

Automated Follicular and Ovarian Tumour Analysis Via Ultrasound Using U-Net Variants: A Review

Mrs. Rajeswari M¹, Dr. Monika G²

¹Department Of Computer Science, Assistant Professor, School Of Arts And Science, Vinayaka Mission Research Foundation Vinayaka Mission Chennai Campus, Chennai

²Associate Professor, Department Of Computer Science, School Of Arts And Science, Vinayaka Mission Research Foundation, Vinayaka Mission Chennai Campus, Chennai

ABSTRACT- Ultrasound is the first-line imaging modality for assessing the ovary, both for monitoring follicles in fertility care and for characterising adnexal and ovarian tumours. Interpretation, however, depends on operator skill and is affected by speckle noise, low contrast, and ambiguous boundaries. Deep learning, and in particular the U-Net family of encoder-decoder networks, has become the dominant approach for automating segmentation of the ovary, follicles, and masses. This paper reviews the literature on automated follicular and ovarian tumour analysis from ultrasound using U-Net variants. It summarises the clinical background, the architectural variants (plain, attention, nested, self-configuring, recurrent, and transformer-hybrid U-Nets), their application to follicle segmentation and adnexal mass segmentation, and the downstream tasks of quantification and classification. The reviewed studies report high overlap with expert outlines, for example Dice coefficients of 0.912 for the ovary and 0.858 for follicles with a composite U-Net, and 0.91 for adnexal masses using a bounding-box-guided U-Net. Evidence nevertheless remains constrained by small, single-centre, and retrospective datasets, limited external validation, device variability, and the scarcity of Indian data. The paper presents a conceptual analysis pipeline, compares key studies, and outlines research directions toward robust, clinically integrated tools.

Keywords- U-Net, Ovarian Ultrasound, Follicle Segmentation, Adnexal Mass, Ovarian Tumour, Medical Image Segmentation, Deep Learning, Polycystic Ovary Syndrome

1 INTRODUCTION

The ovary is examined by ultrasound for two broad purposes. In reproductive medicine, transvaginal ultrasound reveals the number, size, and position of follicles and their response to biological stimuli, which supports infertility assessment and assisted reproduction; manually screening large numbers of images is strenuous and prone to human error [5].

In gynaecological oncology, ultrasound is the main tool for evaluating adnexal masses, which are common and mostly benign, although ovarian cancer is the most lethal gynaecological malignancy and carries a 5-year survival of only about 30% when diagnosed at an advanced stage [1].

Standardised ultrasound risk models such as the IOTA Simple Rules [2], the ADNEX model [3], and the O-RADS US system [4] have improved assessment, but they rely on qualitative judgments that are subject to inter- and intra-observer variability, and expertise in ultrasound interpretation is not universally available [1]. Measuring the ovary and follicles by hand is likewise time-consuming and operator-dependent [6]. These limitations motivate automation, and systematic reviews of artificial intelligence

in gynaecologic imaging document its growing role [7], while reviews of machine learning for ultrasound describe the particular difficulties of the modality [8].

Among deep learning architectures, the U-Net and its descendants have become standard for medical image segmentation because they produce pixel-level masks from relatively modest training sets. Many variants now exist, and several have been tailored to ovarian ultrasound[22],[23].

This review brings the follicular and tumour applications together under one framework. Its objectives are to: (i) outline the clinical and technical challenges of ovarian ultrasound; (ii) describe U-Net variants relevant to the task; (iii) review follicle segmentation and ovarian tumour analysis studies; (iv) compare key studies and metrics; and (v) identify limitations and future directions, including relevance to the Indian setting.

II. REVIEW METHODOLOGY

A structured narrative review was undertaken. Searches combined terms for the anatomy and pathology (ovary,

ovarian follicle, polycystic ovary, adnexal mass, ovarian tumour, ovarian cancer) with terms for the modality and method (ultrasound, transvaginal ultrasound, segmentation, U-Net, attention, deep learning, transformer). Landmark architecture papers were included for context, and priority was given to studies that applied U-Net-type networks to ovarian or adnexal ultrasound, supplemented by selected classification and detection studies for downstream context. Table 1 summarises the scope. As this is a narrative review, no meta-analysis or formal risk-of-bias assessment was performed, and results are reported as stated by the original authors.

