http://doi.org/10.35784/iapgos.5319

received: 24.07.2023 | revised: 06.09.2023 | accepted: 27.09.2023 | available online: 30.09.2023

SEGMENTATION OF CANCER MASSES ON BREAST ULTRASOUND IMAGES USING MODIFIED U-NET

Ihssane Khallassi, My Hachem El Yousfi Alaoui, Abdelilah Jilbab

Mohammed V University in Rabat, Electronic Systems Sensors and Nanobiotechnology, National School of Arts and Crafts, Rabat, Morocco

Abstract. *Breast cancer causes a huge number of women's deaths every year. The accurate localization of a breast lesion is a crucial stage. The segmentation of breast ultrasound images participates in the improvement of the process of detection of breast anomalies. An automatic approach of segmentation of breast ultrasound images is presented in this paper, the proposed model is a modified u-net called Attention Residual U-net, designed to help radiologists in their clinical examination to determine adequately the limitation of breast tumors. Attention Residual U-net is a combination of existing models (Convolutional Neural Network U-net, the Attention Gate Mechanism and the Residual Neural Network). Public breast ultrasound images dataset of Baheya hospital in Egypt is used in this work. Dice coefficient, Jaccard index and Accuracy are used to evaluate the performance of the proposed model on the test set. Attention residual u-net can significantly give a dice coefficient = 90%, Jaccard index = 76% and Accuracy = 90%. The proposed model is compared with two other breast segmentation methods on the same dataset. The results show that the modified U-net model was able to achieve accurate segmentation of breast lesions in breast ultrasound images.*

Keywords: convolutional neural network, segmentation, u-net, residual neural network

SEGMENTACJA MAS NOWOTWOROWYCH NA OBRAZACH ULTRASONOGRAFII PIERSI Z UŻYCIEM ZMODYFIKOWANEGO MODELU U-NET

Streszczenie. *Każdego roku rak piersi powoduje ogromną liczbę zgonów kobiet. Dokładna lokalizacja zmiany piersi jest kluczowym etapem. Segmentacja obrazów ultrasonograficznych piersi przyczynia się do poprawy procesu wykrywania nieprawidłowości piersi. W tym artykule przedstawiono automatyczne podejście do segmentacji obrazów ultrasonograficznych piersi, proponowany model to zmodyfikowany U-net, nazwany Attention Residual U-net, zaprojektowany w celu wspomagania radiologów podczas badania klinicznego, w celu odpowiedniego określenia zasięgu guzów piersiowych. Attention Residual U-net jest połączeniem istniejących modeli (konwolucyjną siecią neuronową U-net, Attention Gate Mechanism i Residual Neural Network). W tym badaniu wykorzystano publiczny zbiór danych obrazów ultrasonograficznych piersi szpitala Baheya w Egipcie. Do oceny wydajności zaproponowanego modelu na zbiorze testowym wykorzystano współczynnik Dice'a, indeks Jaccarda i dokładność. Attention Residual U-net może znacznie przyczynić się do uzyskania współczynnika Dice'a równego 90%, indeksu Jaccarda równego 76% i dokładności równiej 90%. Proponowany model został porównany z dwoma innymi metodami segmentacji piersi na tym samym zbiorze danych. Wyniki pokazują, że zmodyfikowany model U-net był w stanie osiągnąć dokładną segmentację zmian piersiowych na obrazach ultrasonograficznych piersi.*

Słowa kluczowe: konwolucyjna sieć neuronowa, segmentacja, u-net, rezydualna sieć neuronowa

Introduction

Breast cancer is the first causes of death by cancer in women, known as the most detected cancer in the world by recent years, it appears most of the time as a mass in the breast [5, 13]. According to the world health organization early diagnosis of breast cancer can have a high chance of cure and reduce the mortality rate by 40% [7].

For the diagnosis of breast anomalies ultrasound images are often used to create a clear and precise representation of the breast, and they have shown important efficacy in the detection of breast cancer [2, 15].

