Segmentation of the Fascia Lata in Magnetic Resonance Images of the Thigh
Comparison of an Unsupervised Technique with a U-Net in 2D and Patch-wise 3D

## Lis J. Louise P.[1], Klaus Engelke[1,2], Oliver Chaudry[1,2]

[1]Institute of Medical Physics, University of Erlangen-Nuremberg, Germany
[2]Department of Medicine 3, Friedrich-Alexander-Universität Erlangen-Nürnberg
University Hospital Erlangen

INTRODUCTION

Quantification of adipose tissue (AT) in skeletal muscle is of growing interest to understand mechanisms of muscle weakness during ageing and in diseases such as sarcopenia and cachexia. The mid-thigh is the preferred anatomical location for such measurements. It was recently shown that intermuscular adipose tissue (IMAT), the combination of AT among muscles and the agglomeration of larger adipocytes within muscles is a very sensitive parameter to monitor exercise effects, the most widely used intervention to prevent muscle weakness with increasing age [1]. The assessment of IMAT requires an accurate segmentation of the deep fascia or fascia lata (FL), an envelope of fibrous tissue separating subcutaneous adipose tissue from muscles and IMAT (Fig. 1)

For this purpose, the U-Net architecture was implemented and compared for 2D images and patched 3D image stacks along the z direction in magnetic resonance images (MRI). The training data consisted of T_1 MRI data sets from elderly men. To test the performance of the models, they were applied to other MRI thigh datasets of subjects of different age groups and gender and compared to an unsupervised semi-automatic method [2].
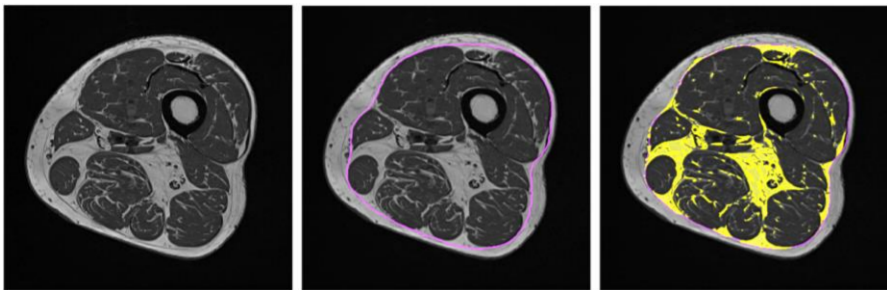
Fig. 1. From left to right:
MRI slice of the thigh
segmented fascia lata (purple)
segmented IMAT (yellow)

METHODS

Data acquisition

- 4 different MRI datasets (D1-D4) from 3 studies
  - **D1**: 43 elderly men (age≥72 years, BMI 24.5±1.9 kg/m^2), 2 visits, 70 scans in total, taken from study [3]
  - **D2**: 17 young men (age≥25 years, BMI 23.4±1.9 kg/m^2) from study [4]
  - **D3**: 17 elderly men (age≥72 years, BMI 27.3±2.2 kg/m^2) from study [4]
  - **D4**: 16 elderly women (age≥71 years, BMI 24.9±1.4kg/m^2) from study [5]
- 3T MAGNETOM Skyra^{fit} Scanner (Siemens Healthineers AG, Erlangen, Germany)
  - 18-channel body receiver array coil
  - T1 weighted Turbo Spin Echo sequence: TR: 844 ms, TE: 14 ms, voxel size: 0.5×0.5×3.0 mm^3 (no slice gap), matrix size: 512×512 in 28 slices
  - Bias field corrected using N4ITK algorithm [6]
- Ground truth segmentation masks were generated by a set level method and manually corrected by medical expert [2].
- Unsupervised masks were automatically generated by a set level method but not corrected by medical expert.

Model

- Training on subset of D1; testing on D2, D3, D4 and a test set from D1
- FL segmentation by a U-Net deep neural network
  - implementation followed the original publication [7], except for the padding per convolution, which did not change the input dimensions;
  - Adam optimization and dropout regularization (dropout rate of 0.5)
- Because of class imbalance the ROI bordered by the FL was segmented instead of the thin FL contour,
- Performance analysis by comparing an implementation in 2D with a patched 3D U-Net
- Implementation in tensorFlow using a Nvidia Geforce GTX 1060 6GB GPU

2D U-Net

- Input: 1960 slices (28 slices per patient)
  - full resolution 512x512 MR images.
  - total number of 2D samples was shuffled, normalized and split at slice level into training (70 %), validation (15 %) and test (15 %) sets.

3D U-Net

- Down sampled to half of the image size in in-plane resolution.
- Input was split at the patient level into 57 patients for training, 7 for validation and 6 for testing.
- Patch size was restricted to 4 slices with a batch size of 1.

