



Multiple U-Net-Based Automatic Segmentations and Radiomics Feature Stability on Ultrasound Images for Patients With Ovarian Cancer

Juebin Jin^{1†}, Haiyan Zhu^{2,3†}, Jindi Zhang³, Yao Ai^{4†}, Ji Zhang⁴, Yinyan Teng⁵, Congying Xie^{4,6*} and Xiance Jin^{4*}

OPEN ACCESS

Edited by:

Wei Wang,
The First Affiliated Hospital of
Sun Yat-Sen University, China

Reviewed by:

Jinghao Duan,
Shandong University, China
Qingtao Qiu,
Shandong Cancer Hospital, China

*Correspondence:

Congying Xie
wzxiecongying@163.com
Xiance Jin
jinxc1979@hotmail.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

Received: 05 October 2020

Accepted: 29 December 2020

Published: 18 February 2021

Citation:

Jin J, Zhu H, Zhang J, Ai Y, Zhang J,
Teng Y, Xie C and Jin X (2021) Multiple
U-Net-Based Automatic
Segmentations and Radiomics
Feature Stability on Ultrasound Images
for Patients With Ovarian Cancer.
Front. Oncol. 10:614201.
doi: 10.3389/fonc.2020.614201

¹ Department of Medical Engineering, Wenzhou Medical University First Affiliated Hospital, Wenzhou, China,

² Department of Gynecology, Shanghai First Maternal and Infant Hospital, Tongji University School of Medicine,

Shanghai, China, ³ Department of Gynecology, Wenzhou Medical University First Affiliated Hospital, Wenzhou, China,

⁴ Department of Radiotherapy Center, Wenzhou Medical University First Affiliated Hospital, Wenzhou, China,

⁵ Department of Ultrasound Imaging, Wenzhou Medical University First Affiliated Hospital, Wenzhou, China,

⁶ Department of Radiation and Medical Oncology, Wenzhou Medical University Second Affiliated Hospital, Wenzhou, China

Few studies have reported the reproducibility and stability of ultrasound (US) images based radiomics features obtained from automatic segmentation in oncology. The purpose of this study is to study the accuracy of automatic segmentation algorithms based on multiple U-net models and their effects on radiomics features from US images for patients with ovarian cancer. A total of 469 US images from 127 patients were collected and randomly divided into three groups: training sets (353 images), validation sets (23 images), and test sets (93 images) for automatic segmentation models building. Manual segmentation of target volumes was delineated as ground truth. Automatic segmentations were conducted with U-net, U-net++, U-net with Resnet as the backbone (U-net with Resnet), and CE-Net. A python 3.7.0 and package Pyradiomics 2.2.0 were used to extract radiomic features from the segmented target volumes. The accuracy of automatic segmentations was evaluated by Jaccard similarity coefficient (JSC), dice similarity coefficient (DSC), and average surface distance (ASD). The reliability of radiomics features were evaluated by Pearson correlation and intraclass correlation coefficients (ICC). CE-Net and U-net with Resnet outperformed U-net and U-net++ in accuracy performance by achieving a DSC, JSC, and ASD of 0.87, 0.79, 8.54, and 0.86, 0.78, 10.00, respectively. A total of 97 features were extracted from the delineated target volumes. The average Pearson correlation was 0.86 (95% CI, 0.83–0.89), 0.87 (95% CI, 0.84–0.90), 0.88 (95% CI, 0.86–0.91), and 0.90 (95% CI, 0.88–0.92) for U-net++, U-net, U-net with Resnet, and CE-Net, respectively. The average ICC was 0.84 (95% CI, 0.81–0.87), 0.85 (95% CI, 0.82–0.88), 0.88 (95% CI, 0.85–0.90), and 0.89 (95% CI, 0.86–0.91) for U-net++, U-net, U-net with Resnet, and CE-Net, respectively. CE-Net based segmentation achieved the best radiomics reliability. In conclusion, U-net based

automatic segmentation was accurate enough to delineate the target volumes on US images for patients with ovarian cancer. Radiomics features extracted from automatic segmented targets showed good reproducibility and for reliability further radiomics investigations.

Keywords: automatic segmentation, U-net, ultrasound images, radiomics, ovarian cancer

INTRODUCTION

Ovarian cancer remains the second most common gynecological malignancy and the leading cause of death in women with gynecological cancer (1). Several imaging modalities, such as computed tomography (CT), ultrasonography (US), positron emission tomography (PET), and magnetic resonance imaging (MRI) have been used as diagnostic and treatment assessment tools for gynecological cancer all over the world (2, 3). US is a well recognized and most common applied image modality for diagnosis and assessment of ovarian cancer due to its advantage characteristics of non-invasive, no radiation, cheap and affordable (4, 5). Recently, the emerging radiomics to find association between clinical characteristics and qualitative and quantitative information extracted from US images, has further expanded the application and importance of US images for gynecological cancer (6).

By converting medical images into quantitative information, which was then analyzed subsequently using conventional biostatistics, machine learning techniques, and artificial intelligence (7), radiomics has been developed rapidly for clinical application to promote precision diagnostics and cancer treatment (8, 9). Multiple processes, such as imaging acquisition, region of interests (ROIs) segmentation, image feature extraction, and modeling, were involved in the radiomics analysis, in which ROI segmentation is the most critical, challenging, and contentious step (7).

