Triplanar CNN Approach for Hippocampus Segmentation and Depression Classification

Andrea C. Salazar-Zozaya
Tecnológico de Monterrey,
School of Engineering and Sciences, Mexico
andiieszozaya@hotmail.com

José Antonio Cantoral-Ceballos
Tecnológico de Monterrey,
School of Engineering and Sciences, Mexico
joseantonio.cantoral@tec.mx

Ricardo Caraza-Camacho
Centro de Neurociencias Cognitivas, México
Tecnológico de Monterrey,
Escuela de Medicina y Ciencias de la Salud, México
rcaraza@tec.mx

Abstract

Major Depressive Disorder (MDD) is the leading cause of disability in the world, affecting approximately 280 million people. Hippocampal volumetric changes have been proposed as a potential biomarker for depression. Despite advancements in neuroimaging studies related to psychiatric disorders, there remains a gap in the utilization of these findings for clinical diagnosis and monitoring of such disorders. This study presents a comprehensive investigation of Major Depressive Disorder (MDD) stage differentiation using MRI data and a U-Net architecture for hippocampal segmentation across axial, coronal, and sagittal orientations. Preliminary results on the CC-359 and validation datasets demonstrate promising segmentation performance, especially in axial and coronal orientations. Future work will include volumetric calculations of the hippocampus for MDD classification according to BDI-II, and the improvement of the model's performance in terms of generalization, ultimately contributing to the advancement of neuroimaging research in MDD disorder.

1. Introduction

Major Depressive Disorder (MDD) is a prevalent psychiatric condition that affects millions of individuals worldwide [1], and includes a wide range of clinical symptoms such as affective, cognitive, and somatic signs [10]. MDD is currently the leading cause of disability worldwide, affecting twice as many women than men and one out of six people at some point in their lives [16].

Imaging studies such as magnetic resonance imaging (MRI), structural resonance imaging (sMRI), and functional magnetic resonance imaging (fMRI) have emerged as promising techniques to outline biomarkers accompanying MDD. Limbic structures such as the hippocampus and the amygdala have emerged as potential biomarkers for the identification of MDD due to their significant role in the neuronal circuitry controlling this disorder [5, 18] and the volumetric changes reported in MRI studies. It has been shown that when MDD is severe, the hippocampus volume decreases [5, 20, 21].

Medical image segmentation for volumetric analysis has shown promising results in improving diagnostic and treatment methods. Manual segmentation is considered the "gold standard", but it can be time-consuming, subjective, and unsuitable for large-scale neuroimaging research [22]. Deep learning and Convolutional Neural Networks (CNN) models have emerged as an effective and accurate method for automatic image segmentation. However, there are still several challenges, such as repeatability, training, computational cost, and poor performance on unseen datasets [22]. Despite the challenges, automated approaches have shown an acceptable correlation with volume but poor absolute approximation for small structures such as the hippocampus [2].

FreeSurfer [5, 20], ITK-SNAP [18], AccuBrain [32], and SACHA software [32] are some of the most widely used software tools for fully-automatic, semi-automatic, or manual segmentation of the amygdala and/or hippocampus. Several new models based on Convolutional Neural Networks (CNN) have also been published, including CAST, a multiscale Convolutional neural network-based Automated hippocampus subfield Segmentation Toolbox

[31], and HippMapp3r, a 3D CNN-based algorithm for hippocampal segmentation in brains with extensive atrophy [9]. These tools, however, have yet to outperform manual segmentation.

Despite rapid advances in neuroimaging technology and major research funding, there is still a lack of widely used clinical neuroimaging studies for the diagnosis or monitoring of neurodevelopmental or neuropsychiatric diseases to help physicians and patients. To overcome this difficulty, it is critical to find brain function patterns that can predict symptoms or behaviors in specific patients, improving our understanding of the complexities of mental diseases. Additionally, integrating multiple types of data, such as genetic, behavioral, imaging, and clinical data, could provide a more comprehensive understanding of psychiatric disorders [4].

This work uses a thorough methodology to examine brain imaging data in the context of depression differentiation. A U-Net algorithm is used to examine input volumes in three orientations (axial, sagittal, and coronal) to capture rich spatial information inherent in each plane. 2D slices extracted from the input volumes are used to train a binary hippocampus segmentation model before being integrated into a 3D volume. Volume analysis of the hippocampus was performed to analyze the inequalities related to depression. Finally, a classifier was used to differentiate the severity of patients based on BDI-II scale and the hippocampal volume.

