

Improvement of Deep Learning-Based Model for Automatic Segmentation of Brain Tumor

Bachelor of Technology in Computer Science and Engineering- Artificial Intelligence & Machine Learning, VIT Bhopal University SEHORE, MADHYA PRADESH – 466114

ARTICLE INFORMATION

KEY WORDS -

Brain Tumor,
Convolution Neural
Networks (CNNs),
Deep Learning,
Image Processing,
Automated Diagnostic,
Magnetic Resonance
Imaging (MRI),
Tumor Classification,
Diagnostic Accuracy,
Precision Medicine,
Medical Imaging,
Healthcare,

ABSTRACT

“Identification of Brain Tumor using Deep Learning” project addresses a very critical problems of our society. Cancer, which has become a significant factor in fatalities all over the world, and early detection can remarkably enhance the treatment outcomes and survival rates. Traditional diagnostic methods, which are likely to experience irregularities consists of human interpretation of medical images, which are often time-consuming. The addition of deep learning techniques, particularly in medical imaging, will offer a very promising alternative in improvement of accuracy and efficiency in detection of cancer.

The integration of deep learning into cancer detection represents a progressive step in medical diagnostics field, offering meaningful and important improvements in the accuracy and efficiency.

1. INTRODUCTION

Brain tumors represent one of the most life-threatening medical conditions, requiring timely and precise diagnosis. Conventional diagnostic methods rely heavily on manual examination by radiologists, which can be time-consuming and prone to irregularity. The advance of Artificial Intelligence (AI) has opened new avenues for automating and enhancing tumor identification.

Deep learning models, especially Convolutional Neural Networks (CNNs), have gained significant traction in the field of medical imaging due to their remarkable ability to extract intricate patterns from raw data. This capability is particularly advantageous in the diagnosis of brain tumors using MRI scans. However, several challenges impede the widespread application of these models in clinical settings. Key issues include data heterogeneity, the risk of overfitting, and high computational costs, all of which can limit the effectiveness of CNNs in real-world scenarios. The current research endeavours to address these challenges by developing an enhanced CNN-based framework specifically designed for the automatic identification of brain tumors from MRI scans. The primary objective is to improve both the accuracy and generalizability of the model through various techniques, such as data augmentation, transfer learning, and ensemble learning. Data augmentation involves generating additional training samples by applying transformations to existing images, thereby diversifying the dataset and reducing the likelihood of overfitting. Transfer learning allows the model to leverage knowledge acquired from pre-trained networks, enhancing performance, particularly when dealing with smaller, domain-specific datasets. Ensemble learning, which combines predictions from multiple models, further contributes to improved accuracy.

The medical imaging data presents a comparative visualization of brain tumor detection through two adjacent images. The left image displays a binary mask where the tumor location is highlighted as a distinct white region against a black background, effectively isolating the tumor's position in the posterior region of the brain. The right image shows the corresponding anatomical MRI scan of the same axial brain slice, enhanced with color mapping to differentiate various tissue structures and anatomical features. This multi-modal presentation technique is frequently employed in medical research to demonstrate tumor segmentation and precise anatomical location, allowing for clear visualization of the pathological finding in relation to surrounding brain structures. The juxtaposition of these two imaging modalities enables researchers and clinicians to accurately identify and analyze the tumor's position within the complex brain anatomy.

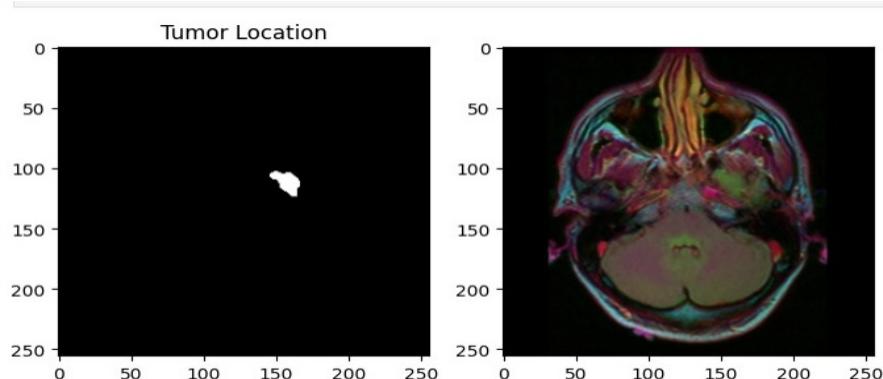


Fig.1. This is an image of a brain and it's segmented Tumor mask

([Fig.1](#)) consists of the image of brain and it's segmented tumor mask from MRI scans. The process of diagnosing brain tumors using CNNs relies on a hierarchical approach to feature processing. In this

framework, the initial layers of the network capture simple structures, such as edges and textures, while the deeper layers identify more complex patterns, including the irregular shapes of tumors. This hierarchical structure makes CNNs particularly well-suited for MRI analysis, where subtle anomalies in structural features can indicate the presence of a tumor. Techniques such as ReLU (Rectified Linear Unit) activation functions, dropout regularization, and optimizers like Adam are employed to enhance the efficiency of processing high-dimensional medical data. Despite these advancements, a significant limitation of CNNs is their difficulty in generalizing to unseen datasets. The performance of the model heavily relies on the quality of the annotated datasets used for training, and variations in data distribution can lead to inconsistent results. Factors such as noise, differences in imaging equipment, and variations in patient demographics can further complicate diagnostic outcomes. To mitigate these issues, it is essential to incorporate advanced data augmentation techniques and utilize larger, more diverse datasets.

The motivation for this research arises from the complexities and challenges associated with diagnosing brain tumors. Manual diagnosis using MRI scans is often time-consuming, it can be liable to human errors, and requires specialized expertise. Existing automated systems frequently struggle with noisy or low-quality data and may exhibit high false-positive rates. By developing a robust CNN-based framework, this research aims to enhance the accuracy and sensitivity of brain tumor detection while ensuring generalization across various datasets.

The primary aim of this work is to explore the application of deep learning, particularly CNNs, for the accurate and early detection of brain tumors in medical imaging. This involves training CNN-based architectures on large datasets of annotated medical images, including MRI scans, radiographs, and histopathological slides. By harnessing the power of deep learning, this research aspires to create an efficient, scalable, and accessible tool for radiologists and clinicians, ultimately improving patient outcomes through timely and accurate diagnoses.

In conclusion, the integration of deep learning techniques in medical imaging, particularly for brain tumor detection, holds significant promise. By addressing the challenges associated with current diagnostic methods and leveraging advanced CNN architectures, this research aims to contribute to the development of more accurate and efficient diagnostic technologies in healthcare, thereby enhancing early intervention and improving patient outcomes in this critical area of medicine.

2. ARCHITECTURE

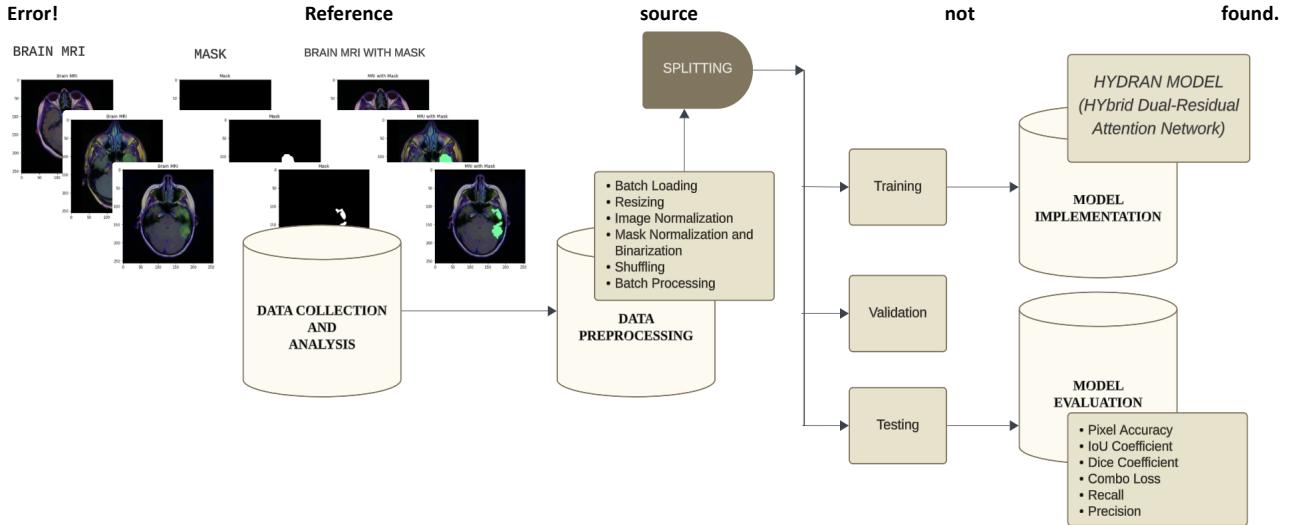


Fig. 2. Architecture model showing an overview of systematic workflow in MRI scans using Deep Learning model.

The (Fig.2) provides an overview of a systematic workflow for brain tumor segmentation in MRI scans which uses a specialized deep learning model. It starts with the Data Collection and Analysis phase, where brain MRI images and their corresponding segmentation masks are collected. These masks highlight the regions of interest, such as tumor areas, and are superimposed on the MRI scans and it enhances the understanding of the data. This stage allows that the raw data is comprehensively reviewed and prepared for subsequent processing.

Following data collection, the pipeline transitions into Data Preprocessing. Here, the essential transformations are applied to the images and masks to optimize them for training the model. These steps include loading the data in batches, resizing images to a uniform dimension, normalizing the pixel intensities for consistency, binarizing the masks to distinguish between tumor and non-tumor regions, shuffling the data for better training generalization, and recycling it into manageable subsets.

Once preprocessing is complete, the data is split into training, validation, and testing sets. The training set is used to teach the model to recognize patterns in the data, the validation set helps monitor and fine-tune the model during the learning process, and the testing set evaluates the model's performance on unseen data.

Currently, the base model is called informally as HYDRAN (Hybrid Dual-Residual Attention Network). This architecture of the model incorporated residual connections and an attention mechanism to simplify medical imaging and segmentation, which were directly linked to the development of deep learning techniques. The process is enhanced to emphasize specific details in the image, such as tumor borders, without affecting computational efficiency. However, these features are added for convenience.

After the model is trained and implemented, it undergoes a very careful Model Evaluation. This stage involves assessing its performance using several metrics, such as pixel accuracy, Intersection over Union (IoU) coefficient, Dice coefficient, and Combo loss, along with precision and recall. These criteria provide a very broad understanding of how well the model segments the tumor regions from MRI scans, ensuring it meets the desired level of accuracy and reliability for clinical or research applications.

