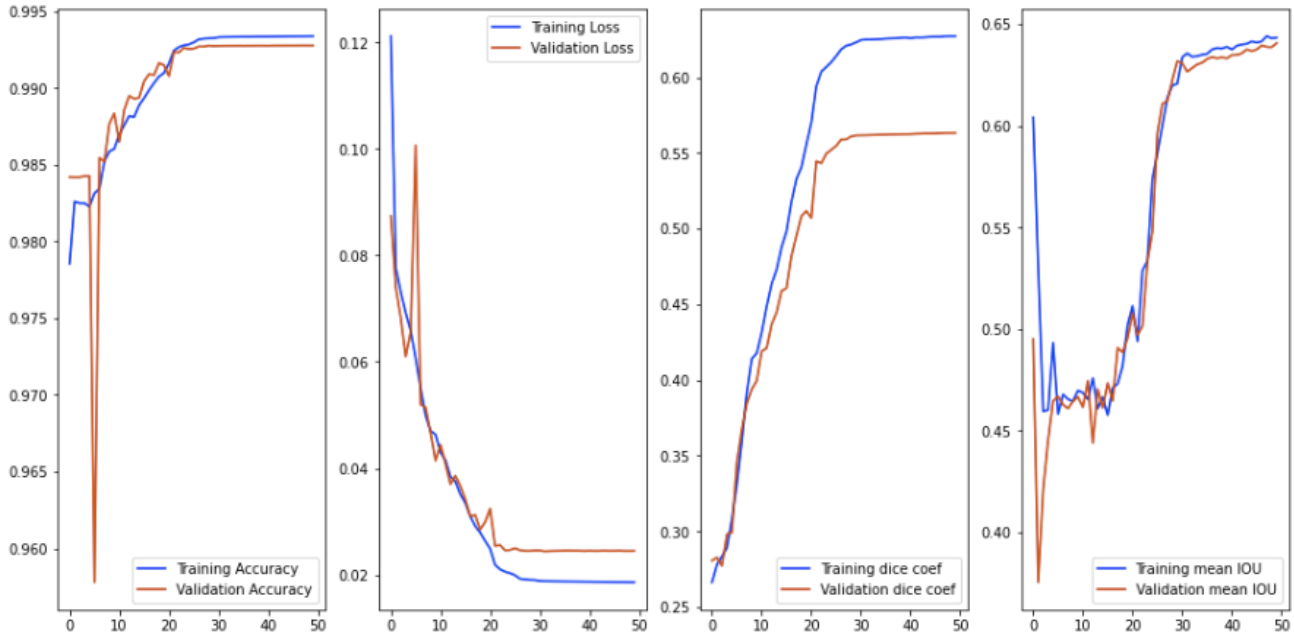
The shrinking encapsulates (Al Nasim *et al.*, 2022):

- Dual 3x3 layer construct – ReLU-activated.
- Maxpool – 2x2 layer (twofold increase in attributes on every iteration)

On the other side, the extending one involves (Al Nasim *et al.*, 2022):

- 2x2 transposed convolutional layers - to upscale the activation map and join it with the respective map of the previous portion of the network (Every upscaling stride is followed by a 50% reduction in the count of attributes).
- Two ReLU-activated convolutional layers of 3x3.
- 1x1 Output Layer – SoftMax-activated.

Figure 10 - Training Graphs of the Reconstructed Model until epoch 50.



3.5. Model Training

The architecture is drafted via Tensorflow and Keras and it was trained over 100 epochs due to time and hardware constraints and limitations. For each epoch, the model processes 250 samples from the training dataset. The batch size is 1 with a learning rate of 0.001. The techniques of Early Stopping, CSV logging and Model Checkpoint were administered to control overfitting (Al Nasim *et al.*, 2022).

4. Results and Discussion

Figure 11 - Results from the Reconstructed Model.

