Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI:10.21474/IJAR01/19571
DOI URL: http://dx.doi.org/10.21474/IJAR01/19571

RESEARCH ARTICLE

CAPSULE NETWORK FOR BRAIN TUMORS SEGMENTATION IN MRI IMAGES

Yasmin Farisya Norazlan[1], Siti Salasiah Mokri[1], Nurul Fatihah Ali[1], Ashrani Aizuddin Abd Rahni[1], Aqilah Baseri Huddin[1], Noraishikin Zulkarnain[1], Ahmad Abid Hakimi Md Salleh[1] and Nor Aziyatul Izni[2]

1. Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Malaysia.
2. Centre of Foundation Studies, Universiti Teknologi MARA, Cawangan Selangor, Kampus Dengkil, 43800 Dengkil, Selangor, Malaysia.

Manuscript Info

Manuscript History
Received: 30 July 2024
Final Accepted: 31 August 2024
Published: September 2024

Key words:-
Segmentation, Brain Glioma, Deep Learning, Capsule Network,U-Net Network

Abstract

Glioma is an abnormal, irregularly shaped cell growth in the brain. Diagnosing glioma often relies on magnetic resonance imaging (MRI), which requires precise segmentation of the tumor for effective analysis. Convolutional Neural Networks (CNN), such as U-Net, is widely used in medical image processing, particularly for image segmentation tasks. However, CNNs have the limitation in effectively learning spatial relationships within images. To address this, Capsule Networks (CapsNet) is introduced, utilizing capsule dynamic routing to better capture spatial hierarchies. This paper aims to investigate the performance of SegCaps, a segmentation model based on Capsule Networks, for brain glioma segmentation in MRI images, compared to the CNN-based U-Net model. Both models were tested on the BraTS2018 glioma dataset, which includes 374 MRI images of brain tumors across four modalities (T1, T1c, T2, FLAIR). The performance of SegCaps and U-Net was evaluated using two key segmentation metrics: Dice coefficient and Jaccard index. The results show that SegCaps outperformed U-Net with a Dice coefficient of 87.96% compared to U-Net's 85.56%, demonstrating a 2.4% improvement. Additionally, SegCaps required fewer parameters than the U-Net model, highlighting its efficiency. In conclusion, SegCaps can be considered as a promising alternative for glioma segmentation in MRI images. Future work could focus on refining the SegCaps model to enhance its performance while reducing computational costs.

Introduction:-

The World Health Organization reported that that cancer was the second leading cause of death, responsible for over 9.6 million deaths in 2018. There are various forms of cancer, including brain tumors, which are known for their aggressiveness and are considered among the deadliest types of cancer. Brain tumor can be classified as either a cancerous (malignant) or non-cancerous (benign) growth within the brain tissues (Wadhwa et al., 2019), (Manan et al., 2022). Malignant brain tumors have irregular shape and high metabolic activity within the cell group structures. On the other hand, benign tumors typically exhibit uniformity in structure and lower cellular activity. Examples of malignant tumors, also known as high-grade tumors, include glioblastoma and astrocytoma. These types of brain

Corresponding Author:-Siti Salasiah Mokri
Address:-Department of Electrical, Electronic and Systems of Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Malaysia.

cancer are extremely dangerous due to their aggressive nature and rapid growth (Bahadure, Ray, &Thethi, 2017). The diagnosis of brain tumour is carried out via medical imaging such magnetic resonance imaging (MRI) (Voon et al., 2023). In the case of brain cancer, tumour segmentation is particularly used to obtain an accurate representation of the tumor's shape, structure, and location within the patient's brain (Sun, Zhang, Chen, & Luo, 2019).

Therefore, accurately identifying the types of brain tumour tissues within their surroundings through segmentation is crucial (Din and Rahni, 2020). Although brain tumors are often manually segmented by experienced radiologists, the manual segmentation approach is prone to errors due to the vast volume of Magnetic Resonance Imaging (MRI) data and is time-consuming (Sun, Zhang, Chen, & Luo, 2019). This has led the researchers to develop automatic brain tumor segmentation techniques, offering valuable support to radiologists in terms of diagnosis, treatment planning and to reduce their daily workload (Liu et al., 2023). However, developing accurate automatic brain tumor segmentation algorithms remains a challenging task. According to Chen, Ding, and Liu (2019), the challenges come from the complex shape and structure of the tumor tissues that vary significantly from one patient to another. In addition, the boundaries between healthy and cancerous tissues are blurry resulting in difficulty to achieve accurate segmentation.

In recent years, several algorithms have been proposed for brain tumor segmentation. Specific to deep learning methods, the commonly used segmentation algorithm is U-Net architecture which is based on convolutional neural network (CNN) (Ronneberger et al., 2015), (Ali et al., 2023), (Iqbal et al., 2022). Although CNNs perform well in brain tumor segmentation tasks, they also have several limitations (Sabour, Frosst, & Hinton, 2017; Hinton, Sabour, &Frosst, 2018). CNNs fail to fully learn spatial relationships within the images due to the utilization of pooling layers. This is a critical drawback for segmentation task because most of the critical information about the tumor is found in the surrounding tissues to accurately define the boundaries.