2. Related Work

Recent advancements in deep learning, including Convolutional Neural Networks (CNN), have shown high efficiency and accuracy across various image-related tasks such as classification [17, 23, 30] and segmentation [6, 13, 29]. In the neuroimaging field, these tools have been used for the segmentation of brain structures [19], including limbic structures, such as the hippocampus [8, 9, 12, 28, 31], which plays a vital role in several brain diseases and psychiatric disorders. This section highlights several key segmentation models for these structures and their contributions.

Efforts to improve hippocampal segmentation resulted in the development of numerous noticeable models. One such model is CAST (Automated hippocampal subfield Segmentation Toolbox), which is a 3D multiscale CNN model for hippocampal subfield automatic segmentation introduced by Yang *et al*. [31]. CAST is notable for its comparable accuracy to state-of-the-art models in terms of dice coefficient and also in enhancing reliability in segmenting small subfields, such as CA2, CA3, and the entorhinal cortex (ERC), as demonstrated by improved intraclass correlation coefficients (ICC).

Another notable work was presented by Thyreau *et al*. [28], who proposed HippoDeep, a hippocampal segmentation model that integrates innovations such as a large and variable training set derived from multiple cohorts, training labels derived in part from the FreeSurfer algorithm output, synthetic data, and a powerful data augmentation scheme were included. Notably, HippoDeep features fast inference times and is freely available online, supporting widespread implementation and use.

Carmo et al. also, introduced a novel method for volumetric hippocampus segmentation using a tri-planar U-Net inspired fully convolutional networks (FCNNs) [7]. By integrating improvements such as residual connections, VGG weight transfer, batch normalization, and patch extraction algorithms, they demonstrated a 96% volumetric dice accuracy in their test data.

In addition to developing hippocampus segmentation techniques, researchers have used these models to diagnose brain disorders. HippMapp3r, developed by Goubran *et al*. [9] is a 3D CNN-based algorithm created for the detection of brain atrophy and lesions associated with aging and neurodegeneration in the human brain. Trained using 259 manually performed segmentations collected from three different experimental setups, HippMapp3r, outperformed five publicly available algorithms (HippoDeep, FreeSurfer, SBHV, volBrain, and FIRST), with an average Dice of 0.89 and a correlation value of 0.95.

Furthermore, using a T1 weighted structural MRI data, Liu et al. [12] proposed a multimodel deep learning framework based on CNN for automatic hippocampal segmentation and Alzheimer's Disease (AD) detection. In this model, a 3D Dense Net was built to learn the characteristics of the 3D patches generated based on the hippocampus segmentation findings for classification. This model classified Alzheimer's disease (AD) and normal control (NC) patients with an accuracy of 88.9%.

These segmentation models have significantly contributed to a better understanding of the hippocampus through their accurate and robust delineation of the structure [3, 8, 10, 14, 15, 33, 34]. These studies address various challenges related to the delimitation of 3D brain structures, brain structure variability, atrophy lesions, and small subfield segmentation. As a result of the implementation of these models, the diagnosis, follow-up, and monitoring of psychiatric disorders and brain diseases. It is expected that deep learning techniques will continue to advance in the future, allowing more efficient segmentation methods for brain structures to create new studies and clinical applications.

3. Materials & Methods

3.1. Data

Three datasets were utilized for the training and evaluation of this model. The Calgary-Campinas Public Dataset and the Validation Datasets were used for training and testing

the U-Net model and the SRPBS_Open dataset is used for the classification of MDD severity stages.

3.1.1 Calgary-Campinas Public Dataset (CC-359)

The CC-359 dataset [25] includes T1-weighted volumes acquired from 359 subjects using MRI scanners manufactured by General Electric (GE), Philips, and Siemens. Additionally, the dataset includes scans acquired at two distinct magnetic field strengths, namely 1.5 Tesla (T) and 3 Tesla (T).

3.1.2 Validation Dataset

The validation dataset was obtained from Jafari et *al.* [11]. This dataset contains 50 t1-weighted MRI scans from epileptic and nonepileptic subjects, with manually generated hippocampal segmentation masks.

3.1.3 SRPBS_Open

The SRPBS Multi-disorder MRI Dataset [27] includes resting-state fMRI EPI images, defaced t1w images, and optional fieldmap data for 1,410 subjects suffering from a variety of disorders, including MDD. It also contains demographic data and thorough clinical rating scales for each illness.

