

Received October 22, 2021, accepted November 7, 2021, date of publication November 11, 2021, date of current version November 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3127565

Automatic Segmentation of Supraspinatus Muscle via Bone-Based Localization in Torso Computed Tomography Images Using U-Net

YUICHI WAKAMATSU¹, NAOKI KAMIYA¹, (Member, IEEE), XIANGRONG ZHOU², HIROKI KATO³, TAKESHI HARA^{2,4}, AND HIROSHI FUJITA²

¹Graduate School of Information Science and Technology, Aichi Prefectural University, Nagakute, Aichi 480-1198, Japan

²Faculty of Engineering, Gifu University, Gifu 501-1193, Japan

³Department of Radiology, Graduate School of Medicine, Gifu University, Gifu 501-1193, Japan

⁴Center for Healthcare Information Technology (C-HiT), Tokai National Higher Education and Research System, Nagoya, Aichi 464-8601, Japan

Corresponding authors: Yuichi Wakamatsu (im202013@cis.aichi-pu.ac.jp) and Naoki Kamiya (n-kamiya@ist.aichi-pu.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (C) under Grant 21K12731.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Review Committees of Gifu University (28-120, June 6, 2020) and Aichi Prefectural University (Jo2020-03, July 13, 2020).

ABSTRACT The supraspinatus tendon is the most frequently torn tendon in the rotator cuff. Rotator cuff reconstruction is more likely to result in re-tear if the muscle has atrophy or fatty degeneration. Thus, atrophy and fatty degeneration of the supraspinatus muscle are predictors of the postoperative course, and volume analysis using three-dimensional segmentation of the supraspinatus muscle is necessary. The supraspinatus muscle is attached to the scapula, making it possible to estimate the region of the muscle based on the position of the scapula. In this paper, we propose a supraspinatus muscle segmentation method based on the scapula position in torso computed tomography (CT) images. Our proposed method consists of supraspinatus muscle localization using a scapula segmentation result and supraspinatus muscle segmentation based on the localization result. U-Net is used for scapula and supraspinatus muscle segmentation. In this experiment, we used torso CT images and pseudo-chest CT images which were generated from the scans of the same patient. The mean Dice values of the segmentation results obtained by applying the proposed method to the torso and pseudo-chest CT images were both 0.881. When localization was not used, the mean Dice values of the segmentation results in the torso and pseudo-chest CT images were 0.000 and 0.850, respectively. The experimental results demonstrate the effectiveness of bone-based localization in supraspinatus muscle segmentation using U-Net.

INDEX TERMS Computed tomography, scapula, segmentation, supraspinatus muscle, U-Net.

I. INTRODUCTION

The supraspinatus muscle is one of the skeletal muscles that make up the rotator cuff. The supraspinatus tendon is the most frequently torn tendon in the rotator cuff [1], and rotator cuff tears cause weakness, pain, and limited/reduced mobility [2]. Patients with rotator cuff tears are treated with reconstructive surgery, but if muscle atrophy or fatty degeneration occurs prior to surgery, they are more likely to re-tear after surgery [3]. Thus, muscle atrophy and fatty degeneration are important factors in predicting the risk of

re-tear after reconstructive surgery. Muscle atrophy and fatty degeneration have been quantitatively assessed by measuring transverse area and volume, and Hounsfield units in muscle regions, respectively [4], [5]. Therefore, for automatic analysis of supraspinatus muscle atrophy and fatty degeneration, three-dimensional (3D) supraspinatus muscle segmentation capable of volume measurement and region identification is necessary.

Skeletal muscles are attached to two bones across a joint. The regions on bones where skeletal muscles are attached are termed the origin and insertion [6]. Since the origin and insertion of each muscle are uniquely determined, the location of the muscle can be estimated by the location of the bones to

The associate editor coordinating the review of this manuscript and approving it for publication was Shen Yin.

which the muscle is attached. In addition, the location of a bone can be used to identify the position of the target region in medical images. For example, the identification of the L3 slice and the segmentation of its muscles have been studied previously [7].

Computational anatomy is focused on skeletal muscles and their site-specific segmentation [8]. Research on site-specific skeletal muscle segmentation utilizes torso computed tomography (CT) images. The conventional method of skeletal muscle segmentation comprises automatic recognition of landmarks (LMs) corresponding to the origin and insertion in the automatic bone segmentation results, representing the position and shape of muscle fibers by lines connecting the LMs, and using these components to generate an overall shape model. This model-based method, which uses a handcrafted feature, has been used for the segmentation of nine skeletal muscles, including the supraspinatus muscle [9], [10]. In addition, in the conventional method, the skeletal muscles are segmented by three consecutive processes: recognition of LMs, representation of muscle fiber, and generation of a shape model. Therefore, this method has two limitations: only skeletal muscles that can easily be modeled are segmented, and the accuracy of each of the three processes affects the overall accuracy of the segmentation result [10]. Automatic segmentation of the skeletal muscles involves machine learning, including deep learning, and thus overcomes these limitations. Erector spinae muscle segmentation using random forest and Bayesian U-Net has achieved an accuracy of 93.0% and 93.4%, respectively [11], [12].

For supraspinatus muscle segmentation in CT images, the model-based method by Katafuchi *et al.*, described earlier [9], and the rotator cuff segmentation method using deep learning by Taghizadeh *et al.* [13] have been proposed. The former method [9] includes segmenting the supraspinatus muscle in torso CT images by automatically recognizing the LMs based on the scapula segmentation results, expressing the position and shape of muscle fibers with lines connecting the LMs, and generating a shape model. The mean Jaccard index of the segmentation results obtained by this method was 0.491. However, the segmentation accuracy depends on the accuracy of the LMs recognition. On the other hand, the latter method [13] uses a modified version of U-Net [14] for segmentation of each of the four skeletal muscles that comprise the rotator cuff. The mean Dice value of the supraspinatus muscle segmentation results by this method was 0.91, making this effective even though the evaluation metrics are different. However, this method segments the muscle only one cross-section perpendicular to the scapular axis and passing through the spinoglenoid notch; hence, the segmentation result is two-dimensional (2D). Although the segmentation in CT images has been studied, 3D segmentation, which is necessary to accurately measure the volume of the supraspinatus muscle, has not been achieved.

