Early Detection of Alzheimer's Disease by Segmenting Hippocampus from MRI of Human Brain Using Deep Learning

D. K. Ramkumar[1*], Dr. N. V. Balaji[2]

Abstract: To investigate a number of neurodegenerative disorders, including Alzheimer's disease, an automatic measurement of hippocampal volume extraction is essential. It is particularly important to examine the features of the hippocampus subfields since they can reveal earlier disease proliferation in the human brain. Due to their complicated structural structure and the requirement for manually labelled high-resolution magnetic resonance images (MRI), segmentation of these subfields is extremely challenging. In the presented paper, we introduce a thoroughly supervised convolutional neural network-based model called VNet for autonomous hippocampal subfield segmentation. The experiments carried out on the challenging data set and their qualitative and quantitative results are reported here. The method has provided improved results in the measures of accuracy and dice score when compared to other cutting-edge methods.

Keywords: Convolution Neural Network, hippocampus, Artificial intelligence, segmentation, brain diseases

1. Introduction

The hippocampus exists in the temporal lobe of the human brain and resembles the structure of a seahorse. Hippocampus is a Greek word that means "great sea monster." It is the name of this brain structure. It is one of the grey matter structures involved in the process of regulating emotions and memory formation and is a transition point between long term and permanent memory. The reduction of its size is known as Alzheimer's disease, and it causes some memory loss. Hence, volume detection in the hippocampus is becoming essential.

Magnetic Resonance Images (MRI) are frequently used to diagnose the disease due to their good contrast. Early detection of changes appearing in the hippocampus is required to treat the diseases [1]. Manual segmentation and atlas-based methods are time-consuming. In the current era, Artificial Intelligence (AI) -enabled machine learning methods play a vital role, especially in medical image analysis.

Convolutional Neural Networks (CNN) are a significant method for image processing because they can extract features from images themselves and train the model with several filters and activation functions. In recent days, CNN has come in different dimensions, such as ResNet, Unet, Deepnat, and Quicknat. Some methods involved in brain tumour grade detection [2], classification of tumours and tissues [3], and bleeding in cerebral structures [4] The stacked layers in CNN are called UNet and VNet.

The hippocampus has substructures like the head, body, and tail. In the earlier days of detection of hippocampus substructure utilizvivo, ultra-high resolution and multimodality MRI images such as T-weighted and T2-weighted [5, 6] were used. Iglesias et al., [5] utilised the parametric method, and Yushkevch et al., [6] utilised the non-parametric method. After the emergence of deep learning, several researchers have been involved in the task of implementing deep learning in hippocampal segmentation and getting good results [7, 8].

Yang et al., [9] analyzed hippocampus segmentation using the CNN model. They provided a toolbox called CAST to segment the subfields of the hippocampus. They are using CNN with residual learning and residual connections. The entire model uses nearly a hundred layers. It produces high dice scores and correlation results. Chen et al., [10] combine triplane voxels as patches that provide six orientations of voxel data. Finally, I got nine plane voxels and used nine updated U Nets, which are based on CNN. Tripler patches are also used by Xie and Gillies [11], Watchinger et al., [12] used a CNN-based model called DeepNat to segment 25 structures of the human brain. They used two 3D CNN models. One classifies foreground and background voxels, and another 3DCNN segments 25 structures from the foreground voxels. They have the MALC dataset and achieved a 0,86 dice score in hippocampus region extraction alone. Roy et al., [13] used QuickNat, which is the base of the CNN model. It achieves a good result like DeepNat and performs faster than Deep Nat.