The purpose of Lesion segmentation in breast ultrasound images is to localize, detect breast tumors in ultrasound images and to assist specialists in their clinical examination instead of taking more time on a single patient because manual analysis for breast ultrasound images needs more works and experience. Manual segmentation depends on manual tracing or visual inspection of the abnormal region requires various medical experts to select exactly the abnormal regions, particularly for large-scale breast ultrasound images which leads to a subjective interpretation between observers that makes the breast segmentation process fastidious and time-consuming task. Automatic segmentation that does not need human intervention which is often used to help radiologists to detect breast masses, reduce errors, increase precision and time control.

Due to the interest in the medical field, the research on automatic segmentation on breast ultrasound images has taken the main focus and has been enhanced during the last decades. Diverse algorithms have been proposed for the segmentation of breast ultrasound images: classical segmentation (edge-based segmentation, threshold-based segmentation, region-based segmentation), machine learning segmentation (unsupervised segmentation, supervised segmentation) [6].

Deep learning segmentation has proven to be very useful for in medical image processing, it is known for its ability to extract high level features by following layer by layer from a raw input, and it leads to excellent segmentation performances. In this paper I will perform automatic segmentation of lesions in breast ultrasound images. The proposed architecture to achieve this goal is based on the popular models including U-Net [16], RESNET [8], Attention Gate [14] which are more used in medical imaging.

The content of the article is structured as bellow, section 1 dedicated to the presentation of the public dataset of breast ultrasound images, the detailed structure of proposed model for segmentation, training and evaluation, section 2 is dedicated to present the results and compare them with two other models. In section 3 the conclusion is presented.

1. Materials and method

1.1. Data collection

The dataset used is the first publicly available breast ultrasound dataset [1] collected in DICOM format at the Baheya hospital, Cairo, Egypt. The data contains 780 images with an average image size of 500*500, acquired in 2018 from 600 female patients between 25 and 75 years old. the images are in PNG format, each images contains its ground truth performed carefully by radiologists. The images are classified into three categories: normal, benign and malignant. This data set is usually in gray level, the tools used in the procedure of digitization are the LOGIQ E9 ultrasound scanner and the LOGIQ E9 agile ultrasound scanner.

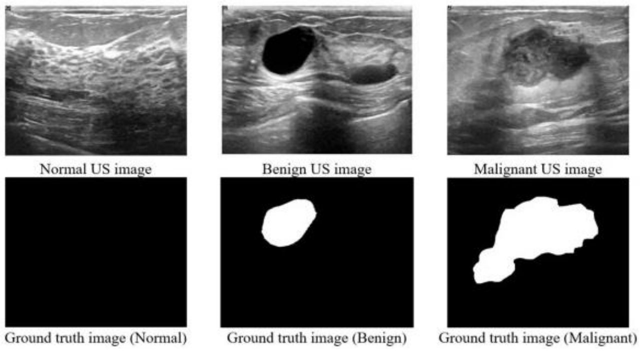


Fig. 1. Breast ultrasound images samples and their ground truth images

1.2. Method

The proposed model for lesion segmentation in breast ultrasound images inspired on mostly common and popular architectures in medical imaging U-Net [16], Attention Gate [14], Residual neural network [8].

1.2.1. Deep convolutional network

U-net

U-net is considered as a convolutional neural network (CNN) most used for semantic segmentation, designed by Olaf Ronneberger et al. in 2015 [16]. Having proven itself in terms of accuracy and speed especially in the medical imaging field. It consists of three sections: the contraction path or encoder, the bottleneck and the expansion path or decoder

The contraction path (encoder) which is used to capture the context and extract the spatial features in the input image, it follows the architecture typical of a convolution network, it is composed of four blocks, each block is composed of two convolution layers (3×3), each layer is followed by a rectified linear unit (Relu), Then a maximum pooling operation of (2×2) with a stride of 2 is applied for down-sampling and keeping the important features that better describes the context of the image. The contraction phase increases the contextual information that defines the nature of an object. During this contraction path the number of feature channels is doubled while the spatial dimension is reduced by half to decrease the number of network parameters.

