Automated Robust Muscle Segmentation in Multi-level Contexts using a Probabilistic Inference Framework

Jinge Wang[1], Guilin Chen[2], Xuefeng Wang[1], Nan Wu[2], and Terry Jianguo Zhang[2]

[1] Human-Oriented Methodology and Equipment (HOME) Laboratory, College of Engineering, Peking University, Beijing 100871, China
wang_xf@pku.edu.cn

[2] Laboratory of Big Data for Spinal Deformities, Beijing Key Laboratory for Genetic Research of Skeletal, State Key Laboratory of Complex Severe and Rare Diseases, Department of Orthopedic Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

Abstract. The paraspinal muscles are crucial for spinal stability, which can be quantitatively analyzed through image segmentation. However, unclear muscle boundaries, severe deformations, and limited training data impose great challenges for existing automatic segmentation methods. This study proposes an automated probabilistic inference framework to reconstruct 3D muscle shapes from thick-slice MRI robustly. Leveraging Fourier basis functions and Gaussian processes, we construct anatomically interpretable shape models. Multi-level contextual observations such as global poses of muscle centroids and local edges are then integrated into posterior estimation to enhance shape model initialization and optimization. The proposed framework is characterized by its intuitive representations and smooth generation capabilities, demonstrating higher accuracy in validation on both public and clinical datasets compared to state-of-the-art methods. The outcomes can aid clinicians and researchers in understanding muscle changes in various conditions, potentially enhancing diagnoses and treatments.

Keywords: Paraspinal muscle · Anatomically interpretable modeling · Bayesian shape inference · MRI.

1 Introduction

Low back pain (LBP) is currently the leading cause of global disability [1]. However, diagnoses and estimations of LBP remain challenging for clinicians due to the low specificity of imaging [2]. LBP is often accompanied by degenerative changes and stability decreases in paraspinal muscles, such as a decrease in cross-sectional area and an increase in fatty infiltration [3]. To better assist in diagnosing and treating LBP, accurate segmentation and 3D reconstruction of

Table 1. Existing paraspinal muscle representations for segmentation.

Method	Modality	Representation	Contextual level		
			Spatial location	Shape	Appearance feature
Engstrom (2011) [8]	MRI	Point-based 3D SSM		✓	
Kamiya (2018) [9]	CT	Landmarks, centerline, and shape model	✓	✓	
Xiao (2018) [10]	MRI	Probabilistic atlas	✓	✓	
Xia (2019) [4]	MRI	CRF-RNN and gradient magnitude map	✓		✓
Huang (2020) [11]	MRI	2D bounding box	✓		
Kawamoto (2023) [12]	CT	Correlation with erector spinae muscle	✓		
Proposed	MRI	Pose correlation, shape model, and edge	✓	✓	✓

muscles from thick-layer lumbar magnetic resonance (MR) images is a crucial step.

Clinical studies on paraspinal muscles are generally conducted through manual annotation [4]. Recently, influenced by the development of deep learning, methods for automatic segmentation of paraspinal muscles have continuously emerged [5,6]. However, due to less distinct edges, irregular shapes, and substantial inter-individual variability, the segmentation of muscles becomes even more challenging compared to bones. Deep-learning-based methods are typically data-driven, which can be problematic due to the cost of high-quality annotations and the scarcity of data, especially for rare pathologies [7]. Some studies have attempted to reduce computational costs and obtain more robust segmentation results by incorporating prior knowledge, as shown in Table 1. As early as 2011, researchers introduced statistical shape models (SSMs) to model the shape of the quadratus lumborum [8], followed by various shape-prior methods based on computational anatomy [9], and probabilistic atlases [10]. The utilization of spatial locations [11,12] and superficial appearance information [4] also enhances segmentation stability. However, most prior-based automatic segmentation methods for paraspinal muscles fail to provide an interpretable muscle representation with contextual completeness; instead, they mainly focus on geometric details. Additionally, global correlations and confidences are often overlooked. This could raise concerns regarding ethical and legal requirements in clinical uses [13]. In other segmentation tasks, studies have attempted to achieve active interpretability by enforcing extracted variables to follow specific variational posterior distributions. However, significant differences in results occur when using different prior hyperparameters [14]. Furthermore, incorporating prior knowledge into conventional segmentation methods increases model complexity and sometimes requires even more labels to represent anatomical features [15].