The purpose of this study was to propose the use of automatic segmentation of the supraspinatus muscle based on the localization of the muscle in torso CT images. To verify

the effectiveness of the proposed method, we used torso CT images and pseudo-chest CT images, which were generated from the same patient. The proposed method and a segmentation method without localization were applied to the two types of CT images, and the segmentation accuracies were compared.

This paper is organized as follows. In Section II, the proposed method is described. In Section III, image details, experimental environment, and evaluation methods are described. In Section IV, we show the experimental results. Then, we discuss the results of the experiments in Section V and conclude the paper in Section VI.

II. METHODS

In the proposed method, two-stage segmentation is performed using two-stage U-Net for segmentation of the supraspinatus muscle in torso CT images. The two-step segmentation method like the proposed method has been proposed in other studies. For example, Liu *et al.* used a two-stage U-Net for segmentation of the whole heart and its substructures [15]. The first stage U-Net segmented the whole heart, and the second stage U-Net segmented the substructures using the segmentation result of whole heart. As shown in this method, the possibility of a method that first segments the entire target organ and then divides it in detail has been suggested. On the other hand, in this study, we focus on the anatomical relationship of muscles attached to bones, and propose a method that localizes the muscle based on the segmentation of the bone to which the muscle is attached in the first stage, and performs muscle segmentation using the localization result in the second stage. An overview of the proposed method is shown in Fig. 1 and consists of two steps: supraspinatus muscle localization (Stage 1) and supraspinatus muscle segmentation (Stage 2). In Stage 1, the supraspinatus muscle is localized based on the segmentation result of the scapula, the bone to which the muscle is attached. Then, in Stage 2, the supraspinatus muscle is segmented in the region including the muscle using the localization results of Stage 1. In these stages, we use U-Net [14] for the segmentation of the scapula and supraspinatus muscle. By using U-Net, which is a general-purpose network for medical image segmentation, we can verify the effectiveness of the bone-based localization in the proposed method. The details of this method are described below.

A. SUPRASPINATUS MUSCLE LOCALIZATION

The supraspinatus muscle is a skeletal muscle of the shoulder girdle, originating from the supraspinatus fossa of the scapula and inserting into the greater tubercle of the humerus. In other words, this muscle is located on the upper part of the scapula. Therefore, the area from the top slice of torso CT images to the bottom slice of the scapula was defined as the region around the supraspinatus muscle because it anatomically contains the muscle. In this stage, the scapula is automatically segmented in the torso CT images. Then, the muscle is localized by cropping the region around the muscle

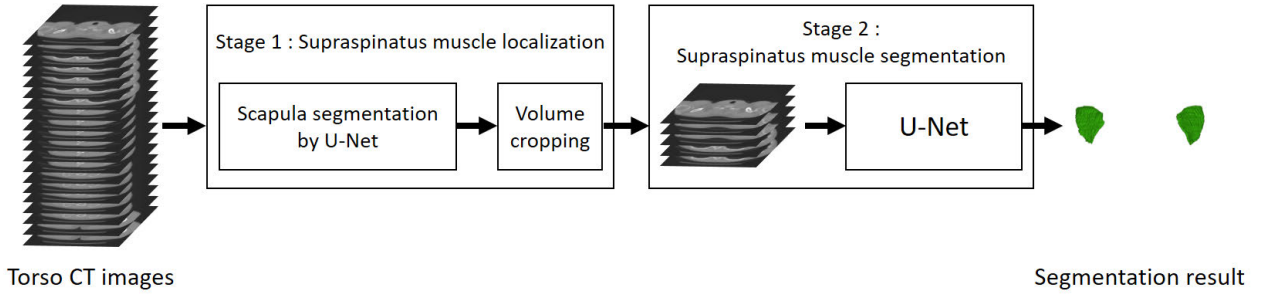


FIGURE 1. Overview of the proposed method for automatic segmentation of the supraspinatus muscle via bone-based localization in torso CT images.

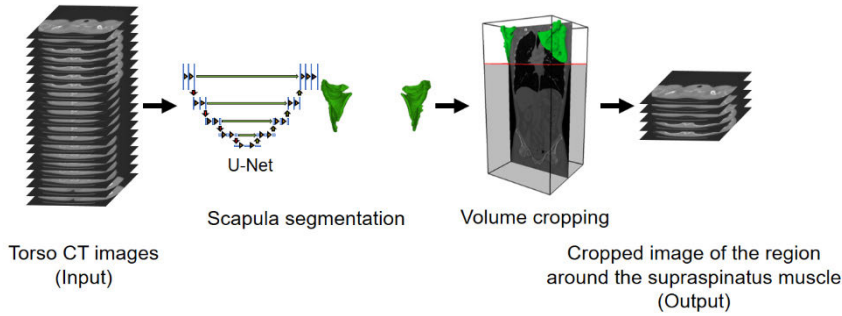


FIGURE 2. Overview of the supraspinatus muscle localization stage in the proposed method.

using the scapula segmentation result. An overview of this stage is shown in Fig. 2.

For scapula segmentation, we use U-Net, which is a network consisting of Encoder and Decoder; in Encoder, the process of applying 3×3 Convolution twice and Max Pooling once is repeated four times. Then, 3×3 Convolution is applied to the feature map twice repeatedly. In Decoder, the process of combining the feature map with Up Convolution and the corresponding feature map in Encoder, and applying 3×3 Convolution twice is repeated four times. Finally, 1×1 Convolution and Sigmoid are applied to obtain the segmentation result. Zero padding is applied before all 3×3 Convolutions, while batch normalization and Rectified Linear Unit (ReLU) are applied afterwards. U-Net utilized all the 2D axial images in torso CT images as input, and segmented the scapula region in the input images. Then, the 3D segmentation result of the scapula was obtained by stacking the 2D segmentation results. The training parameters of this network were as follows: the number of epochs was 100, batch size was 16, optimization function was Adam [16], the learning rate was 3×10^{-4} , and the loss function was a combination of binary cross-entropy (BCE) loss and Dice loss. These loss functions were defined as follows:

$$loss_{BCE} = -\frac{\sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]}{N} \quad (1)$$