Segmentation is the step of extracting or distinguishing a ROI from its background. It is a common and crucial stage in the quantitative and qualitative analysis of medical images, and usually it is one of the most important and earliest steps of image processing (10). Due to the low contrast, speckle noise, low signal noise ratio and artifacts inherently associated with ultrasound images, it presents unique challenges for the analysis on US images, especially for accurate segmentation of different structures and tumor volumes compared with other image modalities, e.g., CT, MRI (11, 12). The image quality of US has a high intra- and inter-observer variability across different institutes and manufactures. It also highly depends on the abundance and experience of operators or diagnosticians. All these render manual segmentation more variable and significantly impact the quantitative (e.g., radiomics) and geometric analyses with US images (13, 14).

The US segmentation problems have been the hot research topics and rapidly evolved over the past few years (11). Currently, no golden standard for tumor segmentation had been established and manual segmentation is usually applied (15). However, except for the inter and intra varieties mentioned

above, the manual segmentation is also quite time consuming and boring. More recently, automatic segmentation techniques based on deep learning have become a main stream and show significant improvement in image classification predictions and recognition tasks (16). A well-known U-net architecture for biomedical imaging segmentation (17), which built upon fully convolutional network (18), has been successfully adapted to segment US images of breast (19), arterial walls (20), and gynecological cancer (21). Studies reported that the reproducibility and reliability of radiomics features could be deeply affected by the segmentation methods for CT (22), MR (23), and PET images (24). However, few studies have reported the reproducibility and stability of US based radiomics features obtained in oncology.

Previously, the feasibility of radiomics based on US images to predict the lymph node status for patients with gynecological cancer had been investigated (6). The purpose of this study is to investigate the accuracy of automatic segmentation algorithms based on multiple U-net models and their effects on radiomics features from US images for patients with ovarian cancer.

MATERIALS AND METHODS

Patients and Images

Patients with ovarian cancer underwent radical hysterectomy and transvaginal US diagnosis at authors' hospital from January 2002 to December 2016 were retrospectively reviewed in this study. The US images were acquired with a transvaginal ultrasonography using Voluson-E8 (GE Healthcare, Wilmington, USA) at 5–9 MHz, Philips (ATL HDI 5000, Netherland) at 4–8 MHz, and Esaote (MyLab classC) at 3–9 MHz or Hitachi (HI Vison Preirus) (Hitachi Ltd, Tokyo, Japan) at 4–8 MHz. All the images were reviewed with a Picture Archiving and Communication Systems (PACS).

Manual segmentation of target volumes was contoured by a radiologist with 7 years of experience in gynecological imaging and was further confirmed by a senior radiologist with > 15 years of experience in gynecological imaging. This retrospective study was approved by the Ethics Committee in Clinical Research (ECCR) of authors' hospital (ECCR#2019059). ECCR waived the need of written informed consent for this retrospective study. Patient data confidentiality was confirmed.

Automatic Segmentation Models

In this work, the classical U-net scheme and its multiple variations were used for the automatic segmentation task. Generally, the U-net is a symmetrical U-shaped model

consisting of an encoder-decoder architecture (17). The left side encoder is a down-sampling used to get feature map, similar to a compression operation, while the right side decoder is an up-sampling used to restore the encoded features to the original image size and to output the results. Skip-connection was added to encoder-decoder networks in order to concatenate the features of high- and low-level together (17). When Resnet is used as a fixed feature encoder to deepen the layers of the network and solve the vanishing gradient, the U-net structure is changed to U-net with Resnet as the backbone (U-net with Resnet) (25). Resnet34 was preferred in this study.

A so-called context encoder network (CE-Net) was also employed in this study, which consists of three major parts: a feature encoder module, a feature decoder module and a context extractor. In CE-Net, Resnet block is used as a fixed feature extractor; a residual multi-kernel pooling (RMP) block and a dense atrous convolution (DAC) block consist of the context extractor module (26). U-net++ is a modified U-net with deeply-supervised encoder-decoder network, in which a series of nested, dense skip pathways are applied to connect the encoder and decoder sub-networks (27). A typical U-net structure was shown in **Figure 1**.