[1*]*Research Scholar, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, India.*
[1*]*Email: ramd.kannan@gmail.com*
[2]*Professor, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, India.*
[2]*Email: balaji.nv@kahedu.edu.in*
Corresponding author: ramd.kannan@gmail.com

Based on U net, but with a different batch size, is nnUnet. This network automatically adjusts its batch size to the resolution and plane. The batch size is changed for 256 × 256 two-dimensional images to 42, and thirty feature maps are taken in the top layer. 30 feature maps at the highest quality layers are taken for 3D images. For various MRI image segmentations of human organs, the batch size and feature maps differ [14]. Multi-planar data augmentation was used by [15]. Utilizing the augmentation method, 2D CNN is used. The network may be able to learn 3D volumes as a result. In those days, [16] presented a semi-supervised technique. In the model, uncertainty aware multiview co-training (UMCT) was added to the method. It uses some unlabeled input using the Bayesian learning method and is not a fully supervised network. It gives a confidence score for each piece of unlabeled data that has labels.

The remaining part of this paper describes the dataset, architecture of the proposed model, results, discussion, and conclusion.

2. Dataset:

This work has taken data from a publicly available dataset called 'Medical Segmentation Decathlon Challenge'. The dataset contains 3D, T1-weighted images of ninety healthy and 105 non-healthy volumes. All brain volumes have human-annotated hippocampus substructures like the head, body, and tail. The work focus is in sagittal view, which has a 1.0 mm^3 voxel size and was taken using a Philips Achieva scanner.
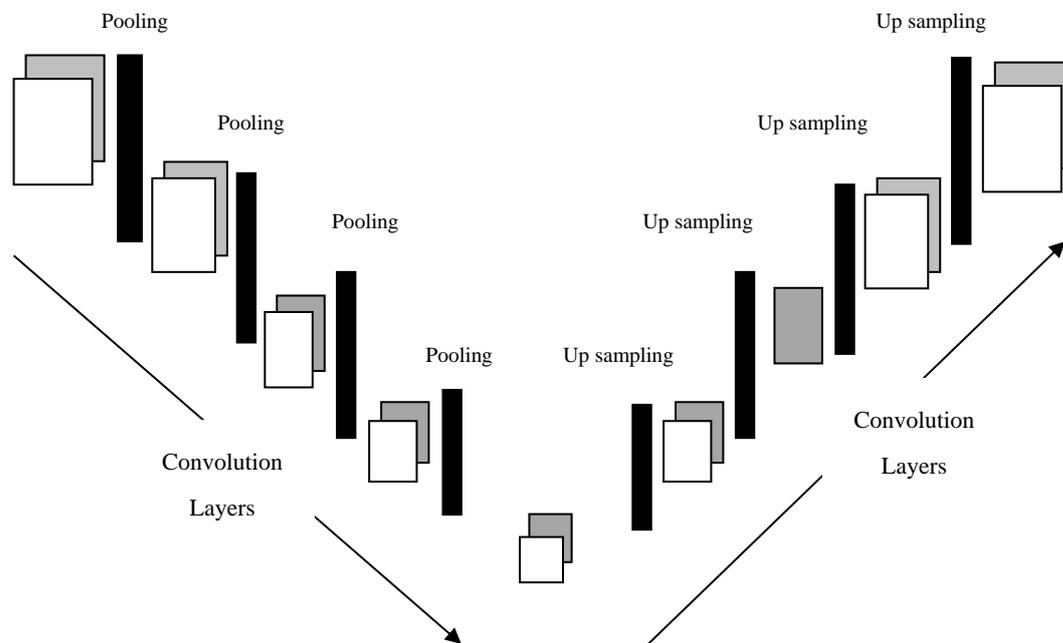

Fig 1: Architecture of VNet with the base of CNN

3. Architecture of Vnet

The architecture of VNet is shown in Figure 1. The network contains both sides, left and right, which are described in the following sections.

3.1 Left Side Process of VNet:

As the architecture of this model resembles the letter 'V', so it is called as 'V-Net'. It has a symmetrical shape in which the left side almost has the same structure as the right side. Though the left side reduces the size of the image, the right side expands the image and produces the original size. Each side of the model consists of different stages and works at various resolutions of the image. Each stage has a maximum of three convolutional layers. On the left side of the architecture, the output of each stage of the convolutional layer is used as the input for the next stage of the convolutional layer. The layers process the image data non-linearly and add it to the output of the previous layer. The learning at each stage is performed by utilising the residual function, which compares the predicted value with the target value.

