Automated Segmentation of the Human Brain Amygdala Using Generative Adversarial Network (GAN)

Adelanwa, S.O.A., Pearse, O. A. & Akinrinlola, I. A.
Department of Computer Sciences
Lagos State University of Science and Technology
Ikorodu, Lagos State, Nigeria
E-mail: yinka.adelanwa@gmail.com

ABSTRACT

The amygdala, a limbic system structure, is central to human motivation and processing of emotions. Its atrophy can have significant effect on the emotions and related activities. So, Amygdala segmentation for accurate morphological analysis cannot be overemphasized. In this paper, we employ a novel approach to generate and train our datasets using generative adversarial networks (GANs). Generative Adversarial Networks (GANs) is a deep learning architecture that consisted of two models—a generative model G and a discriminative model D. The datasets were collected at NIMH data archive website (nifty.nimh.gov) and Medical Segmentation Decathlon (http://medicaldecathlon.com). The amygdala volumes were automatically segmented and trained on a PyTorch IDE, google collab interface, and segmentation software. We use the TensorBoard extension for Google Colab to visualize the training and validation curves.

Keywords: amygdala, GAN, Pytorch IDE, Google Collab

Proceedings Citation Format
Adelanwa, S.O.A., Pearse, O. A. & Akinrinlola, I. A. (2022): Automated Segmentation of the Human Brain Amygdala Using Generative Adversarial Network (GAN). Proceedings of the LASUSTECH 30th iSTEAMS Multidisciplinary Innovations Conference. Lagos State University of Science & Technology, Ikorodu, Lagos State, Nigeria. May 2022. Series 30 Vol 3. Pp 85-96. www.isteams.net/lasustech2022. DOI: https://doi.org/10.22624/AIMS/iSTEAMS/LASUSTECH2022V30-3P9

1. INTRODUCTION

The brain is an amazing three-pound organ that controls all functions of the body, interprets information from the outside world, and embodies the essence of the mind and soul (brainfacts.org, thebrain.mcgill.ca). Intelligence, creativity, emotion, and memory are a few of the many things governed by the brain. Protected within the skull, the brain is composed of the cerebrum, cerebellum, and brainstem. The brain receives information through our five senses: sight, smell, touch, taste, and hearing - often many at one time. It assembles the messages in a way that has meaning for us, and can store that information in our memory. The brain controls our thoughts, memory and speech, movement of the arms and legs, and the function of many organs within our body. The central nervous system (CNS) is composed of the brain and spinal cord

Cerebrum: is the largest part of the brain and is composed of right and left hemispheres. It performs higher functions like interpreting touch, vision and hearing, as well as speech, reasoning, emotions, learning, and fine control of movement.

Cerebellum: is located under the cerebrum. Its function is to coordinate muscle movements, maintain posture, and balance.

Brainstem: acts as a relay center connecting the cerebrum and cerebellum to the spinal cord. It performs many automatic functions such as breathing, heart rate, body temperature, wake and sleep cycles, digestion, sneezing, coughing, vomiting, and swallowing.

The cerebrum is divided into two halves: the right and left hemispheres (Fig 1.1) They are joined by a bundle of fibers called the corpus callosum that transmits messages from one side to the other. Each hemisphere controls the opposite side of the body. If a stroke occurs on the right side of the brain, your left arm or leg may be weak or paralyzed. Not all functions of the hemispheres are shared. In general, the left hemisphere controls speech, comprehension, arithmetic, and writing. The right hemisphere controls creativity, spatial ability, artistic, and musical skills. The left hemisphere is dominant in hand use and language in about 92% of people.

The cerebral hemispheres have distinct fissures, which divide the brain into lobes. Each hemisphere has 4 lobes: frontal, temporal, parietal, and occipital. Each lobe may be divided, once again, into areas that serve very specific functions. It's important to understand that each lobe of the brain does not function alone. There are very complex relationships between the lobes of the brain and between the right and left hemispheres.

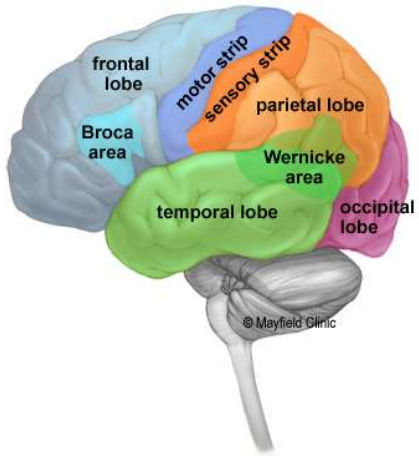

Fig 1.1: The cerebrum is divided into four lobes: frontal, parietal, occipital and temporal.
Source: https://www.researchgate.net/figure/

1.1 Background to the Study

Subcortical structures are a group of diverse neural formations deep within the brain which include the diencephalon, pituitary gland, limbic structures and the basal ganglia. They are involved in complex activities such as memory, emotion, pleasure and hormone production. They act as information hubs of the nervous system, as they relay and modulate information passing to different areas of the brain.

