Holistic Segmentation of Intermuscular Adipose Tissues on Thigh MRI

Jianhua Yao[✉], William Kovacs, Nathan Hsieh, Chia-Ying Liu,
and Ronald M. Summers

Clinical Image Processing Services, Radiology and Imaging Sciences,
Clinical Center, National Institutes of Health, Bethesda, USA
jyao@cc.nih.gov

Abstract. Muscular dystrophies (MD) cause muscles to gradually degenerate into fat. In order to effectively study and track disease progression, it is important to quantify both muscle and fat volumes, especially the intermuscular adipose tissue (IMAT). Existing methods were mostly based on unsupervised pixel clustering and morphological models. We propose a method integrating two holistic neural networks (one for edges and one for regions) and a dual active contour model to accurately locate the fascia lata and segment multiple tissue types on thigh MRIs. The proposed method is robust to image artifacts and weak boundaries, and thus it performs well for severe MD cases. Our method was tested on 104 data sets and achieved Dice coefficients 0.940 and 0.943 for muscle and IMAT in challenging severe cases, respectively.

Keywords: Muscular dystrophy · MRI · Holistic segmentation · Muscle and fat segmentation

1 Introduction

Muscular dystrophies (MD) cause muscles to gradually degenerate into adipose tissue, which can affect the patient's ability to walk and even breathe. Duchenne muscular dystrophy is the most common and severe type of MD and causes muscle weakness in the legs and upper arms [1]. In order to effectively plan the treatment and study disease progression, accurate segmentation of muscle and fat is desirable.

MRI has often been used to evaluate muscle distribution in the thighs. Figure 1 shows examples of T1-weighted thigh MRI of different MD severity levels. As shown in the examples, in severe cases, most muscle tissue has been converted to intermuscular adipose tissue (IMAT). IMAT has the same pixel intensity as subcutaneous adipose tissue (SAT). IMAT and SAT are separated by a boundary called the fascia lata, which normally forms as a stocking-like membrane around the muscles of the thighs. The fascia lata can be obscure on MRI and other fascia may interfere with its identification.

A number of investigations have been published on muscle segmentation on thigh MRI. Most of them worked on normal thighs or thighs with mild MD and were based on unsupervised pixel clustering and morphological models. Origiu et al. [2] developed an active contour model to detect the muscle boundary and a fuzzy c-means method to
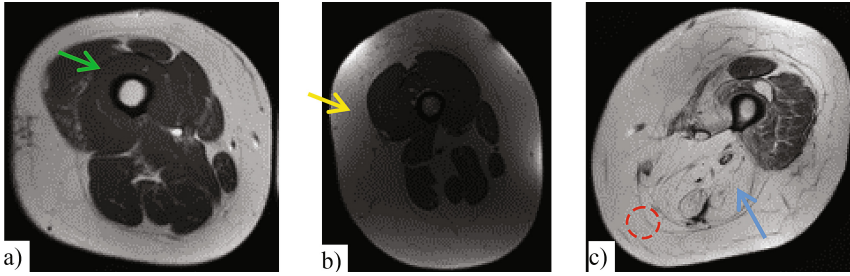
Fig. 1. Muscular dystrophy examples on thigh MRI. (a) mild; (b) moderate; and (c) severe muscular dystrophy. Green, yellow and blue arrows point to muscle, SAT and IMAT, respectively. Red circle indicates the location of fascia lata that separates muscle and SAT. Severe inhomogeneity artifact exists in (b).

distinguish muscle from fat. Similarly, Positano et al. [3] also used fuzzy clustering and edge map to differentiate muscle and IMAT. Chambers et al. [4] used livewire to segment the muscle boundary and fuzzy c-means for pixel classification. Tan et al. [5] developed a random forest classifier to detect the fascia lata near the muscle boundaries, and thus only worked for mild MD cases. Kovacs et al. [6] presented the state-of-the-art method on severe MD tissue segmentation. They developed sophisticated filters and models to detect the fascia lata, remove outliers and incorporate anatomical knowledge. Fuzzy c-means was also adopted for pixel classification.