In this work, we propose a muscle representation method that fully utilizes anatomical contextual information, along with a Bayesian shape inference pipeline that incorporates multi-observation uncertainty, to achieve robust and efficient paraspinal muscle segmentation. This approach provides a new avenue for implementing automated muscle reconstruction and evaluation in clinical studies.

2 Muscle Representations in Multi-level Contexts

Manual observations of anatomical structures tend to follow a holistic-to-local approach [16]. To fully utilize anatomical expertise, we have established interpretable models at different levels, including pose, shape, edges, etc., within a unified framework, as illustrated in Fig. 1.

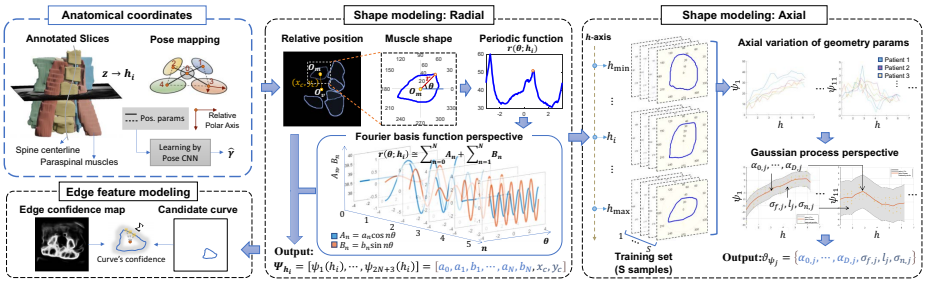

Fig. 1. Overview of the proposed muscle representations.

Pose Correlation. Pose estimation is widely used in full-body, facial, and hand-joint tracking, which has the potential for transfer to our task because paraspinal muscles exhibit strong topological relationships with the spine. We establish an anatomical coordinate system, where the z-coordinate of each slice is transformed into an axial anatomical height h related to the vertebral level. In this coordinate system, the heights of the superior endplates of L1 and S1 are defined as $h_L = 1$ and $h_S = 6$, respectively. By transforming $(h - h_S) / (h - h_L) = (z - z_S) / (z - z_L)$, where z_L and z_S are the axial Cartesian coordinates of the superior endplates of L1 and S1, respectively, h can be obtained. The centroids of each muscle and the vertebral body are regarded as key points. A pose estimation network is employed to learn the relative poses among these key points.

Probabilistic Shape Model. We represent the 3D muscle as multiple radially correlated contours stacked axially. Combining the pose correlations between muscles and vertebrae, we further establish a Fourier-Gaussian-process probabilistic shape model (FGPM). Given h, the muscle contour point $P(x, y)$ can be represented in polar coordinates, where $x = x_c(h) + r(\theta; h)\cos\theta$, and $y = y_c(h) + r(\theta; h)\sin\theta$. (x_c, y_c) represents the relative position of the muscle center, $r(\theta; h)$ represents the muscle radius, where θ is the polar angle. With the Fourier series, we get a smooth approximation of the muscle contour as

$$r(\theta; h) \approx a_0(h) + \sum_{n=1}^{N} \left(a_n(h)\cos n\theta + b_n(h)\sin n\theta\right), \tag{1}$$

where n is the Fourier order, a_0, a_n, and b_n are the coefficients of the basis functions. and a higher N reflects more detailed shape characteristics. The geometric parameter vector $\boldsymbol{\Psi}_h = [\psi_1(h), \cdots, \psi_{2N+3}(h)] = [a_0, a_1, b_1, \cdots, a_N, b_N, x_c, y_c]$, where $\psi_j(h)$ is the jth parameter in $\boldsymbol{\Psi}_h$.

