Surface Muscle Segmentation Using 3D U-Net Based on Selective Voxel Patch Generation in Whole-Body CT Images

Naoki Kamiya 1,*, Ami Oshima 1,*, Xiangrong Zhou 2, Hiroki Kato 3, Takeshi Hara 2, Toshiharu Miyoshi 4, Masayuki Matsuo 3 and Hiroshi Fujita 2

1 Graduate School of Information Science and Technology, Aichi Prefectural University, Aichi 480-1198, Japan
2 Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering, Gifu University, Gifu 501-1194, Japan; zxr@gifu-u.ac.jp (X.Z.); hara@fjt.info.gifu-u.ac.jp (T.H.); fujita@fjt.info.gifu-u.ac.jp (H.F.)
3 Department of Radiology, Graduate School of Medicine, Gifu University, Gifu 501-1194, Japan; hkato@gifu-u.ac.jp (H.K.); matsuo_m@gifu-u.ac.jp (M.M.)
4 Radiology Service, Gifu University Hospital, Gifu 501-1194, Japan; t-miyosi@gifu-u.ac.jp
* Correspondence: n-kamiya@ist.aichi-pu.ac.jp (N.K.); im192004@cis.aichi-pu.ac.jp (A.O.)

Received: 20 May 2020; Accepted: 26 June 2020; Published: 28 June 2020

Abstract: This study aimed to develop and validate an automated segmentation method for surface muscles using a three-dimensional (3D) U-Net based on selective voxel patches from whole-body computed tomography (CT) images. Our method defined a voxel patch (VP) as the input images, which consisted of 56 slices selected at equal intervals from the whole slices. In training, one VP was used for each case. In the test, multiple VPs were created according to the number of slices in the test case. Segmentation was then performed for each VP and the results of each VP merged. The proposed method achieved a segmentation accuracy mean dice coefficient of 0.900 for 8 cases. Although challenges remain in muscles adjacent to visceral organs and in small muscle areas, VP is useful for surface muscle segmentation using whole-body CT images with limited annotation data. The limitation of our study is that it is limited to cases of muscular disease with atrophy. Future studies should address whether the proposed method is effective for other modalities or using data with different imaging ranges.

Keywords: surface muscle; segmentation; voxel patch; whole-body CT; 3D U-Net

1. Introduction

Surface muscle segmentation in whole-body computed tomography (CT) images is essential not only in clinical applications, but also for the anatomical understanding of the human body. The clinical application of surface muscle segmentation includes muscle analysis for amyotrophic lateral sclerosis (ALS), which is an intractable disease characterized by muscular atrophy [1]. A definitive diagnosis of ALS is challenging because effective diagnoses have not been established; exclusion diagnosis is the primary method used [2]. Therefore, the early detection and early differentiation of diseases that cause muscle atrophy are essential. Image feature analysis of skeletal muscle regions has been attempted as an initial study of image discrimination for ALS [1]. In skeletal muscle feature analysis, the texture of skeletal muscles and surface muscle regions of the limbs are analyzed. Analysis of skeletal muscles is limited to the extremities, and it is unknown where the symptoms of ALS appear in the whole body.

The segmentation of surface muscles is also crucial for human anatomical structure recognition by computers. Segmentation studies of anatomical structures in various regions of the human body using CT images have been performed. Zhou et al. used a fully convolutional network (FCN) [3] and segmented a total of 19 regions from 17 organs and 2 regions of interest from torso CT images with an
average matching rate of 87.9% [4]. Additionally, research on the musculoskeletal region has been conducted. First, Klein and colleagues used a two-dimensional (2D) U-Net [5], an effective network for object segmentation, to automatically segment the skeleton from axial sections of whole-body CT slice images at a low dose [6]. As a result, an average dice coefficient (DC) of 0.95 ± 0.01 for the 2D method and 0.92 ± 0.01 for the 2.5D method were reported.

