University of Texas Rio Grande Valley

ScholarWorks @ UTRGV

3-22-2024

Ex-Vivo Hippocampus Segmentation Using Diffusion-Weighted MRI

Haoteng Tang
The University of Texas Rio Grande Valley

Siyuan Dai

Eric M. Zou

Guodong Liu

Ryan Ahearn

See next page for additional authors

Follow this and additional works at: https://scholarworks.utrgv.edu/cs_fac

Part of the Computer Sciences Commons

Authors

Haoteng Tang, Siyuan Dai, Eric M. Zou, Guodong Liu, Ryan Ahearn, Ryan Krafty, Michel Modo, and Liang Zhan

Article

Ex-Vivo Hippocampus Segmentation Using Diffusion-Weighted MRI

Haoteng Tang [1,*], **Siyuan Dai** [2], **Eric M. Zou** [3], **Guodong Liu** [4], **Ryan Ahearn** [5], **Ryan Krafty** [5], **Michel Modo** [5] **and Liang Zhan** [2,*]

1 Department of Computer Science, University of Texas Rio Grande Valley, Edinburg, TX 78539, USA
2 Department of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA 15260, USA
3 Montgomery Blair High School Maryland, Silver Spring, MD 20901, USA
4 Department of Computer Science, University of Maryland, College Park, MD 20742, USA
5 Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15213, USA
* Correspondence: haoteng.tang@utrgv.edu (H.T.); liang.zhan@pitt.edu (L.Z.)

Abstract: The hippocampus is a crucial brain structure involved in memory formation, spatial navigation, emotional regulation, and learning. An accurate MRI image segmentation of the human hippocampus plays an important role in multiple neuro-imaging research and clinical practice, such as diagnosing neurological diseases and guiding surgical interventions. While most hippocampus segmentation studies focus on using T1-weighted or T2-weighted MRI scans, we explore the use of diffusion-weighted MRI (dMRI), which offers unique insights into the microstructural properties of the hippocampus. Particularly, we utilize various anisotropy measures derived from diffusion MRI (dMRI), including fractional anisotropy, mean diffusivity, axial diffusivity, and radial diffusivity, for a multi-contrast deep learning approach to hippocampus segmentation. To exploit the unique benefits offered by various contrasts in dMRI images for accurate hippocampus segmentation, we introduce an innovative multimodal deep learning architecture integrating cross-attention mechanisms. Our proposed framework comprises a multi-head encoder designed to transform each contrast of dMRI images into distinct latent spaces, generating separate image feature maps. Subsequently, we employ a gated cross-attention unit following the encoder, which facilitates the creation of attention maps between every pair of image contrasts. These attention maps serve to enrich the feature maps, thereby enhancing their effectiveness for the segmentation task. In the final stage, a decoder is employed to produce segmentation predictions utilizing the attention-enhanced feature maps. The experimental outcomes demonstrate the efficacy of our framework in hippocampus segmentation and highlight the benefits of using multi-contrast images over single-contrast images in diffusion MRI image segmentation.

Keywords: hippocampus; segmentation; multi-contrast data; anisotropy; diffusion MRI

MSC: 68T07

1. Introduction

The hippocampus, nestled within the human brain's temporal lobe, is a vital structure central to neuroscience and clinical research. Its seahorse-shaped form plays a crucial role in cognitive functions like memory formation and spatial navigation [1]. The landmark case of patient H.M. underscored its importance, revealing memory's dependency on the hippocampus [2]. Furthermore, the discovery of 'place cells' elucidated its role in spatial orientation [3]. Neuro-imaging techniques such as magnetic resonance imaging (MRI) have deepened our understanding, revealing its vulnerability in neurodegenerative disorders [4] and activation patterns during memory tasks [5]. The hippocampus remains a focal point in neuroscience, with ongoing studies unraveling its complexities and implications in various disorders [6–9].

MRI hippocampus segmentation is crucial in neuro-imaging and clinical practice due to the hippocampus's pivotal role in cognitive functions and susceptibility to neurodegenerative diseases. It involves delineating the hippocampus from other brain structures, enabling precise analysis of size, shape, and volume. This accuracy is essential for various reasons:

- Diagnosis and Disease Monitoring: Accurate segmentation aids in diagnosing and monitoring neurodegenerative diseases like Alzheimer's [10,11], epilepsy, and dementia. It helps quantify hippocampal atrophy, a key biomarker in Alzheimer's disease.
- Surgical Intervention and Therapy Planning: In epilepsy surgery, precise segmentation ensures minimal impact on healthy brain tissue [12]. During radiation therapy, protecting the hippocampus from exposure minimizes cognitive impairment risks [13].
- Cognitive Neuroscience Research: Segmentation supports research on memory and spatial navigation, deepening understanding of neural mechanisms underlying cognition [14].
- Personalized Medicine: Variations in hippocampal structure impact disease susceptibility and treatment responses. Accurate segmentation enables tailored treatment plans based on individual neuroanatomy [15].

Most of recent hippocampus segmentation studies focus on utilizing T1-weighted (T1w) or T2-weighted (T2w) MRI scans due to the convenience of data collection and acquisition [16–20]. For example, [16] compared different segmentation methods by utilizing T1w and T2w multispectral MRI data, which highlighted the complexities of hippocampal structure and the benefits of using high-resolution T2w images for better contrast properties in subfield delineation. It also examined the reliability of different MRI sequences and their combinations for accurate hippocampal segmentation. Manjón et al. [17] introduced a novel deep learning-based hippocampus subfield segmentation method, which utilized a variant of the U-NET architecture [21], combining both T1w and T2w images to improve the performance of hippocampus segmentation. However, very few studies focus on utilizing diffusion-weighted MRI (dMRI) on hippocampus segmentation, which provides unique insights offered by dMRI into the microstructural properties of brain tissue. Firstly, dMRI offers insights into the microstructural environment of the hippocampus, which includes the orientation and integrity of white matter tracts and can reveal subtle changes in hippocampal tissue not visible on conventional MRI [22]. Secondly, dMRI performs well in detecting early microstructural changes in the hippocampus associated with neurodegenerative diseases like Alzheimer's [23]. Thirdly, the hippocampus is a hub in the brain's memory network, and its connections to other regions are crucial for its function. Hippocampus segmentation on dMRI scans will facilitate the assessment of these connections through techniques such as tractography, providing a more comprehensive view of hippocampal connectivity and its alterations in various conditions [24]. Achieving a mesoscale resolution for diffusion MRI and tractography further improves the delineation of hippocampal substructures and uncovers more detailed level of intra-regional grey matter connectivity [25,26]. Additionally, dMRI can offer a few insights into the functional aspects of the hippocampus, linking its structural properties to cognitive functions such as memory and learning. This is particularly relevant in understanding how structural changes impact function in diseases affecting the hippocampus [27–29].

