ORIGINAL ARTICLE

Comparison of Different Deep Learning Frameworks for Hippocampus Body and Head Segmentation

Hanie Arabian [1], Alireza Karimian [1*], Reza Rasti [1], Hossein Arabi [2]

[1] *Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran*

[2] *Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva, Switzerland*

*Corresponding Author: Alireza Karimian

Email: karimian@eng.ui.ac.ir

Received: 21 September 2024 / Accepted: 31 May 2025

Abstract

Purpose: Precise hippocampus segmentation from Magnetic Resonance Imaging (MRI) scans is vital in diagnosing various neurological disorders. Traditional segmentation methods face challenges due to the hippocampus's complex structure. This study compares four deep learning frameworks to segment hippocampal parts, including concurrent, separated, ordinal, and attention-based strategies.

Materials and Methods: This research utilized 3D T1-weighted MR images with manually delineated hippocampus head and body labels from 260 participants. The images were randomly split into five folds for experimentation, each time one of those was designated as the test set and the rest as the training set.

Results: The findings indicate that both concurrent and separated frameworks perform better than the ordinal and attention-based frameworks regarding the Dice and Jaccard coefficients. In head segmentation, the separated framework had a Dice similarity of 0.8748, a Jaccard similarity of 0.7794, and a Hausdorff distance of 5.4160. In body segmentation, the concurrent framework had a Dice similarity of 0.8616, a Jaccard similarity of 0.7591, and a sensitivity of 0.8437. Statistical results from the one-way ANOVA test showed a significant difference in performance for the body part (P-value=0.008), but not for the head region (P-value=0.652) between concurrent and separated frameworks. Comparing the concurrent with ordinal and attention-based frameworks showed a significant difference in both body and head regions (P-value<0.001).

Conclusion: Researchers must consider the differences between various frameworks while selecting a segmentation method for their specific task. Understanding the strengths and weaknesses of every framework is essential for deciding on the top-rated segmentation approach for precise applications.

Keywords: Deep Learning; Magnetic Resonance Imaging; Segmentation; Concurrent Framework.

DOI:

1. Introduction

The hippocampus plays a crucial role in memory formation, spatial navigation, emotional regulation, and learning abilities. Damage or dysfunction in this area can lead to memory issues and other cognitive impairments. [1, 2]. Brain magnetic resonance imaging (MRI) is a key imaging method in hippocampal studies, as precise hippocampus segmentation is crucial for diagnosing neurological conditions [3-9]. Despite its significance, hippocampal segmentation remains a challenging task. While manual hippocampus segmentation by medical professionals serves as a benchmark, it is characterized as time-consuming, costly, and susceptible to variability among observers [5, 10, 11].

The intricate structure of the hippocampus, variability in its shape and size, poor image quality, and pathological changes make it difficult to develop a universal segmentation algorithm [12, 13]. Therefore, traditional image segmentation techniques often fail to address these challenges adequately [1, 14]. In the past few years, deep-learning techniques have demonstrated encouraging outcomes in this field [15-17]. The latest hippocampus segmentation techniques utilize advanced deep-learning algorithms. Convolutional Neural Networks (CNNs) have emerged as an innovative approach for segmenting the hippocampus [6, 18-22].

In 2019, Ataloglou *et al.* employed deep learning methods, especially CNNs, to create a fully automated segmentation algorithm for the hippocampus. The researchers tested different training strategies, including transfer learning, to blend multiple datasets for enhanced segmentation accuracy. The developed technique yielded a Dice coefficient of 0.90 and 0.88 for the EADC-ADNI Harp dataset and MICCAI dataset, respectively [22].

In 2020, Goubran *et al.* employed two neural networks, one of which was trained on the whole brain, and the other was trained using the outputs of the first network. The architecture was similar to a U-net, integrating residual blocks and trainable deconvolution kernels. The method proposed by the researchers achieved an average Dice coefficient of 0.89 and a correlation coefficient of 0.95 [21].

In 2020, Basher and colleagues developed a model based on CNNs in 2 dimensions using a small quantity of pre-processed MRI. The researchers noted that their approach was efficient in terms of computation, taking less than 2 seconds to determine voxel counts for an MRI scan. Their results showed correlation coefficients of 0.834 and 0.848 for the left and right hippocampus, respectively [23].

In 2020, Lui *et al.* utilized a complex deep learning framework incorporating a CNN to simultaneously segment the hippocampus and classify Alzheimer's disease using MRI data. They developed a multi-task deep CNN model utilizing a 3D DenseNet for feature extraction. The combined features from the CNN and DenseNet models were then used for disease classification, resulting in a notable Dice similarity coefficient of 87% [24].

