

An Automatic Segmentation of Breast Ultrasound Images Using U-Net Model

Eman Radhi¹, Mohammed Kamil¹

Abstract: Medical imaging, like ultrasound, gives a good visual picture of how an organ works. However, a radiologist has a hard time and takes a long time to process these images, which delays the diagnosis. Several automated methods for detecting and segmenting breast lesions have been developed. Nevertheless, due to ultrasonic artifacts and the intricacy of lesion forms and locations, the segmentation of lesions or tumors from breast ultrasonography remains an open issue. Medical image segmentation has seen a breakthrough thanks to deep learning. U-Net is the most noteworthy deep network in this regard. The traditional U-Net design lacks precision when dealing with complex data sets, despite its exceptional performance in segmenting multimedia medical images. To reduce texture detail redundancy and avoid overfitting, we suggest developing the U-Net architecture by including dropout layers after each max pooling layer. Batch-normalization layers and a binary cross-entropy loss function were used to preserve breast tumor texture features and edge attributes while decreasing computational costs. We used the breast ultrasound dataset of 780 images with normal, benign, or malignant tumors. Our model showed superior segmentation results for breast ultrasound pictures compared to previous deep neural networks. Quantitative measures, accuracy, and IoU values were utilized to evaluate the suggested model's effectiveness. The results were 99.34% and 99.60% for accuracy and IoU. The results imply that the augmented U-Net model that has been suggested has high diagnostic potential in the clinic since it can correctly segment breast lesions.

Keywords: Breast cancer, Ultrasound, Deep learning, Segmentation, U-net.

1 Introduction

Among all cancers, breast cancer affects more women than any other. and over 8% of women will develop this illness at some point in their lives. Early detection is crucial to lowering the death rate (by at least 40%), as the causes of breast cancer are yet unknown [1], and only aware of a few risk factors that can raise the chance of getting breast cancer: aging, heredity, exposure to radiation, dense breast tissue, drinking alcohol, etc. Finding breast cancer's early signs and

¹College of Science, Mustansiriyah University, Baghdad, Iraq;
E-mails: eaer123ali@gmail.com; m80y98@uomustansiriyah.edu.iq

symptoms through clinic screening is crucial for lowering mortality [2]. For the diagnosis of breast cancer, mammography and ultrasound are frequently used. Although mammography is frequently used in clinical practice to evaluate breast masses and improve visualization of tiny and nonpalpable tumors [3]. Breast ultrasound (BUS) is preferred since it doesn't use ionizing radiation, provides real-time visualization, and is inexpensive [4]. Additionally, mammography often has a limited predictive value. As a result, people who have benign tumors undergo unnecessary biopsies. Using ultrasonography as a supplement to mammography has overcome this restriction of mammography [5]. Despite the radiologist's ability, manual analysis of these scans takes time and effort. Computer-aided diagnostic (CAD) systems are being developed to identify breast tumors earlier, allowing for faster diagnosis and treatment. Segmenting tumors in most CAD systems is crucial for diagnostics and follow-up treatment plans [6]. Deep learning (DL) is a class of algorithms and a recent branch of machine learning (ML) appropriate for big data analysis, image segmentation, and other techniques. DL has advanced quickly in medical image analysis in recent years due to the pressing need for illness prevention and treatment. The inherent ability of DL methods to learn complex features directly from the original data is a key benefit when processing large amounts of data efficiently and extracting the most pertinent image features [7]. DL describes various computational models made up of multiple processing layers. These layers primarily study abstract data representations at various levels. It can gradually replace the feature defined manually using the machine learning method and has a strong feature extraction capability. As a result, DL has evolved into a crucial tool for medical picture analysis [8]. Among the most representative DL algorithms is the convolutional neural network (CNN), which has clear benefits in image processing, particularly for issues like target detection and image segmentation. CNN is challenging to train from the start, particularly for medical applications where the production of labeled data requires substantial time and money. The knowledge collected by a network on a separate, often large dataset is transferred to another application instead of complete training. May be performed by retraining the whole network or adjusting a few layers [9]. Numerous studies have shown the usefulness of employing non-medical images (such as images of the natural world) as the source dataset for transfer learning to the field of medical imaging. This benefits the model by making many training pictures available, which at the very least provides a suitable parameter initialization for further training in the new domain. When the target dataset in the new domain is small, the proposed strategy for transfer learning is fine-tuning, which entails keeping the network's shallow layers while updating the deep layers in line with the new dataset [10]. Among CNN's used for semantic object segmentation, U-Net is perhaps the most well-known. The U-net network can solve almost any biomedical segmentation issue [8]. In medical image segmentation, U-Net has demonstrated successful