In dMRI studies, several key metrics are utilized to quantify the diffusion of water molecules in brain tissue, each reflecting different aspects of tissue microstructure, and thereby can be utilized to derive multimodal dMRI images. The most important and widely utilized metrics are Fractional Anisotropy (FA), Mean Diffusivity (MD), Axia Diffusivity (AD) and Radial Diffusivity (RD). FA measures the degree of directional anisotropy of water diffusion in tissue. FA values tend to be high in areas where water diffusion is directionally restricted or oriented (e.g., white matter tracts), while they tend to be low in areas where diffusion is more isotropic (e.g., gray matter or cerebrospinal fluid). MD measures the average rate of water diffusion within a tissue and is derived from the diffusion coefficients along all measured directions. It reflects the overall mobility of water molecules, with higher values indicating more unrestricted diffusion. In brain

tissue, increased MD can be a sign of tissue degeneration or loss, as it suggests a greater ease of water movement, potentially due to loss of barriers like cell membranes. AD measures the rate of water diffusion along the primary axis of white matter fibers reflecting axonal integrity. Low AD values potentially indicate axonal injury or degeneration. RD measures the rate of water diffusion perpendicular to the primary axis of white matter fibers, which is particularly sensitive to changes in the myelin sheath that surrounds axons. All these metrics are important for an accurate hippocampus segmentations as they provide insights into the integrity and microstructure of brain tissues, which highlights the significance of multimodal hippocampus segmentation studies. Particularly, each metric provides insights into different aspects of the hippocampal microstructure. For example, AD and RD can provide information about axonal integrity and myelination, respectively. Compared to segment on each single modal scan, combining different metrics can lead to more accurate segmentation of the hippocampus, since these multimodal dMRI scans can provide complementary information (e.g., provide sufficient contrast information) to distinguish different semantic features [30].

The integration of deep learning methods in multi-contrast MRI image segmentation has been a significant advancement in neuro-imaging studies. These methods have harnessed the power of artificial intelligence to analyze complex datasets, offering more accurate and detailed insights into the structures of human brain or brain subregions (e.g., the hippocampus). One of the most pivotal developments in this field is the development of convolutional neural networks (CNNs), which have become a cornerstone in medical image analysis [31–34]. CNNs are particularly well suited for image data due to their ability to automatically and adaptively learn spatial hierarchies of features from images. In multi-contrast MRI segmentation, CNNs have been used to combine information from different MRI contrast methods (e.g., T1w, or T2w or diffusion MRI scans), to improve the segmentation performance, which is crucial as the different image contrasts provide complementary information about the anatomy of brain [35]. The U-Net, a CNN based architecture, has shown remarkable effectiveness in medical image segmentation tasks [21]. It can effectively captures contextual information while enabling precise localization, making it suitable for complex segmentation tasks, such as segmentation tasks related to regional brain substructures. However, U-Net is designed for single-modal tasks, which might limit its effectiveness to learn relationships across different contrasts of data when applying to the multi-contrast segmentation tasks. Therefore, we propose to address this issue via cross-attention mechanisms, which guides the deep learning framework to focus on the most relevant parts of an image to improve the segmentation performance in multi-contrast dMRI hippocampus segmentation tasks. Cross-attention extends the concept of self-attention [36] to interactions between at least two types of data obtained from different contrasts [37]. It has strong ability to model contextual relationships, and to combine the effective information across multi-contrast MRI images to perform the hippocampus segmentation tasks.

To sum up, our contributions in this work can be summarized as follows: (1). We used mesoscale diffusion MRI data to calculate multiple contrast images to segment the human hippocampus from its surrounding temporal lobe structures. (2). We build up a multimodal deep learning framework with the cross-attention mechanism to improve the performance of hippocampus segmentation. (3). We compared the segmentation performance of each single-contrast diffusion image with different combinations of multi-contrast images.

2. Related Works

2.1. Semantic Segmentation on Hippocampus

The accurate segmentation of the hippocampus is a vital image-processing step to assist the study of the hippocampus and related neurological disorders caused by the impairment of the hippocampus. The early techniques of hippocampus segmentation primarily involved manual delineation, which is time consuming and prone to inter-rater variability [38–40]. The complexity of the shape of hippocampus and its variable appear-

ance across individuals posed additional challenges. To address these limitations, a few semi-automated segmentation methods were developed, which typically involved user intervention for initialization or correction of segmentation results, striking a balance between automation and accuracy [41–43]. Advancements led to fully automated techniques, which were essential for handling large datasets in studies like Alzheimer's disease research. These methods utilized various computational strategies, such as region-growing algorithms, atlas-based approaches and machine learning algorithms [44,45]. The development of machine learning and deep learning marked a significant leap in hippocampus segmentation, where Convolutional Neural Networks (CNNs) as well as CNN-based segmentation architectures (e.g., U-Net) have demonstrated high accuracy and efficiency in segmenting the hippocampus from MRI scans [17,20,21,46–50]. These methods have the ability to learn complex patterns from diverse datasets which yields robust segmentation results. Beyond 2D segmentation studies, [51] introduced a 3D convolution model named DeepHipp, which integrates dense block and attention mechanisms for T1w hippocampus segmentation. This model is designed to improve the efficiency of feature usage by reusing features of each layer learned by the network, which allows the model to focus on the segmentation target and suppress irrelevant regions of the input image, enhancing the accuracy of hippocampus segmentation.

2.2. Deep Neural Networks for Multimodal MRI Hippocampus Segmentation

The field of deep learning for multimodal hippocampus MRI image segmentation has seen notable advancements, where a variety of approaches have been explored to improve the accuracy and efficiency of segmentation processes. Manjón et al. [17] proposed a variant of the of the U-Net architecture, which incorporates multiple resolution levels and a deep supervision approach to capture detailed hippocampal structures from T1w and T2w MRI scans [17]. Additionally, deep CNNs have been employed in the segmentation and classification of the hippocampus in Alzheimer's disease, which offers promising results in automating the segmentation process and potentially aiding in the early diagnosis of Alzheimer's [20]. Most of the deep learning methods for multimodal MRI hippocampus segmentation are based on the T1w and T2w neuro-imaging data [52,53].

3. Methodology

3.1. Proposed Segmentation Framework

The proposed segmentation framework (see Figure 1) integrates cross-attention mechanisms to improve the MRI image segmentation by utilizing multi-contrast MRI data. We will delve into the details of proposed multi-contrast segmentation framework in this section.