In 2022, a new technique was introduced by Hazarika *et al.* for segmenting the hippocampus using a U-Net Convolutional Network. To enhance feature extraction, this adaptation of the U-Net architecture utilized various kernel sizes instead of the typical size of 3×3. As a result of this alteration, the average performance rate for hippocampus segmentation saw a significant increase to 96.5%, surpassing the original U-Net model [1].

In 2023, LI *et al.* proposed a new framework called ESDSA for segmenting the hippocampus parts. This framework includes a spatial self-attention mechanism to improve segmentation. The researchers combined the hippocampus features with the MRI ones for successful hippocampus segmentation. The results of the study demonstrated that ESDSA outperformed other existing methods, achieving a Dice similarity coefficient of 89.37% [12].

Many research efforts have been dedicated to enhancing deep-learning algorithms to segment the hippocampus. Previous studies have primarily emphasized improving loss functions and network configurations. However, the segmentation of specific hippocampal subregions, such as the head and body, has received limited attention. Addressing this gap is crucial, as the subregion-specific segmentation can enhance the diagnostic process and provide more detailed insights into hippocampal pathology.

Our current research focuses on exploring various deep-learning training methods for segmenting the hippocampal parts in MRIs [25]. This study aims to compare distinct deep-learning frameworks, including ordinal, separated, concurrent, and attention-based techniques, to identify the best strategy. To our knowledge, this is the first study to provide a comprehensive evaluation of these frameworks in the context of subregion-specific segmentation. By exploring and assessing the performance of different segmentation approaches, this study will aim to give valuable insights into the best method for accurate and efficient hippocampal segmentation in medical imaging applications.

2. Materials and Methods

2.1. Data Acquisition

This research utilized a set of online and public MR images with manually delineated hippocampus head and body labels. The MR images were in 3D T1-weighted mode with a voxel size of 1 mm3 (TI/TE/TR, 860/3.7/8.0 ms) from 260 participants. The images were obtained from a sample of 90 healthy adults and 105 adults diagnosed with a non-affective psychotic disorder (including 56 schizophrenia, 32 schizoaffective disorder, and 17 schizophreniform disorder) sourced from the Psychiatric Genotype/Phenotype Project data repository at Vanderbilt University Medical Center in Nashville, TN, USA (http://medicaldecathlon.com [26]). The participants did not have any serious medical problems or a history of head injury, or were not under any drugs/ alcohol at the time of the study. Each MR image had a gray-scale mask of the hippocampus as ground truth in which the head and body were segmented [27]. The head part was given the number 1, the body part numbered 2, and the background labeled 0. Using this dataset, the models were trained and assessed.

2.2. Data Preprocessing

Before implementing the selective frameworks, intensity normalization was performed on images using a 90%-cumulative histogram of each subject. All slices were then cropped to 40×56 (in the X-Y plane) or zero-padded based on the largest size of the hippocampus in the available images. Ultimately, the images that went through the procedure were randomly split into five folds (each having 54 images). During each experiment, a given fold was allocated as a test set, whereas all the other four folds became the training set.

