Image Segmentation of Breast Cancer Ultrasounds using U-Net Architecture

Michelle Cheng
Department of Electrical and Computer Engineering
Toronto Metropolitan University
Toronto, Canada

***Abstract*—1 in 8 Canadian women will develop breast cancer [12] and it continues to be one of the most common cancer among Canadian women. Ultrasound imaging is often used for breast cancer patients during the early diagnostic stages when anomalies are found. Analysis of these Ultrasounds can be a difficult task due to the low contrast, high speckle noise and shadows present in ultrasound Imaging. Image segmentation can aid healthcare professionals by extracting a region of interest that can potentially be tumorous. A Convolutional Neural Networks (CNN) is a neural network used for image processing. The U-Net architecture is a fully convolutional neural network used in biomedical image segmentation. Automating the image segmentation of breast cancer ultrasounds can support healthcare professionals with image screening and improve early cancer diagnosis for patients.**

***Index Terms*—Image Segmentation, Convolutional Neural Networks, U-Net, Ultrasound Imaging**

I. Introduction

It is estimated that 1 in 8 Canadian women will develop breast cancer during their lifetime and 1 in 36 die from it.According to Canadian Cancer Statistics, it is one of the most common cancer among Canadian women with an estimated 30,500 women being impacted in this 2024 year alone [12]. As the medical world advances, significant impacts with improved screening and treatment practices continue to be made. This includes the integration of Convolutional Neural networks (CNN) which can be used to aid healthcare professionals with image screening and improve early cancer diagnosis. The rate of when cancerous cells are detected is highly correlated to the survival rate with early detection usually improving survival rates among patients. Ultrasounds are commonly done during the early diagnostic process for breast cancer, often times to observe any anomalies found during a mammogram or regular check. A trained sonographer would analyze the image and this can often be the first indicator to provide more information on the anomaly. At the same time, this is a complex task due to variability between patient cases and the low contrast, high speckle noise in Ultrasound imaging. Medical image segmentation has the potential to aid healthcare professors process the characteristics of an image by extracting a region of interest (ROI). A U-Net architecture achieves this goal as different encoder and decoder layers separate and locate important features in the image.

The following paper is organized as follows: In Section II, a literature Survey of relating research is summarized, Section III provides a dataset description, Section IV describes the Problem Statement and Section V describes the Model Methodology.

II. Literature Survey

Image segmentation is a large topic in the CNN research world. In a paper from Shandong University, the researchers presents a novel automatic way to detect tumors in breast cancer ultrasound images [1]. It highlights the challenges of ultrasound imaging including having lower quality and higher speckled noise. The research defines an automated image detection method, which uses fuzzy entropy principle. The image is first normalized before *fuzzification* is applied to locate the anomalies. They then extracted the objects from the background, increasing the contrast and drew boundary lines using watershed segmentation. The algorithm was tested on a collection of breast cancer ultrasound images obtained with a 5-14 MHz and could accurately detect the tumors.

Another challenge with Ultrasound imaging is the shadows, blurred boundaries and overall noise within the image. Traditional methods like Butterworth and adaptive median filters struggle with locating tumor regions from the shadows. By utilizing text features, extraction from different ultrasound images can be combined with the original creating a RGB representation. In a paper published by Changchun University of Science and Technology, the research evaluates the various preprocessing techniques alongside segmentation algorithms. First-order text features include mean entropy and describes intensity distribution of the pixels in the ROI. Second-order texture features describes complexity of the spatial distribution of the ROI and include autcorrelation and homogeneity. Two experiments were performed, with edge detection and the U-Net architecture. In experiment 1, edge detection and Butterworth high-pass filters, adaptive median filter and image processing filters (First-order and Second-order) are used to preprocess the image. In experiment 2, the U-net network and same preprocessing methods are used to preprocess the image. In the final experiment results, all four preprocessing methods improves the overall Dice-Similarity-Coefficient (DSC) and can effectively improve the image segmentation of breast tumor ultrasound images[2].

Modifications to the U-Net model can also be done for improving image segmentation in ultrasound. U-Net uses encoders and decoders in order to extract the ROI from an image.