Evaluation of Reconstructed Model	
Dataset	BraTS 2020
Batch Size (Training)	1
Epochs Trained	100
Loss	0.0259
Accuracy	0.993
MeanIoU	0.5
Dice Coefficient	0.6098
Precission	0.9936
Sensitivity	0.9916
Specificity	0.9978
Dice coef Necrotic	0.6099
Dice coef Edema	0.686
Dice coef Enhancing	0.6524

Figure 11 visualises the results delivered through the reconstruction of the original model in 100 epochs. It is clearly visible that the model proved to be robust and delivered some high results. The Accurracy achieved during the evaluation is 99%, the loss 0.026. This rectifies the original findings that the model is reliable for its task and can assist doctors in their work. The Mean IoU is 0.5 with a Dice coefficient of 0.61. Precision, Specificity and Sensitivity are also at 99%. The Dice coefficients at the different cancerous types are circulating around the 60% mark.

Figure 12 - Comparison of the Recreated model with the original results reported.

Model	Dataset	Batch Size (Training)	Epochs Trained	Loss	Accuracy	MeanIoU	Precission	Sensitivity	Specificity	Dice Coefficient	Dice coef Necrotic	Dice coef Edema	Dice coef Enhancing
Original U-net	BraTS 2020	1	243	0.0056	0.998	0.8935	0.9973	0.997	0.9983	0.83	0.8846	0.9478	0.8544
Reconstruction	BraTS 2020	1	100	0.0259	0.993	0.5	0.6098	0.9936	0.9916	0.9978	0.6099	0.686	0.6524

From **Figure 12**, it is clearly visible that the original reported results are a bit higher, but that is due to the longer training period. Nonetheless, the recreated architecture does not fall far behind even with less than half of the training epochs. The only major difference comes in the dice coefficients, where the gap is clear.

Some of the challenges faced during the reconstruction and re-training of the model were:

- **Time and hardware limitations** – used less training epochs.
- **Deprecated libraries** from Keras and Tensorflow – resolved by using new functionality classes.
- **Training.log failed to generate labels** for evaluation metrics – could not be fixed.

Figure 13 - Predictions Made with the Reconstructed Model.

