As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 8(04):370-375, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0804062
Available online at: http://www.ijmtst.com/vol8issue04.html

Segmentation of masses using U-Net model

B. Bindu Sri | G. Venkata Sai Praveen Kumar | J. Sravya | D. Pavan Kumar

Departmentof Electronics and Communication Engineering, RVR&JC College of Engineering, Chowdavaram, Guntur, India.
*Corresponding Author Email ID: bankabindusri@gmail.com

To Cite this Article
B. Bindu Sri, G. Venkata Sai Praveen Kumar, J. Sravya and D. Pavan Kumar. Segmentation of masses using U-Net model. International Journal for Modern Trends in Science and Technology 2022, 8(04), pp. 370-375. https://doi.org/10.46501/IJMTST0804062

Article Info
Received: 18 March 2022; Accepted: 15 April 2022; Published: 20 April 2022.

ABSTRACT

The diagnosis of breast cancer in early stage is essential for successful treatment. Deep convolutional neural networks (CNNs) have emerged as a new paradigm for diagnosis of breast cancer. This involves segmentation of masses using deep learning techniques. Segmentation plays a significant role in image analysis and includes detection, feature extraction, classification, and treatment. For segmentation U-Net model is used to segment breast area from the input images. Specifically, this method is constructed in a dual-problem manner, with an additional consideration of important shape and boundary knowledge. This involves detection and segmentation of masses using deep learning techniques. For segmentation U-Net model is used to segment breast area from the given input images. The goal of semantic image segmentation is to label each pixel of an image with a corresponding class of what is being represented. Because we're predicting for every pixel in the image, this task is commonly referred to as dense prediction. The U-Net model would achieve the best segmentation in both high and low resolution.

KEYWORDS: Breast cancer, Segmentation, Deep learning, Computer aided detection, U-Net.

1. INTRODUCTION

Breast cancer is the deadliest disease in women, accounting for more than half a million deaths per year. It is the second most incident type of cancer in the world. Even in the face of advances in treatment, early diagnosis is a crucial point for defining the patient outcome. Nonetheless, mass projection is directly affected by its shape and location, the similarity of density between healthy and disordered tissues, and by the technical and instrumental quality. Besides, this analysis requires the professionals' time and a high degree of attention to analyzing many cases. Convolutional Neural Network (CNN) is used to perform the classification in regions of mass or not.CNN sets are used for segmentation and classification, respectively, in benign or malignant. Nevertheless, most of the time, the authors make a selection of images in the methodologies, however not generally detailing the process used.

2. MATERIALS DESCRIPTION

Breast cancer is one of the most common causes of death among women worldwide. Early detection helps in reducing the number of early deaths. The data reviews the medical images of breast cancer using ultrasound scan. Breast Ultrasound Dataset is categorized into three classes: normal, benign, and malignant images. Breast ultrasound images can produce great results in classification, detection, and segmentation of breast cancer when combined with machine learning.

The data collected at baseline include breast ultrasound images among women in ages between 25 and 75 years old. This data was collected in 2018. The number of patients is 600 female patients. The dataset consists of 780

images with an average image size of 500*500 pixels. The images are in PNG format. The ground truth images are presented with original images. The images are categorized into three classes, which are normal, benign, and malignant. Ultrasound scan is used for examination and early detection of breast cancer. Moreover, it is safe in comparison to other radiology imaging techniques. Breast Ultrasound dataset can be used to train machine learning models which can classify, detect and segment early signs of masses or micro-calcification in breast cancer. Researchers with interest in classification, detection,

and segmentation of breast cancer can utilize this data of breast ultrasound images, combine it with others' datasets, and analyse them for further insights. The data is comprehensive, containing breast cancer states (normal, benign, and malignant). This dataset is – to our best knowledge – the first breast ultrasound dataset publicly available.