$$loss_{Dice} = -\frac{2 \times \sum_{i=0}^N y_i \hat{y}_i + smooth}{\sum_{i=0}^N y_i + \sum_{i=0}^N \hat{y}_i + smooth} \quad (2)$$

$$loss = 0.5 \times loss_{BCE} + loss_{Dice} \quad (3)$$

where N refers to the number of pixels, y_i and \hat{y}_i refer to the ground truth label and the predicted probability of the point i , respectively, and $smooth$ is a constant to avoid division by zero. Data augmentation was applied in the training phase. We used a shear transformation and rotation of random angles from $-\pi/8$ to $+\pi/8$ and from -10° to $+10^\circ$, respectively, a random scaling from -35% to $+35\%$, a translation of random distances with a maximum value of 25% of the side image length, and horizontal flip. After processing with U-Net, we extracted the two regions with the largest volume, which we regarded as the result for the scapula segmentation since the human body has two scapulae.

Next, the region around the supraspinatus muscle was cropped in the torso CT images using the result for scapula segmentation. The lengths of the vertical and horizontal axes of the axial images remained unchanged as only the range in the direction of the body axis was specified for cropping. The cropped image was regarded as the output for this stage.

B. SUPRASPINATUS MUSCLE SEGMENTATION

The second stage is supraspinatus muscle segmentation. In Stage 1, based on the anatomical fact that the supraspinatus muscle attaches to the scapula and humerus, the muscle localization was performed by cropping the CT image using the location information of the scapula. Then, in Stage 2, supraspinatus muscle segmentation is performed in the images localized in Stage 1. For segmentation of the

TABLE 1. Evaluation Results (mean \pm standard deviation).

Input CT images	Method	Dice	Jaccard	Precision	Recall
Torso	With-localization (Proposed)	0.881 \pm 0.087	0.797 \pm 0.114	0.891 \pm 0.064	0.884 \pm 0.129
	Without-localization	0.000 \pm 0.000	0.000 \pm 0.000	0.000 \pm 0.000	0.000 \pm 0.000
Pseudo-chest	With-localization (Proposed)	0.881 \pm 0.087	0.797 \pm 0.114	0.891 \pm 0.064	0.884 \pm 0.129
	Without-localization	0.850 \pm 0.122	0.754 \pm 0.145	0.913 \pm 0.033	0.819 \pm 0.171

Evaluation results of supraspinatus muscle segmentation by the with-localization (proposed) method and the without-localization method in the torso and pseudo-chest CT images. The segmentation accuracies of the proposed method in the torso and pseudo-chest CT images were similar. CT, computed tomography.

supraspinatus muscle, U-Net, which has the same network structure as U-Net in Stage 1, was used. The 2D axial images from the cropped image were the U-Net input, and the 2D segmentation results were the output. The 3D supraspinatus muscle segmentation results were obtained from the 2D results. The training parameters of U-Net and the data augmentation method in this stage were the same as those in Stage 1. The extraction of the two largest volumes was then applied to the U-Net output.

III. EXPERIMENTS

A. IMAGE DETAILS AND EXPERIMENTAL ENVIRONMENT

In this study, we used 30 non-contrast torso CT images obtained using the LightSpeed Ultra 16 (GE Healthcare, Chicago, IL, USA) at Gifu University Hospital. This study was approved by the ethical review committees of Gifu University (28-120, June 6, 2020) and Aichi Prefectural University (Jo2020-03, July 13, 2020). The image size was $512 \times 512 \times 802.1104$ [voxel], and the spatial resolution was $0.625 \times 0.625 \times 0.625$ [mm]. Ground truth was created by manual segmentation of the scapula and supraspinatus muscle, as recommended by an anatomist.

The experiments were conducted on a computer with 4 Tesla V100 (32 GB) graphics processing units (GPU). We used TensorFlow-GPU 1.10.0 [17] and Keras 2.2.2 [18] as machine-learning libraries.

B. EVALUATION METHODS

We used the Dice value, Jaccard index, precision, and recall to evaluate the similarity between the result for supraspinatus muscle segmentation and the ground truth. Definitions of these metrics are as follows:

$$Dice = \frac{2 \times |Res \cap GT|}{|Res| + |GT|} \quad (4)$$

$$Jaccard = \frac{|Res \cap GT|}{|Res \cup GT|} \quad (5)$$

$$Precision = \frac{|Res \cap GT|}{|Res|} \quad (6)$$

$$Recall = \frac{|Res \cap GT|}{|GT|} \quad (7)$$

where Res refers to the segmentation result, GT refers to the ground truth, and the operator $||$ returns the number of voxels

contained in the region. Then, for the supraspinatus muscle localization, the accuracy of slice estimation at the bottom of the scapula was evaluated using Mean Absolute Error (MAE). MAE was defined as follows:

$$MAE = Abs(Idx_{Res} - Idx_{GT}) \quad (8)$$

where Idx_{Res} and Idx_{GT} refer to the index of the slice at the bottom of the scapula in the segmentation result and the ground truth, respectively, and function Abs returns the absolute value. The segmentation accuracy of the supraspinatus muscle was evaluated via 3-fold cross-validation using 20 of the 30 images as training data and 10 images as test data.

The proposed method was compared with a without-localization method. This without-localization method is a method that applies only Stage 2 of the proposed method. Therefore, all axial images of the torso CT images are input to U-Net, and the result of supraspinatus muscle segmentation is output. The training parameters and data augmentation method are the same as those of the proposed method. Then, we compare the segmentation results of the proposed method and the without-localization method to verify the effect of localization in the proposed method. Moreover, torso and pseudo-chest CT images were used as input CT images. The pseudo-chest CT images were generated by cropping the range from the top slice of the images to the bottom slice of the lung in the torso CT images. In addition, to compare the accuracy of the proposed method with the conventional method [13], which achieves 2D segmentation of the supraspinatus muscle, we compare the accuracy with the Dice value in a 2D sagittal-oblique image. The sagittal-oblique section for calculating the Dice value is the standardized sagittal-oblique section [19] defined as the plane perpendicular to the scapular axis and passing through the spinoglenoid notch, as in the conventional method. Sixty sagittal-oblique images were obtained from the left and right sides of 30 torso CT images, but two of them did not show the supraspinatus muscle. Therefore, 58 sagittal-oblique images were used to evaluate the accuracy.

