Original Article

Automatic segmentation of the female pelvic floor muscles on MRI for pelvic floor function assessment

Xiaoqin Zhang[1]^, Yongjia Xiang[2,3], Jie Yao[1], Xin Hu[1], Yangyun Wang[4], Liping Liu[5], Yan Wang[2,6], Yi Wu[1]

[1]Department of Digital Medicine, College of Biomedical Engineering and Medical Imaging, Army Medical University (Third Military Medical University), Chongqing, China; [2]School of Mathematical Sciences, Chongqing Normal University, Chongqing, China; [3]Chongqing Zhence Science and Technology Co., Ltd., Chongqing, China; [4]Department of Urology, Fifth People's Hospital of Shanghai Fudan University, Shanghai, China; [5]Department of Gynecology and Obstetrics, First Affiliated Hospital of Army Medical University, Chongqing, China; [6]Chongqing Key Laboratory of Smart Finance and Big Data Analysis, Chongqing, China

Contributions: (I) Conception and design: Y Wu, Yan Wang, X Zhang; (II) Administrative support: Y Wu; (III) Provision of study materials or patients: L Liu, Yangyun Wang; (IV) Collection and assembly of data: X Hu, X Zhang; (V) Data analysis and interpretation: Y Xiang, J Yao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yi Wu, PhD. Department of Digital Medicine, College of Biomedical Engineering and Medical Imaging, Army Medical University (Third Military Medical University), 30 Gaotanyan Street, Shapingba District, Chongqing 400038, China. Email: wuy1979@tmmu.edu.cn; Yan Wang, PhD. School of Mathematical Sciences, Chongqing Normal University, University Town, Shapingba District, Chongqing 401331, China. Email: 20132095@cqnu.edu.cn.

Background: Pelvic organ prolapse (POP) is a pelvic floor dysfunction disease which affects females. The volume of pelvic floor muscle, especially the levator ani muscle (LAM), is an important indicator of pelvic floor function. However, muscle volume measurements depend on manual segmentation, which is clinically time-consuming. In this work, we present an efficient automatic segmentation model of pelvic floor muscles with magnetic resonance imaging (MRI) based on DenseUnet, to achieve muscle volume calculation and provide a reference for the assessment of pelvic floor function.

Methods: A total of 49 female pelvic floor magnetic resonance (MR) series were retrospectively enrolled from the First Affiliated Hospital of Army Military Medical University between 2013 and 2021, including 21 normal participants and 28 patients with stage 1–4 POP. The LAM, internal obturator muscle (IOM), and external anal sphincter (EAS) were manually segmented. An improved DenseUnet was proposed for automatic segmentation of these 3 muscles. The Dice similarity coefficient (DSC), Hausdorff distance (HD), and average symmetrical surface distance (ASSD) were used to evaluate segmentation results. The segmentation performance of the improved DenseUnet was compared with those of standard DenseUnet, ResUnet, Unet++, and Unet.

Results: The improved DenseUnet showed a good performance. The average DSC and standard deviation of the LAM, IOM, and EAS was 0.758±0.151, 0.716±0.173, and 0.810±0.147, respectively. The average HD was 22.41, 19.00, and 36.01 mm, respectively; and the average ASSD was 3.66, 3.80, and 5.23 mm, respectively. The average DSC and standard deviation of the normal group and POP group was 0.779±0.166 and 0.757±0.154, respectively. There was no significant difference between the muscle volume of the improved DenseUnet and manual segmentation (all P values >0.05). The average total segmentation time for 1 case was 10.18 s on our setup, which is much lower than the manual segmentation time of 45 minutes.

Conclusions: The improved DenseUnet segments the pelvic floor muscles in MRI quickly and efficiently, with good precision and faster speed than those of manual segmentation. This can assist doctors in quickly segmenting pelvic floor muscles, calculating muscle volume, and further evaluating pelvic floor function.

^ ORCID: 0000-0002-1406-7324.

Keywords: Segmentation; pelvic floor muscles; magnetic resonance imaging (MRI); levator ani muscle (LAM); DenseUnet

Submitted Nov 01, 2022. Accepted for publication Apr 14, 2023. Published online May 24, 2023.
doi: 10.21037/qims-22-1198
View this article at: https://dx.doi.org/10.21037/qims-22-1198

Introduction

Pelvic organ prolapse (POP) is a pelvic floor dysfunction disease affecting women. Childbirth injury, prolonged increased abdominal pressure, and obesity are the main causes of the disease. These conditions can result in weak or injured pelvic floor muscles, leading to POP. The decline and displacement of pelvic organs caused by POP can seriously affect women's daily lives and work. About 20% of women have undergone POP surgery during their lifetime (1). Therefore, the condition of the pelvic floor muscles is critical for women. The main muscles of the female pelvic floor muscles include the levator ani muscle (LAM), coccygeus muscle, anal sphincter complex, and urethral sphincter complexes (2,3). These muscles are the main structures of support for the pelvic organs including the bladder, urethra, uterus, vagina, and rectum, and control defecation (4,5). If these muscles become injured or atrophied, they lose the ability to adequately support the pelvic organs. This will result in POP or compromised bowel control, leading to problems such as urinary and fecal incontinence.

The LAM is the main component of the pelvic floor and has the crucial function of providing support to the pelvic viscera. The 3-dimensional (3D) shape of the LAM is funnel-like. Anteriorly to posteriorly, the upper end is attached to the posterior pubic bone, internal obturator muscle (IOM), and lateral ischium, adjacent to the external anal sphincter (EAS) and urethral sphincter. The connections and orientation of the 3 pelvic floor muscles are shown in *Figure 1*. The LAM and EAS dominate pelvic organ support, and their morphological recognition and 3D reconstruction are important for the diagnosis and reconstructive surgery planning of POP. The severity of POP is closely related to the injury and volume of the LAM (6,7). Therefore, segmentation and 3D reconstruction of the LAM and its adjacent EAS and IOM in magnetic resonance imaging (MRI) can provide a morphological reference for doctors to assess the degree of injury and function in the pelvic floor muscles, which is helpful for the surgical planning of POP.

However, the composition, 3D anatomical morphology, and spatial relationship of pelvic floor muscles are complicated. It is difficult for doctors to accurately identify and define their boundaries on tomographic images (8). Moreover, the process is time-consuming and wastes medical resources. Therefore, an automatic segmentation algorithm would be valuable in reducing the manual segmentation time and improving the efficiency of segmentation for pelvic floor muscles (9). Currently, the performance of existing segmentation algorithms for pelvic structures depends on image quality and patients' health status (10). The reliability of the algorithms needs to be improved, particularly for the segmentation of pelvic floor muscles.

Convolutional neural networks (CNNs) have been widely used in the processing of medical images, such as object detection, segmentation, and classification (11-14). Shen *et al.* (15) used a CNN model to segment abdominal skeletal muscle and fat automatically, which performed well in segmentation. However, pelvic structures are very complex and consist of many tissues and organs with different anatomical morphologies and spatial relationships, making them difficult to distinguish and segment. Based on pelvic floor ultrasound, many studies (16-20) have quantitatively assessed the pelvic floor function by automatic detection or segmentation of pelvic floor structures. Bonmati *et al.* (21) presented a self-normalizing CNN to segment the pelvic floor levator hiatus automatically in ultrasound images, achieving an accuracy equivalent to that of a previous semiautomatic approach. van den Noort *et al.* (22,23) trained a CNN to segment the puborectalis muscle and urogenital hiatus in the plane of minimal hiatal dimensions. In addition, they used these segmentations to measure relevant clinical parameters, including hiatus length, width, and area, which showed good agreement with manual segmentation. He *et al.* (24) proposed a distinctive curve-guided fully convolutional network that was accurate and robust for pelvic organ segmentation in computed tomography (CT) images. In order to perform qualitative or quantitative assessments of pelvic function, many

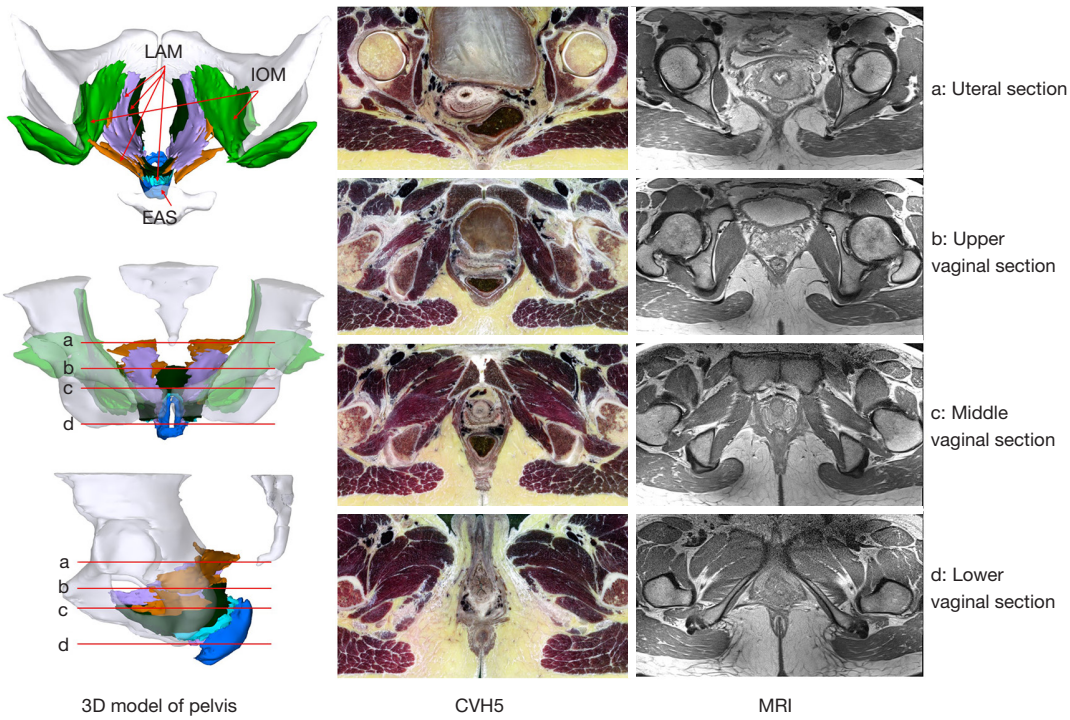

Figure 1 The connections and orientation of the 3 pelvic floor muscles. Column 1 displays different 3D views of CVH5 pelvis, column 2 shows different slices of CVH5 pelvis including the uteral, upper vaginal, middle vaginal, and lower vaginal section, column 3 shows the MR images of corresponding sections in 1 normal case. a: uteral section. b: upper vaginal section. c: middle vaginal section. d: lower vaginal section. LAM, levator ani muscle; IOM, internal obturator muscle; EAS, external anal sphincter; 3D, 3-dimensional; CVH5, the fifth Chinese Visible Human dataset; MRI, magnetic resonance imaging.