Image Preprocessing

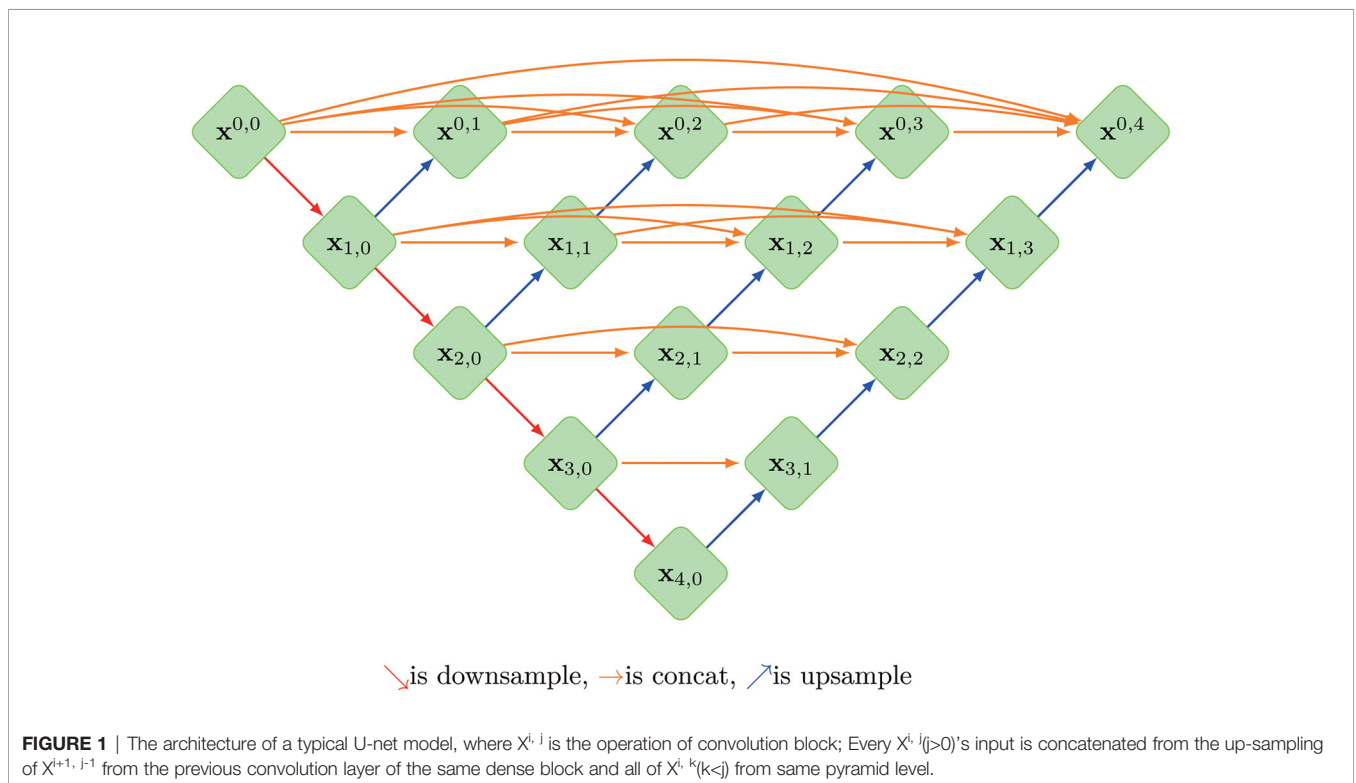
Image clipping was performed on each image set in order to satisfy the size requirement of U-net and to shift the center of clipping box so as to make the training model robust (28). The tumor center minus the offset (a number from 360 to 0 at -60 intervals) was selected as the starting point for a 480 * 512 clipping box. The clipping box should not exceed the image edge. A typical image preprocessing was shown in **Figure 2**.

Radiomics Feature Extraction

After manual and automatic segmentations, the arbitrary gray intensity values on US images were transformed into a standardized intensity range by intensity normalizing. A python 3.7.0 and package Pyradiomics 2.2.0 were used to extract radiomic features from the segmented target volumes. According to different matrices capturing the spatial intensity distributions with four different scales, 79 texture features and 18 first-order histogram statistics were extracted from neighborhood gray-level different matrix (NGLDM), gray level co-occurrence matrix (GLCM), grey-level zone length matrix (GLZLM), and gray-level run length matrix (GLRLM).

Evaluation and Statistical Analysis

The automatic segmentation models were built with the image dataset randomly divided into training sets, validation sets and test sets. The results of automatic segmentation models were evaluated by comparing them with manually segmented targets. Jaccard similarity coefficient (JSC), dice similarity coefficient (DSC), and average surface distance (ASD) were applied during the evaluation of delineation using the four U-net-related models with test data sets (29). The effects of segmentation on the radiomics features were evaluated with Pearson correlation coefficient and intraclass correlation coefficients (ICC), in which the agreement of a certain radiomic feature (e.g., shape features, texture features) between automatic and manual segmentation was evaluated by ICC (30). General statistical analyses were performed in SPSS Statistics (version 20.0.0). Statical significance was considered as a $p < 0.05$.



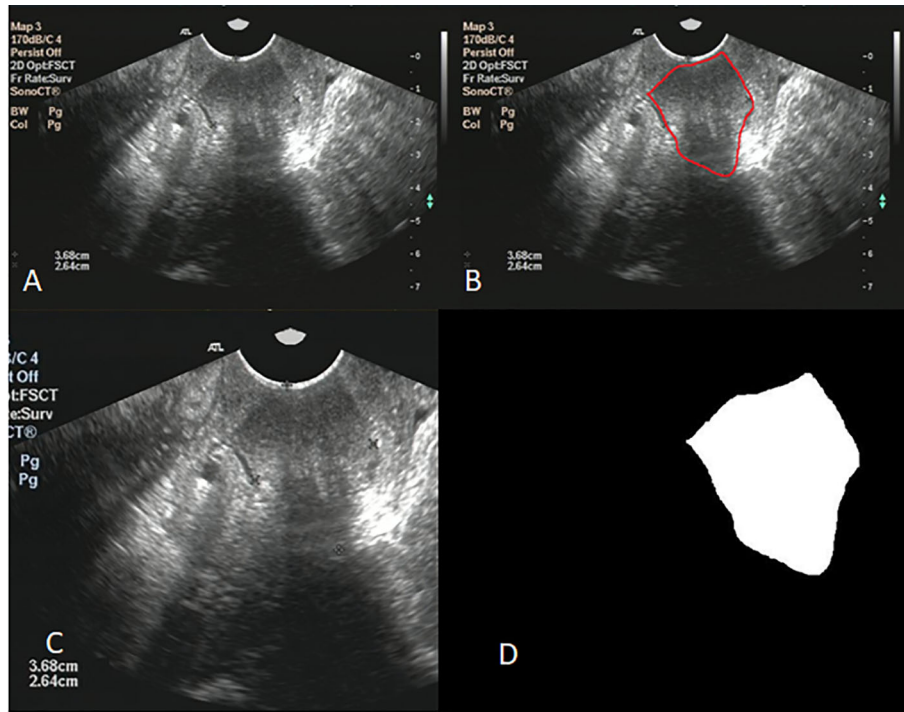


FIGURE 2 | (A) shows the original ultrasound image; (B) shows ovarian tumor segmented by radiologist; (C) shows the image after clipping; (D) shows the mask of ovarian.

