Paper CM10

Using Machine learning for building a demonstrative model to counter the under-diagnosis of dementia

Ashapriya Sebastian, Birmingham City University, Birmingham, UK
Israt Jahan, Birmingham City University, Birmingham, UK
Bahadar Bhatia, Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK
Khalid N. Ismail, Birmingham City University, Birmingham, UK

ABSTRACT

Early and accurate diagnosis of dementia is critical for timely intervention and improved patient outcomes. Neuroimaging techniques like MRI facilitate the identification of structural brain changes associated with dementia pathology. However, manual analysis is limited by inter-rater variability and time intensity. Hence, this work investigates deep learning models for automated MRI segmentation to enable quantitative analysis and aid early dementia detection. Specifically, convolutional neural networks (CNN) like U-Net, SegNet, DeepLab, and Attention U-Net are implemented and evaluated on a brain MRI dataset. The Attention U-Net emerges as the top performer, achieving a Dice coefficient of 0.9262 for segmentation accuracy. Further experiments with Focal and Focal Tversky losses yield coefficients of 0.9297 and 0.9627 respectively, demonstrating the model's adaptability. Comparatively, DeepLab V2 and V3 offer a balance between accuracy (0.84 Dice coefficient) and efficiency (1250ms inference time). Attention U-Net shows immense promise for precise brain region segmentation. The high accuracy is vital for detecting subtle structural changes associated with early dementia.

INTRODUCTION

Dementia is a debilitating group of neurodegenerative disorders affecting millions worldwide. Alzheimer's disease is the most common type, making up 60-70% of cases. Other types include vascular, Lewy body, and frontotemporal dementia. With 55 million affected and 10 million new cases each year, the prevalence of dementia is rising. Early symptoms are memory loss, confusion, and impaired judgment. These can progress to severe challenges in communication, recognition, and self-care. Many dementia sufferers face stigmatization and a lack of support. The impact extends beyond cognitive decline, affecting individuals physically, psychologically, socially, and economically. It poses a significant global economic burden, costing over $1 trillion in 2019 and ranking as the 7th leading cause of death worldwide. As the population ages, dementia's prevalence is set to triple in the next 30 years, necessitating increased funding and support services (World Health Organization, 2023).

Early and accurate diagnosis is vital for effective treatment. Advanced biomarkers and predictive models are being explored in various modalities, including neuroimaging, genetics, cognitive assessments, and speech analysis. Deep learning, particularly convolutional neural networks (CNNs), shows promise in automating the analysis of medical images, achieving high accuracy in MRI-based dementia classification. Multimodal approaches offer additional potential for enhanced accuracy.

These innovations hold the promise of cost-effective screening, early interventions, and better healthcare access for those with dementia. Yet, challenges persist, such as model optimization, reducing bias, and clinical validation before widespread use. The existing attention mechanisms enhance model interpretability for clinicians, showing a path forward in the diagnosis and management of dementia. Recent studies demonstrate the accuracy of CNNs, particularly VoxCNN and U-Net, in Alzheimer's disease identification, offering precise delineation of affected brain structures. This paper investigates deep learning models for automating MRI segmentation through using advanced architectures like SegNet, U-Net, DeepLab and Attention U-Net.

EXPERIMENTAL DESIGN

Earlier and more accurate diagnosis is urgently needed to improve care and treatment planning, especially for conditions like dementia. Manual analysis of MRI scans is limited by inter-rater variability and the time-intensive process of manually tracing brain structures. To address these challenges, automated MRI segmentation using deep learning has emerged as a promising solution.

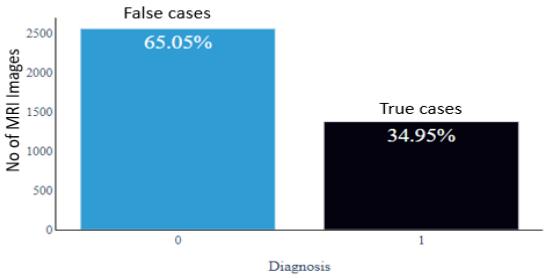

Figure 1: Data Distribution for The Brain MRI Segmentation Dataset.

DATASET

The Brain MRI Segmentation dataset comprises records from 110 patients, with total of 3,929 MRI scans (Buda, 2017). Among these records (scans), approximately 34.95% represent true cases, while 65.05% false cases as shown in Figure 1. This shows there is a need for down-sampling to enhance the class distribution.

DEEP LEARNING MODELS

Several deep learning architectures, including U-Net, SegNet, and Attention U-Net, have shown immense promise for biomedical segmentation tasks. Attention U-Net, for instance, enhances the standard U-Net architecture by integrating attention gates to focus on salient features and suppress irrelevant regions. This attention mechanism improves accuracy, particularly when dealing with limited training data. SegNet, on the other hand, is optimised for pixel-wise segmentation, employing an encoder-decoder structure to map low-resolution feature maps back to the input resolution while preserving boundary details through max pooling indices. Standard U-Net remains a popular choice due to its balanced incorporation of contextual and spatial information, efficient end-to-end learning capabilities, and robustness when trained with data augmentation.

This paper explores the potential benefits of applying Attention U-Net, U-Net, and SegNet to the automated segmentation of MRI scans, specifically to aid in the earlier diagnosis of dementia. Precise segmentation of critical brain structures, such as the hippocampus and ventricles, known to be affected early in dementia, will enable quantitative volume and shape analysis for more accurate diagnosis and tracking of disease progression.

Moreover, a comparative study is conducted, implementing these architectures, and employing various loss functions. This approach will provide valuable insights into performance trade-offs and guide the selection of the most suitable architecture and configuration for the task. The comparative study aims to identify the optimal model that combines MRI segmentation, enabling robust and reproducible identification of mild cognitive impairment, a precursor to dementia. By achieving cost-effective screening, this proposed interpretable approach has the potential to improve healthcare access and stratify patients for clinical trials. The details of these algorithms are described in the following sections.

SEGNET MODEL

SegNet is a deep convolutional neural network designed specifically for semantic pixel-wise segmentation. The core trainable segmentation engine consists of a symmetric encoder-decoder structure followed by a pixel-wise classification layer. There are no fully connected layers and hence it is only convolutional. A decoder upsamples its input using the transferred pool indices from its encoder to produce a sparse feature map(s). It then performs convolution with a trainable filter bank to densify the feature map. The final decoder output feature maps are fed to a soft-max classifier for pixel-wise classification (Badrinarayanan, Kendall and Cipolla, 2017).

