Deep Learning Methods to Identify Intracranial Hemorrhage Using Tissue Pulsatility Ultrasound Imaging

Nhut Minh Phan

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2022

Reading Committee:
Erika Parsons, Chair
Pierre Mourad
Michael Stiber

Program Authorized to Offer Degree:
Computer Science & Software Engineering

University of Washington

Abstract

Deep Learning Methods to Identify Intracranial Hemorrhage Using Tissue Pulsatility Ultrasound Imaging

Nhut Minh Phan

Chair of the Supervisory Committee:
Dr. Erika Parsons
Computer Science & Software Engineering

TBD

TABLE OF CONTENTS

LIST OF FIGURES

GLOSSARY

CRANIUM: the part of the skull that encloses the brain.

CT: Computer Tomography.

IED: Improvised Explosive Device.

IRB: Institutional Review Board.

INTRACRANIAL HEMORRHAGE: bleeding inside the brain.

LSTM: Long Short-Term Memory.

RNN: Recurrent Neural Network.

TBI: Traumatic Brain Injury.

TPI: Tissue Pulsatility Imaging.

CTBI: closed Traumatric Brain Injury.

PTBI: penetrating Traumatic Brain Injury.

WHO: World Health Organization.

ACKNOWLEDGMENTS

I would like to express sincere appreciation to Dr. Pierre Mourad and Dr. Michael Stiber for accepting my request to be on the commitee for this capstone project. I am also grateful for the support received from Dr. John C. Kucewicz and Nina LaPiana. Without their contribution in data collection, data registration, and signal processing, this project would not happen. Finally, I offer my very special thanks to Dr. Erika Parsons for being the Chair of my committee and without whose support and guidance, this project would not have been possible.

DEDICATION

To my parents, sister, and dear wife without whose support I would not have been able to achieve my goals.

Chapter 1

INTRODUCTION

1.1 Background

Brain injury may happen in one of two ways: close brain injury (cTBI) and penetrating brain injury (pTBI) [3]. Closed brain injuries happen when an injury is nonpenetrating and does not cause any break in the skull. The source of these injuries are rapid forward and/or backward movements and shaking of the brain inside the bony skull that results in bruising and tearing of brain tissue and blood vessels. Penetrating brain injuries happen when a foreign object penetrates the skull and then traverses through the brain parenchyma. For instance, a bullet travels through the head, piercing the brain.

What why is TBI important for civilians and battle field? TBI in the battle field: A large percentage of deployed U.S. sodiers (40% to 60% of surviving soldiers) suffer from closed-head injuries caused by the blast effect of IED explosion [16]. These injuries could result in intracranial hemorrhage, causing long-term neurological damages if left untreated. For severe TBI cases, the patients must be evacuated to the nearest combat hospital that has equipment to support neurosergery, airway protection, mechaincal ventilation, among other means for critical care. However, severe cTBI patients often do not survive more than one year post injury [16]. Thus, early diagnosis is critical not only to improve the clinical outcome, but also to provide medical personnel with information to make decision when resources are scarce.

Besides the relevance on the battle field, TBI is a pressing public health and medical problem around the world. According to the World Health Organization (WHO), TBI affects an estimate of 10 million people annually [24]. Low and middle income countries face higher risk factors for causes of TBI due to inadequate health care systems.

Early diagnosis and immediate medical care is extremely important in improving the clinical outcomes for TBI patients [4]. Computer Tomography (CT) and Magnetic Resonance Imaging (MRI) are the current standard methods for identifying intracranial hemorrhage [22]. The main disadvantage of these imaging modalities is the complexity, size, and cost of the required equipment, making them inaccessible in the combat settings and in low income countries. In contrast, ultrasound imaging could be used with relatively affordable equipment that are as small as a standard tablet. An example of such systems is a tablet-like device from Terason (the company website is at https://www.terason.com). Ultrasound imaging has a major drawback: ultrasound waves do not penetrate bones very well, making ultrasound imaging more suitable for infants up to about 18 months old at which age the craniums are yet fused together [1].

A team of researchers from the University of Washington developed a novel ultrasound technique called tissue pulsatility imaging (TPI) that captures the pulsation of the brain tissue as blood infuses the brain during a cardiac cycle [29]. The team collected data from civilian patients who suffer moderate to severe cTBI. The working hypothesis is that the difference in the movements of brain tissue versus TBI lesion allows one to detect intracanial hemorrhage through computer assisted means. This project aims to employ the power of deep learning to produce an algorithm that can automatically identifying intracranial hemorrhage from TPI data.

1.2 Data Collection

The data used in this project were collected from actual humans, thus the data collection is under controlled by the policies of the University of Washington's Human Subjects Division. Data collection was approved by the Institutional Review Board (IRB) through a Zipline application, the Human Subjects Division's e-IRB system.

The data was collected from patients admitted to Harborview Medical Center (HMC) in Seattle, the only Level 1 trauma center in the area capable of providing total care for every aspect of traumatic injuries to the brain [2]. The criteria for selecting patients are:

- Moderate to severe cTBI at HMC;s Neuro ICU, with or without polytrauma
- Patients 18 years or older
- Not prisoners
- Not from Native American or non-U.S. indigenous populations through a tribe, tribe-focused organization, or similar community-based organization

The last two criteria are to simplify the Zipline application. When TBI patients arrived at HMC, the patients received life saving treatment if necessary, including diagnosis of TBI via CT imaging. A research coordinator then screened them for the inclusion criteria. If they had an injury the team is interested in, then the coordinator asked them or their family for consent. Ultrasound raw data were collected using a Terason (Burlington, MA) u3200t, a tablet-based, general purpose scanner with a 4V-2 phased array transducer (64 elements, 2.5 MHz RF sampling frequency, 128 scanlines per frame). The scan rate was fixed by the manufacture at 30 frames per second. A certified medical sonographer used the Terason device in the hospital room to collect ultrasound data. Data could come from patients hours after an injury or days after an injury. Data collection happened after the patients were diagnosed using CT or MRI, and did not interfere or interrupt routine hospital clinical care.

The sonographer collected data through the temporal window of the head (See Figure 1.1) without shaving the head, spending approximately five minutes aiming the ultrasound toward the known intracranial hemorrhage with each of the following orientations: coronal, axial, and oblique. The axial is a horizontal slice of the brain at zero degree; coronal is a 90-degree slice of the brain; oblique is a slice at any angle from one degree to 179 degrees excluding the coronal. Data were collected from the left and right side of the head. The sonographer also used a pulse oximeter to capture phases of the cardiac cycle. Any patient identifying information, including name, age, and gender, was removed from the data. Patients are distinguished by three-digit numbers.

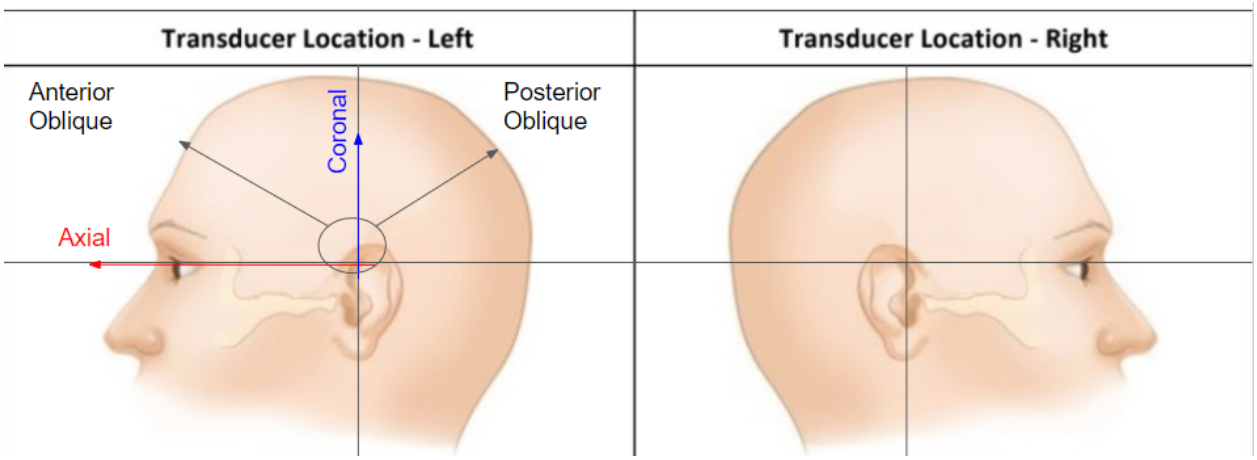

Figure 1.1: The planes at which ultrasound data are collected: left or right side of the head and at axial, coronal, or oblique.