RESULTS

There were 127 patients with ovarian cancer and with transvaginal US images included in the study. The median age of these patients was 56 years old (from 23 to 80 years). A total of 469 US images were analyzed and randomly divided into three groups: training sets (353 images), validation sets (23 images),

and test sets (93 images) for the building of automatic segmentation models. Detailed characteristics of patients and images were presented in **Table 1**. No significant difference among the training, validation, testing sets in terms of age, histological type, and tumor stages was observed.

U-net, CE-Net, U-net++, and U-net with Resnet were applied to delineate automatically the target volumes of ovarian cancer

TABLE 1 | Clinical characteristics of enrolled patients and images.

Category	Patients characteristics	Images			
		Training sets	Validation sets	Testing sets p	
Total number	127	353	23	93	
Age (years)					0.344
Mean	54.65	53.96	56.22	55.53	
Median	56	54	59	56	
Range	23~80	23~80	32~73	23~80	
SD	11.85	11.38	11.01	9.79	
Histological types					0.679
Epithelial	105	308	21	83	
None epithelial	16	38	2	7	
N.A.	2	12	1	5	
Tumor stages					0.691
I	36	87	4	24	
II	15	34	5	11	
III	68	213	13	54	
IV	3	19	1	4	
N.A.	1	5	1	2	

Note. p value is calculated from the univariate association test between sub-groups, one-factor ANOVA for continues variables, Fisher's exact test for categorized variables.

on US images. **Figure 3** presents typical contours achieved by these automatic segmentation models and their comparison with manual contours. Detailed results of segmentation accuracy metrics were presented in **Table 2**. CE-Net and U-net with Resnet achieved a DSC and JSC of 0.87, 0.79, and 0.86, 0.78, respectively. The ASD of CE-Net and U-net with Resnet were 8.54 and 10.00, respectively.

There were 97 features extracted from the delineated target volumes. **Figure 4** shows the hot maps of Pearson correlation and ICC for the comparison between features extracted from automatic segmentations and manual contours. The average Pearson correlation was 0.86 (95% CI, 0.83–0.89), 0.87 (95% CI, 0.84–0.90), 0.88 (95% CI, 0.86–0.91), and 0.90 (95% CI, 0.88–0.92) for U-net++, U-net, U-net with Resnet, and CE-Net,

respectively. The average ICC was 0.84 (95% CI, 0.81–0.87), 0.85 (95% CI, 0.82–0.88), 0.88 (95% CI, 0.85–0.90), and 0.89 (95% CI, 0.86–0.91) for U-net++, U-net, U-net with Resnet, and CE-Net, respectively.

High correlations were observed for most of the features except for some features of shape GLZLM. Detailed results of Pearson correlation and ICC for all the 97 features were presented in **Supplementary Tables 1 and 2**. Further analysis on the shape GLZLM features was shown in **Figure 5**. Sphericity and PerimeterSurfaceRatio were the two shape features that showed weak correlation between automatic and manual segmentations. Excluding these two shape features, the Pearson coefficient and ICC between features extracted by CE-Net and manual segmentation ranged from 0.71–0.98, and 0.70–0.97, respectively.