IV. RESULTS

Table 1 presents the evaluation results of supraspinatus muscle segmentation by applying the with-localization (proposed) method and the without-localization method to

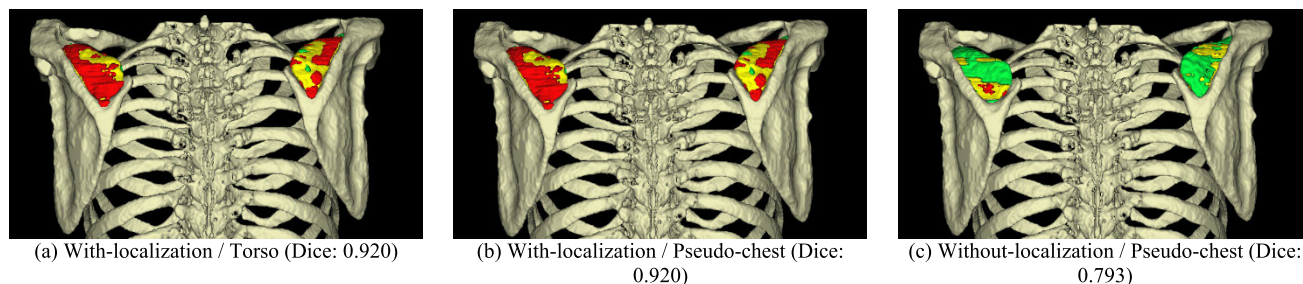


FIGURE 3. 3D-rendered images of the segmentation results. (a) Segmentation result obtained by applying the proposed method to the torso computed tomography (CT) images. (b) Segmentation result obtained by applying the proposed method to the pseudo-chest CT images. (c) Segmentation result obtained by applying the baseline method to the pseudo-chest CT images. The yellow area represents the overlapped area between the segmentation result and the ground truth, the red area represents the over-extracted area, and the green area represents the under-extracted area.

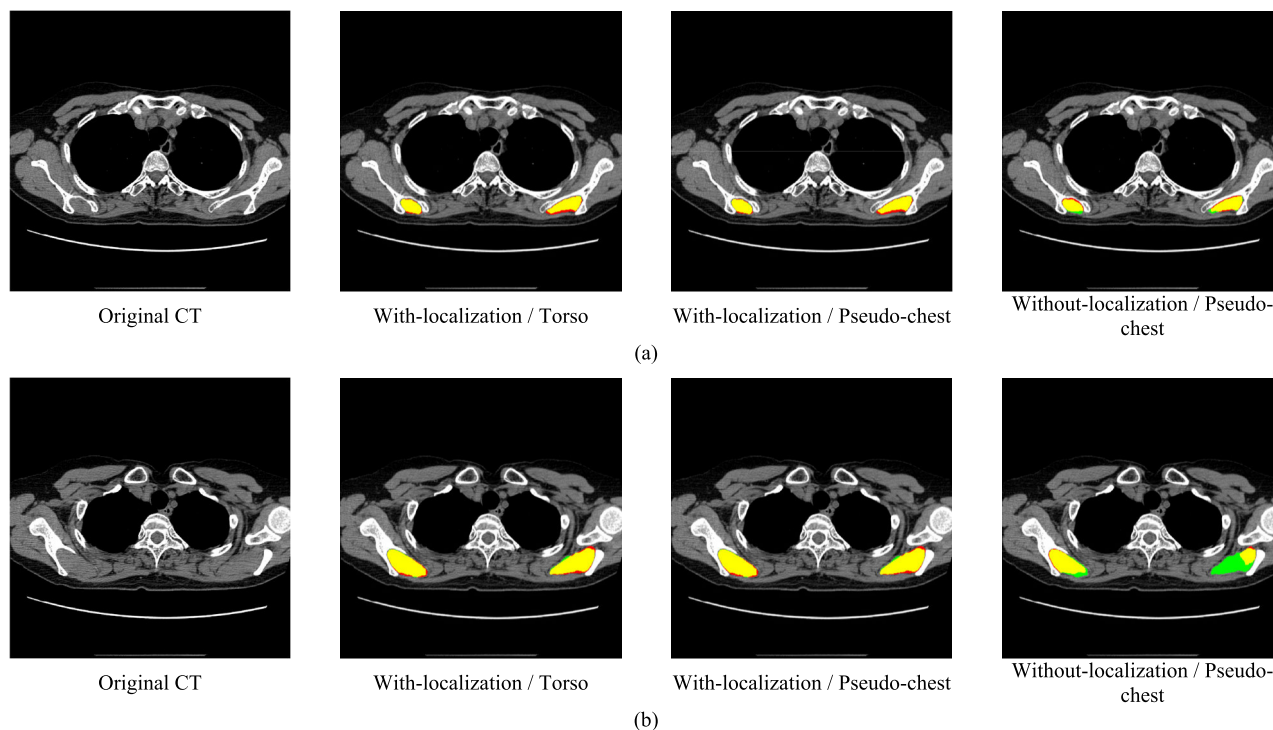


FIGURE 4. Axial images of the segmentation results. The yellow area represents the overlapped area between the segmentation result and the ground truth, the red area represents the over-extracted area, and the green area represents the under-extracted area.

the torso and pseudo-chest CT images. The mean Dice value of the segmentation results by the with-localization method was 0.881 (standard deviation [SD]: 0.087), the mean Jaccard index was 0.797 (SD: 0.114), the mean precision was 0.891 (SD: 0.064), and the mean recall was 0.884 (SD: 0.129) for the torso CT images. On the other hand, the mean Dice value, Jaccard index, precision, and recall were all 0.000 (SD: 0.000) for the without-localization method. In the latter, the pixels in all cases were segmented as the background. The segmentation results of both methods, applied to pseudo-chest CT images, are presented in Table 1. The 3D-rendered images of the supraspinatus muscle segmentation results are presented in Fig. 3, and the axial images of the segmentation results are shown in Fig. 4. Both

segmentation results are of the same case. The Dice values of the results shown in Fig. 3 and Fig. 4 after applying the with-localization method to the torso and pseudo-chest CT images were both 0.920. On the other hand, the Dice value of the result after applying the without-localization method to the pseudo-chest CT images was 0.793. In Fig. 3 and Fig. 4, the overlapped area between the segmentation result and the ground truth are marked in yellow, the over-extracted area in red, and the under-extracted area in green. Fig. 3(a) and (b) show some over-extracted regions in the segmentation results obtained by applying the with-localization method to both CT images. However, Fig. 3(c) shows large under-extracted regions in the segmentation result obtained by applying the without-localization method to the pseudo-chest CT images.