The expansion path (decoder) is the second part of the U-Net architecture that is used to semantically project the discriminative features learned by the encoder onto the pixel space, it retrieves the object details and recovers the initial image size and localizes the lesion.

The expansion path starts with an up-sampling of the feature map generated by the bridge. The up-sampling operation reduces the number of feature channels by half and increases the size of the image, it also considers as concatenation part, using U-Net skip connections concatenates a feature part of the encoder with the decoder, for each block in the encoder the result of the convolution operation before maximum pooling is transferred to the decoder symmetrically.

Each decoder block receives the learned feature representation from the encoder and concatenates it with the output of the up-sampling operation, followed by two convolution layers (3×3), each of them is reinforced by an activation function (Relu).

The concatenation helps the expansion feature to recover the location information of the respective object and acquire the general information that combines the context and location. It is also useful to capture the possible features lost by the maximum pooling.

At the output of the decoder a convolution (1×1) with an activation function (Sigmoid) is applied. During the expansion phase the image size gradually increases while the number of channels gradually decreases.

The bridge or bottleneck connects the encoder to the decoder and completes the information flow, it consists of two convolutional layers (3×3), each layer is followed by an activation function (Relu). This information is important to obtain a segmentation map with high resolution.

The skip connections get the characteristics from the encoder and concatenate them with the decoder, providing the decoder with sufficient context to create an efficient segmentation mask.

The parameters of the U-Net are described as follows:

- the size of the BUS image in the input and in the output of the U-Net is 256×256×1,
- in all convolution operations the ""same"" mode is used place of ""valid"" to have the same input size at the output,
- the activation function (Relu) is applied in all convolution layers, except the final convolution layer the activation function (Sigmoid) is applied,
- the number of filters for the first layer in the encoder is 64, and the number of filters in the contraction phase is doubled after each maximum pooling operation.

Attention gate

A model that contains attention gate units in its architecture allows to focus attention on target regions and to remove insignificant regions from an input image. Attention gate units can be easily implemented in the skip connections of the U-net model while increasing model sensitivity and prediction accuracy. The results show that the attention gate units improve the performances of the U-net model while maintaining the computational efficiency [14].

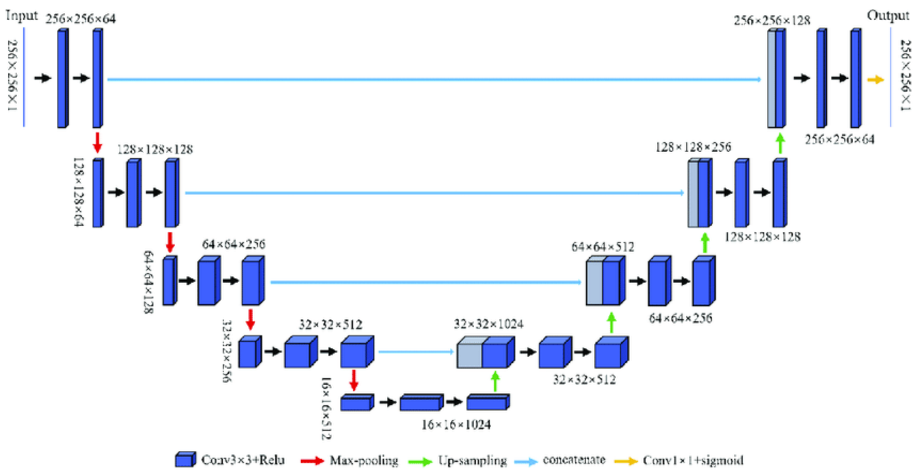


Fig. 2. U-net architecture

In this paper the attention gate mechanism integrated with the skip connections of the U-net model architecture to focus the concentration on the target breast lesion regions and remove the interference of insignificant objects on the generated segmentation map.

The typical architecture of an attention gate unit can be modelled as shown in Fig. 3.