A SegNet architecture consisting of both encoders and decoders are implemented. The encoder portion uses a series of Conv2D layers, Batch Normalisation, Rectified Linear Unit (ReLU) activation, and MaxPooling to progressively downsample the input image. On the other hand, the decoder utilises UpSampling2D and Conv2DTranspose layers to upsample the encoded output, generating an image mask of the same dimensions as the input image. To preserve spatial information lost during downsampling, skip connections link the encoder and decoder at each scale. During the training phase, the model utilises a dice loss function, a commonly used metric for segmentation tasks. Data augmentation techniques are applied to enhance the model's robustness and generalisation by introducing variations in the training dataset. The training process is conducted within a k-fold cross-validation framework, ensuring that all available data is effectively utilised.

U-NET MODEL

U-Net is a convolutional neural network designed for efficient and accurate image segmentation. It comprises two main components: a contracting path to capture context and an expanding path for precise localization.

The contracting path involves repeated 3x3 convolution layers with ReLU activation, followed by 2x2 max pooling to downsample and extract abstract features. The number of feature channels doubles with each downsample.

The expanding path uses up-convolution layers to upsample and concatenate high-resolution features from the

contracting path via skip connections. Two additional convolution layers fuse these features for detailed context propagation.
U-Net can process entire images in a single pass, making it efficient for large images. It's trained with minimal annotated data using real-time data augmentation to handle deformations. A weighted loss function prioritizes separating adjacent objects of the same class.
Its unique features include the symmetric structure, skip connections, and end-to-end processing, providing context and spatial information for precise biological image segmentation with minimal training data. It outputs pixelwise classification results directly from input images (Ronneberger, Fischer and Brox, 2015).
Images are resized to 256x256 pixels and normalised. Data augmentation, including flips, rotations, and elastic deformations, is applied to enhance training. Numpy arrays are converted to PyTorch tensors. The U-Net model adopts an encoder-decoder architecture with convolutional blocks for feature extraction and max pooling for downsampling as shown in Figure 2. The decoder uses transposed convolutions to upsample and concatenate encoder features through skip connections. A weighted dice loss function guides training. Achieving a dice score of 0.86 in 25 epochs demonstrates effective learning. On testing, the model yields a mean dice coefficient of 0.82, closely aligning with actual outlines.

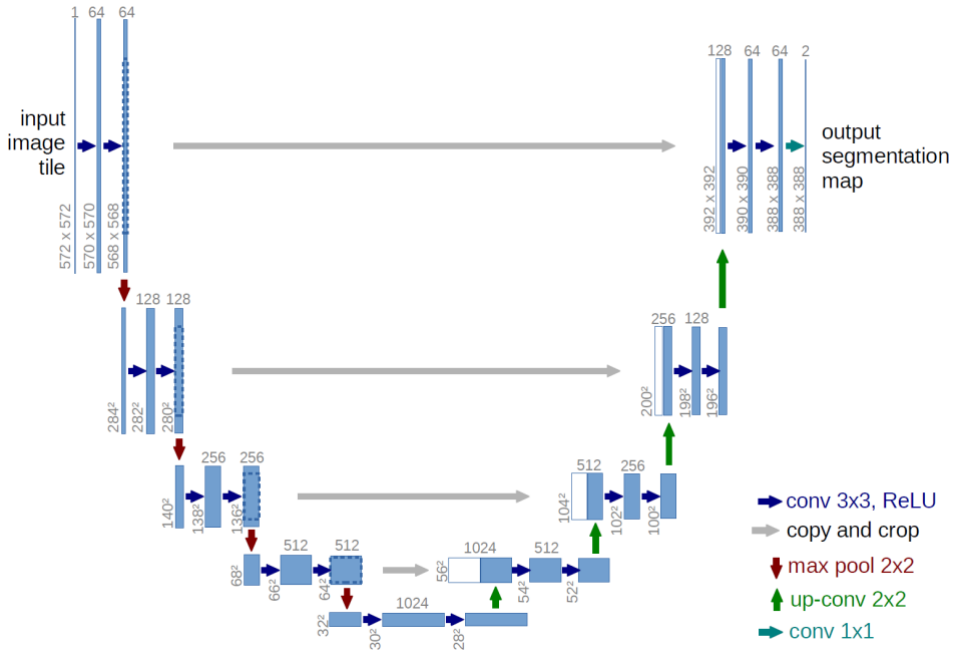

Figure 2: U-Net Model Architecture (Ronneberger, Fischer and Brox, 2015)

DEEPLAB MODELS

DeepLab models employ atrous/dilated convolutions for multi-scale context. Training setups mirror U-Net but show longer training times. DeepLab models achieve superior segmentation performance, with DeepLab v3+ scoring 0.89 in dice coefficient. The DeepLab models integrate atrous/dilated convolutions to capture multi-scale contextual information. **DeepLab v1** applies atrous convolutions with different dilation rates on the final encoder features as shown in Figure 3. This captures features at multiple scales without downsampling. The combined outputs are upsampled to original resolution.