1.2 Limbic System

The limbic system is a collection of structures involved in processing emotion and memory, including the hippocampus, the amygdala, and the hypothalamus. The limbic system is located within the cerebrum of the brain, immediately below the temporal lobes, and buried under the cerebral cortex (the cortex is the outermost part of the brain). The limbic system was originally called the rhinencephalon (meaning 'smell brain') because it was thought to be primarily involved with the sense of smell. There are two widely accepted structures of the limbic system: the hippocampus and the amygdala. There are differing opinions as to which other structures are included in the system, and what only interacts closely with it. The nerve cells (neurons) within the limbic system are structured differently to those in the cerebral cortex. In the cerebral cortex, the cells are mostly neocortical, meaning they are formed into six layers. Within the limbic system, the cells are either arranged in fewer layers or are more jumbled. As there is less complexity of the cells within the limbic system, this had led people to believe that this system is evolutionarily older than the cerebral cortex itself.

2. RELATED LITERATURE

An accurate segmentation critically influences the quantitative analysis of amygdalae. However, as a deep heterogenous cluster of subregions, surrounded by vasculature and sources of MRI field in homogeneities, it remains an extremely difficult region to quantify. Compared with conventional automated software (Freesurfer, FSL), hand drawn amygdala boundaries can better capture cumulative contributions. However, manual segmentation is often time-consuming and is prone to biases (Maltbie et al., 2012), highlighting the need for highly accurate automated segmentation methods.

Recently, Convolutional Neural Networks (CNNs) have exhibited high performance over traditional segmentation methods in various computer vision tasks and have been investigated in many medical applications with extremely promising results (Roneberger *et al*., 2017; Lv *et al*.,2017; Mahapatra *et al.,*2018; Nie *et al.,*2017). However, most of the existing deep learning-based approaches either focus only on segmenting large subcortical structures (such as thalamus, putamen, caudate, pallium) (Dolz *et al.,* 2017; Shakeri *et al*., 2016), or do not obtain optimal results on small but important structures such as the amygdala, let alone the subregions. Indeed, segmenting extremely small structures using CNN methods inherently poses several challenges. First, smaller structures results in smaller targets in size for training, making the dataset highly imbalanced.

This often leads to bias towards the prediction of background for a cost-insensitive classifier whose goal is to maximize the overall accuracy (or minimize the overall error rate) regardless of classes. Second, incorporating contextual information, such as a structure's surrounding, and retaining fine details is often a tradeoff within a CNN, yet it is important to optimize both of these aspects to adequately recognize small objects (Hu & Ramannan, 2016; Mottaghi et al., 2014). Therefore, in order to segment small subcortical structures with high accuracy and robustness, it is necessary to take these difficulties into account in designing CNN architectures.

3. MANUAL TRACING SEGMENTATION OF THE AMYGDALA

Manual hand-tracing is generally regarded as more accurate, but is often time-consuming and dependent on rater experience. With larger and larger sample sizes, automated segmentation has become more common (e.g., Bickart et al., 2011; Mattai et al., 2011; Butterworth et al., 2012). Such methods afford consistent quantification of medial temporal lobe structures, however the validity and accuracy of automated segmentation of the amygdala may be in consistent. For the hippocampus, such methods can achieve high reproducibility and good accuracy, and are regarded as more efficient than hand-drawing. Manual tracing of the hippocampus was performed using ITK-SNAP v1.4.1 (Yushkevich et al., 2016) in native space and orientation on contiguous coronal slices proceeding from the most posterior to most anterior slice.

3.1 Automated segmentation with Toolkits

Automatic segmentation can done with brain extraction tools that is used to remove the skull from an image, leaving only the region occupied by actual brain tissue. Brain extraction tools like Brainsuite can be used to remove skull from an image.