3.1.1. Framework Overview

As shown in Figure 1, the proposed segmentation framework consists of a multi-contrast encoder including K different branches to embed K different contrast maps, and a shared decoder to reconstruct the segmentation predictions from the latent space. A K^2 gated cross-attention unit is inserted between the encoder and the decoder to capture the cross relationships among different contrast maps. We adopt the encoder and decoder setting of the U-Net [21] architecture to serve as the segmentation backbone of our segmentation framework, where we duplicate K different encoder branches to embed K different contrast maps. Each encoder branch takes an contrast map X as input and generate the latent image feature maps F layer by layer. After K feature maps (i.e., $F_1, F_2, \ldots, F_K$) are generated, the K^2 gate cross-attention unit is utilized to generate the gated cross-attention matrix (i.e., A_g) based on these feature maps. Finally, the K feature maps are weighted by the gated cross-attention matrix as an attention-enhanced feature map (i.e., $F_{a_1}, F_{a_2}, \ldots, F_{a_K}$). The attention-enhanced feature map will be forwarded to the next encoder layer or the same decoder layer for segmentation predictions.
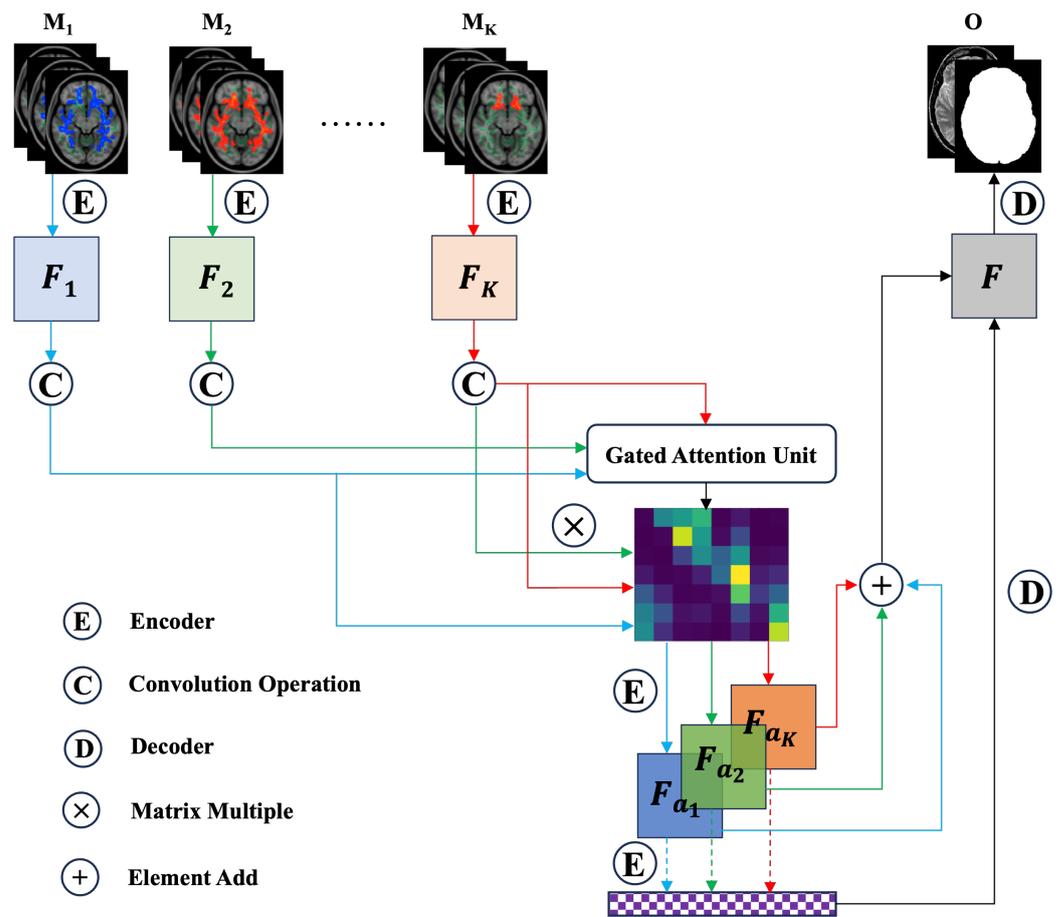
Figure 1. Diagram of the proposed multi-contrast segmentation framework with gated cross-attention unit.

3.1.2. K^2 Gated Cross-Attention Unit

Assume that the dimension of the generated feature map is $F \in \mathcal{R}^{H \times W \times C}$, where H and W are the size of the feature map and C is the channel number. We first reshape the feature map F to a new feature matrix $H \in \mathcal{R}^{N \times C}$, where $N = H \times W$. A linear layer is then utilized on H to adjust the channel number to $\hat{H}^{N \times c}$. As shown in Figure 2, $\hat{H}$ is utilized by the gated cross-attention unit to compute the gated cross-attention matrix. Let $\hat{H}_k$ be the *k-th* feature matrix, where $i = 1, 2, \ldots, K$. The cross-attention matrix between the *i-th* and *j-th* feature matrix can be computed by:

$$A_{i,j} = \hat{H}_i \times \hat{H}_j^T,\tag{1}$$

where i and j are in range of $[1, 2, \ldots, K]$ and the $A_{i,j} \in \mathcal{R}^{N \times N}$. Obviously, there are K^2 cross-attention matrix (When $i = j$, the cross-attention matrix is degraded into the self-attention matrix.). Based on this cross-attention matrix, the gated cross-attention matrix can be computed by:

$$A_g = \sigma(A_{1,1} \otimes A_{1,2} \otimes, \ldots, \otimes A_{1,K} \otimes A_{2,1} \otimes A_{2,2} \otimes, \ldots, \otimes A_{2,K} \otimes, \ldots, A_{K,1} \otimes A_{K,2} \otimes, \ldots, \otimes A_{K,K}),\tag{2}$$

where $\otimes$ denotes the element-wise multiplication operation and σ is a nonlinear activation function (i.e., softmax). After the gated cross-attention matrix is generate, we compute the attention-enhanced feature matrices by:

$$\hat{H}_{a_i} = A_g \times \hat{H}_i \times W,\tag{3}$$

where $i = 1, 2, \ldots, K$. W is trainable parameters of a linear layer to adjust the feature dimension back to the original dimension (i.e., from c to C). We then reshape these attention-enhanced feature matrices (i.e., $\hat{H}_{a_i} \in \mathcal{R}^{N \times C}$) back to the attention-enhanced feature maps (i.e., $F_{a_i} \in \mathcal{R}^{H \times W \times C}$) and feed-forward them to the following encoder layer. Meanwhile, we combine these attention-enhanced feature maps by $F_a = \sum_{k=1}^{K} F_{a_k}$ and feed-forward F_a to the segmentation decoder.
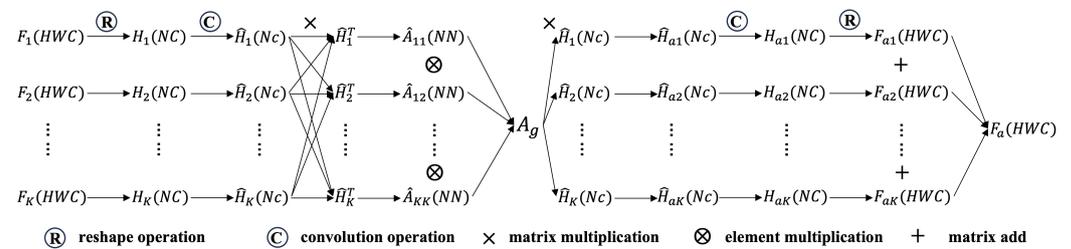

Figure 2. The computation within the gated cross-attention unit.

3.2. Hippocampus Segmentation with the Proposed Framework

In this study, we utilize four different contrast maps (i.e., $K = 4$) obtained from diffusion MRI images for hippocampus segmentation. For the U-Net based segmentation backbone, we adopt all default configurations used in the official implementations (https://github.com/milesial/Pytorch-UNet accessed on 15 March 2024) except for replacing the transposed convolution with the bi-linear interpolation in the decoder side. We deploy the gated cross-attention unit in third and fourth layer of encoder where the feature maps have 1/8 and 1/16 sizes of the original input images, respectively. The segmentation loss L_{seg} is the sum of the binary cross-entropy (BCE) loss and Dice loss, shown as follows:

$$L_{seg}(\hat{y}, y) = \lambda \ell_{BCE} + (1 - \lambda)\ell_{Dice}, \tag{4}$$

where $\hat{y}$ and y are the output segmentation and the segmentation groundtruth, respectively. λ is the weight parameter.