While existing methods would work well with mild disease cases, they may struggle with moderate to severe cases. Fuzzy clustering may fail when the pixel intensity is inconsistent and inhomogeneity artifacts exist (Fig. 1b), which are known problems affecting MR images. Furthermore, fascia lata detection in prior methods relied on the remaining muscles boundaries, which may be unreliable in severe cases.

To address these challenges, we propose a framework that combines deep learning techniques and traditional deformable models. Particularly, we adopt a holistically-nested edge detection (HED) approach for fascia lata detection. The motivation is that the fascia lata does not appear as isolated edges, but instead as a group of edges forming a specific shape, which can be taken advantage of by the holistic image-to-image training and prediction. Furthermore, multi-level output from HED can be used to derive robust segmentation. Our main contributions are threefold: (1) adopt two deep learning networks to classify both edges and regions; (2) integrate deep learning outputs with traditional deformable models; and (3) propose a dual active contour model for accurate fascia lata localization.

2   Method

Our holistic segmentation framework is outlined in Fig. 2. First, a preprocessing step is implemented to mask out the thighs and separate the left and right thighs. They are then fed into a holistic edge neural net for fascia lata detection. The output edge potential map drives a dual active contour model to accurately locate the fascia lata boundaries.
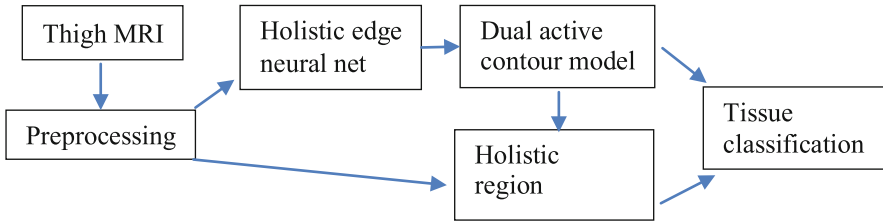
Fig. 2. System framework

A second holistic region neural net combined with the output of the fascia lata detection then classifies muscle, IMAT and SAT regions in the thighs. The algorithm is conducted on 2D slices and can be extended to 3D.

2.1 Preprocessing

In the preprocessing stage, an adaptive threshold is applied to mask thighs from the background. Then the two thighs are separated by connected component analysis. A region is cropped for each thigh for further processing. The cortical bones and bone marrows are segmented using thresholding and morphological closing. The current algorithm does not correct the inhomogeneity artifact that inherent to MRI images.

2.2 Holistic Fascia Lata Detection

As seen in Fig. 1, fascia lata boundaries can be weak. In this setting, traditional edge detectors such as Canny may not reliably detect the fascia lata without leading to over-detection. To extract semantically meaningful boundaries, an image-to-image holistic training and prediction is applied. Holistically-nested Edge Detection, HED, proposed by Xie and Tu [7], is a deep learning model that leverages fully convolutional networks and deeply supervised nets.

The network architecture is illustrated in Fig. 3, which consists of a sequence of deep network layers with side outputs at each stage. The network has 5 stages, with strides of 1, 2, 4, 8, and 16 respectively. The side outputs are fused to give the final prediction, which incorporates coherent contribution at different scales. Each side output is associated with a classifier with a class-balanced cross-entropy loss function,

$$l_{side}(W) = -\beta \sum_{j \in Y^+} logPr(y_j = 1|X) - (1 - \beta) \sum_{j \in Y^-} logPr(y_j = 0|X) \quad (1)$$

Here X is the input image, Y is the ground truth edge map, Y^+ is the set of edge pixels, Y^- is the set of non-edge pixels. W is the network layer parameters. β is the balancing weight between edge and non-edge samples, and $\beta = |Y^-|/|Y|$. Pr is the probability of a pixel belonging to a certain class.