To address the limitations of pooling layers, Sabour, Frosst, & Hinton., (2017) and Hinton, Sabour, &Frosst., (2018) proposed the concept of capsule network, where the spatial information at the neuron level is preserved within the capsules. By grouping the neurons together in capsules, fewer parameters are used in the capsule network as compared to CNN suggesting that such an architecture has the potential to generalize effectively. Following to Sabour, Frosst, & Hinton., (2017) and Hinton, Sabour, &Frosst., (2018), Lalonde and Bagci (2018) successfully implemented a capsule network intended for pathological lung segmentation which is called as SegCaps. This marks the first instance in the literature where capsule network has been applied to the segmentation task for pathological lung. Instead, this paper investigates and compares the performance of SegCaps in segmenting the brain tumour of MRI images with the established CNN based U-Net segmentation model.

Literature Review:-

CNN is composed of multiple layers, including convolutional layers, pooling layers, activation layers, and fully connected layers (Zhou, Ruan, & Canu, 2019), (Zamani et al., 2023), (Thanoon et al., 2023). Convolutional layers are responsible for feature extraction where the kernel moves across the image to generate various feature maps. Next, the pooling layer reduces the size of each feature map by selecting only the maximum or average value from a region, a process known as subsampling. The activation function, such as the commonly used Rectified Linear Unit (ReLU) determines neuron activation by converting negative values to zero and passing positive values as output (He, Zhang, Ren, & Sun, 2015) (Long, J., Shelhamer, E., & Darrell, T., 2015). Following this, the fully connected layer includes neurons connected to all activated neurons from the previous layer, flattening the feature map matrix before inputting it into the output layer. Finally, an activation function, such as softmax, is applied to predict the class of each output. During the CNN training process, the kernel weights are adjusted through backpropagation process.

Although CNN shows impressive results for various tasks in the field of computer vision research, there are some limitations that have been identified. The main limitation of CNN is its disability to distinguish the position, texture, and deformation of images or image parts (Sabour, Frosst, & Hinton, 2017; Hinton, Sabour, &Frosst, 2018).To alleviate this limitation, Sabour, Frosst, & Hinton (2017) and Hinton, Sabour, &Frosst (2018) introduced a new architecture known as capsule network (CapsNet). The core principle of capsule network is that the neurons store feature activations in vectors, or capsules, rather than as scalars. These vectors, represented by multiple capsules, encode information about the spatial orientation, frequency, and other characteristics of the extracted features. In addition, capsule network employs dynamic routing algorithm, which relies on the agreement or similarity between

capsules in the connected layers. This mechanism, known as 'routing-by-agreement,' facilitates the activation and proper arrangement of features between layers.

Lalonde and Bagci (2018) were the first to apply the concept of capsule network for object segmentation named as SegCaps to segment the pathological lungs from Computed Tomography (CT) scans. SegCaps is founded from the work of Sabour et al. (2017). Due to the information loss associated with the locally constrained dynamic routing mechanism, Lalonde and Bagci also developed a deep convolutional-deconvolutional capsule architecture. This approach allows for pixel-level prediction of object labels while retaining near-global contextual information. Additionally, they extended class-targeted masked reconstruction as a regularization strategy. Their work demonstrated excellent performance in pathological lung segmentation, suggesting that capsule networks might be a better approach than traditional CNN architectures.

Capsule network is designed to better model hierarchical relationships. It mimics closely the organization of biological neurons. The main idea is to add structures called “capsules” to convolutional network (CNN) and use the output from several of these capsules to form a more stable representation with respect to various disturbances for higher-level capsules (Sabour, Frosst& Hinton, 2017). The nodes in CNN are neurons that produce scalar outputs, whereas in Capule network, the capsules are the nodes of the network, and their output appears as vectors, which provide more information than scalar outputs. The length of these vectors indicates the likelihood of a specific entity existing in the input image.

In Capsule network, an advanced technique is used to connect capsules between layers, where the network weights are obtained based on an iterative optimization strategy. In this routing technique, the output from the previous capsule is given as input to the next capsule, and during the iterative process, the similarity between the input and the capsule output is compared. Ultimately, the previous capsule is routed to the capsule with the matching output (Sabour et al., 2017). Recently, Capsule network has been applied to various real-world applications in which one of the most significant applications is medical image segmentation (LaLonde et al., 2021) named as SegCaps.

Here, the input for SegCaps model is an MR image with a size of 192 x 192 pixels. First, the input image is passed through a 2D convolutional layer to create 16 feature maps, all of which have the same shape as the input image. This layer produces an initial set of capsules, consisting of a single type of capsule on a 192 × 192 grid, where each capsule is a 16D vector. Next, the 16D capsules are routed to the first convolutional capsule layer, which produces two 16D capsules. As a regularization strategy, Lalonde and Bagci (2018) extended the original capsule implementation by adding input reconstruction. The reconstruction is performed in a convolutional network consisting of three layers with a kernel size of 1 × 1. Then, a squared mean error is calculated between the positive input class and the reconstruction. Regarding the kernel size in the convolutional and deconvolutional capsule layers, a kernel size of 5x5 is used in the convolutional capsules and a kernel size of 4x4 is used in the deconvolutional capsules (Speet& Cassani, 2020).

Methodology:-

In this study, the BraTS 2018 dataset (Menze et al., 2015; Bakas et al., 2017) was used to train and evaluate SegCaps and U-Net model. This dataset is provided by the Brain Tumor Segmentation Challenge (BraTS) 2018, organized by the Center for Biomedical Image Computing and Analytics (CBICA). The data used for training consists of MRI scans from 259 patients diagnosed with high-grade glioma (HGG). On average, the tumor area only accounts for 2.79 percent of the labels in the ground truth images, while the remaining 97.21 percent is background.