Table 1: Review scope and criteria

Element	Criteria
Topic scope	Automated analysis of ovarian follicles and ovarian or adnexal tumours on ultrasound using U-Net or related segmentation networks
Evidence types	Original studies, architecture papers, and reviews; selected classification and detection studies for context
Imaging focus	2D B-mode transvaginal and pelvic ultrasound; 3D and cine data where reported
Exclusions	Non-ultrasound-only studies without relevance to the pipeline; studies without a segmentation or deep learning component; opinion pieces
Synthesis	Thematic synthesis by architecture, task (follicle versus tumour), and evaluation

III. CLINICAL CONTEXT AND TECHNICAL CHALLENGES

I. Follicular Analysis

Follicle monitoring involves identifying the ovary, counting follicles, and measuring their size, tasks that are central to fertility assessment and are carried out manually by sonographers [5], [6].

The follicle is typically a dark fluid-filled structure within a textured ovary, but boundaries are often indistinct and neighbouring follicles may touch. Segmentation therefore underpins automated counts and measurements.

2. Ovarian And Adnexal Tumour Analysis Adnexal masses can be benign, borderline, or malignant, and incidence is about 35% in premenopausal and 17% in postmenopausal women, with roughly 85% benign [1]. Assessment depends on morphology, margins, echogenicity, solid components, acoustic shadowing, and vascularity [1]. Segmenting the mass from surrounding tissue is the first step in a computer-aided pipeline; manual outlining requires expertise, is time-consuming, and is error-prone even for experts [1].

3. Technical Challenges Of Ultrasound

Ultrasound images suffer from poor image quality, low contrast, boundary ambiguity, and complex anatomical shapes [6]. Images also vary by equipment manufacturer, sonographer preferences, acquisition approach, and processing parameters such as depth and gain, and are affected by speckle noise, artefacts, and acoustic shadowing [1]. These factors make ultrasound a demanding modality for machine learning [8] and explain why architectural refinements, careful preprocessing, and rigorous validation matter

1 U-Net And Its Variants

The U-Net, introduced for biomedical image segmentation, is a symmetric encoder-decoder network in which a contracting path captures context and an expanding path restores resolution, with skip connections passing fine detail between them; it relies on strong data augmentation to learn from limited annotated samples [9].

Subsequent work has modified this template to improve boundary accuracy, context modelling, and robustness. Table 2 summarises variants relevant to ovarian ultrasound.

Attention gates aim to focus the network on relevant structures [10]; nested skip pathways aim to reduce the semantic gap between encoder and decoder feature maps [11]; self-configuring frameworks adapt the whole pipeline to a dataset [12]; and transformer hybrids add global context to the local features of convolutions [13], [14].

For ovarian ultrasound, such additions are plausible responses to noisy, low-contrast data, but each should be judged by its measured performance on ovarian data rather than by general medical-imaging results.

The rationale column describes motivation, not reported results, except for CR-UNet and DC-UNet, which were designed and evaluated on ovarian ultrasound.

Table 2: U-Net variants and their relevance

Variant	Key modification	Rationale for ovarian ultrasound
U-Net [9]	Encoder-decoder with skip connections; heavy augmentation	Works with limited annotated data; common baseline for follicles and masses
Attention U-Net [10]	Attention gates on skip connections to highlight salient regions	May help suppress speckle and background in low-contrast images
UNet++ [11]	Nested, dense skip pathways	Narrows the gap between encoder and decoder features; may refine small follicle boundaries
nnU-Net [12]	Self-configuring pipeline adapting preprocessing and training to the dataset	Reduces manual tuning on small, heterogeneous datasets
TransUNet [13]	Transformer encoder combined with U-Net decoder	Captures global context across the ovary
Swin-Unet [14]	Pure Swin-transformer U-shaped network	Global context with hierarchical features
CR-UNet [6]	Spatial recurrent network inside a plain U-Net; self-supervised refinement	Designed for ovary and follicle segmentation in TVUS
DC-UNet [5]	Two contracting paths in a U-Net	Designed for fully automatic follicle segmentation

2. Follicle Segmentation Using U-Net Variants

CR-UNet incorporates a spatial recurrent neural network into a plain U-Net to learn multi-scale and long-range spatial context, and uses self-supervision to iteratively refine segmentation. On 3204 transvaginal ultrasound images from 219 patients, it reported Dice similarity coefficients of 0.912 for the ovary and 0.858 for follicles, the best among the compared methods [6]. This study is notable for segmenting

the ovary and follicles together, reflecting the anatomical relationship between them.