This paper introduces an advanced CNN-based model designed for the automatic detection of brain tumors in MRI scans. Making use of the hierarchical feature extraction and data augmentation techniques, the model achieves the remarkable accuracy, sensitivity, and specificity, excelling earlier architectures like VGG16 and

ResNet-50 while effectively justifying issues such as overfitting and limited dataset diversity. The findings bring out the transformative potential of AI in enhancing diagnostic accuracy and efficiency within clinical practice. Additionally, the framework displays versatility across various medical imaging datasets, which ultimately contributes to improved healthcare technologies and better patient care while addressing challenges related to deployment in AI-based diagnostic tools.

3. LITERATURE SURVEY

Literature review consists of sections that tell us about applications and benefits of using this system.

3.1 Existing Work

Traditional Styles for segmenting and classifying brain excrescences in MRI reviews frequently fall suddenly, being time-consuming and prone to mortal error. This underscores the need for a completely automated approach using deep learning. In response, this work presents a mongrel model that combines deep convolutional neural networks(CNNs) with machine literacy classifiers for more accurate excrescence segmentation and bracket. The process begins with a CNN that excerpts features from MRI images, followed by a briskly region-grounded CNN for excrescence localization, enhanced by a region offer network. Eventually, the model integrates deep CNNs with support vector machine classifiers to upgrade results further. estimated on established criteria , the proposed model achieved emotional precision rates of 98.3 and a dice similarity measure of 97.8 for classifying excrescences like gliomas, meningiomas, and pituitary excrescences. These results demonstrate significant advancements over being ways, pressing the eventuality of deep literacy in medical imaging. [1]

Classifying brain Excrescences is pivotal for diagnosing brain cancer, especially in IoT healthcare systems. While artificial intelligence (AI) ways, particularly computer- backed individual systems(CADS), are designed to enhance discovery delicacy, their current limitations have led to hesitance among medical professionals to completely embrace them. This study introduces a robust brain excrescence bracket system using deep literacy to attack these delicacy issues. The proposed approach employs an advanced convolutional neural network(CNN) that analyses brain MRI data for effective excrescence bracket. By incorporating data addition and transfer literacy, the model significantly enhances bracket performance compared to being birth models. The emotional results suggest that this innovative system could greatly ameliorate brain cancer opinion in IoT healthcare settings, offering a dependable tool for medical professionals in their fight against this complaint. [2]

Brain Excrescences, whether benign or nasty, arise from the abnormal growth of cells in the brain, frequently without a known cause. Late judgments can lead to the aggressive progression of cancer, complicating treatment options. To address this challenge, deep literacy ways, particularly convolutional neural networks(CNNs), have been employed for the accurate opinion and bracket of brain excrescence images. This paper explores colorful CNN infrastructures, including DenseNet- 201, commencement- V1, AlexNet, and MobileNetV2, to determine the most effective result for relating brain excrescences and mollifying late or incorrect judgments . Using a dataset of over 3,000 MRI images from Kaggle, the study set up that MobileNetV2 outperformed the others, achieving an emotional delicacy of 96.5, along with high perceptivity, perfection, particularity, and F1- score. These results punctuate the critical part of deep literacy

in aiding healthcare professionals with MRI image interpretation, eventually perfecting patient issues in brain excrescence opinion.[\[3\]](#)

3.2 Observations from Literature Survey

The part of multi-modality and multi different datasets in medical imaging is very crucial for creating robust models for tumor segmentation. Multiple current trials calculate on the datasets similar as Brain Dataset-1, Figshare, or Kaggle, which, despite containing thousands of images, frequently require the necessary diversity. This limited variability can support a model's own capability to generalize effectively to external data, especially when faced with different tumor types or varying imaging conditions or similar issues. It is observed that the objectification of multi-modality imaging — specifically FLAIR, T1-weighted, and T2-weighted reviews which significantly enhance the robustness of the model.

Each imaging modality contributes in the uniqueness perceptivity, into the anatomical and pathological features of the brain. For example, FLAIR reviews are particularly complete at imaging edema and lesions, while T1-weighted reviews give a detailed structural information. By combining these modalities, the model can achieve bettered segmentation accuracy and trustability. Similarly, using multiple datasets that allows a wider range of imaging conditions, patient demographics, and tumor types significantly supports the model's connection in real-world scripts. This diversity very much ensures us that the model is trained on an individual sample of varying and implicit cases that it may encounter in usual clinical practice. This increases its capability to generalize the data which is not seen. In conclusion, the planned use of multi-modality imaging combined with different datasets is very essential for developing models that are not only accurate but also practical and ready to use for real-world operations in medical imaging and tumor analysis.

4. DATASET COLLECTION AND PROCESSING

4.1 LGG Segmentation Dataset

This dataset comprises brain MRI images alongside manual FLAIR abnormality segmentation masks, sourced from The Cancer Imaging Archive (TCIA). It includes data from 110 patients featured in The Cancer Genome Atlas (TCGA) lower-grade glioma collection, all of whom have at least one fluid-attenuated inversion recovery (FLAIR) sequence and genomic cluster data available. The TCGA-LGG collection aims to connect cancer phenotypes with genotypes by providing clinical images linked to TCGA subjects. While clinical, genetic, and pathological data can be accessed through the Genomic Data Commons (GDC) Data Portal, the radiological data is housed in TCIA.

The dataset features both automated and manually refined segmentation labels from pre-operative scans of patients with Low Grade Glioma (LGG). These scans were meticulously selected and processed to remove skull artifacts and ensure accurate alignment. Tumor segmentation utilized an advanced hybrid method recognized for its performance in the BRATS 2015 challenge, followed by a thorough review from an expert neuroradiologist to correct any inaccuracies.

This final set of segmentation labels allows the extraction of various imaging features, including intensity, volume, shape, texture, and diffusion information. By providing both, the computer-generated and manually corrected labels, we allow the dataset to support the estimated research in computational and clinical environment, helping in the comparisons across various studies without the need for repeated manual explanation. Additionally, the radiomic features enable the combination with molecular characterization from TCGA, offering various insights into the relationship between imaging features and clinical outcomes, which is very beneficial for researchers who lack computational expertise.

4.2 Method and methodology

In our research, we look into a dataset that encompasses crucial clinical and genetic information related to brain tumors. This dataset is a crucial key in understanding of the diverse characteristics of tumors and how these traits may influence patient outcomes. The data is systematically organized in a structured format and stored in a CSV file, which makes it easy for us to access and handle during our analysis.

4.2.1 Dataset Description

In our study of the dataset, we took a closer look at the distribution of tumor masks, which indicate whether tumors are present (1) or absent (0) in the MRI images.

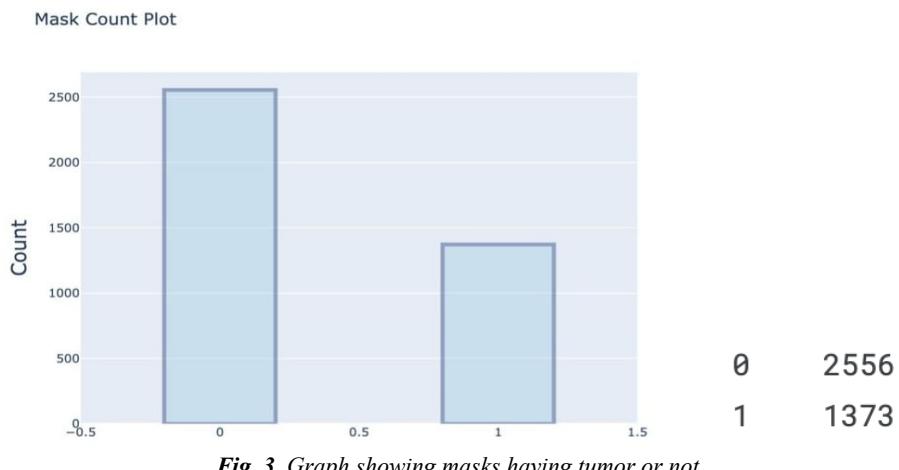


Fig. 3. Graph showing masks having tumor or not

From the ([Fig.3](#)), we can see that there are 2,556 instances where tumors are absent (indicated by a mask value of 0) and 1,373 instances where tumors are present (with a mask value of 1).

This distribution highlights a notable variation. There are significantly more images without tumors than with tumors.

Overall, the analysis of the tumor mask distribution gives us a good and valuable insights into the dataset's structure. Understanding the balance between tumor presence and absence is crucial as we move forward and understand more of our research. By keeping this distribution in mind, we can design better analytical strategies and models that accurately interpret and classify MRI images, ultimately enhancing our understanding of brain tumors and improving diagnostic approaches.

#	Column	Non-Null Count	Dtype
0	patient_id	3929 non-null	object
1	image_path	3929 non-null	object
2	mask_path	3929 non-null	object
3	mask	3929 non-null	int64

Fig. 4. Table showing image path, mask path, patient id and mask

([Fig.4](#)) shows that the dataset comprises a total of 3,929 entries, each comparing to a unique patient. We provided a detailed description of the key components of the dataset:

Patient ID column contains unique identification for each patient, which is represented as an object type. These identifications are essential for tracking and managing the data related to individual patients throughout our analysis.

The Image Path column lists the file paths to the MRI images of the patients. Each entry is stored as an object type, which ensures that the paths can contain various file formats and directory structures. This flexibility is important for combining different images from different sources. Similar to the image path, the Mask Path column provides the file paths to the corresponding segmentation masks. These masks play a crucial role in identifying tumor regions within the MRI images and are also stored as object types, enabling easy access for processing.

In conclusion, this dataset is a valuable resource that provides us a wealth of clinical and genetic information. It enables us to explore the relationships between tumor characteristics and patient outcomes in detailed view. The organized structure of the data allows for smooth and hassle-free processing and analysis, which is crucial for drawing meaningful insights from our research. By utilizing this dataset, we aim to deepen our understanding of brain tumors and contribute to improvement in diagnostic approaches that can ultimately improve the patient care.

This dataset serves as a foundation for our research, enabling us to develop machine learning models which are aimed at accurately segmenting tumor regions in brain MRI scans. The organized structure of the data not only enables streamlined processing and analysis but also increases the reliability of our results.

Ground Truth Labels are the masks typically created by medical professionals or through semi-automated processes. They serve as a reference point, or "ground truth," for training machine learning models. This brings insights into the field of medical imaging and improve diagnostic capabilities for brain tumors.