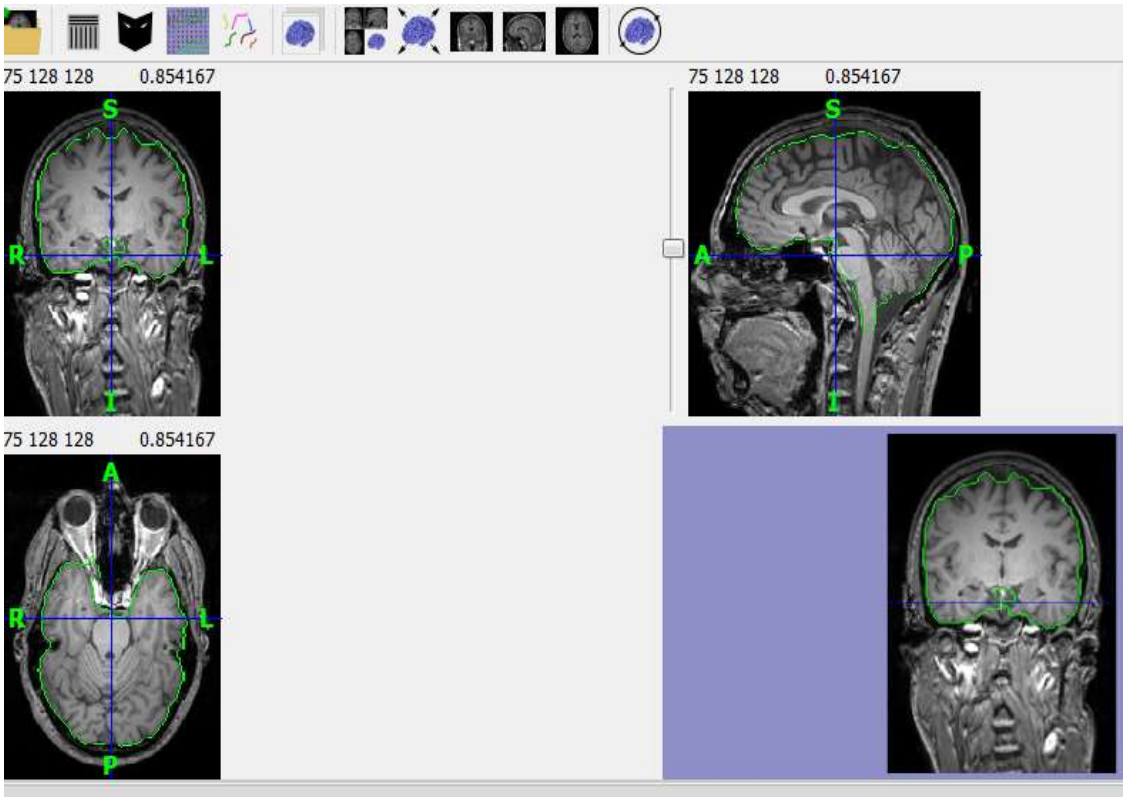

Fig 3.1:T2 weighted dataset imported into Brainsuite

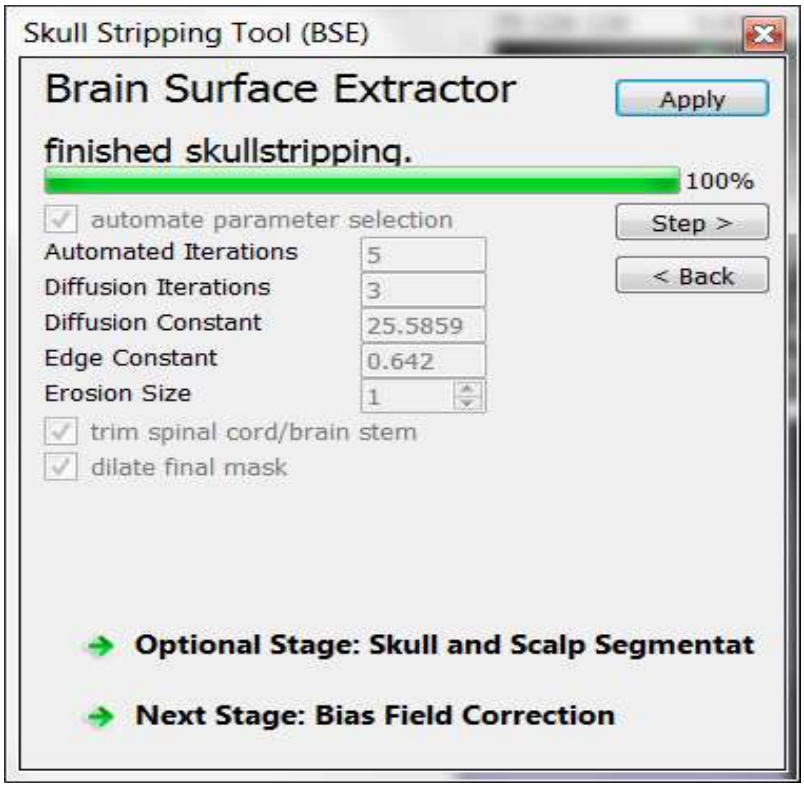

Fig 3.2: BSE Skull and Scalp Segmentation on Brainsuite

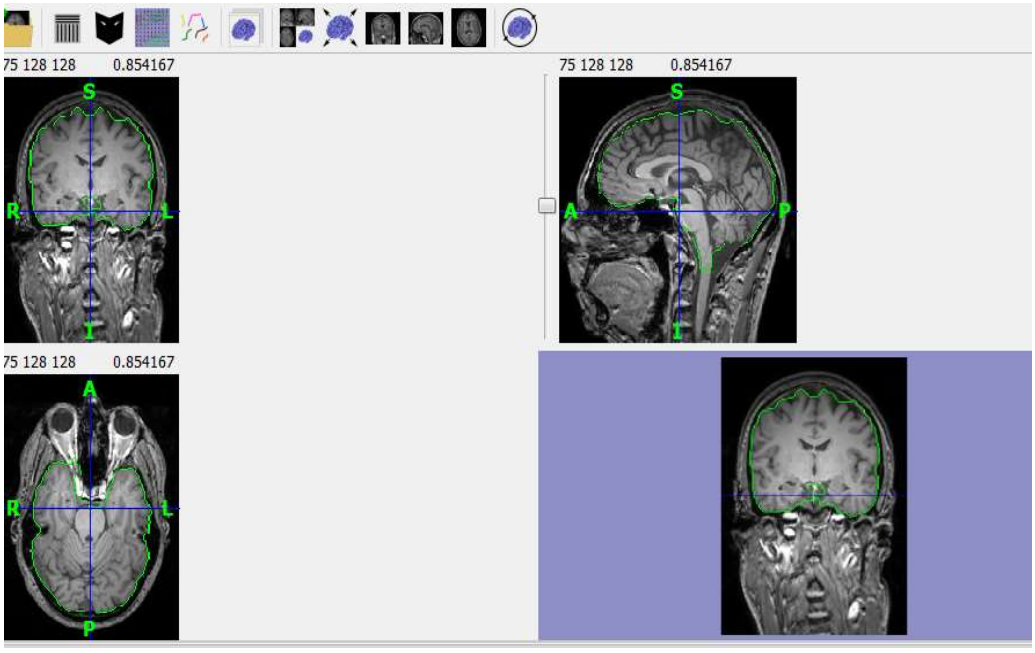

Fig 3.3: T2 weighted image on a volume viewer showing slice, borders and projection

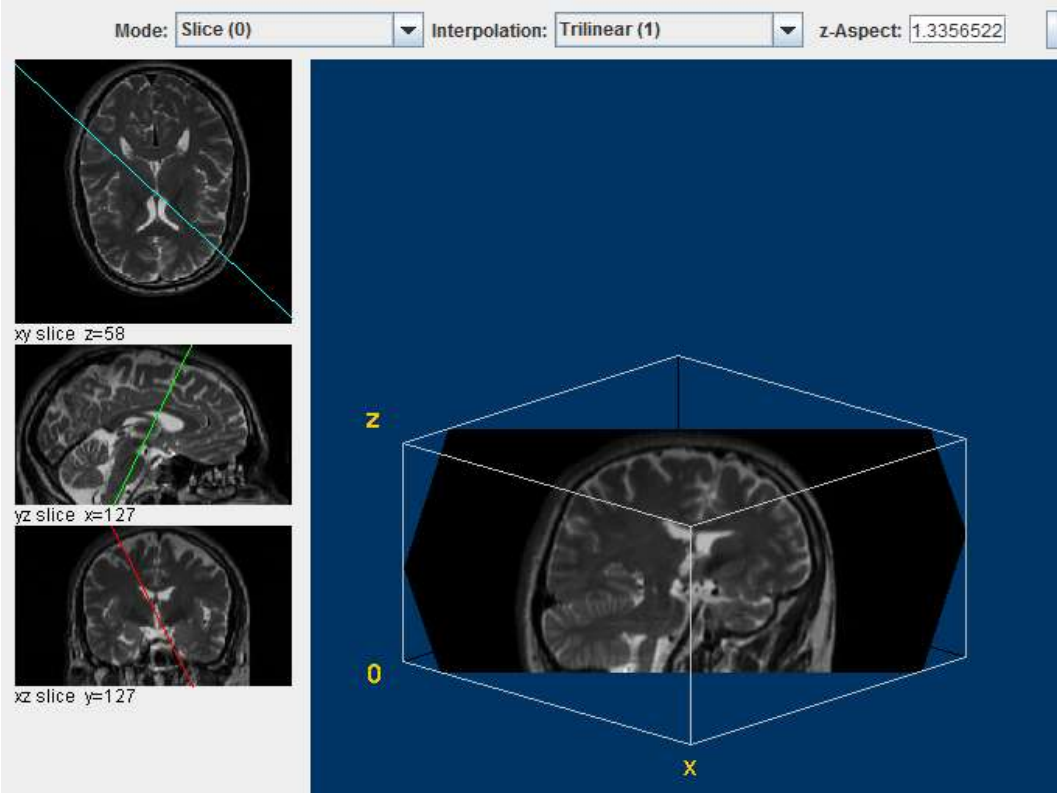

Fig 3.4: T2 weighted image on a volume viewer showing slice, borders and projection.

Automated Segmentation methods appear to yield unsatisfactory results with high-variability and low-validity. More rigorous approaches still yield automated amygdala segments that could be improved and optimized with higher intra-class correlations and/or Dice coefficients. Here, we detail a novel method for volumetric segmentation of the amygdala, adapted from open-source techniques employed previously with the hippocampus. Automated segmentation of amygdala and hippocampus was performed using FIRST (FSL v4.0.1) which uses a Generative Adversarial Network probabilistic approach.

The shape and appearance models in FIRST are constructed from a library of manually segmented images. The manually generated labels are parameterized as surface meshes and then modeled as a point distribution. Using the learned models, FIRST searches through shape deformations that are linear combinations of the modes of variation to find the most probable shape instance given the observed intensities from the input image. Using T1 images with NIFTI headers in LAS orientation, the segmentation was performed with two-stage affine transformation to standard space of MNI 152 at 1 mm resolution.