For the convolution process at each stage, the model uses 5 × 5 × 5 kernels. This model analyses the features by reducing the voxel size by half, which will double the features. As part of the pooling process, the model uses 2 × 2 × 2 convolution kernels to reduce the size of the resulting features. The feature channels are doubled at the end of each stage. In the follow-up of the process, down sampling takes place.

3.2. Activation Function

This model used a generalized activation function, Parametric Rectified Linear Unit (PReLU) which is the optimized version of ReLU. Since the first layer of CNN consists of some frequency based filters like edge detectors that provide positive and negative values, to optimize the values, this function provides some penalty for negative values instead of zero.

$$f(y_i) = \begin{cases} y_i, & if\ y_i > 0 \\ a_i y_i, & if\ y_i \leq 0 \end{cases} \quad (1)$$

where y_i represents input and a_i represents negative slope . The equation (1) can be written as follows,

$$f(y_i) = max(0, y_i) + a_i min(0, y_i) \quad (2)$$

3.3 Right Side Process of VNet:

VNet right side involves in expanding the feature maps to import the necessary information from the image. This process aims to provide two classes of volumetric segmentation results. The three layers of the right side consist of $5 \times 5 \times 5$ kernels for the deconvolution process. With the help of the residual function, the model reconstructs the image from two classes of data. The final convolution layer holds $1 \times 1 \times 1$ kernels to provide the size of the image as an input image.

At the final stage, the output image is like a binary image: the foreground voxels represent the hippocampus, and the remaining voxels represent the background brain portion.

3.4 Training

The training process carried over to the challenge dataset, which contains 3D volumes of 95 normal brain images. The Adam optimizer is used for this model since this optimizer utilizes very little storage. The learning rate that is alpha value and batch size is set to 0.01 and 16 respectively.

Three labels for ground truth were selected: background (0), anterior (1), and posterior (2). The training, validation, and testing percentages of 95 volumes are 60%, 10%, and 20%, respectively.

4. Results and Discussion

4.1 Quantitative Measures

To analyze the performance of the presented method, Dice similarity and accuracy are used. They are defined as,

$$Dice = \frac{2TP}{2TP+FP+FN} \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

The abbreviations are defined as

TP - True Positive

TN - True Negative

FP - False Positive and

FN - False Negative

The definitions of the terms are depicted in Table 1.

Table 1: TP, TN, FP and FN measures

		Actual	
		Positive	Negative
Predicted	Positive	TP	FP
	Negative	FN	TN

4.2 Discussion

The algorithms are coded using Python in Google colab IDE. 56 data volumes were utilized for training, 9 volumes were used as validation data and the remaining were used as testing data volumes. Initially, 30 layers were employed to train the network. The filter size was fixed to $3 \times 3 \times 3$. ReLU activation function with alpha value as 0.3 was used. Though, it produces only 0.87 mean Dice score. Further, the number of layers was increased up to 100, the activation function was changed to PReLU, the optimizer is set to

Adam with alpha value 0.01. Thus results mean Dice value as 0.88. Further convolution filters are increased higher than the filters in the previous layer. At the final fully connect layer the convolution filter size is modified into $1 \times 1 \times 1$.