Prior to restrictions imposed due to COVID-19 virus, the sonographer collected 15 scans from each side of the head equally divided between the three scan planes. The amount of date collected following COVID-19 restriction, the data were limited to 1 axial, 1 coronal and 4 oblique scans from each side of the head.

1.3 Signal Processing

Signal processing is not in the scope of my project. It was done by an ultrasound signal processing expert, Dr. John Kucewicz. This section provide a summary of the Dr. John Kucewicz's signal processing algorithm.

Medical, diagnostic ultrasound transmits and receives short bursts of high frequency sound typically between 1 and 10 MHz. As sound propagates through the body, some fraction of the sound is scattered back towards the ultrasound transducer due to differences in local acoustic impedance. Because of the speed of sound in tissue is relatively constant, two-dimensional images of tissue structure can be created based on the amplitude of the received ultrasound and the time between a sound burst's transmission and its reception. Images of

blood or tissue velocity can be created by transmitting 2 or more bursts of ultrasound and measuring the spatially localized relative temporal shifts in the received ultrasound from burst to burst. In ultrasound parlance, images of structure are referred to as B-Mode, i.e. brightness mode, images, and images of velocity are referred to a Doppler ultrasound or tissue Doppler ultrasound if tissue velocity is being measured.

Tissue Pulsatility Imaging (TPI) is a variation of tissue Doppler designed to measure the naturally occurring pulsatile motion of tissue due to blood flow [30, 29]. During systole, blood accumulates in the arterial vascular causing tissue to expand by a fraction of a percent. Later in the cardiac cycle, accumulated blood flows through the capillary bed into the venous vasculature and back towards the heart allowing tissue to relax to its pre-systolic blood volume. TPI was inspired by plethysmography, an older technology used to measure gross changes in tissue blood volume in, for example, an entire limb. TPI extends this idea to measure local changes in tissue blood volume within the body.

For this project, it is hypothesized that TPI can be used to detect variations in tissue pulsatility localized to areas of the brain experiencing hemorrhage. The measurement of the TPI signal is based on standard ultrasound signal processing methods [13]. Ultrasound data was collected with a Terason (Burlington, MA) u3200T tablet-based ultrasound system and 4V2 phased-array ultrasound transducer. With functionality enabled by the manufacturer, we were able to collect 8 seconds of raw ultrasound data at 30 frames per second for offline analysis in MATLAB (The Mathworks, Natick, MA). The ultrasound data was filtered into 8 frequency sub-bands, quadrature demodulated, lag 1 autocorrelation was used to compute the velocity for each sub-band, and the velocities were averaged weighted by the power of the sub-band yielding an 8 second velocity waveform for every pixel in the image sequence.

Following the velocity calculation, an unpublished vector Doppler method was used to suppress common-mode motion introduced by gross motion of the ultrasound transducer relative to the patient. Tissue displacement was calculated from the time integral of velocity combined with a band-pass filter to emphasize cardiac pulsatility. Lastly, individual cardiac cycles within the 8 seconds were detected and resampled such that all cycles were of uniform

duration. i.e. 30 samples per cardiac cycle. Figure 1.2 summarizes the signal processing steps.

1.4 CT Registration to Create Masks

TBD

1.5 Existing System

A previous student approached the problem of detecting intracranial hemorrhage using a different representation of tissue displacement data, whose description is not in the scope of this writing. The student attempted several different deep learning models, including MobileNetV2 [41], pix2pix [26], and ResNeSt [49]. The architectures of the mentioned models were used unmodified. The loss function used for model optimization was categorical loss entropy. The metrics used for evaluation are precision and recall. The best model was reported to be ResNeSt. with a reported precision score of about 0.07 and recall of about 0.2 after 70 training epochs.

The evaluation results showed that the previous deep learning methods and data preprocessing might not be suitable for the detection of intracranial hemorrhage from displacement data. In addition, the choice of loss function and evaluation metrics is poor. Categorical loss entropy considers the predicted results of all pixel an image equally. The result is that the loss value could be very low, but the predictive power for a class of interest is very low. The precision and recall scores evaluate the quantity of true positive, true negative, false positive, and false negative in relation to each other, but they do not indicate how well a model learn the correct label of a pixel or how well the ground truth overlaps with the detected area.

1.6 Problem Statement and Scope

This capstone project focuses on the core algorithm that detects the regions of intracranial hemorrhage with in a patient's brain tissue. The scope is limited to data preprocessing, skull detection, ventricle detection, brain mass detection, and intracranial hemorrhage diagnosis.