researchers (25-27) have tried to segment pelvic structures such as muscle, adipose tissue, and bone through machine learning or deep CNNs. However, these studies simply divided the pelvic floor muscles as a whole and did not separately segment and quantify the pelvic floor muscles such as the LAM and EAS. Moreover, to date, few studies have used deep learning algorithms to accurately segment the LAM and its adjacent muscles in pelvic MRI.

Based on the Chinese Visible Human datasets (28), we have studied the location, 3D anatomical morphology, and spatial relationship of pelvic floor muscles, including the LAM, anal sphincter complex, IOM, and urethral sphincter complex (4,5,29). We thus were able to identify the boundaries of the pelvic floor muscles in MRI more accurately.

In this retrospective study, we created an improved DenseUnet to automatically segment the female pelvic floor muscles including the LAM and EAS, and its adjacent IOM in pelvic MRI, aiming to segment more of the main pelvic floor muscles and improve the efficiency of segmentation compared with previous research. Based on the automatic and manual segmentation, we reconstructed the 3D models of the 3 muscles to perform 3D morphometric quantification of muscle volume, which is expected to provide a morphological reference for assessing the degree of injury of the pelvic floor muscles and the severity of POP. We present this article in accordance with the STARD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-22-1198/rc).

Methods

Image acquisition

The female pelvic floor MR series were collected from the First Affiliated Hospital of Army Military Medical University between 2013 and 2021. According to the inclusion and exclusion criteria, 21 normal participants and 28 patients with various degrees of POP, including POP patients with stage 1 (1 case), stage 2 (4 cases), stage 3 (21

cases), and stage 4 (2 cases), were enrolled in this study. The POP stages were assessed through the POP quantification (POP-Q) examination (30). The inclusion criteria were as follows: (I) images with no obvious noise or artifacts, (II) patients whose POP stages were determined by 2 gynecologists according to POP-Q examination, and (III) patients without stress urinary incontinence. The exclusion criteria were as follows: (I) images with very poor quality and (II) patients without POP-Q examination. The normal participants in this study comprised patients who did not experience pelvic floor prolapse. Participants enrolled in this study formed a random series. The participant selection process is displayed in Figure S1.

Imaging data were acquired using a 3.0 T MRI scanner, the Siemens 3.0 Trio (Siemens Healthineers, Erlangen, Germany), using a 4-channel surface coil. Approximately 30 minutes before the acquisition, the participants emptied their bladder. All women were placed in a supine position and relaxed their pelvic floor muscles during MRI examination. The MRI scans ranged from the sacral promontory to the lower edge of the perineum. A standardized protocol with the following parameters was used: turbo spin-echo sequences, a repetition time of 8,610 ms, an echo time of 9.8 ms, an axial slice orientation with a field of view of 280 mm × 280 mm, a pixel size of 0.55 mm × 0.55 mm, an image matrix of 512×512, and a slice thickness of 2 mm. The acquisition time of the sequence for 1 person was about 30 minutes.

Manual segmentation

Manual segmentation is recognized as the reference standard to evaluate automatic segmentation. As accurate manual segmentation is time-consuming for professional doctors; increasingly, studies are using manual segmentation as the standard to train deep learning models to achieve automatic segmentation in order to facilitate doctors' accurate manual segmentation and improve segmentation efficiency. In image segmentation, the region of interest (ROI) refers to the area that is segmented manually from the processed image in the form of irregular boundaries. Consequently, the ROIs of the pelvic floor muscles in our study were the LAM, IOM, and EAS. The 3 pelvic floor muscles were manually segmented under the guidance of 2 gynecologists with more than 10 years of clinical experience according to the muscle's anatomical appearance, the dense or loose connective tissue among muscles, and the color differences of the muscles. Amira software, version 5.2.2 (Mercury Computer Systems Inc., Chelmsford, MA, USA), was used in the process of manual segmentation and 3D reconstruction of the 3 muscles.

Image preprocessing

According to the position range of the LAM, EAS, and IOM, all scanned images were extracted and cropped to 512×290 pixels using the Pydicom package in Python (Python Software Foundation, Fredericksburg, VA, USA). Then, these cropped images were randomly divided into training and test sets at a ratio of approximately 8:2. The images of 17 normal cases and 22 POP cases were selected as the training set, and the images of the remaining 4 normal cases and 6 POP cases were used as the test set. There were 1,299 MR images in the LAM training set and 289 in the test set, 933 in the EAS training set and 257 in the test set, and 1,518 in the IOM training set and 383 in the test set. Finally, all images were preprocessed before model training and validation, which included data augmentation, resampling, and normalization.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This retrospective study was approved by the Ethics Committee of the First Affiliated Hospital of Army Military Medical University (No. BKY2021029), and individual consent for this retrospective analysis was waived.

Automatic segmentation framework

An illustration of the proposed pipeline is shown in *Figure 2*. In this study, automatic segmentation of the female pelvic floor muscles was performed using MRI based on the improved DenseUnet. Then, 3D models were reconstructed based on the segmentation from the gynecologists to assess the pelvic floor function.

Inspired by fully convolutional DenseNets (31), we designed DenseUnet with a context extraction module to overcome the shortcomings of Unet, which uses a little contextual and global information under different sensory fields. The segmentation framework of the pelvic floor muscles in MR images with the improved DenseUnet is shown in *Figure 3*. The improved DenseUnet is composed of 3 main parts: the encoder, context extraction, and decoder modules. Codes for the used network can be found in Appendix 1.

The encoder module consists of 4 dense blocks and 4 max-pooling layers. The dense connection of dense blocks improves the ability to reuse features and reduces the loss of

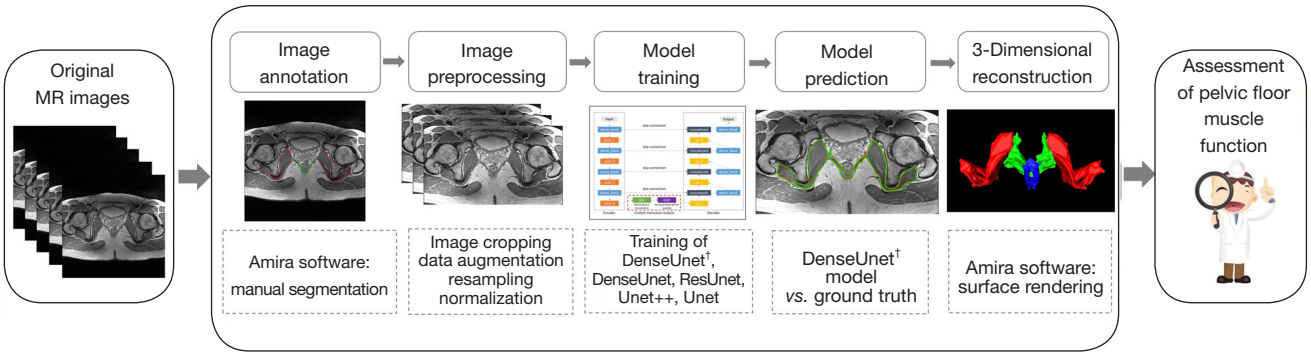

Figure 2 Overview of the proposed study pipeline. The main processes include 5 components: manual segmentation of the 3 muscles, a few preprocessing operations before training, training of 5 models, comparison of model prediction, and 3D reconstruction based on the results of manual segmentation and DenseUnet†. †, denotes the improved DenseUet. MR, magnetic resonance; 3D, 3-dimensional.