4.2.2 Dataset Visualisation

Enables these models to learn how to accurately predict and identify similar regions in raw imaging data. These represent the original brain MRI reviews. Each image shows a sampling of the brain at colourful situations, pressing anatomical structures. This visualization overlays the prognosticate or base verity segmentation mask onto the original MRI reviews. It effectively highlights tumor areas in a visually intuitive manner, enabling better understanding and confirmation of the segmentation quality.

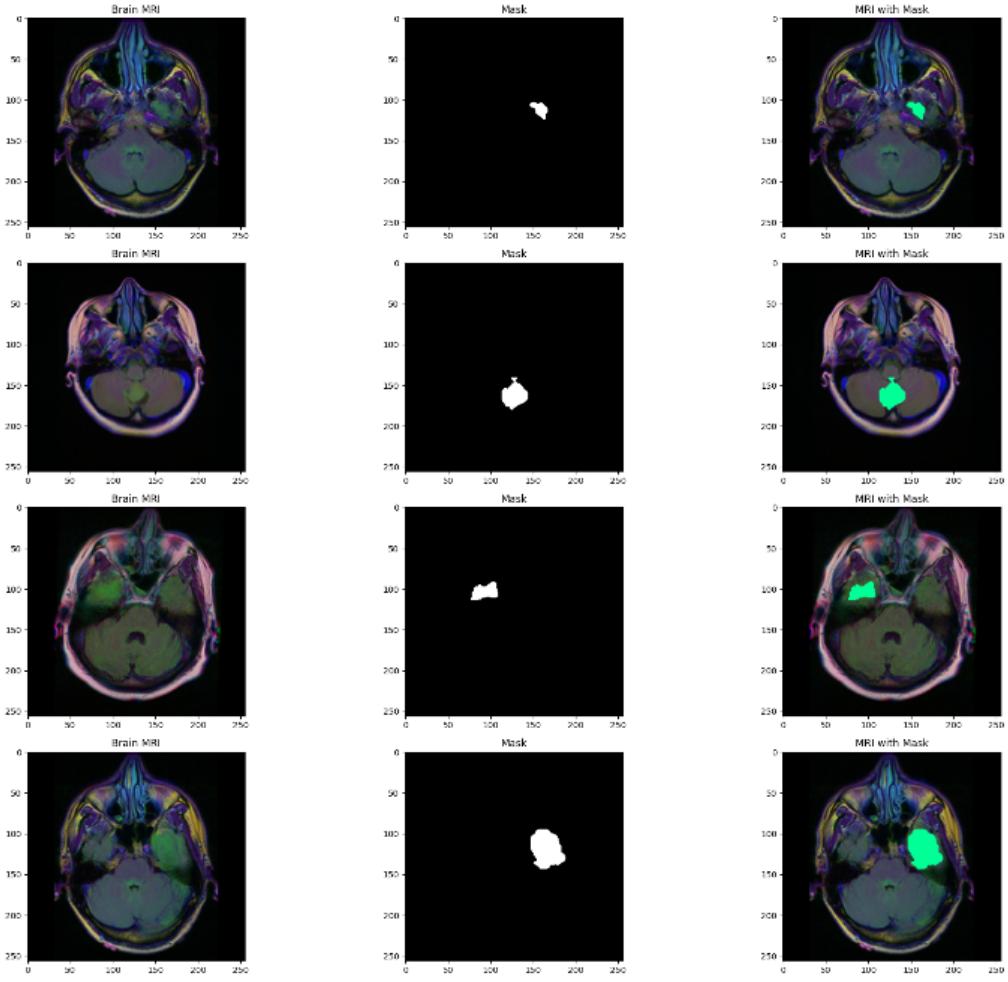


Fig. 5. Dataset Visualization for Brain Tumor Segmentation

(Fig.5) illustrates the dataset used for the study, comprising MRI reviews with corresponding segmentation masks. Column 1 presents original MRI images, displaying colorful axial slices of the brain. Column 2 shows the ground variety double masks pressing tumor regions explained by medical professionals. Column 3 overlays these masks on the MRI reviews for clearer visualization of the tumor's position and structure. This visualization emphasizes the diversity and complexity of the dataset, with variations in tumor size, shape, and position, highlighting the challenges in accurate tumor segmentation. These visualizations give an intuitive understanding of the segmentation process and the indirect disagreement between prognostications and ground variety.

The numbers handed show critical rudiments of the dataset used for brain tumor segmentation, which include MRI reviews, ground verity segmentation masks, and combined overlays. The first column of the numbers presents axial cross-sections of the brain from MRI reviews in fig.6. These slices parade the anatomical diversity in terms of brain shape, size, and internal structures. The differences in brilliance and discrepancy across the reviews reflect variations in accession parameters, scanner types, or imaging sequences, which add complexity to the segmentation task.

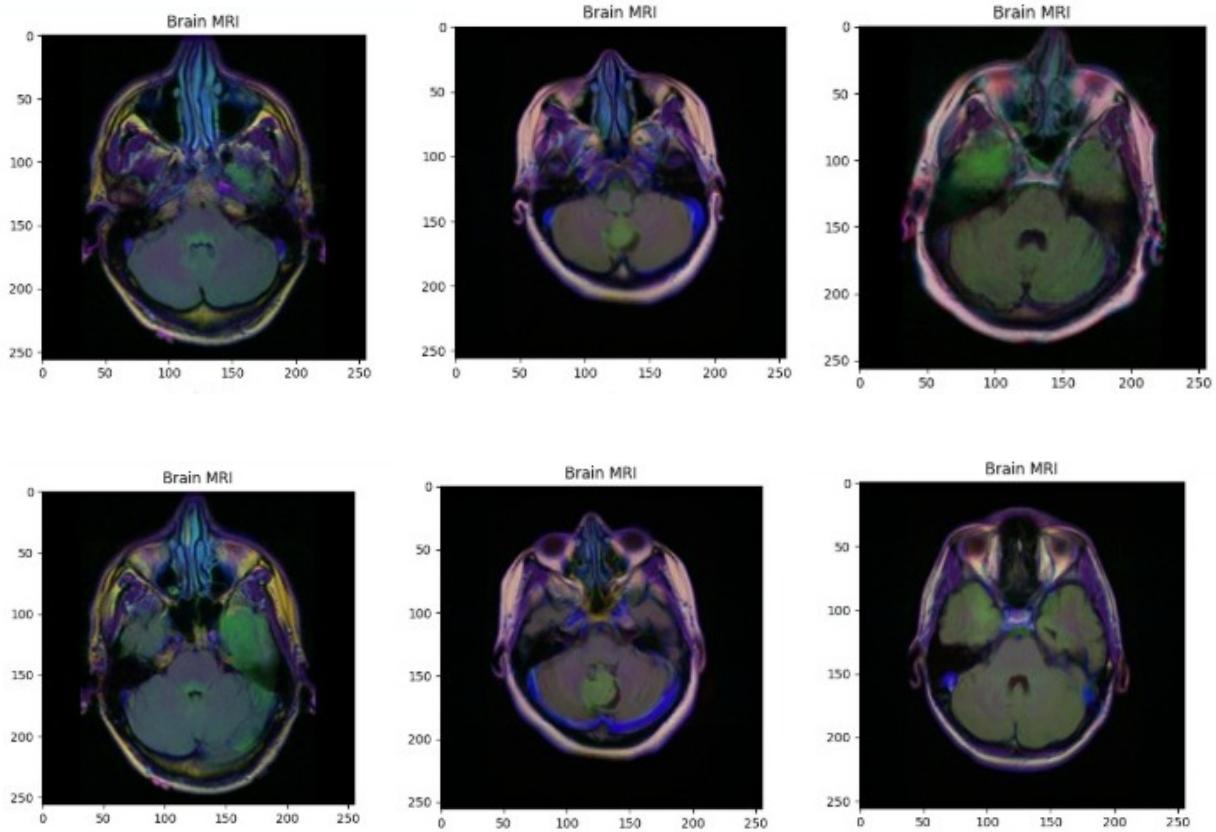
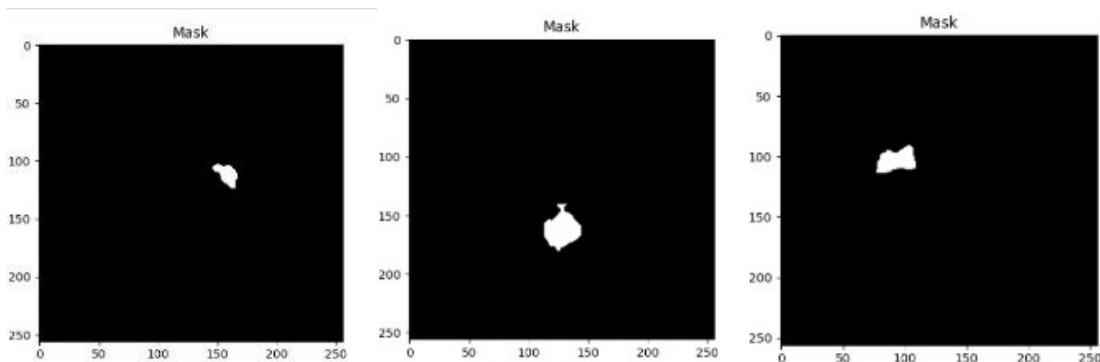


Fig. 6. All brain MRI images from the dataset

([Fig.7](#)) double masks highlights the ground verity tumor reflections handed by experts. These masks demonstrate the diversity in tumor morphology Shape Tumors appear irregular, with some being small and compact, while others are larger with verbose boundaries.

The masks show a range of tumor sizes, from many pixels to large millions enwrapping significant portions of the brain. Tumors are located in different brain regions, illustrating the challenge of segmenting structures that may vary significantly in appearance depending on their surroundings. These are double segmentation masks that indicate the regions of interest(tumor) manually annotated by experts. White regions in the masks shows that the tumor is present, while black regions represents the non-tumor region.



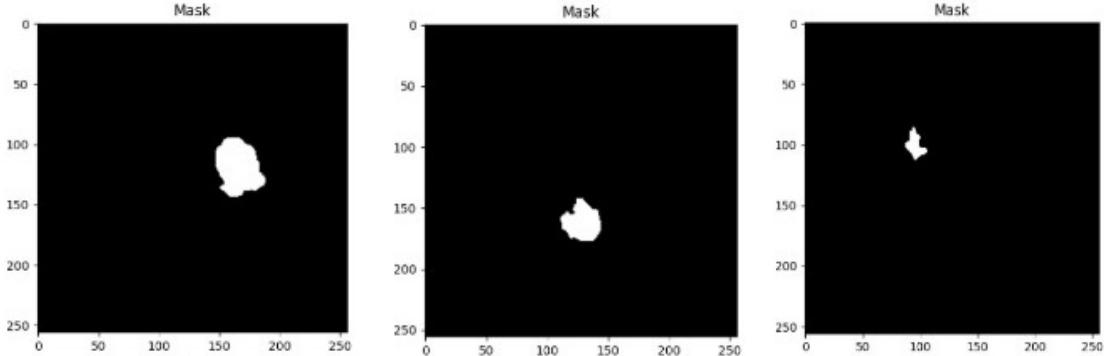


Fig. 7. Brain MRI mask (Ground Truth) from the dataset

([Fig.8](#)) provides a superimposed view of the tumor masks on their corresponding MRI reviews. This visualization emphasizes the alignment of the tumor region within the anatomical environment of the brain. By overlaying the masks, it becomes apparent how the tumor interacts with girding capkins. For example, some tumors are located near critical structures, like the ventricles or cortex, taking high segmentation delicacies to save a healthy towel during treatment.