4. Experiments

4.1. Datasets

Collection and use of human temporal lobes was approved by the Committee for Oversight of Research and Clinical Training Involving Decedents (CORID No.1063). Temporal lobes were obtained post-mortem from subjects who died of issues unrelated to the brain (e.g., septic shock or pancreatic ductal adenocarcinoma). The mean age was (males mean = 58 years, range 45–75 years; females mean = 50 years, range 21–72). The time to fixation after death was <42 hrs. Whole temporal lobes were immersed into 10% buffered formalin (CH_2O equivalent to 4% formaldehyde) for 4 weeks at 4 °C prior to transfer to PBS for 4 weeks.

Diffusion MR images were acquired on $9.4.7T/30$ cm Bruker AV3 HD microimaging scanner equipped with a B-GA12S HP gradient set capable of 660 mT/m maximum gradient strength and a 40 mm quadrature resonator running Paravision 6.0.1 (Bruker Biospin, Billerica, MA, USA).

Multi-shell diffusion MR images were acquired with a 3D diffusion-weighted multi-shot spin-echo EPI sequence with the following parameters: TR = 500 ms, TE = 0.96 ms, diffusion time = 14 ms, diffusion duration δ = 6.5 ms, diffusion spacing $\triangle$ = 13 ms, EPI segments = 30, with 1.2 partial Fourier acceleration in PE1, and a zero-filling acceleration factor 1.2 in the read and PE2 dimensions for a final isotropic resolution of 0.250 mm. A total of 94 images were collected with diffusion-weighted shells having b = 1000, 2000, 4000, 6000 s/mm^2 and 20, 30, 40, 60 directions, respectively ($\sim$68 h total scanning time), as described in detail in [26], Image acquisition was performed at room temperature (21 °C) to provide a high SNR, as well as good water diffusion [25]. Diffusion MR images were

processed using DSI Studio (available at https://dsi-studio.labsolver.org/ accessed on 17 March 2024) [54]. Reconstruction of the diffusion tensor images (DTI) was achieved by performing an eigenvector analysis on the calculated tensor [55,56]. Multi-shell diffusion MRI scans were reconstructed using Generalized Q-sampling Imaging (GQI) [57], with a diffusion sampling length ratio of 0.6 to yield image maps of mean diffusivity (MD), radial diffusivity (RD), axial diffusivity (AD), as well as fractional anisotropy (FA) images [26].

4.2. Implementation Details

Since our dataset consists of 10 3D multi-contrast diffusion MRI scans, we take 2 different scans each time as our validation set and the other 8 scans are utilized for training. We repeat this process 5 times to conduct five-fold cross-validations for all the experiments, and report the mean performance as well as the standard deviation (`std`). After the data partitions, we slice 3D diffusion MRI scans to 2D image slices along the Z-axis. In the training phase, we first applied data augmentation techniques on the fly to reduce potential overfitting, including random scaling (0.8 to 1.2), random rotation ($\pm 15°$), random intensity shift of (± 0.1), and intensity scaling of (0.9 to 1.1). Since the size of different image samples are different, we cropped or padded each image to a size of 512×512. The training iterations were set to 300 epochs with a linear warmup of the first 5 epochs. We trained the model using the Adam optimizer with a batch size of 32 and synchronized batch normalization. The initial learning rate was set to 1×10^{-3} and decayed by $(1 - \frac{current\_epoch}{max\_epoch})^{0.9}$. We also regularized the training with an l_2 weight decay of 1×10^{-5}.

In the inference phase, we only applied padding operations to the input image if its size can not be divisible by the down-sample rate of the model. All experiments were conducted based on Python 3.11.5 and PyTorch 1.7.1 and were deployed on a server with 2 NVIDIA A100 GPUs.

4.3. Baselines and Evaluation Metrics

We compared our approach with six segmentation baselines, i.e., U-Net [21], U^{2}Net [58], DeepLabv3+ [59], Attention U-Net [60], NNU-Net [61] and IVD-Net [62]. The U-Net model is a convolutional neural network initially developed for biomedical image segmentation. It features a distinctive architecture with a contracting path to capture context and a symmetric expanding path for precise localization, making it particularly effective for medical image tasks. The U^{2}Net model is a deep learning architecture that features a novel nested U-structure that enhances the learning of local and global contextual information. The Attention U-Net model is an advanced version of the traditional U-Net architecture used for medical image segmentation, which incorporates attention gates to enhance the ability of U-Net to focus on specific areas of interest. The DeepLabv3+ is an advanced semantic image segmentation model that builds upon the previous DeepLab frameworks. It introduces an encoder-decoder structure, utilizing atrous convolutions to efficiently capture multi-scale contextual information, and an improved Atrous Spatial Pyramid Pooling module, allowing it to effectively segment objects at multiple scales with enhanced boundary definition. The IVD-Net is a variant of U-Net for multi-contrast MRI data segmentation.

We adopted three metrics to assess the performance of segmentation models, including the mean intersection over union (mIoU), Dice similarity coefficient (DSC), Hausdorff distance (HD). Specifically, mIoU, and DSC are two overlap-based metrics, each ranging from 0 to 1 and a larger value indicating better performance. HD is a shape distance-based metric, which can be used to measure the dissimilarity between the surfaces/boundaries of the segmentation result and the ground-truth. As for HD, a lower value indicates a better segmentation result.

5. Results and Discussions

5.1. Comparative Experimental Results

Table 1 provides the performance of our framework and six competing baseline methods, including U-Net [21], DeepLabv3+ [59], U^{2}Net [58], Attention U-Net [60], NNU-

Net [61], and IVD-Net [62] on our multi-contrast diffusion MRI data for hippocampus segmentation. Since some baseline methods (i.e., U-Net, DeepLabv3+, U²Net, Attention U-Net and NNU-Net) are designed for the single contrast input, we train these models by joint dataset including 4 contrasts of image slices. It shows that our framework outperforms all competing methods substantially and consistently in terms of DSC and mIoU, indicating the superiority of our model in hippocampus segmentation. Meanwhile, comparing to the multi-contrast based method, i.e., IVD-Net, our model achieves superior segmentation results, which tends to show the strong ability of cross-attention for hippocampus segmentation based on multi-contrast diffusion MRI images. We visualize the qualitative segmentation results in Figure 3, where we present three multi-contrast image slices with their segmentation ground-truth. Meanwhile, the visualized segmentation predictions produced by our method and other two classic baselines (i.e., NNU-Net and U-Net) are also provided in Figure 3.

Table 1. Quantitative results of different methods on Hippocampus dataset. The best results are shown in **bold**. The DSC and mIoU are in percentage form. The HD is in mm.