2.3. Different Deep Learning Frameworks

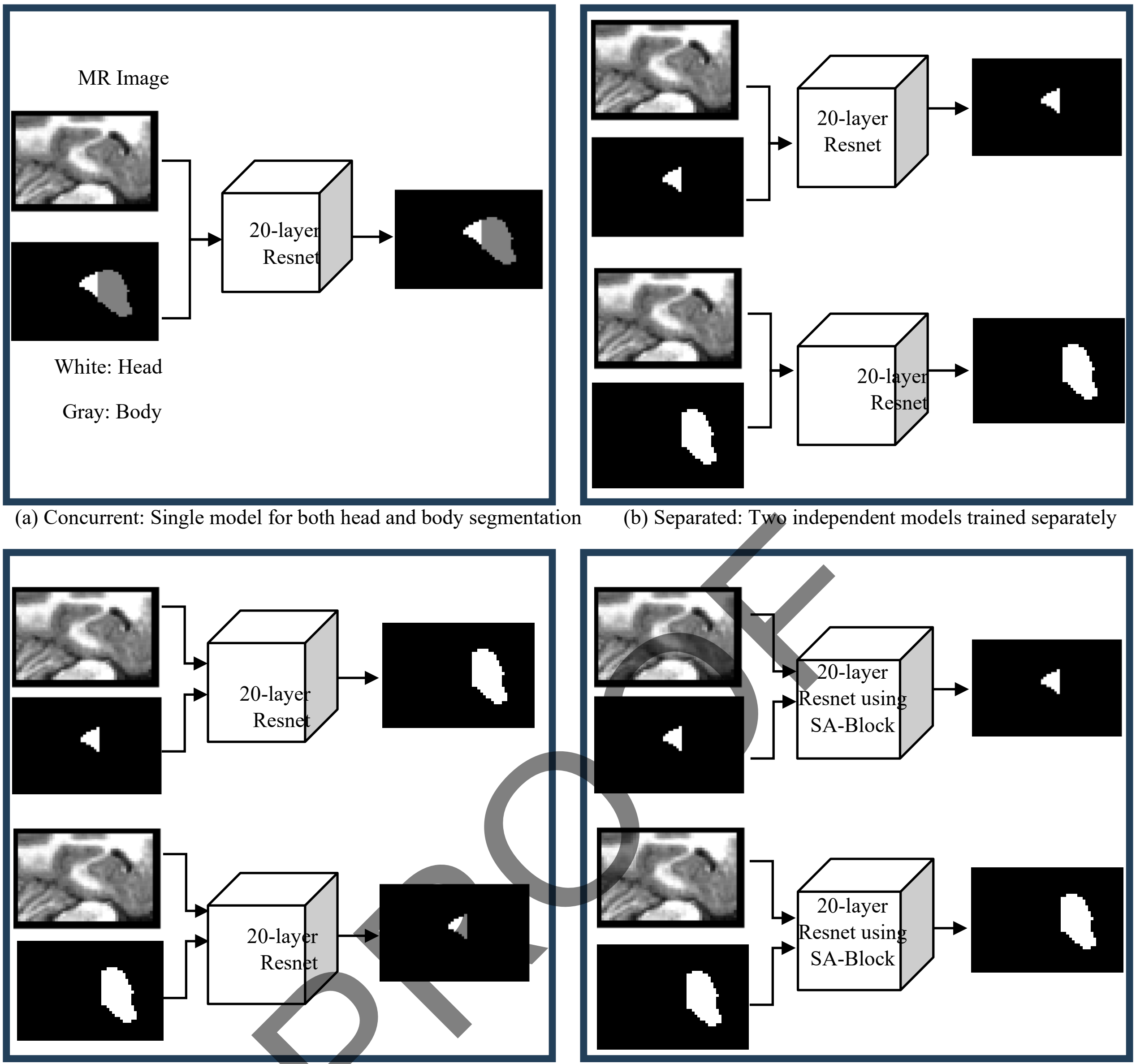
This study assessed four varying methods for training and implementing deep learning. Hippocampus segmentation was carried out using a 20-layer residual neural network with dilated convolutional kernels (Resnet) on the NiftyNet platform, an open-source Convolutional Neural Network (CNN) platform that utilizes TensorFlow for medical image analysis and therapy [28].

In the previous study, we used the Dice and Cross-Entropy loss functions for neural networks. Due to the superior performance of the Dice loss function compared to the Cross-Entropy loss function, we chose to continue using Dice loss in our expansion study for consistency and to build upon the results of our initial work [25].

Some important network parameters are set: Dice_NS loss functions and the Adam optimizer were used. The learning rate varied from 0.01 to 0.005. The activation function was Leakyrelu. Due to the RAM capacity, the batch size could be changed between 100 and 120. L2 regularization was used to update the weights and control overfitting. Figure 1 shows the examined hippocampal segmenting frameworks: separated, concurrent, ordinal, and attention-based.

Concurrent strategy trains a single model to segment the hippocampus head and body simultaneously. This process takes the images from the training dataset as input to the network, and the network learns the segmentation of the hippocampus head and body according to the label of each input image that distinguishes the head and body.

The separated strategy includes independently training the two ResNet models for the hippocampus head and body segmentation. This process involves extracting two binary images from each labeled image to explicitly represent the hippocampus's head and body. Subsequently, the model will be trained for head segmentation, and the training images with head labels corresponding to the hippocampus will be input into



(a) Concurrent: Single model for both head and body segmentation

(b) Separated: Two independent models trained separately

(c) Ordinal: Sequential segmentation: head → body (or vice versa)

(d) Attention-based: Attention module enhances feature

Figure 1. Different deep learning strategies for hippocampus segmentation during the training phase: (a) Concurrent strategy processes the head and body simultaneously in a single model. (b) Separated strategy trains two distinct models for the head and body independently. (c) Ordinal strategy processes segmentation of head and body sequentially, with one model's output informing the next. (d) Attention-based strategy integrates spatial attention mechanisms for improved feature extraction and segmentation accuracy

the network. Body segmentation will be performed using body labels similar to those of its predecessor.