Moreover, in the slices shown in Fig. 4(a), a large segmentation error was not observed in any of the segmentation results; however, in the slices shown in Fig. 4(b), there was under-extraction in the segmentation result obtained by applying the without-localization method to the pseudo-chest CT images.

The MAE of the bottom slice of the scapula in the supraspinatus muscle localization was 0.433 slices in the torso CT images and 0.633 slices in the pseudo-chest CT images. The error was 0 slice in 17 cases for the torso CT images and 15 cases for the pseudo-chest CT images, and the maximum error was 3 slices.

The Dice values of the proposed method and the conventional method [13] in the 2D sagittal-oblique images are shown in Table 2. The mean Dice value of the proposed method is 0.863, while the mean Dice value of the conventional method is 0.91. Fig. 5(a) shows the sagittal-oblique image of the torso CT images, and Fig. 5(b) shows the sagittal-oblique image of the segmentation result of the proposed method. The Dice value of the segmentation result in the 2D image shown in Fig. 5(b) is 0.973. The segmentation accuracy of the proposed method was not as good as that of the conventional method, although the comparison was not made under the same conditions because the positions of the arms in the CT images were different between the two methods.

TABLE 2. Evaluation results in the sagittal-oblique image.

Method	Dice	
	Mean	SD
Proposed method	0.863	0.158
Conventional method [13]	0.91	0.096

Evaluation results of supraspinatus muscle segmentation by the proposed method in the 2D sagittal-oblique images. The segmentation accuracy of the conventional method [13] is also shown.

The mean Jaccard index of the 3D segmentation results obtained by applying the proposed method to the torso CT images was 0.797, but that obtained by applying the model-based method was 0.491 [9]. Therefore, the with-localization method outperformed the conventional method of 3D supraspinatus muscle segmentation.

V. DISCUSSION

We propose a supraspinatus muscle segmentation method that localizes the muscle based on the scapula position and inputs the axial CT images from the top of the images to the bottom of the scapula to U-Net for segmentation even when the field of view of the CT images differ owing to the localization of the muscle. Therefore, we consider that the segmentation accuracies were similar, regardless of the field of view of the input images. On the other hand, when we applied the without-localization method, which inputs all slices of the pseudo-chest CT images to U-Net, the mean Dice value of the segmentation results was lower than that

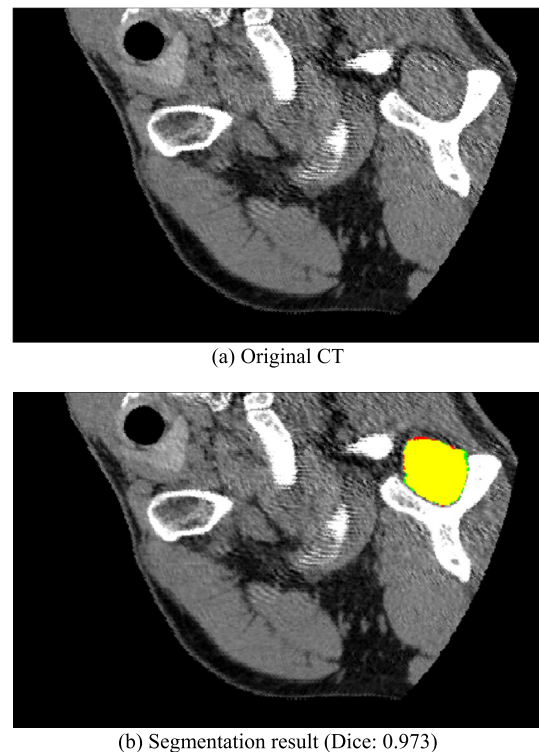


FIGURE 5. Sagittal-oblique images of the original CT (a) and the segmentation result (b). The yellow area represents the overlapped area between the segmentation result and the ground truth, the red area represents the over-extracted area, and the green area represents the under-extracted area.

obtained when applying the with-localization method. Furthermore, all pixels were segmented as the background in all cases of the segmentation results in the without-localization method (mean Dice value: 0.000). Fig. 6 shows the loss values for the training data during U-Net training for supraspinatus muscle segmentation. Fig. 6(a) shows the loss values for the with-localization method, (b) shows the loss values when the without-localization method was applied to the torso CT images, and (c) shows the loss values when the without-localization method was applied to the pseudo-chest CT images. The loss value of the with-localization method converges to a value close to the minimum value of -1 . Moreover, when the without-localization method was applied to the pseudo-chest CT images, the loss value converged to a value close to that of the with-localization method but higher. On the other hand, when the without-localization method was applied to the torso CT images, the loss value decreased to -0.3 , then increased, and finally converged to -0.2 . As described above, although the same parameters were used in all experiments, there were differences in the changes of loss values during training depending on the range of input CT images. Table 3 shows the mean number of input axial images of U-Net for supraspinatus muscle segmentation in the with-localization and without-localization method and the axial images in which the supraspinatus muscle is included for each case. In deep learning, when the number of pixels