performance and efficient segmentation. Due to the variety and complexity of BUS images, which are frequently hypoechoic, with hazy borders, low signal-to-noise ratios, and asymmetric intensity, U-Net segmentation has been continually refined to enhance segmentation for these image types [11]. For example, M. Amiri et al. [12] suggested employing a lesion detection stage before the segmentation stage. A collection of 163 breast images from BUS imaging was used in the methods. They used a U-Net to find the lesions and then segmented the region that was found using another U-Net. The Dice score result was 97%. M. Byra et al. suggested in [13] a CNN with a selective kernel (SK). The SKs were designed to combine feature maps obtained with dilated and standard convolutions and alter the network's receptive fields via an attention mechanism. The suggested approach was created and assessed using US images of 882 breast masses. The mean Dice score for the SK-U-Net was 0.826, and the accuracy was 0.979. BUS images were segmented by Y. Tong et al. using an "Improved U-net based on Mixed Attention Loss Function" (Improved U-net MALF) [14]. First, using the U-net network structure for attention, the convolution module for the coding path was replaced with the residual convolution module and extended residual convolution module. The traditional cross-entropy loss function was combined with four other attention loss functions as a second phase. Additionally, the texture consistency index was employed to determine the percentage of loss values for the four attention functions in the total loss value to emphasize the tumor target. 316 images were gathered to validate the process. Based on the upgraded U-net MAL network, sensitivity and specificity were 85% and 97.9% respectively. In [11], Y. Guo et al. proposed an extended training technique to get enlarged of U-Net by expanding the loss function (binary cross-entropy function) to a ternary cross-entropy function. 192 clinical BUS images were employed. The average IOU coefficient is $82.7 \pm 0.02\%$, and the average Dice coefficient is $90.5 \pm 0.02\%$. Furthermore, Y. Yan et al. segmented of breast cancers in ultrasound images using the attention enhanced U-net with hybrid dilated convolution (AE U-net with HDC) model. The experiment's results showed that accuracy was 95.61%, and recall (Recall) was 80.48% [15]. To improve the segmentation performance of various tumor sizes, N. S. Punn and S. Agarwal introduced "residual cross-spatial attention-guided inception U-Net" (RCA-IUNet) for cancer segmentation using BUS image with minimal training settings. The U-Net topology with residual inception depth-wise separable convolution is used in the RCA-IUNet model, which also employs hybrid pooling layers. The models are trained and tested using datasets from BUSIS and BUSI. On the dice, the result was 98.4% accuracy was 99% [16]. The motivation behind this work is to improve breast tumor segmentation accuracy by utilizing U-Net advantages in medical image segmentation, where the training mode is tuned to better fit the properties of BUS images. Following is the organization of the rest of the paper: Section 2 provides a general description of the U-Net; Section 3 presents the

dataset; Section 4 describes the experiments; Section 5 contains the results and discussion and the conclusion is given in Section 6.

2 U-Net Architectures

Ronneberger proposed the convolutional neural network known as U-Net [17]. This architecture uses a series of symmetric upsampling convolutional layers followed by hierarchical downsampling convolutional layers. The decoding part also passes semantic information through the concatenation of feature maps from the encoder network [18]. Two pathways make up the U-net network: The contracting path is the first and uses a typical CNN design. A ReLU activation unit, a max-pooling layer, and two successive 3×3 convolutions make up each block in the contracting route. Several iterations of this arrangement are made [19]. The enlarged route is the name given to the second segment of the U-net. The feature map is upsampled in this section's steps using 2×2 up-convolution. Then, a feature map is clipped and concatenated onto the upsampled feature map from the matching layer in the contracting route. Two subsequent 3×3 convolutions and ReLU activation come next [20]. The feature map is reduced to the appropriate number of channels, and a further 1×1 convolution is performed to get the segmented image. The margins of the pixels must be cropped because they contain the least amount of contextual information and must be ignored [17]. The last convolutional layer employs a Sigmoid activation function. Adding skip connections to the U-architecture Net is undoubtedly its most innovative element. Before the encoder's pooling operation, all layers provide the convolutional layer's output to the decoder. The upsampling operation results are appended to these feature maps, and the resulting map is propagated to the next layers. With the help of these skip connections, the network can retrieve the spatial information lost during pooling processes [20]. The overall U-net architecture is shown in Fig. 1.