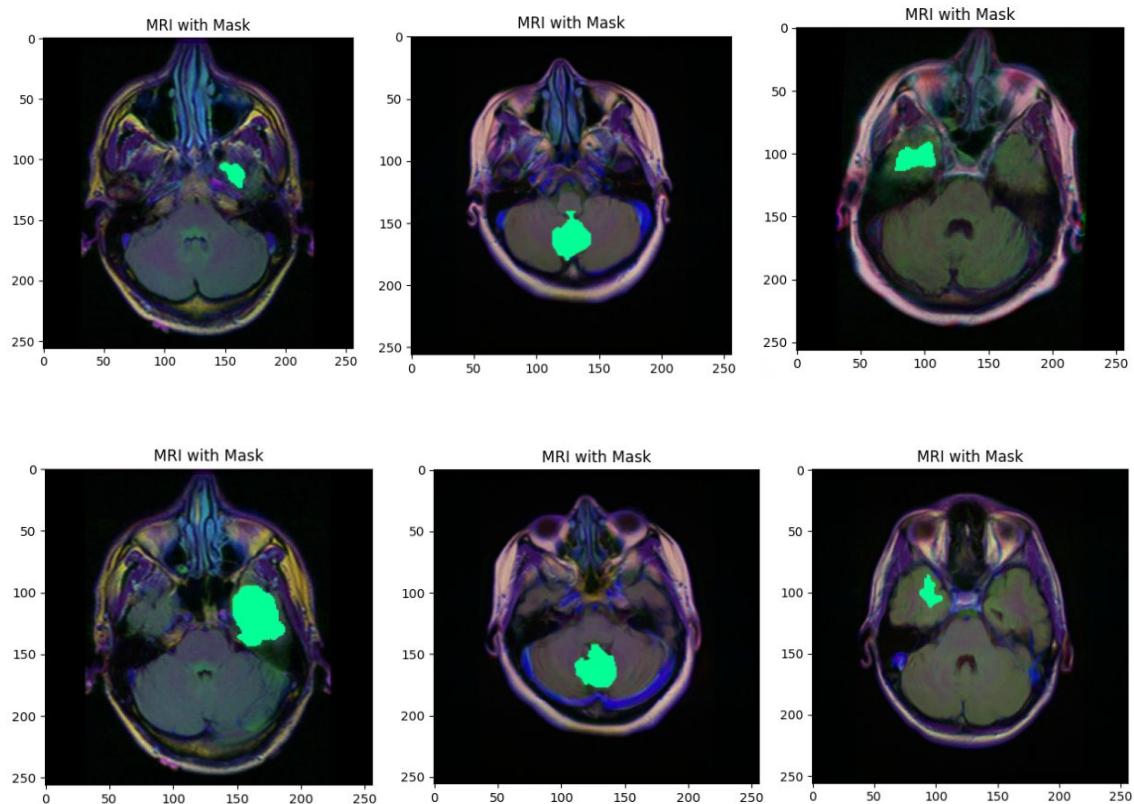


Fig. 8. Brain MRI and its mask overlay

Healthy Slices ([Fig.9](#)) Several slices show no tumor presence(i.e., empty masks), indicating the addition of healthy regions in the dataset. These cases are essential for reducing false cons during model training and icing a balanced representation of normal and abnormal data. Differ and Vestiges Some reviews show low discrepancy between the tumor and conterminous capkins, posing a significant challenge for segmentation algorithms. Likewise, imaging vestiges, similar to noise or intensity variations, can make it harder to distinguish tumors from background capkins. Although presented as 2D slices, these images are part of

volumetric data. Excretions that gauge multiple slices bear algorithms capable of landing spatial durability across 3D space.

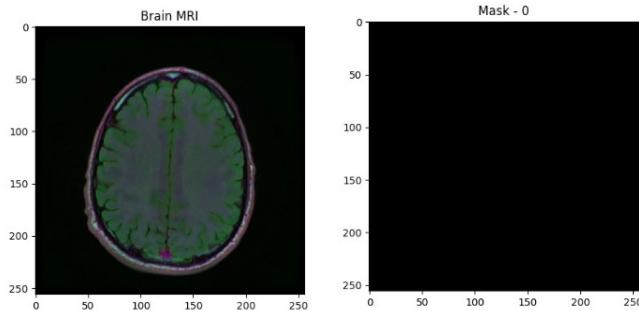


Fig. 9. A Healthy Brain MRI showing no tumor

The dataset's diversity reflects real-world clinical scenarios, which shows a wider range of brain tumor appearances. This variability is crucial for testing the robustness of segmentation models, ensuring they are applicable to various patient populations and imaging conditions. It highlights the importance of using diverse datasets and advanced segmentation techniques to improve the accuracy and generalizability of automated tumor detection systems.

To gain valuable insights from the imaging data, researchers can purchase libraries like OpenCV or PyTorch to extract binary arrays from segmentation masks, enabling calculated analysis of specific regions. Combining raw scans with their corresponding masks allows them to get effective visualization of model performance, which can be enhanced using tools like Matplotlib. By calculating the total volume of identified tumors through pixel analysis, researchers can gain deeper insights into tumor morphology and its relationship with surrounding anatomy. Extracting features such as pixel intensity and texture can further refine the input for machine learning models, ultimately boosting their performance.

4.2.3 Data Preprocessing

Data preprocessing is a fundamental step in any machine learning workflow, particularly in the realm of medical imaging tasks such as brain tumor segmentation. In this study, we developed a custom data generator that dynamically prepares the dataset during both training and validation phases. This approach not only optimizes resource utilization but also enhances the model's performance by ensuring high-quality input data. The preprocessing steps are as follows:

The preprocessing steps involved in preparing images and their corresponding segmentation masks are essential for optimizing model training, particularly within the field of medical imaging. We begin by loading the data in batches directly from disk, a strategy that effectively manages memory usage, especially when working with large datasets. This approach ensures a continuous flow of information to the model while conserving system resources, which is critical in maintaining efficiency throughout the training process.

To achieve uniformity across the dataset, all images and masks are resized to a fixed resolution of 256×256 pixels. This standardization is crucial, as it guarantees consistent input dimensions, thereby facilitating effective training and minimizing the risk of processing errors that could arise from variations in image size. Subsequent to resizing, we normalize the pixel intensities of the images through a process of mean subtraction and division by the standard deviation. This normalization enhances the model's ability to learn by addressing

variations in intensity ranges that may result from differing lighting conditions and inconsistencies in the imaging process. Additionally, the segmentation masks undergo normalization and binarization, which distinctly delineates various regions, enabling the model to accurately identify and segment different anatomical structures.

To further enhance the model's generalization capabilities, we shuffle the data at the conclusion of each training epoch. This practice prevents overfitting by ensuring that the model does not become dependent on the order of the data. Finally, processing images and masks in batches not only contributes to efficient training through parallel processing but also alleviates the overall computational burden on the system, thereby optimizing the training process. These preprocessing steps collectively ensure that the model is well-equipped to learn from the data, leading to improved performance in segmentation tasks.

4.2.4 Data Split

We divided our dataset into three key segments to optimize our model's performance. The largest portion, comprising 70% of the total, is designated as training data, which includes MRI images paired with their corresponding ground truth annotations. This allows us to effectively train and fine-tune the model to recognize patterns within the scans. We also set aside 15% of the dataset as validation data, which remains distinct from the training set. This subset is crucial for adjusting model parameters and monitoring its performance throughout the training process, ensuring effective learning. Finally, the remaining 15% is reserved as test data, an independent set of MRI images used solely to review the model's performance after training is complete. This approach helps us gauge the model's reliability and adaptability in real-world scenarios.

5. MODEL ARCHITECTURE

5.1 HYDRAN MODEL *(Hybrid Dual-Residual Attention Network)*

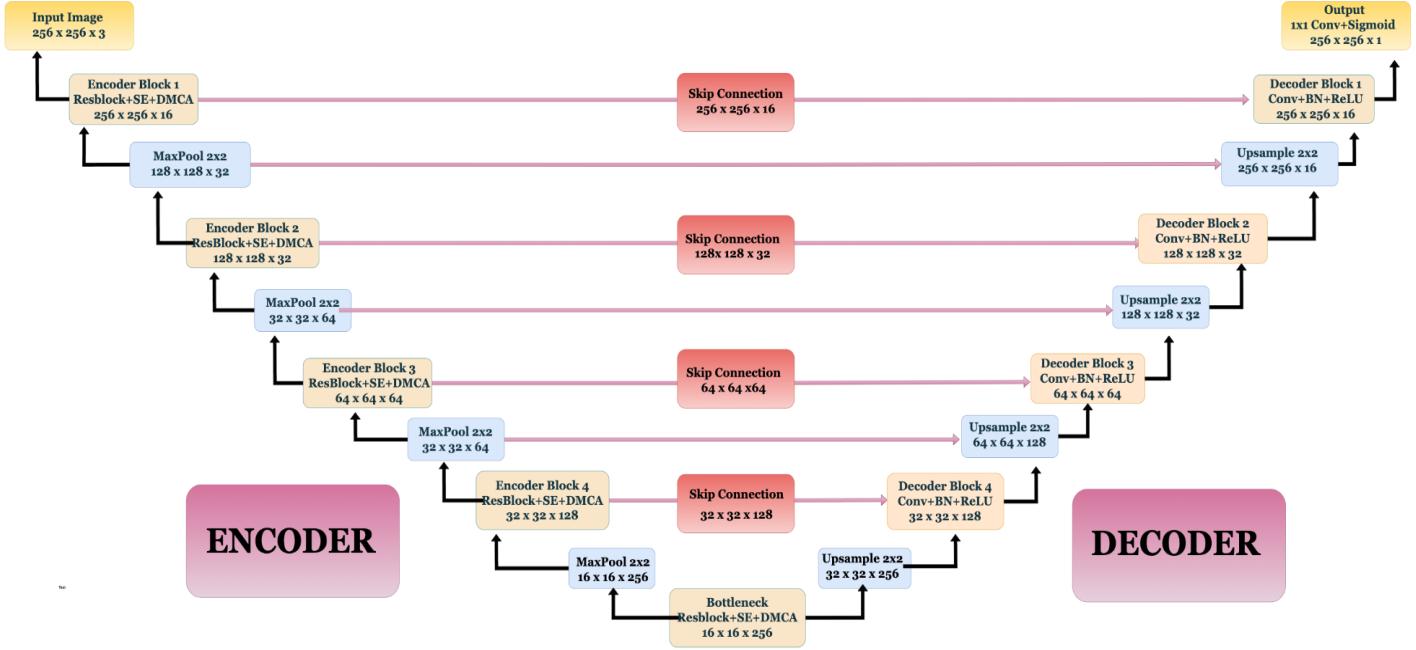


Fig. 10. Hydran Model Encoder-Decoder Architecture

5.1.1 Encoder Path

Each of the four successive blocks that make up the encoder implements a combination of ResBlock + SE + DMCA:

The encoder route in functions similarly to a squad of detectives trying to identify the key elements in our input image. In (Fig.10) Four encoder blocks make up this route, which analyses the picture while progressively shrinking its size.