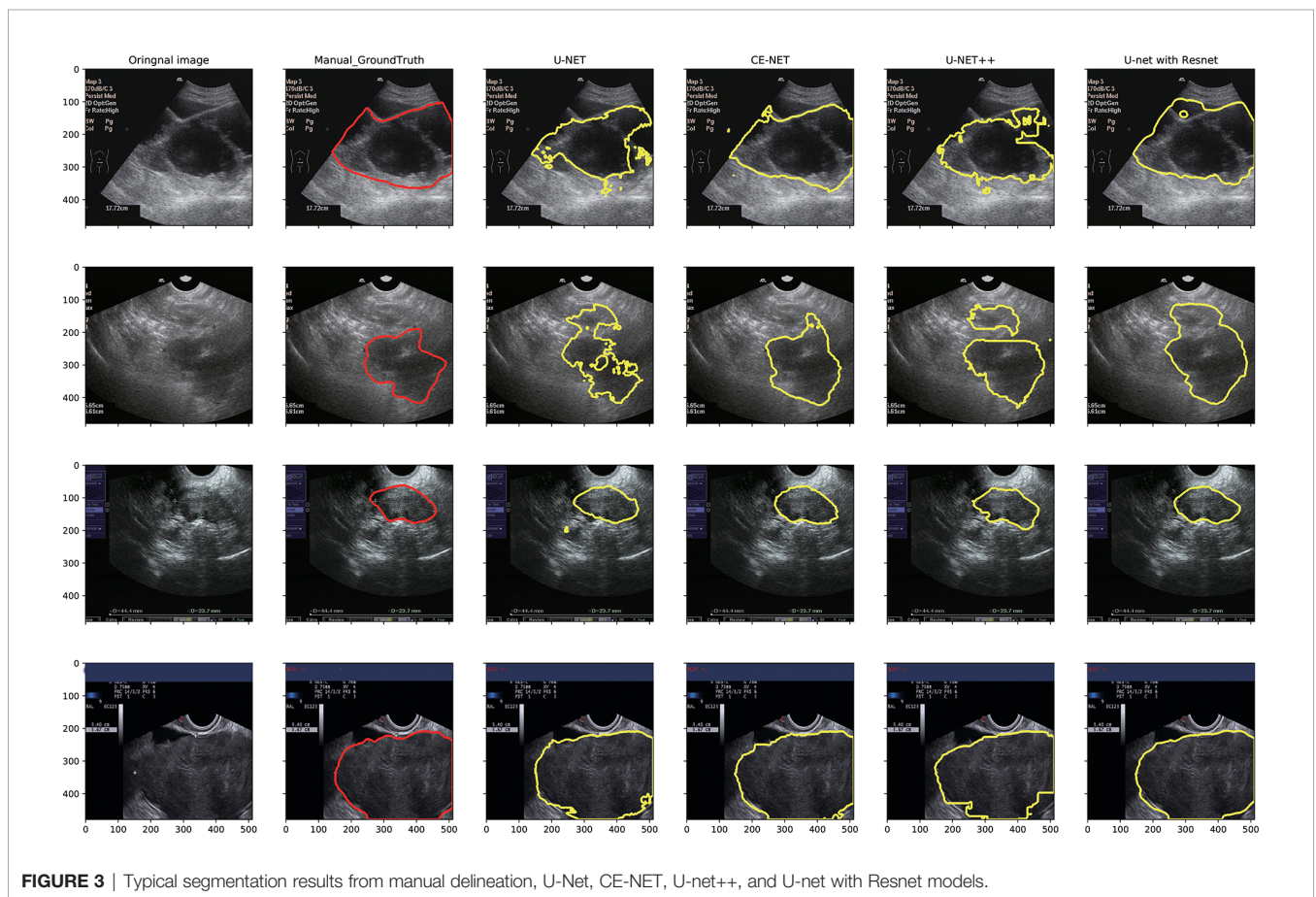


FIGURE 3 | Typical segmentation results from manual delineation, U-Net, CE-NET, U-net++, and U-net with Resnet models.

TABLE 2 | Automatic segmentation accuracy metrics for U-net-related models.

Models	Evaluation metrics					
	JSC (95%CI)		DSC (95%CI)		ASD (95%CI)	
U-net	0.71	0.68–0.75	0.81	0.79–0.85	10.57	9.25–11.89
CE-Net	0.79	0.76–0.82	0.87	0.85–0.90	8.54	7.21–9.86
U-net++	0.72	0.68–0.75	0.82	0.79–0.85	10.15	8.90–11.40
U-net with Resnet	0.78	0.75–0.82	0.86	0.83–0.89	10.00	8.03–11.97

Note. JSC, Jaccard similarity coefficient; DSC, Dice similarity coefficient; ASD, Average surface distance.

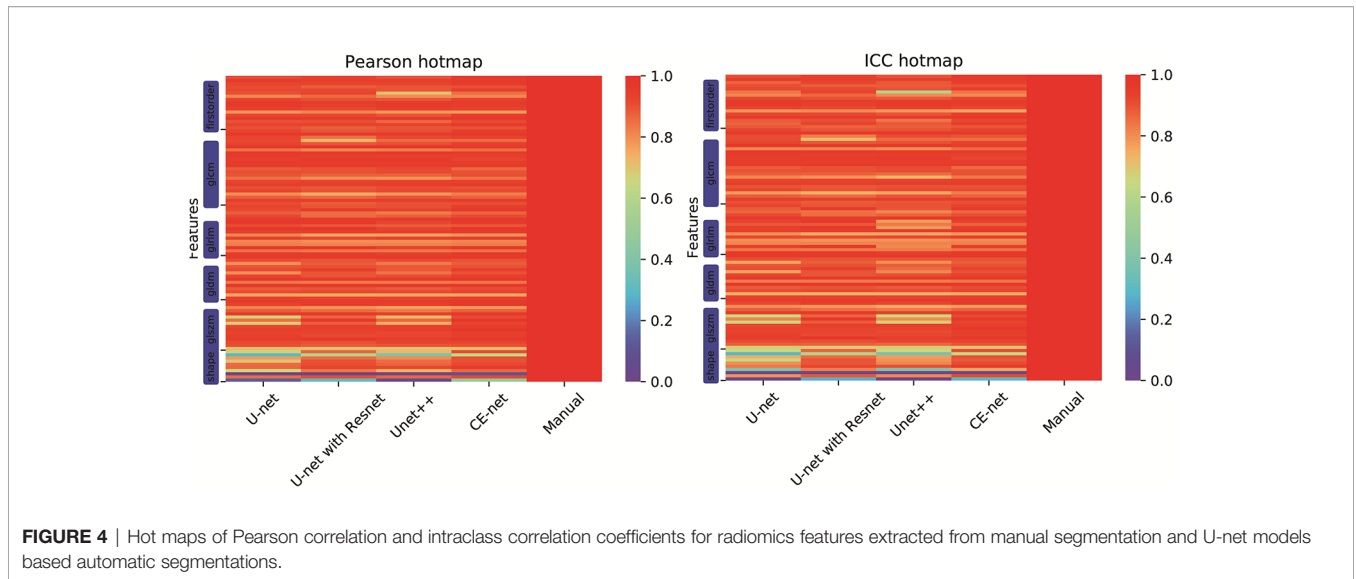


FIGURE 4 | Hot maps of Pearson correlation and intraclass correlation coefficients for radiomics features extracted from manual segmentation and U-net models based automatic segmentations.