3 Datasets

Curating datasets for medical imaging are more challenging than doing so for datasets for traditional computer vision. The creation of medical imaging datasets is made difficult by pricey imaging equipment, complex image acquisition processes, professional annotation requirements, and privacy issues. As a result, there are very few public benchmark datasets for medical imaging, and even those datasets only have a few photos [20]. We used the BUS dataset provided by [21]. The collected data contains BUS images of women aged between 25 and 75. This information was compiled in 2018. There are a total of 600 female patients. The collection contains 780 images with an average resolution of 500 by 500 pixels. Images are saved in PNG format. The images are grouped into three categories: normal 133, benign 487, and malignant 210; each

An Automatic Segmentation of Breast Ultrasound Images Using U-Net Model

image has its mask image. The scanning technique uses the LOGIQ E9 and LOGIQ E9 Agile ultrasound systems. These tools are frequently used in advanced imaging in radiology, cardiology, and vascular surgery. They produce images at a size of 1280×1024 . Fig. 2 shows a sample of the BUS image and mask.

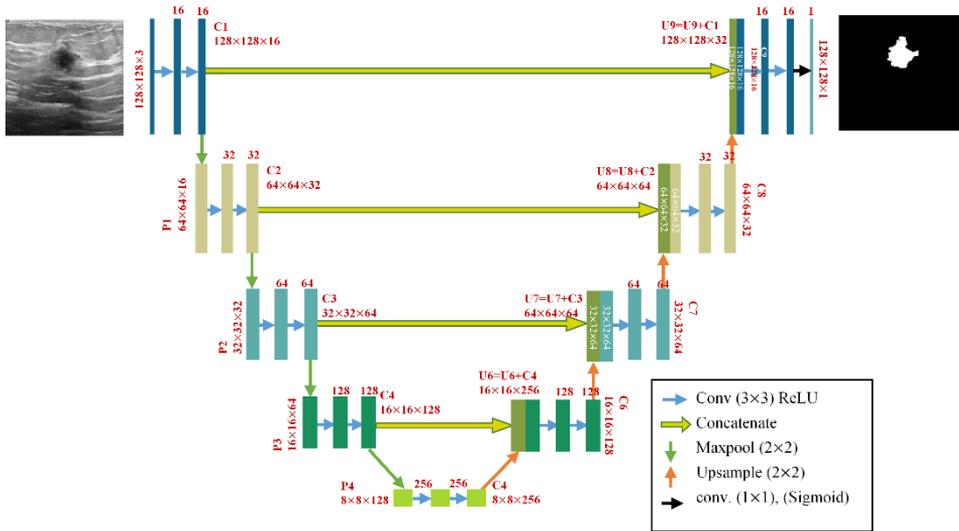


Fig. 1 – U-net network structure.

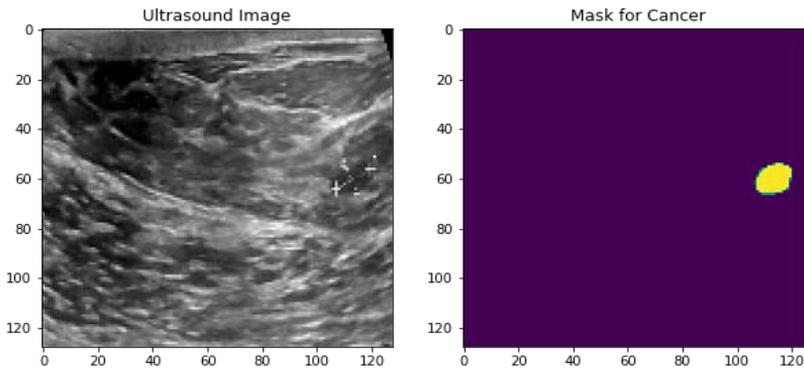


Fig. 2 – Sample of BUS image.

4 Experiments

The studies were carried out at Google Colab using the Python programming language. The network models were implemented with Keras and TensorFlow.

4.1 Pre-processing and training procedure

This experiment uses a dataset (780) image divided into a training set (702) image, validation, or testing (78) image. Initially, we reorganized the dataset using a code that created two files, one including all the original images (normal, benign, and malignant) and the other containing the masks corresponding to each image. The masks corresponding to the benign and malignant tumor pictures were binary, but the masks corresponding to the normal images were of the RGB before their conversion to binary images. The input images and their masks were resized to 128×128 to reduce the computational cost. Batch normalization was applied after each block to reduce mean and variance issues and stabilize the U-Net network's layers. After each max pooling layer, dropout layers were added to reduce texture redundancy and stop overfitting. Semantic segmentation's task is to determine whether a pixel is a point of interest or just background by making predictions about its characteristics. A pixel-wise binary classification problem is what this issue finally comes down to. So, we used the binary cross-entropy function for the network loss function and reduced it. The model was trained using the Adam optimizer after we reduced the binary cross-entropy loss. Adam adaptively calculates various learning rates for various parameters based on estimates for the first and second moments of the gradients. Adam has been frequently utilized in assessing deep learning models as the default option because it incorporates the benefits of both AdaGrad [22] and RMSProp [23]. Additionally, the early stopping technique, besides the dropout layer, is used to prevent the overfitting issue, which terminates the training process immediately after the testing error stops improving. For input and output shapes to be the same, kernels were also padded by (padding = "same"). The improved U-net architecture is illustrated in Fig. 3.