To capture both axial correlation and variability in the cross-sectional shape of muscles, we model the variations in muscle geometry along h using $2N+3$ independent Gaussian processes (GPs) that encompass all parameters of $\boldsymbol{\Psi}_h$. The axial representation of $\psi_j(h)$ is given by

$$\psi_j(h) \sim \mathcal{GP}\left(m_j(h), k_j\left(h, h'\right)\right), \tag{2}$$

where $m_j(h)$ is the mean function determining the overall trend of the GP, and $k_j(h, h')$ is the kernel function representing the covariance at heights h and h'. For simplicity and flexibility, $m_j(h) = \sum_{d=0}^{D} \alpha_{d,j} h^d$, where $\alpha_{d,j}$ are the coefficients of the mean function, and D denotes the highest polynomial degree. The Matérn kernel, chosen for its adaptability and computational efficiency, serves as the primary kernel function. Additionally, a dynamic noise kernel is designed considering that different observations have varying confidences. The final kernel function is defined as

$$k_j(h, h') = \sigma_{f,j}^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} \parallel \Delta h \parallel}{l_j} \right)^{\nu} K_\nu \left(\frac{\sqrt{2\nu} \parallel \Delta h \parallel}{l_j} \right) + \delta(h, h') w(h) \sigma_{n,j}^2, \tag{3}$$

where $\Delta h = h - h'$, $\nu = 3/2$, $\Gamma(\nu)$ denotes the gamma function, K_ν is the modified Bessel function, and the kernel parameters $\sigma_{f,j}$, $l_j > 0$ control the smoothness of the kernel and adapt to different data scales. $\delta(h, h')$ is the Kronecker delta function, $w(h)$ is the noise weight of observation h, and $\sigma_{n,j}^2$ is the noise variance. Finally, the parameters of the mean and kernel functions form the GP parameter set $\vartheta_{\psi_j} = \{\alpha_{0,j}, \cdots, \alpha_{D,j}, \sigma_{f,j}, l_j, \sigma_{n,j}\}$.

Edge Confidence. Subtle edge and irregular texture information significantly influence muscle classification. Leveraging a lightweight neural edge detector [17], we focus on regions in images that may constitute muscle edges. By combining these regions with shape models, we provide confidence in shape model-generated edges. Let event $I_{\text{edge}}(u, v)$ denote the occurrence of pixel (u, v) as a muscle edge. The detector's output at pixel (u, v) is the probability $P_\beta(I_{\text{edge}}(u, v))$, where β represents the detector's weights. Given height h, let $L_{\boldsymbol{\Psi}_h}$ denote the point set of a closed curve generated by $\boldsymbol{\Psi}_h$. The curve confidence is defined as

$$C_{\boldsymbol{\Psi}_h} = P_{\boldsymbol{\beta}}\left(I_{\text{edge}} \mid L_{\boldsymbol{\Psi}_h}\right) = \sum_q P_{\boldsymbol{\beta}}\left(I_{\text{edge}}(u_q, v_q) \mid (u_q, v_q) \in L_{\boldsymbol{\Psi}_h}\right) / \left|L_{\boldsymbol{\Psi}_h}\right|, \tag{4}$$

where q is the index of the pixel (u_q, v_q) on the curve $L_{\boldsymbol{\Psi}_h}$, and $|L_{\boldsymbol{\Psi}_h}|$ represents the number of pixels in the set $L_{\boldsymbol{\Psi}_h}$.

3 Probabilistic Inference Framework

Based on muscle representations, we propose a probabilistic inference framework, while detections from multi-level contexts are treated as observations. The framework, as shown in Fig. 2, involves the following steps:

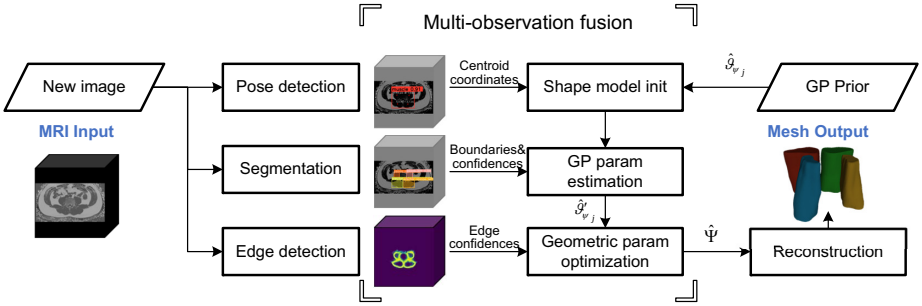

Fig. 2. Pipeline of the proposed probabilistic inference framework.