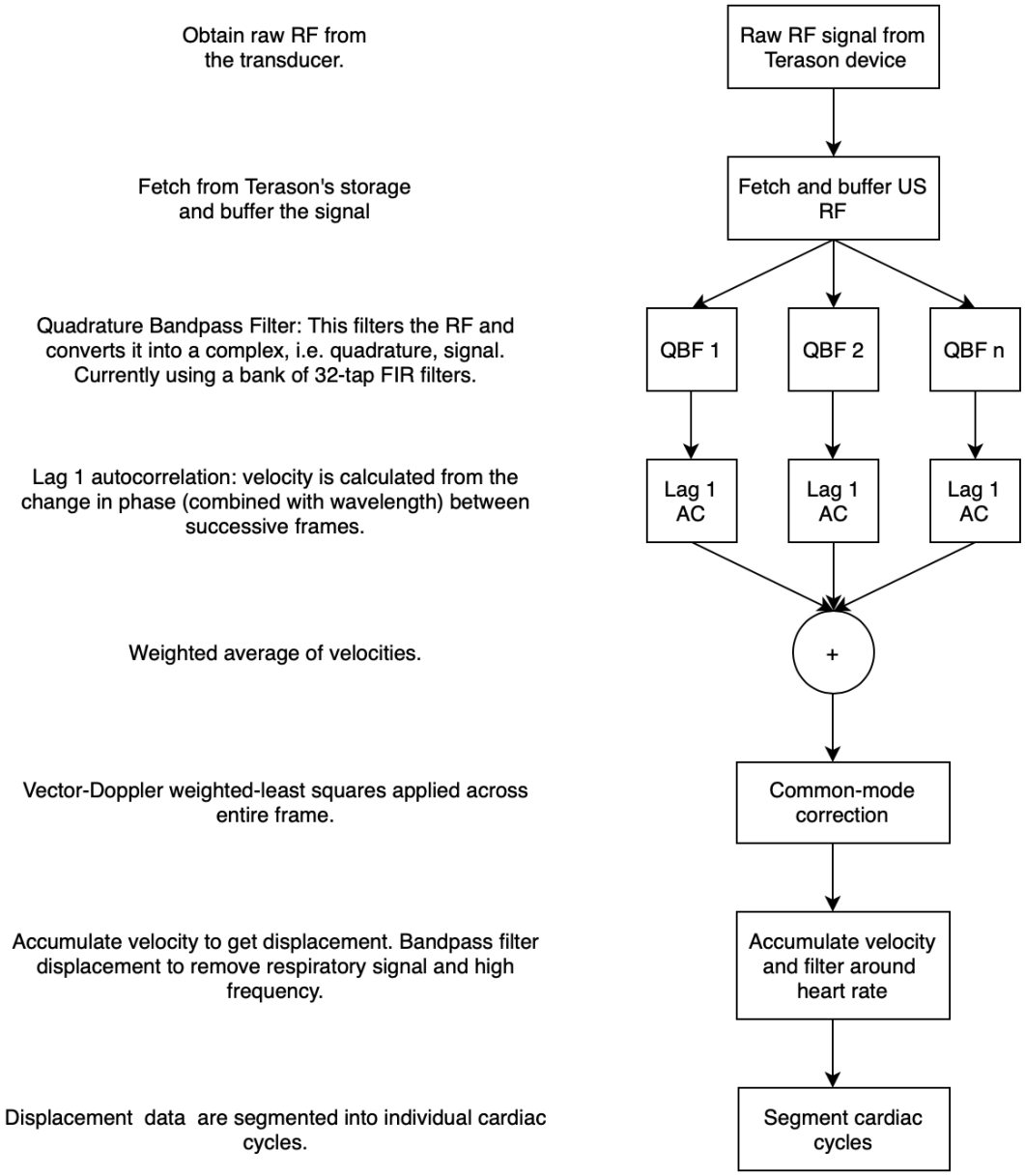

Figure 1.2: The raw RF data are processed using various signal processing techiques to produce displacement data that is synchronized with the heart's cardiac cycle.

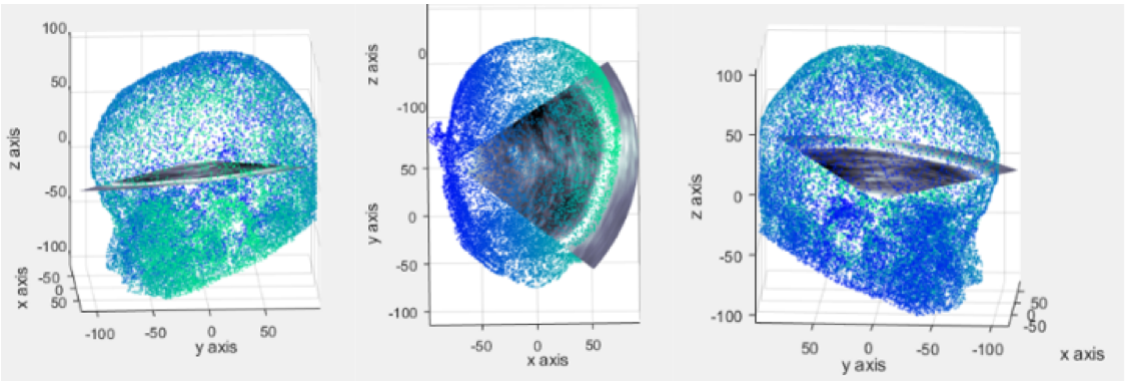

Figure 1.3: Registering a 2D ultrasound scan plane to the 3D CT data. The 2D B-mode scane is shown relative to a point cloud representing the skin surface of the head in the 3D CT data.

1.6.1 Data Preprocessing

TBD

1.6.2 Skull Detection

TBD

1.6.3 Brain Mass Detection

TBD

1.6.4 Hemorrhage Diagnosis

TBD

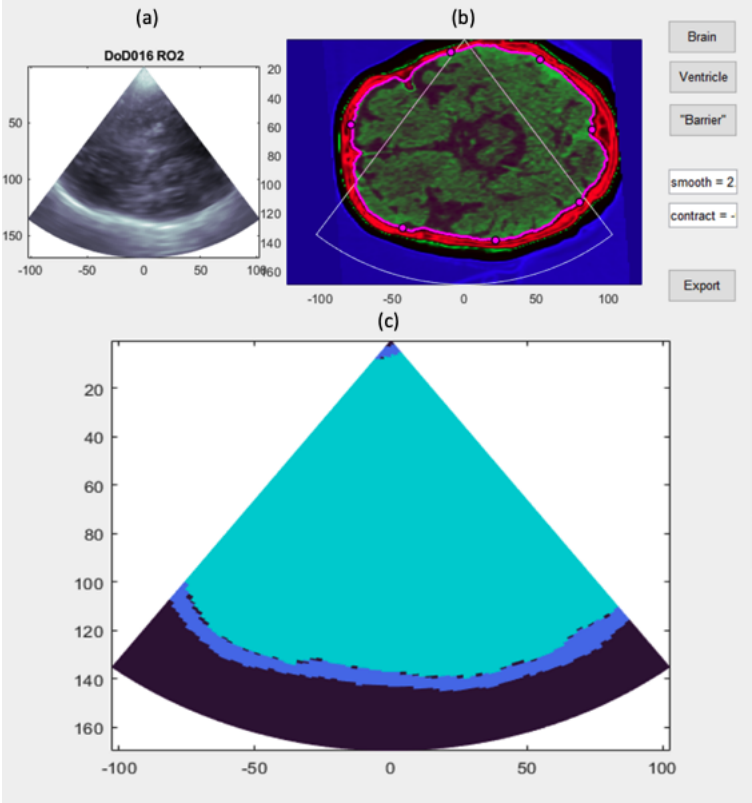

Figure 1.4: CT brain mask constructed by registration with B-Mode ultrasound for patient # 16 at the right oblique plane. (a) B-mode ultrasound data. (b) An overlay of the ultrasound plane on the CT image. (c) Resulting brain mask (cyan) and skull mask (blue).

Chapter 2

RELATED WORK

2.1 Overview

Medical image segmentation aims to identify structures such as organs or lesions in an image; it plays a key role in computer aid diagnosis. The two categories of image segmentation tasks are semantic segmentation and instant segmentation [31]. Semantic segmentation involves pixel-level classification, assigning a corresponding category for each pixel in an image. Instant segmentation is more complex, distinguishing instances on the basis of specific categories in addition to outputting the basic semantic segmentation's result. This project focuses on semantic segmentation; for the remaining sections of this writing, image segmentation means semantic segmentation. Depends on the availability of labeled data, different flavors of deep learning (or general machine learning) could be applied to solve a problem, supervised learning, weakly supervised learning, and unsupervised learning. When the data are carefully labeled, supervised learning is the popular choice. The disadvantage of this method is that it is difficult to obtain a large number of labeled medical images. Unsupervised learning does not require labeled data but the difficulty of learning is significantly higher. On the other hand, when a dataset contains both labeled and unlabeled images, weakly supervised learning is another option. Deep learning has been used widely for medical image segmentation with huge success compared to early approaches, including edge detection, template matching techniques, statistical shape models, active contours, etc [31]. This chapter discusses the recent progress in the field of medical image segmentation. The discussion focuses on the applications of deep-learning based medical image segmentation, supervised deep learning techniques, and data augmentation.

2.2 Application

Popular medical image segmentation tasks includes, but not limited to, liver and liver-tumor segmentation [32, 46], brain and brain-tumor segmentation [35, 11], optic disc segmentation [10, 15], cell segmentation [40, 43], lung segmentation and segmentation of pulmonary nodules [47, 39]. The data used for medical diagnosis are commonly 2D gray scale or three-channel RBG images and 3D volumetric data. Deep learning practitioners get their data mostly from the four popular imaging modalities, X-ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and ultrasound [31]. When the structures of interest is at cellular scale, the images often come from Electron Microscopy (EM) [40].