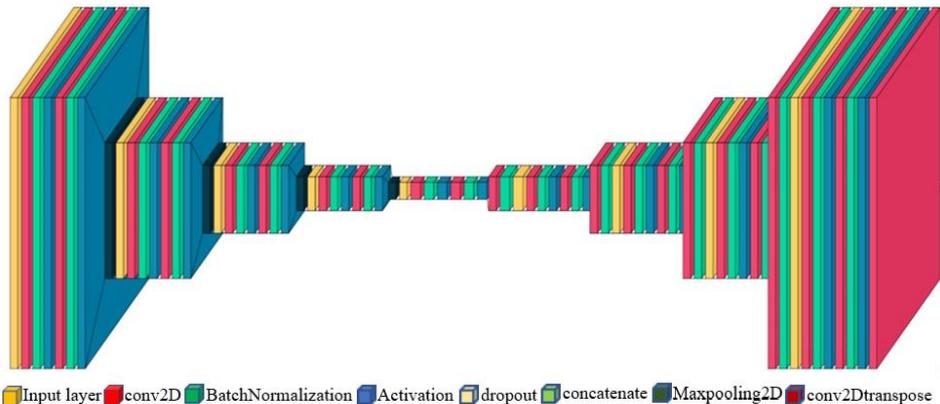


Fig. 3 – Improved U-net network structure.

4.2 Evaluation metric

A good evaluation of an image processing model is equally as important as designing one. Some of the metrics for evaluating image segmentation are presented in this section. The resulting confusion matrix and its corresponding true positive (T_p), true negative (T_N), false positive (F_p), and false negative (F_N) values were the basis for many of these measurements [24]. The accuracy metric compares the proportion of properly predicted samples to the overall sample count. These samples are usually pixels or voxels in image processing. The "accuracy," or the ratio of correctly categorized items to all other elements, may be calculated as [25]:

$$Accuracy = \frac{T_p + T_N}{T_p + T_N + F_p + F_N}. \quad (1)$$

When image processing involves unbalanced data distributions, accuracy is less helpful and is seldom considered alone [19]. When semantic segmentation is done, points of interest usually cover a small percentage of the entire image. As a result, the accuracy metric is insufficient and frequently generates a misleading appearance of superiority due to the perfection of background detection. The intersection over union (IoU) measure, commonly known as the Jaccard index, has been extensively utilized to evaluate and compare image segmentation and object localization techniques [20]. IoU is a method for assessing detection efficiency. It compares the training model's segmentation areas to the test images. The overlap between the two surfaces is divided by the surface's overall area [26]. The IoU is expressed as follows:

$$IoU = \frac{T_p}{T_p + F_p + F_N}. \quad (2)$$

A prediction result is considered correctly positive if the IoU is more than 0.5 [27]. Using the IoU as the metric, we can penalize both under- and over-segmentation and emphasize precise segmentation.

5 Results and Discussion

Quantitative and qualitative assessments have been performed. In quantitative analysis, the numerical evaluation of performance is examined. The visual quality of the results is examined during qualitative analysis. The segmentation performance scores for the proposed U-Net on the BUS images test set are shown in **Table 1**.

The dataset's training was terminated after 100 epochs since convergence was evident, as shown in Fig. 4. Training on the dataset was terminated at a loss rate of 2.65%, and the validation loss rate reached 2.04%. When the procedure was finished, the IoU scores for model training and validation were 91.55% and

99.60%, respectively. On the training set, the model’s accuracy was 98.78%, and on the validation set, it was 99.341%. During training, the loss rate declines significantly during the first 20 epochs, then gradually lowers until it reaches 3.66% after 60 epochs. The loss on the test set evolves similarly, except that it converges at 2.39%. Since the validation curve’s loss is just 0.61% less than the training curve’s loss after training, the loss curves are almost identical. The IoU value grows continuously from epoch 0 to 5 and then converges to 91.55% throughout training. During testing, the IoU value grows significantly after fewer than five epochs, then continues to climb marginally until leveling at approximately 99.60% after epoch 9. U-Net has already achieved 99% accuracy rate after only 40 epochs of training and testing.