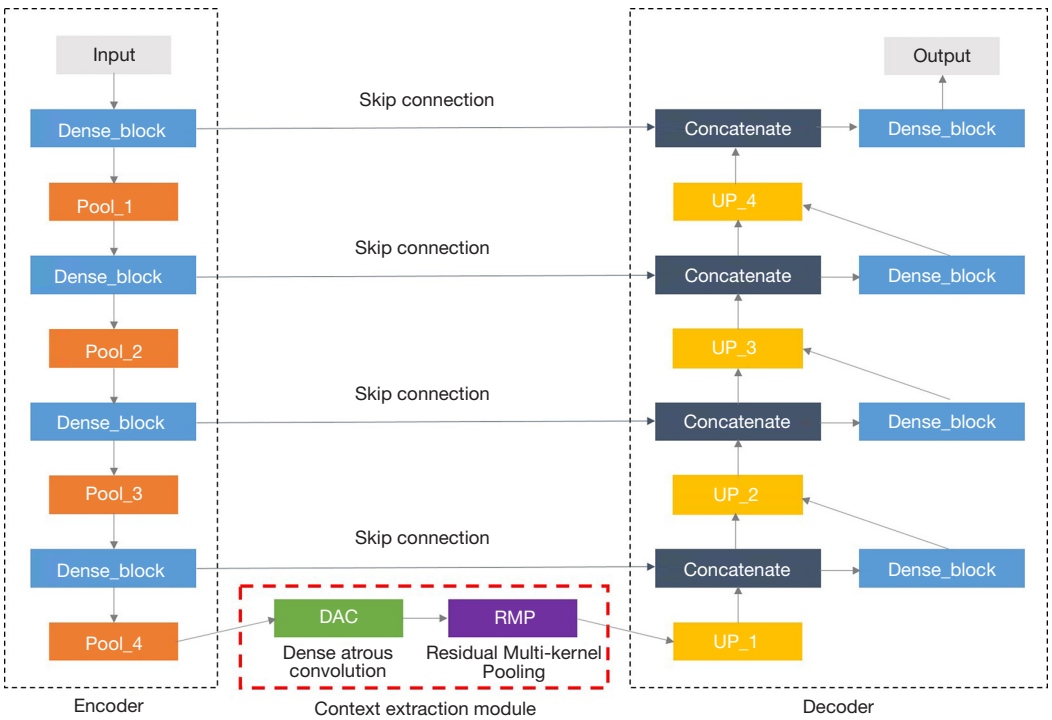

Figure 3 Illustration of the improved DenseUnet for segmentation. The framework is composed of 3 modules: encoder, context extraction, and decoder. The context extraction module consisting of a DAC module and a RMP module was optimized according to the characteristics of pelvic floor muscles. DAC, dense atrous convolution; RMP, residual multikernel pooling.

image information during model training. The max-pooling layers downsample the images to obtain more abstract and advanced image features.

The context extraction module consists of a dense atrous convolution (DAC) module and a residual multikernel pooling (RMP) module, which was first proposed by Gu *et al.* (32). These modules explore a large amount of contextual information from multiple receptive fields. We improved these 2 modules to better extract the characteristics of the pelvic floor muscles. In each branch of the DAC module, the size of the convolution kernel was 1×1 or 3×3; the number of convolutional kernels was 512;

and the expansion rate was 1, 3, or 5. To extract features from different scales, the sensing fields of each branch from left to right were 3, 7, 9, and 19, respectively. In the designed DAC module, the outputs of each branch were spliced on the channel axis to perform feature fusion and obtain contextual information. In the RMP module, the size of the convolutional kernels in max-pooling was 2×2, 3×3, 5×5, and 6×6, and the step size was 1. To reduce the number of parameters, a 1×1 convolutional kernel was connected to the pooling layer. Finally, the inputs and branch outputs were spliced together on the channel axis. Different sizes of convolutional kernels in max pooling allow the module to extract image features at different scales. Furthermore, considering the relative invariance of the LAM position, we set the step size of the max-pooling layer to 1 and used zero padding to keep the image size constant. This helps the network to better segment the LAM in the first few slices of each case. The input and output image sizes of the context extraction module remained the same so that they could be easily added to other modules. Moreover, this module can extract contextual semantic information to produce higher-order feature maps, thereby improving the segmentation framework accuracy.

The decoder module consists of 4 dense blocks and 4 upsampling layers. This module restores the images to input resolution. Upsampling does not completely restore image information. Therefore, we fused the shallow texture features with the deep semantic features through skip connection, thereby recovering the boundary information of some images and further improving the segmentation accuracy of the model.

The DenseUnet proposed in this study was implemented using the Keras framework with TensorFlow as the backend. The experiments were conducted on a server with Ubuntu 18.04.4 LTS (Canonical, London, UK) configured with an Intel Core i9-9820X CPU (3.30 GHz) and 2 NVIDIA GeForce (Nvidia, Santa Clara, CA, USA) RTX 2080Ti GPUs. Data augmentation (random rotation, random scaling, horizontal flipping, etc.) was performed as soon as the images were input into the model. Then, the images were resampled to 256×256 pixels, and the image value were normalized between 0 and 1. Adam was chosen as the optimizer. Dice loss was selected as the loss function, while the rectified linear unit (ReLU) function with Kaiming normal initialization was selected as the convolution kernel activation function (33). The parameters were as follows. As the default, the learning rate was set as 3×10^{-4}, the batch size was 2, and the number of steps in 1 epoch represented the number of training pictures divided by 2. Both the parameters γ and β for batch normalization use L_2 regularization with regularization factor 10^{-4}. The exponential decay rates for the estimation of first and second moments in Adam optimizer are 0.9 and 0.999, respectively. The parameter epoch was chosen according to the model training experiments. We set the value of epoch as 30, 60, and 90, and compared the Dice loss of segmentation results obtained by the proposed network on the training set. The experimental results showed that the Dice loss was relatively lower when the epoch was set as 60. In addition, in order to avoid the loss function not decreasing due to a too-large learning rate, a learning rate decay was adopted; that is, after 40 rounds, the learning rate of each round was multiplied by $e^{-0.1}$.

Evaluation

The segmentation results were evaluated by comparing the results with the reference standard from manual segmentation. In this study, we used Dice similarity coefficient (DSC) (34) as an important criterion for evaluating the segmentation performance. DSC is typically used to calculate the similarity between 2 sets of samples. Here, it represents the similarity between the manually segmented image and corresponding image segmented by the algorithm. This metric is defined as follows:

$$\mathrm{DSC}=\frac{2|X\cap Y|}{|X|+|Y|} \quad [1]$$

Here, X denotes the manually segmented image (ground truth), and Y denotes the corresponding image segmented by the algorithm.

In order to better evaluate segmentation performance, we also used Hausdorff distance (HD) and average symmetric surface distance (ASSD) (35) to measure the quantitative segmentation results. These 2 indicators are distance measures that are used to measure the differences in region boundaries. For HD and ASSD, the smaller the values are, the better the performance. These metrics are defined as follows:

$$\mathrm{HD}=\max\left(\max_{x\in X}\left\{\min_{y\in Y}\|x-y\|\right\},\max_{y\in Y}\left\{\min_{x\in X}\|y-x\|\right\}\right) \quad [2]$$

$$\mathrm{ASSD}=\frac{1}{|X|+|Y|}\times\left(\sum_{x\in X}d(x,Y)+\sum_{y\in Y}d(y,X)\right) \quad [3]$$

Here, x is the pixel in X, and y is the pixel in Y. d represents the Euclidean distance between the 2 pixels.

Table 1 Basic characteristics of the normal group and POP group

Characteristic	Normal group (n=21)	POP group (n=28)	P value*
Age (years)	35.0 (19.5)	51.0 (21.5)	0.001
BMI (kg/m^2)	24.8 (3.5)	24.3 (3.0)	0.628
Pregnancy	1.0 (1.0)	3.0 (1.0)	0.000
Vaginal birth	1.0 (1.5)	2.0 (1.0)	0.000
Abortion	0.0 (1.0)	1.0 (2.0)	0.007
Number of total images	1,189	1,411	0.075

Values are reported in medians and interquartile ranges, unless indicated otherwise. Numbers in parentheses are the interquartile ranges. *, the Mann-Whitney test was used for the P value computation. P<0.05 is considered statistically significant. POP, pelvic organ prolapse; BMI, body mass index.

We resized the output segmentation images to the original sizes of input ground truth, and recorded the DSC, HD, and ASSD of the improved DenseUnet, standard DenseUnet, and 3 other models [Unet (36), Unet++ (37), and ResUnet (38)] for the automatic segmentation of the LAM, EAS, and IOM. In addition, we compared the segmentation and reconstruction performances of the improved DenseUnet on different slices, including the uteral, upper vaginal, middle vaginal, and lower vaginal sections. The times required for manual and automatic segmentation were recorded.

Statistical analysis

Statistical analysis was performed using SPSS 24.0 software (IBM Corp., Armonk, NY, USA). The quantitative segmentation results and segmentation times were reported in means and standard deviations. According to the normality test, the basic characteristics of participants and muscles' volume were not normally distributed data, so they are reported in medians with interquartile ranges. The Mann-Whitney test was used for the P value computation to compare the muscle volume between manual and automatic segmentation. A P value <0.05 was considered statistically significant.

Results

Data on age, body mass index (BMI), number of pregnancies, vaginal birth, abortions, and total images in the normal and POP group were collected. The basic characteristics of the 49 participants are listed in *Table 1*. Except for age, there were significant differences in these characteristics between the normal group and the POP group (P<0.05).

The segmentation performance of Unet, Unet++, ResUnet, standard DenseUnet, and the improved DenseUnet for the LAM, EAS, and IOM is displayed in *Table 2*. There were no adverse events in our study, and there were no indeterminate or missing data in our results. The average DSC and HD of the improved DenseUnet for segmenting the LAM were 0.758 and 22.410 mm, respectively, which were better than those of the 3 other segmentation models. However, Unet had better segmentation results on IOM, with DSC, HD, and ASSD values of 0.809, 35.996 mm, and 4.989 mm, respectively. For the EAS, ResUnet achieved the highest DSC of 0.729.

Furthermore, as shown by the qualitative segmentation results in *Figure 4*, the improved DenseUnet model had different segmentation results for the 3 muscles in different sections. Muscles in the upper section of the pelvis, such as the LAM and IOM, had a higher DSC. Their segmentation contours were similar to the results of manual segmentation. Only some of the finer protrusions or depressions were not accurately segmented. The DSC of the muscles in the lower section of the pelvis was lower. Generally, the closer to the perineum, the lower the segmentation accuracy was.

The quantitative segmentation results of the improved DenseUnet for the different groups are shown in *Table 3*. These results showed that from POP2 to POP4, the heavier the degree of prolapse was, the lower the segmentation DSC of the pelvic floor muscles. This aspect was particularly evident in the automatic segmentation of the LAM and EAS. Overall, the segmentation performance of the normal group was better than that of the POP group.