The encoder uses the CNN to reduce spatial dimensions of the image and increase the number of feature maps. A decoder then reconstructs the image by adding spatial dimensions and decreasing the feature maps. In the study conducted by Technical University of Cluj-Napoca, the encoder has three convolutional layers with stride 2 and kernel 2x2. This applies to the decoder as well. The image data is also normalized after the second and third convolutional layers and regularization is applied on all layers. The dataset for this study includes 780 ultrasound images split between three groups: healthy breast tissue, benign tumors and malignant tumors. Image augmentation was also done to expand the dataset with 3900 images and their corresponding masks created. The experiment results achieved a accuracy of 98.23% on the training set and 94.28% on the testing set which were promising towards the use in medical image analysis[3].

Fine-Tuning U-Net also plays a role on the overall outcome as seen in a study which focused on refining the U-Net architecture for ultrasound image segmentation by altering specific layers within the U-Net. Adjustments in different layers led to different segmentation results. One of the findings include modifying the number of filters in the convolutional layers which effect the ability for the model to find fine details. The paper also highlights the use of pooling operations for feature reduction and spatial preservation, as well as skip connections which effect the communication between encoder and decoder layers. These factors also played a role in the overall image segmentation accuracy. The over conclusion of the study shows that the fine-tuning of shallow layers for small data sets improves the results in ultrasound image segmentation.

Image segmentation is also effective for identifying information in other medical cases. One example is in Gynaecological Ultrasound images. In a study conducted by INESC TEC, a U-Net architecture is used to segment gynaecological ultrasound images and it notes the enhancements that can be made to improve the accuracy in image segmentation using the U-Net approach. One of the continuing challenges with ultrasound segmentation is the complexity of the image and speckle noise. However, certain modifications can be made to enhance the segmentation. One conclusion drawn was the increase in the depth of a U-Net with more layers which improves the details captured. Filter size also affect how the final image is captured with smaller filters showing finer details and larger filters showing broader patterns. Pooling operations affect the information preservation and reduction done by the encoders and decoders. By adjusting operations such as max pooling or average pooling for the problem, it can reduce the dimension of the data while ensure proper image segmentation of the ROI. Skip connections also affect the information flow between encoder and decoders, along with data augmentation which can significantly improve the results of a U-Net. This includes techniques such as rotation, scaling and intensity normalization to create a more diverse training set. All of these modifications can improve the U-Net to handle more challenges cases.

Another case of image segmentation with U-Net can be seen in a study which was focused on segmentation lung x-rays in COVID-19 infected patients. The paper was published from Jiangsu university of Science and Technology, which breaks down the U-Net into three parts: Coding module, encoding decoding and prediction. The first part focuses on the down-sampling operation of an image, before passing to the encoding module to obtain one feature image after another, totalling to five effective feature layers. The decoding module uses the characteristics obtained in the five layers to produce a final blend of all the features within all layers. The prediction is completed with SoftMax function to classifying the feature image. This SoftMax function classifies the pixel which ensures that the output segmentation has the same input image.

The application of CNN Architecture and U-Net Architecture is also seen in Brain Tumour Detection. Brain tumors are an abnormal cell growth in a brain which can cause increased intracranial pressure. CNN uses strong object detection and when coupled with image segmentation ability in U-Net can achieve accurate localization of tumors within medical images. In a study conducted by Romaiah Institute of Technology, a 98% accuracy rate was achieved when combining the two architectures together. The brain MRI images were pre-processed including resizing to 256 pixels x 256 pixel and the normalization of the images. The CNN detects the tumour regions by picking up patterns seen in images with tumors present. The U-Net refines the segmentation results. The model is trained on BITE' dataset and performance is measured using quantitative measurement such as dice coefficient, accuracy and correct tumor localization compared to the target mask.

The U-net architecture can also be expanded by combining with other modules. In a paper published from Alexandria University, a ConvMixer module and convolutional layers are used to improve the image segmentation accuracy. This also reduces the number of parameters compared to a traditional U-net. In the study, the encoders, decoders and skip connections are used just like a traditional U-net architecture. Three ConvMixer blocks are used at the bottom where they mix the spatial and channel locations. This purposed model has a reduced number of parameters of just 1.77 million instead of the traditional 31.1. million parameters. Hyperparameter tuning of 3 parameters are also mentioned in this paper which includes: depth, number of channels and kernel size.