Table 1
*Quantitative outcomes for U-Net on BUS images
 for the training and validation sets.*

	Images No.	Loss %	IoU %	Accuracy %
training	702	2.65	91.55	98.78
validation	78	2.04	99.60	99.34

Grouping the same areas of an image is the fundamental idea behind semantic segmentation. However, symmetrical regions often become asymmetrical in genuine medical imaging because of various types of noise, artifacts, and other abnormalities. Therefore, it is difficult to differentiate the region of interest from the background in medical images. Consequently, we often end up with a collection of fractured segmented areas rather than a continuous segmented region. On the other hand, because of textures and turbulence, a simple backdrop might sometimes resemble a region of interest. Both situations result in information loss and misclassifications. To make encoder feature maps more resistant to disturbances, we suggest that they should be given additional nonlinear operations. We apply batch normalization after each convolution block to maintain orderly gradient levels, speed up convergence, and reduce the effect of internal shift variables so that network parameters do not change rapidly during backpropagation. Dropout layers are used after each max pooling layer to reduce overfitting. During training, the important concept is randomly dropping units (along with their connections) from the neural network. Units are prevented from co-adapting too much as a result. Using the binary cross-entropy loss function helped the retention of breast tumor edge detail and characteristics, which aided in higher segmentation accuracy. Our improved U-net substantially increased segmentation accuracy without raising the computational cost by reusing shallow characteristics to extract more detailed features from BUS tumors.

An Automatic Segmentation of Breast Ultrasound Images Using U-Net Model

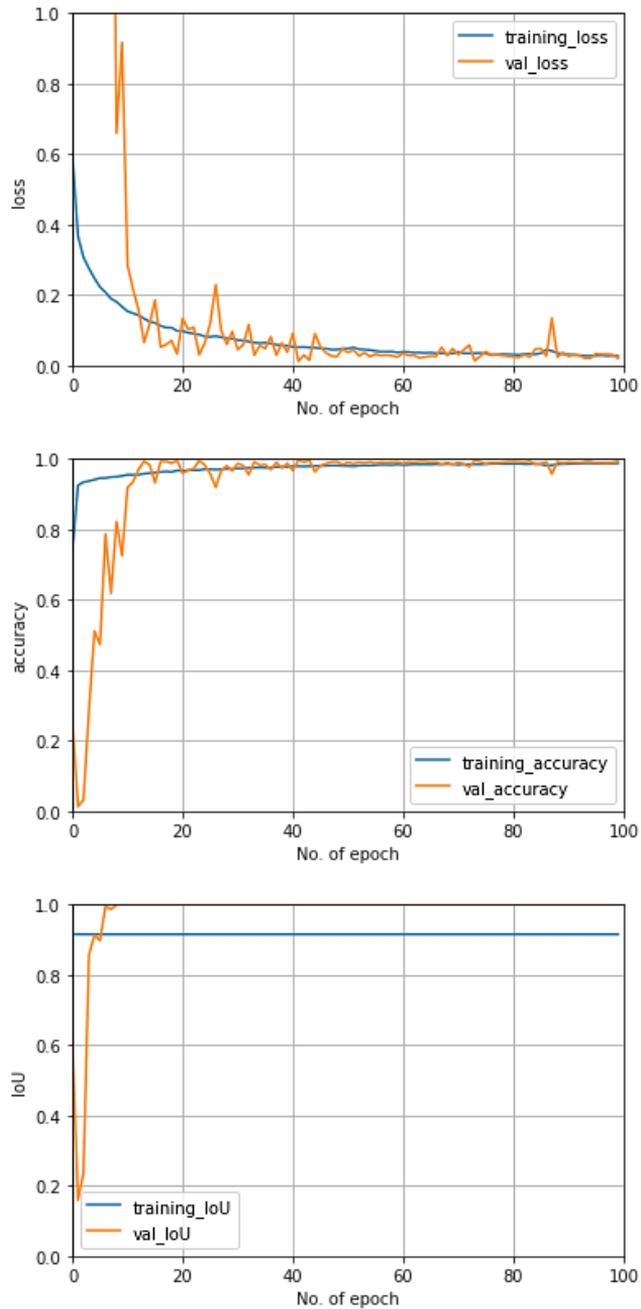


Fig. 4 – Training and validation performance progress with the number of epochs.

Additionally, it reduced the gradient explosion or gradient disappearance brought on by the deepening of network layers. Fig. 5 shows a random selection of BUS images from the test set to simplify understanding of the results by comparing mask prediction from our model to the actual mask. As is evident from the figure, our technique offers distinct advantages over previous models since it is more sensitive to regions of varying sizes and attentive to nuances.