For skeletal muscle segmentation using CT images, model-based methods have been proposed [7]. In a study by Kamiya et al., a three-dimensional (3D) shape model was used to identify the psoas major muscle, and in recent years, skeletal muscle segmentation using deep learning has been performed [8–12]. Hiasa and Sakamoto proposed a site-specific segmentation method for a total of 19 skeletal muscles, in the thigh and hip regions, using deep learning [8,9]. Hashimoto et al. used a 2D U-Net to identify the psoas major muscle in a low-dose CT [10], and Lee et al. proposed a method for identifying skeletal muscle and fat from 2D slices in abdominal CT images using an improved fully convolutional network (FCN) [11]. Kamiya et al. identified the erector spinae muscle using an FCN and showed that large-scale skeletal muscle segmentation by deep learning is possible [12].

Notably, in a CT image, the skeletal muscle region, including the surface muscle, has a similar distribution of CT values to the visceral organs. Compared to muscle segmentation using magnetic-resonance imaging (MRI), this is a challenging task. These are among the primary reasons that surface muscle segmentation for the whole body has not been performed (other than that by the authors of this study in a previous study using a 2D U-Net) [13]. As a result, the mean DC, which is the segmentation match rate with the ground truth, was 0.749 (standard deviation [SD]: 0.032) in three experiments for five test cases. Since the CT images used in this study are composed of 3D volumes, we propose a 3D U-Net as the 3D surface muscle segmentation method, which maintains continuity between slices. However, the 3D-based segmentation method using deep learning requires large graphics processing unit (GPU) memory. Therefore, in order to perform 3D deep learning of the skeletal muscles with limited GPU memory, it is necessary to select and limit the image data size properly.

Another task involved with the segmentation of surface muscles is gathering the annotation data required for machine learning. Creating correctly labeled annotation data requires more than 1300 slices per case in a whole-body CT image, meaning it takes an exceptionally long time to generate a correct image from all slices. Therefore, in the present study, one set of image pairs was created by selecting 56 slices from the whole body per case, which is defined as a voxel patch (VP). This is the method for automatic surface muscle segmentation used in this study. For example, one VP per case is used for training, and N sets of VPs are used for the test according to the total number of slices in each test case, which enables automatic segmentation of the surface muscles from limited annotation data.

2. Materials and Methods

2.1. Methodology Overview

Figure 1 shows the methodological flow. In this study, surface muscles were obtained by the following three steps using whole-body CT images as the input. First, a training VP for each training case and test VPs were created by reducing the input image and selecting slices. Next, using the test VP, the surface muscles were identified by a 3D U-Net. Finally, slices were combined from each test VP, and the image was enlarged in order to obtain the original size of the surface muscle segmentation results.

In the first step, after reducing the input images, 56 slices were selected at equal intervals from all of the slices in each training case, as a “training VP.” These 56 slices were selected due to the limitation of the memory capacity of the GPU, and the details were described in the Method. For the “test VP,” 56 slices were selected in a similar method as the training VP, and N sets of test VPs were created by sequentially offsetting the slice selected for the first slice. We segmented the surface muscle using a 3D U-Net under the training VP in the second step. Furthermore, the test VPs were predicted.
third step, we combined the predicted slices from the obtained segmentation results, and obtained the surface muscle segmentation results. The following sections describe the details of each process.

The experimental GPU environment in this research used four Tesla V100 (32 GB) GPUs and has a total memory of 128 GB per GPU (NVIDIA Corporation, California, USA). Keras (ver. 2.1.2) [15] and Tensorflow-GPU (ver. 1.13.1) were the software environments used.

![Figure 1. Schematic diagram of the proposed method. The proposed method consists of three steps. First, a voxel patch is created by image reduction and slice selection. Next, a three-dimensional U-Net is used to segment the surface muscle in each test voxel patch. Finally, slices are combined from each test voxel patch, and the image is enlarged in order to obtain the original size of the surface muscle segmentation results. VP, voxel path; 3D, three-dimensional.](image)

### 2.2. Voxel Patch Generation by Image Reduction and Slice Selection

In this section, a VP composed of limited slices was created by reducing the input image and selecting 56 slices. A VP is a voxel set that summarizes information about the body axis. Generally, a deep convolutional neural network (DCNN) requires a large number of training images, but it is challenging to create a large number of training images for skeletal muscles, especially for surface muscles, which have substantial individual differences. Therefore, we propose a VP-based method. In training, one VP is generated for each case. In the test, multiple VPs are created according to the number of slices in the test case. Then, segmentation is performed for each VP, and finally, the results of each VP are merged. The details of how to generate training and test VPs are described below.