Each MRI scan is in NIfTI file format and consists of 155 axial slices with a resolution of 240 x 240 pixels. Additionally, all patients were scanned with four MRI sequences: T1-weighted (T1), contrast-enhanced T1-weighted (T1ce), T2-weighted (T2), and Fluid-Attenuated Inversion Recovery (FLAIR), as shown in Fig.1. Fig.2 provides an example of a FLAIR MRI scan with the ground truth mask. All images have been skull-stripped and resampled to an isotropic resolution of 1 mm.

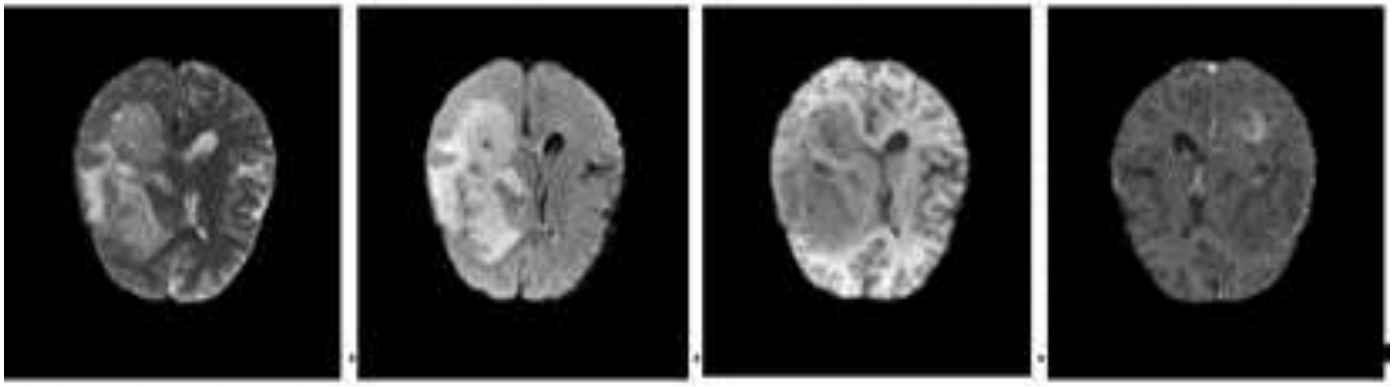

Fig. 1:- Brats2018 subfiles, from left to right, T1ce, T1, T2 and FLAIR.

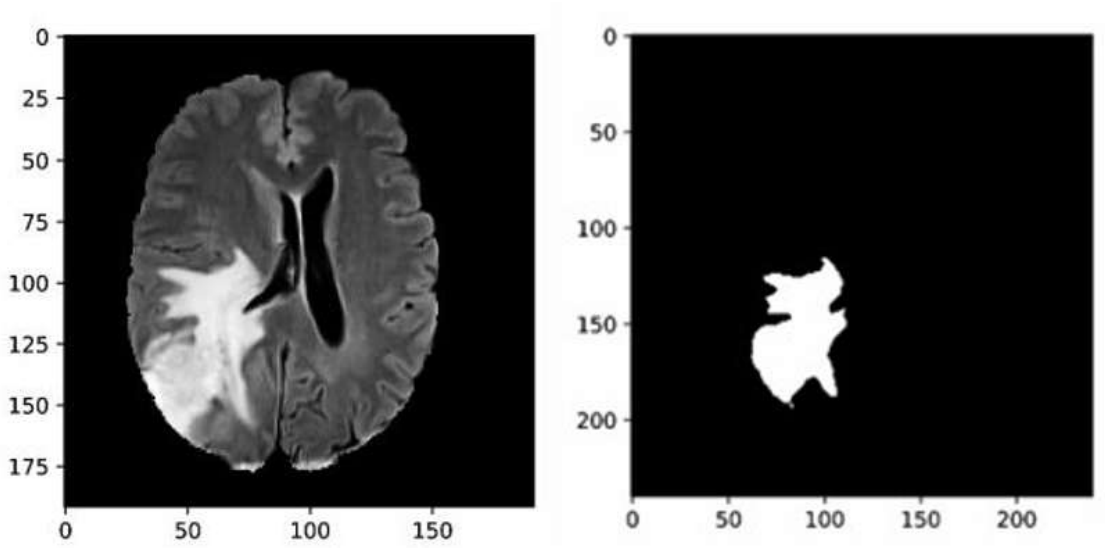

Fig. 2:- (Left) a single slice of LAIR MRI scan. (Right) the ground truth segmentation mask.

In addition, as BraTS dataset comes from various medical institutions, the images are produced by different MRI scanners. As a result, MRI images will have varying intensities depending on the MRI scanner used by each medical institution. This will benefit the generalizability and robustness of the tested algorithms. Finally, the ground truth segmentation masks were manually obtained by the radiologists according to the Brain Tumor Segmentation Benchmark (Menze et al., 2015). The ground truth images are the same size as the actual MRI scans, but with background and healthy brain tissue labeled as zero and tumor tissue labeled as one.