2.3 Supervised Learning

Supervised learning is the most popular method for medical image segmentation tasks. This section reviews the existing methods that deep learning researchers have attempted to solve medical image segmentation problems. Specifically, I will discuss the two important components of a deep learning model, the backbone network and the network block.

2.3.1 Backbone network

The backbone network is the high-level architecture of a deep learning model. It draws a big picture of the way different pixel-level computational operations are chained together to process the input and output a segmentation map. In the last decade, researchers have proposed various backbone architectures for deep learning, U-Net, 3D Net, Recurrent Neural Network (RNN), skip connection, cascade of 2D and 3D models.

U-Net One of the most popular deep learning architectures in the field of general image segmentation is the encoder-decoder structure. U-Net [40], fully convolution network (FCN) [33], and Deeplab [8] are examples of the encoder-decoder structure. As the name suggest, encoder-decoder architectures have two main parts, the encoder and decoder. The encoder

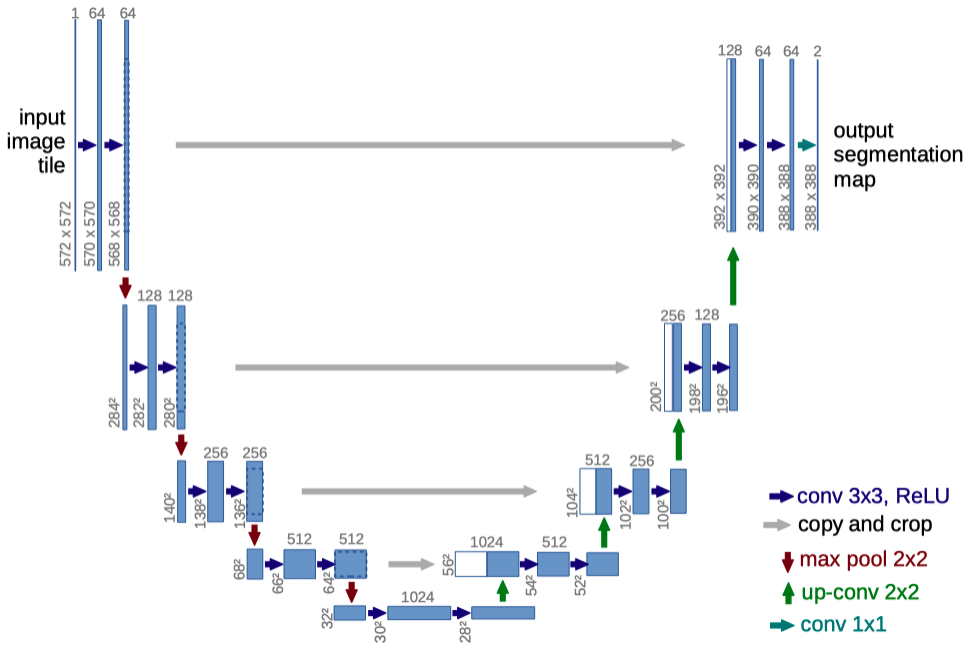

Figure 2.1: The U-Net architecture. The left and right side are the encoder and decoder, respectively[40].

reduces the dimensions of the inputs and extracts image features while the decoder restores these features to the original size and output the segmentation maps. U-Net has been used widely for medical image segmentation and has become the benchmark for most medical image segmentation tasks[]. Many modern architectures were inspired by U-Net [48, 50, 27, 5]. Figure 2.1 shows the architecture of the U-Net. The left side of the U-Net works as a encoder. It consists of a series of convolutional filters that extract features and max pooling layers that shrink the inputs. The right side is the decoder that recovers the original image dimensions by four up-convolution operations. The U-Net is perfectly symmetric, allowing skip connections to concatenate the input features of the encoder to the features of the decoder. This structure effectively fuses low-level (high resolution) and high-level (low resolution) image features through skip connections. By these connections, the U-Net allows segmentation of medical images, which often contain noise and show blurred boundaries.

3D Net Motivated by the success of U-Net, the author of the U-Net paper extended the architecture to work with 3D medical data [51]. Such data are very common in the form of volumetric CT and MRI data. Inherently, this architecture has more parameters and requires much higher computer resource than the U-Net. Due to insufficient computing capability, the authors limited the 3D U-Net to only three down-sampling layers, reducing the model's segmentation accuracy when tested. Figure 2.2 depicts this architecture. Milletari et al. proposed a similar structure called V-Net, [36] as shown in Figure 2.3. V-Net was developed to allow stacking more network layers, providing better feature representation. One problem that deep neural networks face is vanishing gradient [21]. The gradients (derivative) of a neural network are found using back propagation during the training process. By the chain rule, the derivatives of the hidden layers are multiplied from the output layer to the input layer to compute the derivative of the initial layer. If the derivatives are small, multiplication causes the gradient to decrease quickly, causing vanishing gradient. Residual connection is one of the solution for this problem [21]. The residual connection (identity), as shown in Figure 2.4, is essentially a skip connection that brings weights from an earlier stage to a later stage. These weights do not go through the activation function, so do not suffer from vanishing gradient. The V-Net employs residual connection in the design, allowing it to have four down-sampling layers.

Recurrent Neural Network Recurrent neural network (RNN) is intended for sequence problems such as voice recognition, language translation, and time-series sequences to name a few. One of the most popular RNNs is the Long Short-Term Memory (LSTM) network, developed by Hochreiter et al. in 1997 [23]. It can mitigate the vanishing gradient problem by introducing self-loop. RNN has been used to model the time dependence of image sequences. Alom et al. proposed the Recurrent Convolutional Neural Network based on U-Net (RU-Net), as shown in Figure 2.5, that improves feature representation for medical image segmentation tasks by recursive residual convolutional layers [5]. Gao et al. proposed the Fully Convolutional Structured LSTM (FCSLSTM), a combination of LSTM and CNN,

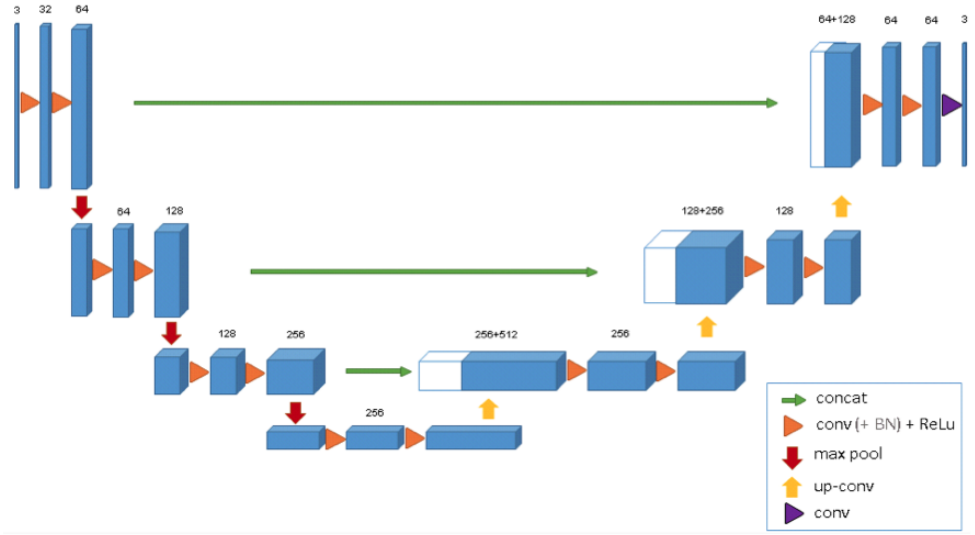

Figure 2.2: The 3D U-Net architecture[51].

to model the temporal relationship between different brain MRI slices [17]. Another work from Bai et al. joined FCN with RNN to extract the spatiotemporal information from aortic image sequences [6].