3. METHODOLOGY

We have chosen ultrasound imaging for this process and the data set used is Breast Ultrasound Images dataset. Ultrasound imaging is also accompanied by X-ray mammograms to assess if a mass on the breast that appears on mammogram images is a stable or less harm tissue cyst that contains fluid. In the systems of ultrasound image screening, high-frequency sound waves can be produced, which permeate the human body. During the bounce waves off the tissue of the boundary of the human body, distinctive echoes will be created, which a computer utilizes to produce an image called a sonogram. Since the cysts of fluid-filled have a distinguished "sound signature" instead of a solid mass, experts may utilize ultrasound precisely to identify cysts that are existed in the breast. Some palpable lesions (lumps) that can be assessed by ultrasound images, which can assist radiologists. In contrast, these lumps are not easy to find on mammogram images, particularly within the dense breasts of the women. Research of women with lumps showed that ultrasound images were very efficient in the identification of benign that may have removed, which required more than 50% of the biopsies. More so, some other works showed that the ultrasound images might also help in classifying lesions of non-palpable solid into benign or malignant. Further research suggests that the combination of ultrasound combined with x-ray mammography images might enhance breast cancer screening accuracy and allow early-stage tumor detection in women with dense breasts. More research is required to determine the efficacy of ultrasound as a screening method used in combination with mammography. While ultrasound may be helpful as an alternative to mammography, but when used alone, it will be affected by some limitations on breast cancer identification. Predominantly the ultrasound images are not able to identify small tumors that are smaller than five millimeters is about one-quarter inch—also, abnormalities or microcalcifications associated with proven kinds of breast cancers. However, current progress in the technology if ultrasound images can help in resolving some of the limitations and expand their use in identifying cancer of the breast. But, at this point in development, their ultimate efficiency in the detection of breast cancer cannot be predicted.

4. SEGMENTATION

The main aim of the project is to segment the mass abnormalities from the breast ultra sound images Image segmentation is a set of procedures used to split a screened image into several regions. In processing medical images, image segmentation plays a pivotal role in segmenting the representation of tissues of interest from the background. As ultrasound (US) images are always crowded with image noises, it becomes practically challenging to segment tumors. In the medical examination of image segmentation, the task is typically accomplished by labor-intensive human efforts to perform tracing. Such a strategy is laborious, involves specific skillset, requires substantial experience, and consumes a significant amount of time. Therefore, developments in CNN systems are also deemed very important because radiologists can accurately and efficiently diagnose breast cancer.

There are different segmentation methods which are currently used in mammogram segmentation, such as classical segmentation, machine learning segmentation and deep learning segmentation. The classical segmentation method depends on digital image processing and mathematics to segment the image which includes the following:

i. Edge-based segmentation methods (EBS), such as canny edge detection, active contour, Sobel, energy minimisation, and contour
ii. Threshold-based segmentation methods (TBS) includes Otsu thresholding, morphological thresholding, adaptive thresholding, manual thresholding, Kittler's optimal thresholding, and global and local thresholding
iii. Region-based segmentation (RBS), such as watershed, rough set theory, partial region growing, and marker controller.

Machine learning segmentation methods which include the following:

i. Unsupervised machine learning methods (USML), such as fuzzy *C*-clustering, *k*-means clustering, novel clustering, and hierarchical *k*-clustering
ii. Supervised machine learning methods (SML), such as support vector machine (SVM) and extreme learning machine.

Deep learning segmentation methods which include the following:

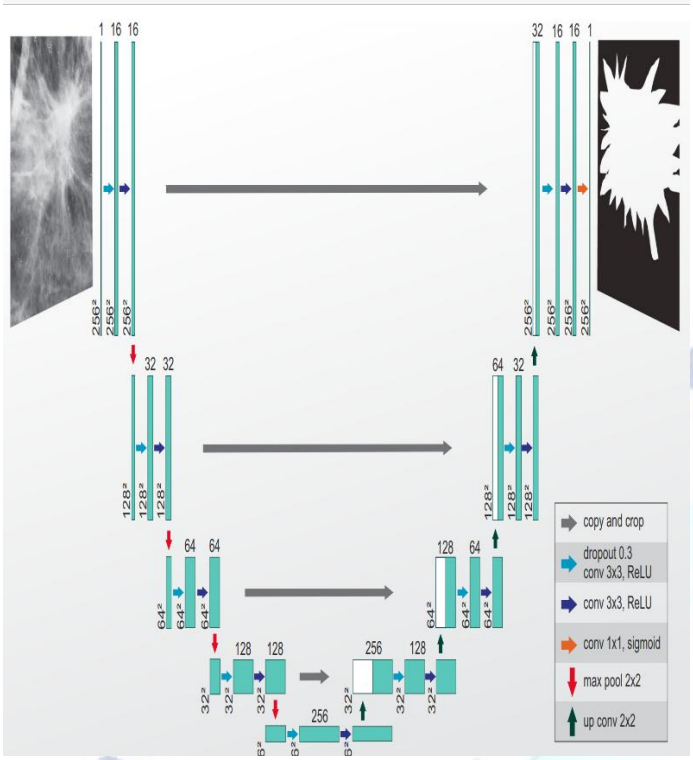
i. Deep learning segmentation (DL), such as SegNet, U-Net, and fully convolutional neural networks (FCN).

A convolutional neural network (CNN) is a specific type of artificial neural network to analyze data. CNNs apply to image processing. CNNs are considered the most effective architecture for image segmentation. In CNNs, there could be multiple hidden layers, which perform feature extraction from the image by doing calculations. The CNN cannot process the whole image at once. It scans the image, looking at a small "filter" of several pixels each time until it has mapped the entire image. The final output layer has a large receptive field and corresponds to the height and width of the image, while the number of channels corresponds to the number of classes. The convolutional layers classify every pixel to determine the context of the image, including the location of objects. CNNs are a fundamental example of deep learning.

5. U-NET MODEL

The architecture looks like a 'U' which justifies its name. This architecture consists of three sections: The contraction, The bottleneck, and the expansion section. The contraction section is made of many contraction blocks. Each block takes an input applies two 3X3 convolution layers followed by a 2X2 max pooling. The number of kernels or feature maps after each block doubles so that architecture can learn the complex structures effectively. The bottommost layer mediates between the contraction layer and the expansion layer. It uses two 3X3 CNN layers followed by 2X2 up convolution layer. But the heart of this architecture lies in the expansion section. Similar to contraction layer, it also consists of several expansion blocks. Each block passes the input to two 3X3 CNN layers followed by a 2X2 upsampling layer. Also, after each block number of feature maps used by convolutional layer get half to maintain symmetry. However, every time the input is also get appended by feature maps of the corresponding contraction layer. This action would ensure that the features that are learned while contracting the image will be used to reconstruct it. The number of expansion blocks is as same as the number of contraction block. After that, the resultant mapping passes through another 3X3 CNN layer with the number of feature maps equal to the number of segments desired. It consists of encoder and decoder paths.

- The left-hand side is the contraction path (Encoder) where we apply regular convolutions and max pooling layers.
- In the Encoder, the size of the image gradually reduces while the depth gradually increases starting from 128x128x3 to 8x8x256.
- This basically means the network learns the "WHAT" information in the image, however it has lost the "WHERE" information



- The right-hand side is the expansion path (Decoder) where we apply transposed convolutions along with regular convolutions
- In the decoder, the size of the image gradually increases and the depth gradually decreases. Starting from 8x8x256 to 128x128x1
- Intuitively, the Decoder recovers the "WHERE" information (precise localization) by gradually applying up-sampling
- To get better precise locations, at every step of the decoder we use skip connections by concatenating the output of the transposed convolution layers with the feature maps from the Encoder at the same level.
- After every concatenation we again apply two consecutive regular convolutions so that the model can learn to assemble a more precise output.
- This is what gives the architecture a symmetric U-shape, hence the name UNET