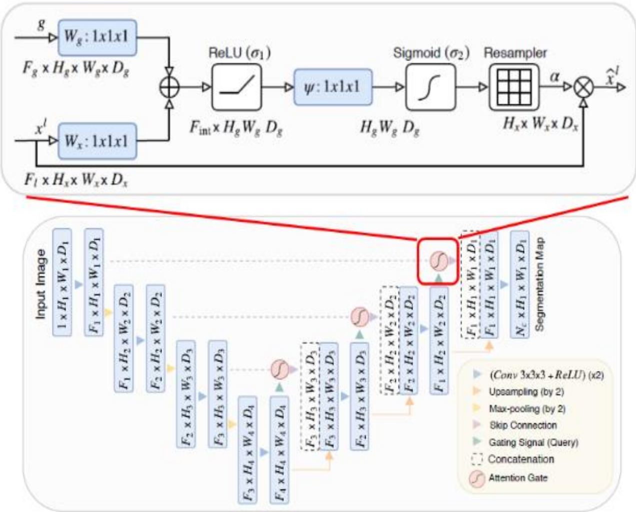


Fig. 3. Schematic of attention gate unit

Residual neural network

Residual network constitutes residual blocks proposed in 2015 by Kaiming He et al [8]. The growth of the depth of a deep neural network can hinder its training process and cause the problem of precision degradation which influence negatively of the performance of the model.

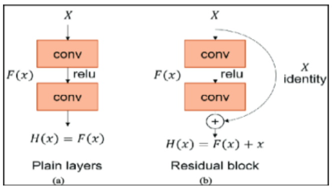


Fig. 4. Residual convolutional block

Replace the basic convolution blocks (plain layer) in the primitive U-Net architecture with residual convolution blocks (Fig. 4) by adding a skip connection between two convolution layers (3×3) helps to build a deeper network without worrying about the problem of gradient loss.

The skip connections used in the network contribute to a better flow of information between the different layers which contributes to a better gradient flow during the network formation.

1.2.2. Architecture of the proposed model

The proposed model for lesion segmentation in breast ultrasound images can be modelled with a combination of the popular existing models U-Net [16], Residual neural network [8] and t the Attention Gate Module [14].

To achieve more exciting performance of lesion segmentation in breast ultrasound images, we integrate attention gates and residual convolution blocks into the U-Net architecture to give the "Residual Attention U-Net" model that that can extract more accurate information about dense features and efficiently restore spatial information and location details.

The input size for the network is 256×256×1, where the size of each image is 256×256 and the number of channels is 1. The contracted path is formed by 4 residual blocks that replace the simple convolution blocks at the architecture of the U-Net, and each residual convolution block contains two convolution layers (3×3) reinforced by a rectified linear unit (Relu), then a skip connection adds the input of the block with the convolution output, followed by a rectified linear unit (Relu).

The size of the first layer is 256×256×64 in the contraction path. During the contraction phase the number of feature map channels is doubled while the feature map size is decreased by half. The bridge connects the decoder to the encoder.

The expansion path contains 4 residual blocks to replace the simple convolution blocks in U-Net and 4 attention gates to improve the information about salient features while disambiguating the responses of irrelevant and noisy features.

The gating signal and feature map obtained from the fourth residual block are combined via the attention gate unit, obtaining more relevant tumor location information.

Before the individual residual block in the expansion path, there is an up-sampling operation that reduces the number of feature channels by half and doubles the size of the image, then a concatenation performed between the output of up-sampling and the output of the attention gate.

At the last layer of the expansion path using a convolution (1×1) with the activation function (sigmoid). The output size for the network is 256×256×1.