The qualitative 2D segmentation and 3D reconstruction results of the 3 structures on different sections in the normal and POP group with different stages are shown in *Figure 5*. The 3D models of the 3 pelvic floor muscles reconstructed

Table 2 Quantitative segmentation results of improved DenseUnet and the compared models for 3 structures in the testing set

Models	LAM			EAS			IOM		
	DSC (95% CI)	HD (95% CI)	ASSD (95% CI)	DSC (95% CI)	HD (95% CI)	ASSD (95% CI)	DSC (95% CI)	HD (95% CI)	ASSD (95% CI)
Unet	0.749±0.164 (0.731, 0.768)	22.993±22.266 (20.415, 25.571)	3.732±5.190 (3.132, 4.333)	0.717±0.172 (0.696, 0.738)	22.295±34.542 (18.051, 26.538)	4.129±7.736 (3.179, 5.079)	0.809±0.140 (0.795, 0.823)	35.996±47.235 (31.238, 40.755)	4.989±10.907 (3.890, 6.088)
Unet++	0.749±0.169 (0.729, 0.769)	23.874±26.245 (20.835, 26.912)	4.719±12.926 (3.223, 6.216)	0.708±0.170 (0.688, 0.729)	17.817±13.361 (16.176, 19.458)	3.399±2.869 (3.046, 3.751)	0.807±0.137 (0.794, 0.821)	37.698±48.430 (32.820, 42.577)	5.223±10.223 (4.194, 6.253)
ResUnet	0.753±0.162 (0.734, 0.771)	22.884±22.311 (20.301, 25.467)	3.593±4.813 (3.036, 4.150)	0.729±0.169 (0.708, 0.750)	18.015±26.890 (14.712, 21.319)	3.092±2.990 (2.725, 3.459)	0.807±0.152 (0.792, 0.823)	38.770±52.676 (33.464, 44.076)	5.798±12.988 (4.490, 7.107)
DenseUnet	0.756±0.161 (0.737, 0.774)	23.953±30.851 (20.381, 27.525)	3.810±5.645 (3.156, 4.463)	0.710±0.177 (0.688, 0.732)	18.095±13.972 (16.378, 19.811)	3.721±4.552 (3.162, 4.280)	0.806±0.153 (0.790, 0.821)	38.939±53.018 (33.598, 44.279)	5.631±11.807 (4.441, 6.820)
DenseUnet†	0.758±0.151 (0.741, 0.775)	22.410±24.342 (19.592, 25.229)	3.661±5.308 (3.047, 4.276)	0.716±0.173 (0.695, 0.738)	19.003±26.068 (15.801, 22.205)	3.803±7.004 (2.942, 4.663)	0.810±0.147 (0.795, 0.825)	36.009±47.494 (31.225, 40.793)	5.226±12.210 (3.996, 6.456)

All values are described as means ± standard deviations. Numbers in parentheses are the confidence intervals. The units of HD and ASSD are in millimeters. †, denotes the improved DenseUet. LAM, levator ani muscle; EAS, external anal sphincter; IOM, internal obturator muscle; DSC, Dice similarity coefficient; CI, confidence interval; HD, Hausdorff distance; ASSD, average symmetrical surface distance.

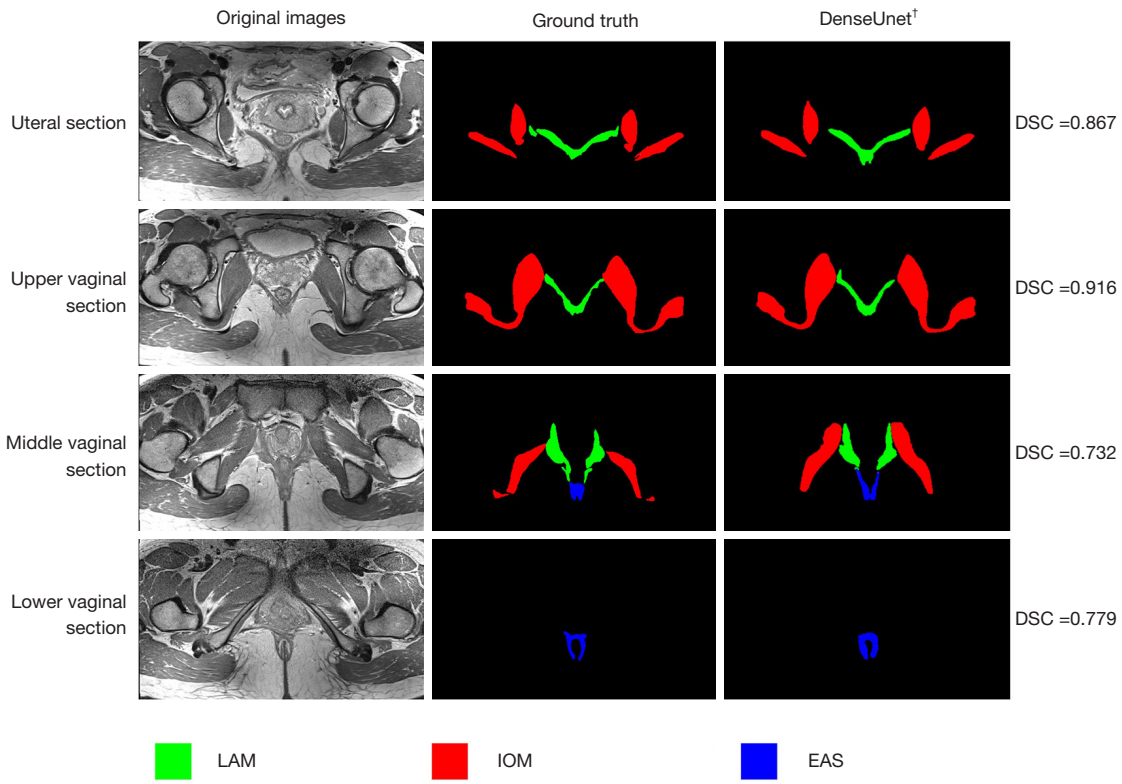

Figure 4 Qualitative segmentation results of the improved DenseUnet on different slices including the uteral, upper vaginal, middle vaginal, and lower vaginal sections. The 4 sections that have different 2D shapes and special relationships can represent the whole pelvic floor anatomy including the uterus, vagina, the LAM, EAS, and IOM. The slices with relative better segmentation performance in each section are shown. †, the improved DenseUet. DSC, Dice similarity coefficient; 2D, 2-dimensional; LAM, levator ani muscle; IOM, internal obturator muscle; EAS, external anal sphincter.

Table 3 Quantitative segmentation results on DSC (95% CI) of the improved DenseUnet for 3 structures in different groups

Muscle type	Normal	POP2	POP3	POP4
LAM	0.767±0.170 (0.739, 0.796)	0.764±0.182 (0.687, 0.841)	0.748±0.106 (0.726, 0.769)	0.740±0.152 (0.679, 0.801)
EAS	0.736±0.160 (0.708, 0.763)	0.792±0.129 (0.732, 0.852)	0.705±0.177 (0.667, 0.743)	0.557±0.193 (0.464, 0.650)
IOM	0.818±0.159 (0.795, 0.840)	0.865±0.035 (0.848, 0.882)	0.790±0.148 (0.766, 0.814)	0.833±0.054 (0.811, 0.854)
Average	0.779±0.166 (0.764, 0.794)	0.802±0.141 (0.767, 0.838)	0.755±0.149 (0.739, 0.771)	0.727±0.175 (0.685, 0.768)

DSC values are described as means ± standard deviations. Numbers in parentheses are the 95% CIs. DSC, Dice similarity coefficient; CI, confidence interval; POP2, patients with stage 2 pelvic organ prolapse; POP3, patients with stage 3 pelvic organ prolapse; POP4, patients with stage 4 pelvic organ prolapse; LAM, levator ani muscle; EAS, external anal sphincter; IOM, internal obturator muscle.