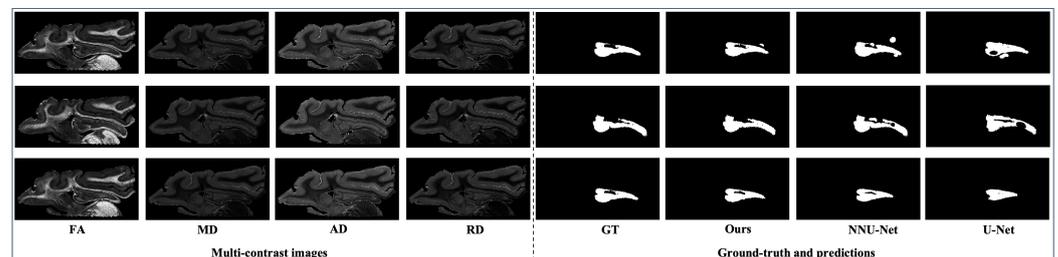| Methods	Multi-Contrast Diffusion MRI of Hippocampus		
	DSC	mIoU	HD
U-Net [21]	81.55 ± 0.95	83.94 ± 1.06	16.56 ± 0.82
DeepLabv3+ [59]	83.32 ± 1.49	83.62 ± 1.79	14.29 ± 1.16
U²Net [58]	83.67 ± 0.75	85.25 ± 0.68	11.30 ± 0.95
Attention U-Net [60]	85.77 ± 2.24	87.24 ± 1.26	8.64 ± 1.80
IVD-Net [62]	84.75 ± 1.61	85.91 ± 1.77	9.79 ± 1.04
NNU-Net [61]	88.19 ± 1.43	91.06 ± 0.77	6.69 ± 0.85
Ours	$\mathbf{89.74 \pm 1.32}$	$\mathbf{92.27 \pm 0.97}$	$\mathbf{6.26 \pm 1.11}$

Figure 3. Visualization of the hippocampus segmentation results based on multi-contrast MRI images produced by our model, as well as by NNU-Net and U-Net. GT represents ground-truth annotations.

5.2. Analytical Experimental Results

We conduct four different sets of analytical experiments. Firstly, we compare the hippocampus segmentation results by utilizing single-contrast and multi-contrast MRI scans to validate the importance of multi-contrast representations in the segmentation tasks. Meanwhile, we compare the segmentation results of our framework with and without cross-attention equipped to show the superiority of the attention mechanism. We also show the segmentation results of 4 single contrasts of diffusion MRI images (i.e., FA, MD, AD, RD) to compare the segmentation performance provided by each contrast of images. Finally, we perform a grid search experiment on the loss weight parameter (i.e., λ).

5.2.1. Comparisons: Single-Contrast vs. Multi-Contrast

In this experiment, we adopt the U-Net architecture as a foundation for segmentation tasks, utilizing fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD) as individual input contrasts. Given U-Net's absence of an attention mechanism, we deploy our proposed framework without attention mechanisms, utilizing these four contrasts as multi-contrast input. The comparison between segmentation results obtained from these two experimental settings underscores the compelling advantages of multi-contrast

images over single-contrast data. Table 2 serves as a testament to this, illustrating the notable improvements in hippocampus segmentation achieved through the integration of multiple contrasts. This enhancement is indicative of the complementary information offered by each contrast, collectively enriching the feature representation and enhancing the segmentation accuracy. Generally, the utilization of multi-contrast imaging holds significant promise in neuro-imaging research and clinical applications [63–65]. By leveraging the diverse information encapsulated within different contrasts, we can gain a more comprehensive understanding of the underlying tissue microstructure and pathology [34].

Table 2. Hippocampus segmentation results based on single contrast and multi-contrast diffusion MRI. wo/Attention represents without the attention mechanism. The best results are shown in **bold**.

Methods	Contrast Map	Evaluation Metrics		
		DSC	**mIoU**	**HD**
U-Net	FA	80.53 ± 1.20	82.21 ± 0.98	15.85 ± 1.79
	MD	79.20 ± 1.38	79.96 ± 1.66	16.26 ± 2.09
	AD	77.69 ± 2.62	79.55 ± 1.83	19.05 ± 1.13
	RD	78.02 ± 1.03	79.01 ± 0.84	17.04 ± 0.85
Ours wo/ Attention	All Modals	$\mathbf{87.72 \pm 1.29}$	$\mathbf{90.05 \pm 1.08}$	$\mathbf{8.52 \pm 1.06}$

5.2.2. Segmentation with Attention Mechanisms

To elucidate the significance of attention mechanisms, we conduct a comparative analysis between the conventional U-Net and a modified version incorporating self-attention (U-Net + Self-attention) (The self-attention is deployed in the third and fourth layer of the U-Net encoder). This comparison aims to highlight the impact of attention mechanisms on segmentation performance. Simultaneously, we explore the efficacy of attention mechanisms within the context of multi-contrast imaging. Employing our proposed framework, we conduct experiments comparing segmentation results obtained with and without cross-attention mechanisms. The outcomes of these comparative experiments are provided in Figure 4a, shedding light on the relative performance of attention-enabled frameworks. Our comprehensive analysis reveals a consistent trend across all comparative experiments: segmentation backbones equipped with attention mechanisms consistently outperform those without. This compelling evidence underscores the indispensable role of attention mechanisms in enhancing segmentation accuracy and robustness [36]. These findings carry profound implications for the field of medical image analysis, suggesting that attention mechanisms can serve as a powerful tool for improving the efficacy of segmentation algorithms.
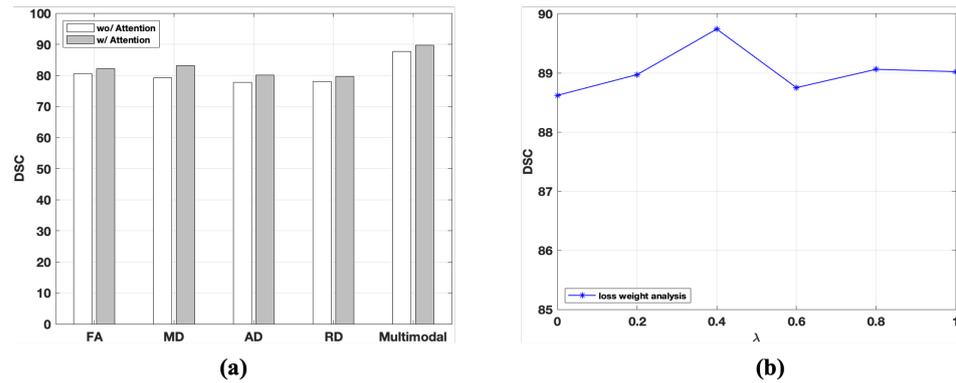
Figure 4. (**a**). Result comparisons: with attention mechanism and without attention mechanism. U-Net and U-Net with self-attention methods are compared based on four different data contrasts including FA, MD, AD, and RD. Our proposed method is compared with its variant (i.e., the proposed method without cross-attention) based on the multi-contrast data. wo/Attention and w/Attention represent without and with the attention mechanisms, respectively. (**b**). Loss weight analysis.