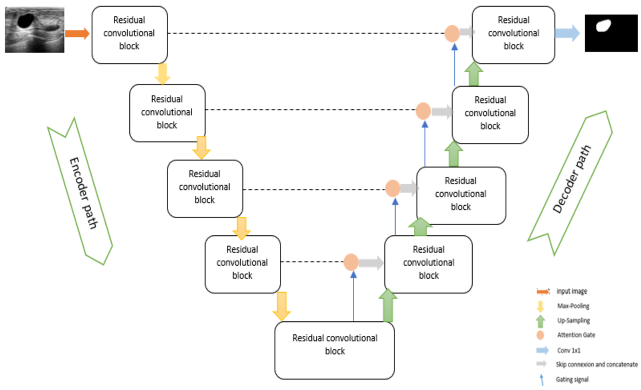


Fig. 5. Attention Residual U-net

1.2.3. Training

The model is trained on the public data set [1] which are partitioned respectively into training data represent 80% and test data which represent 20%. In the training of the network the dimension of the image and its corresponding mask were a 256×256 pixels, the Adam optimizer [10] was used to train the network, using (Accuracy, DICE coefficient [13], Jaccard index [9]) as measures of the accuracy of the segmentation procedure. While (DICE loss [12]) was used as a loss function that was repropagated through the network. The batch size was fixed at eight and the model was trained in 100 epochs and the learning rate was initially 6.1284e-8. The training and evaluation are executed in python using tensorflow [11] and Keras [3] libraries. The number of trainable parameters is 34,540,867.

1.2.4. Data augmentation

Since the number of medical images is limited, data augmentation techniques were used by applying different deformation on the training data using ImageDataGenerator.

The application of these deformation to the original image does not change the target region but only provides a new perspective. The image augmentation technique helps to increase the accuracy and robustness of the proposed model.

The deformation operations used in this work are: rotation by 50 degrees, flipping and shifting.

1.2.5. Experiment

This experiment was performed on a game computer equipped with a NIVIDA GPU GTX 1080 TI with 11GB and up to 24GB of GDDR5 memory intel core i7 7700K CPU @5.00 GHz overclocked processor, 32GB with 3200MHZ frequency of DRR4 Ram, running under Windows 10.

1.2.6. Evaluation metrics

The model performance was evaluated using the three-evaluation metrics commonly used to validate the segmentation task of medical images.

Dice coefficient

It considers as spatial overlap index between two segmentations; it presents the degree of similarity between the segmented output and the ground truth.

$$dice\ coefficient = \frac{2*|X \cap Y|}{|X| + |Y|} \tag{1}$$

We can also notice that the Dice Loss is formulated as:

$$\text{Dice Loss} = \text{1-Dice Coefficient} \tag{2}$$

X refers to the set of pixels in the lesion region manually delineated by the radiologist, while Y refers to the lesion region automatically generated by the model.

Jaccard index or intersection-over-union

IOU is used to compare the similarity between two se segmentations, presents the ratio of the intersection and union of the ground truth and the segmented output of the model.

$$Iou = \frac{|X \cap Y|}{|X| \cup |Y|} \tag{3}$$

X refers to the set of pixels in the lesion region manually delineated by the radiologist, while Y refers to the lesion region automatically generated by the model.

Accuracy

Common evaluation index for medical image segmentation, it considers as a ratio between the number of correctly predicted pixels and the total number of pixels in the image.

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \tag{4}$$

TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

2. Results and discussion

The 3 models (U-Net, Attention U-Net, Residual Attention U-Net) are trained by the same dataset and parameters training and evaluated by the same evaluation metrics.

The Attention Res U-net model achieves exiting performance in lesion segmentation in breast ultrasound images, and it allows to focus attention on significant and important salient regions and disambiguate feature responses and irrelevant regions. The segmentation map generated by the U-net model imports a lot of unnecessary background information that interferes with the target object because at the beginning of the encoder path of the u-net network the feature representation is quite poor.

Table 1. Comparison of the performance of the three models after 100 epochs

model	dice%	iou%	accuracy%
Attention-Residual U-net	90%	76%	90%
Attention U-net	70%	63%	90%
U-net	52%	37%	90%

According to the table above, the best performance was obtained by the Attention Res U-Net model.