Loss function

Linear combination of dice loss (DL) and weighted cross entropy (WCE)

$$Loss(x, y) = \alpha WCE + (1 - \alpha)DL$$

with *DL* equal to 1 - DSC

$$DSC(y, p) = \frac{2\sum_i y_i p_i + s}{\sum_i y_i + \sum_i p_i + s}$$

and

$$WCE(x, p) = -(\beta \cdot y \log(p) + (1 - y)\log(1 - p))$$

Training

- Initial weights obtained using He initialization for the 2D model. For the 3D model pre-trained model on the gold standard segmented cross sectional area of the thigh was used for initialization
- Grid search with cross validation found optimum values for α = 0.2 and a learning rate of 2e-4. Same parameters were used for the 3D model
- Training of the 2D model took 2 hours for 12 epochs.
- The last 6 epochs were trained at a learning rate of 2e-5 to refine the results

Assessment of Accuracy compared to Gold standard

1. Dice similarity coefficient (DSC) of the ROI enclosed by the FL
2. Surface Dice similarity coefficient of the FL
3. Hausdorff distance (HD) of the ROI enclosed by the FL

$$HD(A, B) = \max\big(h(A, B), h(B, A)\big)$$

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

4. Average symmetric surface distance (ASD)

$$ASD(A, B) = Average\left(\min_{b \in B}\|a - b\|, \min_{a \in A}\|b - a\|\right)$$

RESULTS

2D U-Net

- Table 1.: Accuracy of U-Net 2D and of unsupervised classical segmentation results compared to gold standard
- 2D U-net was trained on D1
- Fig. 2: Representative images of U-Net 2D results for each data set.

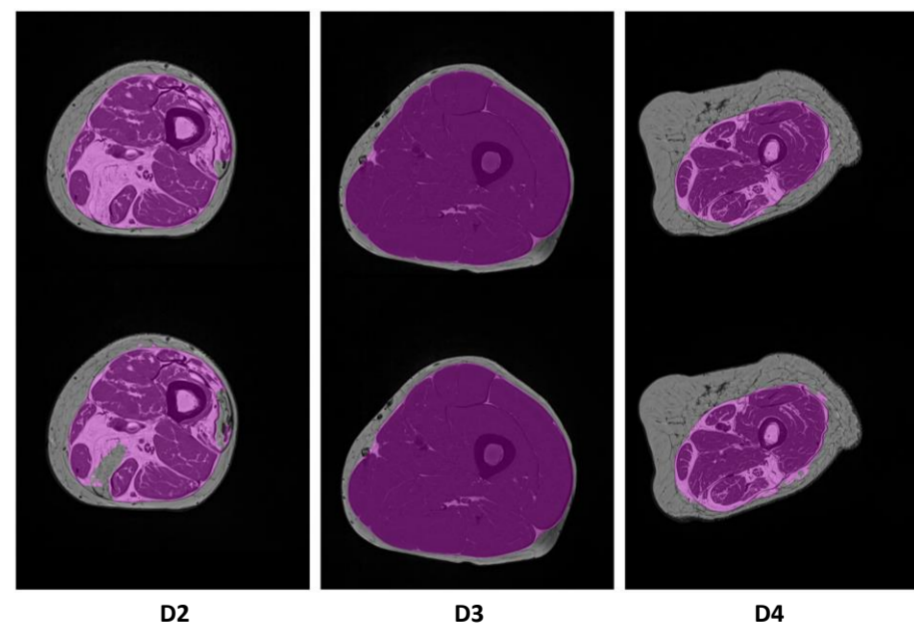| Dataset	D1	D2	D3	D4
HD [pixel] Unsup.	15.050 ± 16.714	12.951 ± 12.648	6.123 ± 3.072	10.590 ± 6.117
HD [pixel] U-Net 2D	2.860 ± 0.935	7.139 ± 7.799	4.807 ± 2.547	10.135 ± 5.494
DSC Unsup.	0.976 ± 0.075	0.985 ± 0.030	0.998 ± 0.004	0.987 ± 0.012
DSC U-Net 2D	0.996 ± 0.001	0.992 ± 0.008	0.994 ± 0.002	0.986 ± 0.008
Surface Dice Unsup.	0.767 ± 0.212	0.768 ± 0.152	0.869 ± 0.192	0.751 ± 0.193
Surface Dice U-Net 2D	0.692 ± 0.104	0.546 ± 0.117	0.522 ± 0.118	0.454 ± 0.104
ASD [pixel] Unsup.	2.557 ± 7.786	1.727 ± 3.373	0.333 ± 0.503	1.287 ± 1.212
ASD [pixel] U-Net 2D	0.359 ± 0.167	1.256 ± 2.062	0.965 ± 2.078	1.399 ± 0.912
Number of samples	196	442	442	416
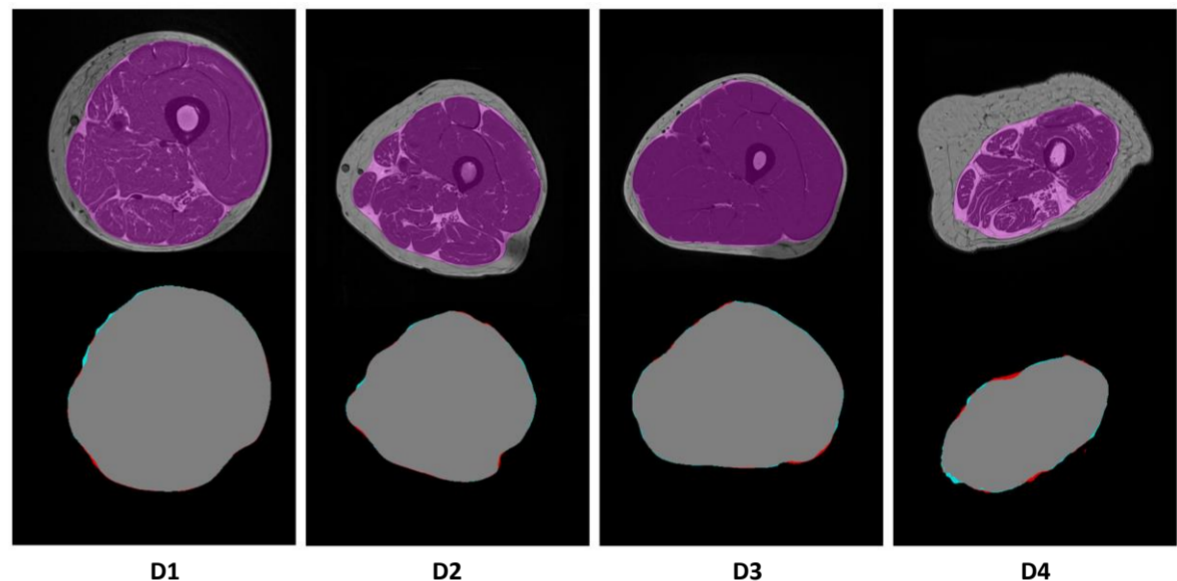
Fig. 2. Top row: MR image with overlay of the 2D U-Net prediction mask (magenta). Bottom row: gold standard mask (red), prediction mask (cyan) and intersection of both masks (gray).