The ordinal strategy trains a model first for body (or head) segmentation, and then the obtained segmentation probability map is used as prior knowledge for head (or body) segmentation. First, the training set images should generate the results of the parallel network to obtain probability maps of the hippocampus head and body and then use them as the prior knowledge of the ordinal training network. Similar to the separated strategy, two independent networks must be trained for head and body segmentation. The first network for body segmentation takes the images of the training set and the probability map obtained from the head as input, and then the network is trained according to the body label. Similarly, for head segmentation, MR images and probability maps of the body are the input of the network, and it is trained according to the network head label.

The attention-based framework trains two separate models to segment the head and body by incorporating an attention module into the Resnet architecture to concentrate solely on the specific targeted structure (whether the head or body). The attention module reduces the importance of irrelevant regions in the input image, allowing the ResNet model to extract only the relevant features from the target structure.

2.4. Evaluation

Five performance criteria were calculated for the test images of each fold to assess the performance of the different hippocampus head and body segmentation strategies [29]. These standard segmentation criteria are the Sorensen-Dice similarity coefficient or $dice$ (Equation 1), the Jaccard Similarity or d_J (Equation 2), the Hausdorff distance or d_H (Equation 3), the Absolute Volume Difference or AVD (Equation 4), and Sensitivity or Se (Equation 5).

$$dice(A,B) = \frac{2|A \cap B|}{|A| + |B|} \tag{1}$$

$$d_J(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

$$d_H(A,B) = max\left\{\sup_{a \in A} d(a,B), \sup_{b \in B} d(A,b)\right\} \tag{3}$$

$$AVD(A,B) = \frac{||A| - |B||}{|A| + |B|} \tag{4}$$

$$Se(A,B) = \frac{|A \cap B|}{|A|} \tag{5}$$

Where A and B denote the reference mask and the estimated masks, respectively. Also, $|A|$ and $|B|$ represent the cardinality of images A and B. Besides, $d(a,B)$ quantifies the minimum distance from a point $a \in A$ to the subset $b \in B$.

The higher the Dice and Jaccard similarity value (close to 1), the more similar the output is to the gold standard. The lower the Hausdorff distance and AVD value (close to 0), the less error the output has compared to the gold standard.

The mean and standard deviations of the mentioned criteria were reported for each method. In comparing the performance criteria, one-way ANOVA was done to check the significant differences among different strategies at the level of P-value≤0.05

3. Results

After applying each approach to test sets, all performance aspects were determined for every five-fold test set. The performance of concurrent, separated, ordinal, and attention-based frameworks (the mean and the standard deviation of each criterion on all five sets) in hippocampus head and body segmentation is reported in Tables 1 and 2, respectively. In addition, the minimum and maximum value of each criterion is reported.

The mean of the Dice coefficient of body segmentation for concurrent, separated, ordinal, and attention-based frameworks is 0.8714, 0.8748, 0.8721, and 0.8534, respectively. The mean of the Dice coefficient of head segmentation for concurrent, separated, ordinal, and attention-based frameworks is 0.8616, 0.8581, 0.8527, and 0.8502, respectively.

The mean of Hausdorff distance of body segmentation for concurrent, separated, ordinal, and attention-based frameworks is 5.5347, 5.4160, 5.5113, and 5.8247, respectively. The mean of the Hausdorff distance of head segmentation for concurrent, separated, ordinal, and attention-based frameworks is 5.9543, 6.0413, 6.2137, and 6.2531, respectively.

Figure 2 shows the box plots of Dice and Jaccard similarity measures for all four hippocampal head and body segmentation frameworks. The minimum to maximum range, mean, and outliers from Dice and Jaccard similarity criteria for each framework can be seen separately for head and body segmentation.

The one-way ANOVA test was performed on Dice similarity measures to compare the performance of different frameworks in the head and body of the hippocampus segmentation. The P-values are reported in Table 3.