Firstly, a detection-segmentation network based on appearance information is employed to obtain initial muscle boundaries and confidences. After a confidence threshold and outlier removal based on pose detection, we initialize the shape model with valid muscle boundaries. The top and bottom slices for analysis are uniformly specified by the dataset. If no observations are available on these slices, the muscle centroid generated by pose detection at these locations serves as the origin to generate prior geometric parameters. Let $(\boldsymbol{h}^{\mathrm{o}}, \boldsymbol{\psi}_j^{\mathrm{o}})$ be noisy annotated observations and $\boldsymbol{\psi}_j^*$ be the generated geometric parameters at $\boldsymbol{h}^*$. Assuming the average confidence of the predicted observations at $\boldsymbol{h}^*$ is the same as in the historical data, the noise weights for all future points are set to the harmonic mean of the noise weights of the observation points. GP parameters are optimized, and the posterior of $\boldsymbol{\psi}_j^*$ is given by $\boldsymbol{\psi}_j^* | \boldsymbol{h}^{\mathrm{o}}, \boldsymbol{\psi}_j^{\mathrm{o}}, \boldsymbol{h}^* \sim \mathcal{N}(\bar{\boldsymbol{\psi}}_j^*, \mathrm{cov}(\boldsymbol{\psi}_j^*))$, where

$$\bar{\boldsymbol{\psi}}_j^* = \boldsymbol{\mu}_j^* + \boldsymbol{K}_j^{*,\mathrm{o}} \boldsymbol{K}_j^{\mathrm{o},\mathrm{o}-1} \left(\boldsymbol{\psi}_j^{\mathrm{o}} - \boldsymbol{\mu}_j^{\mathrm{o}} \right), \tag{5}$$

$$\mathrm{cov}(\boldsymbol{\psi}_j^*) = \boldsymbol{K}_j^{*,*} - \boldsymbol{K}_j^{*,\mathrm{o}} \boldsymbol{K}_j^{\mathrm{o},\mathrm{o}-1} \boldsymbol{K}_j^{\mathrm{o},*}, \tag{6}$$

in which $\boldsymbol{K}_j^{\mathrm{o},\mathrm{o}}$, $\boldsymbol{K}_j^{\mathrm{o},*}$, $\boldsymbol{K}_j^{*,\mathrm{o}}$, and $\boldsymbol{K}_j^{*,*}$ correspond to the covariance matrices that $\boldsymbol{K}_{j,(u,v)}^{x,y} = \hat{k}_j(\boldsymbol{h}_u^x, \boldsymbol{h}_v^y)$, and $\boldsymbol{\mu}_j^{\mathrm{o}}$ and $\boldsymbol{\mu}_j^*$ arise from the mean functions $\hat{m}_j(\boldsymbol{h}^{\mathrm{o}})$ and $\hat{m}_j(\boldsymbol{h}^*)$, respectively. Parameters of $\hat{m}_j$, $\hat{k}_j$, and $\hat{\sigma}_{n,j}$ come from the updated $\hat{\boldsymbol{\vartheta}}_j$.