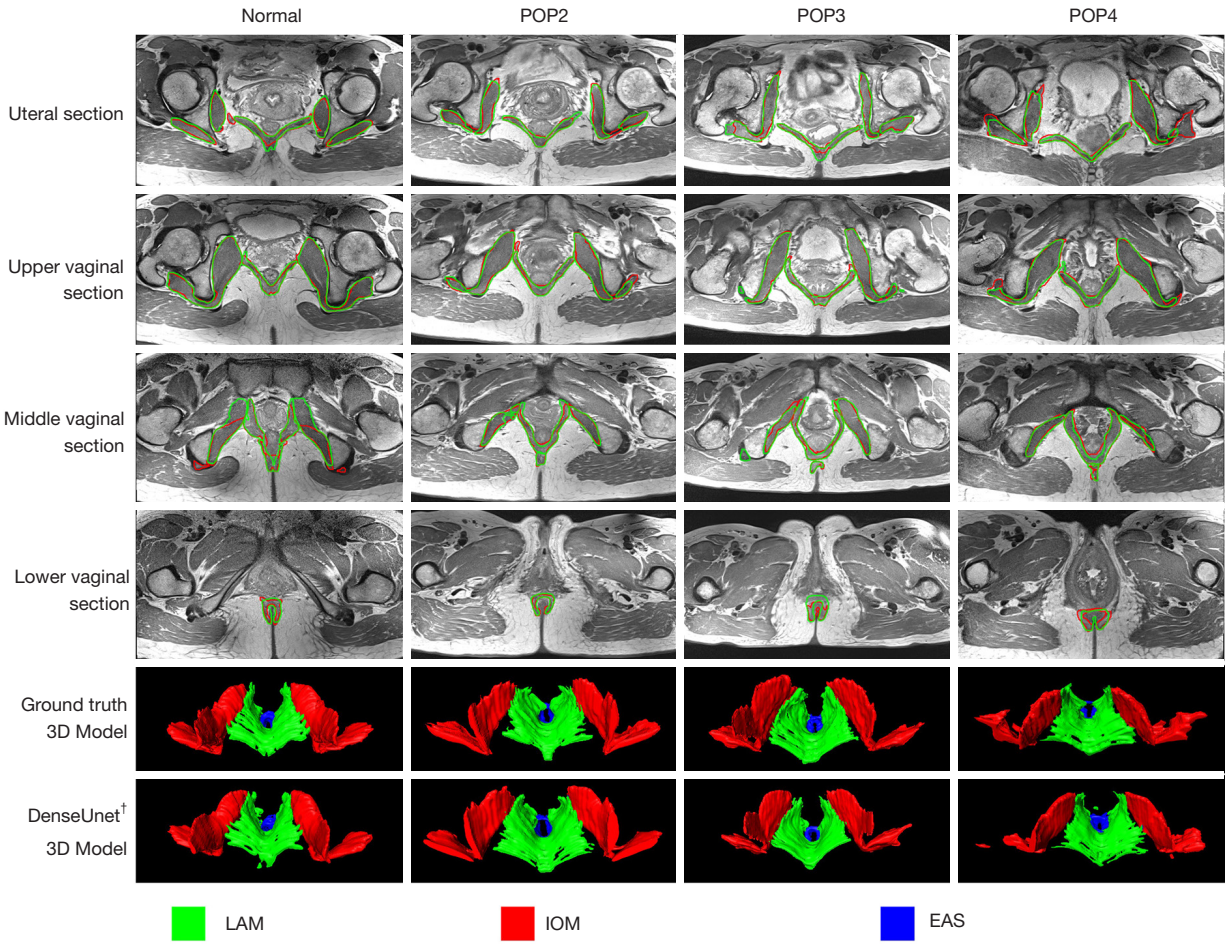

Figure 5 Qualitative 2D segmentation and 3D reconstruction results of the 3 structures on different sections in the normal and POP groups with different stages. The columns show different the normal and POP groups. Rows 1–4 show the segmentation results of different sections, where the red contour represents manual segmentation (ground truth), and the green contour represents automatic segmentation. Rows 5 and 6 represent the 3D models reconstructed by segmentation results of the ground truth and the improved DenseUnet, respectively. †, the improved DenseUnet. POP2, patients with stage 2 pelvic organ prolapse; POP3, patients with stage 3 pelvic organ prolapse; POP4, patients with stage 4 pelvic organ prolapse; 3D, 3-dimensional; LAM, levator ani muscle; IOM, internal obturator muscle; EAS, external anal sphincter; 2D, 2-dimensional.

Table 4 Volume comparison between the manual and the improved DenseUnet segmentation

Muscle type	Volume of ground truth	Predicted volume	P value*
LAM	30,560 [5,961]	26,781 [7,188]	0.247
EAS	11,179 [4,679]	10,928 [2,571]	0.853
IOM	86,154 [27,199]	81,605 [24,474]	0.739

Values are reported in medians with interquartile ranges. Numbers in the parentheses are the interquartile ranges. The unit of volume is mm^3. *, the Mann-Whitney test was used for the P value computation. P<0.05 is considered statistically significant. LAM, levator ani muscle; EAS, external anal sphincter; IOM, internal obturator muscle.

Table 5 Mean segmentation time comparison for 1 case between manual, improved DenseUnet, and compared models

Muscle type	Manual (s)	DenseUnet† (s)	ResUnet (s)	Unet++ (s)	Unet (s)
LAM	722.5±224.37	3.32±0.31	3.08±0.28	3.66±0.44	3.76±0.32
EAS	411.20±154.22	3.21±0.34	2.99±0.32	3.50±0.49	3.64±0.34
IOM	1,532.00±469.16	3.65±0.42	3.36±0.38	4.08±0.57	4.05±0.40

All values are reported in means ± standard deviations. †, denotes the improved DenseUnet. LAM, levator ani muscle; EAS, external anal sphincter; IOM, internal obturator muscle.

from the improved DenseUnet segmentation results were similar to those reconstructed from manual segmentation. The effect of automatic segmentation on the lateral side of the IOM near the femur was poor.

There was no significant difference between the muscle volume of the improved DenseUnet and manual segmentation (*Table 4*). *Table 5* shows the average automatic segmentation time of the improved DenseUnet, ResUnet, Unet++, and Unet for the 3 pelvic floor muscles in each case of the test set. The total segmentation time of the proposed DenseUnet models for 1 case was about 10.18 s, which was much lower than the manual segmentation time of 45 minutes. The training times of the improved DenseUnet and the compared models for the 3 structures are displayed in Table S1.

Discussion

Medical significance

In this study, an improved DenseUnet was proposed for the automatic segmentation of the pelvic muscles, including the LAM, EAS, and IOM, in MR images. The segmentation accuracy for the LAM of this model was higher than that of the other 3 segmentation models. Compared to manual segmentation, the proposed model shortened the segmentation time, it saved about 44 minutes of clinical time, which is significant. To our knowledge, this is the first study to achieve segmentation of 3 pelvic floor muscles, which can provide a reference for function and injury assessment of the pelvic floor muscle and diagnosis of POP.

Segmentation performance

We found that segmentation performance is closely related to the 2D morphology and spatial relationship of the pelvic floor muscles in MRI. Overall, the upper pelvic floor's segmentation accuracy was better than that of the lower pelvic floor (*Figure 4*). This is because the morphology of the pelvic floor muscles varies greatly among different sections. The structures of the upper pelvic floor are relatively simple, and there are few adjacent structures. The LAM is confused only with the coccygeus muscle. However, the muscles in the lower section of the pelvic floor are complicated, with many adjacent structures. The LAM is closer to the IOM and is adjacent to the internal anal sphincter, perineal body, and bulbospongious muscles. The gray values of these muscles are close, so they are not easily segmented.

Moreover, the larger the muscle volume was, the higher the segmentation accuracy. For example, the DSC of IOM was higher (0.810). On the contrary, the smaller the structure was, the lower the segmentation accuracy. For example, the DSC of the EAS was only 0.716 (*Table 2*). This may be because the larger the volume is, the easier the CNN identification and segmentation, and thus the higher

the accuracy. This may also be the reason why many studies (39-41) choose to automatically segment the larger limb muscles instead of the pelvic floor muscles. The EAS is also commonly injured, and the pathology can greatly vary. The segmentation performance of the EAS could be improved with more data augmentation (i.e., zooming) and more targeted model improvements for future work.

Meanwhile, we found that the more severe the POP and pelvic floor muscle injuries were, the worse the segmentation accuracy (*Table 3*). The automatic segmentation DSC of the LAM in the normal group reached 0.767, but in the POP group, it was slightly lower at 0.740. Due to the injury, deformation, and displacement of the LAM in the POP group, the levator ani hiatus can widen (*Figure 5*), which may easily lead to prolapse of pelvic organs such as the uterus and bladder. This suggests that the injured and deformed structures cause a decrease in the accuracy of automatic segmentation. All the P values of the volume comparison between manual and automatic segmentation were greater than 0.05 (*Table 4*). This indicates that the volume of the pelvic floor muscles can be quickly quantified using automatic segmentation.

Model comparison

The improved DenseUnet achieved the best segmentation DSC for the LAM compared to the other models. This may be due to the improvement that the model included contextual information under different sensory fields in the images. Therefore, the LAM can be accurately segmented from complex pelvic floor muscles. ResUnet performed better for the segmentation of the EAS. This is perhaps because the large number of residual structures in this model effectively alleviates the problem of gradient disappearance and can train deeper networks. The EAS is small and different in shape. Deeper networks may be conducive to the segmentation of the EAS. For the IOM segmentation, which is large and has more slices, the segmentation results of each model are better. However, Unet performed relatively better in terms of DSC and best in the HD and ASSD evaluation metrics (*Table 2*). Since Unet has the lowest number of parameters and has a basic segmentation ability, it is suitable for segmenting this type of muscle. Overall, we improved the DenseUnet model specifically for the characteristics of the LAM changes, so the improved DenseUnet is more sensitive to the LAM. Moreover, from the results, it can be seen that ResUnet is more suitable for segmenting small muscles such as the EAS, while Unet is more suitable for segmenting larger muscles such as the IOM.

Related research

CNNs have been widely used in the field of image segmentation, and studies on pelvic tissue and organ segmentation have mainly focused on male pelvic organs. Wang *et al.* (42,43) and Almeida *et al.* (44) have trained deep learning models to achieve accurate segmentation of the prostate and other risk-associated organs, such as bladder and rectum, on male pelvic CT images, aiming to help doctors in planning radiation therapy for patients with prostate cancer. Liu *et al.* (45,46) used the 3D U-Net algorithm to accurately and automatically detect and segment pelvic bone metastases in patients with pelvic lymph nodes and prostate cancer based on pelvic lymph nodes on diffusion-weighted images. Lei *et al.* (47) developed an anchor-free mask CNN-based approach to simultaneously segment multiple organs including the prostate, bladder, rectum, and urethra on transrectal ultrasound images, potentially enabling the autoplanning and autoevaluation of prostate brachytherapy. To improve the speed and accuracy of radiotherapy dose prediction for organs at risk, Weston *et al.* (48) proposed a 3D U-Net variant architecture to segment the whole abdomen and pelvis into 33 unique organ and tissue structures on CT images. Other studies (49,50) have constructed CNN models to automatically segment at-risk organs such as the anal canal, bladder, small intestine, rectum, femoral head, and spinal cord in CT images and multiple MRI sequences of patients with prostate or cervical cancer.