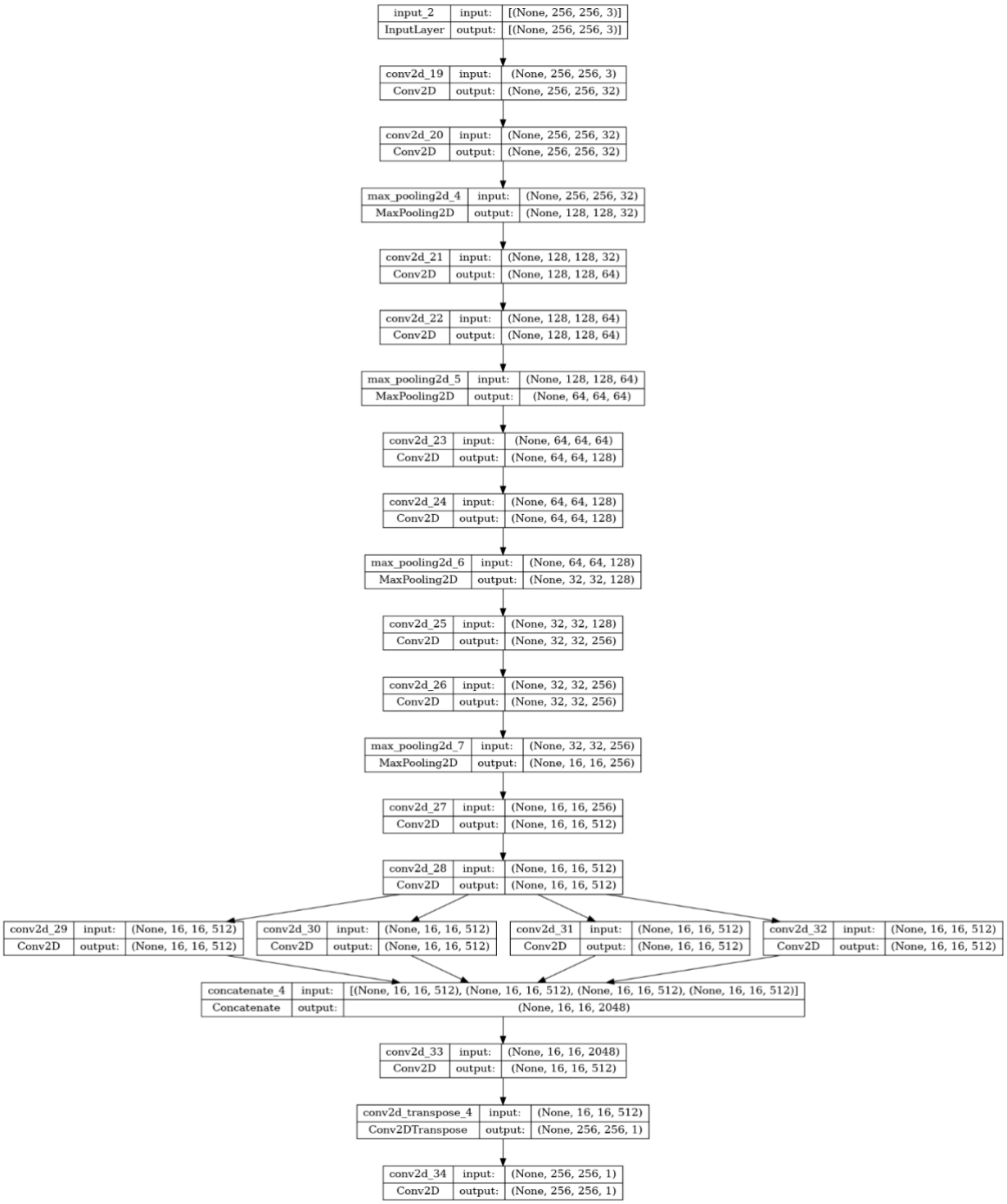

Figure 3: DeepLab v1 Model Architecture

DeepLab v2 simplifies this using atrous spatial pyramid pooling with dilated convolutions at multiple rates in parallel as shown in Figure 4. The features are merged and upsampled.

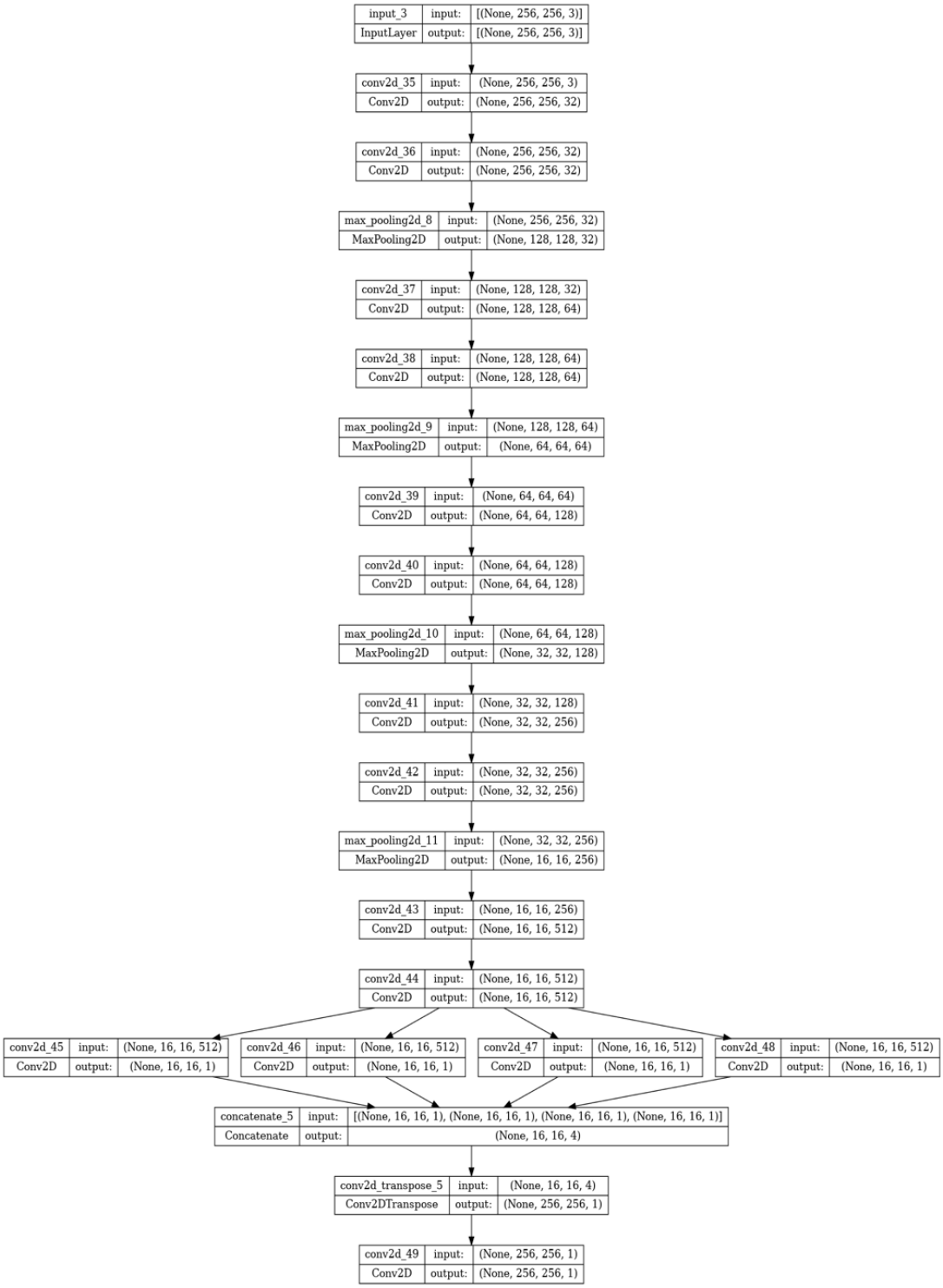

Figure 4: DeepLab v2 Model Architecture

DeepLab v3 incorporates image-level features by averaging and upsampling the final encoder output as shown in Figure 5. This global context supplements the ASPP features.