5.2.3. Comparisons of Single Contrast of Diffusion MRI

We select three baseline methods including U-Net, DeepLabv3+ and Attention U-Net to compare the segmentation results provided by each of four modals of diffusion MRI in Figure 5. The results suggest that FA images consistently exhibit superiority in hippocampus segmentation compared to other contrasts such as MD, AD, and RD. However, it is important to acknowledge that the superiority of MD, AD, and RD cannot be conclusively determined solely from our findings. This ambiguity arises from the observation that the performance of these contrasts varies depending on the specific deep learning framework employed. Therefore, the relative efficacy of MD, AD, and RD warrants further investigation within different methodological contexts to ascertain their respective strengths and limitations in hippocampus segmentation.
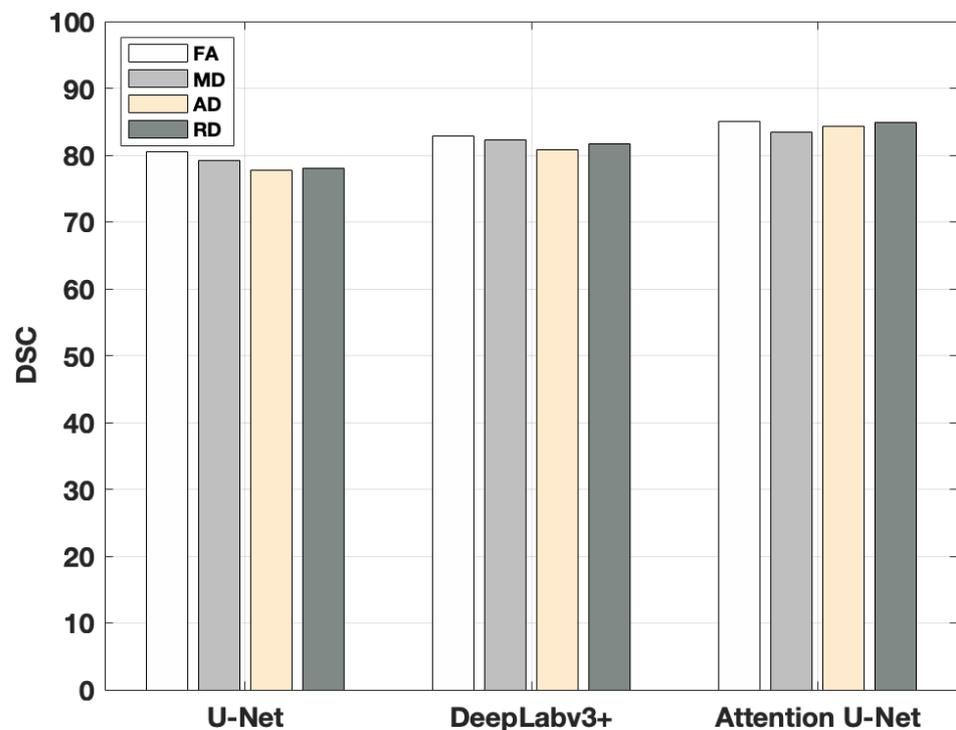

Figure 5. Result comparisons: four different contrasts deployed on U-Net, DeepLabv3+ and Attention U-Net.

5.2.4. Parameter Analysis

A grid search experiment is performed to determine the optimal value of the loss weight, λ. Particularly, we set the search space of λ as $\{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$. Figure 4b shows that the best value of the loss weight is 0.4, and it shows that the segmentation performance of our framework is consistent under different loss weights.

6. Conclusions

This paper presents a novel multimodal deep-learning framework incorporating cross-attention mechanisms tailored specifically for the segmentation of the hippocampus. Leveraging mesoscale diffusion MRI data, we harness its potential to compute a spectrum of contrast images crucial for segmenting the human hippocampus from surrounding brain structures. In our experimental evaluation, we assess the performance of our framework in hippocampus segmentation, demonstrating its efficacy in this segmentation task. Moreover, our experimental findings provide compelling evidence supporting the superiority of multi-contrast images over their single-contrast counterparts in diffusion MRI image segmentation. Furthermore, our experimental results shed light on the advantages offered by the fractional anisotropy (FA) contrast of dMRI images, where we observe that the FA contrast consistently performs better than the other image contrasts (i.e., MD, AD and RD) in the hippocampus segmentation task. This may potentially inform future developments in diffusion MRI analysis for hippocampus segmentation.

Author Contributions: H.T. took charge of conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, validation, visualization, writing—original draft, writing—review & editing. S.D. took charge of conceptualization, methodology, and writing—review & editing. E.M.Z. took charge of conceptualization, formal analysis, writing—review & editing. G.L. took charge of conceptualization, methodology, formal analysis, writing—review & editing. R.A. and R.K. took charge of data curation, formal analysis, investigation, resources, writing—review & editing. M.M. took charge of data curation, formal analysis, investigation, resources, supervision, writing—review & editing. L.Z. took charge of data curation, formal analysis, funding acquisition, investigation, resources, supervision, writing—review & editing. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data is available upon request.

Conflicts of Interest: The authors state that they have no conflicts of interest to declare.

References

1. O'Keefe, J.; Nadel, L. *The Hippocampus as a Cognitive Map*; Clarendon Press: Oxford, UK, 1978.
2. Scoville, W.B.; Milner, B. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* **1957**, *20*, 11. [CrossRef] [PubMed]
3. O'Keefe, J.; Dostrovsky, J. The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* **1971**, *34*, 171–175. [CrossRef]
4. Braak, H.; Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol.* **1991**, *82*, 239–259. [CrossRef]
5. Squire, L.R.; Stark, C.E.; Clark, R.E. The medial temporal lobe. *Annu. Rev. Neurosci.* **2004**, *27*, 279–306. [CrossRef] [PubMed]

6. Berdugo-Vega, G.; Arias-Gil, G.; López-Fernández, A.; Artegiani, B.; Wasielewska, J.M.; Lee, C.C.; Lippert, M.T.; Kempermann, G.; Takagaki, K.; Calegari, F. Increasing neurogenesis refines hippocampal activity rejuvenating navigational learning strategies and contextual memory throughout life. *Nat. Commun.* **2020**, *11*, 135. [CrossRef] [PubMed]

7. Topolnik, L.; Tamboli, S. The role of inhibitory circuits in hippocampal memory processing. *Nat. Rev. Neurosci.* **2022**, *23*, 476–492. [CrossRef]

8. Tzilivaki, A.; Tukker, J.J.; Maier, N.; Poirazi, P.; Sammons, R.P.; Schmitz, D. Hippocampal GABAergic interneurons and memory. *Neuron* **2023**, *111*, 3154–3175. [CrossRef]

9. Poh, J.H.; Vu, M.A.T.; Stanek, J.K.; Hsiung, A.; Egner, T.; Adcock, R.A. Hippocampal convergence during anticipatory midbrain activation promotes subsequent memory formation. *Nat. Commun.* **2022**, *13*, 6729. [CrossRef]

10. Petersen, R.C.; Aisen, P.S.; Beckett, L.A.; Donohue, M.C.; Gamst, A.C.; Harvey, D.J.; Jack, C.R.; Jagust, W.J.; Shaw, L.M.; Toga, A.W.; et al. Alzheimer's disease neuro-imaging initiative (ADNI): Clinical characterization. *Neurology* **2010**, *74*, 201–209. [CrossRef]

11. Frisoni, G.B.; Fox, N.C.; Jack Jr, C.R.; Scheltens, P.; Thompson, P.M. The clinical use of structural MRI in Alzheimer disease. *Nat. Rev. Neurol.* **2010**, *6*, 67–77. [CrossRef]

12. Duncan, J.S.; Winston, G.P.; Koepp, M.J.; Ourselin, S. Brain imaging in the assessment for epilepsy surgery. *Lancet Neurol.* **2016**, *15*, 420–433. [CrossRef]