The visual results of two sample images are showed in Figure 2. The column 1 shows the ground truth images column 2 and column 3 show the results of VNet and corrected VNet respectively.
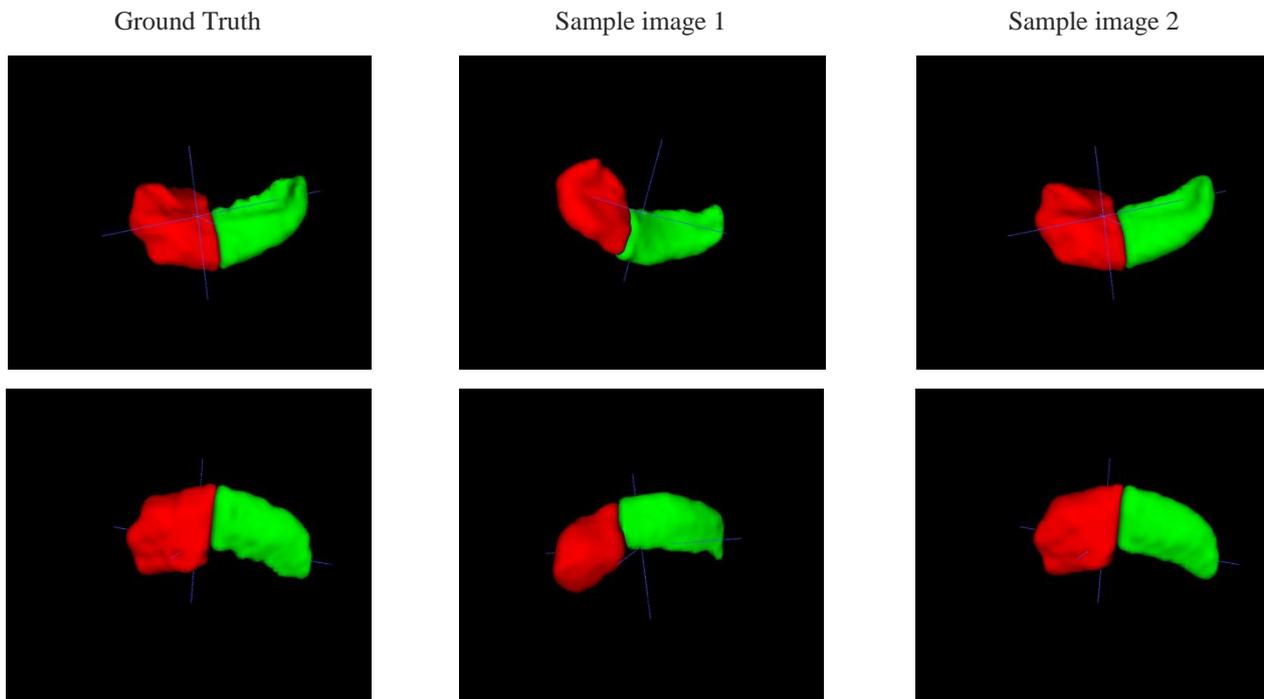
| Ground Truth | Sample image 1 | Sample image 2 |

Fig 2: Segmented hippocampus regions The column 1 shows the ground truth images, column 2 and column 3 show the results VNet and proposed VNet respectively. Row 1 depicts the sample image 1 and row 2 show the sample image 2.

56 data volumes are used as a training data set, 9 volumes are used as a validation data set, and the remaining 19 volumes are taken as testing images. Some sample cases which are visually good, average and bad are tested initially to validate the performance of the proposed work. They are showed in Figure 3(a-c).

The Figure 3a depicts the average Dice scores of the cases in segmenting Anterior portion of the hippocampus. The best cases show 0.89 and 0.85 dice scores with ±0.01 and ±0.04 standard deviation. The worst case shows 0.50 dice score with 0.31 standard deviation which reveals that the proposed method segments perfectly in some images and fails in others.

The Figure 3b shows the average Dice scores of the sathe approach. Uncertainty aware multiview co-training (UMCT) was added to the method. It uses some unlabeled input using the Bayesian learning method and is not a fully supervised network. It offers a confidence score for each piece of unlabeled data that receives labels. In the beginning, a few sample cases with good, average, and poor visual quality are checked to confirm the efficacy of the suggested model. Figure 3 (a-c) illustrates them.

Figure 3a displays the average Dice scores for each case when segmenting the anterior region of the hippocampus. Dice scores of 0.89 and 0.85, with standard deviations of 0.01 and 0.04 respectively, indicate the best circumstances. In the worst case, the dice score is 0.50 with a standard deviation of 0.31, showing that the suggested method only successfully segments certain slices while failing to do so in others. In the segmentation of posterior hippocampus, Case 3 displays a noteworthy outcome because the anterior portion's segmentation is 0.89. Since the suggested method's standard deviation is 0.35 in Case 4, it accurately detects the voxels in some slices. In the other two instances, the suggested technique offers comparable results in the segmentation of the anterior region; the corresponding standard deviations are 0.29 and 0.31.