Finally, the posterior distribution of the shape is merged with the edge confidence map, and the final posterior shape is inferred through Maximum A Posteriori (MAP) estimation. The Bayesian optimization objective for inferring $\boldsymbol{\Psi}_h$

of the MAP distribution $P(\boldsymbol{\Psi}_h|C_{\boldsymbol{\Psi}_h})$ is

$$\begin{aligned}\hat{\boldsymbol{\Psi}}_h &= \arg\max_{\boldsymbol{\Psi}_h} P_{\hat{\boldsymbol{\vartheta}}_*}(\boldsymbol{\Psi}_h|h)\, P_{\hat{\boldsymbol{\beta}}}(I_{\text{edge}} \mid L_{\boldsymbol{\Psi}_h})/P_{\hat{\boldsymbol{\beta}}}(I_{\text{edge}}) \\ &= \arg\min_{\boldsymbol{\Psi}_h} -\left[\ln P_{\hat{\boldsymbol{\vartheta}}_*}(\boldsymbol{\Psi}_h|h) + \ln P_{\hat{\boldsymbol{\beta}}}(I_{\text{edge}} \mid L_{\boldsymbol{\Psi}_h})\right].\end{aligned} \tag{7}$$

Compared to existing techniques, the muscle representation method established in this study offers a more intuitive expression of anatomical features with fewer parameters. It possesses strong descriptive capabilities for shape variations and enables automatic inference.

4 Results and Discussion

4.1 Datasets and Experimental Settings

To evaluate model performance, we utilized a total of 100 lumbar MRI scans (2530 slices) from both public and clinical datasets. Dataset 1 [18] comprises 54 Fat Fraction (FF) sequences from healthy individuals, where erector spinae and multifidus muscles are not distinguished. Among these, 36 were used for training, nine for validation, and nine for testing. Dataset 2, sourced from a hospital, consists of 46 patients with spinal deformities, with manually annotated left and right psoas, erector spinae, and multifidus muscles. Thirty-two were randomly selected for training, eight for validation, and six for testing, with experiments repeated six times. For FGPM, the maximum Fourier order N was set to six and the highest degree D for the GP mean function was set to two. Initial guesses of the kernel function parameters were set as $\sigma_f = 1$, $l = 0.25$, and $\sigma_n = 0.05$, and were updated in training. The Polack-Ribiere flavor of conjugate gradients is used to compute search directions, and a line search using quadratic and cubic polynomial approximations and the Wolfe-Powell stopping criteria is used together with the slope ratio method for guessing initial step sizes. We trained two YOLOv8-s [19] models for pose and segmentation observation.

We employed common evaluation metrics such as Dice Similarity Coefficient (DSC) and 95th percentile Hausdorff Distance (HD95) [20]. Comparative algorithms include the self-configured nnU-Net [21], BATFormer based on CNN and Transformer considering edge features [22], and BayeSeg modeling topology, shape, and appearance information through probabilistic graphical models [14]. Optimization methods, training epochs, random seeds, etc., for all compared methods were set based on their original papers. Additionally, we conducted ablation experiments to assess the impact of pose detection, shape modeling, probabilistic observations, and edge-based optimization in our framework on the results.

4.2 Segmentation Results

Comparison of training time and segmentation results are displayed in Table 2. The training time was calculated based on the time when the optimal epoch occurs. With minimal training time, our proposed method demonstrates a higher

Table 2. Comparison of training time and segmentation accuracy on dataset 1 and 2 (Mean±Std).

Method	Training time (h)	DSC(%) ↑		HD95(mm) ↓	
		Dataset 1	Dataset 2	Dataset 1	Dataset 2
nnU-Net (2d) [21]	3.25	87.12±2.92	85.78±3.40	14.03±9.78	21.81±18.54
nnU-Net (3d) [21]	4.90	86.67±3.00	83.34±12.05	14.81±10.08	16.34±12.98
BATFormer [22]	6.59	84.86±3.92	87.14±3.92	19.01±9.80	9.65±5.51
BayeSeg [14]	13.86	81.70±5.86	69.05±8.29	18.43±11.07	53.97±29.81
Ours - Dyn. Noise - Edge opt.	1.30	90.38±3.46	87.21±4.83	4.21±1.20	6.37±2.67
Ours - Dyn. Noise	1.96	90.81±2.95	88.34±4.18	4.01±0.94	6.17±3.02
Ours	1.97	**91.71±2.74**	**89.61±3.04**	**3.49±0.71**	**4.96±2.46**

average DSC and lower HD95, indicating excellent segmentation performance. As a baseline method, nnU-Net2d performs better overall than 3d, indicating that the key information in the dataset lies in 2D, suggesting that our approach of reflecting axial contextual information solely through stochastic processes is reasonable. While BATFormer and BayeSeg both incorporate more prior knowledge, they still exhibit muscle misclassification, leading to significant fluctuations in HD95. In comparison, our method is relatively more robust. Ablation experiment results suggest that fusion optimization based on edge information and dynamic noisy observations contribute to performance improvement.

