



International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 20 No. 3 (2024)



ijerst.editor@gmail.com
editor@ijerst.com

Research Paper**A HYBRID NETWORK FOR SEGMENTATION OF OVARY AND FOLLICLE IN ULTRASOUND IMAGES****Vodapelli Pranathi***Department of cyber security, Malla Reddy Engineering College (A), Hyderabad-500100*pranathipranu37@gmail.com

Abstract: Transvaginal ultrasonography (TVUS) is a common method used to treat infertility. Sonographers can use manual measures to evaluate if the ovary and follicles are viable. In contrast, this procedure is particularly time-timewasting and operator-supported. This research offers a unique combination system, CR-Unet, for segmenting the ovary and follicles simultaneously in TVUS. The CR-Unet is a basic U-Net that incorporates a spatial recurrent neural network (RNN). It can efficiently realize multi-scale and long-term spatial environments to overcome the assignment restrictions, which include poor picture condition, less difference, periphery ambiguity, and complicated anatomical structures. In addition, we apply a deep supervision approach to improve the efficacy and efficiency of model training. Self-supervision is also utilized to improve segmentation results on a regular basis. Experiments using 3204 TVUS pictures from 219 patient role showed that the recommended strategy surpassed earlier literature techniques for segmenting the ovary and follicles, with Dice Similarity Coefficients (DSC) of 0.931 and 0.885, correspondingly.

Received: 02-06-2024

Accepted: 10-07-2024

Published: 18-07-2024

I. INTRODUCTION

TVUS is a noninvasive imaging approach for quantitative ovarian assessment that is effective, real-time, simple to use, safe and easily available. It is now an essential approach for tracking ovarian follicular development and calculating the mass and quantity of follicles in in-vitro fertilization (IVF) during the ovulation, induction, or controlled ovarian hyperstimulation (COH) phases [1]. It has become one of the most often used imaging modalities in routine obstetric and gynaecological examinations to assess the causes of infertility and plan future treatments and therapies. Sonographers currently measure the size and form of the ovary and follicles every few days to determine their physiological status.

This approach, on the other hand, is tremendously time-timewasting and operator-supported. As a result, automated ovary and follicle segmentation techniques are in great challenge in modern experimental manner for

ovarian quantification evaluation. As seen in Figure 1, automated segmentation of the ovary and follicles in TVUS remains difficult for the following reasons [2]. US pictures are of low quality because to the prevalence of speckle noise and acoustic shadowing. 2) TVUS generally lacks anatomical boundaries due to imaging defects and the ultrasound probe's acoustic wave path being comparable to the anatomical construction. 3) The ovaries and follicles are random and change completely in shape. 4) The quantity and quantity of follicles in a single copy varies greatly, ranging from 1.5 to more than 32 mm for normofollicular images and 3 to 25 for multi follicular photos. 5) Follicles, especially those with a diameter of less than 5 mm, are mistaken with other matters or shapes (such as tiny veins, vesicles, or hydrops). The fundamental problem of computerized follicle breakdown in ovarian ultrasound pictures is accurately detecting true follicles during significant spot noise. Several division

approaches have been exploited to recover the shapes of things from ultrasound checkup pictures [3].

Seatiawati et al. [4] clustered ultrasound images with a improved non-parametric condition function and Particle Swarm Optimization. Using a contrast enhancement may lead to a more accurate evaluation of follicles. Machine Learning algorithms are used to categorize all collected follicular characteristics. Adiwijaya et al. [5] used the area evolution method and Euclidian distance to find and size follicles with 78% accuracy. There is no classification in this situation; just follicle detection and measurement are conducted.

Wisesty et al. present an ultrasound-based variant of the backpropagation technique for identifying polycystic ovarian syndrome [6-7] categorize ultrasound pictures using three machine learning procedures: KNN with Euclidian distance, NN-LVQ and SVM. The ultrasound picture was processed initially with a filter of low frequencies, histogram average, binarization, morphology, and edge detection to produce a binary image. Before classification, the image's characteristics are manually extracted using Gabor Wavelet. The greatest result was 82.55 percentage when SVM was employed 78.81 percentage when KNN with $k = 5$ was employed.

II. METHODOLOGY

Figure 1 demonstrates the procedure employed in the suggested technique. We employed the U-Net topology as the strength for the spatial RNN zones. Multiple combinations of integrating RNN elements were tested to increase adaptability to our segmentation task. Furthermore, by including RNN elements into the U-Net, instruction becomes additional tough and computationally costly [33]. As a consequence, intensive supervision was utilized to make training easier and increase segmentation performance.