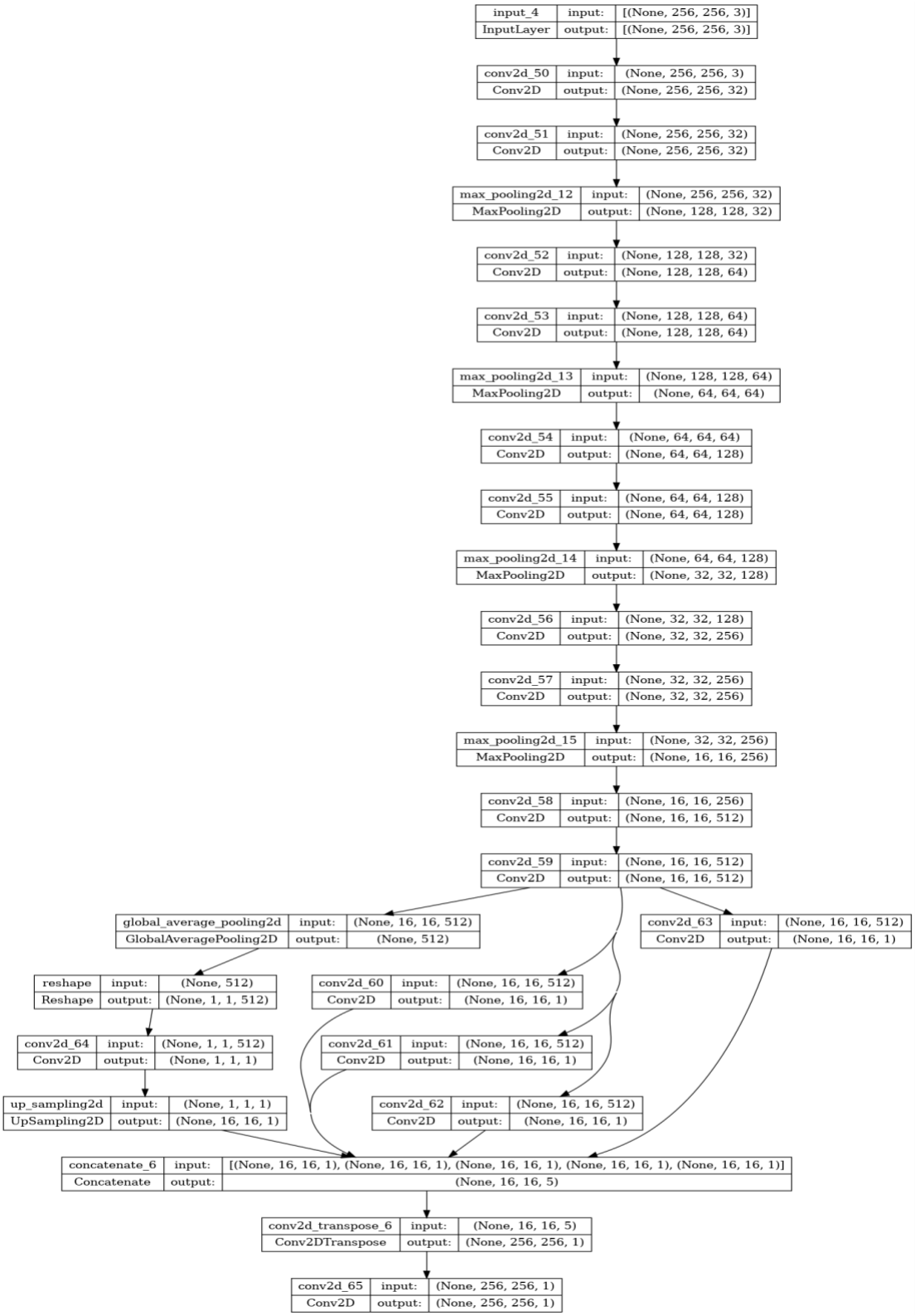

Figure 5: DeepLab v3 Model Architecture

DeepLab v3+ further refines this by using a decoder module. Low-level encoder features are merged with upsampled ASPP output and processed through additional convolutions before final upsampling.

Attention U-Net

The Attention U-Net is a convolutional neural network tailored for precise medical image segmentation with limited training data. It improves the traditional U-net by incorporating attention gates, which enable the network to emphasize crucial features and suppress irrelevant regions in images.
The key innovation of the Attention U-Net is the inclusion of attention gates in the decoder pathway. These gates use a gating signal from the encoder path to create attention coefficients, which filter out background noise and enhance localisation cues with spatial specificity. The model is trained end-to-end with standard techniques, eliminating the need for additional modules. During inference, the attention gates automatically enhance relevant regions, providing precise (Oktay et al., 2018).
In the Attention U-net code, convolutional layers are structured using ConvBlock layers, enhancing code readability and flexibility for the encoder architecture. These ConvBlocks consist of two consecutive 3x3 convolution layers without padding, preserving spatial resolution and enabling the learning of complex hierarchical features.
After each convolution, batch normalization and ReLU activation are applied, stabilizing activations and introducing non-linearity for complex feature learning.

Dice Loss
"The Dice score coefficient (DSC) is a measure of overlap widely used to assess segmentation performance when a gold standard or ground truth is available. The 2-class variant of the Dice loss, denoted DL2, can be expressed as

$$\mathrm{DL2} = 1 - \sum_{n=1}^{N} p_n r_n + \epsilon \Big/ \sum_{n=1}^{N} p_n + r_n + \epsilon - \sum_{n=1}^{N} (1-p_n)(1-r_n) + \epsilon \Big/ \sum_{n=1}^{N} 2 - p_n - r_n + \epsilon$$

The term is used here to ensure the loss function stability by avoiding the numerical issue of dividing by 0, i.e., R and P empty" (Sudre et al., 2017).