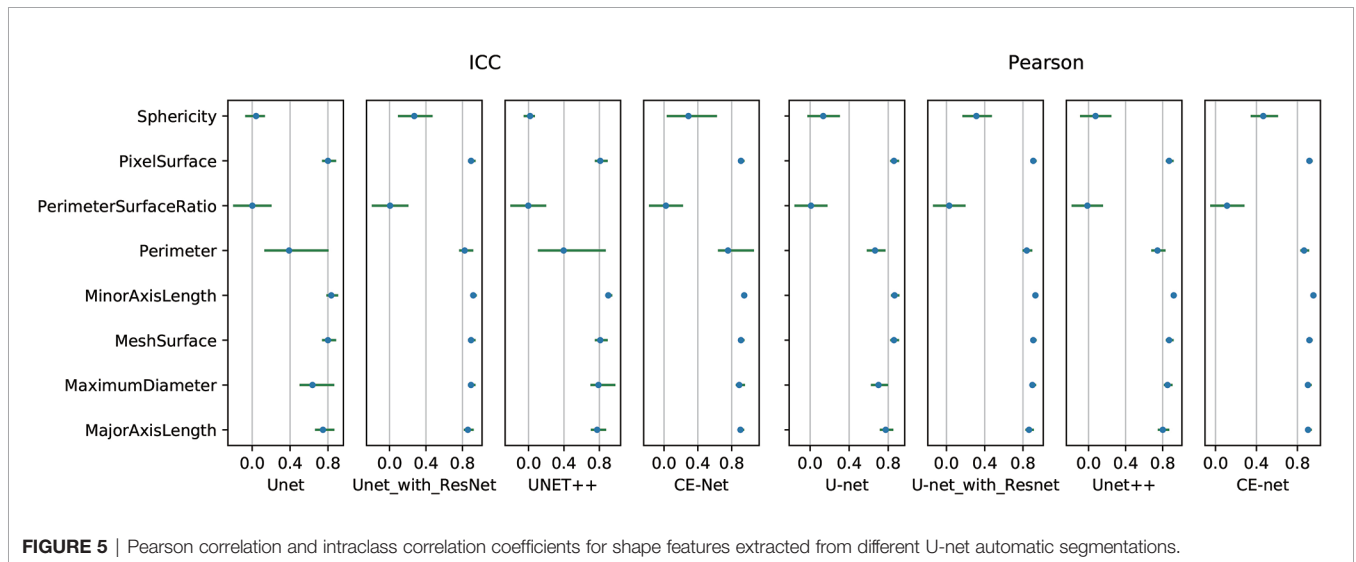


FIGURE 5 | Pearson correlation and intraclass correlation coefficients for shape features extracted from different U-net automatic segmentations.

DISCUSSION

Automatic segmentation of target volumes for ovarian cancer on US images were generated with multiple U-net models. The segmentation accuracy and its effects on radiomics features were evaluated in this study. CE-Net and U-net with Resnet models achieved a relatively higher accuracy on target delineation. Except for some shape features, most features extracted with automatic segmentation algorithms achieved high Pearson correlation and ICC in correlation with features extracted from manual contours.

US is a standard imaging modality for lots of diagnostic and monitoring purposes, had has been significantly investigated with deep learning based automatic segmentation (31). Yang et al. (32) used a fine-grained recurrent neural network to segment prostate US images automatically and achieved a high DSC around 0.92. Ghavami et al. (33) also proposed convolutional neural networks

(CNNs) to automatically segment transrectal US images of prostate and got a mean DSC of 0.91 ± 0.12 . Automatic segmentations on cardiac and carotid artery US images were proposed by Chen et al. (34) and Mechon-Lara et al. (35) using deep learning methods. Amiri et al. (36) fine-tuned the U-Net on breast US images and got a mean DSC of 0.80 ± 0.03 . Similarly, a DSC of 0.83 to 0.90 was achieved on US images of ovarian cancer using different U-net models in this study.

U-net is a structure for medical image segmentation with superior skip connections design for different stages of the network, which had inspired the development of many variations (37). Marques et al. (21) explored different U-Net architectures with various hyperparameters in their automatic segmentations on the transvaginal US images of ovary and ovarian follicles, and indicated that architecture takes into account the spatial context of ROI is important for a better performance (21). In this study, Unet++, U-net, CE-Net, and

U-net with Resnet were applied for the automatic segmentations. As shown in **Table 2**, CE-Net and U-net with Resnet exhibited higher mean DSC and JSF, and lower mean ASD compared with U-net++ and U-net, where CE-Net achieved the best performance.

In radiomics analysis, usually the ROI contoured is the region analyzed. The reproducibility and reliability of radiomics features were highly impacted by the segmentation methods. Parmar et al. (38) demonstrated that semi-automatic segmentation (ICC: 0.85 ± 0.15) provided a better feature extraction reproducibility than manual segmentation (ICC, 0.77 ± 0.17) in CT images for 20 non-small cell lung cancer patients. Heye et al. (23) achieved an ICC of 0.99 with a semiautomatic segmentation on dynamic contrast material-enhanced MR images. In this study, a highest Pearson correlation and ICC of 0.90 (95% CI, 0.88–0.92) and 0.89 (95% CI, 0.86–0.91) were achieved with CE-Net automatic segmentation. Similarly, Lin et al. (39) achieved an ICC of 0.70–0.99 on first-order apparent diffusion coefficient radiomics parameters using U-net automatic segmentation for cervical cancer.