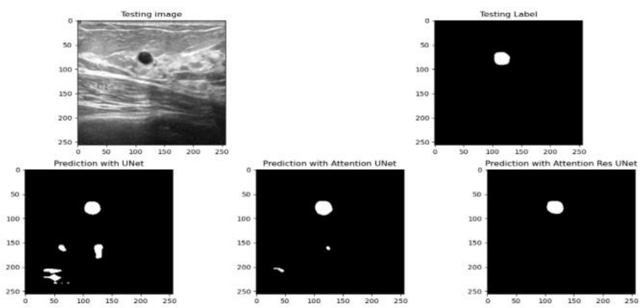


Fig. 6. The difference between ground truth mask with the prediction output of the three models

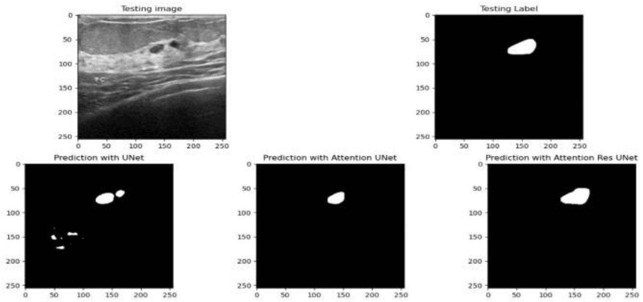


Fig. 7. The difference between ground truth mask with the prediction output of the three models

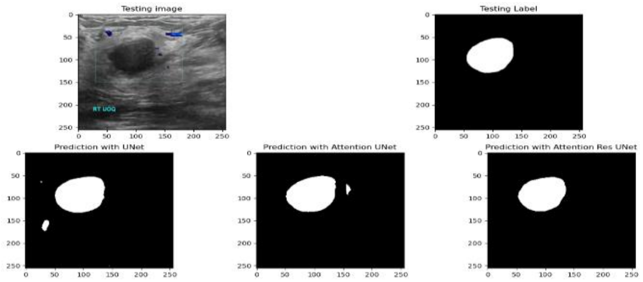


Fig. 8. The difference between ground truth mask with the prediction output of the three models

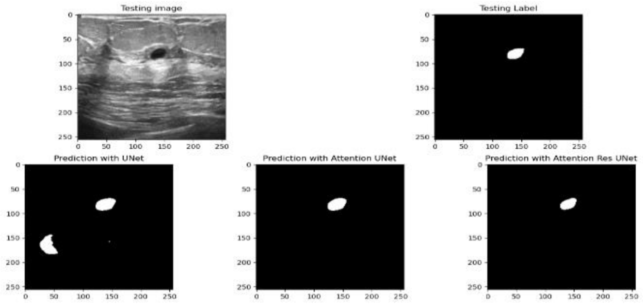


Fig. 9. The difference between ground truth mask with the prediction output of the three models

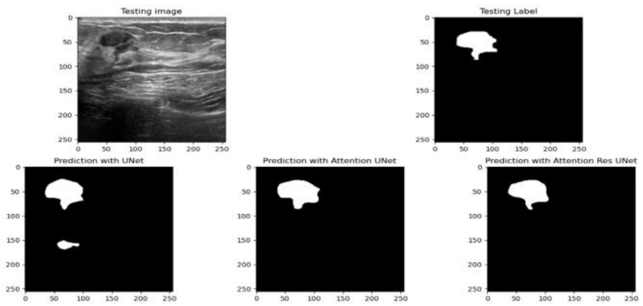


Fig. 10. The difference between ground truth mask with the prediction output of the three models

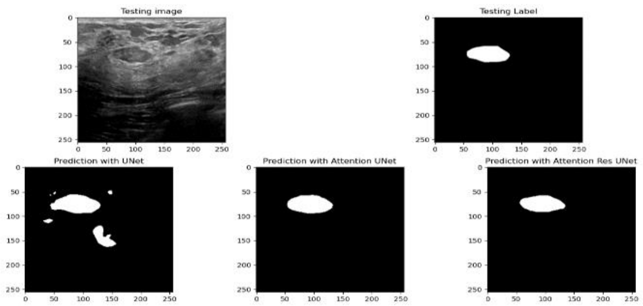


Fig. 11. The difference between ground truth mask with the prediction output of the three models

2.1. Attention residual U-net learning curve

A learning curve is simply a graph showing the progression over time of a specific learning-related metric during the training of a machine learning model. It is simply a mathematical representation of the learning process.