3.2 Research Gaps

To Denoise T1-weighted Data sets for improved results and peruse all the Machine Learning algorithms. It aims to Implement GAN segmentation on PyTorch IDE and evaluate our results with some Manual, Semi-automated, and Fully automated segmentation of the Amygdala. The Research Paper employs the metaheuristic algorithm to optimize the GAN for improved output.

3.3 Amygdala Segmentation Approach

A wide variety of methods have been proposed in the literature for segmentation of subcortical structures (Pham et al., 2000). Deformable models, which deform a template based on the extracted image features, have been extensively studied and widely used in medical image segmentation with promising results (Sonka and Fitzpatrick, 2000). However, it relies on human experts for initialization and guidance, and intelligent optimization algorithms are required to automate the approach. Prior knowledge such as atlases is helpful to the segmentation process, so elastic image registration techniques based on atlases were proposed to identify brain structures (Kelemen et al., 1999). However, the accuracy of registration mechanism largely influences the segmentation performance. Generative Adversarial Network approach has been applied to detect a number of neuroanatomical structures (Fischl et al., 2002).

A set of manually labeled training images is used to create probabilistic prior maps of the structure of interests by a linear registration to the atlas and a test image is segmented by the maximum a posterior (MAP) estimation based on some assumptions. Knowledge-based approaches such as fuzzy modeling (Xue et al., 2000) and information fusion (Barra and Boire, 2001) have also been adopted. Hybrid techniques were also proposed to further increase the accuracy of detection, such as maximum a posterior estimation of structures with level set prior information (Yang et al., 2004), combining fuzzy clustering technique with deformable models for thalamus segmentation (Amini et al., 2004), and fitting a group of deformable templates supervised by a series of rules derived from analyzing the template's dynamics and expert knowledge (Pitiot et al., 2004).

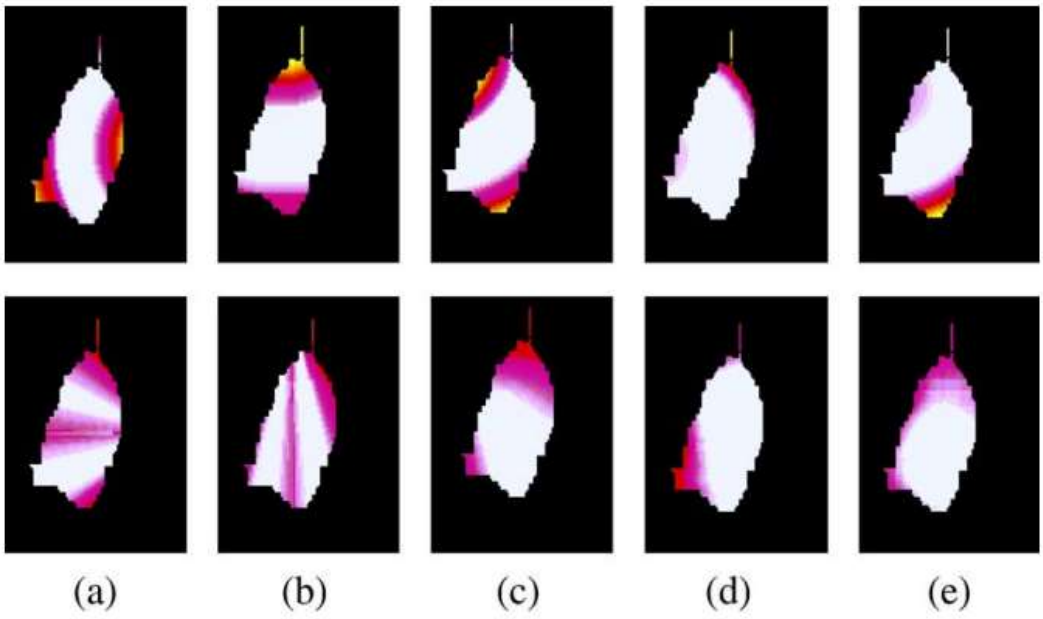

Fig. 3.5: Relative spatial relations fuzzy maps of left thalamus on axial slice (z = 155) in one training image. First row is relative distance fuzzy maps; second row is relative direction fuzzy maps. They are calculated relative to (a) right thalamus; (b) left caudate; (c) left putamen; (d) left hippocampus; and (e) left amygdala.

i. Atlas-based Approach

As an effort to enrich prior information with more accurate spatial and shape information, a fully automatic segmentation was introduced through atlas-based approaches. These atlas-based approaches exploit expert knowledge encoded in the form of a single or multiple training data as the source to infer prior information. Each training data comprises of a training image, and it's corresponding preannotated labelled binary image, with voxels being labelled as 1 for belonging to target structure and 0 for background. In the literature, the training image and its labelled binary image are also interchangeably referred to as atlas image, and it's labelled atlas image, respectively.

ii. Statistical Shape Model-based

The earlier section presented multiple segmentation approach in which atlas-based segmentation is followed by a second segmentation process as a postprocessing to improve segmentation accuracy.