To efficiently train the capsule network, the MRI data was preprocessed using the BraTS preprocessing helper function to prepare the data for deep learning training. This function performs the following operations: cropping the data to focus on the regions containing the brain and tumor. This cropping process reduces the data size while preserving the most critical parts of each MRI volume and its corresponding labels, normalizing each modality of each volume independently by subtracting the mean and dividing by the standard deviation of the cropped brain region, and splitting the 344 images for training and 30 images for testing.

The training of the network was carried out by taking the preprocessed input data and feeding it into a deep neural network, specifically using a capsule neural network. This network comprises two main components: an encoder and a decoder, with a total of six layers. To train the network, 20 epochs were conducted using the standard Adam optimization method, along with early stopping to prevent overfitting and underfitting.

Data augmentation was carried out based on the input model, employing two different methods in a hybrid approach. The batch size was typically set to 75% of the training data (excluding the validation set). Despite

requiring substantial memory, the results showed excellent performance, even in fewer epochs. The initial learning rate was set to 0.001, as it proved effective across various training processes.

Factors such as the patient's position in the scanner, the scanner itself, and other unknown issues can cause brightness variations in MR images, known as bias fields. These low-frequency artifacts can degrade the MR images and, if left uncorrected, lead to inaccurate image processing results. Before segmentation or classification, preprocessing is necessary to correct the influence of bias fields. This correction was achieved by applying the SimpleITK N4 bias field correction filter to all images in the dataset. The BRATS database provides different target structures like edema, enhancing tumors, non-enhancing tumors, and background. The study utilized images in color, grayscale, or intensity with a size of 128 × 128 pixels. Automatic brain tumor detection in patients involves two critical stages: image segmentation and edge detection.

To evaluate the performance of the brain glioma segmentation, the Dice Similarity Coefficient (DSC) is used (Li et al., 2019; Havaei et al., 2014; Kamnitsas et al., 2017; Myronenko, 2018; Nazarudin et al., 2023). It is employed to measure the similarity between the predicted segmentation mask and the ground truth segmentation mask and can be expressed as:

$$DSC = \frac{2|G \cap P|}{|G| + |P|} \tag{1}$$

Here, G represents the cardinality of the pixel set in the ground truth segmentation, and P represents the cardinality of the pixel set in the predicted segmentation. The Dice similarity coefficient ranges from 0 to 1. If there is no overlap between the two sets, the Dice coefficient will be 0. When the ground truth segmentation and the predicted segmentation are perfectly identical, the Dice coefficient will be equal to 1. In addition, the Jaccard coefficient is defined as the size of the intersection divided by the size of the union of the sets. In the context of image segmentation, such as in MRI tumor segmentation, it compares the overlap between the predicted segmentation and the ground truth, with values ranging from 0 (no overlap) to 1 (perfect overlap). The Jaccard index corresponded to the Dice an in the following equation:

$$Jaccard = \frac{|G \cap P|}{|G| + |P| - |G \cap P|} \tag{2}$$

Training loss is a measure of the error or discrepancy between the predicted segmentation and the actual ground truth labels during training. It reflects how well the model fits the training data. According to Mao et al., (2023) the cross-entropy loss function, particularly binary cross-entropy (BCE), is commonly used training loss for model training. Binary cross-entropy is calculated by comparing the predicted pixel vector, $y_i$with the actual pixel vector$\hat{y}$, and it is represented by the following formula,

$$LOSS_{BCE} = -\sum_{i=1}^{N} y_i log\hat{y} \tag{3}$$

However, an imbalanced dataset in terms of class representation may cause the loss function to behave incorrectly. To address class imbalance in the dataset, the weighted binary cross-entropy W_{BCE} loss function is used. This approach is effective because the ground truth segmentation mask typically contains many background pixels and very few tumor regions (Long, Shelhamer, & Darrell, 2015). W_{BCE} can be represented as follows, where w_i represents the weights applied to both classes. This loss function was also used in the original implementation of SegCaps by LaLonde and Bangchi (2018).

$$LOSS_{WBCE} = -\sum_{i=1}^{N} w_i y_i log\hat{y} \tag{4}$$

Based on the work by Dong et al. (2017), the U-Net model for brain tumor segmentation uses a Dice-based loss function. Dice Loss (DL) is essentially the inverse of the Dice similarity coefficient and is defined as:

$$LOSSDSC = \frac{2|G \cap P|}{1 - |G| + |P|} \quad (5)$$

This study is conducted using the Python programming language within the Visual Studio Code platform. The following Python libraries are used:

SimpleITK 1.2.4 for loading NIfTI files
NumPy 1.18.3 for data structures and other operations
Matplotlib 3.2.1 for visualization
Scikit-image 0.16.2 for image processing
Scikit-learn 0.22.2 for ML-related functions
Keras 2.3.1 and TensorFlow 1.15.2 for building all models
TensorBoard 1.15.0 for visualizing training logs

Results and Discussions:-

The training process of SegCaps consists of several epochs, during which the model learns from the training data and adjusts its parameters to improve performance. The number of epochs is set to 20 in which the model is trained over 20 complete passes through the training dataset. This allows the model to learn complex patterns and capture important features for accurate tumor segmentation. Throughout the training, two key metrics are monitored to assess the model's progress: training loss and validation dice score.