DC-UNet restructures the U-Net with two contracting paths to segment follicles with higher accuracy. It was trained and tested on approximately forty-two thousand annotated 2D ultrasound images from the USOVA3D Training Set 1 and was reported to outperform other U-Net-based models on the same dataset [5]. Earlier work explored U-Net-based approaches for segmenting ovarian follicles in gynaecological ultrasound [15]. Together these studies show that purpose-built U-Net variants can reach high overlap with expert annotation for follicles, and that public datasets such as USOVA3D support comparison between methods.

Reliable follicle masks enable downstream measurements such as follicle count and size, which are the quantities clinicians use to follow ovarian response during fertility treatment [5]. They also provide the basis for decision support in conditions assessed by ovarian morphology, although the studies reviewed here evaluate segmentation quality rather than clinical outcomes.

IV. OVARIAN TUMOUR AND ADNEXAL MASS ANALYSIS

Segmentation Of Masses

A U-Net guided by a clinician-drawn bounding box segmented adnexal masses on transvaginal ultrasound; on a test set from a cohort of 54 masses (36 benign and 18 malignant) from 53 patients, it achieved a median Dice coefficient of 0.91 against expert outlines.

An unsupervised fuzzy c-means step then separated internal hypoechoic and hyperechoic components, which correlated well with clinical review and could support radiomic analysis by component [1]. The authors noted that U-Net segmentation does not need very large training sets, especially with augmentation, because it operates at the pixel level [1].

Other studies provide context. A study of ultrasound images from 127 patients with ovarian cancer evaluated multiple U-Net variants, reporting mean Dice scores between 0.81 and 0.87 and examining the stability of radiomic features extracted from the segmentations [16].

Another reported a segmentation and classification approach in which the best U-Net had a median Dice of 0.85 on an external test set of 184 masses [17]. Work on adnexal mass segmentation with synthetic ultrasound data has also identified internal components such as cyst locules, solid elements, and papillary projections [18]. Most ultrasound radiomics classification studies for adnexal masses have relied on manual segmentation [1], which segmentation networks could replace.

V. CLASSIFICATION AND DETECTION

Segmentation is usually a step toward diagnosis. A deep convolutional network trained on pelvic ultrasound images from ten hospitals across China, with 34,488 images from 3,755 patients with ovarian cancer and 541,442 images from 101,777 controls, exceeded the average diagnostic level of radiologists, matched expert readers, and improved radiologists' accuracy, although prospective evaluation was recommended [19]. An international multicentre study of transformer-based models, using 17,119 images from 3,652 patients across 20 centres in eight countries, reported above-expert-level accuracy and strong generalisation, and addressed the earlier lack of external validation [20]. These studies are classification studies rather than U-Net segmentation studies, but they show where segmentation fits: it localises the lesion, and supports quantification and explanation alongside end-to-end classification.

1. Proposed Analysis Pipeline And Evaluation

Figure 1 gives a conceptual pipeline synthesised from the reviewed work. It is not an implemented system. Images are acquired by transvaginal or pelvic ultrasound, preprocessed (speckle reduction, normalisation, resizing, augmentation, and an optional clinician-provided region of interest), and segmented by a U-Net variant to produce ovary, follicle, or tumour masks. Quantification yields follicle counts and sizes or tumour shape, margins, internal components, and radiomic features, which feed decision support aligned with clinical systems such as O-RADS [4]. A clinician reviews the output.

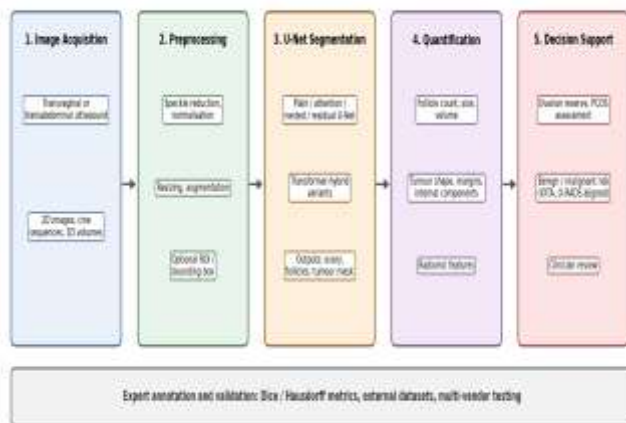


Figure 1: Conceptual pipeline for automated follicular and ovarian tumour analysis from ultrasound

Segmentation quality is usually reported with the Dice similarity coefficient [21], which measures overlap between predicted and reference masks, and with boundary-distance measures such as the Hausdorff distance, reported in one

study as a ratio to the effective diameter of the mass [1]. Overlap metrics alone can hide boundary errors, so both types of measure are useful. Table 3 compares the key segmentation studies, as reported by their authors.