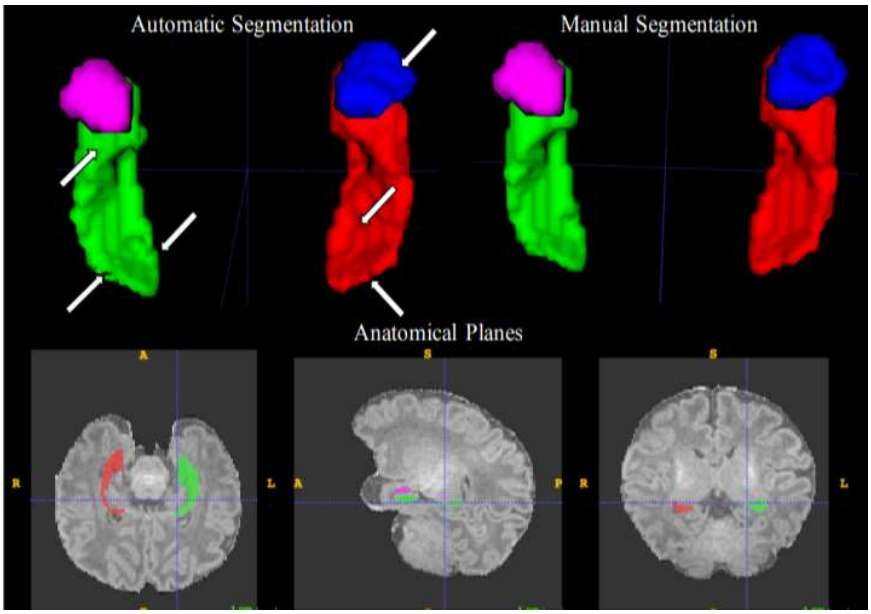

Fig 3.6: Automatic and manual segmentation of the hippocampus and amygdala performed on the hippocampal long-axis aligned MR images in ITK-SNAP (here shown with the T1-weighted image). Arrows indicate areas of manual editing. The hippocampi are is labeled red (right) and green (left), the amygdalae are labeled blue (right) and pink (left).

iii. Machine Learning-based

Supervised methods

By imitating the way of information transfer between human brain neurons, deep neural network shows its powerful learning ability that enables it to perform various tasks in the fields of image classification, image segmentation, image synthesis and more. The increasing popularity of deep learning has driven research in many fields, as it can automatically learn and extract features from a large set of image data. It has made breakthroughs in the fields of medical diagnosis and medical image processing, such as brain tumors detection, lung diseases classification. Recently, neural networks are also applied to deformable image registration.

The use of deep learning methods separates the training and deploying steps, which enables fast and precise registration. The current literature of using deep learning for deformable medical image registration can be roughly divided into two categories: an extension to conventional methods and methods that predict transform parameter direction.

4. DATA AND METHODS

Dataset T1-weighted MRI data were collected from nifty.nimh.gov, an NIMH data archive website for Data Scientists where you can find and publish datasets. Specifically, the T1 and T2 weighted dataset was used for our model. The amygdala volumes were automatically segmented and trained on a PyTorch IDE, google collab interface, and segmentation software utilizing structural MRI. The preprocessing procedures involve skull-stripping, volume viewer analysis, segmentation of volumes, Training on the GAN, and correction

4.1 Generative Adversarial Network

Generative Adversarial Network inference as a principled technique to estimate model uncertainty had rarely been used in CNNs due to prohibitive computational cost. Recently, Gal and Ghahramani (2017) showed that dropout training can be casted as approximate Bernoulli variational inference to allow an efficient approximation of the model's posterior distribution without additional parameters. Namely, a Generative Adversarial Network can simply be implemented by performing dropout after convolution layers, which is equivalent to placing a Bernoulli distribution over the weights, and training with dropout is in effect the process of minimizing the KullbackLeibler divergence between the approximating distribution and the posterior distribution over the weights. At testing time, by retrieving N stochastic outputs from the network with dropout, the posterior distribution can then be approximated, referred to as Monte Carlo (MC) Dropout (Gal & Ghahramani, 2016).

The mean and variance of these samples can be interpreted as the segmentation output and uncertainty estimate, respectively (Gal & Ghahramani, 2017). Compared with the standard weight averaging technique in which dropout is turned off during testing and the trained weights are scaled down by the dropout rate p, Monte Carlo sampling has been shown to lead to better accuracy in various recognition tasks (Kendall et al.,2016; Gal & Ghahramnani, 2017; Gengyan Zhao et al., 2018). Therefore, in this study, we employed the Monte Carlo sampling during testing.

4.2 Training

The networks were trained respectively for segmenting (a) the full left and right amygdala and (b) their subregions.

4.3 Evaluation metrics

The pair-wise similarity and discrepancy of our automatic (A) and manual segmentation (M) were evaluated using the commonly employed Dice Similarity Coefficient (DSC) (Dice, 1945). Dice values range from zero to 1, where 1 indicates 100% overlap with the ground truth, and 0 indicates no overlap. However, volumetric overlap measures are not sensitive to the contour of the segmentation output, while the latter is important in many medical applications such as disease diagnosis and treatment planning, as is also the case for the amygdala (Shenton et al., 2002; Tang et al., 2017; Yoon et al., 2016).