The average results for each volume are displayed in Figure 3c. A dice result of more than 0.7 is noteworthy. As a result, we make sure that this model can effectively segregate the hippocampus into its several sub-regions. By offering more training tests and increasing the number of epochs, the model gets improved further.
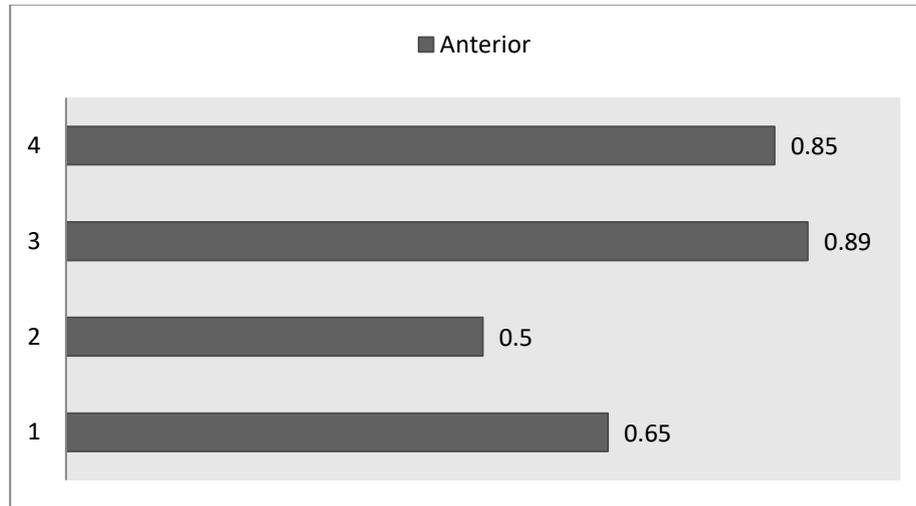
(a)
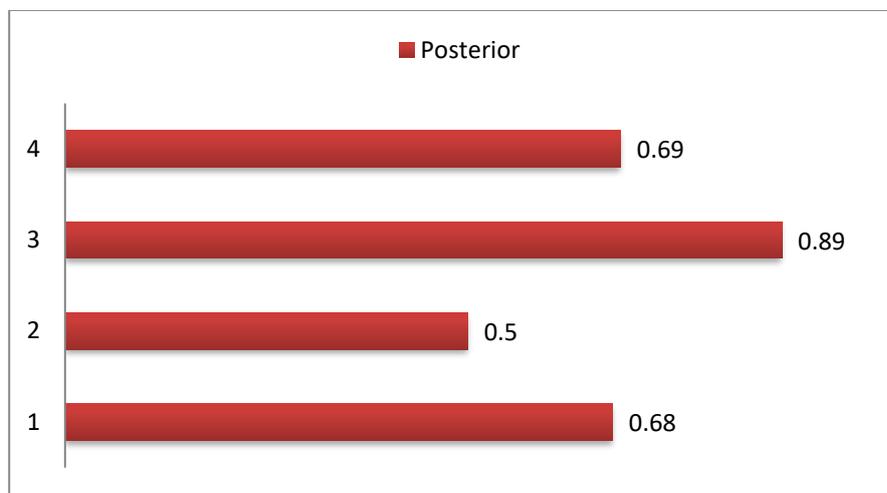
(b)

(c)

Fig 3: Dice scores of some sample cases in segmenting anterior, posterior and both regions.

Table 2 contains a list of the obtained dice scores (mean). The table displays the outcomes of the proposed method applied to the challenging data set and compares them with some deep learning models applied to the same dataset. In Table 2, row 1 displays the outcomes of the suggested approach, row 2 displays those of 2D Unet, row 3 displays those of Nicola Altini, whose code was made available on Github, and row 4 displays those of Ying da Xia et al. (2020). A related model, 2D Unet, with less architectural change, delivers 0.6% less results than the suggested approach. Results from Ying da Xia et al.'s semi-supervised technique were 0.9% inferior to those of the suggested model.