To qualitatively compare the segmentation and reconstruction performance of different algorithms, we selected representative samples and slices from Dataset 1 and 2, as shown in Fig. 3. It's evident that the results in the last column closely match the expert manual annotations in the second column, showing high overall overlap and edge integrity. Furthermore, the reconstruction results from rows 3 and 6 exhibit smooth 3D geometric shapes close to manual annotations. In contrast, other major appearance-based methods show issues such as over-segmentation, under-segmentation, or misclassification.

4.3 Discussion and Future Work

Our work offers a comprehensive segmentation and reconstruction process, with significant advantages over common segmentation methods. Firstly, paraspinal muscles exhibit similarity in appearance, and our modeling approach fully leverages multi-level anatomical expertise to avoid misclassification errors. We represent the topological relationships between muscles as a global pose estimation problem, allowing estimation of approximate center positions even in cases of severe fat infiltration, based on the positions of other muscles. Unlike probabilistic atlases or traditional PDMs, we propose a probabilistic shape model based on Fourier basis functions and Gaussian processes, which has clear anatomical significance. For instance, Gaussian processes are correlated with vertebral IDs rather than Cartesian coordinates, and low-order Fourier terms reflect overall muscle size while high-order terms reflect local shape details. This approach re-

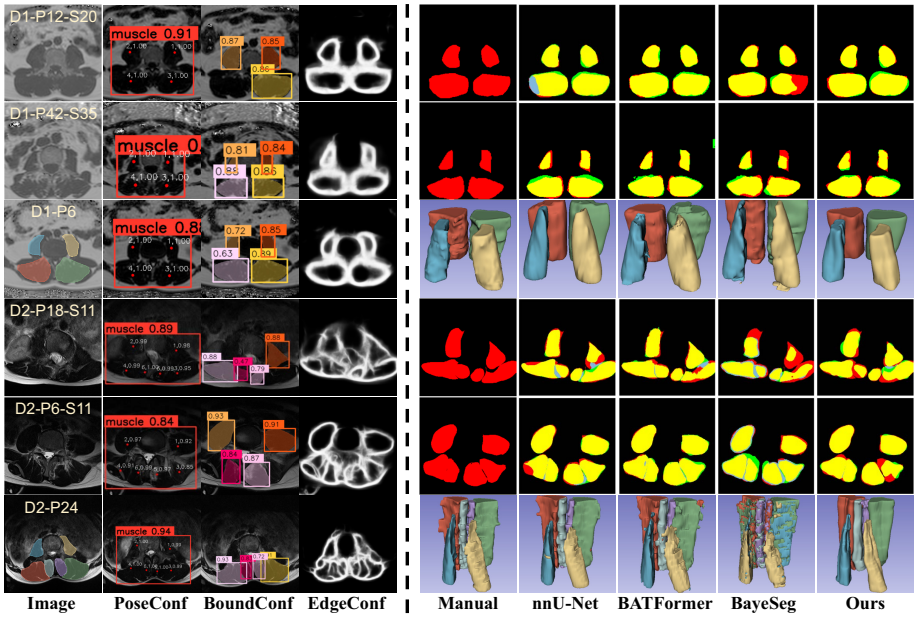

Fig. 3. Visualization comparison of segmentation and reconstruction results. Dx-Py-Sz represents the z-th slice of the y-th person in the Dataset x. The intermediate results of pose estimation, segmentation, and edge detection are shown in the second to fourth columns, respectively. In the segmentation result comparison on the right side, true positive (TP), false positive (FP), false negative (FN), and areas with incorrect prediction labels are indicated by yellow, green, red, and blue, respectively.

quires minimal parameters to represent muscles, adapts well to data, and does not require landmark alignment through registration. Additionally, our method incorporates common segmentation networks to enhance the representation of deep features but treats them only as a source of uncertain observations.