4. Discussion

According to Table 1, the separated framework in head segmentation achieved a higher value in terms of Dice similarity, Jaccard similarity, and also a lower Hausdorff distance which was equal to 0.8748±0.0370

Table 1. Comparison of hippocampus' Head segmentation performance (mean±SD) between the different frameworks using the selected evaluation metrics, including the Dice, Jaccard, Hausdorff, Absolute Volume Difference, and Sensitivity

Framework	Dice	Jaccard	Hausdorff	AVD	Sensitivity
Concurrent	0.8714±0.0392	0.7742±0.0592	5.5347±0.7679	0.0770±0.0612	0.8535±0.0588
[min-max]c	[0.6620-0.9342]	[0.4948-0.8766]	[3.8730-9.6954]	[0-0.2884]	[0.5851-0.9545]
Separated	0.8748±0.0370	0.7794±0.0566	.4160±0.7134	0.0803±0.0619	0.8531±0.0564
[min-max]	[0.7390-0.9294]	[0.5860-0.8682]	[3.6056-7.8102]	[0.0014-0.3103]	[0.6540-0.9589]
Ordinal	0.8721±0.0384	0.7752±0.0578	5.5113±0.7716	0.0782±0.0603	0.8511±0.0582
[min-max]	[0.6560-0.9291]	[0.4880-0.8676]	[4.000-10.0499]	[0.0011-0.3153]	[0.5575-0.9540]
Attention-based	0.8534±0.0531	0.7478±0.0761	5.8247±0.9786	0.0943±0.0805	0.8446±0.0615
[min-max]	[0.5995-0.9294]	[0.4281-0.8681]	[4.000-9.3274]	[0.0015-0.5232]	[0.6521-0.9702]

Table 2. Comparison of hippocampus' Body segmentation performance (mean±SD) between the different frameworks using the selected evaluation metrics, including the Dice, Jaccard, Hausdorff, Absolute Volume Difference, and Sensitivity

Framework	Dice	Jaccard	Hausdorff	AVD	Sensitivity
Concurrent	0.8616±0.0410	0.7591±0.0600	5.9543±0.8179	0.0769±0.0701	0.8437±0.0548
[min-max]	[0.6539-0.9216]	[0.4858-0.8546]	[4.5826-10.3441]	[0.0005-0.8164]	[0.6010-0.9502]
Separated	0.8581±0.0421	0.7537±0.0606	6.0413±0.7932	0.0793±0.0600	0.8358±0.0572
[min-max]	[0.5497-0.9310]	[0.3790-0.8708]	[4.4721-9.7980]	[0.0006-0.4082]	[0.4948-0.9569]
Ordinal	0.8527±0.0399	0.7453±0.0574	6.2137±0.7576	0.0854±0.0647	0.8272±0.0541
[min-max]	[0.5962-0.9140]	[0.4247-0.8416]	[4.4721-8.1854]	[0-0.4816]	[0.5738-0.9412]
Attention-based	0.8502±0.0399	0.7415±0.0580	6.2531±0.7886	0.0968±0.0682	0.8187±0.0579
[min-max]	[0.6179-0.9232]	[0.4471-0.8573]	[4.3589-10.2470]	[0-0.3217]	[0.5238-0.9285]

Table 3. P-values of one-way ANOVA test

Framework 1	Framework 2	Dice	
		Body	Head
Concurrent	Separated	0.008	0.652
	Ordinal	<0.001	1.000
	Attention-based	<0.001	<0.001
Separated	Ordinal	<0.001	1.000
	Attention-based	<0.001	0.009
Ordinal -based	Attention-based	0.424	0.003

(95% CI; 0.870 to 0.879), 0.7794±0.0566 (95% CI; 0.772 to 0.786), and 5.4160±0.7134 (95% CI; 5.329 to 5.503), respectively. Furthermore, the concurrent framework in head segmentation had the highest sensitivity value of 0.8535±0.0588 and the lowest Absolute Volume Difference of 0.0770±0.0612 (95% CI; 0.070 to 0.084). According to Table 2, the concurrent framework in body segmentation achieved a higher value in terms of Dice, Jaccard similarity, and sensitivity, which were equal to 0.8616±0.0410 (95% CI; 0.857 to 0.867), 0.7591±0.0600 (95% CI; 0.752 to 0.766), and 0.8437±0.0548 (95% CI; 0.837 to 0.850), respectively. In addition, it achieved a lower value in terms of the Hausdorff distance and the Absolute Volume of difference, which were equal to 5.9543±0.8179 (95% CI; 5.854 to 6.054) and