D1   D2   D3   D4

3D U-Net

- Table 2 and Fig 3: Corresponding results for U-Net 3D after up-sampling

Dataset	D1	D2	D3	D4
HD [pixel]	4.180 ± 1.714	6.459 ± 3.813	5.204 ± 3.532	19.622 ± 44.141
DSC	0.994 ± 0.002	0.992 ± 0.002	0.994 ± 0.004	0.982 ± 0.010
Surface Dice	0.496 ± 0.106	0.471 ± 0.082	0.502 ± 0.095	0.355 ± 0.096
ASD [pixel]	0.658 ± 0.214	0.892 ± 0.437	0.811 ± 0.622	1.998 ± 1.749
Number of samples	7	17	17	16

Fig. 3. Comparison of the 3D segmentation (upper row) with the 2D segmentation (bottom row) for the same patients in D2, D3 and D4.

D2   D3   D4

Assessment of manual correction time

Dataset	D2	D3	D4
Unsupervised [min]	4.3 ± 1.7	2.4 ± 1.1	6.8 ± 2.9
2D U-Net [min]	1.8 ± 0.9	1.0 ± 0.4	4.6 ± 2.5

DISCUSSION

Two different U-Net models were implemented for the segmentation of the FL. Although the models were only trained on images of elderly men, both approaches also delivered excellent results for young men. The models were able to predict the FL segmentation with higher accuracy than the unsupervised method and halved the time for final manual corrections by a medical expert.
The patched 3D approach showed slightly lower accuracy in the test datasets but produced smoother, less convex segmentations.
Results for D4 indicated that higher subcutaneous and intramuscular AT content caused poorer segmentation. Data from both genders and 3D information should be included in the training data to gain best results.

[1] Grimm A et al. Feasibility of dixon magnetic resonance imaging to quantify effects of physical training on muscle composition—A pilot study in young and healthy men. EJR 2019
[2] Chaudry O et al. Segmentation of the fascia lata and reproducible quantification of intermuscular adipose tissue (IMAT) of the thigh. MAGMA 2020
[3] Kemmler W et al. Effects of high intensity dynamic resistance exercise and whey protein supplements on osteosarcopenia in older men with low bone and muscle mass. Final results of the randomized controlled FrOST study. Nutrients. 2020
[4] Kemmler W et al. Effects of combined whole-body electromyostimulation and protein supplementation on local and overall muscle/fat distribution in older men with sarcopenic obesity: the randomized Controlled Franco-nia Sarcopenic Obesity. CTI 2018
[5] Kemmler W et al. Prevalence of sarcopenia in Germany and the corresponding effect of osteoarthritis in females 70 years and older living in the community: results of the FORMoSA study. CIA 2015
[6] Tustison NJ et al. N4ITK: Improved N3 bias correction. IEEE 2010
[7] Ronneberger O et al. U-Net: Convolutional networks for biomedical image segmentation. MICCAI 2015