Focal Loss
Focal Loss is a specialised loss function for one-stage object detection tasks with severe class imbalance. It is an extension of the cross-entropy (CE) loss. Here is the equation:

$$FL(pt) = -(1 - pt)^{\gamma} * \log(pt)$$

Where:
- pt is a modified probability value.
- γ is a focusing parameter that controls how much easy examples are downweighted.
- FL down-weights easily classified examples and emphasises hard negatives.

An α-balanced variant of FL can be used for better class balance:

$$FL(pt) = -\alpha t (1 - pt)^{\gamma} * \log(pt)$$

This loss function helps improve training on imbalanced datasets by giving more importance to challenging examples.

Focal Tversky Loss
Focal Tversky Loss (FTL) is a specialised loss function designed for medical image segmentation, particularly in scenarios where class imbalance is prevalent, such as identifying lesions in medical images. It addresses the limitations of traditional loss functions like Dice Loss (DL) by focusing on hard-to-segment regions while mitigating issues associated with false positives and false negatives.
The FTL is derived from the Tversky similarity index (TI), which is a measure of the similarity between the predicted and ground truth labels. The TI allows for flexible balancing between false positives (FP) and false negatives (FN). Here's the equation for TI:

$$TIc = \Sigma (pic * gic) / (\Sigma (pic * gic) + \alpha * \Sigma (pic^{-} * gic) + \beta * \Sigma (pic * gic^{-}) + \varepsilon)$$

- pic represents the probability that pixel i belongs to class c.
- gic is the ground truth label for pixel i for class c.
- pic^{-} and gic^{-} are the complementary probabilities and labels for class c^{-} (non-lesion class).
- α and β are hyperparameters that control the emphasis on FP and FN, respectively.
- ε is a small value for numerical stability.

The FTL introduces a focal parameter γ to control the balance between easy background and hard region of interest (ROI) training examples. The FTL equation is as follows:

$$FTLc = \Sigma (1 - TIc)^{(1/\gamma)}$$

When γ > 1, FTL focuses more on less accurate predictions, which have been misclassified. The loss is unaffected if the Tversky index is high, indicating accurate predictions. The value of γ can be tuned; in practice, γ = 4/3 is found to work well. Deep supervision is employed, where intermediate layers are also supervised using FTL, but the last layer uses TI to provide a strong error signal for stable convergence.

RESULTS AND DISCUSSIONS

In SegNet model, Figure 6 shows the training and validation metrics (I.e., accuracy, loss, dice coefficient, IoU) plotted over epoch count for each cross-validation fold. The training and validation curves are close together with small gaps, indicating low variance and that the model generalises well across folds. The validation metrics steadily improve over epochs, demonstrating the model is learning effectively. Moreover, the implementation tracks key segmentation metrics, such as the dice coefficient, Intersection over Union (IoU), and accuracy, on the validation dataset after each training epoch. Model checkpoints are saved whenever there is an improvement in the validation metrics. Furthermore, training curves are generated to analyse the model's convergence over the training epochs. In the final stages, the trained model is loaded, and sample predictions are visualised as shown in Figure 7.

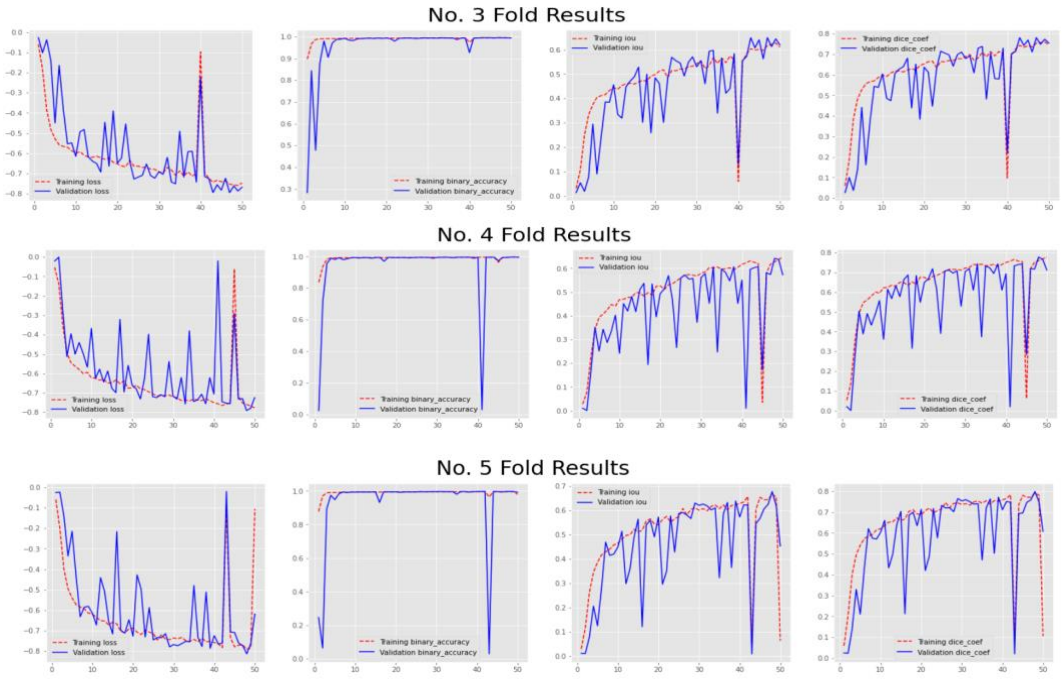

Figure 6: Results obtained for training and validation metrices for the SegNet Model

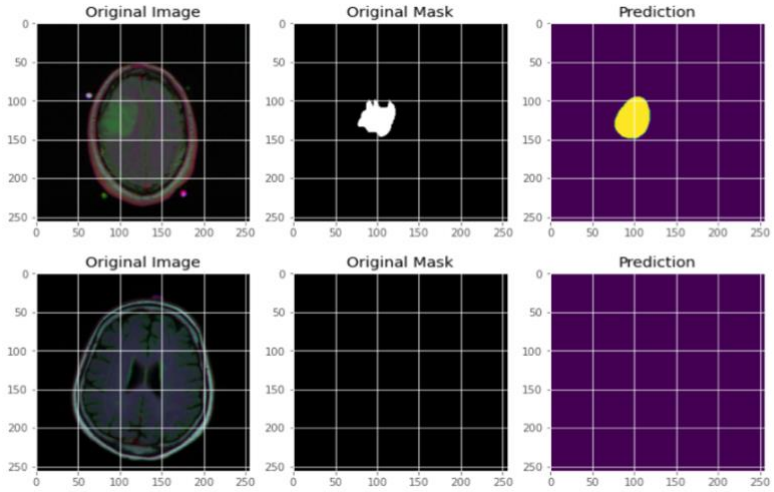

Figure 7: Original image, masked image, and final prediction output for the SegNet Model