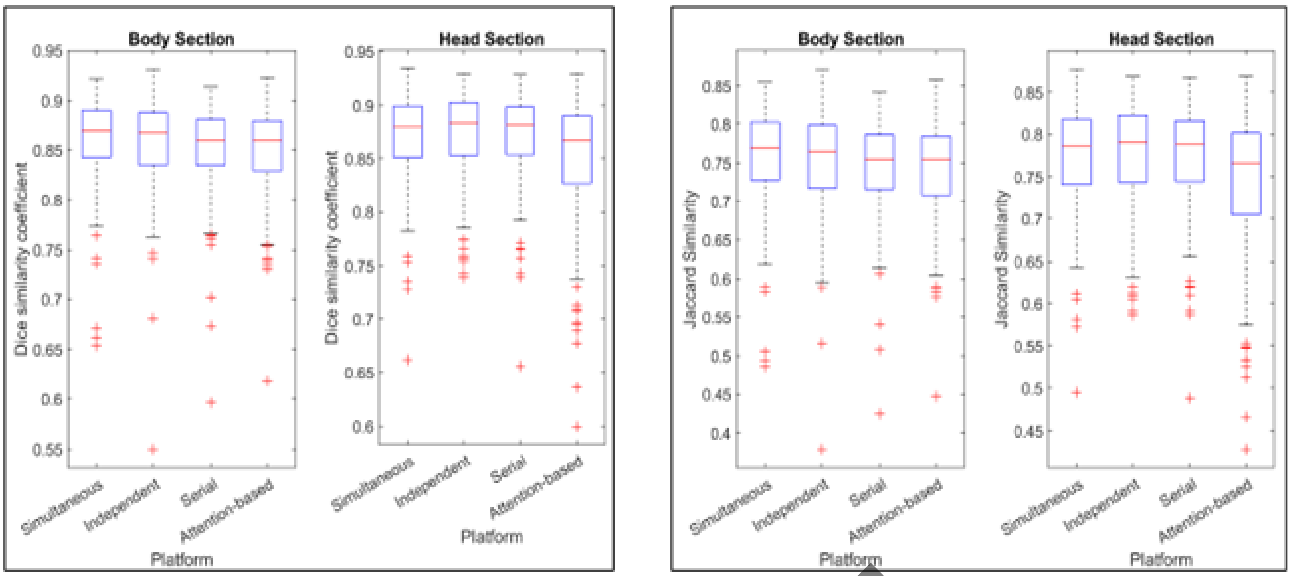


Figure 2. Box plots of the Dice and Jaccard similarity measures of examined frameworks for the hippocampus head and body segmentation. Left) Box plot of Dice measures. Right) Box plot of Jaccard measures

0.0769±0.0701 (95% CI; 0.068 to 0.085), respectively.

As reported in Tables 1 and 2, the Attention-based framework has the lowest value of Dice, Jaccard, and sensitivity, and the highest value of Hausdorff distance and Absolute Volume of Difference for both head and body segmentation. This underperformance of the attention-based framework could be attributed to the need for larger datasets in these mechanisms to capture spatial dependencies effectively. The attention mechanisms in general require sufficiently large dataset to efficiently detect the optimum attention field and dedicate the resources of the network only to the target region [6, 18]. Therefore, the limited sample size in this study may have affected the ability of this mechanism. Furthermore, the added complexity of the attention layers can lead to suboptimal feature extraction and reduce the model's ability to discriminate hippocampal subregions accurately.

As Figure 2 shows, in the worst case, the Dice criterion in body segmentation by the separated framework has reduced to 0.55, and in the head segmentation by the attention-based framework, it has reduced to 0.6. Also, in the worst case, the Jaccard similarity in body segmentation by the separated framework has reduced to less than 0.4, and in the head segmentation by the attention-based framework, it has reduced to less than 0.45.

The box plots of the Dice and Jaccard criteria showed that the lowest outlier distance to the average in body and head segmentation is related to the concurrent and separated frameworks, respectively. In addition, there was no noticeable difference in the maximum Dice and Jaccard similarity value for hippocampal head segmentation between the examined frameworks.

The one-way ANOVA test results in Table 3 show significant differences in Dice coefficients between the different segmentation methods for certain body and head regions. Specifically, for the comparison between the concurrent and separated methods, the P-value is 0.008 for the body region, indicating a significant difference in performance. However, there is no significant difference for the head region with a P-value of 0.652.

Similarly, the comparison between the concurrent and ordinal methods shows significant differences for both body and head regions, with P-values of <0.001. The comparison between the concurrent and attention-based methods also shows significant differences for both regions, with P-values of <0.001 for both cases.