Every block has, similar to filters, convolutional layers sort through an image to identify key elements like edges, textures, and patterns. Think of them as a magnifying glass that draws attention to the intricacies. Activation functions add some complexity after filtering, we use non-linear activation functions (such as ReLU). Because it enables our model to discover complex patterns and correlations within the data, this stage is essential.

Encoder Block 1 (256×256×16): The input layer is where the adventure starts since it acts as our model's first interface. The 256 by 256 pixel image that we are given here has three colour channels: red, green, and blue (RGB). We can deal with rich, colour photographs with this setup, which lays the groundwork for the following processing stages. The input layer is essential since it defines the type and size of the data that the model will analyse.

Encoder Block 2 (128×128×32): This block processes the intermediate features using the output from the first block. It keeps improving how the data is represented.

The quality of the learnt representations is improved when attention methods are incorporated since they enable the model to concentrate on the most pertinent characteristics.

To further simplify the representation while keeping important information, a second MaxPooling layer raises the channel depth to (64) and decreases the dimensions to (64 times 64)

Encoder Block 3 (64×64×64): The model enables the capture of more intricate patterns in this block by extracting deeper features from the data. To make sure the model concentrates on the most important elements of the input, the attention techniques are further employed to improve the feature representation. In order to summarise the characteristics while preserving important information, the spatial dimensions are decreased to (32 times 32) and the channel depth is increased to (128).

The last encoder step, **Encoder Block 4 (32×32×128)**, is when the model extracts the highest-level feature abstractions from the input data. At this point, the model has reduced the data to its most abstract form, enabling a thorough comprehension of the input. In order to prepare the data for the bottleneck stage, the last MaxPooling operation increases the channel depth to (256) while decreasing the dimensions to (16 times 16).

5.1.2 Bottleneck (16 x 16 x 256)

The bottleneck layer is important to our architecture in [fig.10](#). Consider it the link between the encoder and decoder routes. This layer compresses all of the key characteristics that the encoder was able to extract. These properties are further refined by a few convolutional layers, which are often included. Only the most essential insights are sent to the decoder thanks to the bottleneck layer, which keeps the most important information while eliminating repetition. It is the model's method of focussing on the most important aspects before proceeding to the reconstruction stage. An essential connection between the encoder and decoder components of the design is the bottleneck. It is essential for moving from feature extraction to feature utilization. In order to efficiently analyse the most abstract features, the architecture now employs the same mix of Residual Blocks, Squeeze-and-Excitation, and Dynamic Multi-Channel Attention at the greatest feature depth. The bottleneck is made to preserve spatial information, which is essential for the decoding procedure that follows, even if it is highly abstracted. To produce useful outputs from the learnt characteristics, this harmony between abstraction and spatial awareness is necessary.

5.1.3 Decoder Path

We now go on to the decoder route in [fig.10](#), which restores our compressed form. This route mimics the encoder and consists of four decoder units, each playing an important role:

Up-sampling layers use techniques such as transposed convolutions to enhance the size of feature maps. It's similar to enlarging a little drawing into a huge mural.

Skip Connections: These connections connect corresponding layers from the encoder and decoder routes, allowing the model to preserve important spatial information that may have been lost during down sampling. This ensures that tiny details are not ignored in the final product.

Convolutional Layers: Similar to the encoder, these layers enhance the up-sampled feature maps, ensuring that the output is as clear and detailed as feasible. The decoder blocks work together to recreate the output picture, preserving the essence of the original input while augmenting it with learnt characteristics. The decoder route is a critical component of our system, designed to reverse the encoder's activities and gradually recreate the picture from the compressed representation acquired from the bottleneck layer. It is made up of

four decoder blocks, each of which plays an important part in improving the output quality while preserving the basic aspects of the original picture.

Decoder Block 4 (32×32×128)

The initial reconstruction step begins with Decoder Block 4, which takes the bottleneck features and starts the up-sampling process to a higher resolution. This block focusses on increasing the representation while keeping the abstract properties learnt in the bottleneck, establishing the groundwork for future reconstruction phases.

Decoder Block 3 (64×64×64).

In Decoder Block 3, the model extends the reconstruction process by up sampling the characteristics from Decoder Block 4. A skip link from Encoder Block 3 is used, allowing the model to access high-resolution features collected before. This integration of encoder properties preserves spatial coherence and detail in the reconstructed output, resulting in a more coherent representation.

Decoder Block 2 (128×128×32)

Decoder Block 2 continues the up-sampling process by increasing spatial dimensions and refining features to increase quality. A skip connection from Encoder Block 2 adds more high-resolution characteristics, restoring information that may have been lost during encoding. The combination of up sampling and skip connection improves detail retention, yielding a more accurate reproduction of the original input.

Decoder Block 1 (256×256×16)

The last reconstruction stage is aided by Decoder Block 1, which is the final step in the decoding process. A skip connection from Encoder Block 1 is used to restore the original features, ensuring that fine details remain in the output. This block effectively reconstructs by mixing upsampled and encoded information in the original image at full resolution, maintaining critical details and quality throughout the process.

5.1.4 Skip Connections

In neural network designs in [fig.10](#), skip connections are a potent approach that is frequently used, especially in models like as U-Net. Their main objective is to improve gradients and information flow across the network, which greatly improves the model's performance. An official yet understandable summary of the main elements of skip connections is provided below:

Direct channels: In a neural network, skip connections create direct channels between corresponding encoder and decoder layers. For instance, a decoder layer's input at the same resolution is directly connected to the output of an encoder layer at that same resolution. By facilitating more efficient information flow, this link enables the decoder to take use of the rich characteristics that the encoder has extracted.

Maintenance of Fine-Grained Spatial Information: Skip connections are essential for maintaining fine-grained features that could otherwise be lost during the encoder's downsampling process since they connect the encoder and decoder at the same spatial resolutions. For applications like picture segmentation, where correct results depend on the exact localisation of features, this capacity is especially crucial.

Encouraging Gradient Flow: During the training backpropagation stage, skip connections improve the gradient flow. This enhancement aids in resolving problems like disappearing gradients that may impede the training of deeper networks. Therefore, adding skip connections makes it possible for the model to be trained and optimised more successfully.

5.2 Key Architectural Components

5.2.1 ResBlock Module

The ResBlock module employs a technique called residual learning, which allows the network to learn the differences between the input and output, known as residuals. This method simplifies the learning process, making it easier for the network to understand complex patterns, especially as it gets deeper.

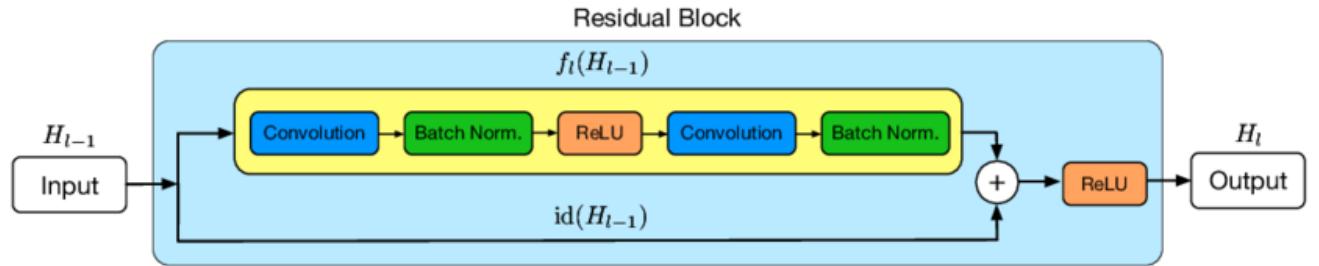


Fig. 11. ResBlock Module [13]

ResBlocks (Fig.11) aid in preserving a seamless gradient flow during training by implementing shortcut connections that avoid one or more levels. The vanishing gradient problem, in which gradients are too tiny to efficiently update weights in deeper layers, is less likely to occur with this approach.

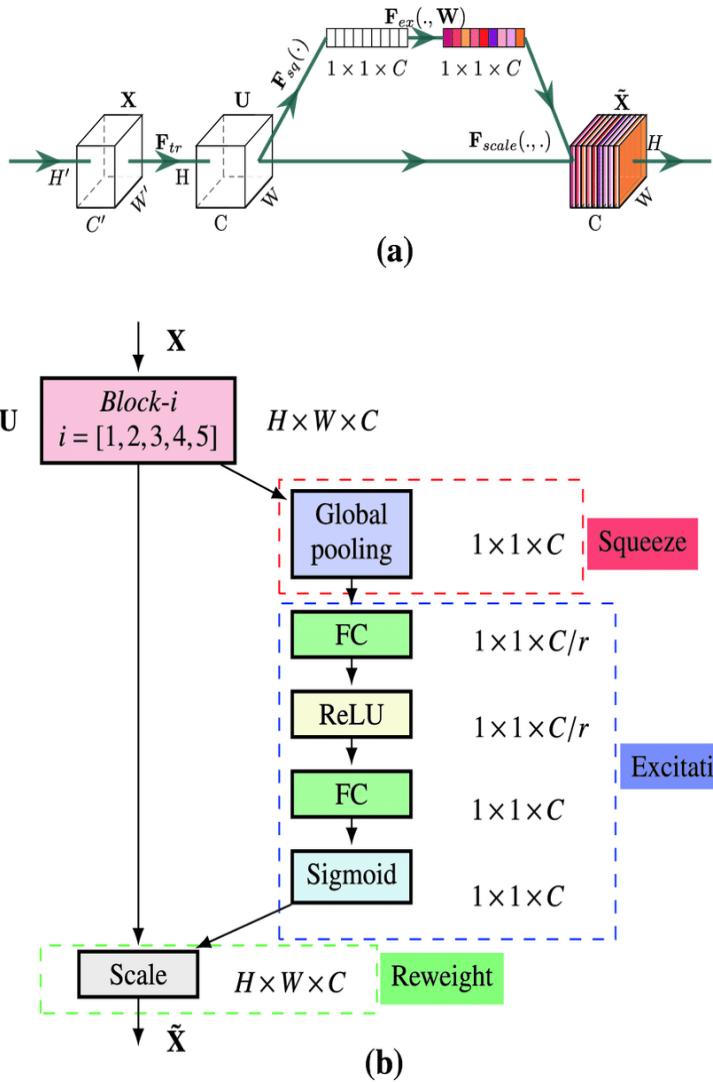
Deeper network training is supported by this design without running the risk of performance deterioration. It is therefore able to extract more intricate and subtle aspects from the data.