Skip connection As aforementioned, the skip connection of U-Net fuses low-resolution and high resolutions features to improve feature representation. A disadvantage of the skip connection is that there is a large semantic gap between low-resolution and high-resolution features, leading to blurred feature maps [31]. Ibtehaz et al. improved the original U-Net multiple ways, including replacing the ordinary skip connection with convolution operations on the encoder's features before fusing with the decoder's features [25]. They tested the model on five different dataset, including EM images for cell segmentation and endoscopy images for colon cancer detection, and reported better performance than U-Net. Two additional work by Seo et al. [42] and Chen et al. [9] also added convolution operations to the skip connection to improve performance of liver-lesion segmentation tasks.

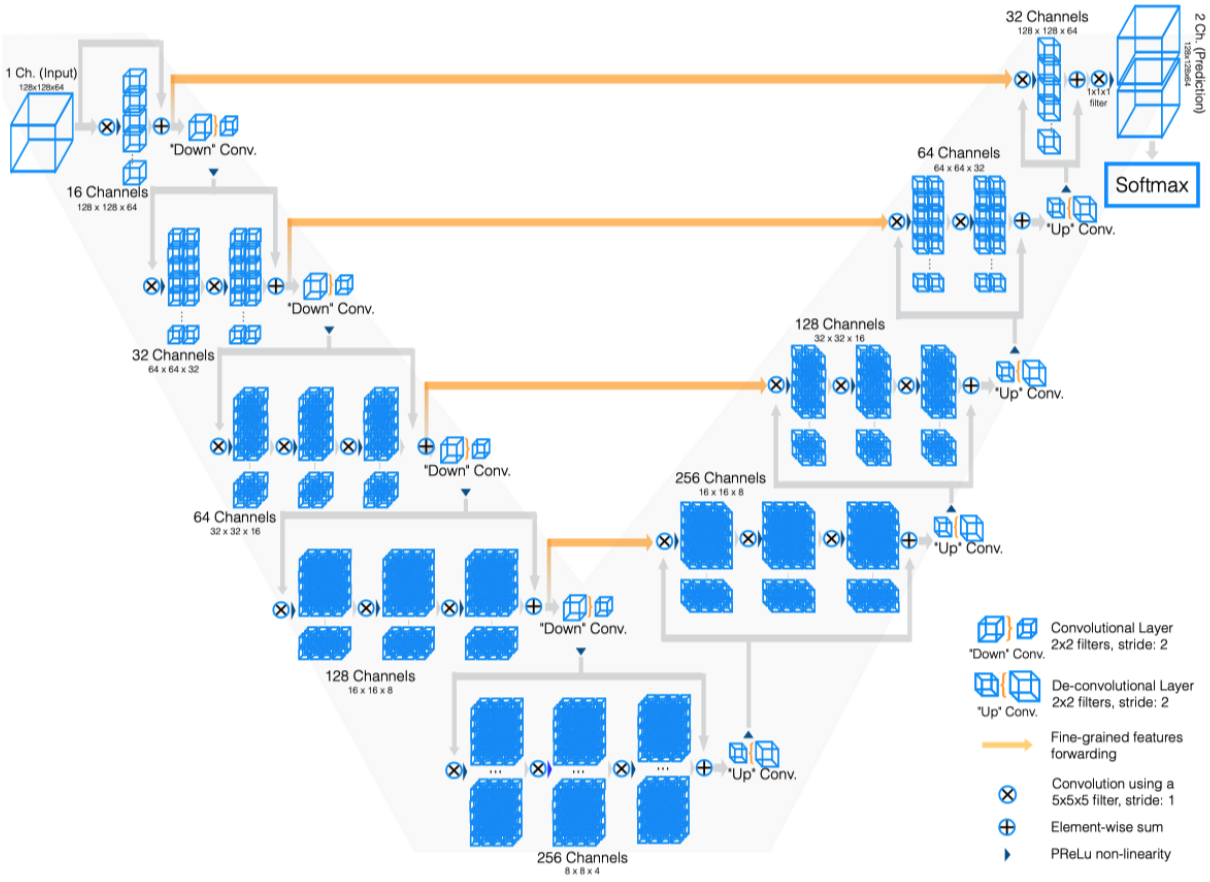

Figure 2.3: The V-Net architecture[36].

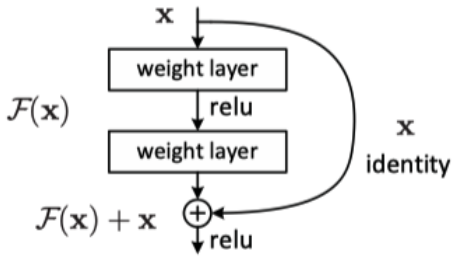

Figure 2.4: Residual connection[21].

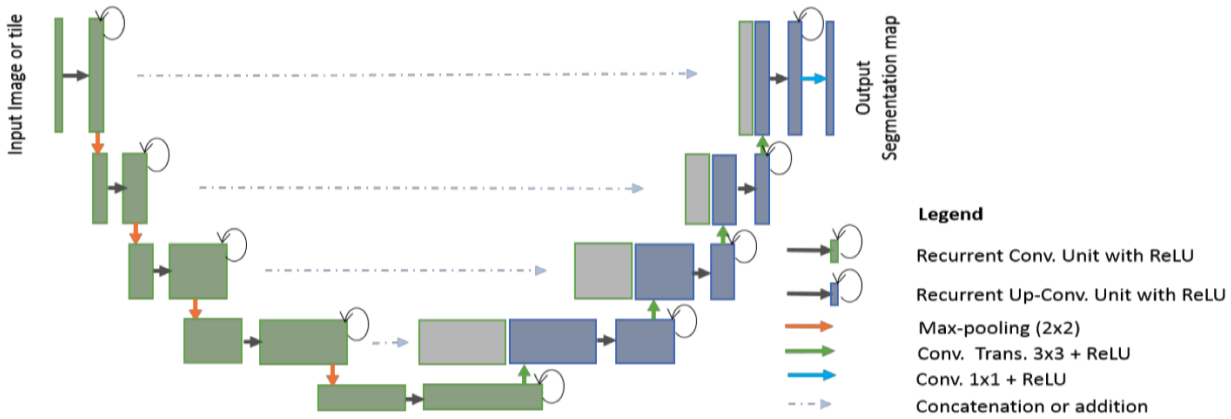

Figure 2.5: RU-Net architecture with convolutional encoding and decoding units using recurrent convolutional layers (RCL) based on U-Net architecture.[5].

Cascade Models The accuracy of segmentation tasks could be improved by training multiple models and chaining their input/output together. This type of architecture has three main categories, coarse-fine segmentation, detection segmentation, and mixed segmentation [31]. In this writing, I am interested mostly in the coarse-fine segmentation type. In coarse-fine segmentation networks, a cascade of two 2D networks is used. The first network provides a coarse segmentation, and the second network refines the coarse segmentation with details. Many researchers have developed this type of networks for liver and liver tumor segmentation using 3D. Christ et al. proposed the Cascaded fully convolutional neural networks (CFCNs) for liver and liver tumor segmentation[12]. In their work, they used an FCN to segment the liver, then fed the result to a second FCN for liver tumor segmentation. Other coarse-fine segmentation networks for liver and liver tumor segmentation are described in [45, 28, 14].