3.2. Methodology

The methodology of this work was inspired by Carmo et *al.* [7] and consists of a comprehensive analysis of input volumes in all three orientations (axial, sagittal, and coronal) using CNNs trained on patches specific to each orientation. Binary hippocampus segmentations are conducted on each 2D slice extracted from the input volume and then concatenated into a single volumetric representation. Subsequently, the volume of the hippocampus is calculated and then fed into a classifier that differentiates between four stages of severity of MDD according to BDI-II stages: minimal, mild, moderate, and severe.

3.3. Architecture

The architecture for this model consists of a 2D U-Net. The architecture comprises encoder and decoder blocks. The encoder blocks comprise convolutional layers, which are downsampled via max pooling. The U-Net model's encoder blocks reduce spatial dimensions while increasing the feature map's depth. In contrast, decoder blocks use transposed convolutional layers for upsampling and skip connections to preserve spatial information. Batch normalization, dropout, and concatenation with skip connections ensure that features propagate and integrate properly during decoding. Also, the U-Net's bottleneck layer acts as a feature representation central point, allowing for the extraction of abstract features from input images.

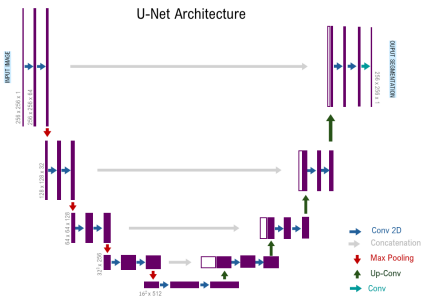


Figure 1. U-Net Architecture.

The model is constructed by connecting encoder and decoder blocks, with the encoder gradually reducing spatial dimensions while increasing the depth of feature maps, and the decoder upsampling and integrating those features. The final output layer uses transposed convolutional layers with a sigmoid activation function, which is appropriate for binary segmentation applications. This highlights the U-Net's capacity to capture hierarchical features effectively, making it suited for diverse image segmentation applications, such as medical image analysis. "Fig. 1" shows the architecture of this model.

3.4. Training

The model was trained separately on axial, sagittal, and coronal MRI images. Each dataset was augmented and batched to facilitate efficient training. The learning phase of the model was led by Leaky ReLu activation layers, which used the Adam optimizer in conjunction with a binary cross-entropy loss function. The dataset was divided into training and validation sets, with a validation split of 20%.

3.5. Volume Calculation

After training the CNN and concatenating the resulting binary mask slices into a single volume, the volume is calculated by summing the volumes of the segmented voxels [24]. The volume, V, of an object represented by a set of voxels, S, is computed as follows:

$$V = \sum_{u \epsilon S} RowSpacing(u) \times ColSpacing(u) \times SliceSpacing(u) \tag{1}$$

where *RowSpacing(u)* is the distance between the center of voxel and the adjacent voxel in its row, and *ColSpacing(u)* and *SliceSpacing(u)* are similar distances in the column and z-direction, respectively.

3.6. Evaluation Metrics

For the analysis of Image Segmentation, valid metrics are needed. Dice coefficient (DICE) [26], also known as over-

Dataset	CC359 Dataset	Validation Dataset
Axial Plane	0.8997	0.7547
Coronal Plane	0.8854	0.7167
Sagittal Plane	0.8173	0.7382

Table 1. DSC values for the CC-359 Dataset and an independent validation dataset

lapping was used to measure reproducibility, this metric is defined by

$$DICE = \frac{2\left|S_g^1 \bigcap S_t^1\right|}{\left|S_g^1\right| + \left|S_t^1\right|} = \frac{2TP}{2TP + FP + FN}, \quad (2)$$

where TP is the true positive values, FP false positives, and FN are the false negatives.