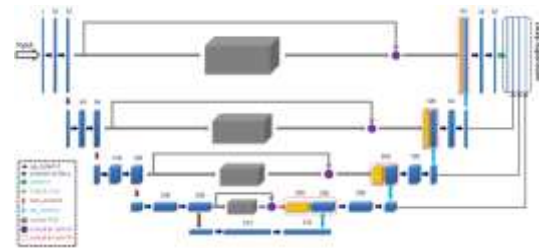


Fig. 1. The suggested pipeline for ovarian follicle separation.

The support is a universal U-Net, with personalized spatial RNN components inserted among the encoder and decoder. The proposed network, which uses four spatial RNN components, is known as CR-Unet. While the upper and base spatial RNN sections are removed, a new report of the recommended association, known as CR-Unet-V2, is constructed. After implementing RNN, the extra loss topics in the deep administration approach are applied at every decoder layer. The number of channels for each module is indicated by its number.

One benefit of the U-Net structure with decoders and encoders is that it utilizes four pooling layers to provide multi-scale attribute identification [8]. The decoder aggregates the encoder's attributes for every sheet. In contrast to the FCNN [9], which employs addition for similar pixels, feature fusion via splice may result in thicker feature blocks. However, none of them considered the spatial relationships between pixels in characteristic maps.

In typical, contextual knowledge, such as artifacts and shadowing noise, is useful when working with US pictures. Typically, our assignment entails detecting multiple minor follicles on a picture; these follicles, remarkably individuals that are physically close to one another, may be spatially connected. Furthermore, the ovary and follicles may have a spatial relationship. Spatial RNNs [10] are used in every layer among the encoder and decoder of the U-Net to create a supplementary powerful separation network by leveraging spatial context information [11].

The RNN is commonly benefited to handle 1D progressive data using three arrays of nodes: (1) the contribution node for statistics collecting, (2) the secreted node for changing inside states constructed on previous and current data, and (3) the yield node for production creation. A basic RNN may be used instantly for text, audio, and signal processing, but it needs be improved to handle 2D or 3D input (for example, pictures).

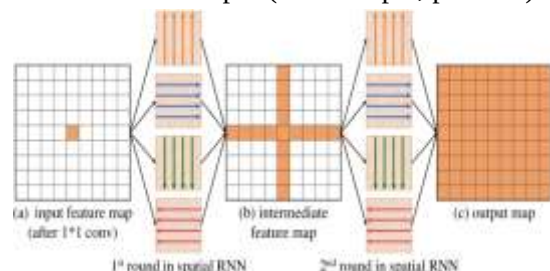


Figure. 2: The Spatial RNN module

Every feature slice in the encoder is convoluted with 11 kernels before being input into the spatial RNN, which performs four directional procedures. The outputs are then concatenated and twisted using the 11 kernel. As a result, the amount of channels is identical as the RNN response. Once again, the technique is followed to create the last feature map, at which time every pixel contains global spatial data. For example, Visin et al. decreased complexity by moving an RNN geographically (or horizontally) across a characteristic map. When identifying objects, Bell et al. employed spatial RNNs to extract spatially contextual properties.

Other varieties of RNNs exist now, such as gated recurrent units (GRU) [12], long short-term memory (LSTM) [13], and basic tanh RNN [14]. These RNNs are rarely employed directly in image analysis, although they are commonly used to improve the findings. Le et al. have expressed that by exchanging rectified linear units (ReLU) for the tanh activation function, basic RNNs may nominate network education more accessible and quicker than GRUs or LSTMs. As a consequence, we constructed our composite network's spatial RNN module using a basic RNN and ReLU. Figure 4 illustrates our

spatial RNN component, that receives feature plots as input and generates spatial RNN descriptions.

Using the feature data of a particular layer in the encoder as contribution, the RNN component conducts an input-to-hidden transformation first, resulting in a latest feature map. Following that, it combines regional spatial data in every principal target (up, down, left, and right) using four distinct data translations before concatenating and convolving the results with an 11 kernel to get the hidden-to-hidden conversion. As a consequence, an transitional characteristic map (Figure 2(b)) is generated, with every point including evidence about four ways.

Lastly, the preceding steps are continued to link regional spatial data to universal contextual data, yielding the end product map (Figure 2(c)). This section will show how to translate information in a spatial RNN component. As previously indicated, data translation occurs in four routes (right, left, down and up). We'll utilize information transformation to the left field as a representative (and likewise for other targets). We perform the following algorithm n times, using h_i as the attribute at pixel $(i,.)$.

$$h_{i,j} = \max(W_h^{left} h_{i,j-1} + h_{i,j}, 0)$$

Where n is the feature map's width(or height), and W_h^{left} is the weight parameter for the left direction in the recurrent translation layer. Given that the size of the input features to the spatial RNN component can impair the system's presentation, we suggest a CR-Unet-V2 modified that removes spatial RNN components from the primary (highest spatial RNN segment) and final layers (lowermost spatial RNN component) (Figure 1).