13. Gondi, V.; Hermann, B.P.; Mehta, M.P.; Tomé, W.A. Hippocampal dosimetry predicts neurocognitive function impairment after fractionated stereotactic radiotherapy for benign or low-grade adult brain tumors. *Int. J. Radiat. Oncol. Biol. Phys.* **2012**, *83*, e487–e493. [CrossRef]

14. Moser, M.B.; Rowland, D.C.; Moser, E.I. Place cells, grid cells, and memory. *Cold Spring Harb. Perspect. Biol.* **2015**, *7*, a021808. [CrossRef]

15. Klöppel, S.; Peter, J.; Ludl, A.; Pilatus, A.; Maier, S.; Mader, I.; Heimbach, B.; Frings, L.; Egger, K.; Dukart, J.; et al. Applying automated MR-based diagnostic methods to the memory clinic: a prospective study. *J. Alzheimer's Dis.* **2015**, *47*, 939–954. [CrossRef]

16. Seiger, R.; Hammerle, F.P.; Godbersen, G.M.; Reed, M.B.; Spurny-Dworak, B.; Handschuh, P.; Klöbl, M.; Unterholzner, J.; Gryglewski, G.; Vanicek, T.; et al. Comparison and reliability of hippocampal subfield segmentations within FreeSurfer utilizing T1-and T2-weighted multispectral MRI data. *Front. Neurosci.* **2021**, *15*, 666000. [CrossRef]

17. Manjón, J.V.; Romero, J.E.; Coupe, P. A novel deep learning based hippocampus subfield segmentation method. *Sci. Rep.* **2022**, *12*, 1333. [CrossRef]

18. Samara, A.; Raji, C.; Li, Z.; Hershey, T. Comparison of Hippocampal Subfield Segmentation Agreement between 2 Automated Protocols across the Adult Life Span. *Am. J. Neuroradiol.* **2021**, *42*, 1783–1789. [CrossRef]

19. Kulaga-Yoskovitz, J.; Bernhardt, B.C.; Hong, S.J.; Mansi, T.; Liang, K.E.; Van Der Kouwe, A.J.; Smallwood, J.; Bernasconi, A.; Bernasconi, N. Multi-contrast submillimetric 3 Tesla hippocampal subfield segmentation protocol and dataset. *Sci. Data* **2015**, *2*, 1–9. [CrossRef]

20. Liu, M.; Li, F.; Yan, H.; Wang, K.; Ma, Y.; Shen, L.; Xu, M.; Initiative, A.D.N.; et al. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. *Neuroimage* **2020**, *208*, 116459. [CrossRef]

21. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.

22. Assaf, Y.; Pasternak, O. Diffusion tensor imaging (DTI)-based white matter mapping in brain research: A review. *J. Mol. Neurosci.* **2008**, *34*, 51–61. [CrossRef]

23. Qin, Y.Y.; Li, M.W.; Zhang, S.; Zhang, Y.; Zhao, L.Y.; Lei, H.; Oishi, K.; Zhu, W.Z. In vivo quantitative whole-brain diffusion tensor imaging analysis of APP/PS1 transgenic mice using voxel-based and atlas-based methods. *Neuroradiology* **2013**, *55*, 1027–1038. [CrossRef]

24. Van Essen, D.C.; Ugurbil, K.; Auerbach, E.; Barch, D.; Behrens, T.E.; Bucholz, R.; Chang, A.; Chen, L.; Corbetta, M.; Curtiss, S.W.; et al. The Human Connectome Project: A data acquisition perspective. *Neuroimage* **2012**, *62*, 2222–2231. [CrossRef]

25. Ly, M.; Foley, L.; Manivannan, A.; Hitchens, T.K.; Richardson, R.M.; Modo, M. Mesoscale diffusion magnetic resonance imaging of the ex vivo human hippocampus. *Hum. Brain Mapp.* **2020**, *41*, 4200–4218. [CrossRef]

26. Modo, M.; Sparling, K.; Novotny, J.; Perry, N.; Foley, L.M.; Hitchens, T.K. Mapping mesoscale connectivity within the human hippocampus. *Neuroimage* **2023**, *282*, 120406. [CrossRef]

27. Jones, D.K. *Diffusion MRI: Theory, Methods, and Applications*; Oxford University Press: Oxford, UK , 2010. [CrossRef]

28. Modo, M.; Hitchens, T.K.; Liu, J.R.; Richardson, R.M. *Detection of Aberrant Hippocampal Mossy Fiber Connections: Ex Vivo Mesoscale Diffusion MRI and Microtractography with Histological Validation in a Patient with Uncontrolled Temporal Lobe Epilepsy*; Technical report; Wiley Online Library: New York, NY, USA, 2016.

29. Ke, J.; Foley, L.M.; Hitchens, T.K.; Richardson, R.M.; Modo, M. Ex vivo mesoscopic diffusion MRI correlates with seizure frequency in patients with uncontrolled mesial temporal lobe epilepsy. *Hum. Brain Mapp.* **2020**, *41*, 4529–4548. [CrossRef]

30. Hett, K.; Ta, V.T.; Catheline, G.; Tourdias, T.; Manjón, J.V.; Coupé, P. Multimodal hippocampal subfield grading for Alzheimer's disease classification. *Sci. Rep.* **2019**, *9*, 13845. [CrossRef]

31. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [CrossRef]

32. Jia, H.; Tang, H.; Ma, G.; Cai, W.; Huang, H.; Zhan, L.; Xia, Y. A convolutional neural network with pixel-wise sparse graph reasoning for COVID-19 lesion segmentation in CT images. *Comput. Biol. Med.* **2023**, *155*, 106698. [CrossRef]

33. Fu, X.; Sun, Z.; Tang, H.; Zou, E.M.; Huang, H.; Wang, Y.; Zhan, L. 3D bi-directional transformer U-Net for medical image segmentation. *Front. Big Data* **2023**, *5*, 1080715. [CrossRef]

34. Dai, S.; Ye, K.; Zhao, K.; Cui, G.; Tang, H.; Zhan, L. Constrained Multiview Representation for Self-supervised Contrastive Learning. *arXiv* **2024**, arXiv:2402.03456.

35. Menze, B.H.; Jakab, A.; Bauer, S.; Kalpathy-Cramer, J.; Farahani, K.; Kirby, J.; Burren, Y.; Porz, N.; Slotboom, J.; Wiest, R.; et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* **2014**, *34*, 1993–2024. [CrossRef]

36. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.

37. Fang, F.; Zhou, T.; Song, Z.; Lu, J. MMCAN: Multi-Modal Cross-Attention Network for Free-Space Detection with Uncalibrated Hyperspectral Sensors. *Remote Sens.* **2023**, *15*, 1142. [CrossRef]

38. Duvernoy, H.M.; Cattin, F.; Risold, P.Y.; Salvolini, U.; Scarabine, U. *The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI*; Number 3; Springer: Berlin/Heidelberg, Germany, 2005.