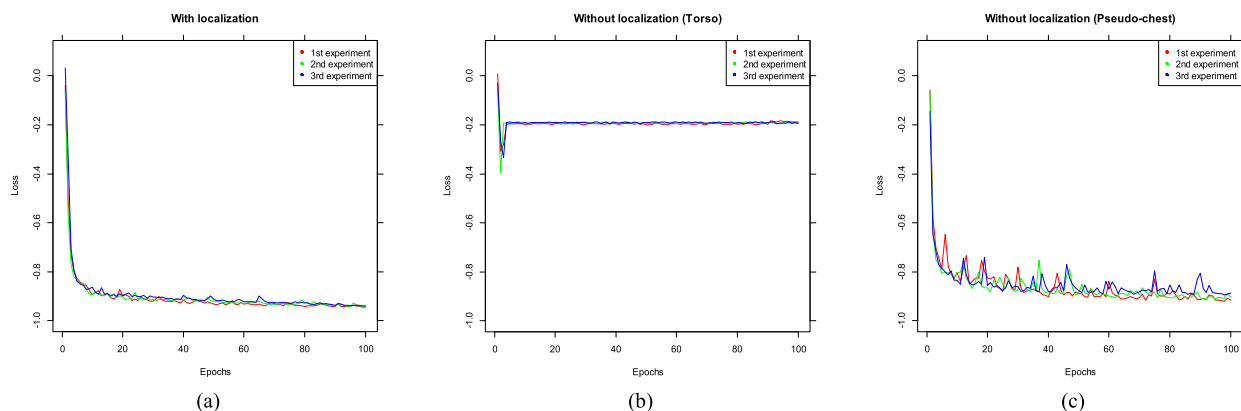


FIGURE 6. The loss value for training data during the training of U-Net for supraspinatus muscle segmentation. (a) The loss value for the with-localization method. (b) The loss value when the without-localization method was applied to the torso CT images. (c) The loss value when the without-localization method was applied to the pseudo-chest CT images.

TABLE 3. Mean number of input axial images.

Image	(a) Number of axial images [slices/case]	(b) Number of axial images included the supraspinatus muscle [slices/case]	Ratio of (b) to (a) [%]
Torso CT	953.9	94.9	9.9
Pseudo-chest CT	415.3	94.9	22.9
Cropped images of the region around the supraspinatus muscle	241.6	94.9	39.3

The mean number of the input axial images of the U-Net for supraspinatus muscle segmentation and the axial images including the supraspinatus muscle. The number of images is the mean number of images for each case. The cropped image of the region around the supraspinatus muscle has the highest ratio of (b) to (a) compared with the torso and the pseudo-chest CT images. CT, computed tomography.

in the target and background regions are unbalanced, the segmentation results may be inaccurate [20]. Therefore, we considered that the with-localization method achieves higher segmentation accuracy than the without method because the number of axial images is the lowest in the input image of U-Net.

In the experiment, the without-localization method segmented all pixels as background in the supraspinatus muscle segmentation using all axial images of the torso CT images as input. On the other hand, the MAE of the proposed method was 0.433 for the muscle localization by scapula segmentation using all axial images of torso CT images as input. These results show that although the range of the input image was the same, there were cases where segmentation was successful and cases where it failed depending on the target region. The reason for this phenomenon is considered to be the difference between the number of slices that include the scapula and that include the supraspinatus muscle in the torso CT images. This is the same cause of the difference in segmentation accuracy when the without-localization method was applied to torso CT images and pseudo-chest CT images. Since the upper end of the scapula is not included in the torso CT images used in this study, the mean number of slices including the scapula is 241.6, which is the same as that of the cropped image in Table 3. Moreover, the mean number

of slices of the torso CT images is 953.9 slices. Therefore, the ratio of the number of slices including the scapula to the total number of slices in the torso CT images is 25.3%. This value is higher than the 22.9% ratio of the number of slices including the supraspinatus muscle to the total number of slices in the input CT images for the muscle segmentation in the pseudo-chest CT images, where successful segmentation results were obtained. Thus, we consider that the scapula segmentation using U-Net in the torso CT images was successful. On the other hand, the ratio of the number of slices including the supraspinatus muscle to the total number of slices in the input CT images for the muscle segmentation in the torso CT images that failed segmentation was as low as 9.9%. This difference in the ratio of the number of slices including the target region to the total number of slices in the input images is considered to have caused the difference in the segmentation results.

The mean Dice value in the sagittal-oblique images of the supraspinatus muscle segmentation results by the proposed method was 0.863. This value is lower than that of results by the conventional method (0.91). Therefore, when evaluating only the sagittal-oblique section, it is better to perform segmentation using only that section as in the conventional method, instead of obtaining a virtual sagittal-oblique section from 3D segmentation results as in

the proposed method. On the other hand, the evaluation of the supraspinatus muscle as a 3D volume can only be achieved by the proposed method. In addition, in the conventional method, the sagittal-oblique image to be input to the modified U-Net is determined manually, but in the proposed method, the segmentation of the supraspinatus muscle can be achieved fully automatically from the input of the torso CT image. Thus, the proposed method is the first to achieve 3D supraspinatus muscle segmentation. Furthermore, the proposed method achieves robust supraspinatus segmentation for different imaging ranges by localization, while being fully automatic.

There are several limitations of this study. First, there is a need for further studies on the image-cropping method for supraspinatus muscle localization. In the proposed method, the CT images were cropped in the axial direction until the bottom slice of the scapula. The muscle was then segmented using the U-Net in the cropped CT image. However, since the attachment site of the supraspinatus muscle to the scapula (supraspinatus fossa) is located close to the upper edge of the scapula, the range in which the muscle appears can be narrowed much further than in the proposed method by localizing the muscle based on its origin and insertion. The recognition of origin and insertion has been addressed for other skeletal muscles [21]. Therefore, the automatic segmentation of the supraspinatus muscle based on its origin and insertion remains a challenge. Second, this study did not examine the effectiveness of bone-based localization in the segmentation of other skeletal muscles. In particular, for segmentation of those with a small cross-sectional area in the axial plane, it may be more effective to combine cropping in the axial direction, as in the proposed method, with cropping in the sagittal and coronal directions in bone-based localization. Therefore, it is necessary to select the image-cropping method according to the position and shape of the target skeletal muscle. Third, the accuracy of bone segmentation in bone-based localization affects the accuracy of muscle segmentation. In the proposed method, the CT images are cropped to the edge of the bone using the bone segmentation result, and then input to U-Net for muscle segmentation. Thus, if the accuracy of the bone segmentation is low, it may not be possible to crop the image to include the whole muscle. In this study, the supraspinatus muscle, which is a muscle located on the upper part of the scapula, was targeted, and the crop range was from the top of the CT images to the bottom of the scapula. Therefore, we consider that the segmentation accuracy of the scapula does not affect the segmentation accuracy of the supraspinatus muscle. However, if the proposed method is applied to other muscles, the segmentation accuracy of the bone may become important. For example, if the infraspinatus muscle, which is located in the lower part of the scapula, is to be segmented, a part of the muscle may be out of the range from the top of the CT images to the bottom of the scapula if the scapula segmentation accuracy is low. Therefore, depending on the position of the target muscle in relation to the bone, it is necessary to consider a more accurate

bone-based slice estimation method or an expansion of the crop range.