2.3.2 Netwok Block

Within the larger backbone network, there are the network blocks that perform the computations on the inputs to generate the features maps. For instance, The network blocks of the vanilla U-Net model are simple convolutional layers. Researchers have devised a variety

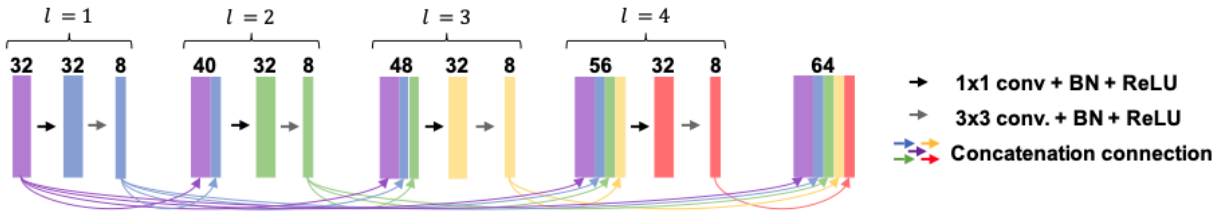

Figure 2.6: The dense block with four layers. The feature maps from previous layers are concatenated together as input to the following layers. BN stands for batch normalization [19].

of network blocks to effectively extract features maps within a larger network from medical images. In this section, I will discuss three of these network blocks: dense connection, inception, and attention mechanism.

Dense Connection In a dense connection network, the input of the previous layers are fed into subsequence layers. Guan et al. replaced the last block in a layer of the U-Net model with a dense block as shown in Figure 2.6 [19]. They named this architecture fully dense UNet (FD-UNet). The architecture was showed to outperform U-Net for removing artifacts from reconstructed 2D photoacoustic tomography images. Zhou et al. [50] extended this idea to the larger model as shown in Figure 2.7 and named the new architecture UNet++. This network contains intermediate layers and skip connections, resulting in a highly flexible feature fusion scheme. The downside is that the number of parameters is much higher than the basic U-Net due to dense connection. UNet++ outperforms U-Net in cell segmentation, brain tumor segmentation, liver segmentation, and lung nodule segmentation.

Inception Deep CNN networks often perform better than shallows ones, but they suffer from problems, including vanishing gradient, difficulty of network convergence, large memory usage [31]. Gu et al. proposed CE-Net by introducing the inception structure into the U-Net architecture before the middle layer [18]. Figure 2.8 depicts this structure. The inception

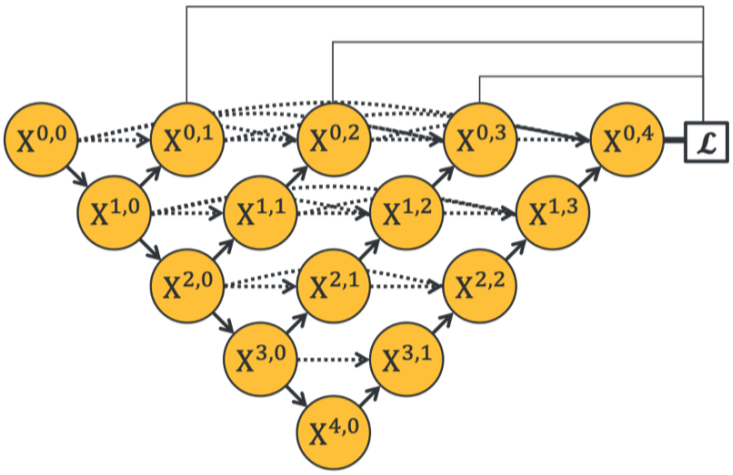

Figure 2.7: The U-Net++ architecture. Solid up and down arrows represent the encoding and decoding path, respectively, while dotted arrows represents skip connections. The yellow circles are convolutional blocks [19].

structure contains various atrous convolution blocks to extract features on a wide reception field. The inception structure improves the problems of deep CNN but is complex, leading to difficulty of model modification [31].

Attention Mechanism Another network block that has been shown to improve a network's performance for medical image segmentation is the attention block. The basic of the attention block is that it can assign different weights to the input features according to the importance. Oktay et al. incorporated the attention mechanism into the U-Net and proposed attention U-Net [38]. Figure 2.9 shows the overall architecture of attention U-Net. It is very similar to the original U-Net besides the addition of the attention gate, as shown in Figure 2.10. This network block modifies the output of the encoder before it is concatenated to the decoder features. The attention gate effectively controls the feature importance of input pixels.

The U-Net and many of its variations have two important limitations. The first one

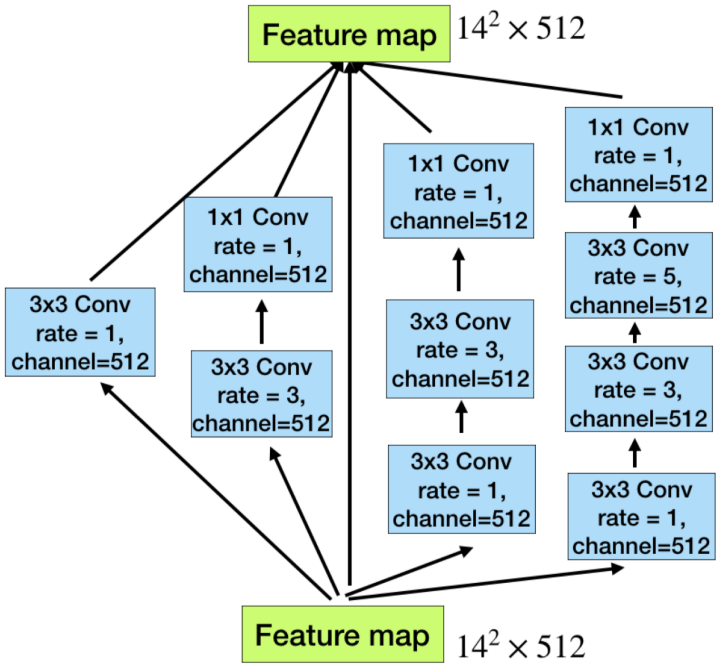

Figure 2.8: Inception structure or dense atrous convolutional block from CE-Net [18].

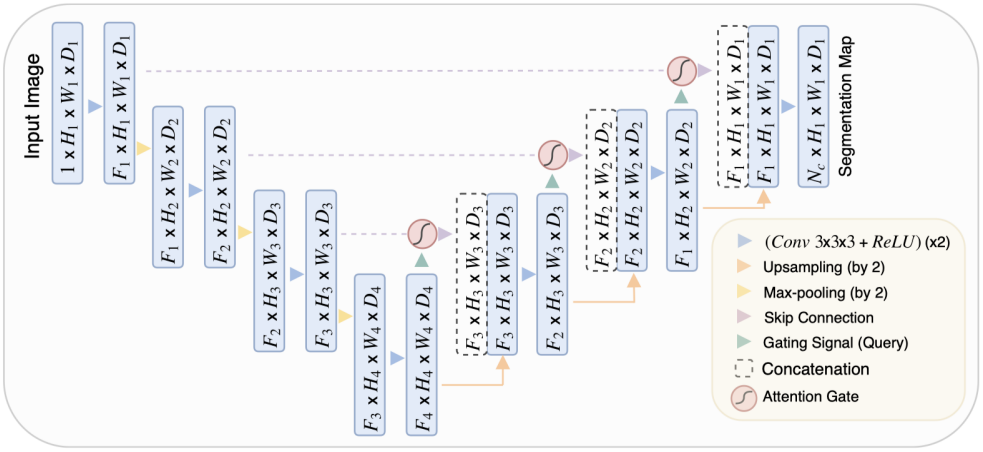

Figure 2.9: The architecture of the attention U-Net [38].