The fused output is a weighted combination of all side outputs and its loss function is the distance between the fused prediction and the ground truth edge map. The network parameters W and combination weight h are learned by minimizing the sum of the loss function at each layer via back-propagation and stochastic gradient decent.
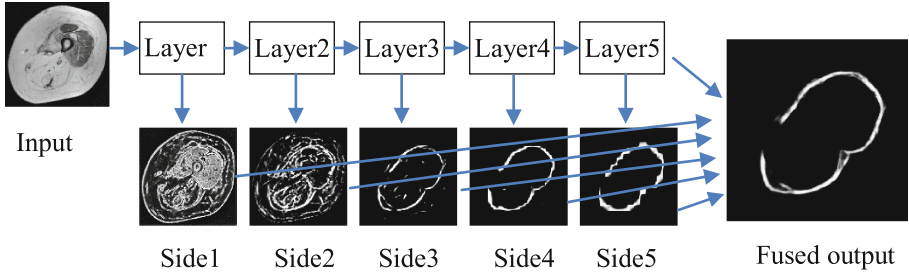
Fig. 3. Holistic fascia lata detection

Given a test image X_t, the edge map predictions can be obtained by forwarding the convolutional neural network (CNN) with fitted parameters and weights

$$\left(Y_{fused}, Y_{side1}, Y_{side2}, Y_{side3}, Y_{side4}, Y_{side5}\right) = CNN(X_T, W, h) \qquad (2)$$

Examples of side outputs and fused output are shown in Fig. 3. The fused edge map indicates the probability that a pixel belongs to the fascia lata.

The network is trained using manually labelled fascia latae. Since it is very difficult to manually locate the precise boundary, we employ a Canny edge detector to assist ground truth generation. An operator first traces a contour close to the fascia lata. Then the Canny edges closest to the manual tracing are used for the HED training. If no Canny edge exists in the neighborhood, the pixel with the maximum gradient is used. Figure 4 shows the training sample generation.
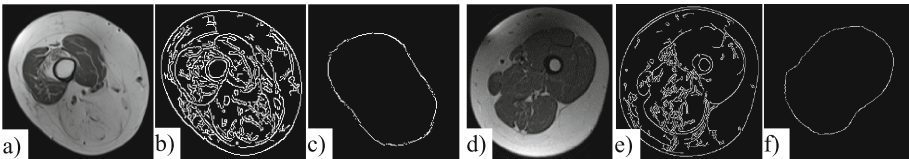

Fig. 4. Fascia lata training samples. (a,d) training images; (b,e) Canny edges; (c,f) ground truth fascia lata for training. Note that Canny edge parameters are set to over-detect the edges so that fascia lata boundaries can be included.

2.3 Dual Active Contour Model for Fascia Lata Segmentation

The holistic edge image represents the probability of a pixel being on the fascia lata. As shown in Figs. 3 and 5b, it forms a thick band around the fascia lata and some weak boundaries may be missing. In order to obtain a continuous and precise fascia lata segmentation, we employ a dual active contour model (DACM) on the holistic edge potential map.
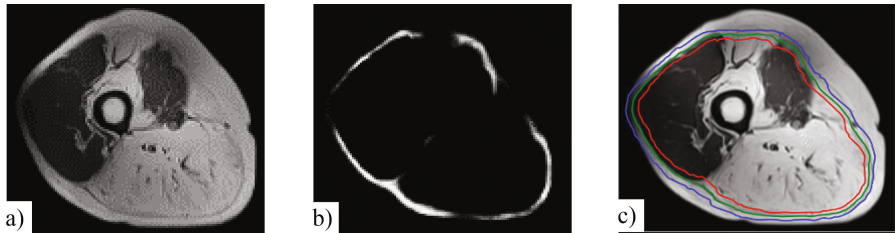
Fig. 5. Dual active contour model for fascia lata segmentation. (a) thigh MRI image; (b) edge potential image; and (c) DACM results, red: interior contour, blue: exterior contour, green: fascia lata segmentation.

The DACMs are driven by internal forces, edge potential forces, and constraints between the exterior and interior contours. The energy function is written as,

$$E(C_E) = w_i I(C_E) + w_p P^+(C_E) + w_c S(C_E, C_I)$$
$$E(C_I) = w_i I(C_I) + w_p P^-(C_I) + w_c S(C_E, C_I)$$