From the graphs of accuracy and loss, we are able to see that in each learning strategy, the model was quickly trained, as the trend in accuracy is up and the loss trend is down.

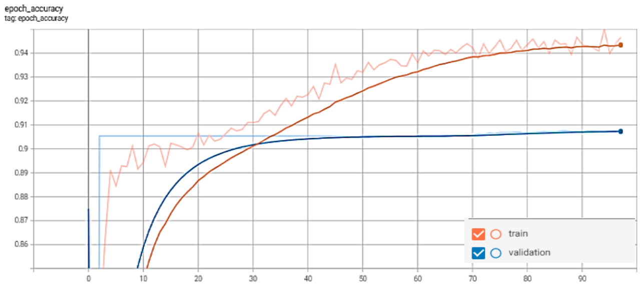


Fig. 12. Learning curves with accuracy over epochs

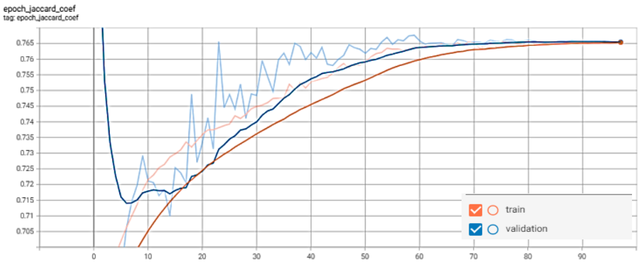


Fig. 13. Learning curves with jaccard index over epochs

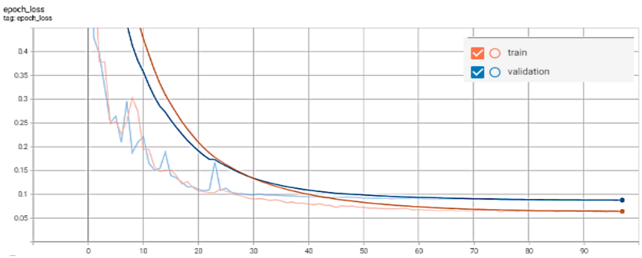


Fig. 14. Learning curves with loss over epochs

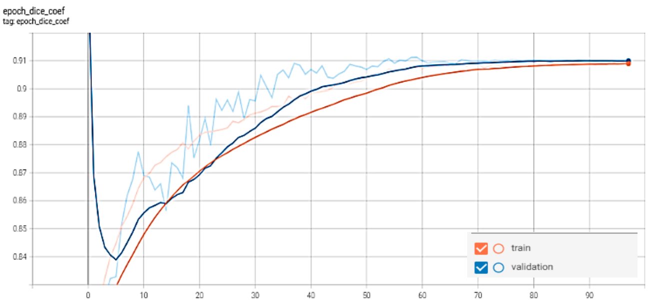


Fig. 15. Learning curves with dice over epochs

3. Conclusion

In this paper, an automated approach for lesion segmentation in breast ultrasound images is presented. The proposed method directly takes a breast ultrasound image to obtain lesion segmentation map.

The proposed model based on a combination of existing popular models U-Net, Residual neural network and Attention Gate module . To achieve more promising performance of lesion segmentation in breast ultrasound images, we integrate attention gates and residual convolution blocks into the U-Net architecture to produce the "Residual Attention U-Net" model that can extract more accurate dense feature information and well restore spatial information and location details.

The proposed model Residual Attention U-Net achieved 90% in terms of DICE and 76% in terms of Jaccard index and 90% in terms of Accuracy. Residual Attention U-Net model has been compared to two models (U-Net and Attention U-Net). Even though the proposed model shows good performance, there is certain points that can be developed in the future, also the proposed model can be applied on other types of medical images for segmentation.