To assess the effectiveness of the proposed method, experiments were conducted using the BRATS 2018 dataset, which consists of 374 cases of brain tumor MRI images. This dataset includes four modalities (T1, T1c, T2, FLAIR) and accuracy of the segmentation is measured using the Dice score, which quantifies the overlap between the predicted segmentation and the ground truth labels for three tumor regions: the whole tumor, the tumor core and enhancing core. The accuracy of the segmentation is measured using the Dice score, which quantifies the overlap between the predicted segmentation and the ground truth.

First, SegCaps was trained and validated on a 90-10 percent split of the dataset. The training process lasted for 20 epochs, with a batch size of 10. The Adaptive Moment Estimation (Adam) optimizer (Kingma & Ba, 2014) was used, with an initial learning rate of 0.001. In addition, the U-Net model was also trained using a 90-10 percent split of the dataset. The number of epochs required to train the model was 20, with a batch size of 30. The U-Net model was trained using a learning rate of 0.001.

The training loss is a measure of the error or difference between the predicted segmentation and the actual ground truth labels during training. It reflects how well the model fits the training data. Fig. 3 and Fig.4 show the training loss during U-Net and SegCaps training respectively. It is observed that the training loss of U-Net is higher than that of SegCaps, reaching a minimum loss value of 0.05 at epoch 19 compared to SegCaps, which has a minimum value of 0.01. This indicates that U-Net requires more time to learn the data as compared to SegCaps. Additionally, the validation Dice score of U-Net is lower than that of SegCaps, reaching its maximum at epoch 20.

In Fig. 4, SegCaps achieves a minimum training loss value of less than 0.01 at epoch 19, indicating that the model achieves a good fit with the training set. The trend of decreasing training loss and increasing validation Dice score is clear. This suggests that the model consistently learns and improves the segmentation performance as the training progresses.To avoid the issue of overfitting, where the model performs well on training data but fails to generalize to unseen data, early stopping was employed. A criterion was set to halt training if the validation dice score did not show improvement for 10 consecutive epochs. This mechanism ensures that the model is stopped at the optimal point, preventing overfitting and allowing for better generalization. In this way, the saved model has the highest accuracy on the validation set, ensuring that the most optimal model is used for testing.

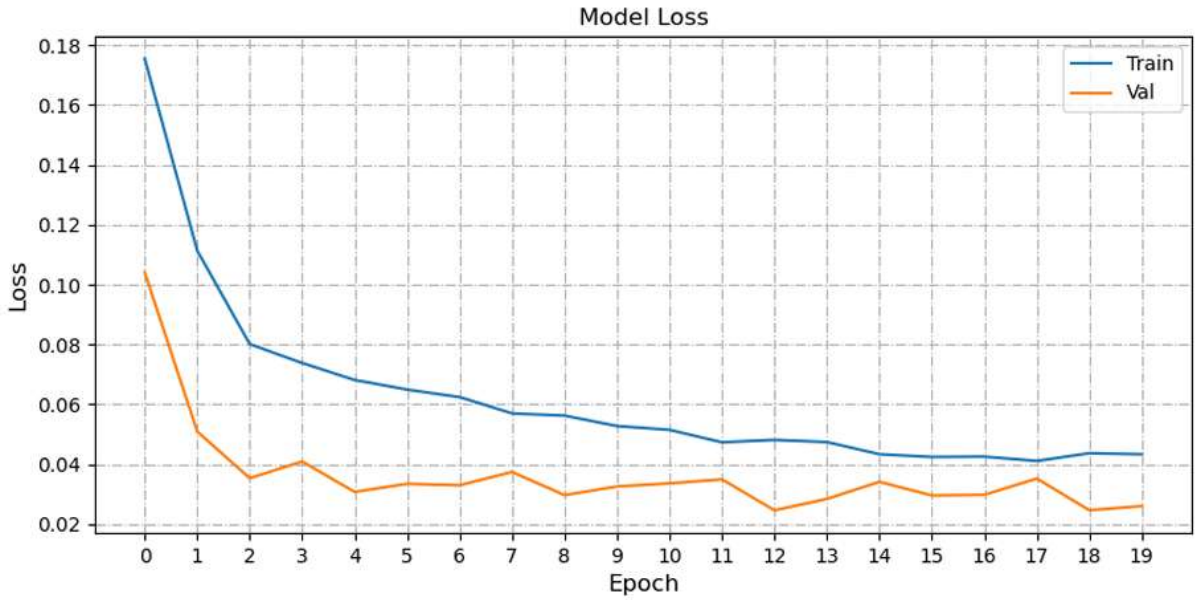

Fig. 3:- Plot of training loss for U-Net model.