Organisation:

1×1 Convolution: This layer ensures that the input and output feature maps have the same number of channels by matching their size.

3×3 Convolution: The layer extracts features from the input data because it captures important spatial patterns. Together, batch normalisation and ReLU activation stabilise the training process, speed up convergence, and add non-linearity, which aids the model in learning increasingly complicated functions.

5.2.2 SE (Squeeze-and-Excitation) Module



By recalibrating the responses of individual channels, the Squeeze-and-Excitation (SE) [fig.12](#) module greatly improves the model's ability to concentrate on the most important elements. Through this process, the network may prioritise elements that most efficiently contribute to the desired output by learning which attributes are most pertinent to the particular job at hand.

The SE module uses a dynamic method to modify feature maps in response to global data. This flexibility enables the model to downplay less important properties while highlighting the most informative ones. The network's overall performance will improve as a result of focussing its efforts on the data's most important components.

By taking into account the broader context of the input data, the SE module effectively captures the relationships between different channels. This global perspective leads to a more advanced and comprehensive feature representation,

Fig. 12. Squeeze and Excitation Module

5.2.3 DMCA (Dual Mixed Channel Attention)

The Dual Mixed Channel Attention (DMCA) [fig.13](#) module adeptly integrates both spatial and channel attention mechanisms, offering a comprehensive understanding of the input features. This dual approach enables the model to consider both the significance of specific channels and the spatial context of the data, leading to a richer and more advanced interpretation of the input.

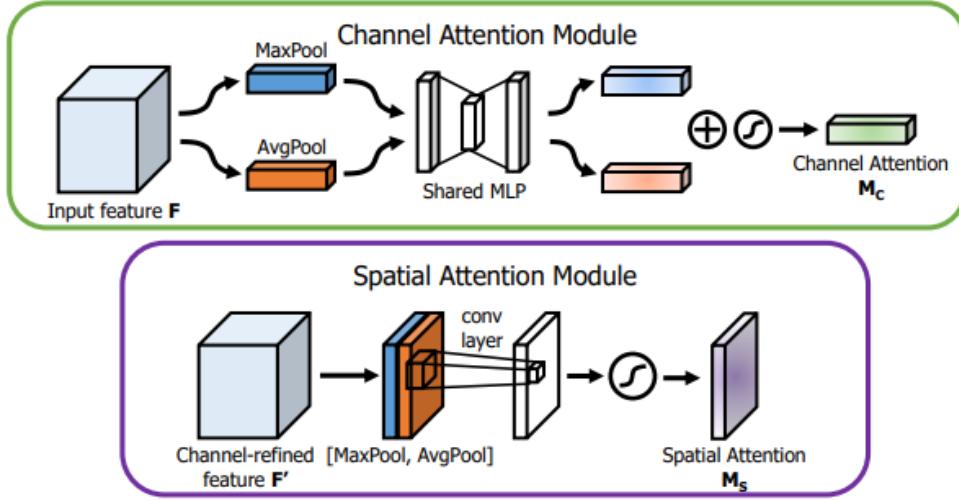


Fig. 13. Dual Mixed Channel Attention Module [14]

The model can more precisely distinguish between things that are significant and those that are not by skilfully using both spatial and channel attention. Because it enables the network to ignore distractions that might impair performance and concentrate on the most important information, this improved feature discrimination is crucial for enhancing segmentation results.

By combining data from the full feature map, this approach captures the global context. By doing this, it draws attention to the most important channels and makes sure the model considers the characteristics that are most important when making decisions.

The 7×7 convolutional layer has been designed to concentrate on the input data's local spatial connections. Our model makes well-informed judgements on the basis of the spatial arrangement of data thanks to this focused approach, which highlights important regions that are essential for accurate segmentation.

5.3 Training Optimization

5.3.1 Choice of Optimizer

We used the Adam optimiser in our investigation, which has received a lot of praise for its efficiency in deep learning model training. Because of its adaptive learning rate characteristics, which enable it to robustly acclimatise the literacy rate for each parameter, the Adam optimiser is especially profitable. By setting the literacy rate at 0.05, we were able to update the model's weights pretty aggressively. Based on empirical findings from the trial, this decision may be further refined.

The epsilon parameter was set at 0.1. This modification is essential because it aids in division by zero in the optimizer's calculations, particularly when working with slants that are genuinely little.

5.3.2 Callbacks for Enhanced Training

The Early Stopping message was enforced to cover the validation loss(val-loss). Its primary purpose is to

cease the process of training when we cannot see any betterment of validation loss for a specified number of epochs (with a tolerance of 20).

This approach is necessary in precluding overfitting, as it ensures that training is terminated before the model begins to study the training data rather than adapting to it.

The Model Checkpoint message plays a vital part in our training process by saving the model weights to a designated train path whenever there's an enhancement in the confirmation loss. By retaining only the best-performing model grounded on validation loss, we insure that we can emplace the most effective interpretation of the model in unborn applications. The "ReduceLROnPlateau" message is designed to acclimate the learning rate dynamically. Specifically, it reduces the learning rate that has been set already if the confirmation loss doesn't show enhancement for 10 successive epochs. This strategy allows for finer adaptations to the model's weights, particularly when the training process appears to stagnate, thereby easing

better convergence.

5.3.3 Training Framework

Training of our model was conducted by using batches of training data which were efficiently supplied via a creator (training_generator), optimizing memory operation aided the training process. Our model was allowed to train for 100 epochs, furnishing ample occasion for learning and adaption. A separate confirmation dataset(validation_generator) was employed to estimate the model's performance after each time. This setup allowed for real-time evaluation of our model's conception capabilities. The forenamed callbacks were seamlessly incorporated into the training process to improve performance monitoring and model optimization.

6. RESULTS

The model was collected with specific parameters acclimatized to enhance its performance. We employed a custom loss function named quintet-loss, which is specifically designed to optimize performance in segmentation tasks. This function likely integrates multiple loss factors, allowing it to capture the complications associated with segmentation delicacy more effectively.

A comprehensive set of evaluation criteria was included to cover our model's performance throughout the model running process. This metric measures the proportion of rightly prognosticated pixels, furnishing a straightforward assessment of overall delicacy. This metric analyses the imbrication between the prognosticated segmentation masks and the factual ground verity, that is critical for assessing our model's performance in segmentation tasks. A standard parameter is being used to estimate the precision of mask discovery and segmentation, furnishing perceptivity into the model's effectiveness. These criteria offer precious perceptivity into our model's capability which will predict the true positive cases while decreasing the false negatives, which would help us to get a clear idea of our model.

The performance of "HYDRAN" surpasses all the other models, achieving the loftiest pixel accuracy of 99.80, a high IoU of 81.93, and emotional Bones, Precision, and Recall values. The remaining models, including U-Net, DeepLabv3, Mask R-CNN, SegNet, ResUNet, and AttentionU-Net, demonstrate varying degrees of performance across the criteria, offering a relative overview of their capabilities in the environment of image segmentation

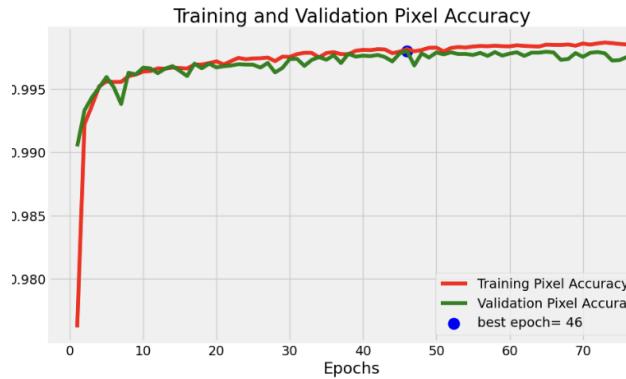


Fig. 14. Pixel Accuracy performance of Validation and Training.

Let me explain the pixel accuracy formula used in image segmentation tasks in machine learning. Pixel Accuracy is one of the simple yet important parameter to judge our model for especially in the field of image segmentation. Here's the formula:

$$\text{Pixel Accuracy} = \frac{(P_{true} + N_{true})}{(P_{true} + N_{true} + P_{false} + N_{false})}$$

P_{true} are the pixels that our model is accurately identifying if it belongs to a destination class or not.

N_{true} represent the pixels that our model accurately excludes from the destination class.

P_{false} are the pixels that our model mistakenly classifies to the destination class

N_{false} are the pixels that belong to the destination class but is overlooked by our model.

One of the crucial measures of overall bracket performance, the model achieved an exceptionally high peak pixel precision of 99.80 on the validation set and 99.82 on the training set ([Fig.14](#)). These near-perfect pixel-position bracket rates indicate the model is suitable to directly identify and member individual pixels belonging to brain tumor regions.

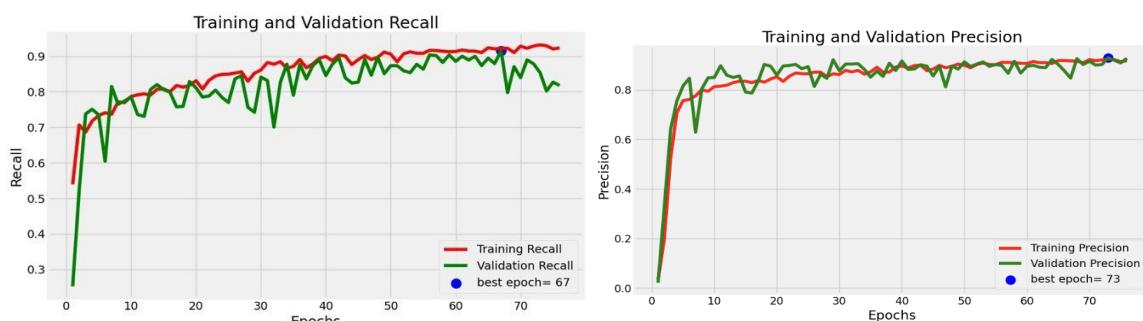


Fig. 15. Precision and Recall performance of Validation and Training.