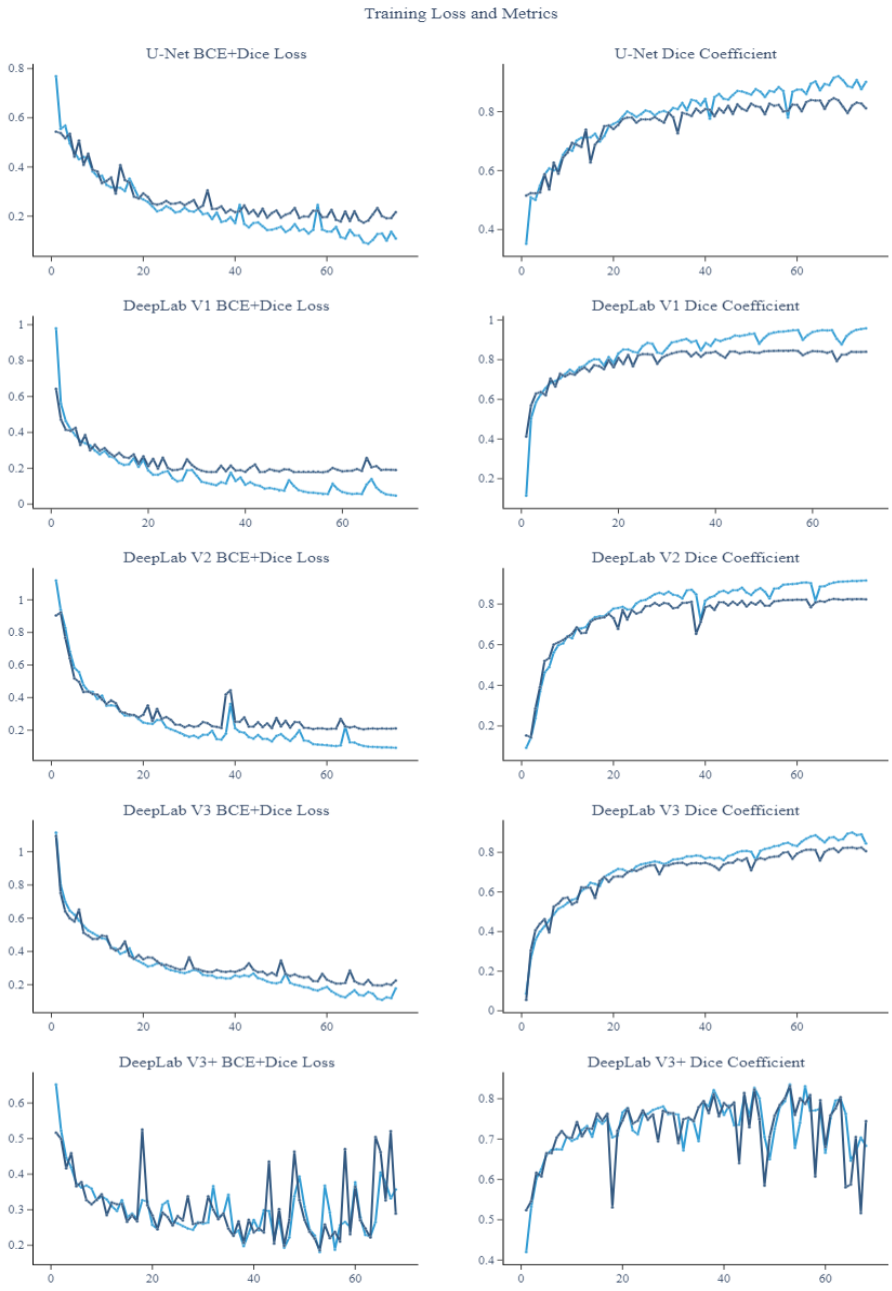

Figure 8: Training Loss and Metrics for U-Net, DeepLab v1, DeepLab v2, DeepLab v3 and DeepLab v3+ Models

The subplot visualisation compares training loss, validation loss, and dice coefficient for 5 semantic segmentation models: UNet, DeepLab V1/V2/V3, and DeepLab V3+ as shown in Figure 8. Metrics are plotted over epochs to showcase convergence during training. DeepLab V3+ achieves the best dice coefficient, followed by DeepLab V3 and V2. UNet and DeepLab V1 have higher loss and lower dice coefficient. In the final stages, the trained model is loaded, and sample predictions are visualised as shown in Figure 9.