Table 3: Comparison of key segmentation studies (values as reported by the original authors; datasets and protocols differ, so results are not directly comparable)

Study	Task and model	Data	Reported result
Li et al. [6]	Ovary and follicle; U-Net with spatial RNN (CR-UNet)	3204 TVUS images, 219 patients	Dice 0.912 (ovary), 0.858 (follicles)
Sarkar et al. [5]	Follicle; two-contracting-path U-Net (DC-UNet)	About 42,000 annotated 2D images, USOVA3D Set 1	Outperformed other U-Net-based models on same dataset
Whitney et al. [1]	Adnexal mass; U-Net with bounding box, plus fuzzy c-means	54 masses, 53 patients, single centre	Median Dice 0.91 [0.78, 0.96]
Jin et al. [16]	Ovarian cancer; multiple U-Net variants	127 patients	Mean Dice 0.81 to 0.87 (as summarised in [1])
Barcroft et al. [17]	Adnexal mass; segmentation plus classification	External test set of 184 masses	Best U-Net median Dice 0.85 (as summarised in [1])

2. Challenges And Limitations

Data size and diversity. Several studies use modest cohorts, such as 54 masses in the adnexal mass study, which was also single-centre and retrospective with strict exclusion criteria, including exclusion of images with incomplete borders or visible Doppler or measurement markings [1]. Such selection limits applicability to routine, imperfect images.

Domain shift. Differences between ultrasound systems, operators, and acquisition settings affect image appearance [1], and lack of external validation was a recognised gap in ovarian cancer deep learning until multicentre international validation appeared [20]. Segmentation models in particular need to be tested across vendors and centres.

Annotation and variability. High-quality masks require expert annotation, which is costly, and expert outlines themselves vary [1]. Datasets such as USOVA3D help benchmarking for follicles [5], but comparable public resources for tumours are scarcer.

Clinical integration. Most reviewed work evaluates overlap with expert outlines rather than clinical outcomes. Studies of classification have recommended prospective evaluation [19], and similar evidence is needed for segmentation-based tools, together with attention to interpretability, workflow fit, and regulation.

Fair comparison. Studies differ in datasets, splits, preprocessing, and metrics, so reported scores in Table 3 cannot be ranked directly. Head-to-head comparison of variants on common data is rare, and the choice of variant should be justified empirically.

3. Future Research Directions

Priorities include: (i) multi-centre, multi-vendor datasets with expert annotation, including Indian populations and routine-quality images; (ii) benchmarking of U-Net variants, including attention, nested, and transformer hybrids, on shared ovarian data with common protocols; (iii) joint models that segment ovary, follicles, and masses and link the masks to quantification and classification; (iv) video and 3D approaches that use sequences of images instead of single frames; (v) label-efficient learning, such as semi-supervised and self-supervised methods, to reduce annotation cost; (vi) uncertainty estimation and explainability, so that clinicians can see where the model is unsure; (vii) lightweight models for portable and point-of-care ultrasound devices; and (viii) prospective studies of clinical impact. The shortage of expert ultrasound examiners noted internationally [20] suggests such tools may be especially valuable where specialist access is limited, as in many parts of India, provided they are validated locally.

VI. CONCLUSION

U-Net variants have become the principal tools for automating segmentation of ovarian follicles and tumours on ultrasound. Purpose-built networks report high overlap with expert annotation for the ovary and follicles, and bounding-box-guided U-Nets achieve strong agreement for adnexal masses, creating a basis for follicle quantification and radiomic analysis. At the same time, small and single-centre datasets, device variability, limited external validation, and a lack of clinical outcome evidence mean the field is not yet ready for unsupervised use. Progress will depend on larger, more diverse annotated datasets, fair benchmarking of architectures, transparent reporting, and prospective evaluation within clinical workflows.