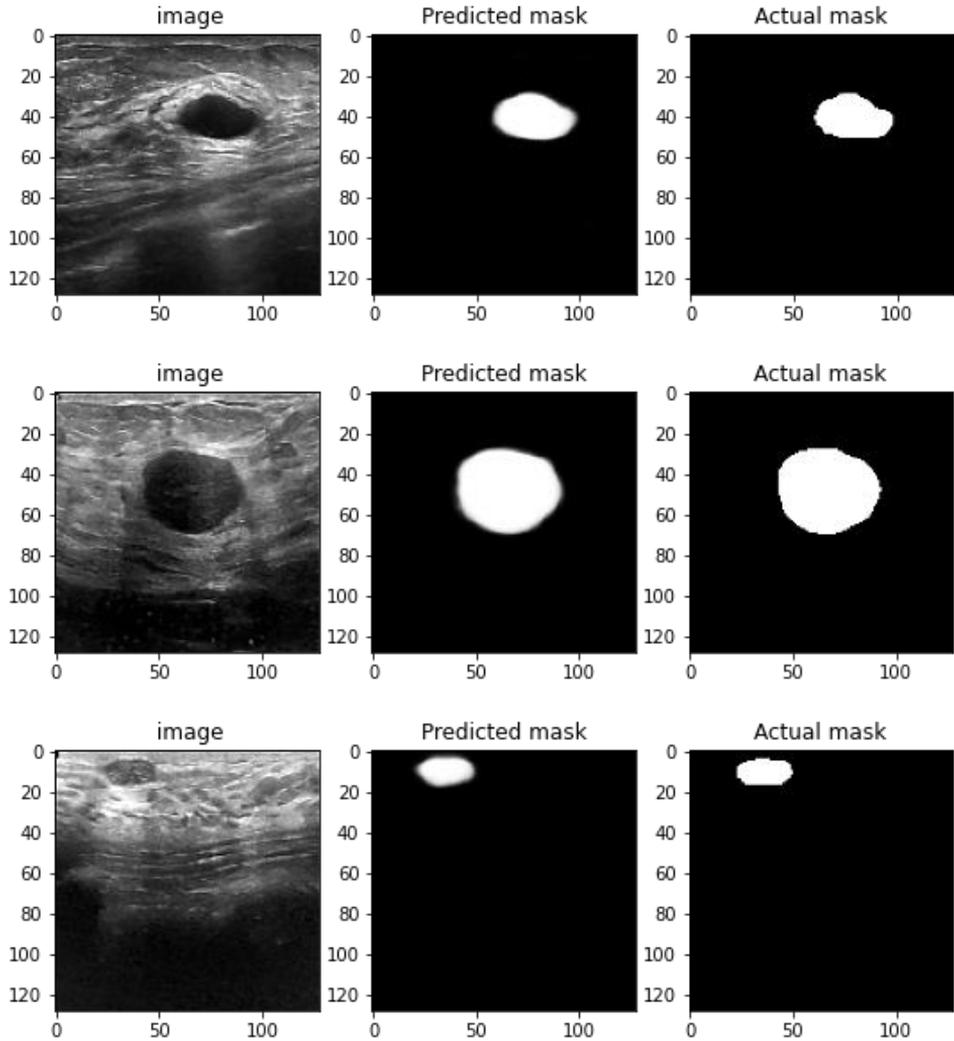


Fig. 5a – Samples of results in BUS images.