VI. UNDERSTANDING CONVOLUTION, MAX POOLING AND TRANSPOSED CONVOLUTION

6.1 Convolution Operation:

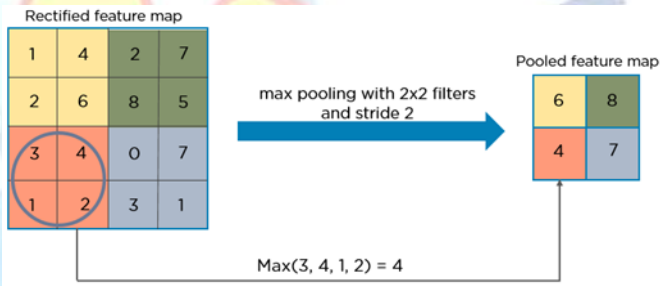
There are two inputs to a convolutional operation i) A 3D volume (input image) of size (nin x nin x channels ii) A set of 'k' filters (also called as kernels or feature extractors) each one of size (f x f x channels), where f is typically 3 or 5. The output of a convolutional operation is also a 3D volume (also called as output image or feature map) of size (nout x nout x k). The relationship between nin and nout is as follows:

$$n_{out} = \left\lfloor \frac{n_{in} + 2p - k}{s} \right\rfloor + 1$$

n_{in}: number of input features
n_{out}: number of output features
k: convolution kernel size
p: convolution padding size
s: convolution stride size

6.2 Maxpooling Operation:

In simple words, the function of pooling is to reduce the size of the feature map so that we have fewer parameters in the network.



Basically, from every 2x2 block of the input feature map, we select the maximum pixel value and thus obtain a pooled feature map. Note that the size of the filter and strides are two important hyper-parameters in the max pooling operation. The idea is to retain only the important features (max valued pixels) from each region and throw away the information which is not important. By important, I mean that information which best describes the context of the image. A very important point to note here is that both convolution operation and specially the pooling operation reduce the size of the image. This is called as **down sampling**. In the above example, the size of the image before pooling is 4x4 and after pooling is 2x2. In fact, down sampling basically means converting a high-resolution image to a low-resolution image. Thus, before pooling, the information which was present in a 4x4 image, after pooling, (almost) the same information is now present in a 2x2 image. Now when we apply the convolution operation again, the filters in the

next layer will be able to see larger context, i.e. as we go deeper into the network, the size of the image reduces however the receptive field increases.

Intuitively we can make the following conclusion of the pooling operation. By down sampling, the model better understands **"WHAT"** is present in the image, but it loses the information of **"WHERE"** it is present.

6.3 Need for upsampling:

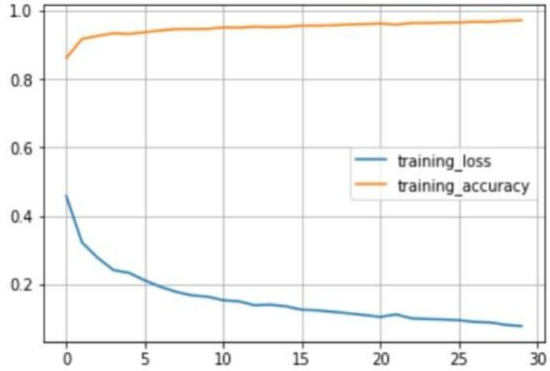
As stated previously, the output of semantic segmentation is not just a class label or some bounding box parameters. In-fact the output is a complete high-resolution image in which all the pixels are classified. Thus, if we use a regular convolutional network with pooling layers and dense layers, we will lose the "WHERE" information and only retain the "WHAT" information which is not what we want. In case of segmentation, we need both "WHAT" as well as "WHERE" information. Hence there is a need to up sample the image, i.e., convert a low-resolution image to a high-resolution image to recover the "WHERE" information. In the literature, there are many techniques to up sample an image. Some of them are bi-linear interpolation, cubic interpolation, nearest neighbor interpolation, unpooling, transposed convolution, etc. However, in most state-of-the-art networks, transposed convolution is the preferred choice for up sampling an image.