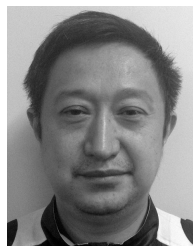
VI. CONCLUSION

We proposed a supraspinatus muscle segmentation method that uses bone-based localization in torso CT images. In this method, the muscle was localized based on the segmentation result of the scapula, to which the muscle attaches, by U-Net. We conclude that this method provides a higher accuracy for supraspinatus muscle segmentation than a method performed without localization. The proposed method is an extension of the 2D cross-sectional localization based on bone location and skeletal muscle segmentation [7] to 3D localization and skeletal muscle segmentation. However, further studies on image cropping for supraspinatus muscle localization, such as cropping based on origin and insertion and in the sagittal and coronal directions, are needed.

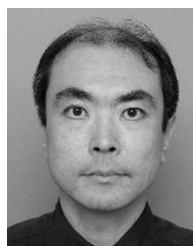
REFERENCES

- [1] J. D. Wening, R. F. Hollis, R. E. Hughes, and J. E. Kuhn, "Quantitative morphology of full thickness rotator cuff tears," *Clin. Anatomy*, vol. 15, no. 1, pp. 18–22, Jan. 2002, doi: [10.1002/ca.1086](https://doi.org/10.1002/ca.1086).
- [2] L. H. Frich, L. R. Fernandes, H. D. Schröder, E. K. Hejbøl, P. V. Nielsen, P. H. Jørgensen, A. Stensballe, and K. L. Lambertsen, "The inflammatory response of the supraspinatus muscle in rotator cuff tear conditions," *J. Shoulder Elbow Surg.*, vol. 30, no. 6, pp. e261–e275, Jun. 2021, doi: [10.1016/j.jse.2020.08.028](https://doi.org/10.1016/j.jse.2020.08.028).
- [3] A. R. Jensen, A. J. Taylor, and J. Sanchez-Sotelo, "Factors influencing the reparability and healing rates of rotator cuff tears," *Current Rev. Musculoskeletal Med.*, vol. 13, no. 5, pp. 572–583, Oct. 2020, doi: [10.1007/s12178-020-09660-w](https://doi.org/10.1007/s12178-020-09660-w).
- [4] X. Feng, T. Todd, C. R. Lintzenich, J. Ding, J. J. Carr, Y. Ge, J. D. Browne, S. B. Kritchevsky, and S. G. Butler, "Aging-related geniohyoid muscle atrophy is related to aspiration status in healthy older adults," *J. Gerontol. A, Biol. Sci. Med. Sci.*, vol. 68, no. 7, pp. 853–860, Jul. 2013, doi: [10.1093/geron/gls225](https://doi.org/10.1093/geron/gls225).
- [5] K. Engelke, O. Museyko, L. Wang, and J.-D. Laredo, "Quantitative analysis of skeletal muscle by computed tomography imaging—State of the art," *J. Orthopaedic Transl.*, vol. 15, pp. 91–103, Oct. 2018, doi: [10.1016/j.jot.2018.10.004](https://doi.org/10.1016/j.jot.2018.10.004).
- [6] G. Iazzetti, "The skeletal muscles," in *Human Anatomy*, Surrey, Bengaluru, India: TAJ Books International LLP, 2009, pp. 33–35.
- [7] J. Castiglione, E. Somasundaram, L. A. Gilligan, A. T. Trout, and S. Brady, "Automated segmentation of abdominal skeletal muscle on pediatric CT scans using deep learning," *Radiol. Artif. Intell.*, vol. 3, no. 2, pp. 1–9, Mar. 2021, doi: [10.1148/ryai.2021200130](https://doi.org/10.1148/ryai.2021200130).
- [8] H. Fujita, T. Hara, X. Zhou, C. Muramatsu, N. Kamiya, M. Zhang, D. Fukuoka, Y. Hatanaka, T. Matsubara, A. Teramoto, Y. Uchiyama, H. Chen, and H. Hoshi, "Model construction for computational anatomy: Progress overview FY2009-FY2013," in *Proc. 5th Int. Symp. Project Comput. Anatomy*, Mar. 2014, pp. 25–35.
- [9] T. Katafuchi, N. Kamiya, X. Zhou, H. Kato, K. Azuma, C. Muramatsu, T. Hara, T. Miyoshi, M. Matsuo, and H. Fujita, "Improvement of supraspinatus muscle recognition methods based on the anatomical features on the scapula in torso CT image," in *Proc. Int. Forum Med. Imag. Asia (IFMIA)*, Jan. 2017, pp. 315–316.
- [10] N. Kamiya, "Deep learning technique for musculoskeletal analysis," in *Deep Learning in Medical Image Analysis: Challenges and Application*, G. Lee and H. Fujita, Eds. Cham, Switzerland: Springer, 2020, pp. 165–176.
- [11] N. Kamiya, J. Li, M. Kume, H. Fujita, D. Shen, and G. Zheng, "Fully automatic segmentation of paraspinal muscles from 3D torso CT images via multi-scale iterative random forest classifications," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 11, pp. 1697–1706, Nov. 2018, doi: [10.1007/s11548-018-1852-1](https://doi.org/10.1007/s11548-018-1852-1).