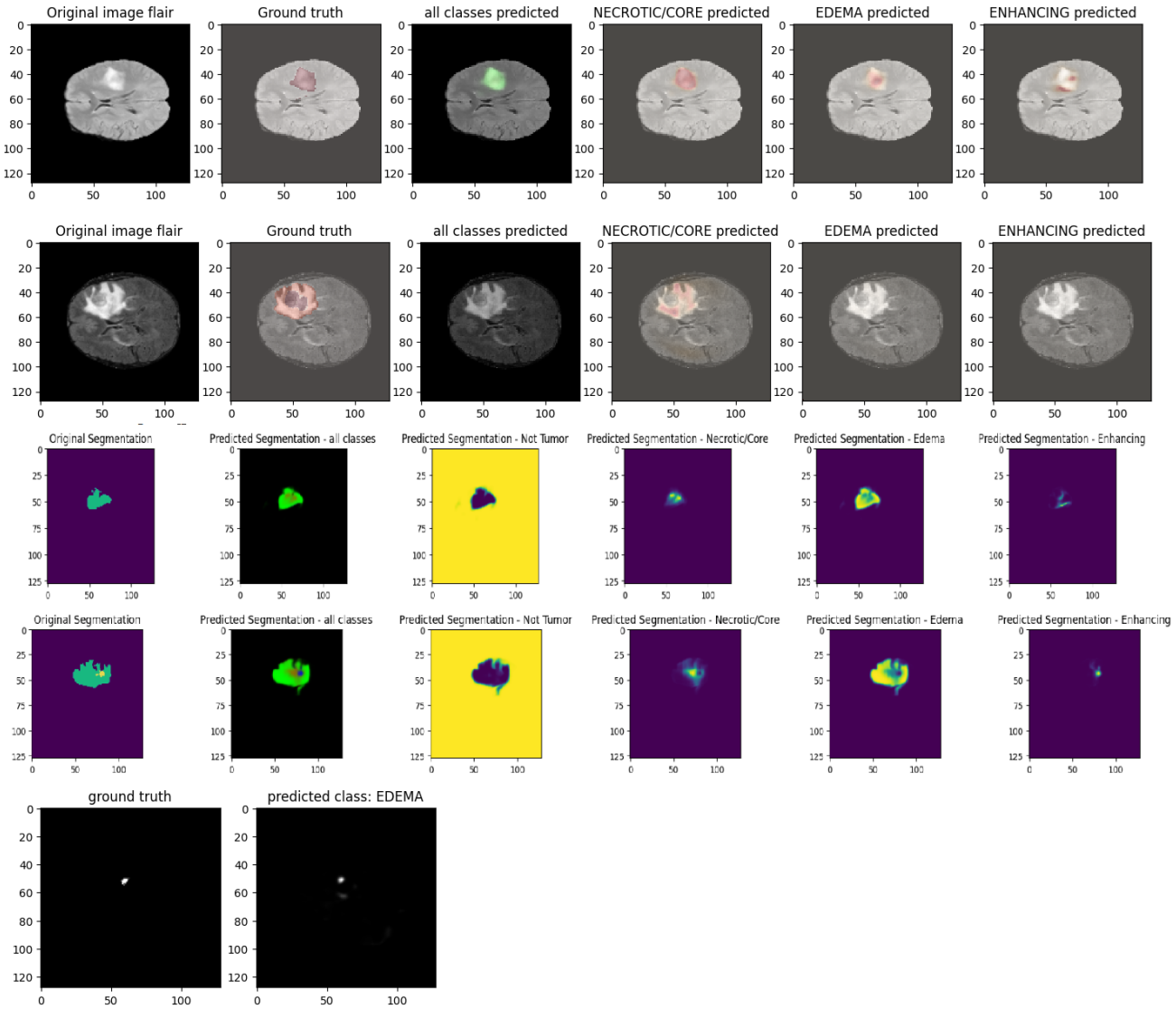


Figure 13 portrays some of the predictions made with the reconstructed model, which are still accurate, although the evaluation results were slightly lower than the original model.

This work has managed to confirm and recreate the findings by the authors of the original paper and proved that the model is robust and can be implemented for the use of automated MRI image segmentation of BTs and to assist physicians in their work.

5. Conclusion

In conclusion, the integration of computer vision and deep learning methodologies offers a substantial leap forward in detecting and diagnosing brain tumours in their early stages. By leveraging CNNs and volumetric segmentation architectures like U-Net, clinicians and researchers can more accurately and time-efficiently delineate carcinoma boundaries while simultaneously automating labour-intensive processes. Moreover, the fusion of multiple MRI sequences uplifts diagnostic precision by capturing diverse information sources with a single tool. The Multimodal approach discussed in this work also provides some astonishing results by improving decision-making by combining two sources of data, but it needs to be further researched to be deemed as a robust tool. This work has managed to recreate the discussed U-Net model and prove its findings even in half of the training time that was used originally. This allowed for the model to be deemed replicable and robust in its predictions. Such approaches result in better decision-making and empowers professionals to focus on personalised treatment planning and monitoring, rather than employing both theirs and the patient's time in time-consuming tasks. The techniques also uplift the physicians' work-process by accurately defining the borders of the malformations. This research also proves that DL techniques can be implemented in routine practice and stand to deliver faster and more targeted interventions.

6. Bibliography

Bauer, S. et al. (2013) A survey of MRI-based medical image analysis for brain tumor studies. *Physics in Medicine and Biology*, 58(13). [Accessed: 19 December 2024].

Bohaju, J. (2020) *Brain Tumor*. Available at: https://www.kaggle.com/datasets/jakeshbohaju/brain-tumor [Accessed: 20 December 2024].

Brownlee, J. (2019) *Better deep learning*. v1.81. [Online]. Available at: https://machinelearningmastery.com/better-deep-learning/ [Accessed: 23 November 2023].

Curci, A. & Esposito, A. (2024) Detecting Brain Tumors through Multimodal Neural Networks. *International Conference on Pattern Recognition Applications and Methods*, 1, 995–1000. Science and Technology Publications, Lda. [Accessed: 19 December 2024].

Dong, H. et al. (2017) Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. *Communications in Computer and Information Science*, 723, 506–517. Springer Verlag. [Accessed: 19 December 2024].

Habib, A.B. et al. (2020) Performance Analysis of Different 2D and 3D CNN Model for Liver Semantic Segmentation: A Review. *Lecture Notes in Electrical Engineering*, 633 LNEE, 166–174. Springer, Singapore. [Accessed: 20 December 2024].

Hambarde, P. et al. (2020) Prostate lesion segmentation in MR images using radiomics based deeply supervised U-Net. *Biocybernetics and Biomedical Engineering*, 40(4), 1421–1435. Elsevier. [Accessed: 20 December 2024].