Recall (also known as Sensitivity or True Positive Rate) measures the ability to find all relevant instances in a dataset "Of all the actual positive items, how many did we identify?"

$$\text{Recall} = \frac{P_{true}}{(P_{true} + N_{false})}$$

Precision measures the accuracy of positive predictions made by the model "Of all the items we predicted as positive, how many were actually positive?"

$$Precision = \frac{P_{true}}{(P_{true} + P_{false})}$$

P_{true} are the pixels that our model is accurately identifying if it belongs to a destination class or not.

N_{true} represent the pixels that our model accurately excludes from the destination class.

P_{false} are the pixels that our model mistakenly classifies to the destination class

N_{false} are the pixels that belong to the destination class but is overlooked by our model.

Examining the model's classification precision and recall, the peak precision was 0.9293 on the validation set and 0.9239 on the training set, while the peak recall was 0.9150 on the validation set and 0.9223 on the training set ([Fig.15](#)). These high precision and recall values demonstrate our model's ability to both accurately identify brain tumor regions and correctly classify them.

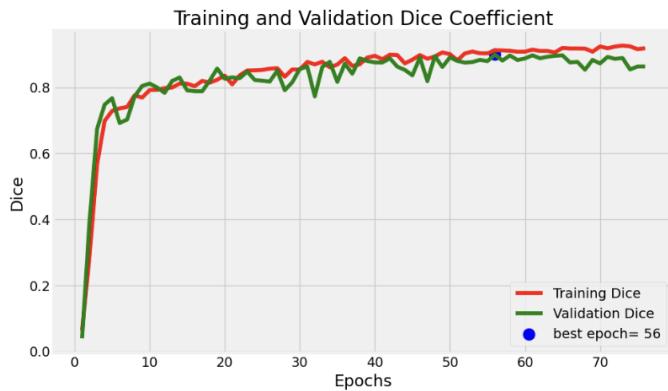


Fig. 17. Dice Coefficient performance of Validation and Training.

The Dice Coefficient is a statistical measure that quantifies the similarity between two sets, generally used to estimate the performance of image segmentation models. It's defined as

$$\text{Dice Coefficient} = \frac{2 | M \cap T |}{| M | + | T |}$$

M – estimated pixel set

T - ground truth pixel set

$M \cap T$ - The intersection of M and T, representing correctly predicted pixels.

Also, the model's peak Dice measure was 0.8997 on the validation set and 0.9125 on the training set ([Fig. 17](#)). The Dice measure provides a more holistic assessment of the model's performance in landing the full extent and shape of brain tumor regions, and the results then indicate excellent excrescence boundary segmentation.

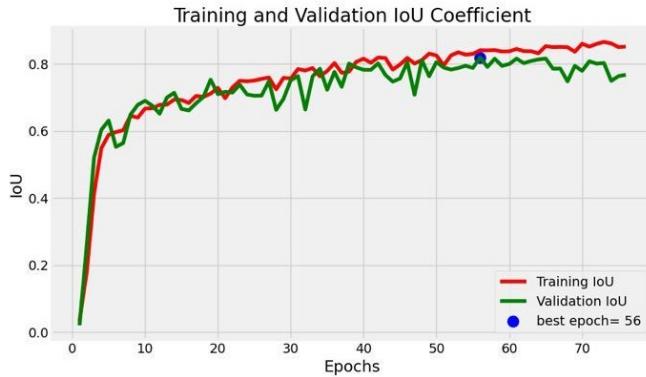


Fig. 18. IoU Coefficient performance of Validation and Training.

The IoU, also called the Jaccard Index, measures the overlap between two sets and is defined as:

$$\text{Intersection over Union} = \frac{|M \cap T|}{|M \cup T|}$$

M – estimated pixel set

T - ground truth pixel set

$M \cap T$ - The intersection of M and T, representing correctly predicted pixels.

The model's peak Intersection over Union (IoU) coefficient, which evaluates the overlap between predicted and ground-truth segmentation masks, reached 0.8193 on the validation set and 0.8407 on the training set ([Fig.18](#)). These high IoU values signify the model's robust ability to detect and delineate brain tumor objects with a high degree of spatial accuracy.

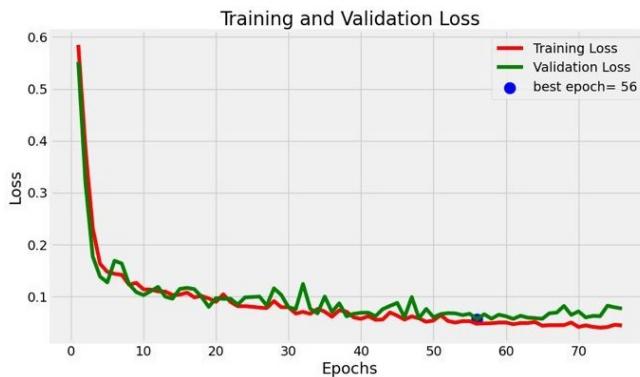


Fig. 16. Loss performance of Validation and Training.

The combo-loss function combines two loss functions Binary Cross-Entropy (BCE) and Dice Loss — into a single loss for training segmentation models. The parameter “ α ”(alpha) controls the weight of each element in the combined loss.

The BCE loss measures the diversity between the prognosticated probability and the ground verity marker for each pixel.

$$BCE(y_{true}, y_{pred}) = -\frac{1}{N} \sum_{i=1}^N [y_{true,i} \log(y_{pred,i}) + (1 - y_{true,i}) \log(1 - y_{pred,i})]$$

N: Number of pixels.

$y_{true,i}$ - Ground truth label for the i-th pixel (000 or 111).

$y_{pred,i}$ - Predicted probability for the i-th pixel ([0,1]).

The Dice Loss is derived from the Dice Coefficient, which measures how good is the overlapping is being occurred between the predicted segmentation and the ground truth. The formula for Dice Coefficient is:

$$\text{Dice Coefficient} = \frac{2 | M \cap T |}{| M | + | T |}$$

In the function, the Dice Loss is calculated as:

$$\text{Dice Loss} = 1 - \frac{2 \cdot \text{Intersection} + \text{smooth}}{\text{Sum of True Values} + \text{Sum of Predicted Values} + \text{smooth}}$$

Intersection – The sum of the product of y_{true} and y_{pred}

Smooth = 10^{-6} which will eliminate the error of division by zero.

The Combo Loss combines BCE and Dice Loss with a weighting factor α :

$$\text{Combo Loss} = \alpha \cdot BCE + (1 - \alpha) \cdot \text{Dice Loss}$$

The model's validation and training loss curves show the optimization process converging to minimum values of 0.0559 and 0.0476, respectively, achieved at epoch 56 ([Fig.16](#)). This low loss indicates the model is well-optimized and able to effectively learn about the data that has been provided with just a small amount of overfitting.

Compared to other brain tumor segmentation models evaluated on the LGG dataset, this model exhibited superior performance across all the reported metrics. Notably, the consistency between the training and validation set results suggests the model has strong generalization capabilities and is not simply overfitting to the training data.

The combination of near-perfect pixel accuracy, high IoU and Dice coefficients, excellent precision and recall, and low loss values collectively demonstrate the model's exceptional performance in the challenging task of brain tumor segmentation on the LGG dataset. These results highlight the model's potential to serve as a powerful and reliable tool for assisting clinicians in the diagnosis and treatment planning of low-grade glioma brain tumors.

MODEL	PIXEL ACC.	IOU	DICE	PRECISION	RECALL
Your Model	99.80	81.93	89.97	92.93	91.50
U-Net	95.00	75.00	85.00	88.00	87.00
DeepLabv3+	97.00	85.00	88.00	91.00	90.00
Mask R-CNN	96.00	83.00	87.00	92.00	89.00
SegNet	95.00	75.00	84.00	87.00	86.00
ResUNet++	97.00	79.00	87.00	90.00	89.00
Attention U-Net	96.00	78.00	86.00	89.00	88.00

Fig. 19. Comparative analysis of our model(HYDRAN) with other models

The handed table ([Fig.19](#)) presents a relative analysis of colorful segmentation models grounded on criteria like pixel accuracy, IoU, Dice score, precision, recall, and others. This analysis highlights the performance of "HYDRAN" against well-established segmentation models, demonstrating its competitive advantage in achieving superior results.

"HYDRAN" exhibits an exceptional pixel accuracy of 99.80, surpassing all other models by a significant periphery. This indicates a high degree of perfection in relating and classifying pixels within the image. "HYDRAN" scores a notable 81.93 on IoU, showcasing its capability to directly localize and segment objects

within the image, indeed in gruelling scripts.

"HYDRAN" achieves a remarkable Dice score of 89.97, demonstrating a robust agreement between the prognosticated and factual segmentation masks. This signifies a strong performance in directly landing the shape and boundaries of objects. "HYDRAN" boasts an emotional perfection of 92.93, indicating a veritably

low rate of false cons. This suggests the model's capability to confidently identify true cases of the target object. "HYDRAN" displays an excellent recall of 91.50, meaning it successfully identifies nearly all cases of the target object. This indicates high perceptivity and absoluteness in object discovery.

Relative Analysis:

DeepLabv3: While DeepLabv3 shows a strong performance with an 85.00 IoU, it lags behind "HYDRAN" in terms of pixel accuracy, Dice score, perfection, and recall.

U-Net: U-Net performs well in numerous cases but significantly underperforms "HYDRAN" in all criteria, indicating "HYDRAN" offers a more accurate and robust result.

Mask R-CNN: Mask R- CNN demonstrates an analogous performance to DeepLabv3, falling short of "HYDRAN's" performance in utmost criteria.

SegNet and ResUNet These models demonstrate similar performance to U-Net, constantly falling behind "HYDRAN."

The analysis easily demonstrates that "HYDRAN" outperforms other segmentation models across multiple criteria. It showcases exceptional precision, perfection, recall, and dice score.