The proposed method may not be suitable for extreme scenarios, such as a few axial discontinuities or muscles whose geometric shapes are difficult to represent using polar coordinates and low-order Fourier basis functions. When accurate vertebral numbering cannot be automatically obtained due to spinal defects, manual intervention is required to determine the range for segmentation. In the future, we plan to refine this method on clinical datasets containing thousands of samples and extend it to more muscle groups. Furthermore, these smooth 3D reconstruction results can be integrated with biomechanical models and innovative devices to provide personalized guidance for robot-assisted lumbar pain rehabilitation.

5 Conclusion

The quality and quantity of muscles maintain the normal function of the musculoskeletal system. In pathological conditions, irregular shape changes and fat infiltration pose challenges for accurate segmentation of paraspinal muscles. Our work offers a comprehensive segmentation and reconstruction process. We describe a muscle from multi-level contexts and establish anatomically interpretable models of global pose, probabilistic shape, and edge confidence. After representation learning, we propose a probabilistic inference framework to handle unseen samples. Initially, we place a shape model based on pose estimation results and optimize GP parameters through segmentation observations with dynamic noise representing confidence. Next, we solve shape model parameters through MAP estimation using edge observations, yielding smooth 3D muscle reconstruction results. The method is expected to be applicable to multiple downstream clinical scenarios.

Acknowledgments. This study was funded by National Key R&D Program of China (2023YFC2507700), Beijing Nova Program, National Natural Science Foundation of China (12272003, 82072391, 82172382), National High Level Hospital Clinical Research Funding (2022-PUMCH-D-004, 2022-PUMCH-C-033), CAMS Innovation Fund for Medical Sciences (CIFMS, 2021-I2M-1-051) and the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (No.2019PT320025).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ferreira, M.L., de Luca, K., Haile, L.M., Steinmetz, J.D., Culbreth, G.T., Cross, M., Kopec, J.A., Ferreira, P.H., Blyth, F.M., Buchbinder, R., et al.: Global, regional, and national burden of low back pain, 1990–2020, its attributable risk factors, and projections to 2050: A systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol. **5**(6), e316–e329 (2023)
2. Knezevic, N.N., Candido, K.D., Vlaeyen, J.W.S., Van Zundert, J., Cohen, S.P.: Low back pain. Lancet **398**(10294), 78–92 (2021)
3. Goubert, D., Van Oosterwijck, J., Meeus, M., Danneels, L.: Structural changes of lumbar muscles in non-specific low back pain. Pain Physician **19**(7), E985–E999 (2016)
4. Xia, W., Fortin, M., Ahn, J., Rivaz, H., Battié, M.C., Peters, T.M., Xiao, Y.: Automatic paraspinal muscle segmentation in patients with lumbar pathology using deep convolutional neural network. In: Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.T., Khan, A. (eds.) MICCAI 2019. LNCS, vol. 11765, pp. 318–325. Springer, Cham (2019) https://doi.org/10.1007/978-3-030-32245-8_36
5. Roshanzamir, P., Rivaz, H., Ahn, J., Mirza, H., Naghdi, N., Anstruther, M., Battié, M.C., Fortin, M., Xiao, Y.: Joint paraspinal muscle segmentation and inter-rater labeling variability prediction with multi-task transunet. In: International Workshop on Uncertainty for Safe Utilization of Machine Learning in Medical Imaging. pp. 125–134. Springer (2022)