6.4 Transposed Convolution:

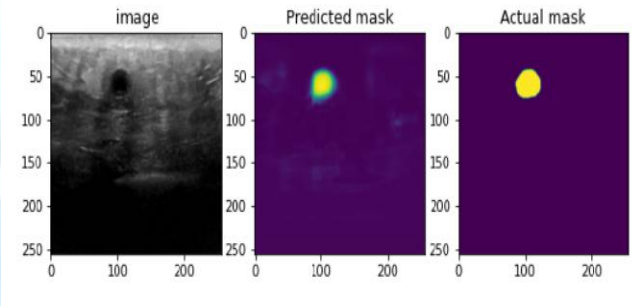
Transposed convolution (sometimes also called as deconvolution or fractionally strided convolution) is a technique to perform up sampling of an image with learnable parameters. However, on a high level, transposed convolution is exactly the opposite process of a normal convolution i.e., the input volume is a low resolution image and the output volume is a high resolution image. By just taking the transpose of the filter matrix, we can reverse the convolution process, hence the name transposed convolution.

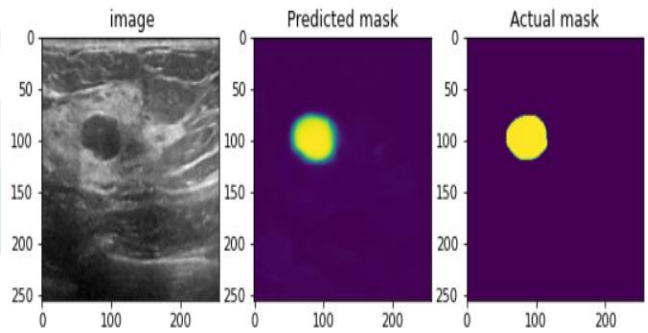
7. RESULTS

This work presented the segmentation of breast ultrasound images.



According to the images produced by the models, we perceive that the network tends to classify sets of pixels with a higher intensity as masses. This behavior is desired, but in dense breasts, usually of young women, the network ends up producing false positives.





Therefore, a methodology adapted to breast density may contribute to the improvement of classification accuracy. Furthermore, the inclusion of clinical data in the training process, such as age, family history, smoking, weight, and use of hormone replacement, is an aspect to be considered in future models. Another future work would be the use of more data to improve the accuracy of the model.

8. CONCLUSION

Image segmentation is an important problem. U Net contributed significantly in such research. Many new architectures are inspired by U Net. But still, there is so much to explore. There are so many variants of this architecture in the industry and hence it is necessary to understand the first one to understand them better. After completion of segmentation process, we get an image with partition of masses with specific color by using UNET model. Then we use the resultant image for processing to differentiate malignant (noncancerous) and benign (cancerous). After diagnosis of mammography one can say that ,will the patient suffer from breast cancer or not We may save a life of a woman from being died with cancer.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Al-antari MA et al.: A fully integrated computer-aided diagnosis system for digital x-ray mammograms via deep learning detection, segmentation, and classification. International Journal of Medical Informatics 117:44–54, 2018

[2] Al-masni MA et al.: Simultaneous detection and classification of breast masses in digital mammograms via a deep learning yolobased cad system. Computer Methods and Programs in Biomedicine 157:85–94, 2018

[3] Felzenszwalb PF, Huttenlocher DP: Efficient graph- based image segmentation. International Journal of Computer Vision 59(2):167–181, 2004

[4] Caldas FAA et al.: Controle de qualidade e artefatos em mamografia. Revista Radiologia Brasileira 38(4):295–300, 2005

[5] Dromain C et al.: Computed-aided diagnosis (CAD) in the detection of breast cancer. European journal of radiology,2013.