Research into other improvements for medical image segmentation also include CNN-Transformer based approaches. Fully Convolutional Neural Networks and U-Net pose a challenge when capturing long-range dependencies. Transformer-based models can capture global context but often miss local details. Medical imaging such as ultrasound have high speckle noise which led to the study of a hybrid model integrating CNN and Transformers. ConvMixer modules are employed in the CNN while the Transformer branch uses Progressive Atrous Spatial Pyramid Pooling (PASPP) modules to improve feature extraction. PASPP uses dilated convolution with varying dilation rates to improve the receptive fields, which allows more information to be found in the image without increasing computational complexity. By combining both, the chances of obtaining local and global content are increased which leads to increased accuracy in the segmentation.

The transformer model can also be combined further with Convmixer to produce a ConvmixFormer-Unet. The transformer-based model accounts for long-range dependencies within the image but can be computationally intensive. CNN's are effective in local dependencies but often fall to locate global contexts. A ConvmixFormer-Unet would have ConvMixer's local context capture and global context capture of Transformer models. The ConvmixFormer block in the encoder and ConvMix Transpose blocks in the decoder show improved performance and reduced parameter count. This module utilizes a multi-scale ConvMix former block in the encoder stage, channel-mix layer for global dependencies and depth wise convolutions in the decoder. This study used three datasets: Gland Segmentation, International Skin Imaging Collaboration and Data Science Bowl. The experiment results indicate that the ConvmixerFormer U-net hybrid model improves the performance while reducing number of parameters. By combining convolutional operations with various kernel sizes and obtaining global information with channel-mix layers, the medical image segmentation can be improved.

III. Dataset Description

The Breast Cancer Ultrasound Image Dataset is a open source dataset found on Kaggle.com. It contains a variety of resolutions, tumor types and forms. This includes breast lesions such as ductal carcinoma in situ (DCIS), invasive ductal carcinoma (IDC), fibroadenomas, cysts and more. The dataset is split between the following folders below and each tumor has a corresponding target mask. The mask separates the ROI indicated in white as depicted in Fig.1 Benign tumor with Mask Target.

- Benign - 437 benign tumor patients
- Malignant - 210 malignant tumor patients
- Normal -133 total normal patients images

Fig. 1. Example of Benign Tumor Image with Corresponding Target Mask from the Dataset.

IV. Problem Statement

The objective of this paper is to automate the image segmentation of breast cancer tumors from a series of ultrasound images. Ultrasound imaging has low contrast and high speckle noise in the image which can make detection of the anomalies hard to analyze. Medical image segmentation can process the characteristics of an image by extracting the ROI from the rest of the image. This in return, can improve the diagnostic and screening process for patients. A CNN is used notably for finding classes and patterns within image data. The U-Net architecture is a CNN that is famous in biomedical imaging segmentation including, its ability to handle inputs with various size and ratios.

V. Model Methodology

The U-Net architecture was created by Olaf Ronneberger, Philipp Fischer and Thomas Brox. It is a fully convolutional neural network that is famous for its applications in localization and feature capturing. Its ability to do pixel-level localization is what makes it a popular architecture for biomedical image segmentation.

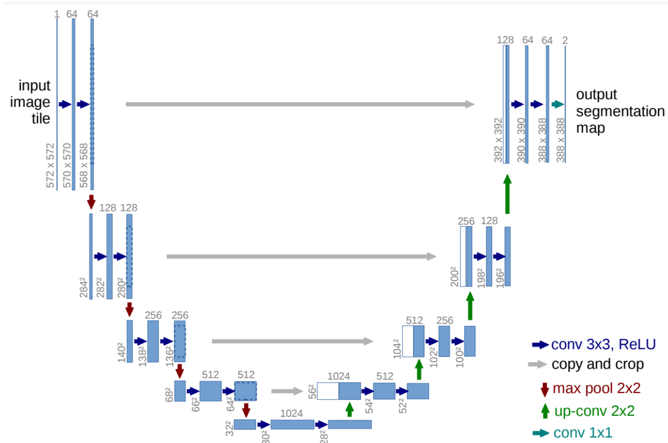

Fig. 2. U-Net Architecture as illustrated in [11]