However, CNN applications for female pelvic floor muscle segmentation and POP analysis have not been extensively studied. This may be due to the irregular anatomical morphology of the female pelvic floor muscles, especially the LAM and EAS, for which, the 2D and 3D morphology and spatial adjacency are much more complicated. There is considerable controversy regarding their anatomical division and 3D morphology. Moreover, due to fertility reasons, there are certain differences in the structure and morphology of the pelvic organs and pelvic floor muscles in different women. Therefore, it is difficult for doctors, let alone algorithms, to segment female pelvic floor muscles.

Feng *et al.* (51) presented a CNN-based solution for segmenting 4 female pelvic organ structures, including the uterus, rectum, bladder, and the LAM, from 24 sagittal

MR series. The DSC of the LAM segmentation was 0.608. The reason for this low segmentation accuracy may be due to the limited number of cases. In our study, the improved DenseUnet model achieved an average DSC of 0.758 for the segmentation of the LAM. The relatively high segmentation accuracy may be due to the effectiveness of the improved model and the larger number of normal participants and patients with POP. This further indicates that female pelvic floor muscles are particularly complicated, and it is difficult to achieve precise segmentation.

Yan *et al.* (52) proposed a new Scale- and Slice-aware Net for the 3D dense segmentation of 54 organs and musculoskeletal structures in a pelvic MRI cohort of 27 patients. The average DSC of the 54 organs and musculoskeletal structures was 0.749. However, because of the difficulty in identifying rare categories, such as the ovary, adductor magnus, and vagina, the model did not perform well on these structures. According to each doctor's experience, because of the smaller structure and worse anatomical regularity, the segmentation of the pelvic floor muscles is more difficult than that of the pelvic organs. In our study, the average DSCs of the LAM, EAS, and IOM were 0.758, 0.716, and 0.810, respectively, further supporting the advantages of our model. To our knowledge, this is the first study to report a good segmentation performance of the 3 pelvic floor muscles.

This study had several limitations. First, we only performed automatic segmentation of the LAM, EAS, and IOM. However, pelvic floor dysfunction is associated with many small pelvic floor muscles, such as the compressor urethrae and the main part of the urethral sphincter. Other important pelvic floor muscles should be considered in future research. Second, the thin sectional MR images in this study needed to be scanned with a unique and special sequence, while the number of patients with different stages varies. Thus, the cases were limited for model training and testing. More cases of different stages should be included in future studies. Finally, the proposed model should be more reliable and robust for clinical applications. The segmentation accuracy of the severe POP group was worse because the LAM is injured and deformed in severe POP. In future work, we intend to establish a multicenter training set and add MRI images of different quality to improve the compatibility and generalizability of the model.

Conclusions

We created and improved the automatic segmentation model DenseUnet based on the characteristics of pelvic floor muscles. The segmentation performance of LAM was better than that reported in previous studies on pelvic floor MRI. The average segmentation time for 1 case was about 44 minutes quicker than that of manual segmentation. The improved DenseUnet can quickly and efficiently segment the LAM, EAS, and IOM in pelvic MR images, which can assist doctors in quickly performing pelvic floor muscle identification, segmentation, 3D reconstruction, and automatic volume calculation to evaluate pelvic floor function. This model may be helpful for completing the diagnosis and reconstructive surgery planning of POP, reducing the doctors' burden, and saving medical resources.

Acknowledgments

Funding: This work was supported by the National Natural Science Foundation of China (Nos. 31971113 and 11901071), the Chongqing Science and Technology Talent Project (No. CQYC201905037), and the Graduate Education and Teaching Reform Research Project of Chongqing (No. yjg183144).

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-22-1198/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-22-1198/coif). YX is currently an employee of Chongqing Zhence Science and Technology Co., Ltd. The study was conducted when he was a postgraduate student in the School of Mathematical Sciences, Chongqing Normal University. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as

revised in 2013) and was approved by the Ethics Committee of the First Affiliated Hospital of Army Military Medical University (No. BKY2021029). Requirement for patient consent was waived due to the retrospective nature of this study. This diagnostic study was not registered on any clinical trial platform.

References

1. Smith FJ, Holman CD, Moorin RE, Tsokos N. Lifetime risk of undergoing surgery for pelvic organ prolapse. Obstet Gynecol 2010;116:1096-100.
2. Hagen S, Stark D, Glazener C, Dickson S, Barry S, Elders A, Frawley H, Galea MP, Logan J, McDonald A, McPherson G, Moore KH, Norrie J, Walker A, Wilson D; . Individualised pelvic floor muscle training in women with pelvic organ prolapse (POPPY): a multicentre andomized controlled trial. Lancet 2014;383:796-806.
3. Giarenis I, Robinson D. Prevention and management of pelvic organ prolapse. F1000Prime Rep 2014;6:77.
4. Wu Y, Dabhoiwala NF, Hagoort J, Hikspoors JPJM, Tan LW, Mommen G, Hu X, Zhang SX, Lamers WH. Architecture of structures in the urogenital triangle of young adult males; comparison with females. J Anat 2018;233:447-59.
5. Wu Y, Dabhoiwala NF, Hagoort J, Shan JL, Tan LW, Fang BJ, Zhang SX, Lamers WH. 3D Topography of the Young Adult Anal Sphincter Complex Reconstructed from Undeformed Serial Anatomical Sections. PLoS One 2015;10:e0132226.
6. DeLancey JO, Morgan DM, Fenner DE, Kearney R, Guire K, Miller JM, Hussain H, Umek W, Hsu Y, Ashton-Miller JA. Comparison of levator ani muscle defects and function in women with and without pelvic organ prolapse. Obstet Gynecol 2007;109:295-302.
7. Nagle AS, Barker MA, Kleeman SD, Haridas B, Mast TD. Passive biomechanical properties of human cadaveric levator ani muscle at low strains. J Biomech 2014;47:583-6.
8. Hoyte L, Ye W, Brubaker L, Fielding JR, Lockhart ME, Heilbrun ME, Brown MB, Warfield SK; . Segmentations of MRI images of the female pelvic floor: a study of inter- and intra-reader reliability. J Magn Reson Imaging 2011;33:684-91.
9. Ma Z, Jorge RN, Tavares JM. A shape guided C-V model to segment the levator ani muscle in axial magnetic resonance images. Med Eng Phys 2010;32:766-74.
10. Ma Z, Jorge RN, Mascarenhas T, Tavares JM. Segmentation of female pelvic cavity in axial T2-weighted MR images towards the 3D reconstruction. Int J Numer Method Biomed Eng 2012;28:714-26.
11. Zhang L, Le Lu, Nogues I, Summers RM, Liu S, Yao J. DeepPap: Deep Convolutional Networks for Cervical Cell Classification. IEEE J Biomed Health Inform 2017;21:1633-43.
12. Vigneault DM, Xie W, Ho CY, Bluemke DA, Noble JA. Ω-Net (Omega-Net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks. Med Image Anal 2018;48:95-106.
13. Li X, Chen H, Qi X, Dou Q, Fu CW, Heng PA. H-DenseUNet: Hybrid Densely Connected Unet for Liver and Tumor Segmentation From CT Volumes. IEEE Trans Med Imaging 2018;37:2663-74.
14. Hu D, Jian J, Li Y, Gao X. Deep learning-based segmentation of epithelial ovarian cancer on T2-weighted magnetic resonance images. Quant Imaging Med Surg 2023;13:1464-77.
15. Shen H, He P, Ren Y, Huang Z, Li S, Wang G, Cong M, Luo D, Shao D, Lee EY, Cui R, Huo L, Qin J, Liu J, Hu Z, Liu Z, Zhang N. A deep learning model based on the attention mechanism for automatic segmentation of abdominal muscle and fat for body composition assessment. Quant Imaging Med Surg 2023;13:1384-98.
16. Noort F, Sirmacek B, Slump CH. Recurrent U-net for automatic pelvic floor muscle segmentation on 3D ultrasound. Available online: https://doi.org/10.48550/arXiv.2107.13833
17. Williams H, Cattani L, Vercauteren T, Deprest J, D'hooge J. Automatic Tomographic Ultrasound Imaging Sequence Extraction of the Anal Sphincter. Simplifying Medical Ultrasound. ASMUS 2021. New York, NY, USA: Springer International Publishing, 2021:35-44.
18. Williams H, Cattani L, Van Schoubroeck D, Yaqub M, Sudre C, Vercauteren T, D'Hooge J, Deprest J. Automatic Extraction of Hiatal Dimensions in 3-D Transperineal Pelvic Ultrasound Recordings. Ultrasound Med Biol 2021;47:3470-9.