This is because the primary layer of the U-Net, with the significant input dimensions and the lowest receptive disciplines, removes low-altitude interpretations, which may be inadequate for encoding spatial environments

using RNN components. In contrast, the previous layer with huge receptive disciplines encodes the maximum features, making spatial RNNs' feature representation redundant.

Incorporating RNN to preserve long-range spatial dependences improves the likelihood of the gradient disappearance for immediate sequence actions, making it more challenging to fit all local sequence sections into the global loss task. By means of deep management, supplementary losses are inserted at every layer after RNN integration to replace the gradient flow; each decoder layer's output is up-sampled to the identical resolution as the input pictures and then recommended to construct the supplementary loss functions. From top to bottom, the loss weights for every layer were set to [1.0, 0.8, 0.4, 0.1]. Following the model's training, we use nature-supervision to improve the breakdown results for every test picture.

Every cycle uses the probability prediction map to produce a label for each testing image. Each pixel in our investigation contains three possibility expectation values: background, follicle, and ovary. For every position, the most important worth is put to 1, while the other values are decided to 0. This is one of the easiest techniques to generate labels from forecast results. The self-supervised algorithm is set to seven iterations. To boost the effectiveness of training and safeguard convergence, we employed the Logarithmic Dice loss function.

III. RESULTS AND DISCUSSION

Figure 3 shows a qualitative visual assessment of the ovary and follicle division effects achieved with the recommended technology and two existing cutting-edge methods: Deeplabv3 [15] and U-Net. On both ovarian and follicular separation, we can demonstrate that our approaches outperform all available cutting-edge techniques. On an ultrasound picture, the ovary and follicles are mistaken for adjoining tissues or constitutes, resulting in segmentation

collapse, as explained in the second and fifth columns of Figure 3. (a).

Alternative common occurrence is that the system inaccurately perceives artifacts to be segments (3rd, 4th, and 5th columns in Figure 3 (b)(d)), or that directs are considered as artifacts that should be ignored (3rd, 5th, and 7th columns in Figure 3(c)). Without a good identification of the universal semantic data underlying these things, it is hard to differentiate and section them effectively. Based on the data, we can infer that our technique (CR-Unet) outperforms other approaches (particularly the U-Net) in terms of target recognition and segmentation, as well as avoiding false positives.

Furthermore, while the CR-Unet performs marginally enhanced than the CR-Unet-V2, the CR-Unet-V2 trains twice as fast as the CR-Unet, and it is frequently simple to detect and part a pretty larger object (e.g., the ovary or large follicles), as shown in Figure 3. However, segmentation of small objects remains a difficult task, especially when several small targets appear in a picture.

The visual comparisons in Figure 3(e)-(h) show that our techniques can distinguish and section small things. As shown in Figure 3(e), all approaches correctly identify and section large follicles in the ovary, but only our approaches efficiently detect and sector minor follicles. Likewise, in Figure 3(h), the majority of techniques fail to differentiate and segment the follicles; only the suggested technique (CR-Unet) achieves accurate ovarian and follicle segmentation.

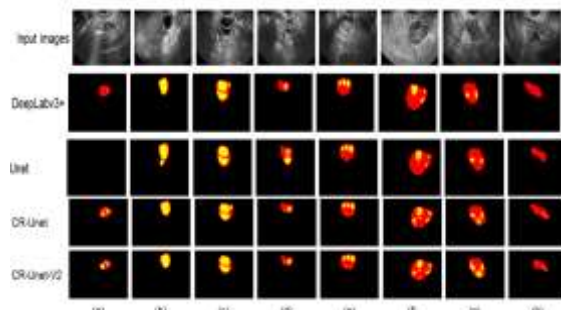


Figure 3. Qualitative comparison of ovarian and follicular segmentation results produced by our methods (CR-Unet and CR-Unet-V2) and other methods (Deeplabv3+ and U-Net) against ground truth (GT)

IV. CONCLUSION

This article describes an innovative composite deep learning architecture for autonomous ovarian and follicular segmentation. Our proposed approach makes use of global contextual information by adding spatial RNN components into a universal U-Net. We evaluated our approach using a 2D TVUS dataset and got an complete identification rate of 0.90 for follicles, as well as an overall Dice correspondence of 0.931 for the ovary and 0.885 for the follicles. Our methodology overtook further cutting-edge technologies in trials.