4. Results & Discussion

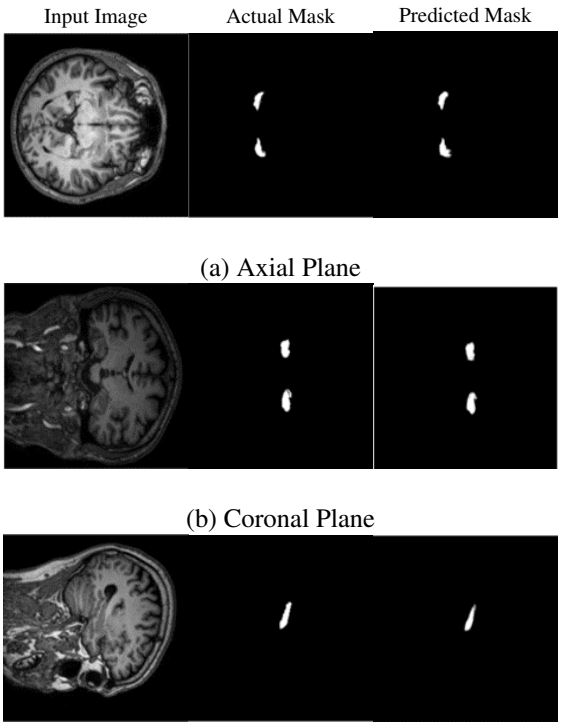
The proposed method's segmentation performance was evaluated on a test set from the CC-359 Dataset and a validation dataset using the Dice Score Coefficient (DSC). The DSC obtained is presented in Table 1.

The model showed promising segmentation results on the CC-359 dataset from the three orientations. The axial orientation achieved a DSC of 0.8997 showing that brain structures were accurately segmented. Similarly, the model achieved a DSC of 0.8854 in the coronal orientation, showcasing its effectiveness in segmenting this brain structure. However, a slight decrease in performance was observed in the sagittal orientation, which obtained a DSC of 0.8173. Figure 2 shows some predicted masks from this dataset.

The model's performance remained positive after validation with an independent dataset, despite some somewhat lower DSC. This could be caused by some of the dataset's patients suffering from temporal lobe epilepsy, which can produce atrophic hippocampi and alterations in structure and volume.

Comparing our results with existing methods, this approach exhibits competitive performance. Notably, CAST [31] achieved Dice similarity coefficients (DSC) of 0.80 for the left hippocampus and 0.78 for the right hippocampus, while HippoDeep achieved a DSC of 0.85 and HippMapp3r reached 0.89. These comparison results demonstrate the effectiveness of our proposed methodology in hippocampus segmentation.

This work emphasizes the benefits of accounting for different imaging viewpoints in brain structure segmentation tasks, with axial and coronal views showing better performance due to richer anatomical details and clearer boundaries. However, issues with the sagittal view might be due to image acquisition complications or anatomical variances in the images. Differences between the CC-359 dataset and the validation dataset highlight the importance of data diversity, especially including data from subjects with alterations in the hippocampus structure.



(a) Axial Plane

(b) Coronal Plane

(c) Sagittal Plane

Figure 2. Three sample images showcasing input images, actual masks, and predicted masks across axial (a), coronal (b), and sagittal (c) planes. Data sourced from the validation set obtained from the CC-359 Dataset.

5. Conclusion

This study presents a throughout investigation into brain imaging data concerning Major Depressive Disorder stage differentiation, using a U-Net architecture to segment the hippocampus across axial, coronal, and sagittal orientations. The models achieved promising results on both the CC-359 and the validation dataset using 2D slice analysis. The model was particularly effective in axial and coronal segmentations, demonstrating its ability to capture detailed anatomical information. Although there was a minor decrease in the performance of the sagittal segmentations, the model's robustness highlights its potential use in MDD diagnosis. It is important to note that these findings represent preliminary results, and future work will include volume calculations of the hippocampus to differentiate MDD stages. Additionally, further research is warranted to optimize the model's performance by addressing the challenges regarding dataset diversity and generalization. These ongoing efforts aim to improve the model's clinical application

and help advance neuroimaging studies in psychiatric diseases.

References

[1] Depression. https://www.who.int/news-room/fact-sheets/detail/depression, 2012. Last accessed October 2022. 1

[2] Theophilus N Akudjedu, Leila Nabulsi, Migle Makelyte, Cathy Scanlon, Sarah Hehir, Helen Casey, Srinath Ambati, Joanne Kenney, Stefani O'Donoghue, Emma McDermott, et al. A comparative study of segmentation techniques for the quantification of brain subcortical volume. *Brain imaging and behavior*, 12(6):1678–1695, 2018. 1

[3] Bonnie Alexander, Nellie Georgiou-Karistianis, Richard Beare, Lotta M Ahveninen, Valentina Lorenzetti, Julie C Stout, and Yifat Glikmann-Johnston. Accuracy of automated amygdala mri segmentation approaches in huntington's disease in the image-hd cohort. *Human brain mapping*, 41(7): 1875–1888, 2020. 2