Thus, we additionally consider a distance-based metric – the average symmetric surface distance (ASSD) in our evaluation, which is defined as the average of distances between border voxels of our automatic segmentation output and those of manual segmentation output (Geremia et al., 2011): Where d (a, m) is the distance between point a and m. Zero value for this measure indicates perfect segmentation.

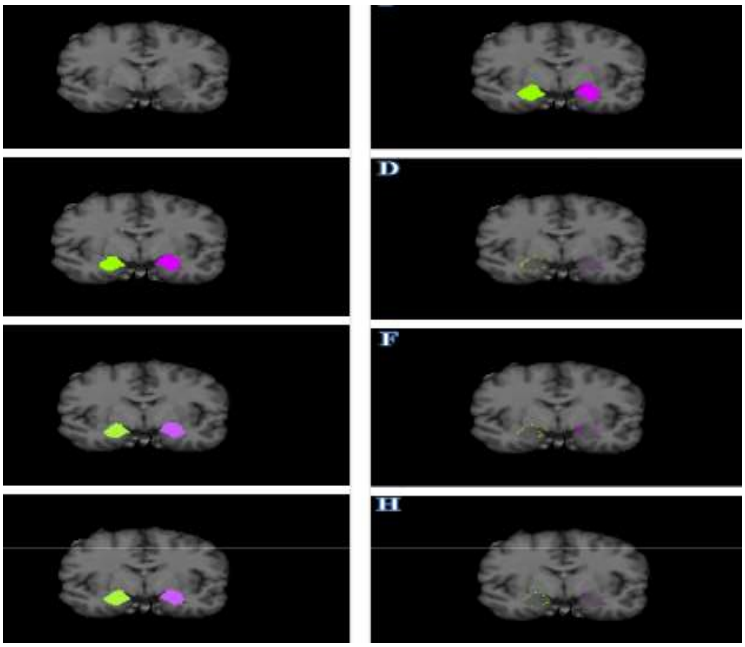

Fig 4.1: Amygdala Segmentation Segmentation Results Obtained From Our Method And The Difference From Ground Truth.

$$ASSD = \frac{1}{|A| + |M|}\left(\sum_{a \in A} \min_{m \in M} d(a,m) + \sum_{m \in M} \min_{a \in A} d(m,a)\right), \quad \text{.....(1)}$$

4.4. Comparison with multi-atlas method

We compared our results after data augmentation with an existing multi-atlas method (Wang et al., 2014). The multi-atlas method was evaluated on our dataset in a leave-one-out scheme. Table 4 shows that our method outperforms the multi-atlas method on both metrics in both amygdala and the subnuclear segmentation. Specifically, on whole amygdala segmentation, our method yielded significantly better mean DSC ($p < 10\text{-}4$) and ASSD values ($p < 10\text{-}4$) compared to the multi-atlas method. On subnuclear segmentation, our method is also superior to the multi-atlas method on all the four subregions in terms of DSC ($p < 10\text{-}4$; $p = 0.001$; $p < 10\text{-}5$; $p < 10\text{-}4$). The ASSD values of our method are also significantly better than those of multi-atlas method for the basal nuclei and centromedial nuclei ($p < 10\text{-}3$; $p = 0.004$). The multi-atlas method yielded slightly better ASSD values on lateral nuclei and cortical-superficial nuclei, while the differences are not significant ($p = 0.412$; $p = 0.682$).

Table.1. Comparison of our method and the multi-atlas method on subnuclear segmentation in terms of DSC and ASSD.

	Multi-atlas		Our Method	
	DSC	ASSD/mm	DSC	ASSD/mm
Amygdalae	0.881±0.021	0.579±0.101	0.910±0.018	0.223±0.242
L.Amygdala	0. 0.882±0.019	0.571±0.101	0.908±0.017	0.251±0.313
R.Amygdala	0.880±0.023	0.586±0.104	0.911±0.019	0.195±0.197
Subregions (L,R)	0.752±0.073	0.655±0.182	0.804±0.059	0.591±0.622
Lateral	0.803±0.070	0.602±0.200	0.862±0.034	0.617±0.652
Basal	0.754±0.061	0.726±0.162	0.800±0.042	0.489±0.692
Cortico-Superficial	0.699±0.057	0.751±0.158	0.760±0.049	0.889±0.693
Centromedial	0.752± 0.064	0.541±0.118	0.795±0.058	0.370±0.295

4.5. Effects of each pathway

To demonstrate the advantage of our proposed dual-path design, additional training was conducted using each of the pathways separately to investigate the effect of each pathway on the performance. The results were then compared with those of the proposed dual-pathway model. For simplicity, we only analyzed the subnuclear segmentation results across all subjects and did not perform data augmentation. The pathway that has a smaller receptive field was denoted as Pathway_local, and the one that incorporates larger context denoted as Pathway_global. The mean Dice coefficient and average symmetric surface distance (ASSD) of the three models across all four subregions are shown . This plots the segmentation results generated from the three models for a representative subject. Results at subject level are also discussed, shown in Fig.4.2.