$$(3)$$

where C_E and C_I are exterior and interior contours, $I(C)$ is the internal spline force, $P^+(C)$ and $P^-(C)$ are positive and negative edge potential forces, $S(C_E, C_I)$ is the constraint between the two contours, and w_i, w_p and w_c are weights for the three forces. The internal forces keep the contour smooth and continuous. Directional gradient is applied to compute the edge potential force. For a point on the holistic edge image, the directional gradient is computed as the gradient along the line between that point and the center of the cortical bone (obtained in the preprocessing step). Positive $P^+(C)$ and negative $P^-(C)$ directional gradients are used for exterior and interior contour evolution respectively. The constraint between the dual contours is the thickness of DACM (the distance between two contours). If the thickness at some points is greater than one standard deviation beyond the mean, a constraint force will be exerted to pull the two contours closer. The weights for different forces are kept constant throughout the evolution. The initialization of the DACM is as follows. From the center of the femur bone, a ray is shot outwards in every direction (360°). The first point with probability >0.5 on the holistic edge image hit by the ray is set as one point on the initial interior contour. The initial interior contour is expanded outward until hitting a point with probability 0, which is then set as an initial exterior contour point. Outliers are removed if their distances to the bone are beyond two standard deviations from mean. Missing initial contour points are interpolated by neighbors.

The average of converged C_I and C_E is set as the final fascia lata contour. Figure 5 demonstrates the DACM for fascia lata segmentation.

2.4   Holistic Tissue Classification

All thigh pixels outside the fascia lata are considered to be subcutaneous adipose tissue (SAT). The regions inside the fascia lata are classified as either muscle or IMAT (cortical bones and bone mallows have already been masked out in the preprocessing

stage). For now, we ignore other tissues such as vessel, tendon and nerve since they only count for a very small portion of the thigh.

Using architecture similar to the holistic fascia lata detector, a holistic region neural network is designed to classify muscle and IMAT regions in an image-to-image training and prediction manner. The training images are cropped around the fascia lata with manually delineated muscle regions as foreground and IMAT as background. The fused output returns the probability map of a pixel belonging to the muscle region. The images are normalized using the cortical bone intensity (lowest intensity) and 95% highest intensity in the image to address the intensity inconsistency issue with MRI images. A threshold of 0.5 is applied on the holistic output to differentiate muscle and IMAT. Figure 6 shows several side outputs of the holistic muscle neural net and the final tissue classification result.
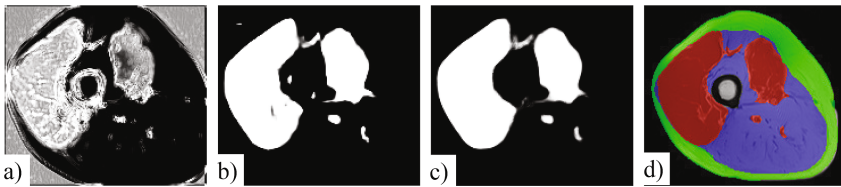

Fig. 6. Holistic tissue classification. (a) side output 1; (b) side output 3; (c) fused muscle potential image; (d) final tissue classification, red: muscle; blue: IMAT; green: SAT. The original MR image is shown in Fig. 5(a).

2.5 Implementation Details

The holistic network is implemented using the publicly available Caffe library [8]. The network is first loaded with pre-trained VGG-16 model [9], and then fine-tuned with our thigh training data. The hyper-parameters include: learning rate ($1e^{-6}$), momentum (0.9 for edge, 0.5 for region), weight decay (0.0002). All images are resized to 400 * 400 for training and prediction. The left and right thigh images are also mirrored to augment the training set. The training takes about 4 h on two NVidia Titan X GPUs with 12 GB memory. The testing takes about 0.6 s.

3 Results

T1 weighted thigh MRIs from 104 patients (mean age:40 year, 38 male, 66 female) were acquired for this research. The MR parameters included TR 600 ms, TE 20 ms, flip angle 150, slice thickness 8 mm, and in-plane resolution 0.98 mm. Among the studies, there were 16 mild/normal, 23 moderate and 65 severe cases. We randomly selected 25 cases (9 mild, 8 moderate, and 8 severe) for quantitative training and testing, and the rest for visual inspection and qualitative assessment. For the 25 cases, an operator manually traced the fascia lata and classified the tissues. 15 cases (322 thigh images) were used for training and 10 cases (180 thigh images) for independent testing. For a subset of the test set (60 thigh images), a second operator conducted

manual segmentation independently for the inter-observer variability assessment. The results reported in this section are for the test set performance.