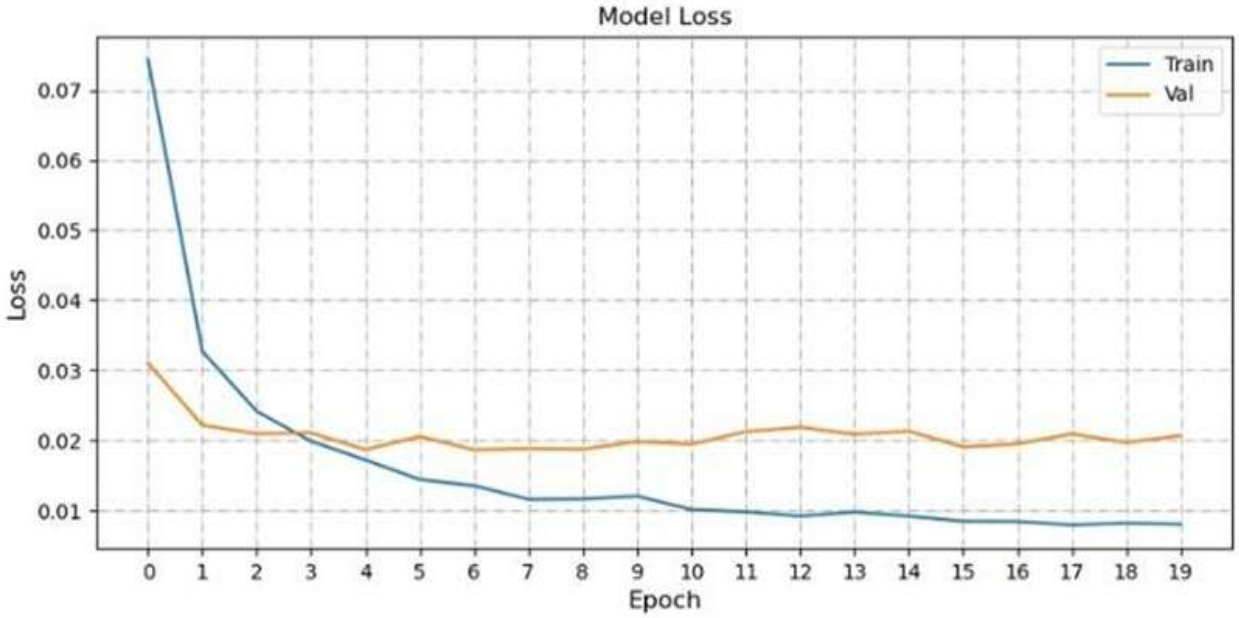

Fig. 4:- Plot of training loss for SegCaps model.

Based on the evaluation of the validation Dice score, the model at epoch 20 was selected as the final model for testing. This model demonstrated the highest accuracy and proved its capability to segment brain tumors accurately. Using this model, the performance evaluation of the capsule network was extended to an independent test set, assessing its effectiveness in segmenting brain tumors on previously unseen MRI images.

Visually, Figure 5 shows the segmentation results for both methods. In this figure, it shows that the predictions generated by U-Net are not as accurate and detailed as the images produced by the SegCaps model. Although most of the tumor tissues are accurately segmented, the results demonstrate that U-Net is unable to capture the fine contours of the entire tumor area, while SegCaps appears to handle this irregularity well. Due to U-Net's inability to capture the tumor's contour details, U-Net produces suboptimal predictions, indicating that the model also includes small portions of healthy tissue in the segmentation. It is observed that SegCaps does not ignore certain internal

regions within the tumor core. This may explain the improved performance of SegCaps compared to U-Net, as the ignored internal tumor areas are treated as background. Regarding U-Net's performance in isolating the tumor core regions, the internal core areas appear to be overlooked. This finding explains the lower and less efficient Dice coefficient achieved by U-Net.