When comparing the separated and ordinal methods, there are significant differences for the body region with a P-value of less than 0.001, but not for the head region with a P-value of 1.000. The comparison between the separated and attention-based methods shows significant differences for both body and head regions, with P-values of less than 0.001 and 0.009, respectively.

In conclusion, the head region significantly differs between the ordinal and attention-based methods, with a P-value of 0.003, while there is no difference in the body region, with a P-value of 0.424.

This study utilized a publicly available dataset of MR images that did not provide information on the age, gender, or ethnic diversity of its participants. While this limitation does not diminish the methodological validity of this study, future studies should consider more diverse populations to account for the influence of factors such as ethnicity, age, or health conditions on changes in hippocampal structure. Further research should examine the generalizability of the proposed segmentation techniques to broader populations using datasets from diverse populations.

5. Conclusion

This study compared four deep learning frameworks—concurrent, separated, ordinal, and attention-based—for hippocampus head and body segmentation in MRI images. The results showed that the concurrent and separated frameworks outperformed the sequential and attention-based approaches in segmenting the head and body of the hippocampus. The attention-based framework showed lower segmentation performance, which can probably be attributed to its reliance on larger datasets and increased architectural complexity. These findings highlight the effectiveness of structured training strategies for hippocampal segmentation, and researchers should consider these differences when choosing a segmentation method for their specific application. Further research is needed to optimize attention-based methods using larger and more diverse datasets and to use multimodal imaging techniques to improve segmentation performance and generalizability.

Acknowledgment

The authors received no financial support for the research, authorship, and/or publication of this article.

References

1- Ruhul Amin Hazarika, Arnab Kumar Maji, Raplang Syiem, Samarendra Nath Sur, and Debdatta Kandar, "Hippocampus segmentation using U-net convolutional network from brain magnetic resonance imaging (MRI)." *Journal of Digital Imaging,* Vol. 35 (No. 4), pp. 893-909, (2022).

2- S. Yazdani, R. Yusof, A. Karimian, A. H. Riazi, and M. Bennamoun, "A Unified Framework for Brain Segmentation in MR Images." *Computational and Mathematical Methods in Medicine,* Vol. 2015 (No. 1), p. 829893, 2015/01/01 (2015).

3- Sarah Atwi, Dallan McMahon, Helen Scharfman, and Neil J MacLusky, "Androgen modulation of hippocampal structure and function." *The Neuroscientist,* Vol. 22 (No. 1), pp. 46-60, (2016).

4- A. Basher *et al.*, "Hippocampus Localization Using a Two-Stage Ensemble Hough Convolutional Neural Network." *IEEE Access,* Vol. 7pp. 73436-47, (2019).

5- Vanderson Dill, Alexandre Rosa Franco, and Márcio Sarroglia Pinho, "Automated methods for hippocampus segmentation: the evolution and a review of the state of the art." *Neuroinformatics,* Vol. 13pp. 133-50, (2015).

6- Reza Karimzadeh, Emad Fatemizadeh, and Hossein Arabi, "Attention-based deep learning segmentation: Application to brain tumor delineation." in *2021 28th National and 6th International Iranian Conference on Biomedical Engineering (ICBME)*, (2021): *IEEE*, pp. 248-52.

7- P Sharifian, A Karimian, H Arabi, and H Zaidi, "Anatomical brain regions segmentation from MR images using global and local deep learning networks." in *2023 IEEE Nuclear Science Symposium, Medical Imaging Conference and International Symposium on Room-Temperature Semiconductor Detectors (NSS MIC RTSD)*, (2023): *IEEE*, pp. 1-1.

8- Frauke Nees and Sebastian T Pohlack, "Functional MRI studies of the hippocampus." *The Hippocampus in Clinical Neuroscience,* Vol. 34pp. 85-94, (2014).

9- Hossein Arabi and Habib Zaidi, "Recent advances in positron emission tomography/magnetic resonance imaging technology." *Magnetic Resonance Imaging Clinics,* Vol. 31 (No. 4), pp. 503-15, (2023).

10- Sandra González-Villà, Arnau Oliver, Sergi Valverde, Liping Wang, Reyer Zwiggelaar, and Xavier Lladó, "A review on brain structures segmentation in magnetic resonance imaging." *Artificial Intelligence in Medicine,* Vol. 73pp. 45-69, 2016/10/01/ (2016).