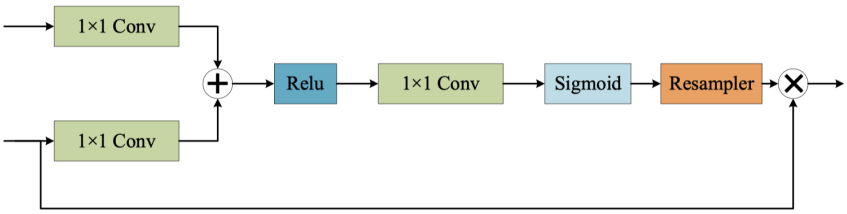

Figure 2.10: Attention gate from Attention U-Net[38], redrawn by Lei et al. [31].

is low effectiveness: the encoder and decoder depend on local operators; the second is low efficiency: a large amount of feature maps are generated due to doubling of output channels at each step [48]. Wang et al. proposed Non-local U-Net to solve these problems [48]. Non-local U-Net contains global aggregation blocks based on self-attention operator to aggregate global information without a deep encoder.

2.3.3 Data Augmentation

The quality of the data dictates the segmentation results of deep learning models. However, in the medical field, it is difficult to build high-quality datasets since the cost of data acquisition

and labeling is very high [31]. In the absence of large labeled datasets, researchers have to rely on data augmentation. Conditional generative adversarial nets (cGAN) is frequently used to generate additional data for various medical segmentation tasks. Guibas et al. proposed an architecture that contains a GAN and a cGAN to generate synthetic images of retinal fundi images [20]. Another work by Mahapatra et al. employed a cGAN to synthesize realistic chest X-ray images with different disease characteristics by conditioning the model on real samples [34].

Chapter 3

METHOD

3.1 Data and Preprocessing

Chapter 4

EXPERIMENT AND RESULT

Chapter 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

5.2 Limitation

Talk about the unique differences of bTBI and other cTBI as point out in this paper [16] and that data available to the study are related to cTBI instead of bTBI.

From an interview with the sonographer on the team, one of the many difficulties in this study is that the team do not know the exact time of the injury. If the more data could be collected, she believes that it would be best if data were collected as soon as possible after a patient is admitted. This way, the data would be "fresh", more representative of the data that would be scanned by the medics on the battle field.

5.3 Future work

BIBLIOGRAPHY

[1] Cranial ultrasound. https://www.healthlinkbc.ca/tests-treatments-medications/medical-tests/cranial-ultrasound. Accessed: 2022-02-21.

[2] Trauma care. https://www.uwmedicine.org/specialties/emergency-medicine/trauma-care. Accessed: 2022-02-22.

[3] Traumatic brain injury. https://www.hopkinsmedicine.org/health/conditions-and-diseases/traumatic-brain-injury. Accessed: 2022-02-20.

[4] Brain Tumor Foundation Inc 2000. Management and prognosis of severe traumatic brain injury.

[5] Zahangir Alom, Tarek M Taha, and Vijayan K Asari. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. *arXiv*, page 12, 2018.

[6] Wenjia Bai, Hideaki Suzuki, Chen Qin, Giacomo Tarroni, Ozan Oktay, Paul M. Matthews, and Daniel Rueckert. Recurrent neural networks for aortic image sequence segmentation with sparse annotations. *arXiv:1808.00273 [cs]*, August 2018. arXiv: 1808.00273.

[7] J Campbell, J Clark, D White, and C Jenkins. Pulsatile echo-encephalography. *Acta Neurologica Scandinavica. Supplementum*, 45:1–57, 1970.

[8] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking Atrous Convolution for Semantic Image Segmentation. 2017.

[9] Xueying Chen, Rong Zhang, and Pingkun Yan. Feature Fusion Encoder Decoder Network For Automatic Liver Lesion Segmentation. 2019.

[10] Jun Cheng, Jiang Liu, Yanwu Xu, Fengshou Yin, D. W. K Wong, Ngan-Meng Tan, Dacheng Tao, Ching-Yu Cheng, Tin Aung, and Tien Yin Wong. Superpixel Classification Based Optic Disc and Optic Cup Segmentation for Glaucoma Screening. *IEEE transactions on medical imaging*, 32(6):1019–1032, 2013.

[11] Venkateswararao Cherukuri, Peter Ssenyonga, Benjamin C Warf, Abhaya V Kulkarni, Vishal Monga, and Steven J Schiff. Learning Based Segmentation of CT Brain Images: Application to Postoperative Hydrocephalic Scans. *IEEE transactions on biomedical engineering*, 65(8):1871–1884, 2018.

[12] Patrick Ferdinand Christ, Mohamed Ezzeldin A. Elshaer, Florian Ettlinger, Sunil Tatavarty, Marc Bickel, Patrick Bilic, Markus Rempfler, Marco Armbruster, Felix Hofmann, Melvin D'Anastasi, Wieland H. Sommer, Seyed-Ahmad Ahmadi, and Bjoern H. Menze. Automatic Liver and Lesion Segmentation in CT Using Cascaded Fully Convolutional Neural Networks and 3D Conditional Random Fields. *arXiv:1610.02177 [cs]*, 9901:415–423, 2016. arXiv: 1610.02177.

[13] David Evans and W. Norman McDicken. Chapter 11, Signal Processing for Colour Flow Imaging. In *Doppler Ultrasound: Physics, Instrumentation, and Signal Processing*, pages 245–287. Chichester, England: John Wiley and Sons, Ltd, 2nd edition, 2000.

[14] Xiaorui Feng, Chaoli Wang, Shuqun Cheng, and Lei Guo. Automatic Liver and Tumor Segmentation of CT Based on Cascaded U-Net. In *Proceedings of 2018 Chinese Intelligent Systems Conference*, Lecture Notes in Electrical Engineering, pages 155–164. Springer Singapore, Singapore, 2018.

[15] Huazhu Fu, Jun Cheng, Yanwu Xu, Damon Wing Kee Wong, Jiang Liu, and Xiaochun Cao. Joint Optic Disc and Cup Segmentation Based on Multi-Label Deep Network and Polar Transformation. *IEEE transactions on medical imaging*, 37(7):1597–1605, 2018.

[16] Ling G, Bandak F, Armonda R, Grant G, and Ecklund J. Explosive blast neurotrauma. *J Neurotrauma*, June 2009.

[17] Yang Gao, Jeff M. Phillips, Yan Zheng, Renqiang Min, P. Thomas Fletcher, and Guido Gerig. Fully convolutional structured LSTM networks for joint 4D medical image segmentation. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1104–1108, Washington, DC, April 2018. IEEE.

[18] Zaiwang Gu, Jun Cheng, Huazhu Fu, Kang Zhou, Huaying Hao, Yitian Zhao, Tianyang Zhang, Shenghua Gao, and Jiang Liu. CE-Net: Context Encoder Network for 2D Medical Image Segmentation. *IEEE Transactions on Medical Imaging*, 38(10):2281–2292, October 2019.

[19] Steven Guan, Amir A. Khan, Siddhartha Sikdar, and Parag V. Chitnis. Fully Dense UNet for 2-D Sparse Photoacoustic Tomography Artifact Removal. *IEEE Journal of Biomedical and Health Informatics*, 24(2):568–576, February 2020.

[20] John T. Guibas, Tejpal S. Virdi, and Peter S. Li. Synthetic Medical Images from Dual Generative Adversarial Networks. *arXiv:1709.01872 [cs]*, January 2018. arXiv: 1709.01872.

[21] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. *arXiv*, December 2015. arXiv: 1512.03385.

[22] Jeremy J. Heit, Michael Iv, and Max Wintermark. Imaging of intracranial hemorrhage. *Journal of Stroke*, 19(1):11–27, 2017.

[23] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, November 1997. _eprint: https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf.