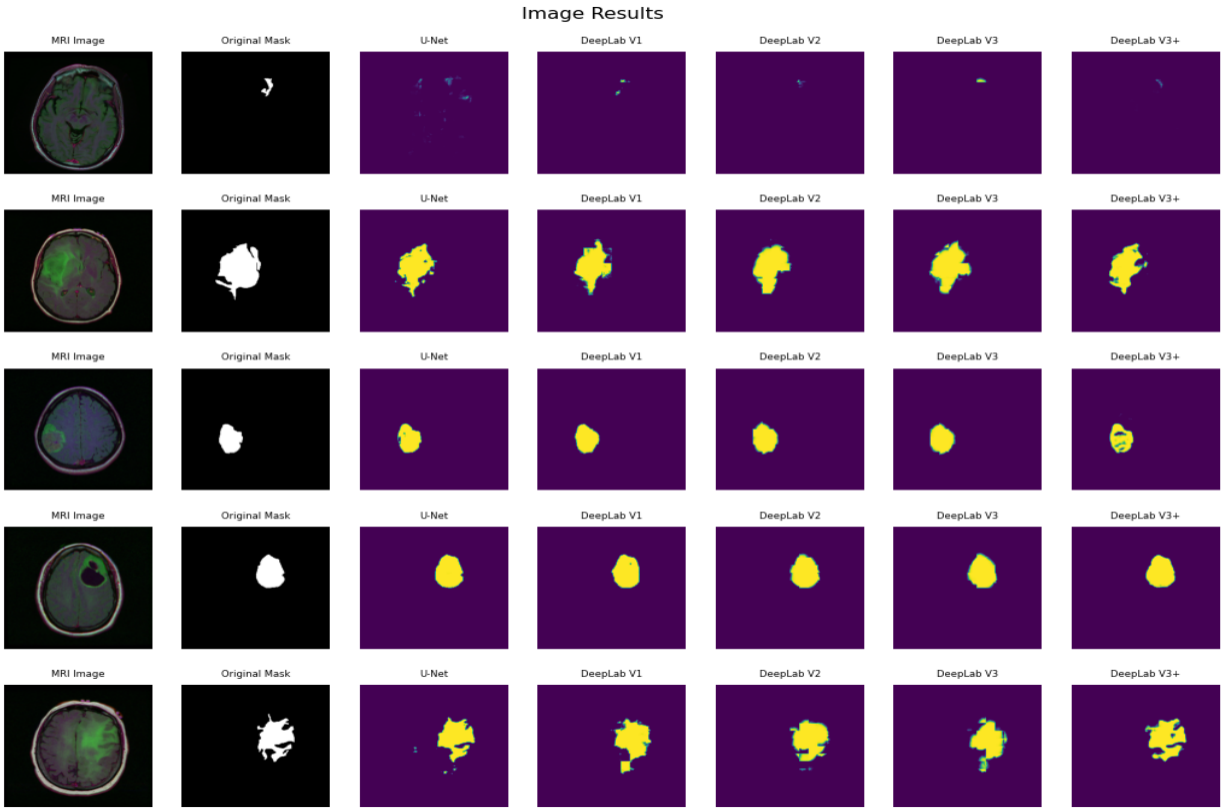

Figure 9: Image results with MR images, original mask, and predicted output of U-Net, DeepLab v1, DeepLab v2, DeepLab v3 and DeepLab v3+ Models

In the attention U-Net model, ConvBlocks stack and downsample feature maps in the encoder pathway utilising doubling output filter channels to capture richer features in deeper layers. In the decoder pathway, ConvBlocks fuse encoder and decoder features via skip connections. The attention gates play a crucial role by selectively propagating salient features, suppressing background noise. Data is pre-processed, and augmentation is applied to ensure robustness. The model employs a weighted dice loss for effective medical image segmentation. Training involves Adam optimisation, batch normalisation, and early stopping, resulting in a significant improvement in the validation metric. Attention visualization shows the model's focus on relevant regions. The model achieves a 0.91 dice score on test data, indicating robust generalisation as shown in Figure 10. This success highlights the benefits of integrating attention mechanisms for precise medical image analysis, overcoming challenges posed by limited training data. The dice loss implemented computes the Sørensen–Dice similarity coefficient between the predicted segmentation mask and the ground truth mask for each batch. During training, the dice score is computed on validation batches to quantify the overlap between model predictions and targets. Tracking this metric over epochs provides validation that the weighted dice loss is effectively optimising the network to improve segmentation accuracy. The final trained model achieves a high average dice score of 91% on unseen test data. This indicates the model has learned robust feature representations that translate to accurately segmenting and delineating organs from the medical images.

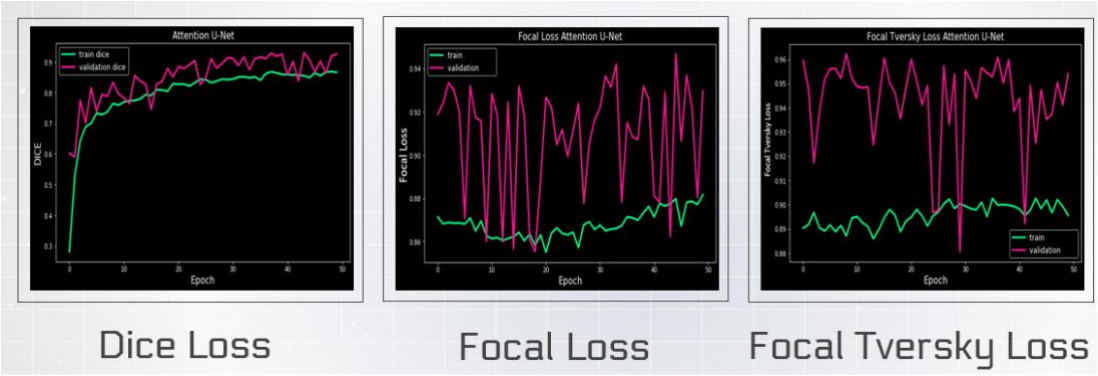

Figure 10: Different Loss Functions used with the attention U-Net.

Focal loss tackles the large class imbalance between foreground and background pixels in medical images by dynamically downweighting the contribution of easy examples during training. This prevents the model from getting overwhelmed by the abundant easy background patches and focuses model capacity on hard or rare examples. With a standard cross-entropy loss, the vast number of easy negative pixels dominate the gradient signal. Focal loss normalises this through an exponential modulation factor, so the model is forced to dedicate attention to hard foreground patches needed for precise segmentation (Lin et al., 2017). The validation dice metric improves from 0.91 to 0.93 by switching the weighted dice loss to focal loss as shown in Figure 10. This shows the benefits of focusing model capacity on hard foreground examples.

In medical image processing, FTL is particularly useful because it addresses class imbalance, where the lesion class is much smaller than the background class. By emphasising hard-to-segment regions and balancing FP and FN detections, FTL improves the accuracy of segmentation models as shown in Figure 10 (Abraham and Khan, 2018).

The computational efficiency and segmentation accuracy using dice coefficient is evaluated as shown in Table 1. The evaluation of the attention U-net models with dice loss, focal loss and focal Tversky loss are evaluated and shown in Table 2. The Dice Coefficient is vital for assessing semantic segmentation quality, especially in early dementia detection where precise brain region segmentation is crucial. Among the models tested, the "Attention U-Net" stood out with a Dice Coefficient of 0.9262, signifying exceptional segmentation accuracy. This accuracy is vital for spotting subtle structural changes in early-stage dementia-related brain images.
"DeepLab V2" and "DeepLab V3" also delivered impressive results with Dice Coefficients of 0.8437 and 0.8466, highlighting the effectiveness of deep convolutional neural networks (CNNs) in accurately segmenting brain regions.
"U-Net" and "DeepLab V3+" struck a balance between accuracy and computational efficiency, achieving Dice Coefficients of 0.7433 and 0.8306, respectively.
In contrast, "DeepLab V1" had the lowest Dice Coefficient at 0.5261, indicating suboptimal performance, and "SegNet," while moderately accurate, lagged top-performing models.
In real-world medical applications, faster inference times are critical. "DeepLab V2" and "DeepLab V3" provided competitive inference times of around 1287.89ms and 1250.20ms, maintaining high accuracy. The "Attention U-Net" excelled in accuracy but had a slightly longer inference time at 3257.4ms.
However, the "SegNet" was the slowest at 4750ms, potentially limiting its use in real-time clinical settings.
Ultimately, the choice of segmentation model depends on the trade-off between accuracy and computational efficiency. The "Attention U-Net" adapts well to different loss functions, achieving remarkable Dice Coefficients, making it valuable for early dementia detection. But its longer inference time may be a concern in real-time settings.