[4] Ashley N Anderson, Jace B King, and Jeffrey S Anderson. Neuroimaging in psychiatry and neurodevelopment: why the emperor has no clothes. *The British Journal of Radiology*, 92 (1101):20180910, 2019. 2

[5] Stephanie SG Brown, John W Rutland, Gaurav Verma, RE Feldman, Judy Alper, Molly Schneider, Bradley N Delman, JM Murrough, and Priti Balchandani. Structural mri at 7t reveals amygdala nuclei and hippocampal subfield volumetric association with major depressive disorder symptom severity. *Scientific reports*, 9(1):1–10, 2019. 1

[6] Joseph Bullock, Carolina Cuesta-Lázaro, and Arnau Quera-Bofarull. Xnet: a convolutional neural network (cnn) implementation for medical x-ray image segmentation suitable for small datasets. In *Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*, pages 453–463. SPIE, 2019. 2

[7] Diedre Carmo, Bruna Silva, Clarissa Yasuda, Letícia Rittner, and Roberto Lotufo. Extended 2d consensus hippocampus segmentation. *arXiv preprint arXiv:1902.04487*, 2019. 2, 3

[8] Diedre Carmo, Bruna Silva, Clarissa Yasuda, Letícia Rittner, and Roberto Lotufo. Hippocampus segmentation on epilepsy and alzheimer's disease studies with multiple convolutional neural networks. *Heliyon*, 7(2), 2021. 2

[9] Maged Goubran, Emmanuel Edward Ntiri, Hassan Akhavein, Melissa Holmes, Sean Nestor, Joel Ramirez, Sabrina Adamo, Miracle Ozzoude, Christopher Scott, Fuqiang Gao, et al. Hippocampal segmentation for brains with extensive atrophy using three-dimensional convolutional neural networks. Technical report, Wiley Online Library, 2020. 2

[10] Yael Jacob, Laurel S Morris, Gaurav Verma, Sarah B Rutter, Priti Balchandani, and James W Murrough. Altered hippocampus and amygdala subregion connectome hierarchy in major depressive disorder. *Translational psychiatry*, 12(1): 209, 2022. 1, 2

[11] Kourosh Jafari-Khouzani, Kost V Elisevich, Suresh Patel, and Hamid Soltanian-Zadeh. Dataset of magnetic resonance images of nonepileptic subjects and temporal lobe epilepsy patients for validation of hippocampal segmentation techniques. *Neuroinformatics*, 9:335–346, 2011. 3

[12] Manhua Liu, Fan Li, Hao Yan, Kundong Wang, Yixin Ma, Li Shen, Mingqing Xu, Alzheimer's Disease Neuroimaging Initiative, et al. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease. *Neuroimage*, 208:116459, 2020. 2

[13] Ange Lou, Shuyue Guan, and Murray Loew. Dc-unet: rethinking the u-net architecture with dual channel efficient cnn for medical image segmentation. In *Medical Imaging 2021: Image Processing*, pages 758–768. SPIE, 2021. 2

[14] Nikita Nogovitsyn, Roberto Souza, Meghan Muller, Amelia Srajer, Stefanie Hassel, Stephen R Arnott, Andrew D Davis, Geoffrey B Hall, Jacqueline K Harris, Mojdeh Zamyadi, et al. Testing a deep convolutional neural network for automated hippocampus segmentation in a longitudinal sample of healthy participants. *NeuroImage*, 197:589–597, 2019. 2

[15] Assaf Oshri, Joshua C Gray, Max M Owens, Sihong Liu, Erinn Bernstein Duprey, Lawrence H Sweet, and James MacKillop. Adverse childhood experiences and amygdalar reduction: high-resolution segmentation reveals associations with subnuclei and psychiatric outcomes. *Child maltreatment*, 24(4):400–410, 2019. 2

[16] Christian Otte, Stefan M Gold, Brenda W Penninx, Carmine M Pariante, Amit Etkin, Maurizio Fava, David C Mohr, and Alan F Schatzberg. Major depressive disorder. *Nature reviews Disease primers*, 2(1):1–20, 2016. 1

[17] Şaban Öztürk and Umut Özkaya. Residual lstm layered cnn for classification of gastrointestinal tract diseases. *Journal of Biomedical Informatics*, 113:103638, 2021. 2

[18] Koen van Oijen Mark Rijpkema Bozena Goraj Robbert Jan Verkes Richard Oude Voshaar Guillén Fernández Jan Buitelaar Indira Tendolkar Philip van Eijndhoven, Guido van Wingen. Amygdala volume marks the acute state in the early course of depression. *Biological Psychiatry*, 65(9):812–818, 2009. 1