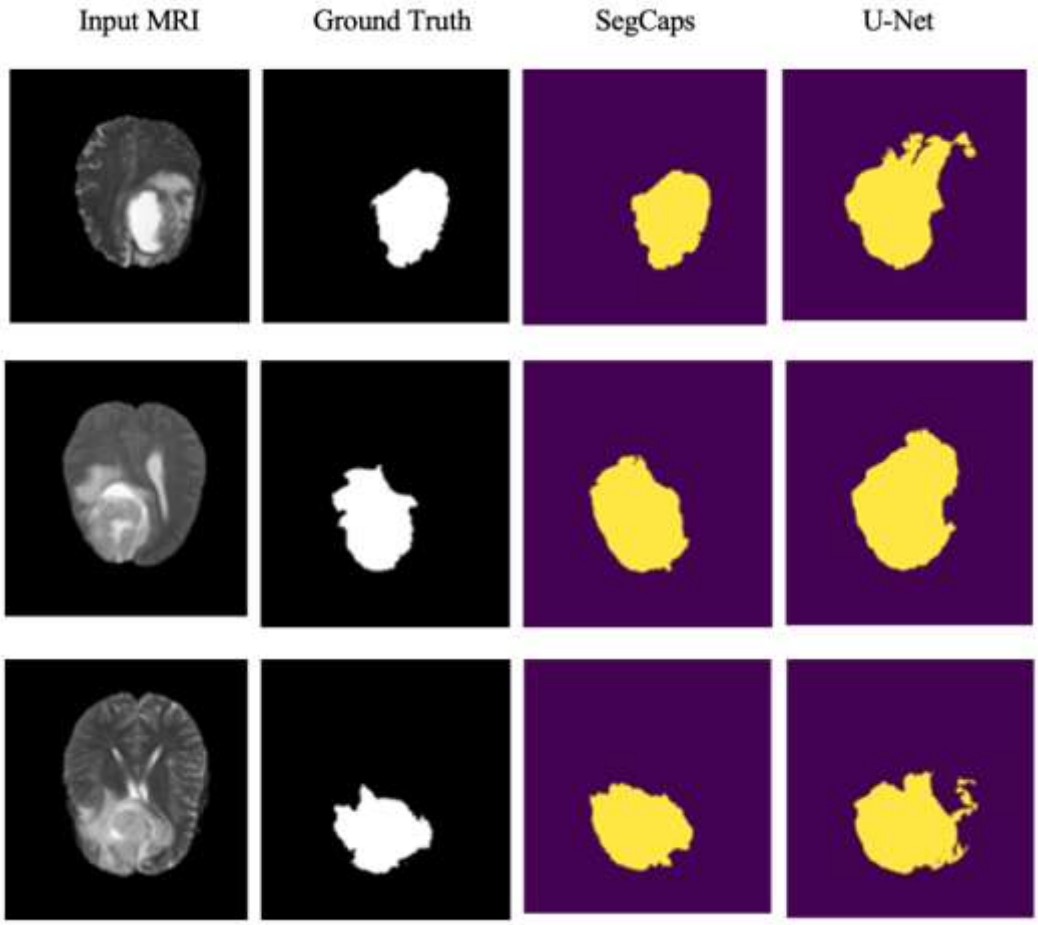

Fig. 5:- Whole tumor segmentation results produced by SegCaps and U-Net. In each row, from left to right, represents the input MRI scan, the ground truth, the SegCaps prediction, and the U-Net prediction.

Table 1 shows the segmentation results obtained from the Dice Coefficient and Jaccard Index. It can be concluded that SegCaps outperforms U-Net, with SegCaps achieving a Dice Coefficient of 0.832255 compared to U-Net's 0.786869. Similarly, SegCaps has a Jaccard Index of 0.740718, which is higher than U-Net's Jaccard Index of 0.689731. In terms of percentage improvement, SegCaps demonstrates a 5.77% improvement in the Dice Coefficient over U-Net. Likewise, SegCaps outperforms U-Net in the Jaccard Index with a 7.39% increase. These percentage improvements underscore the enhanced capability of SegCaps in capturing more accurate tumor boundaries and spatial relationships compared to U-Net, making it a more effective model for MRI brain tumor segmentation. The performance of both models was also evaluated on three different groups of MRI brain images. The purpose of evaluating the models on multiple groups is to account for variability that may arise from various factors such as imaging protocols, equipment used, and patient characteristics. This approach allows for a more comprehensive assessment of the model's segmentation accuracy.

Table 2 shows the performance between SegCaps and U-Net across multiple batches of MRI brain images. SegCaps consistently outperforms U-Net, achieving higher Dice coefficients in all three batches: 0.852 in the first batch, 0.793 in the second, and 0.727 in the third batch. On the other hand, U-Net scores 0.840, 0.727, and 0.662, respectively. The main reason for an improved performance achieved by SegCaps lies in its ability to retain spatial hierarchies within the image via the capsules. Capsule Networks preserve the relationships between features at different levels of the image unlike U-Net that depends on pooling layers and this mechanism faces the risk of losing

crucial spatial information. This is critical in medical image segmentation, where the accurate delineation of tumor boundaries is often dependent on subtle changes in the surrounding tissue.

Another important factor contributing to the enhanced performance of SegCaps is its efficiency in parameter usage. SegCaps uses 95.5% fewer parameters than U-Net, but it achieves better segmentation performance. This not only makes SegCaps computationally more efficient but also reduces the risk of overfitting, as fewer parameters need to be trained. Consequently, the model is more likely to generalize well to unseen data, making it particularly suitable for real-world medical applications where computational resources and time are often limited. Table 1 shows the segmentation results obtained from the Dice Coefficient and Jaccard Index. It can be concluded that SegCaps outperforms U-Net, with SegCaps achieving a Dice Coefficient of 0.832255 compared to U-Net's 0.786869. Similarly, SegCaps has a Jaccard Index of 0.740718, which is higher than U-Net's Jaccard Index of 0.689731.

Table 1:- Comparison of dice coefficients and jaccard index between SegCaps and U-Net.

Model	**SegCaps**	**U-Net**	**% of Improvement**
Dice	0.832255	0.786869	5.77%
Jaccard	0.740718	0.689731	7.39%

Table 2:- Comparison of SegCaps and U-Net for three batches.

Model	**Dice coefficient** First Batch	**Dice coefficient Second Batch**	**Dice coefficient Third Batch**
SegCaps	0.852127	0.792881	0.726944
U-Net	0.839912	0.727087	0.661855
% of Improvement	1.45%	9.05%	9.83%

Conclusion:-

In summary, this paper underscores the effectiveness of SegCaps in accurately segmenting tumor regions, consistently outperforming U-Net. SegCaps achieves superior performance with significantly fewer parameters—95.5% less than the basic U-Net model—highlighting its efficiency in parameter usage. Despite these advantages, the model's high computational cost and memory demands, particularly due to dynamic routing, present challenges when handling larger capsules. This points to the need for further research aimed at reducing both memory consumption and processing time to enhance the model's overall efficiency.

Acknowledgement:-

The authors would like to acknowledge the Ministry of Higher Education Malaysia (MOHE) for the Fundamental Research Grant Scheme (FRGS) Project Code: FRGS/1/2024/TK07/UKM/02/6 and the UniversitiKebangsaan Malaysia.