Table 2: Dice measures comparison over similar deep learning models

Method	Mean Dice	Anterior	Posterior
Results of the proposed method over challenging data set	**0.8815**	**0.8935**	**0.8695**
2D Unet	0.8761	0.8852	0.8670
Results of Nicola Altini @ github over challenging data set	0.8727	0.8821	0.8634
Xia et al.[16] (2020)	0.8734	0.8797	0.8671

Table 3 reports the results of the presented model's comparison with a few Deep Learning models that were presented at the Decathlon Challenge 2021. [15] used certain augmentation approaches in the presented models to precisely identify the voxels of the hippocampus and its sub-regions, outperforming the suggested model by less than 0.4%. When compared to [17] the suggested model achieves results that are comparable for both overall and anterior section segmentation tasks.

Table 3: Dice measure comparison over the Decathlon challenge 2021 results

Method	Mean Dice	Anterior	Posterior
Perslev et al., [15]	0.8899	0.8968	0.8831
Yu et al., [17]	0.8866	0.8937	0.8796
Results of the proposed method over challenging data set	0.8815	0.8935	0.8695

The results of the proposed technique applied to the non-challenging data set, which was maintained by us with the gathering of data from nearby MRI centers, are shown in Figure 4. The suggested model can generate significant results that are greater than 0.94 dice with the non-challenging dataset, according to the visual examination of Figure 4. In that as well, the model does better at segmenting the anterior hippocampus.
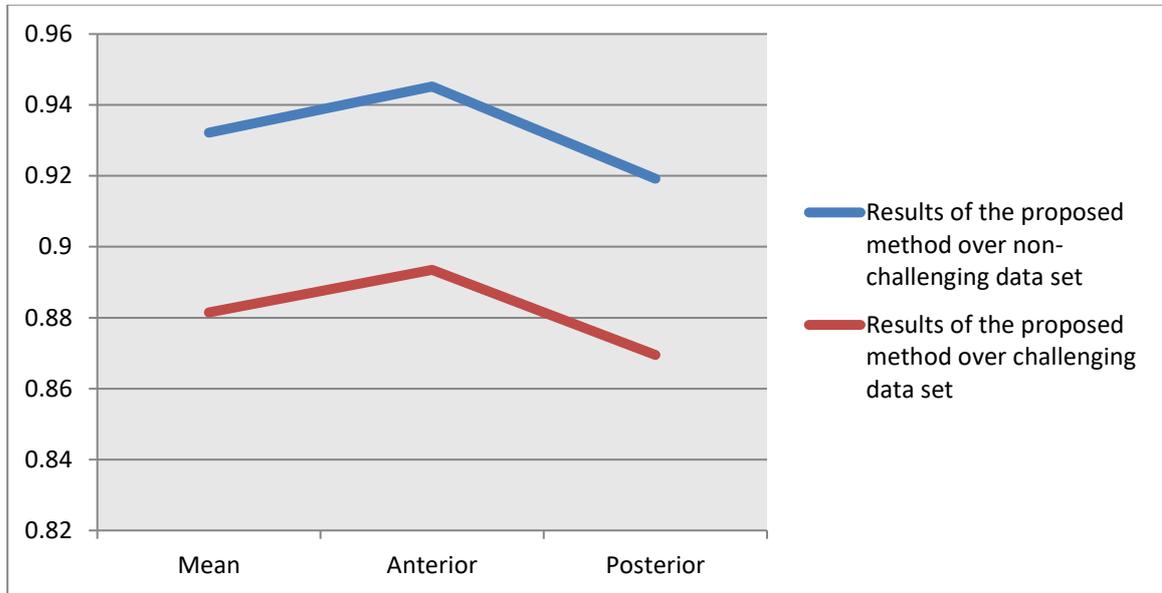
Fig 4: Results of proposed model employed over challenging and non-challenging dataset