6. Wesselink, E.O., Elliott, J.M., Coppieters, M.W., Hancock, M.J., Cronin, B., Pool-Goudzwaard, A., Weber II, K.A.: Convolutional neural networks for the automatic segmentation of lumbar paraspinal muscles in people with low back pain. Sci. Rep. **12**(1), 13485 (2022)
7. Huang, S., Xu, T., Shen, N., Mu, F., Li, J.: Rethinking few-shot medical segmentation: a vector quantization view. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3072–3081, IEEE (2023)
8. Engstrom, C.M., Fripp, J., Jurcak, V., Walker, D.G., Salvado, O., Crozier, S.: Segmentation of the quadratus lumborum muscle using statistical shape modeling. J. Magn. Reson. Imaging **33**(6), 1422–1429 (2011)
9. Kamiya, N.: Muscle segmentation for orthopedic interventions. In: Intelligent Orthopaedics: Artificial Intelligence and Smart Image-guided Technology for Orthopaedics, pp. 81–91. Springer (2018)
10. Xiao, Y., Fortin, M., Battié, M.C., Rivaz, H.: Population-averaged MRI atlases for automated image processing and assessments of lumbar paraspinal muscles. Eur. Spine J. **27**, 2442–2448 (2018)
11. Huang, J., Shen, H., Chen, B., Wang, Y., Li, S.: Segmentation of paraspinal muscles at varied lumbar spinal levels by explicit saliency-aware learning. In: Martel, A.L., Abolmaesumi, P., Stoyanov, D., Mateus, D., Zuluaga, M.A., Zhou, S.K., Racoceanu, D., Joskowicz, L. (eds.) MICCAI 2020. LNCS, vol. 12266, pp. 652–661. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59725-2_63
12. Kawamoto, M., Kamiya, N., Zhou, X., Kato, H., Hara, T., Fujita, H.: Simultaneous learning of erector spinae muscles for automatic segmentation of site-specific skeletal muscles in body CT images. IEEE Access **12**, 15468–15476 (2024)
13. Zhang, J., Zhang, Z.M.: Ethics and governance of trustworthy medical artificial intelligence. BMC Med. Inform. Decis. Mak. **23**(1), 1–15 (2023)
14. Gao, S., Zhou, H., Gao, Y., Zhuang, X.: BayeSeg: Bayesian modeling for medical image segmentation with interpretable generalizability. Med. Image Anal. **89**, 102889 (2023)
15. Liu, L., Wolterink, J.M., Brune, C., Veldhuis, R.N.J.: Anatomy-aided deep learning for medical image segmentation: a review. Phys. Med. Biol. **66**(11), 11TR01 (2021)
16. Rokszin, A.A., Győri-Dani, D., Linnert, S., Krajcsi, A., Tompa, T., Csifcsák, G.: The interplay of holistic shape, local feature and color information in object categorization. Biol. Psychol. **109**, 120–131 (2015)
17. Soria, X., Li, Y., Rouhani, M., Sappa, A.D.: Tiny and efficient model for the edge detection generalization. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1364–1373, IEEE (2023)
18. Burian, E., Rohrmeier, A., Schlaeger, S., Dieckmeyer, M., Diefenbach, M.N., Syväri, J., Klupp, E., Weidlich, D., Zimmer, C., Rummeny, E.J., et al.: Lumbar muscle and vertebral bodies segmentation of chemical shift encoding-based water-fat MRI: the reference database MyoSegmenTUM spine. BMC Musculoskelet. Disord. **20**(1), 1–7 (2019)
19. YOLOv8: A new state-of-the-art computer vision model, https://yolov8.com/. Last accessed 7 Mar 2024
20. Wang, J., Chen, F., Ma, Y., Wang, L., Fei, Z., Shuai, J., Tang, X., Zhou, Q., Qin, J.: XBound-Former: Toward cross-scale boundary modeling in Transformers. IEEE Trans. Med. Imaging **42**(6), 1735–1745 (2023)
21. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods **18**(2), 203–211 (2021)

22. Lin, X., Yu, L., Cheng, K.T., Yan, Z.: BATFormer: Towards boundary-aware lightweight Transformer for efficient medical image segmentation. IEEE J. Biomed. Health Inform. **27**(7), 3501–3512 (2023)