References:-

1. Wadhwa, A., Bhardwaj, A., & Singh Verma, V. (2019): A review on brain tumor segmentation of MRI images. Magnetic Resonance Imaging, 61: 247–259.
2. Manan, H. A., Yahya, N., Nor Shafiza, A. W., & Yusoff, A. N. (2022): Distribution pattern of brain tumour types and location in patients who underwent MRI scans: a survey of 1240 patients in a Tertiary Malaysian University Hospital.
3. Bahadure, N. B., Ray, A. K., &Thethi, H. P. (2017): Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. International journal of biomedical imaging, 2017(1): 9749108.
4. Voon, N. S., Manan, H. A., & Yahya, N. (2023): Role of resting-state functional MRI in detecting brain functional changes following radiotherapy for head and neck cancer: a systematic review and meta-analysis. Strahlentherapie und Onkologie, 199(8): 706-717.
5. Sun, L., Zhang, S., Chen, H., & Luo, L. (2019): Brain tumor segmentation and survival prediction using multimodal MRI scans with deep learning. Frontiers in neuroscience, 13: 810.
6. Din, N. K. A. M., & Rahni, A. A. A. (2020, March): Evaluation of a deep learning-based brain tumour segmentation method. In Journal of Physics, IOP Publishing, 1497(1): 012009.
7. Liu, Z., Tong, L., Chen, L., Jiang, Z., Zhou, F., Zhang, Q.,& Zhou, H. (2023): Deep learning-based brain tumor segmentation: a survey. Complex & intelligent systems, 9(1): 1001-1026.

8. Chen, S., Ding, C., & Liu, M. (2019): Dual-force convolutional neural networks for accurate brain tumor segmentation. Pattern Recognition, 88: 90-100.
9. Ronneberger, O., Fischer, P., & Brox, T. (2015): U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, Springer International Publishing, part III 18: 234-241.
10. Ali, N. F. B., Mokri, S. S., Abd Halim, S., Zulkarnain, N., Abd Rahni, A. A., &Mustaza, S. M. (2023): Comparison Review on Brain Tumor Classification and Segmentation using Convolutional Neural Network (CNN) and Capsule Network. International Journal of Advanced Computer Science and Applications, 14(4).
11. Iqbal, M. J., Iqbal, M. W., Anwar, M., Khan, M. M., Nazimi, A. J., & Ahmad, M. N. (2022): Brain tumor segmentation in multimodal MRI using U-Net layered structure. Comput Mater Contin, 74(3): 5267-5281.
12. Sabour, S., Frosst, N., & Hinton, G. E. (2017): Dynamic routing between capsules. Advances in neural information processing systems, 30.
13. Hinton, G. E., Sabour, S., &Frosst, N. (2018, February): Matrix capsules with EM routing. In International conference on learning representations.
14. LaLonde, R., Xu, Z., Irmakci, I., Jain, S. & Bagci, U. 2021: Capsules for biomedical image segmentation. Medical Image Analysis 68: 101889.
15. Zhou, T., Ruan, S., & Canu, S. (2019): A review: Deep learning for medical image segmentation using multi-modality fusion. Array, 3: 100004.
16. Zamani, N. S. M., Hoe, E. Y. C., Huddin, A. B., Zaki, W. M. D. W., & Abd Hamid, Z. (2023): Deep learning for an automated image-based stem cell classification. JurnalKejuruteraan, 35(5): 1181-1189.
17. Thanoon, M. A., Zulkifley, M. A., Mohd Zainuri, M. A. A., &Abdani, S. R. (2023): A review of deep learning techniques for lung cancer screening and diagnosis based on CT images. Diagnostics, 13(16): 2617.
18. He, K., Zhang, X., Ren, S., & Sun, J. (2016): Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778.
19. Long, J., Shelhamer, E., & Darrell, T. (2015): Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, 3431-3440.
20. Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., & Van Leemput, K. (2014). The multimodal brain tumor image segmentation benchmark (BRATS). IEEE transactions on medical imaging, 34(10): 1993-2024.
21. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., ... &Jambawalikar, S. R. (2018): Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv preprint arXiv:1811.02629.
22. Li, X., Sun, X., Meng, Y., Liang, J., Wu, F., & Li, J. (2019): Dice loss for data-imbalanced NLP tasks. arXiv preprint arXiv:1911.02855.
23. Havaei, M., Jodoin, P. M., & Larochelle, H. (2014): Efficient interactive brain tumor segmentation as within-brain kNN classification. In 2014 IEEE 22nd international conference on pattern recognition, 556-561.
24. Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D. & Glocker, B. 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical Image Analysis 36: 61–78.
25. Myronenko, A. (2019). 3D MRI brain tumor segmentation using autoencoder regularization. In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II 4 (pp. 311-320). Springer International Publishing.
26. Nazarudin, A. A., Zulkarnain, N., Mokri, S. S., Zaki, W. M. D. W., Hussain, A., Ahmad, M. F., & Nordin, I. N. A. M. (2023): Performance analysis of a novel hybrid segmentation method for polycystic ovarian syndrome monitoring. Diagnostics, 13(4): 750.
27. Mao, A., Mohri, M., & Zhong, Y. (2023). Cross-entropy loss functions: Theoretical analysis and applications. In International conference on Machine learning, 23803-23828.
28. Dong, H., Yang, G., Liu, F., Mo, Y., & Guo, Y. (2017). Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In Medical Image Understanding and Analysis: 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11–13, 2017, Proceedings 21 (pp. 506-517). Springer International Publishing.
29. Kingma, D. P. (2014): Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
30. Speet, S. & Cassani, G. 2020. Brain Tumor Segmentation Using a Capsule Network.