39. Fel, J.; Ellis, C.; Turk-Browne, N. Automated and manual segmentation of the hippocampus in human infants. *Dev. Cogn. Neurosci.* **2023**, *60*, 101203. [CrossRef]

40. Dalton, M.A.; Zeidman, P.; Barry, D.N.; Williams, E.; Maguire, E.A. Segmenting subregions of the human hippocampus on structural magnetic resonance image scans: An illustrated tutorial. *Brain Neurosci. Adv.* **2017**, *1*, 2398212817701448. [CrossRef]

41. Carmichael, O.T.; Aizenstein, H.A.; Davis, S.W.; Becker, J.T.; Thompson, P.M.; Meltzer, C.C.; Liu, Y. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. *Neuroimage* **2005**, *27*, 979–990. [CrossRef]

42. Nobakht, S.; Schaeffer, M.; Forkert, N.D.; Nestor, S.; Black, S.E.; Barber, P.; Initiative, A.D.N. Combined atlas and convolutional neural network-based segmentation of the hippocampus from MRI according to the ADNI harmonized protocol. *Sensors* **2021**, *21*, 2427. [CrossRef]

43. Plassard, A.J.; McHugo, M.; Heckers, S.; Landman, B.A. Multi-scale hippocampal parcellation improves atlas-based segmentation accuracy. In Proceedings of the Medical Imaging 2017: Image Processing, Orlando, FL, USA, 11–16 February 2017 ; SPIE: Bellingham, WA, USA, 2017; Volume 10133, pp. 666–672.

44. Chupin, M.; Gérardin, E.; Cuingnet, R.; Boutet, C.; Lemieux, L.; Lehéricy, S.; Benali, H.; Garnero, L.; Colliot, O. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus* **2009**, *19*, 579–587. [CrossRef]

45. Morra, J.H.; Tu, Z.; Apostolova, L.G.; Green, A.E.; Avedissian, C.; Madsen, S.K.; Parikshak, N.; Toga, A.W.; Jack Jr, C.R.; Schuff, N.; et al. Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. *Neuroimage* **2009**, *45*, S3–S15. [CrossRef]

46. Henschel, L.; Conjeti, S.; Estrada, S.; Diers, K.; Fischl, B.; Reuter, M. Fastsurfer-a fast and accurate deep learning based neuro-imaging pipeline. *NeuroImage* **2020**, *219*, 117012. [CrossRef]

47. Tang, H.; Ma, G.; Guo, L.; Fu, X.; Huang, H.; Zhan, L. Contrastive brain network learning via hierarchical signed graph pooling model. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**. [CrossRef]

48. Kim, M.; Wu, G.; Shen, D. Unsupervised deep learning for hippocampus segmentation in 7.0 Tesla MR images. In Proceedings of the Machine Learning in Medical Imaging: 4th International Workshop, MLMI 2013, Held in Conjunction with MICCAI 2013, Nagoya, Japan, 22 September 2013; Proceedings 4; Springer: Berlin/Heidelberg, Germany, 2013; pp. 1–8.

49. Zheng, Q.; Liu, B.; Gao, Y.; Bai, L.; Cheng, Y.; Li, H. HGM-cNet: Integrating hippocampal gray matter probability map into a cascaded deep learning framework improves hippocampus segmentation. *Eur. J. Radiol.* **2023**, *162*, 110771. [CrossRef]

50. Liu, Y.; Yan, Z. A combined deep-learning and lattice Boltzmann model for segmentation of the hippocampus in MRI. *Sensors* **2020**, *20*, 3628. [CrossRef] [PubMed]

51. Wang, H.; Lei, C.; Zhao, D.; Gao, L.; Gao, J. DeepHipp: Accurate segmentation of hippocampus using 3D dense-block based on attention mechanism. *BMC Med. Imaging* **2023**, *23*, 158. [CrossRef]

52. Wu, Z.; Gao, Y.; Shi, F.; Ma, G.; Jewells, V.; Shen, D. Segmenting hippocampal subfields from 3T MRI with multi-modality images. *Med. Image Anal.* **2018**, *43*, 10–22. [CrossRef] [PubMed]

53. Jiang, H.; Guo, Y. Multi-class multimodal semantic segmentation with an improved 3D fully convolutional networks. *Neurocomputing* **2020**, *391*, 220–226. [CrossRef]

54. Yeh, F.C.; Verstynen, T.D.; Wang, Y.; Fernández-Miranda, J.C.; Tseng, W.Y.I. Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS ONE* **2013**, *8*, e80713. [CrossRef] [PubMed]

55. Jiang, H.; Van Zijl, P.C.; Kim, J.; Pearlson, G.D.; Mori, S. DtiStudio: Resource program for diffusion tensor computation and fiber bundle tracking. *Comput. Methods Programs Biomed.* **2006**, *81*, 106–116. [CrossRef] [PubMed]

56. Zhan, L.; Leow, A.D.; Jahanshad, N.; Chiang, M.C.; Barysheva, M.; Lee, A.D.; Toga, A.W.; McMahon, K.L.; De Zubicaray, G.I.; Wright, M.J.; et al. How does angular resolution affect diffusion imaging measures? *Neuroimage* **2010**, *49*, 1357–1371. [CrossRef]

57. Jin, Z.; Bao, Y.; Wang, Y.; Li, Z.; Zheng, X.; Long, S.; Wang, Y. Differences between generalized Q-sampling imaging and diffusion tensor imaging in visualization of crossing neural fibers in the brain. *Surg. Radiol. Anat.* **2019**, *41*, 1019–1028. [CrossRef]

58. Qin, X.; Zhang, Z.; Huang, C.; Dehghan, M.; Zaiane, O.R.; Jagersand, M. U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognit.* **2020**, *106*, 107404. [CrossRef]

59. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.

60. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B.; et al. Attention U-Net: Learning where to look for the pancreas. In Proceedings of the Medical Imaging with Deep Learning (MIDL), Amsterdam, The Netherlands, 4–6 July 2018.

61. Isensee, F.; Jaeger, P.F.; Kohl, S.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **2021**, *18*, 203–211. [CrossRef] [PubMed]

62. Dolz, J.; Desrosiers, C.; Ben Ayed, I. IVD-Net: Intervertebral disc localization and segmentation in MRI with a multi-modal UNet. In *Proceedings of the International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 130–143.

63. Ye, K.; Tang, H.; Dai, S.; Guo, L.; Liu, J.Y.; Wang, Y.; Leow, A.; Thompson, P.M.; Huang, H.; Zhan, L. Bidirectional mapping with contrastive learning on multimodal neuro-imaging data. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Vancouver, ON, Canada, 8–12 October 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 138–148.

64. Guo, L.; Zhang, Y.; Tang, H.; Mackin, S.R.; Thompson, P.M.; Huang, H.; Zhan, L. Investigating the effect of neuropsychiatric symptoms on Alzheimer's diagnosis using multi-modal brain networks. *Alzheimer's Dement.* **2023**, *19*, e080376. [CrossRef]

65. Tang, H.; Guo, L.; Fu, X.; Wang, Y.; Mackin, S.; Ajilore, O.; Leow, A.D.; Thompson, P.M.; Huang, H.; Zhan, L. Signed graph representation learning for functional-to-structural brain network mapping. *Med. Image Anal.* **2023**, *83*, 102674. [CrossRef] [PubMed]

66. Brown, S.T.; Buitrago, P.; Hanna, E.; Sanielevici, S.; Scibek, R.; Nystrom, N.A. Bridges-2: A platform for rapidly-evolving and data intensive research. In *Practice and Experience in Advanced Research Computing*; ACM: New York, NY, USA, 2021; pp. 1–4.