Huang, G. et al. (2016) Densely Connected Convolutional Networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-January, 2261–2269. Institute of Electrical and Electronics Engineers Inc. [Accessed: 20 December 2024].

Kazerooni, A.F. et al. (2020) *BraTS2020 Dataset (Training + Validation)*. Available at: https://www.kaggle.com/datasets/awsaf49/brats20-dataset-training-validation/data [Accessed: 21 December 2024].

Kazerooni, A.F. et al. (2023) *The Brain Tumor Segmentation (BraTS) Challenge 2023: Focus on Pediatrics (CBTN-CONNECT-DIPGR-ASNR-MICCAI BraTS-PEDs)*. [Accessed: 19 December 2024].

Latif, G. et al. (2018) Automatic multimodal brain image classification using MLP and 3D glioma tumor reconstruction. *2017 9th IEEE-GCC Conference and Exhibition, GCCCE 2017*. Institute of Electrical and Electronics Engineers Inc. [Accessed: 19 December 2024].

Li, M. et al. (2019) Brain Tumor Detection Based on Multimodal Information Fusion and Convolutional Neural Network. *IEEE Access*, 7, 180134–180146. Institute of Electrical and Electronics Engineers Inc. [Accessed: 20 December 2024].

Litjens, G. et al. (2017) A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis*, 42, 60–88. Elsevier B.V. [Accessed: 19 December 2024].

Ma, X. & Jia, F. (2020) Brain tumor classification with multimodal MR and pathology images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11993 LNCS, 343–352. Springer. [Accessed: 19 December 2024].

Al Nasim, M.A. et al. (2022) Brain Tumor Segmentation using Enhanced U-Net Model with Empirical Analysis. *Proceedings of 2022 25th International Conference on Computer and Information Technology, ICCIT 2022*, 1027–1032. Institute of Electrical and Electronics Engineers Inc. [Accessed: 19 December 2024].

NHS (2023) *Brain tumours*. Available at: https://www.nhs.uk/conditions/brain-tumours/ [Accessed: 19 December 2024].

Pereira, S. et al. (2016) Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Transactions on Medical Imaging*, 35(5), 1240–1251. Institute of Electrical and Electronics Engineers Inc. [Accessed: 19 December 2024].

Soomro, T.A. et al. (2023) Image Segmentation for MR Brain Tumor Detection Using Machine Learning: A Review. *IEEE Reviews in Biomedical Engineering*, 16, 70–90. Institute of Electrical and Electronics Engineers Inc. [Accessed: 19 December 2024].

Ullah, M.S. et al. (2024) Multimodal brain tumor segmentation and classification from MRI scans based on optimized DeepLabV3+ and interpreted networks information fusion empowered with explainable AI. *Computers in Biology and Medicine*, 182, 109183. Pergamon. [Accessed: 20 December 2024].

7. Word Count

2904 words

8. GAI Declaration

I declare that no GAI was used for the development of this work.

The software used is:

- Microsoft Word and its grammar/spelling corrector
- Microsoft Excel – for the development of comparison tables and graphs
- Mendeley Reference Manager
- Google, Google Scholar and Discover@Bolton to narrow down and uplift research

9. Appendices

9.1. List of Abbreviations Used

1. **AI** - Artificial Intelligence
2. **AUC** – Area Under ROC-Curve
3. **BT** – Brain Tumour
4. **CNN** – Convolutional Neural Network
5. **CT** – Computed Tomography
6. **DL** – Deep Learning
7. **IoU** – Intersection over Union
8. **FCNN** – Fully Connected Neural Network
9. **FLAIR** – Fluid-Attenuated Inversion Recovery
10. **GAI** – Generative Artificial Intelligence
11. **ML** – Machine Learning
12. **MNN** – Multimodal Neural Network
13. **MRI** - Magnetic Resonance Imaging
14. **NN** – Neural Network
15. **PET** - Positron Emission Tomography
16. **ReLU** – Rectified Linear Unit

9.2. Links to Access the Dataset and Code Created to Recreate the Model.

Dataset – https://www.kaggle.com/datasets/awsaf49/brats20-dataset-training-validation/data

Google Colab Code Notebook – https://colab.research.google.com/drive/1U-w_uNyWdtCXOlw62hAwQ2kijIGrNsr-?usp=sharing