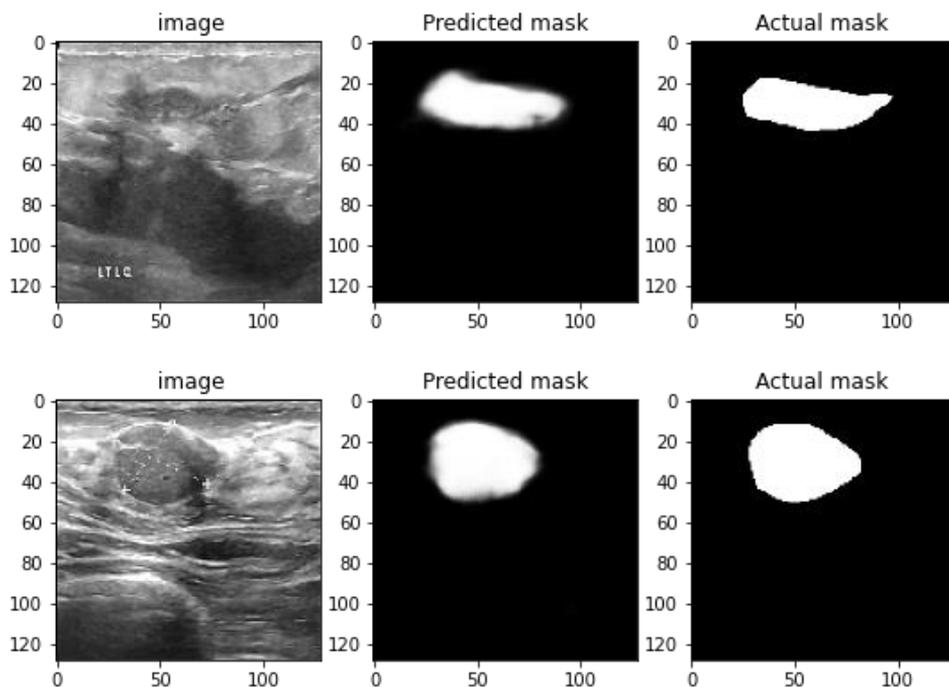


Fig. 5b – *Samples of results in BUS images.*

Finally, **Table 2** compares different studies conducted on BUS images. In general, our strategy performed better than other methods [11, 13, 14, 16], indicating the reliability and robustness of the proposed model.

Table 2
Comparison of our proposed model to current models.

Reference	Images	Accuracy %	IoU%
M. Byra [13]	882	97.90	-
Y. Tong [14]	316	95.90	-
Y. Guo [11]	192	-	82.70
N.S. Punn [16]	562	99.00	-
Proposed model	780	99.34	99.60

6 Conclusion

This work presents a machine learning-based automated segmentation technique for BUS images. The segmentation findings are in close conformity with the visibly segmented masks. This means that the proposed technique has

the same ability as medical experts to identify different tissues from BUS images. To resolve the issues of uneven brightness, low contrast, and poor quality of BUS images, we combined batch normalization, dropout layer, and binary cross-entropy loss to propose an improved U-net. More precise characteristics of ultrasonography breast tumors were retrieved without increasing the calculation's expense and resolving overfitting issues. The method was tested on a dataset of 780 images, and the findings showed that segmentation performance had improved. The comparative analysis demonstrates that the suggested strategy outperforms the existing research methods. Our research findings and methodologies might be an essential first step in developing deep learning algorithms for breast mass detection. In the future, we want to examine the applicability of other deep-learning techniques for mass breast segmentation.

7 References

- [1] M. Muhammad, D. Zeebaree, A. M. A. Brifcani, J. Saeed, D. A. Zebari: Region of Interest Segmentation Based on Clustering Techniques for Breast Cancer Ultrasound Images: A Review, *Journal of Applied Science and Technology Trends*, Vol. 1, No. 3, June 2020, pp. 78–91.
- [2] M. Xian, Y. Zhang, H. D. Cheng, F. Xu, B. Zhang, J. Ding: Automatic Breast Ultrasound Image Segmentation: A Survey, *Pattern Recognition*, Vol. 79, July 2018, pp. 340–355.
- [3] E. A. Radhi, M. Y. Kamil: Breast Tumor Detection Via Active Contour Technique, *International Journal of Intelligent Engineering and Systems*, Vol. 14, No. 4, 2021, pp. 561–570.
- [4] Y. Xu, Y. Wang, J. Yuan, Q. Cheng, X. Wang, P. L. Carson: Medical Breast Ultrasound Image Segmentation by Machine Learning, *Ultrasonics*, Vol. 91, January 2019, pp. 1–9.
- [5] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, D. A. Zebari: Machine Learning and Region Growing for Breast Cancer Segmentation, *Proceedings of the IEEE International Conference on Advanced Science and Engineering (ICOASE)*, Zakho – Duhok, Iraq, April 2019, pp. 88–93.
- [6] N. S. Punna, S. Agarwal: RCA-IUnet: A Residual Cross-Spatial Attention-Guided Inception U-Net Model for Tumor Segmentation in Breast Ultrasound Imaging, *Machine Vision and Applications*, Vol. 33, No. 2, March 2022, p. 27.
- [7] R. R. Kadhim, M. Y. Kamil: Comparison of Breast Cancer Classification Models on Wisconsin Dataset, *International Journal of Reconfigurable and Embedded Systems*, Vol. 11, No. 2, July 2022, pp. 166–174.
- [8] G. Du, X. Cao, J. Liang, X. Chen, Y. Zhan: Medical Image Segmentation based on U-Net: A Review, *Journal of Imaging Science Technology*, Vol. 64, No. 2, March – April 2020, p. 020508.
- [9] F. Ponzio, G. Urgese, E. Ficarra, S. Di Cataldo: Dealing with Lack of Training Data for Convolutional Neural Networks: The Case of Digital Pathology, *Electronics*, Vol. 8, No. 3, February 2019, p. 256.
- [10] M. Amiri, R. Brooks, H. Rivaz: Fine-Tuning U-Net for Ultrasound Image Segmentation: Different Layers, Different Outcomes, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 67, No. 12, December 2020, pp. 2510–2518.
- [11] Y. Guo, X. Duan, C. Wang, H. Guo: Segmentation and Recognition of Breast Ultrasound Images based on an Expanded U-Net, *PLoS ONE*, Vol. 16, No. 6, June 2021, p. e0253202.