Table 1: Comparison of different models with time, dice coefficient and IoU

Model	Time (ms)	Dice Coefficient	IoU
U-Net	2069.442	0.743	0.579
SegNet	4750.87	0.761	0.629
DeepLab V1	2730.870	0.526	0.312
DeepLab V2	1287.892	0.843	0.751
DeepLab V3	**1250.195**	0.846	0.765
DeepLab V3+	1430.632	0.830	0.608
Attention U-Net	3257.4	**0.926**	**0.88**

Table 2: Comparison of Attention U-Net with Dice loss, Focal loss and Focal Tversky loss

Model	Dice Loss	Focal Loss	Focal Tversky Loss
Attention U-Net	0.926	0.929	**0.962**

CONCLUSION AND FUTURE WORK

This study, is an extensive exploration of deep learning models, prominently featuring the Attention U-Net, for the critical task of brain MRI segmentation in early dementia detection. Notably, the Attention U-Net emerged as a standout performer, achieving a remarkable Dice Coefficient of 0.9262, a testament to its exceptional accuracy in medical image segmentation. This investigation also ventured into the realm of loss functions, revealing the model's versatility by achieving Dice Coefficients of 0.9297 with the Focal loss and a striking 0.9627 with the Focal Tversky loss. This adaptability becomes particularly crucial when detecting subtle structural changes in brain images associated with early-stage dementia.
For the future, a wide array of exciting research avenues beckons. One such path involves the exploration of ensemble models, which combine the Attention U-Net with complementary architectures like DeepLab, with the aim of enhancing model robustness and facilitating deployment across various MRI scanners. Furthermore, a detailed investigation into the interplay between the Attention U-Net and different loss functions is imperative, with the goal of identifying the most optimal combinations tailored to specific segmentation objectives.
Augmenting the Attention U-Net with DenseNet connections presents another intriguing avenue for improvement, as it can enhance information flow and feature reuse, leading to more effective medical image analysis. The growing

relevance of Transformer-based encoder modules in the field of medical imaging also warrants thorough exploration, as they may introduce novel capabilities for feature extraction and segmentation tasks.
Integrating additional layers of complexity by incorporating genetic, cognitive, and demographic data alongside neuroimaging information offers a multidimensional approach to bolster early dementia risk assessment. Additionally, access to larger longitudinal datasets holds immense potential for the development of deep learning systems capable of predicting dementia progression over time, thereby enabling patient stratification and personalised care plans.
Equally vital for clinical adoption is ensuring model interpretability. Advancing techniques like activation mapping and prediction explanation can bolster trust in these models, fostering their acceptance within the clinical community. Finally, rigorous validation of model performance against clinical experts using diverse real-world datasets remains a crucial prerequisite before contemplating full-scale deployment within healthcare systems.
The landscape of deep learning in dementia detection and brain region segmentation offers a multitude of opportunities for advancement. These future directions, spanning ensemble models to loss function optimisation, network architecture enhancements, multimodal data integration, dataset expansion, interpretability refinements, and rigorous clinical validation, collectively hold immense potential to revolutionise computer-aided diagnosis and management of dementia. The remarkable adaptability of the Attention U-Net, as underscored by its performance with different loss functions, reaffirms its significance in the context of early dementia detection and precise brain region segmentation, particularly when dealing with subtle structural changes in brain images that demand the utmost precision.

REFERENCES

Abraham, N. and Khan, N.M. (2018) 'A Novel Focal Tversky loss function with improved Attention U-Net for lesion segmentation', *Proceedings - International Symposium on Biomedical Imaging*, 2019-April, pp. 683–687. Available at: https://doi.org/10.1109/ISBI.2019.8759329.
Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) 'SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), pp. 2481–2495. Available at: https://doi.org/10.1109/TPAMI.2016.2644615.
Lin, T.Y., Goyal, P., Girshick, R., He, K. and Dollar, P. (2017) 'Focal Loss for Dense Object Detection', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), pp. 318–327. Available at: https://doi.org/10.1109/TPAMI.2018.2858826.
Oktay, O., Schlemper, J., Folgoc, L. Le, Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B. and Rueckert, D. (2018) 'Attention U-Net: Learning Where to Look for the Pancreas'. Available at: https://arxiv.org/abs/1804.03999v3 (Accessed: 20 June 2023).
Ronneberger, O., Fischer, P. and Brox, T. (2015) 'U-Net: Convolutional Networks for Biomedical Image Segmentation', *IEEE Access*, 9, pp. 16591–16603. Available at: https://doi.org/10.1109/ACCESS.2021.3053408.
Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S. and Cardoso, M.J. (2017) 'Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10553 LNCS, pp. 240–248. Available at: https://doi.org/10.1007/978-3-319-67558-9_28.
World Health Organization (2023) *Dementia*. Available at: https://www.who.int/news-room/fact-sheets/detail/dementia (Accessed: 27 August 2023).

RECOMMENDED READING

Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) 'SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation', IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), pp. 2481–2495. Available at: https://doi.org/10.1109/TPAMI.2016.2644615 .
Oktay, O., Schlemper, J., Folgoc, L. Le, Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B. and Rueckert, D. (2018) 'Attention U-Net: Learning Where to Look for the Pancreas'. Available at: https://arxiv.org/abs/1804.03999v3 .
Ronneberger, O., Fischer, P. and Brox, T. (2015) 'U-Net: Convolutional Networks for Biomedical Image Segmentation', IEEE Access, 9, pp. 16591–16603. Available at: https://doi.org/10.1109/ACCESS.2021.3053408 .

CONTACT INFORMATION

Dr. Khalid Ismail
Birmingham City University
15 Bartholomew Row, Birmingham
B5 5JU
Email: khalid.ismail@bcu.ac.uk