REFERENCES

1. G. E. Sarty, W. Liang, M. Sonka, and R. A. Pierson, "Semiautomated segmentation of ovarian follicular ultrasound images using a knowledge-based algorithm," *Ultrasound Med. Biol.*, vol. 24, no. 1, pp. 27–42, 1998, doi: 10.1016/S0301-5629(97)00213-5.
2. B. Cigale and D. Zazula, "Segmentation of ovarian ultrasound images using cellular neural networks," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 18, no. 4, pp. 563–581, 2004, doi: 10.1142/S0218001404003368.
3. Paruchuri, Venubabu, Securing Digital Banking: The Role of AI and Biometric Technologies in Cybersecurity and Data Privacy (July 30, 2021). Available at

SSRN:

<https://ssrn.com/abstract=5515258> or <http://dx.doi.org/10.2139/ssrn.5515258>

4. Y. Deng, Y. Wang, and Y. Shen, "An automated diagnostic system of polycystic ovary syndrome based on object growing," *Artif. Intell. Med.*, vol. 51, no. 3, pp. 199–209, 2011, doi: 10.1016/j.artmed.2010.10.002.
5. G. Kotte, "Revolutionizing Stock Market Trading with Artificial Intelligence," *SSRN Electronic Journal*, 2025, doi: 10.2139/ssrn.5283647.
6. Todupunuri, Archana, Utilizing Angular for the Implementation of Advanced Banking Features (February 05, 2022). Available at SSRN: <https://ssrn.com/abstract=5283395> or <http://dx.doi.org/10.2139/ssrn.5283395>
7. E. Setiawati, Adiwijaya, and A. B. W. Tjokorda, "Particle Swarm Optimization on follicles segmentation to support PCOS detection," 2015 3rd Int. Conf. Inf. Commun. Technol. ICoICT 2015, pp. 369–374, 2015, doi: 10.1109/ICoICT.2015.7231453.
8. Paruchuri, Venubabu, Leveraging Generative AI to Streamline Account Approval Processes and Improve the Precision of Risk Assessment in Financial Services (September 30, 2024). Available at SSRN: <https://ssrn.com/abstract=5473867> or <http://dx.doi.org/10.2139/ssrn.5473867>
9. B. Cahyono, Adiwijaya, M. S. Mubarak, and U. N. Wisesty, "An implementation of convolutional neural network on PCO classification based on ultrasound image," 2017 5th Int. Conf. Inf. Commun. Technol. ICoICT 2017, vol. 0, no. c, pp. 3–6, 2017, doi: 10.1109/ICoICT.2017.8074702.
10. G. KOTTE, "Overcoming Challenges and Driving Innovations in API Design

- for High-Performance Ai Applications,” Journal Of Advance And Future Research, vol. 3, no. 4, 2025, doi: 10.56975/jaafr.v3i4.500282.
11. R. M. Dewi, Adiwijaya, U. N. Wisesty, and Jondri, “Classification of polycystic ovary based on ultrasound images using competitive neural network,” J. Phys. Conf. Ser., vol. 971, no. 1, 2018, doi: 10.1088/1742-6596/971/1/012005.
12. Paruchuri, Venubabu, Optimizing Financial Operations with Advanced Cloud Computing: A Framework for Performance and Security (September 30, 2020). Available at SSRN: <https://ssrn.com/abstract=5515238> or <http://dx.doi.org/10.2139/ssrn.5515238>
13. B. Purnama, U. N. Wisesti, Adiwijaya, F. Nhita, A. Gayatri, and T. Mutiah, “A classification of polycystic Ovary Syndrome based on follicle detection of ultrasound images,” 2015 3rd Int. Conf. Inf. Commun. Technol. ICoICT 2015, pp. 396–401, 2015, doi: 10.1109/ICoICT.2015.7231458.
14. O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9351, pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4_28.
15. Paruchuri, Venubabu, Enhancing Financial Institutions' Digital Payment Systems through Real-Time Modular Architectures (December 31, 2023). Available at SSRN: <https://ssrn.com/abstract=5473846> or <http://dx.doi.org/10.2139/ssrn.5473846>
16. F. Milletari, N. Navab, and S. A. Ahmadi, “V-Net: Fully convolutional neural networks for volumetric medical image segmentation,” Proc. - 2016 4th Int. Conf. 3D Vision, 3DV 2016, pp. 565–571, 2016, doi: 10.1109/3DV.2016.79.
17. N. Rahaman et al., “S2RMs: Spatially Structured Recurrent Modules,” 2020, [Online]. Available: <http://arxiv.org/abs/2007.06533>.
18. A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” Artif. Intell. Rev., pp. 1–70, 2020, doi: 10.1007/s10462-020-09825-6.
19. M. Z. Alom et al., “A state-of-the-art survey on deep learning theory and architectures,” Electron., vol. 8, no. 3, pp. 1–67, 2019, doi: 10.3390/electronics8030292.
20. S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
21. Z. Alom, S. Member, M. Hasan, and C. Yakopcic, “Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation.”
22. H. Li et al., “CR-Unet: A Composite Network for Ovary and Follicle Segmentation in Ultrasound Images,” pp. 1–10.