References

[1] Al-Dhabyani W. et al.: Dataset of breast ultrasound images. Data in Brief 28, 2020, 104863.

[2] Balagalla B et al.: A Review On Ultrasound Image Pre-Processing, Segmentation And Compression For Enhanced Image Storage And Transmission. 11-th International Research Conference General, Sir John Kotelawala Defence University, 2018, 106–111.

[3] Chollet F.: Kears. https://github.com/fchollet/keras, 2015.

[4] Dice L. R.: Measures of the Amount of Ecologic Association Between Species. Ecology 26(3), 1945, 297–302.

[5] El Fouhi M. M. et al.: Profil épidémiologique et anatomopathologique du cancer de sein au CHU Ibn Rochd, Casablanca. Pan Afr Med J. 37, 2020, 41.

[6] Epimack M. et al.: Breast Cancer Segmentation Methods: Current Status and Future Potentials. BioMed Research International 2021, 9962109.

[7] Fondation lalla salma, détection précoce 2022.

[8] He K. et al.: Deep Residual Learning for Image Recognition. 2016, arXiv:1512.03385.

[9] Jaccard P.: The Distribution of the Flora in the Alpine Zone. The New Phytologist 11(2), 1912, 37–50.

[10] Kingma D. P., Ba J. L.: Adam: A Method For Stochastic Optimization. 3rd International Conference for Learning Representations, San Diego, 2015, arXiv:1412.6980.

[11] Mart´ın A. et al.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. 2016, arXiv:1603.04467.

[12] Milletari F. et al.: V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. 2016, arXiv:1606.04797.

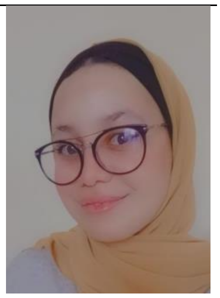
[13] Nations Unies, ONU Info: Le cancer du sein est désormais le plus fréquent au monde, 2021 [https://news.un.org/fr/story/2021/02/1088502].

[14] Oktay O. et al.: Attention U-Net: Learning Where to Look for the Pancreas, 2018, arXiv:1804.03999.

[15] Prasad S. N., Houserkova D.: A Comparison Of Mammography And Ultrasonography In The Evaluation Of Breast Masses. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 151(2), 2007, 315–322 [http://doi.org/10.5507/bp.2007.054].

[16] Ronneberger O. et al.: U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015, arXiv:1505.04597.

M.Sc. Ihssane Khallassi
e-mail: ihssanekhallassi639@gmail.com



Received the master's degree in Electrical Engineering from the National School of Arts and Crafts of Rabat, Mohammed V University, Rabat, Morocco, in 2022. She is a research student of sciences and Technology of the engineer in Ecole National Supérieure d'Informatique et d'Analyse des Systèmes ENSIAS. She is part of E2SN research team at ENSAM RABAT. She is interested in the segmentation of cancer masses on breast ultrasound images using modified u-net.

http://orcid.org/0009-0006-7965-0269

Prof. Abdelilah Jilbab
e-mail: a_jilbab@yahoo.fr

Abdelilah Jilbab graduated in electronic and industrial computer aggregation in 1995. He received the Ph.D. degree in computer and telecommunication from Mohammed V-Agdal University, Rabat, Morocco, in 2009. He is a Professor at ENSAM, Mohammed V University, Rabat, Morocco. Since 2003, he is a member of the laboratory LRIT (unit associated with the CNRST, FSR, Mohammed V University, Rabat, Morocco). His domains of interest include signal processing and embedded systems.

http://orcid.org/0000-0002-1577-9040

Prof. My Hachem El Yousfi Alaoui
e-mail: h.elyousfi@um5r.ac.ma

He is a Professor at ENSAM, Mohammed V University, Rabat, Morocco. He is a member of Electromechanical, Diagnosis and Control Research Team – Electrical Engineering Laboratory – ENSAM, Mohamed V University in Rabat, Morocco. He is a member of E2SN research team at ENSAM RABAT.

http://orcid.org/0000-0003-4285-0540