[24] Adnan A. Hyder, Colleen A. Wunderlich, Prasanthi Puvanachandra, G. Gururaj, and Olive C. Kobusingye. The impact of traumatic brain injuries: A global perspective. *PubMed*, 22(5):341–353, 2007.

[25] Nabil Ibtehaz and M. Sohel Rahman. MultiResUNet : Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation. *Neural Networks*, 121:74–87, January 2020. arXiv: 1902.04049.

[26] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. *arXiv:1611.07004 [cs]*, November 2018. arXiv: 1611.07004.

[27] Debesh Jha, Michael A. Riegler, Dag Johansen, Pål Halvorsen, and Håvard D. Johansen. DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation. *arXiv:2006.04868 [cs, eess]*, June 2020. arXiv: 2006.04868.

[28] Krishna Chaitanya Kaluva, Mahendra Khened, Avinash Kori, and Ganapathy Krishnamurthi. 2D-Densely Connected Convolution Neural Networks for automatic Liver and Tumor Segmentation. *arXiv:1802.02182 [cs]*, January 2018. arXiv: 1802.02182.

[29] John C. Kucewicz, Barbrina Dunmire, Nicholas D. Giardino, Daniel F. Leotta, Marla Paun, Stephen R. Dager, and Kirk W. Beach. Tissue pulsatility imaging of cerebral vasoreactivity during hyperventilation. *Ultrasound in Medicine & Biology*, 34(8):1200–1208, 2008.

[30] John C. Kucewicz, Barbrina Dunmire, Daniel F. Leotta, Heracles Panagiotides, Marla Paun, and Kirk W. Beach. Functional Tissue Pulsatility Imaging of the Brain During Visual Stimulation. *Ultrasound in Medicine & Biology*, 33(5):681–690, May 2007.

[31] Tao Lei, Risheng Wang, Yong Wan, Bingtao Zhang, Hongying Meng, and Asoke K. Nandi. Medical Image Segmentation Using Deep Learning: A Survey. *arXiv*, December 2020. arXiv: 2009.13120.

[32] Wen Li, Fucang Jia, and Qingmao Hu. Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks. *Journal of computer and communications*, 3(11):146–151, 2015. Publisher: Modern Science Publishers.

[33] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3431–3440. IEEE, 2015. ISSN: 1063-6919.

[34] Dwarikanath Mahapatra, Behzad Bozorgtabar, Jean-Philippe Thiran, and Mauricio Reyes. Efficient Active Learning for Image Classification and Segmentation Using a Sample Selection and Conditional Generative Adversarial Network. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, Lecture Notes in Computer Science, pages 580–588, Cham, 2018. Springer International Publishing.

[35] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, Levente Lanczi, Elizabeth Gerstner, Marc-Andre Weber, Tal Arbel, Brian B Avants, Nicholas Ayache, Patricia Buendia, D. Louis Collins, Nicolas Cordier, Jason J Corso, Antonio Criminisi, Tilak Das, Herve Delingette, Cagatay Demiralp, Christopher R Durst, Michel Dojat, Senan Doyle, Joana Festa, Florence Forbes, Ezequiel Geremia, Ben Glocker, Polina Golland, Xiaotao Guo, Andac Hamamci, Khan M Iftekharuddin, Raj Jena, Nigel M John, Ender Konukoglu, Danial Lashkari, Jose Antonio Mariz, Raphael Meier, Sergio Pereira, Doina Precup, Stephen J Price, Tammy Riklin Raviv, Syed M. S Reza, Michael Ryan, Duygu Sarikaya, Lawrence Schwartz, Hoo-Chang Shin, Jamie Shotton, Carlos A Silva, Nuno Sousa, Nagesh K Subbanna, Gabor Szekely, Thomas J Taylor, Owen M Thomas, Nicholas J Tustison, Gozde Unal, Flor Vasseur, Max Wintermark, Dong Hye Ye, Liang Zhao, Binsheng Zhao, Darko Zikic, Marcel Prastawa, Mauricio Reyes, and Koen Van Leemput. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2015. Place: PISCATAWAY Publisher: IEEE.

[36] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *arXiv:1606.04797 [cs]*, June 2016. arXiv: 1606.04797.

[37] Fausto Milletary. *Hough Voting Strategies for Segmentation, Detection and Tracking*. PhD thesis, Technischen Universität München, November 2017.

[38] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv*, page 10, 2018.

[39] Yuya Onishi, Atsushi Teramoto, Masakazu Tsujimoto, Tetsuya Tsukamoto, Kuniaki Saito, Hiroshi Toyama, Kazuyoshi Imaizumi, and Hiroshi Fujita. Multiplanar analysis for pulmonary nodule classification in CT images using deep convolutional neural network and generative adversarial networks. *International journal for computer assisted radiology and surgery*, 15(1):173–178, 2019.

[40] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science, pages 234–241, Cham, 2015. Springer International Publishing. ISSN: 0302-9743.

[41] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. MobileNetV2: Inverted Residuals and Linear Bottlenecks. *arXiv:1801.04381 [cs]*, March 2019. arXiv: 1801.04381.

[42] Hyunseok Seo, Charles Huang, Maxime Bassenne, Ruoxiu Xiao, and Lei Xing. Modified U-Net (mU-Net) With Incorporation of Object-Dependent High Level Features for Improved Liver and Liver-Tumor Segmentation in CT Images. *IEEE Transactions on Medical Imaging*, 39(5):1316–1325, May 2020.

[43] Tzu-Hsi Song, Victor Sanchez, Hesham EIDaly, and Nasir M Rajpoot. Dual-Channel Active Contour Model for Megakaryocytic Cell Segmentation in Bone Marrow Trephine Histology Images. *IEEE transactions on biomedical engineering*, 64(12):2913–2923, 2017.

[44] Strandness and Summer. *Hemodynamics for Surgeons*. New York: Grune and Strattion, 1975.

[45] Wei Tang, Dongsheng Zou, Su Yang, and Jing Shi. DSL: Automatic Liver Segmentation with Faster R-CNN and DeepLab. In *Artificial Neural Networks and Machine Learning – ICANN 2018*, Lecture Notes in Computer Science, pages 137–147, Cham, 2018. Springer International Publishing.

[46] Refael Vivanti, Ariel Ephrat, Leo Joskowicz, Naama Lev-Cohain, Onur A Karaaslan, and Jacob Sosna. Automatic Liver Tumor Segmentation in Follow-Up CT Scans: Preliminary Method and Results. In *Patch-Based Techniques in Medical Imaging*, Lecture

Notes in Computer Science, pages 54–61, Cham, 2016. Springer International Publishing. ISSN: 0302-9743.

[47] Shuo Wang, Mu Zhou, Zaiyi Liu, Zhenyu Liu, Dongsheng Gu, Yali Zang, Di Dong, Olivier Gevaert, and Jie Tian. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. *Medical image analysis*, 40:172–183, 2017.

[48] Zhengyang Wang, Na Zou, Dinggang Shen, and Shuiwang Ji. Non-local U-Net for Biomedical Image Segmentation. *arXiv*, February 2020. arXiv: 1812.04103.

[49] Hang Zhang, Chongruo Wu, Zhongyue Zhang, Yi Zhu, Haibin Lin, Zhi Zhang, Yue Sun, Tong He, Jonas Mueller, R. Manmatha, Mu Li, and Alexander Smola. ResNeSt: Split-Attention Networks. *arXiv:2004.08955 [cs]*, December 2020. arXiv: 2004.08955.

[50] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Transactions on Medical Imaging*, 39(6):1856–1867, June 2020.

[51] Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *arXiv:1606.06650 [cs]*, June 2016. arXiv: 1606.06650.