REFERENCES

1. H. M. Whitney, R. Yoeli-Bik, J. S. Abramowicz, L. Lan, H. Li, R. E. Longman, E. Lengyel, and M. L. Giger, "AI-Based Automated Segmentation for Ovarian/Adnexal Masses and Their Internal Components on Ultrasound Imaging," *Journal of Medical Imaging*, vol. 11, no. 4, 044505, 2024.
2. D. Timmerman et al., "Simple Ultrasound-Based Rules for the Diagnosis of Ovarian Cancer," *Ultrasound in Obstetrics & Gynecology*, vol. 31, no. 6, pp. 681–690, 2008.
3. B. Van Calster et al., "Evaluating the Risk of Ovarian Cancer Before Surgery Using the ADNEX Model to Differentiate Between Benign, Borderline, Early and Advanced Stage Invasive, and Secondary Metastatic Tumours: Prospective Multicentre Diagnostic Study," *BMJ*, vol. 349, g5920, 2014.
4. R. F. Andreotti et al., "O-RADS US Risk Stratification and Management System: A Consensus Guideline from the ACR Ovarian-Adnexal Reporting and Data System Committee," *Radiology*, vol. 294, no. 1, pp. 168–185, 2020.
5. M. Sarkar, A. Mandal, and A. Tudu, "DC-UNet: Looking for Follicles in the Ovarian Ultrasound Images," *Franklin Open*, vol. 8, 100149, 2024.
6. H. Li, J. Fang, S. Liu, X. Liang, X. Yang, Z. Mai, M. T. Van, T. Wang, Z. Chen, and D. Ni, "CR-Unet: A Composite Network for Ovary and Follicle Segmentation in Ultrasound Images," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 974–983, 2020.
7. P. Shrestha et al., "A Systematic Review on the Use of Artificial Intelligence in Gynecologic Imaging – Background, State of the Art, and Future Directions," *Gynecologic Oncology*, vol. 166, no. 3, pp. 596–605, 2022.
8. L. J. Brattain, "Machine Learning for Medical Ultrasound: Status, Methods, and Future Opportunities," *Abdominal Radiology*, vol. 43, no. 4, pp. 786–799, 2018.
9. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, LNCS vol. 9351, 2015, pp. 234–241.
10. O. Oktay et al., "Attention U-Net: Learning Where to Look for the Pancreas," *arXiv:1804.03999*, 2018.
11. Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for*

- Clinical Decision Support, LNCS vol. 11045, 2018, pp. 3–11.
12. F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, “nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation,” *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021.
13. J. Chen et al., “TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation,” *arXiv:2102.04306*, 2021.
14. H. Cao et al., “Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation,” *arXiv:2105.05537*, 2021.
15. S. Marques et al., “Segmentation of Gynaecological Ultrasound Images Using Different U-Net Based Approaches,” in *Proc. IEEE International Ultrasonics Symposium (IUS)*, 2019, pp. 1485–1488.
16. J. Jin et al., “Multiple U-Net-Based Automatic Segmentations and Radiomics Feature Stability on Ultrasound Images for Patients with Ovarian Cancer,” *Frontiers in Oncology*, vol. 10, 614201, 2021.
17. J. F. Barcroft et al., “Machine Learning and Radiomics for Segmentation and Classification of Adnexal Masses on Ultrasound,” *npj Precision Oncology*, vol. 8, 41, 2024.
18. C. Lebbos et al., “Adnexal Mass Segmentation with Ultrasound Data Synthesis,” in *Simplifying Medical Ultrasound*, S. Aylward et al., Eds. Cham: Springer, 2022, pp. 106–116.
19. Y. Gao et al., “Deep Learning-Enabled Pelvic Ultrasound Images for Accurate Diagnosis of Ovarian Cancer in China: A Retrospective, Multicentre, Diagnostic Study,” *The Lancet Digital Health*, vol. 4, no. 3, 2022.
20. F. Christiansen et al., “International Multicenter Validation of AI-Driven Ultrasound Detection of Ovarian Cancer,” *Nature Medicine*, 2025, doi: 10.1038/s41591-024-03329-4.
21. L. R. Dice, “Measures of the Amount of Ecologic Association Between Species,” *Ecology*, vol. 26, no. 3, pp. 297–302, 1945. [22] V Vetrivel, Dr. K Malarvizhi, Dr. D Kesavan “Customer satisfaction and trust of private bank ATM’s in Rural Areas,” *International Journal of Commerce and Management Research*, Vol 5, Issue 5, 2019, Pages 97-99
22. Hariharan, S., & Vetrivel, V. (2024). The impact of service quality in the success of traditional catering businesses in Coimbatore. In *Proceedings of the International Conference on Digital Transformation in Business: Navigating the New Frontiers Beyond Boundaries* (DTBNNF 2024) (pp. 123–138). Atlantis Press. https://doi.org/10.2991/978-94-6463-433-4_10