According to the results discussed in this section, the updated activation function model produces good results compared to the existing model. In some volumes, the model can extract the hippocampus regions efficiently, but in some models, it can't extract the tail part significantly. For five volumes of challenging data sets, the proposed method suited perfectly and produced more than 95% of the dice score for both regions extraction. The method produces a 90% dice score for 10 volumes and a below-90% dice score for four volumes. Though, the proposed method is working perfectly for the non-challenging dataset.
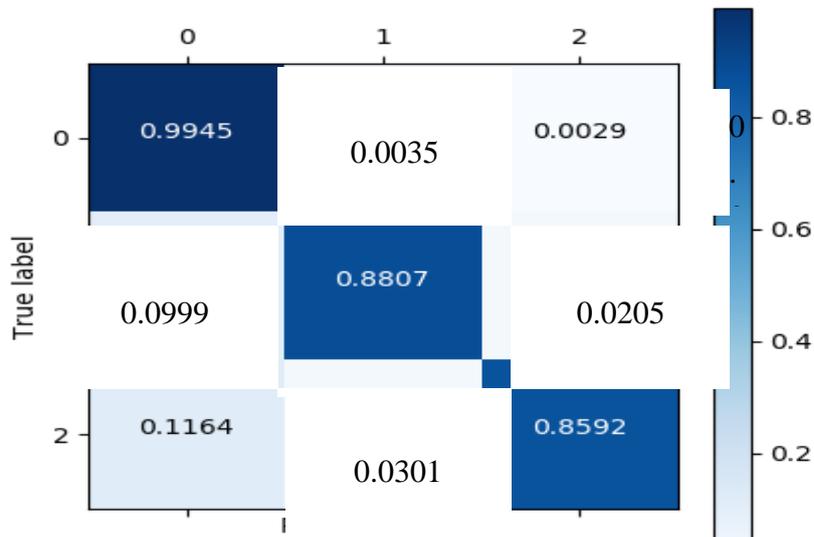

Fig 5: Confusion matrix

The confusion matrix is given in Figure 5, which shows the predicted voxels against the three labels. As mentioned above, label 0 represents the background voxels, label 1 represents the anterior voxels, and label 2 represents the posterior voxels. The total number of voxels is suppressed to a 0–1 value. As per the representation of the confusion matrix, the error is minimal in the detection of background and anterior region voxels. The overall accuracy of the presented method is 0.989, and the rate of misclassification is 0.0221.

As per the above discussion, the proposed method is supportable for the diagnosis of hippocampus related diseases.

5. Conclusion

In this research work, we propose a deep-learning algorithm called VNet, which is the base of CNN, to segment the hippocampus and its sub-regions. This paper investigates the layers and suitable activation functions.

This deep learning based model simplifies the segmentation process, thus reducing the need for lots of manually labelled data. The presented method has been evaluated using two kinds of datasets: publicly available datasets and locally maintained datasets. The model produces 98% accuracy, which ensures its overall performance. It is easier for the physician to diagnose hippocampus related diseases.

Reference

[1] Y. Mu, and F.H. Gage, 2011. Adult hippocampal neurogenesis and its role in Alzheimer's disease. *Molecular neurodegeneration*, *6*(1), pp.1-9.

[2] Y. Fan W. Huang Z. Lin W. Zhu J. Zhou, and J. Wong, 2015. Brain tumor grading based on neural networks and convolutional neural network, 37th *International Conference of IEEE Eng. in Medicine & Biology Society*, pp. 699-702.

[3] L. Zhao, and K. Jia, 2016. Multiscale CNNs for brain tumor segmentation and diagnosis. *Computational and mathematical methods in medicine*, *2016*.

[4] Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, V.C. Mok, L. Shi, and P.A. Heng, 2016. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. *IEEE transactions on medical imaging*, *35*(5), pp.1182-1195.