The U-Net gets its name by its U-shape as seen in Fig.2 U-Net Architecture. The contracting path on the left of the path is symmetrical to the expansive path on the right side, with the left being the encoder block and right as decoder block. The encoder block captures the information and reduces the resolution of the image input. Decoder on the other hand, use the information obtained by using skip connections to make the segmentation map. An image is passed through the 3x3 convolutional layers with the ReLU activation function. The original input image size is 572x572x1. The encoder has a decreasing image size with 2x2 max polling operation with stride 2. The filters are also increasing and being doubled at every down sampling step. When the decoder is reached, the filters decrease and are halved each time. Skip connections are used to retain the information from previous layers. A final 1x1 convolution is used to map the feature vector giving the desired ROI with a output map 388x388x2.

Obtaining the image segmentation of the breast cancer ultrasounds will require the following steps below.

1) Image preprocessing
2) Training

3) Testing
4) Error results between model output and target mask.

REFERENCES

[1] L. Zhang, Y. Ren, C. Huang and F. Liu, ”A novel automatic tumor detection for breast cancer ultrasound Images,” 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Shanghai, China, 2011, pp. 401-404

[2] S. Cai, Y. Zhu, J. Zhang and T. Liu, ”A Study on the Combination of Image Preprocessing Method Based on Texture Feature and Segmentation Algorithm for Breast Ultrasound Images,” 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2022, pp. 760-764

[3] A. Ciobotaru, D. Gota, T. Covrig and L. Miclea, ”Reliable Breast Cancer Detection from Ultrasound Images using Image Segmentation,” 2024 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), Cluj-Napoca, Romania, 2024, pp. 1-5

[4] M. Amiri, R. Brooks and H. Rivaz, ”Fine-Tuning U-Net for Ultrasound Image Segmentation: Different Layers, Different Outcomes,” in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 67, no. 12, pp. 2510-2518

[5] S. Marques et al., ”Segmentation of gynaecological ultrasound images using different U-Net based approaches,” 2019 IEEE International Ultrasonics Symposium (IUS), Glasgow, UK, 2019, pp. 1485-1488

[6] C. Wang, J. Zhu, K. Snu, D. Li, Z. Wang and H. Yuan, ”A COVID-19 medical image Segmentation method based on U-NET,” 2021 IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE), Shanghai, China, 2021, pp. 1-4

[7] S. R. Totad, Meeradevi, B. J. Sowmya, P. Boppudi, N. Yaswanth and I. Nivedith, ”Brain Tumour Detection and Segmentation using CNN Architecture and U-Net Architecture,” 2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS), Kalaburagi, India, 2023, pp. 1-5

[8] S. AbdElhakem, S. Basiony and M. Torki, ”ConvMixer-UNet: A Lightweight Network for Breast Lesion Segmentation in Ultrasound Images,” 2023 20th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA), Giza, Egypt, 2023, pp. 1-5

[9] T. -T. Tran, D. -T. Vu, T. -H. Nguyen and V. -T. Pham, ”A CNN-Transformer-based Approach for Medical Image Segmentation,” 2023 International Conference on System Science and Engineering (ICSSE), Ho Chi Minh, Vietnam, 2023, pp. 22-27

[10] V. -Q. Nguyen, Q. -H. Nguyen and T. -T. Tran, ”ConvmixFormer-Unet: A New Approach for Medical Image Segmentation based on Convmixer and Transformer,” 2023 12th International Conference on Control, Automation and Information Sciences (ICCAIS), Hanoi, Vietnam, 2023, pp. 662-667

[11] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015.

[12] Canadian Cancer Society / Société canadienne du cancer, “Breast cancer statistics,” Canadian Cancer Society, https://cancer.ca/en/cancer-information/cancer-types/breast/statistics