- [12] Y. Wakamatsu, N. Kamiya, X. Zhou, T. Hara, and H. Fujita, "Relationship between number of annotations and accuracy in segmentation of the erector spinae muscle using Bayesian U-Net in torso CT images," in *Proc. Int. Forum Med. Imag. Asia (IFMIA)*, Jan. 2021, pp. 1–4, doi: [10.1117/12.2590780](https://doi.org/10.1117/12.2590780).
- [13] E. Taghizadeh, O. Truffer, F. Becce, S. Eminian, S. Gidoïn, A. Terrier, A. Farron, and P. Büchler, "Deep learning for the rapid automatic quantification and characterization of rotator cuff muscle degeneration from shoulder CT datasets," *Eur. Radiol.*, vol. 31, no. 1, pp. 181–190, Jul. 2020, doi: [10.1007/s00330-020-07070-7](https://doi.org/10.1007/s00330-020-07070-7).
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent (MICCAI)*, Munich, Germany, Oct. 2015, pp. 234–241, doi: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [15] T. Liu, Y. Tian, S. Zhao, X. Huang, and Q. Wang, "Automatic whole heart segmentation using a two-stage U-Net framework and an adaptive threshold window," *IEEE Access*, vol. 7, pp. 83628–83636, 2019, doi: [10.1109/ACCESS.2019.2923318](https://doi.org/10.1109/ACCESS.2019.2923318).
- [16] D. P. Kingma and J. L. Ba, "ADAM: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, May 2015, pp. 1–5. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [17] TensorFlow. Google. Accessed: Feb. 20, 2019. [Online]. Available: <https://www.tensorflow.org>
- [18] FlowKeras. Chollet. Accessed: Feb. 20, 2019. [Online]. Available: <https://keras.io>
- [19] A. Terrier, J. Ston, A. Dewarrat, F. Becce, and A. Farron, "A semi-automated quantitative CT method for measuring rotator cuff muscle degeneration in shoulders with primary osteoarthritis," *Orthopaedics Traumatol., Surg. Res.*, vol. 103, no. 2, pp. 151–157, Apr. 2017, doi: [10.1016/j.otsr.2016.12.006](https://doi.org/10.1016/j.otsr.2016.12.006).
- [20] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: Achievements and challenges," *J. Digit. Imag.*, vol. 32, no. 4, pp. 582–596, 2019, doi: [10.1007/s10278-019-00227-x](https://doi.org/10.1007/s10278-019-00227-x).
- [21] N. Kamiya, M. Kume, G. Zheng, X. Zhou, H. Kato, H. Chen, C. Muramatsu, T. Hara, T. Miyoshi, M. Matsuo, and H. Fujita, "Automated recognition of erector spinae muscles and their skeletal attachment region via deep learning in torso CT images," in *Proc. 6th MICCAI Workshop Comput. Methods Clin. Appl. Musculoskeletal Imag.*, Sep. 2018, pp. 1–10, doi: [10.1007/978-3-030-11166-3_1](https://doi.org/10.1007/978-3-030-11166-3_1).



XIANGRONG ZHOU received the M.S. and Ph.D. degrees in information engineering from Nagoya University, Japan, in 1997 and 2000, respectively. From 2000 to 2002, he continued his research in medical image processing as a Postdoctoral Researcher at Gifu University, Japan, where he is currently an Associate Professor at the Department of Electrical, Electronic and Computer Engineering. His research interests include medical image analysis, medical image visualization, and pattern recognition.



HIROKI KATO graduated from the School of Medicine, Gifu University, in 1999. He received the Ph.D. degree from the Department of Radiology, Gifu University, Japan, in 2007. In 1999 and 2000, he became a Resident at the Department of Radiology, Gifu University Hospital, and Jichi Medical University Hospital, respectively. In 2002 and 2004, he became a Clinical Fellow and Assistant Professor at the Department of Radiology, Gifu University Hospital, respectively. He is currently an Associate Professor at the Department of Radiology, Graduate School of Medicine, Gifu University. His main research interests include diagnostic imaging of computed tomography and magnetic resonance imaging. His specialized fields include head and neck imaging, genitourinary imaging, and musculoskeletal imaging.



TAKESHI HARA received the B.S. and M.S. degrees from the Faculty of Engineering, Gifu University, Japan, in 1994 and 1995, respectively, and the Ph.D. degree from Gifu University, in 2000. In 1995, he was a Research Assistant at the Faculty of Engineering, Gifu University, where he became an Associate Professor, in 2001 and 2017. He was also an Associate Professor at the Department of Intelligent Image Information, Gifu University. From 2008 to 2009, he was a Visiting Associate Professor at the Department of Radiology, The University of Chicago. Since 2019, he has been a Professor at the Faculty of Engineering and a Chair of the Artificial Intelligence Research Promotion Center, Medical Research Department, Gifu University.



HIROSHI FUJITA received the B.S. and M.S. degrees in electrical engineering from Gifu University, Japan, in 1976 and 1978, respectively, and the Ph.D. degree from Nagoya University, in 1983. In 1978, he became a Research Associate at the Gifu National College of Technology, where he also became an Associate Professor, in 1986. From 1983 to 1986, he was a Visiting Researcher at the K. Rossmann Radiologic Image Laboratory, The University of Chicago. He became an Associate Professor at the Faculty of Engineering, Gifu University, in 1991, where he also became a Professor, in 1995. Since 2002, he has been a Professor and a Chair of Intelligent Image Information at the Graduate School of Medicine, Gifu University. He is currently a Research Professor at Gifu University and also a Visiting Professor at Zhengzhou University, China.



YUICHI WAKAMATSU received the B.S. degree in information science and technology from Aichi Prefectural University, Japan, in 2020, where he is currently pursuing the M.S. degree. His research interest includes medical image processing.



NAOKI KAMIYA (Member, IEEE) received the B.S. degree in information science from Gifu University, Japan, in 2005, and the M.S. and Ph.D. degrees in regeneration and advanced medical sciences from the Graduate School of Medicine, Gifu University, in 2007 and 2011, respectively. In 2010, he became an Assistant Professor at the Department of Information and Computer Engineering, Toyota National College of Technology. He became an Assistant Professor at the School of Information Science and Technology, Aichi Prefectural University, in 2015, where he also became a Lecturer, in 2017. He was a Visiting Scholar at the University of Bern Institute for Surgical Technology and Biomechanics, in 2017. He is currently an Associate Professor at the School of Information Science and Technology, Aichi Prefectural University.