However, a few of shape textures showed worse correlation, as shown in **Figures 4** and **5**. This may be caused by artifacts resulted from less optimal automatic segmentation algorithms as shown in **Figure 3**, which could be improved by manual correction during clinical practice. This also indicated that automatic segmentation for US images needs further investigation to improve the reliability and reproducibility of delineated volumes and radiomics features. Future evaluation of the reliability and reproducibility may be focused on prediction modeling level instead of at the level of radiomics features.

CONCLUSIONS

U-net based automatic segmentation was accurate enough to delineate the target volumes on US images for patients with ovarian cancer. Radiomics features extracted from automatic segmented ROI showed high reliability and reproducibility for further radiomics investigations.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin* (2017) 67:7–30. doi: 10.3322/caac.21387
2. Williams AD, Cousins C, Soutter WP, Mubashar M, Peters AM, Dina R, et al. Detection of Pelvic Lymph Node Metastases in Gynecologic Malignancy. *Am J Roentgenol* (2001) 177:343–8. doi: 10.2214/ajr.177.2.1770343
3. Alcázar JL, Arribas S, Mínguez JA, Jurado M. The Role of Ultrasound in the Assessment of Uterine Cervical Cancer. *J Obstet Gynecol India* (2014) 64:311–6. doi: 10.1007/s13224-014-0622-4
4. Meys EMJ, Kaijser J, Kruitwagen RFP, Slangen BFM, Van Calster B, Aertgeerts B, et al. Subjective assessment versus ultrasound models to diagnose ovarian cancer: A systematic review and meta-analysis. *Eur J Cancer* (2016) 58:17–29. doi: 10.1016/j.ejca.2016.01.007
5. Kyriazi S, Kaye SB, Desouza NM. Imaging ovarian cancer and peritoneal metastases - current and emerging techniques. *Nat Rev Clin Oncol* (2010) 7:381–93. doi: 10.1038/nrclinonc.2010.47
6. Jin X, Ai Y, Zhang J, Zhu H, Jin J, Teng Y, et al. Noninvasive prediction of lymph node status for patients with early-stage cervical cancer based on

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee in Clinical Research (ECCR) of the First Affiliated Hospital of Wenzhou Medical University. The ethics committee waived the requirement of written informed consent for participation.

AUTHOR CONTRIBUTIONS

CX and XJ conceptualized and designed the study. XJ and CX provided administrative support. CX and HZ provided the study materials or patients. JJ and HZ collected and assembled the data. YA, JZ, and YT analyzed and interpreted the data. JJ and XJ wrote the manuscript. JJ, HZ, JZ, YA, JDZ, YT, CX, and XJ gave the final approval of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was partially funded by the Wenzhou Municipal Science and Technology Bureau (2018ZY016, 2019) and National Natural Science Foundation of China (No.11675122, 2017).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.614201/full#supplementary-material>

- radiomics features from ultrasound images. *Eur Radiol* (2020) 30:4117–24. doi: 10.1007/s00330-020-06692-1
7. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* (2016) 278:563–77. doi: 10.1148/radiol.2015151169
8. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol* (2016) 34:2157–64. doi: 10.1200/JCO.2015.65.9128
9. Liu Z, Zhang XY, Shi YJ, Wang L, Zhu HT, Tang Z, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Clin Cancer Res* (2017) 23:7253–62. doi: 10.1158/1078-0432.CCR-17-1038
10. Shi C, Cheng Y, Wang J, Wang Y, Mori K, Tamura S. Low-rank and sparse decomposition based shape model and probabilistic atlas for automatic pathological organ segmentation. *Med Image Anal* (2017) 38:30–49. doi: 10.1016/j.media.2017.02.008
11. Noble JA, Boukerrouj D. Ultrasound image segmentation: A survey. *IEEE Trans Med Imaging* (2006) 25:987–1010. doi: 10.1109/TMI.2006.877092