11- Hossein Arabi and Habib Zaidi, "Single annotator versus multi-annotator: Challenge of segmenting two neighboring hippocampus head and body with high precision." *Biomedical Signal Processing and Control,* Vol. 97p. 106667, (2024).

12- Jia-Ni Li, Shao-Wu Zhang, Yan-Rui Qiang, and Qin-Yi Zhou, "A novel cross-layer dual encoding-shared decoding network framework with spatial self-attention mechanism for hippocampus segmentation." *Computers in Biology and Medicine,* Vol. 167p. 107584, (2023).

13- Hossein Arabi and Habib Zaidi, "Segmentation of the Hippocampus Head and Body: Comparison of Single Annotator and Multi-annotator." in *2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, (2021): *IEEE*, pp. 1-3.

14- H. Arabi and H. Zaidi, "Whole-body bone segmentation from MRI for PET/MRI attenuation correction using shape-based averaging." (in eng), *Med Phys,* Vol. 43 (No. 11), p. 5848, Nov (2016).

15- Carlos E Cardenas, Jinzhong Yang, Brian M Anderson, Laurence E Court, and Kristy B Brock, "Advances in auto-segmentation." in *Seminars in radiation oncology*, (2019), Vol. 29 (No. 3): *Elsevier*, pp. 185-97.

16- Hossein Arabi, Azadeh AkhavanAllaf, Amirhossein Sanaat, Isaac Shiri, and Habib Zaidi, "The promise of artificial intelligence and deep learning in PET and SPECT imaging." *Physica Medica,* Vol. 83pp. 122-37, (2021).

17- Hossein Arabi and Habib Zaidi, "Applications of artificial intelligence and deep learning in molecular imaging and radiotherapy." *European Journal of Hybrid Imaging,* Vol. 4 (No. 1), p. 17, (2020).

18- Seyed Reza Seif, Alireza Karimian, Hossein Arabi, and Habib Zaidi, "Impact of Attention Modules in Deep Learning-based Semantic Segmentation: Evaluation for Liver Lesion Segmentation from CT Images." in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, (2022): *IEEE*, pp. 1-3.

19- Parisa Talebi, Elham Saeedzadeh, Mohsen Bakhshandeh, and Hossein Arabi, "A Novel Attention-based Neural Network for Automated Lung Lesion Delineation from 4DCT Images." in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, (2022): *IEEE*, pp. 1-3.

20- H. Arabi and H. Zaidi, "Contrastive Learning vs. Self-Learning vs. Deformable Data Augmentation in Semantic Segmentation of Medical Images." *J Imaging Inform Med,* pp. 1-14, Jun 10 (2024).

21- Maged Goubran *et al.*, "Hippocampal segmentation for brains with extensive atrophy using three-dimensional convolutional neural networks." *Human Brain Mapping,* Vol. 41 (No. 2), pp. 291-308, (2020).

22- Dimitrios Ataloglou, Anastasios Dimou, Dimitrios Zarpalas, and Petros Daras, "Fast and precise hippocampus segmentation through deep convolutional neural network ensembles and transfer learning." *Neuroinformatics,* Vol. 17pp. 563-82, (2019).

23- A. Basher, B. C. Kim, K. H. Lee, and H. Y. Jung, "Automatic Localization and Discrete Volume Measurements of Hippocampi From MRI Data Using a Convolutional Neural Network." *IEEE Access,* Vol. 8pp. 91725-39, (2020).

24- Manhua Liu *et al.*, "A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease." *Neuroimage,* Vol. 208p. 116459, (2020).

25- Hanieh Arabian, Alireza Karimian, Reza Rasti, Hossein Arabi, and Habib Zaidi, "Deep Attention-based Seminal Segmentation: A Practical Deep Learning Framework for Accurate Segmentation of the Hippocampus from Magnetic Resonance Images." in *2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, (2021): *IEEE*, pp. 1-3.

26- Michela Antonelli *et al.*, "The medical segmentation decathlon." *Nature communications,* Vol. 13 (No. 1), p. 4128, (2022).

27- Amber L Simpson *et al.*, "A large annotated medical image dataset for the development and evaluation of segmentation algorithms." *arXiv preprint arXiv:1902.09063,* (2019).

28- Eli Gibson *et al.*, "NiftyNet: a deep-learning platform for medical imaging." *Computer Methods and Programs in Biomedicine,* Vol. 158pp. 113-22, 2018/05/01/ (2018).

29- Abdel Aziz Taha and Allan Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool." *BMC Medical Imaging,* Vol. 15 (No. 1), p. 29, 2015/08/12 (2015).

.