[19] Farheen Ramzan, Muhammad Usman Ghani Khan, Sajid Iqbal, Tanzila Saba, and Amjad Rehman. Volumetric segmentation of brain regions from mri scans using 3d convolutional neural networks. *IEEE Access*, 8:103697–103709, 2020. 2

[20] Darren W Roddy, Chloe Farrell, Kelly Doolin, Elena Roman, Leonardo Tozzi, Thomas Frodl, Veronica O'Keane, and Erik O'Hanlon. The hippocampus in depression: more than the sum of its parts? advanced hippocampal substructure segmentation in depression. *Biological Psychiatry*, 85(6):487–497, 2019. 1

[21] Claire Cury Christophe Vandendrie Florence Gressier Patrick Hardy Bruno Falissard Olivier Colliot Denis Ducreux Emmanuelle Corruble Romain Colle, Marie Chupin. Depressed suicide attempters have smaller hippocampus than depressed patients without suicide attempts. *Journal of Psychiatric Research*, 61:13–18, 2015. 1

[22] Mahender Kumar Singh and Krishna Kumar Singh. A review of publicly available automatic brain segmentation methodologies, machine learning models, recent advancements, and

their comparison. *Annals of Neurosciences*, 28(1-2):82–93, 2021. 1

[23] Satya P Singh, Lipo Wang, Sukrit Gupta, Balazs Gulyas, and Parasuraman Padmanabhan. Shallow 3d cnn for detecting acute brain hemorrhage from medical imaging sensors. *IEEE Sensors Journal*, 21(13):14290–14299, 2020. 2

[24] Milan Sonka and J. Michael Fitzpatrick. *Handbook of Medical Imaging*. SPIE Press, 2009. 3

[25] Roberto Souza, Oeslle Lucena, Julia Garrafa, David Gobbi, Marina Saluzzi, Simone Appenzeller, Letícia Rittner, Richard Frayne, and Roberto Lotufo. An open, multi-vendor, multi-field-strength brain mr dataset and analysis of publicly available skull stripping methods agreement. *NeuroImage*, 170:482–494, 2018. 3

[26] Abdel Aziz Taha and Allan Hanbury. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, 15(1):1–28, 2015. 3

[27] S.C. Tanaka, A. Yamashita, N. Yahata, et al. A multi-site, multi-disorder resting-state magnetic resonance image database. *Sci Data*, 8:227, 2021. 3

[28] Benjamin Thyreau, Kazunori Sato, Hiroshi Fukuda, and Yasuyuki Taki. Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing. *Medical image analysis*, 43:214–228, 2018. 2

[29] Kuo-Kun Tseng, Ran Zhang, Chien-Ming Chen, and Mohammad Mehedi Hassan. Dnetunet: a semi-supervised cnn of medical image segmentation for super-computing ai service. *The Journal of Supercomputing*, 77:3594–3615, 2021. 2

[30] Samir S Yadav and Shivajirao M Jadhav. Deep convolutional neural network based medical image classification for disease diagnosis. *Journal of Big data*, 6(1):1–18, 2019. 2

[31] Zhengshi Yang, Xiaowei Zhuang, Virendra Mishra, Karthik Sreenivasan, and Dietmar Cordes. Cast: A multi-scale convolutional neural network based automated hippocampal subfield segmentation toolbox. *NeuroImage*, 218:116947, 2020. 2, 4

[32] Weina Zhao, Yishan Luo, Lei Zhao, Vincent Mok, Li Su, Changhao Yin, Yu Sun, Jie Lu, Lin Shi, and Ying Han. Automated brain mri volumetry differentiates early stages of alzheimer's disease from normal aging. *Journal of Geriatric Psychiatry and Neurology*, 32(6):354–364, 2019. 1

[33] Quan Zhou, Siman Liu, Chao Jiang, Ye He, Xi-Nian Zuo, Chinese Color Nest Consortium, et al. Charting the human amygdala development across childhood and adolescence: Manual and automatic segmentation. *Developmental Cognitive Neuroscience*, 52:101028, 2021. 2

[34] Hancan Zhu, Feng Shi, Li Wang, Sheng-Che Hung, Meng-Hsiang Chen, Shuai Wang, Weili Lin, and Dinggang Shen. Dilated dense u-net for infant hippocampus subfield segmentation. *Frontiers in neuroinformatics*, 13:30, 2019. 2