Figure 7 shows results for fascia lata segmentation done via manual segmentation, active contour model, the state-of-the art model based method [6], and our proposed method. Table 1 compares the performance metrics including Dice coefficient of the region enclosed by the fascia lata, mean and max contour distance for active contour model, state-of-the-art model based method [6], inter-observer variability, and our proposed method. Classic active contour models work well on mild cases, but fail to detect the fascia lata in severe cases. The state-of-the-art model-based method misses a few boundaries and has trouble with severe inhomogeneity artifact.
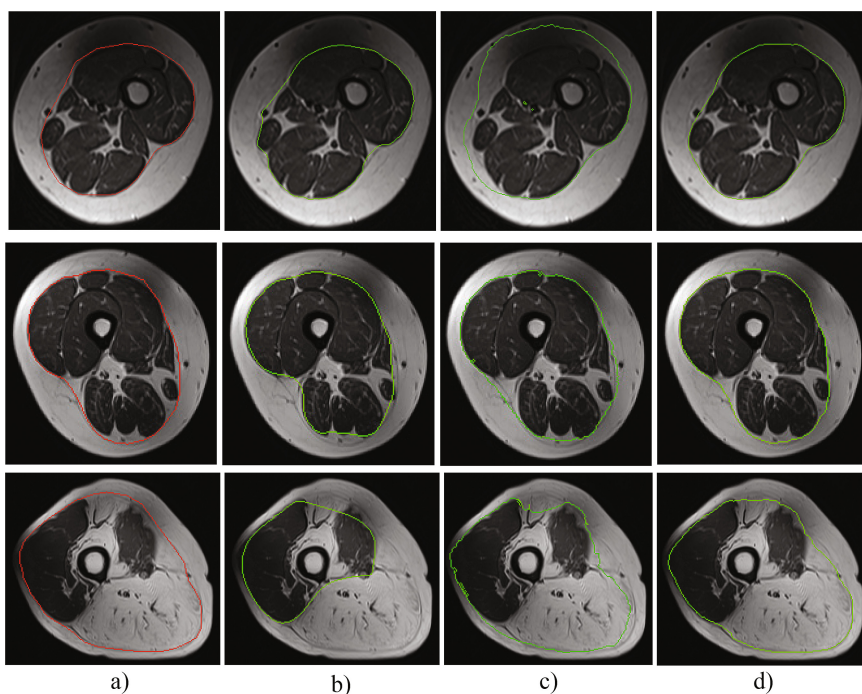

a)   b)   c)   d)

Fig. 7. Examples of fascia lata segmentation. (a) manual segmentation; (b) active contour model; (c) state-of-the-art model-based method, and (d) our proposed method.

Table 1. Fascia lata segmentation performance

	Dice coefficient			Mean dist (mm)			Max dist (mm)		
	Mild	Mod	Sev	Mild	Mod	Sev	Mild	Mod	Sev
ACM	0.957 ± 0.02	0.838 ± 0.29	0.686 ± 0.23	2.7 ± 1.5	16.1 ± 35.7	22.3 ± 25.	11.3 ± 6.1	31.9 ± 49.	58.2 ± 36.
Model-based [6]	0.944 ± 0.03	0.931 ± 0.06	0.929 ± 0.1	3.7 ± 2.7	3.5 ± 5.3	6.9 ± 15.0	16.2 ± 10.	14.5 ± 12	19.0 ± 26
Inter-Obs	0.977 ± 0.01	0.975 ± 0.01	0.969 ± 0.01	**1.0 ± 0.7**	**1.1 ± 0.7**	**1.3 ± 0.7**	**5.1 ± 4.3**	**5.3 ± 3.9**	**5.7 ± 2.9**
Our Method	**0.978 ± 0.01**	**0.976 ± 0.01**	**0.974 ± 0.02**	1.2 ± 0.6	1.4 ± 0.8	1.3 ± 1.1	6.1 ± 3.6	6.7 ± 4.1	6.6 ± 6.4