[5] J.E. Iglesias, J.C. Augustinack, K. Nguyen, C.M. Player, A. Player, M. Wright, N. Roy, M.P. Frosch, A.C. McKee, L.L. Wald, and B. Fischl, 2015. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. *Neuroimage*, *115*, pp.117-137.

[6] P.A. Yushkevich, J.B. Pluta, H. Wang, L. Xie, S.L. Ding, E.C. Gertje, L. Mancuso, D. Kliot, S.R. Das, and D.A. Wolk, 2015. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. *Human brain mapping*, *36*(1), pp.258-287.

[7] M. Goubran, E.E. Ntiri, H. Akhavein, M. Holmes, S. Nestor, J. Ramirez, S. Adamo, M. Ozzoude, C. Scott, F. Gao, and A. Martel, 2020. *Hippocampal segmentation for brains with extensive atrophy using three-dimensional convolutional neural networks* (Vol. 41, No. 2, pp. 291-308). Hoboken, USA: John Wiley & Sons, Inc..

[8] M. Liu, F. Li, H. Yan, K. Wang, Y. Ma, L. Shen, M. Xu, and Alzheimer's Disease Neuroimaging Initiative, 2020. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. *Neuroimage*, *208*, p.116459.

[9] Z. Yang, X. Zhuang, V. Mishra, K. Sreenivasan, and D. Cordes, 2020. CAST: A multi-scale convolutional neural network based automated hippocampal subfield segmentation toolbox. *NeuroImage*, *218*, p.116947.

[10] Y. Chen, B. Shi, Z. Wang, P. Zhang, C.D. Smith,. and J. Liu, 2017, April. Hippocampus segmentation through multi-view ensemble ConvNets. In *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)* (pp. 192-196). IEEE.

[11] Z. Xie, and D. Gillies, 2018. Near real-time hippocampus segmentation using patch-based canonical neural network. *arXiv preprint arXiv:1807.05482*.

[12] C. Wachinger, M. Reuter, and T. Klein, 2018. DeepNAT: Deep convolutional neural network for segmenting neuroanatomy. *NeuroImage*, *170*, pp.434-445.

[13] A.G. Roy, S. Conjeti, N. Navab, C. Wachinger, and Alzheimer's Disease Neuroimaging Initiative, 2019. QuickNAT: A fully convolutional network for quick and accurate segmentation of neuroanatomy. *NeuroImage*, *186*, pp.713-727.

[14] F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P.F. Jaeger, S. Kohl, J. Wasserthal, G. Koehler, T. Norajitra, S.Wirkert, and K.H. Maier-Hein, 2018. nnu-net: Self-adapting framework for u-net-based medical image segmentation. *arXiv preprint arXiv:1809.10486*.

[15] M. Perslev, E.B. Dam, A. Pai, and C. Igel, 2019. One network to segment them all: A general, lightweight system for accurate 3d medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22* (pp. 30-38). Springer International Publishing.

[16] Y. Xia, D. Yang, Z. Yu, F. Liu, J. Cai, L. Yu, Z. Zhu, D. Xu, A. Yuille, and H. Roth, 2020. Uncertainty-aware multi-view co-training for semi-supervised medical image segmentation and domain adaptation. *Medical image analysis*, *65*, p.101766.

[17] Q. Yu, D. Yang, H. Roth, Y. Bai, Y. Zhang, A.L. Yuille, and D. Xu, 2020. C2fnas: Coarse-to-fine neural architecture search for 3d medical image segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4126-4135).

[18] Dr. Bhushan Bandre. (2013). Design and Analysis of Low Power Energy Efficient Braun Multiplier. International Journal of New Practices in Management and Engineering, 2(01), 08 - 16.

Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/12

[19] Chinthamu, N. ., Gooda, S. K. ., Venkatachalam, C. ., S., S. ., & Malathy, G. . (2023). IoT-based Secure Data Transmission Prediction using Deep Learning Model in Cloud Computing. International Journal on Recent and Innovation Trends in Computing and Communication, 11(4s), 68–76. https://doi.org/10.17762/ijritcc.v11i4s.6308