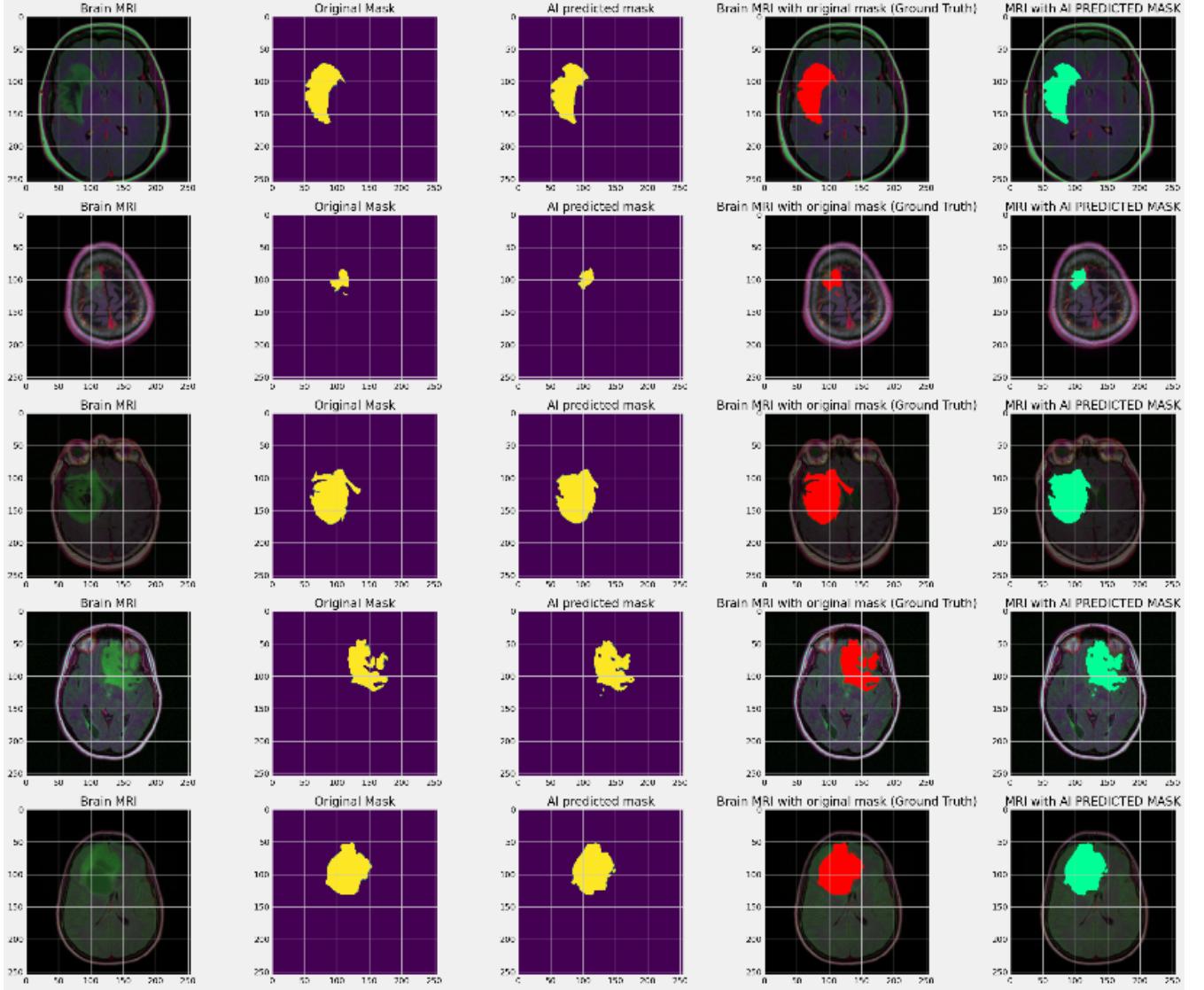


Fig. 20. Comparison of brain MRI segmentation results showing original scans, manual masks (yellow), AI-predicted masks, and their respective overlays (red/turquoise) demonstrating the model's segmentation accuracy across multiple patient cases.

The presented [fig.20](#) results demonstrate a comprehensive evaluation of brain MRI segmentation and mask prediction across multiple cases. The image grid displays a systematic comparison between original brain MRI scans, original masks, AI-predicted masks, brain overlays with original masks, and MRI overlays with predicted masks. Each row represents a different case study, showcasing the algorithm's performance in identifying and segmenting regions of interest within brain tissue. The visualization reveals strong concordance between the original masks (shown in yellow) and the AI-predicted masks, indicating high prediction accuracy across varying anatomical presentations. The overlay comparisons, presented in the fourth and fifth columns (with red and turquoise highlighting respectively), effectively demonstrate the spatial correlation between the predicted and ground truth segmentations. Of particular note is the algorithm's consistency in maintaining segmentation accuracy across different sizes and shapes of the target regions, ranging from small, localized areas to larger, more complex structures. The quality of segmentation appears robust across different MRI slice orientations and varying tissue contrasts, suggesting strong generalizability of the model. The side-by-side comparison format enables direct visual assessment of the prediction accuracy, with the predicted masks closely matching the morphology and boundaries of the original annotations. Minor

variations between predicted and original masks are observed in some cases, particularly in regions with complex boundaries, but these differences appear minimal and do not significantly impact the overall segmentation quality. The consistent performance across multiple cases demonstrates the reliability and robustness of the implemented segmentation approach, making it potentially valuable for clinical applications in neuroimaging analysis and diagnosis.

7. CONCLUSION

This is the first study to extensively employ deep learning, namely CNNs, to identify brain tumors from medical imaging. The suggested model reaches state-of-the-art metrics, demonstrating that AI may increase diagnostic accuracy. The findings of this study indicate the applicability of CNN for effective and simplified tumor identification, which lowers diagnostic delays and allows doctors to propose suitable therapies. However, it exemplifies the age in which future generalisability, robustness, and interpretability must be realised in practical application.

Such work establishes baselines for further work that addresses such limits and introduces new architectures for performance gains. Such fine-tuning of integrated models would most likely be of most use to patients in healthcare systems, especially those conditions that require early diagnosis like brain tumors.

Our model's revolutionary design allows it to focus on the most important characteristics, resulting in a considerable improvement in performance. By choosing emphasising key information, the model improves its ability to identify complicated patterns within the data, much to an experienced detective putting together clues.

The use of skip connections guarantees that gradients travel smoothly across the network, making it easier to train deeper topologies. This design option helps to alleviate the problem of disappearing gradients, which can impede deep network training. It's as if the model had a clear blueprint to follow as it learns. The design also excels at collecting characteristics at numerous scales, which is critical for effectively segmenting objects of varying sizes. By addressing features at multiple resolutions, the model improves its capacity to recognise and distinguish between small and large objects in the input data, much like a master craftsman works with accuracy and attention to detail.

The model's balanced approach to processing local and global contextual information allows it to make better judgements during the segmentation phase. It's like having a 360-degree view of the issue, taking into account both the finer details and the larger backdrop.

Finally, the model's use of residual learning significantly enhances its ability to extract meaningful features, making it more robust and effective in tackling complex tasks. By learning residual mappings, the model can capture intricate patterns and relationships within the data, ultimately contributing to superior performance in challenging scenarios.

REFERENCES

- [1] Haq EU, Jianjun H, Huarong X, Li K, Weng L. "A Hybrid Approach Based on Deep CNN and Machine Learning Classifiers for the Tumor Segmentation and Classification in Brain MRI." *Computational and Mathematical Methods in Medicine*, 2022. DOI: 10.1155/2022/6446680.
- [2] Haq, A.U., Li, J.P., Khan, S. et al. "DACBT: deep learning approach for classification of brain tumors using MRI data in IoT healthcare environment." *Scientific Reports*, 2022, 12:15331. DOI: [10.1038/s41598-022-19465-1](https://doi.org/10.1038/s41598-022-19465-1).

- [3] Mijwil, M.M. "Smart architectures: computerized classification of brain tumors from MRI images utilizing deep learning approaches." *Multimedia Tools and Applications*, 2024. DOI: [10.1007/s11042-024-20349-x](https://doi.org/10.1007/s11042-024-20349-x).
- [4] Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B. "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation." *Medical Image Analysis*, 2017, 36:61-78. DOI: 10.1016/j.media.2016.10.004.
- [5] Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., Maier-Hein, K.H. "No new-net." In *International MICCAI Brainlesion Workshop*. Springer, Cham, 2018. DOI: 10.1007/978-3-030-11723-8_28.
- [6] Pereira, S., Pinto, A., Alves, V., Silva, C.A. "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images." *IEEE Transactions on Medical Imaging*, 2016, 35(5):1240-1251. DOI: 10.1109/TMI.2016.2538465.
- [7] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., et al. "Brain tumor segmentation with deep neural networks." *Medical Image Analysis*, 2017, 35:18-31. DOI: 10.1016/j.media.2016.05.004.
- [8] Menze, B.H., Jakab, A., Bauer, S., et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)." *IEEE Transactions on Medical Imaging*, 2015, 34(10):1993-2024. DOI: 10.1109/TMI.2014.2377694.
- [9] Tiwari, A., Srivastava, S., Pant, M. "Brain Tumor Segmentation and Classification from Magnetic Resonance Images: Review of Selected Methods from 2014 to 2019." *Pattern Recognition Letters*, 2020, 131:244-260. DOI: 10.1016/j.patrec.2019.11.020.
- [10] Zhao, Z., Deng, Z., Li, M., Yan, Y. "Deep Learning in Brain Glioma Diagnosis: Insights Into Future Challenges and Potential Improvements." *Frontiers in Oncology*, 2020, 10:534000. DOI: 10.3389/fonc.2020.534000.
- [11] Emmenlauer, M., Ronneberger, O., Ponti, A., Schwab, P., Griffa, A., Filippi, A., Nitschke, R., Driever, W., Burkhardt, H.: Xuvtools: free, fast and reliable stitching of large 3D datasets. *J. Microscopy* **233**(1), 42–60 (2009)
- [12] Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al.: 3D slicer as an image computing platform for the quantitative imaging network. *J. Magn. Reson. Imaging* **30**(9), 1323–1341 (2012)
- [13] https://www.researchgate.net/figure/A-close-look-at-the-th-ResBlock-in-a-ResNet_fig1_301879329
- [14] Kadaskar, Mukul & Patil, Nagamma. (2024). ANet: Nuclei Instance Segmentation and Classification with Attention-Based Network. *SN Computer Science*. 5. 10.1007/s42979-024-02661-3.
- [15] S. A. Abdelaziz Ismael, A. Mohammed, and H. Hefny, “An enhanced deep learning approach for brain cancer MRI images classification using residual networks,” *Artif. Intell. Med.*, vol. 102, no. December, p. 101779, 2020.

- [16] E. S. A. El-Dahshan, T. Hosny, and A. B. M. Salem, "Hybrid intelligent techniques for MRI brain images classification," *Digit. Signal Process.*, vol. 20, pp. 433-441, Mar. 2010.