Figure 8 shows results of tissue classification. Table 2 compares the performance of the fuzzy C-means method, inter-observer variability and our proposed method. We note that our method works robustly for severe disease cases and cases with severe inhomogeneity artifact. Our method also performs comparably (sometimes better than) with inter-observer variability.
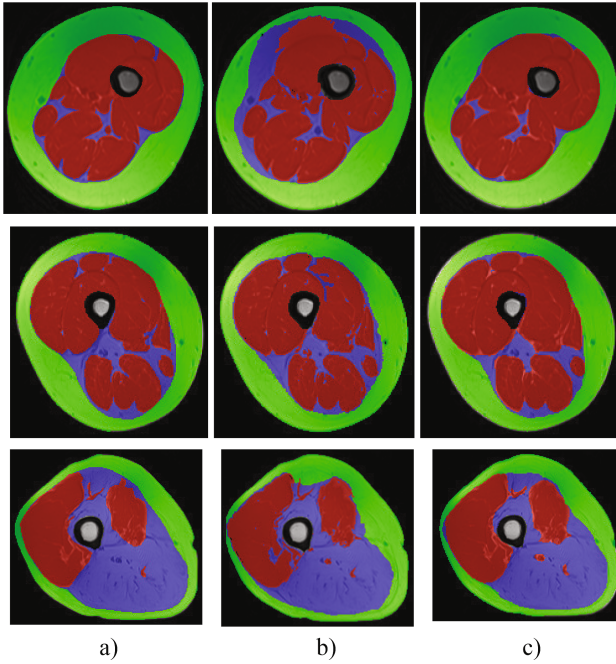

Fig. 8. Examples of tissue classification. (a) manual segmentation; (b) fuzzy c-mean clustering; and (c) our approach. Red: muscle, blue: IMAT, green: SAT. Original MRI images in Fig. 7.

Table 2. Tissue classification performance (Dice Coefficient).

	Muscle			IMAT			SAT		
	Mild	Mod	Sev	Mild	Mod	Sev	Mild	Mod	Sev
Fuzzy C-means	0.96 ± 0.02	0.917 ± 0.06	0.801 ± 0.19	0.434 ± 0.15	0.558 ± 0.13	0.849 ± 0.18	0.887 ± 0.05	0.942 ± 0.02	0.912 ± 0.09
Inter-Obs	0.945 ± 0.02	0.925 ± 0.02	0.811 ± 0.19	**0.654 ± 0.01**	0.696 ± 0.05	0.922 ± 0.03	**0.947 ± 0.01**	**0.946 ± 0.01**	**0.946 ± 0.03**
Our Method	**0.978 ± 0.01**	**0.969 ± 0.02**	**0.940 ± 0.07**	0.646 ± 0.1	**0.748 ± 0.11**	**0.943 ± 0.03**	0.945 ± 0.02	0.945 ± 0.03	0.944 ± 0.03

4   Conclusion

The two (edge and region) holistic neural networks built on top of fully connected network and deeply-supervised network are capable of detecting weak boundaries such as the fascia lata and handle image artifact. Integrating the holistic network with dual active contour models allow us to achieve highly accurate classification for severe

muscle dystrophy cases on MRI. The proposed method outperforms the state-of-the-art model-based methods and unsupervised clustering techniques.

References

1. Mercuri, E., Muntoni, F.: Muscular dystrophies. Lancet **381**, 845–860 (2013)
2. Orgiu, S., et al.: Automatic muscle and fat segmentation in the thigh from T1-Weighted MRI. J. MRI **43**(3), 601–610 (2016)
3. Positano, V., et al.: Accurate segmentation of subcutaneous and intermuscular adipose tissue from MR images of the thigh. J. MRI **29**, 677–684 (2009)
4. Chambers, O., et al.: Computer-based assessment for facioscapulohumeral dystrophy diagnosis. Comput. Methods Progr. Biomed. **120**(1), 37–48 (2015)
5. Tan, C., et al.: Accurate thigh inter-muscular adipose quantification using a data-driven and sparsity-constrained deformable model. In: IEEE International Symposium on Biomedical Imaging (2015)
6. Kovacs, W., et al.: Identification of muscle and subcutaneous and intermuscular adipose tissue on thigh MRI of muscular dystrophy. In: IEEE International Symposium on Biomedical Imaging (ISBI) (2016)
7. Xie, S., Tu, Z.: Holistically-nested edge detection. In: IEEE International Conference on Computer Vision (2015)
8. Jia, Y., et al.: Caffe: convolutional architecture for fast feature embedding. In: The 22nd ACM International Conference on Multimedia (2014)
9. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations (2015)