The input images were resized from 512 × 512 pixels to 256 × 256 pixels in an axial cross-section with nearest-neighbor interpolation. Next, 56 slices were selected at equal intervals from each of the whole-body slices. Our GPU environment is four Tesla V100 (32 GB), and is not necessarily poor, but the VP size that can be generated depends on the total GPU memory capacity (128 GB). In this study, the VP size was set to maintain the resolution of the cross-section while maintaining the number of slices in the body axis direction in the whole-body CT images. When using the original image size (512 × 512 pixels) without resizing, the number of images in the body axis direction for which a VP can be constructed is only 14 slices. In a whole-body CT image consisting of approximately 1500 images per case, it is thought that the number of images will be too small for training. Therefore, we defined the VP size as 256 × 256 × 56 voxels.
The VP for training used 56 slices selected at equal intervals from the first slice in the axial direction of the whole-body CT image; one VP per case was used for training. The selected slices vary in anatomical spatial position between training cases, so the effect of increasing the variation in training is expected.

In our method, the above-mentioned training VP consisted of 56 slices per case. Similarly, the test VP was also constructed with 56 slices per VP, so using VPs of N sets gives a segmentation result for all slices in a test case (Figure 1). The test VPs consisted of 56 slices selected at equal intervals in the same manner as the training VPs. Moreover, N sets of VPs were created by selecting the first slice without consecutive duplication from the axial cross-section in the whole-body CT-image. However, this slice selection method has a problem: the test VP cannot be constructed when the number of slices is not a multiple of 56. At that time, the slices used in other test VP sets could be selected at equal intervals from the head again, and the test VP is composed of 56 slices. Note that the slices are rearranged craniocaudally in one VP set. As described above, all slices can be covered by combining all test VPs; therefore, surface muscle segmentation can be performed.

2.3. Automatic Segmentation of Surface Muscles by 3D U-Net Using Voxel Patches

This study used a 3D U-Net. A 2D U-Net [6], which was proposed by Ronneberger et al., won the Dental X-ray Image Segmentation and Cell Tracking Challenges. The authors of this study had previously performed segmentation of the surface muscle with 2D U-Net using all slices, and the mean DC for five test cases was 0.749 (SD: 0.032) in three experiments [13].

On the other hand, continuity by slice is not considered. The surface muscles of this segmentation target are composed of 3D volume data in CT images. Therefore, we used a 3D U-Net, which is proposed by Çiçek et al. [12], and trained under the consecutiveness between slices. Additionally, the segmentation of multiple organs by Zhou et al. compared the segmentation accuracy of 2D images with 3D images as the input images [16]. In Zhou’s method, multiple organs are first detected with a bounding box. Then, segmentation was used to detect the limits of the 3D images. As a result, the intersection over union of segmentation accuracies were 79% and 67% in the 3D and 2D DCNNs, respectively. Thus, that study showed it is useful for the segmentation approach of a DCNN to use 3D CT images.

For these reasons, in this study, a 3D based training was also carried out for the segmentation of surface muscles (Figure 2). Voxel Patch (VP) is input, and the segmentation result of the surface muscle is obtained as an output. The architecture consists of an encoder and decoder. In the encoder, the $3 \times 3 \times 3$ convolution, batch normalization [17], and rectified linear unit (ReLU) as an activation function are repeated eight times. Moreover, max pooling is used three times for down-sampling. After concatenating the corresponding cropped feature map from the encoding layers, $3 \times 3 \times 3$ convolutions with batch normalization and ReLU are applied two times in the encoder. Two convolutions are performed with a drop out of 50% because of over-training. Furthermore, three $2 \times 2 \times 2$ up-convolutions and ReLU are used for up-sampling. In the last layer, we obtained the segmentation results through $1 \times 1 \times 1$ convolution and sigmoid.