19. Sindhwani N, Barbosa D, Alessandrini M, Heyde B, Dietz HP, D'Hooge J, Deprest J. Semi-automatic outlining of levator hiatus. Ultrasound Obstet Gynecol 2016;48:98-105.
20. Williams H, Cattani L, Yaqub M, Sudre C, Vercauteren T, Deprest J, D'hooge J. Automatic C-Plane Detection in Pelvic Floor Transperineal Volumetric Ultrasound. Medical Ultrasound, and Preterm, Perinatal and Paediatric Image Analysis. ASMUS PIPPI 2020. New York, NY, USA: Springer International Publishing, 2020:136-45.
21. Bonmati E, Hu Y, Sindhwani N, Dietz HP, D'hooge J, Barratt D, Deprest J, Vercauteren T. Automatic segmentation method of pelvic floor levator hiatus in ultrasound using a self-normalizing neural network. J Med Imaging (Bellingham) 2018;5:021206.
22. van den Noort F, van der Vaart CH, Grob ATM, van de Waarsenburg MK, Slump CH, van Stralen M. Deep learning enables automatic quantitative assessment of puborectalis muscle and urogenital hiatus in plane of minimal hiatal dimensions. Ultrasound Obstet Gynecol 2019;54:270-5.
23. van den Noort F, Manzini C, van der Vaart CH, van Limbeek MAJ, Slump CH, Grob ATM. Automatic identification and segmentation of slice of minimal hiatal dimensions in transperineal ultrasound volumes. Ultrasound Obstet Gynecol 2022;60:570-6.
24. He K, Cao X, Shi Y, Nie D, Gao Y, Shen D. Pelvic Organ Segmentation Using Distinctive Curve Guided Fully Convolutional Networks. IEEE Trans Med Imaging 2019;38:585-95.
25. Zavala Bojorquez JA, Jodoin PM, Bricq S, Walker PM, Brunotte F, Lalande A. Automatic classification of tissues on pelvic MRI based on relaxation times and support vector machine. PLoS One 2019;14:e0211944.
26. Hemke R, Buckless CG, Tsao A, Wang B, Torriani M. Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment. Skeletal Radiol 2020;49:387-95.
27. Khwaja R, Dessouky R, Heffler MA, Xi Y, Neeland IJ, Chhabra A. Multi-compartment mesenchymal tissue segmentation in pelvic MRI examinations of women: Anthropomorphic and clinical correlations. Eur J Radiol 2019;112:37-43.
28. Zhang SX, Heng PA, Liu ZJ, Tan LW, Qiu MG, Li QY, et al. The Chinese Visible Human (CVH) datasets incorporate technical and imaging advances on earlier digital humans. J Anat 2004;204:165-73.
29. Wu Y, Dabhoiwala NF, Hagoort J, Tan LW, Zhang SX, Lamers WH. Architectural differences in the anterior and middle compartments of the pelvic floor of young-adult and postmenopausal females. J Anat 2017;230:651-63.
30. Bump RC, Mattiasson A, Bø K, Brubaker LP, DeLancey JO, Klarskov P, Shull BL, Smith AR. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. Am J Obstet Gynecol 1996;175:10-7.
31. Jégou S, Drozdzal M, Vazquez D, Romero A, Bengio Y. The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Honolulu, HI, USA: IEEE, 2017:1175-83.
32. Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, Zhang T, Gao S, Liu J. CE-Net: Context Encoder Network for 2D Medical Image Segmentation. IEEE Trans Med Imaging 2019;38:2281-92.
33. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. 2015 IEEE International Conference on Computer Vision (ICCV). IEEE Computer Society, 2015:1026-34.
34. Sørensen TA. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Biol Skr 1948;5:1-34.
35. Fechter T, Adebahr S, Baltas D, Ben Ayed I, Desrosiers C, Dolz J. Esophagus segmentation in CT via 3D fully convolutional neural network and random walk. Med Phys 2017;44:6341-52.
36. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Cham: Springer, 2015:234-41.
37. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. Unet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. IEEE Trans Med Imaging 2020;39:1856-67.
38. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016:770-8.
39. Ni R, Meyer CH, Blemker SS, Hart JM, Feng X. Automatic segmentation of all lower limb muscles from high-resolution magnetic resonance imaging using a cascaded three-dimensional deep convolutional neural network. J Med Imaging (Bellingham) 2019;6:044009.
40. Mesbah S, Shalaby AM, Stills S, Soliman AM, Willhite A, Harkema SJ, Rejc E, El-Baz AS. Novel stochastic

framework for automatic segmentation of human thigh MRI volumes and its applications in spinal cord injured individuals. PLoS One 2019;14:e0216487.
41. Hiasa Y, Otake Y, Takao M, Ogawa T, Sugano N, Sato Y. Automated Muscle Segmentation from Clinical CT Using Bayesian U-Net for Personalized Musculoskeletal Modeling. IEEE Trans Med Imaging 2020;39:1030-40.
42. Wang S, He K, Nie D, Zhou S, Gao Y, Shen D. CT male pelvic organ segmentation using fully convolutional networks with boundary sensitive representation. Med Image Anal 2019;54:168-78.
43. Wang S, Nie D, Qu L, Shao Y, Lian J, Wang Q, Shen D. CT Male Pelvic Organ Segmentation via Hybrid Loss Network With Incomplete Annotation. IEEE Trans Med Imaging 2020;39:2151-62.
44. Almeida G, Figueira AR, Lencart J, Tavares JMRS. Segmentation of male pelvic organs on computed tomography with a deep neural network fine-tuned by a level-set method. Comput Biol Med 2021. [Epub ahead of print]. doi: 10.1016/j.compbiomed.2021.105107.
45. Liu X, Han C, Cui Y, Xie T, Zhang X, Wang X. Detection and Segmentation of Pelvic Bones Metastases in MRI Images for Patients With Prostate Cancer Based on Deep Learning. Front Oncol 2021;11:773299.
46. Liu X, Sun Z, Han C, Cui Y, Huang J, Wang X, Zhang X, Wang X. Development and validation of the 3D U-Net algorithm for segmentation of pelvic lymph nodes on diffusion-weighted images. BMC Med Imaging 2021;21:170.
47. Lei Y, Wang T, Roper J, Jani AB, Patel SA, Curran WJ, Patel P, Liu T, Yang X. Male pelvic multi-organ segmentation on transrectal ultrasound using anchor-free mask CNN. Med Phys 2021;48:3055-64.
48. Weston AD, Korfiatis P, Philbrick KA, Conte GM, Kostandy P, Sakinis T, Zeinoddini A, Boonrod A, Moynagh M, Takahashi N, Erickson BJ. Complete abdomen and pelvis segmentation using U-net variant architecture. Med Phys 2020;47:5609-18.
49. Huang S, Cheng Z, Lai L, Zheng W, He M, Li J, Zeng T, Huang X, Yang X. Integrating multiple MRI sequences for pelvic organs segmentation via the attention mechanism. Med Phys 2021;48:7930-45.
50. Ju Z, Wu Q, Yang W, Gu S, Guo W, Wang J, Ge R, Quan H, Liu J, Qu B. Automatic segmentation of pelvic organs-at-risk using a fusion network model based on limited training samples. Acta Oncol 2020;59:933-9.
51. Feng F, Ashton-Miller JA, DeLancey JOL, Luo J. Convolutional neural network-based pelvic floor structure segmentation using magnetic resonance imaging in pelvic organ prolapse. Med Phys 2020;47:4281-93.
52. Yan C, Lu JJ, Chen K, Wang L, Lu H, Yu L, Sun M, Xu J. Scale- and Slice-aware Net (S(2) aNet) for 3D segmentation of organs and musculoskeletal structures in pelvic MRI. Magn Reson Med 2022;87:431-45.

Cite this article as: Zhang X, Xiang Y, Yao J, Hu X, Wang Y, Liu L, Wang Y, Wu Y. Automatic segmentation of the female pelvic floor muscles on MRI for pelvic floor function assessment. Quant Imaging Med Surg 2023;13(7):4181-4195. doi: 10.21037/qims-22-1198

Appendix 1

Codes for the used network

```
import numpy as np
import os
import skimage.io as io
import skimage.transform as trans
from scipy import ndimage
from tensorflow.keras.models import *
from tensorflow.keras.layers import *
from tensorflow.keras.optimizers import *
from tensorflow.keras.callbacks import ModelCheckpoint, LearningRateScheduler
from tensorflow.keras import backend as keras
from tensorflow.keras import backend as K
import tensorflow as tf
from tensorflow.keras.applications import *

smooth = 1
def dice_coef(y_true, y_pred):
  y_true_f = K.flatten(y_true)
  y_pred_f = K.flatten(y_pred)
  intersection = K.sum(y_true_f * y_pred_f)
  return (2. * intersection + smooth) / (K.sum(y_true_f * y_true_f) + K.sum(y_pred_f * y_pred_f) + smooth)

def dice_coef_loss(y_true, y_pred):
  return 1. - dice_coef(y_true, y_pred)

from tensorflow.keras import regularizers

def DenseUnet(pretrained_weights = None,input_size = (256,256,1)):
  inputs = Input(input_size)

  conv1_1 = Conv2D(64, 3, activation='relu', padding='same', kernel_initializer='he_normal')(inputs)
  BatchNorm1_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
            beta_regularizer=regularizers.l2(1e-4))(conv1_1)
  ReLU1_1 = Activation('relu')(BatchNorm1_1)
  conv1_2 = Conv2D(64, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU1_1)
  drop1_2 = Dropout(0)(conv1_2)#
  # Merge1 = merge([conv1_1,drop1_2], mode = 'concat', concat_axis = 3)
  Merge1 = Concatenate(axis=3)([conv1_1, drop1_2])

  pool1 = MaxPooling2D(pool_size=(2, 2))(Merge1)

  conv2_1 = Conv2D(128, 3, activation='relu', padding='same', kernel_initializer='he_normal')(pool1)
  BatchNorm2_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
            beta_regularizer=regularizers.l2(1e-4))(conv2_1)
  ReLU2_1 = Activation('relu')(BatchNorm2_1)
  conv2_2 = Conv2D(128, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU2_1)
```