An Automatic Segmentation of Breast Ultrasound Images Using U-Net Model

- [12] M. Amiri, R. Brooks, B. Behboodi, H. Rivaz: Two-Stage Ultrasound Image Segmentation Using U-Net and Test Time Augmentation, *International Journal of Computer Assisted Radiology and Surgery*, Vol. 15, No. 6, June 2020, pp. 981 – 988.
- [13] M. Byra, P. Jarosik, A. Szubert, M. Galperin, H. Ojeda-Fournier, L. Olson, M. O’Boyle, C. Comstock, M. Andre: Breast Mass Segmentation in Ultrasound with Selective Kernel U-Net Convolutional Neural Network, *Biomedical Signal Processing and Control*, Vol. 61, August 2020, p. 102027.
- [14] Y. Tong, Y. Liu, M. Zhao, L. Meng, J. Zhang: Improved U-Net MALF Model for Lesion Segmentation in Breast Ultrasound Images, *Biomedical Signal Processing and Control*, Vol. 68, July 2021, p. 102721.
- [15] Y. Yan, Y. Liu, Y. Wu, H. Zhang, Y. Zhang, L. Meng: Accurate Segmentation of Breast Tumors Using AE U-Net with HDC Model in Ultrasound Images, *Biomedical Signal Processing and Control*, Vol. 72, Part A, February 2022, p. 103299.
- [16] N. S. Pun, S. Agarwal: RCA-IUnet: a residual cross-spatial attention-guided inception U-Net model for tumor segmentation in breast ultrasound imaging," *Machine Vision Applications*, vol. 33, no. 2, pp. 1 – 10, 2022.
- [17] O. Ronneberger, P. Fischer, T. Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation, *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, October 2015, pp. 234 – 241.
- [18] V. Ummadi: U-Net and its Variants for Medical Image Segmentation: A Short Review, *arXiv:2011.01118 [eess.IV]*, November 2020, pp. 1 – 42.
- [19] N. Siddique, S. Paheding, C. P. Elkin, V. Devabhaktuni: U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications, *IEEE Access*, Vol. 9, June 2021, pp. 82031 – 82057.
- [20] N. Ibtihaz, M. S. Rahman: MultiResUNet: Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation, *Neural Networks*, Vol. 121, January 2020, pp. 74 – 87.
- [21] W. Al-Dhabyani, M. Goma, H. Khaled, A. Fahmy: Dataset of Breast Ultrasound Images, *Data in Brief*, Vol. 28, February 2020, p. 104863.
- [22] J. Duchi, E. Hazan, Y. Singer: Adaptive Subgradient Methods for Online Learning and Stochastic Optimization, *Journal of Machine Learning Research*, Vol. 12, January 2011, pp. 2121 – 2159.
- [23] T. Tieleman, G. Hinton: Lecture 6.5-rmsprop: Divide the Gradient by a Running Average of Its Recent Magnitude, *COURSERA: Neural Networks for Machine Learning*, Vol. 4, No. 2, October 2012, pp. 26 – 31.
- [24] M. Y. Kamil: A Deep Learning Framework to Detect Covid-19 Disease Via Chest X-ray and CT Scan Images, *International Journal of Electrical and Computer Engineering*, Vol. 11, No. 1, February 2021, pp. 844 – 850.
- [25] E. Fernandez-Moral, R. Martins, D. Wolf, P. Rives: A New Metric for Evaluating Semantic Segmentation: Leveraging Global and Contour Accuracy, *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, Changshu, China, June 2018, pp. 1051 – 1056.
- [26] D. Müller, A. Ehlen, B. Valeske: Convolutional Neural Networks for Semantic Segmentation as a Tool for Multiclass Face Analysis in Thermal Infrared, *Journal of Nondestructive Evaluation*, Vol. 40, No. 1, March 2021, p. 9.
- [27] E. A. Radhi, M. Y. Kamil: Breast Tumor Segmentation in Mammography Image Via Chan-Vese Technique, *Indonesian Journal of Electrical Engineering and Computer Science*, Vol. 22, No. 2, May 2021, pp. 809 – 817.