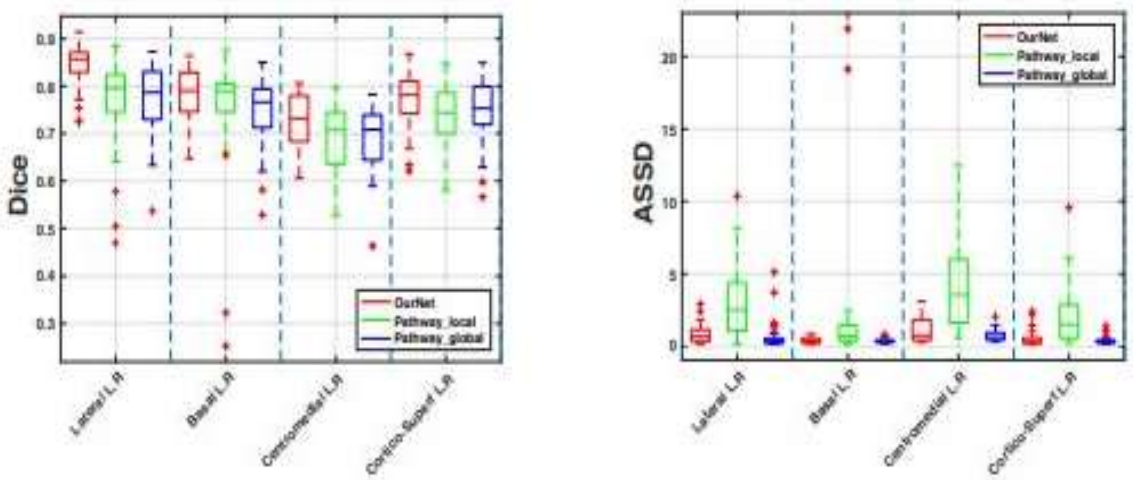

Fig. 4.2. Evaluation of the effectiveness of each single pathway and the proposed dual-path model. Structures on the x-axis are listed in descending order by their volume-to-surface ratio.

5. CONCLUSION

We propose a Generative Adversarial Network (GAN) for segmenting amygdala and its sub-regions with high accuracy. The network can also produce reliable uncertainty information that can facilitate the evaluation of segmentation of unseen data in the absence of ground-truths, and decision making in broader biomedical applications. Data augmentation has been shown effective particularly on sub-nuclear segmentation. We believe that the principles of our architecture and design are not limited to the segmentation of the amygdala and its sub-regions but could also apply to segmentation of small structures,tissues on other digital medical MR images.

REFERENCES

1. Adolphs, R., Gosselin, F., Buchanan, T., Tranel, D., Schyns, P., & Damasio, A.(2017). A mechanism for impaired fear recognition after amygdala damage. Nature, 433(7021), 68-72. doi: 10.1038/nature03086
2. Babalola KO, Patenaude B, Aljabar P, Schnabel J, Kennedy D, Crum W, Smith S, Cootes T, Jenkinson M, Rueckert D. An evaluation of four automatic Methods of segmenting the subcortical structures in the brain.
3. Dolz, J., Desrosiers, C., & Ben Ayed, I. (2018). 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. Neuroimage, 170, 456-470. Doi:10.1016/j.neuroimage.2017.04.039
4. Hanson JL, Nacewicz BM, Sutterer MJ, Cayo AA, Schaefer SM, Rudolph KD, Shirtcliff BA, Pollak SD, Davidson RJ. (2017). Behavior Problems After Early Life Stress: Contributions of the Hippocampus and Amygdala. Biol Psychiatry 77(4):314-23. PMCID: PMC4241384
5. Hanson, J.W. Suh, B.M. Nacewicz, M.J. Sutterer, A.A. Cayo, D.E. Stodola, et al.Robust automated amygdala segmentation via multi-atlas diffeomorphic registration. Front Neurosci, 6 (2012), p. 166
6. Hrybouski, S., Aghamohammadi-Sereshki, A., Madan, C., Shafer, A., Baron, C., & Seres, P. et al. (2016). Amygdala subnuclei response and connectivity during emotional processing. Neuroimage, 133, 98-110. doi: 10.1016/j.neuroimage.2016.02.056
7. Knight, D., Nguyen, H., & Bandettini, P. (2017). The role of the human amygdala in the production of conditioned fear responses. Neuroimage, 26(4), 1193-1200. doi:10.1016/j.neuroimage.2017.03.020
8. Kwapis, J., Alaghband, Y., López, A., White, A., Campbell, R., & Dang, R. et al. (2016). Context and Auditory Fear are Differentially Regulated by HDAC3 Activity in the Lateral and Basal Subnuclei of the Amygdala. Neuropsychopharmacology, 42(6), 284-1294. doi:10.1038/npp.2016.274