https://dx.doi.org/10.21037/qims-22-1198

```
drop2_2 = Dropout(0)(conv2_2)#
# Merge2 = merge([conv2_1,drop2_2], mode = 'concat', concat_axis = 3)
Merge2 = Concatenate(axis=3)([conv2_1, drop2_2])
pool2 = MaxPooling2D(pool_size=(2, 2))(Merge2)
conv3_1 = Conv2D(256, 3, activation='relu', padding='same', kernel_initializer='he_normal')(pool2)
BatchNorm3_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
            beta_regularizer=regularizers.l2(1e-4))(conv3_1)
ReLU3_1 = Activation('relu')(BatchNorm3_1)
conv3_2 = Conv2D(256, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU3_1)
drop3_2 = Dropout(0)(conv3_2)#
# Merge3 = merge([conv3_1,drop3_2], mode = 'concat', concat_axis = 3)
Merge3 = Concatenate(axis=3)([conv3_1, drop3_2])

pool3 = MaxPooling2D(pool_size=(2, 2))(Merge3)

conv4_1 = Conv2D(512, 3, activation='relu', padding='same', kernel_initializer='he_normal')(pool3)
BatchNorm4_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
            beta_regularizer=regularizers.l2(1e-4))(conv4_1)
ReLU4_1 = Activation('relu')(BatchNorm4_1)
conv4_2 = Conv2D(512, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU4_1)
drop4_2 = Dropout(0)(conv4_2)#
# Merge4 = merge([conv4_1,drop4_2], mode = 'concat', concat_axis = 3)
Merge4 = Concatenate(axis=3)([conv4_1, drop4_2])
drop4 = Dropout(0.5)(Merge4)
pool4 = MaxPooling2D(pool_size=(2, 2))(drop4)

# conv5_1 = Conv2D(1024, 3, activation='relu', padding='same', kernel_initializer='he_normal')(pool4)
# BatchNorm5_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
#              beta_regularizer=regularizers.l2(1e-4))(conv5_1)
# ReLU5_1 = Activation('relu')(BatchNorm5_1)
# conv5_2 = Conv2D(1024, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU5_1)
# drop5_2 = Dropout(0)(conv5_2)#
# # Merge5 = merge([conv5_1,drop5_2], mode = 'concat', concat_axis = 3)
# Merge5 = Concatenate(axis=3)([conv5_2, drop5_2])
# drop5 = Dropout(0.5)(Merge5)
# DAC
branch1 = Conv2D(512, 3, activation='relu', padding= 'same', dilation_rate=1)(pool4)
branch2 = Conv2D(512, 3, activation='relu', padding='same', dilation_rate=3)(pool4)
branch2 = Conv2D(512, 1, activation='relu', padding='same', dilation_rate=1)(branch2)

branch3 = Conv2D(512, 3, activation='relu', padding='same', dilation_rate=1)(pool4)
branch3 = Conv2D(512, 3, activation='relu', padding='same', dilation_rate=3)(branch3)
branch3 = Conv2D(512, 1, activation='relu', padding='same', dilation_rate=1)(branch3)

branch4 = Conv2D(512, 3, activation='relu', padding='same', dilation_rate=1)(pool4)
branch4 = Conv2D(512, 3, activation='relu', padding='same', dilation_rate=3)(branch4)
branch4 = Conv2D(512, 3, activation='relu', padding='same', dilation_rate=5)(branch4)
branch4 = Conv2D(512, 1, activation='relu', padding='same', dilation_rate=1)(branch4)
```


https://dx.doi.org/10.21037/qims-22-1198

```
# pool4 = Concatenate(axis=3)([branch1, branch2, branch3, branch4, pool4])
pool4 = branch1 + branch2 + branch3 + branch4

# RMP

up6 = Conv2D(512, 2, activation='relu', padding='same', kernel_initializer='he_normal')(
  UpSampling2D(size=(2, 2))(pool4))
# merge6 = merge([drop4,up6], mode = 'concat', concat_axis = 3)
# drop4_at = cbam_block(drop4, name = "4")
merge6 = Concatenate(axis=3)([drop4, up6])

conv6_1 = Conv2D(512, 3, activation='relu', padding='same', kernel_initializer='he_normal')(merge6)
BatchNorm6_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
            beta_regularizer=regularizers.l2(1e-4))(conv6_1)
ReLU6_1 = Activation('relu')(BatchNorm6_1)
conv6_2 = Conv2D(512, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU6_1)
drop6_2 = Dropout(0)(conv6_2)#

# Merge6 = merge([conv6_1,drop6_2], mode = 'concat', concat_axis = 3)
Merge6 = Concatenate(axis=3)([conv6_1, drop6_2])

up7 = Conv2D(256, 2, activation='relu', padding='same', kernel_initializer='he_normal')(
  UpSampling2D(size=(2, 2))(Merge6))
# merge7 = merge([Merge3,up7], mode = 'concat', concat_axis = 3)
# Merge3_at = cbam_block(Merge3, name = '3')
merge7 = Concatenate(axis=3)([Merge3, up7])

conv7_1 = Conv2D(256, 3, activation='relu', padding='same', kernel_initializer='he_normal')(merge7)
BatchNorm7_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
            beta_regularizer=regularizers.l2(1e-4))(conv7_1)
ReLU7_1 = Activation('relu')(BatchNorm7_1)
conv7_2 = Conv2D(256, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU7_1)
drop7_2 = Dropout(0)(conv7_2)#
# Merge7 = merge([conv7_1,drop7_2], mode = 'concat', concat_axis = 3)
Merge7 = Concatenate(axis=3)([conv7_1, drop7_2])


up8 = Conv2D(128, 2, activation='relu', padding='same', kernel_initializer='he_normal')(
  UpSampling2D(size=(2, 2))(Merge7))
# merge8 = merge([Merge2,up8], mode = 'concat', concat_axis = 3)
# Merge2_at = cbam_block(Merge2, name = '2')
merge8 = Concatenate(axis=3)([Merge2, up8])

conv8_1 = Conv2D(128, 3, activation='relu', padding='same', kernel_initializer='he_normal')(merge8)
BatchNorm8_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
            beta_regularizer=regularizers.l2(1e-4))(conv8_1)
ReLU8_1 = Activation('relu')(BatchNorm8_1)
conv8_2 = Conv2D(128, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU8_1)
```


https://dx.doi.org/10.21037/qims-22-1198

```
drop8_2 = Dropout(0)(conv8_2)#
# Merge8 = merge([conv8_1,drop8_2], mode = 'concat', concat_axis = 3)
Merge8 = Concatenate(axis=3)([conv8_1, drop8_2])

up9 = Conv2D(64, 2, activation='relu', padding='same', kernel_initializer='he_normal')(
  UpSampling2D(size=(2, 2))(Merge8))
# merge9 = merge([Merge1,up9], mode = 'concat', concat_axis = 3)
# Merge1_at = cbam_block(Merge1, name = '1')
merge9 = Concatenate(axis=3)([Merge1, up9])

conv9_1 = Conv2D(64, 3, activation='relu', padding='same', kernel_initializer='he_normal')(merge9)
BatchNorm9_1 = BatchNormalization(axis=3, gamma_regularizer=regularizers.l2(1e-4),
            beta_regularizer=regularizers.l2(1e-4))(conv9_1)
ReLU9_1 = Activation('relu')(BatchNorm9_1)
conv9_2 = Conv2D(64, 3, activation='relu', padding='same', kernel_initializer='he_normal')(ReLU9_1)
drop9_2 = Dropout(0)(conv9_2)#
# Merge9 = merge([conv9_1,drop9_2], mode = 'concat', concat_axis = 3)
Merge9 = Concatenate(axis=3)([conv9_1, drop9_2])

conv9 = Conv2D(2, 3, activation='relu', padding='same', kernel_initializer='he_normal')(Merge9)
conv10 = Conv2D(1, 1, activation='sigmoid')(conv9) # sigmoid
# conv10 = Conv2D(1, 1, activation = 'softmax')(conv9)#sigmoid

model = Model(inputs, conv10)
model.compile(optimizer=Adam(lr=3e-4), loss=dice_coef_loss, metrics=[dice_coef, 'acc'])

model.summary()

if (pretrained_weights):
  model.load_weights(pretrained_weights)

return model
```


https://dx.doi.org/10.21037/qims-22-1198

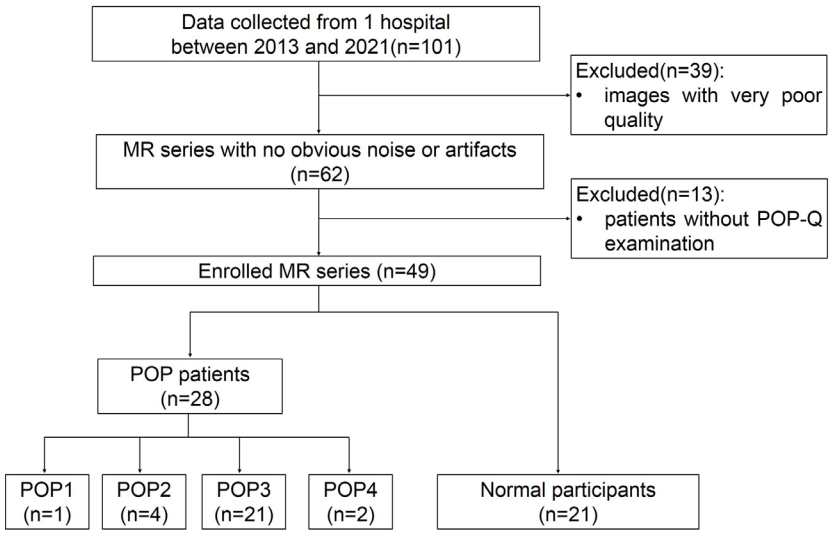

Figure S1 Flowchart of data collection. MR, magnetic resonance; POP, pelvic organ prolapse; POP-Q, pelvic organ prolapse quantification; POP 1, patients with stage 1 pelvic organ prolapse; POP2, patients with stage 2 pelvic organ prolapse; POP3, patients with stage 3 pelvic organ prolapse; POP4, patients with stage 4 pelvic organ prolapse.

Table S1 Training time of the improved DenseUnet and compared models for 3 structures

Models	LAM	EAS	IOM
Unet	0:26'54"	0:19'6"	0:31'4"
Unet++	0:56'17"	0:40'33"	1:5'3"
ResUnet	1:45'13"	1:16'43"	2:1'42"
DenseUnet	1:51'24"	0:24'32"	0:45'2"
DenseUnet†	2:26'4"	0:38'26"	1:0'39"

†, denotes the improved DenseUet. All values are described as hour:minute'second. LAM, levator ani muscle; EAS, external anal sphincter; IOM, internal obturator muscle.

https://dx.doi.org/10.21037/qims-22-1198