12. Shi C, Cheng Y, Liu F, Wang Y, Bai J, Tamura S. A hierarchical local region-based sparse shape composition for liver segmentation in CT scans. *Pattern Recognit* (2016) 50:88–106. doi: 10.1016/j.patcog.2015.09.001
13. Zhao J, Zheng W, Zhang L, Tian H. Segmentation of ultrasound images of thyroid nodule for assisting fine needle aspiration cytology. *Heal Inf Sci Syst* (2013) 1:5. doi: 10.1186/2047-2501-1-5
14. Yazbek J, Amey L, Testa AC, Valentin L, Timmerman D, Holland TK, et al. Confidence of expert ultrasound operators in making a diagnosis of adnexal tumor: effect on diagnostic accuracy and interobserver agreement. *Ultrasound Obstet Gynecol* (2010) 35:89–93. doi: 10.1002/uog.7335
15. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: The process and the challenges. *Magn Reson Imaging* (2012) 30:1234–48. doi: 10.1016/j.mri.2012.06.010
16. Nair AA, Tran TD, Reiter A, Lediju Bell MA. A Deep Learning Based Alternative to Beamforming Ultrasound Images. *ICASSP IEEE Int Conf Acoust Speech Signal Process - Proc* (2018) 2018-April:3359–63. doi: 10.1109/ICASSP.2018.8461575
17. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Lect Notes Comput Sci (Incl Subser Lect Notes Artif Intell Lect Notes Bioinforma)* (2015) 9351:234–41. doi: 10.1007/978-3-319-24574-4_28
18. Shelhamer E, Long J, Darrell T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans Pattern Anal Mach Intell* (2017) 39:640–51. doi: 10.1109/TPAMI.2016.2572683
19. Yap MH, Goyal M, Osman FM, Martí R, Denton E, Juetta A, et al. Breast ultrasound lesions recognition: end-to-end deep learning approaches. *J Med Imaging (Bellingham Wash)* (2019) 6:11007. doi: 10.1117/1.JMI.6.1.011007
20. Yang J, Faraji M, Basu A. Robust segmentation of arterial walls in intravascular ultrasound images using Dual Path U-Net. *Ultrasonics* (2019) 96:24–33. doi: 10.1016/j.ultras.2019.03.014
21. Marques S, Carvalho C, Peixoto C, Pignatelli D, Beires J, Silva J, et al. Segmentation of gynaecological ultrasound images using different U-Net based approaches. *IEEE Int Ultrason Symp IUS* (2019) 2019-October:1485–8. doi: 10.1109/ULTSYM.2019.8925948
22. Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med Phys* (2011) 38:915–31. doi: 10.1118/1.3528204
23. Heye T, Merkle EM, Reiner CS, Davenport MS, Horvath JJ, Feuerlein S, et al. Reproducibility of dynamic contrast-enhanced MR imaging part II. Comparison of intra- and interobserver variability with manual region of interest placement versus semiautomatic lesion segmentation and histogram analysis. *Radiology* (2013) 266:812–21. doi: 10.1148/radiol.12120255
24. Hatt M, Lee JA, Schmidtlein CR, El Naqa I, Caldwell C, De Bernardi E, et al. Classification and evaluation strategies of auto-segmentation approaches for PET: Report of AAPM task group No. 211. *Med Phys* (2017) 44:e1–e42. doi: 10.1002/mp.12124
25. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV (2016). p. 770–8. doi: 10.1109/CVPR.2016.90
26. Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, et al. CE-Net: Context Encoder Network for 2D Medical Image Segmentation. *IEEE Trans Med Imaging* (2019) 38:2281–92. doi: 10.1109/TMI.2019.2903562
27. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. Unet++: A nested u-net architecture for medical image segmentation. *Lect Notes Comput Sci (Incl Subser Lect Notes Artif Intell Lect Notes Bioinforma)* (2018) 11045 LNCS:3–11. doi: 10.1007/978-3-030-00889-5_1
28. Aithal S, Krishna PK. Two Dimensional Clipping Based Segmentation Algorithm for Grayscale Fingerprint Images. *Soc Sci Electron Publ* (2017) 1:51–6. doi: 10.5281/zenodo.1037627
29. Heimann T, Van Ginneken B, Styner MA, Arzhaeva Y, Aurich V, Bauer C, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans Med Imaging* (2009) 28:1251–65. doi: 10.1109/TMI.2009.2013851
30. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* (2016) 15:155–63. doi: 10.1016/j.jcm.2016.02.012
31. Yang X, Yu L, Wu L, Wang Y, Ni D, Qin J, Heng PA, et al. Fine-grained recurrent neural networks for automatic prostate segmentation in ultrasound images. In: *31st AAAI Conference on Artificial Intelligence, AAAI*. San Francisco, CA (2017).
32. Yang X, Yu L, Li S, Wang X, Wang N, Qin J, Heng PA. Towards automatic semantic segmentation. In: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, Cham (2017). doi: 10.1007/978-3-319-66182-7_81
33. Ghavami N, Hu Y, Bonmati E, Rodell R, Gibson E, Moore C, et al. Integration of spatial information in convolutional neural networks for automatic segmentation of intraoperative transrectal ultrasound images. *J Med Imaging* (2018) 6:011003. doi: 10.1117/1.jmi.6.1.011003
34. Chen H, Zheng Y, Park JH, Heng PA, Zhou SK. Iterative multi-domain regularized deep learning for anatomical structure detection and segmentation from ultrasound images. In: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, Cham (2016). doi: 10.1007/978-3-319-46723-8_56
35. Menchón-Lara RM, Sancho-Gómez JL. Fully automatic segmentation of ultrasound common carotid artery images based on machine learning. *Neurocomputing* (2015) 45:215–28. doi: 10.1016/j.neucom.2014.09.066
36. Amir M, Brooks R, Rivaz H. Fine-Tuning U-Net for Ultrasound Image Segmentation: Different Layers, Different Outcomes. *IEEE Trans Ultrason Ferroelectr Freq Control* (2020) 67:2510–18. doi: 10.1109/TUFFC.2020.3015081
37. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. “3D U-net: Learning dense volumetric segmentation from sparse annotation”, In: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, Cham (2016). doi: 10.1007/978-3-319-46723-8_49
38. Parmar C, Velazquez ER, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One* (2014) 9:e102107. doi: 10.1371/journal.pone.0102107
39. Lin YC, Lin CH, Lu HY, Chiang HJ, Wang HK, Huang YT, et al. Deep learning for fully automated tumor segmentation and extraction of magnetic resonance radiomics features in cervical cancer. *Eur Radiol* (2019). doi: 10.1007/s00330-019-06467-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Jin, Zhu, Zhang, Ai, Zhang, Teng, Xie and Jin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.