The number of epochs and the batch sizes was 600 and 1, respectively, which could be executed with the memory size (128GByte) of the experiment environment. We defined the learning rate as $1 \times 10^{-4}$, and the adaptive moment estimation [18] was used as an optimization function. The loss function is (1–[the mean of the DC]).
Individuals included in the study. The images were 512×512 pixels with 1305–1726 slices per scan. The spatial resolution was 0.625 mm, and the density was 0.625 mm, and the density resolution was 12 bits. These data were scanned by using a recommended protocol (tube voltage 120 KVp with a current of Auto mA) for detailed examination in our hospital. In this study, we defined surface muscles as skeletal muscles that exist outside of the body cavity and are visible on non-contrast CT images taken under the above imaging conditions. Ground truth was created by Ms. Ami Oshima, who modified each result of the surface muscle segmentation using a conventional method [1] under the advice of an anatomist. We used the graph cut-based interactive method implemented in the common software platform “PLUTO” [19]. It took about 45 min per case to modify the VP label used for training. On the other hand, test case correction took about 30 h per case for all slices, which was used as ground truth.

Assuming the automatically segmented set of voxels as AS and the manually defined ground truth as GT, we used both volume overlap metrics to evaluate the present method. We computed the

![Figure 2. Representation of a three-dimensional U-Net architecture and details. Voxel Patch (VP) is input, and the segmentation result of the surface muscle is obtained as an output. The blue box is a feature map with the number of channels shown at the top. Details of the encoder and decoder are described above, and other parameters are described in the main text. VP, voxel patch; 3D, three-dimensional; Conv., convolution; BN, batch normalization; ReLU, rectified linear unit; Concat., concatenation.](image_url)
DCs [20], which quantifies the match of two sets by normalizing the size of their intersection over the average of their sizes. The DC is defined as follows:

$$DC = \frac{2|AS \cap GT|}{|AS| + |GT|}$$  \hspace{1cm} (1)

where the operator $|\cdot|$ returns the number of voxels contained in a region.

We randomly selected training (30 cases) and validation (3 cases), performed tests on eight cases three times, and evaluated the average DC value.

3. Results

Table 1 shows the DC of three experiments in eight test cases. The mean DC was 0.900 (SD: 0.022) for surface muscle segmentation. Here, SD is obtained from the DC values of all experimental results, and the results obtained for each of the three tests. The segmentation results of Case E (Figure 3) obtained the highest DC (0.937) in the first experiment and comparison with our 2D U-Net based approach [13]. Figure 4 shows an example of the characteristic segmentation results of the surface muscles in each section. Blood vessels were over-extracted in the surface muscles of the limbs and areas not adjacent to the organs, but no major mis-extraction was observed in other tissues, and the segmentation was successful. On the other hand, over-extraction of the mammary gland region and the abdominal organs adjacent to the surface muscles and unextracted in a small region of the rectus abdominis and oblique muscles were observed. Despite those remaining issues, the conventional 2D U-Net for surface muscle segmentation [13] has a DC of 0.749, whereas the proposed method had a higher DC value (0.900).

Table 1. Dice coefficients of three experiments in eight test cases (mean: 0.900, SD: 0.022, across all three experiments) with no significant difference in DC value between cases.

| Case | 1st DC Value in Each Experiment | 2nd | 3rd | Mean |
|------|---------------------------------|-----|-----|------|
| A    | 0.887                           | 0.883 | 0.882 | 0.884 |
| B    | 0.896                           | 0.890 | 0.891 | 0.892 |
| C    | 0.872                           | 0.859 | 0.859 | 0.864 |
| D    | 0.920                           | 0.921 | 0.919 | 0.920 |
| E    | 0.937                           | 0.933 | 0.934 | 0.935 |
| F    | 0.888                           | 0.886 | 0.886 | 0.887 |
| G    | 0.918                           | 0.917 | 0.917 | 0.917 |
| H    | 0.908                           | 0.902 | 0.905 | 0.905 |
| Average | 0.903 | 0.899 | 0.899 | 0.900 |

2 SD: 0.020 0.023 0.022

1 DC, dice coefficient. 2 SD, standard deviation.
4. Discussion

In this study, automatic segmentation of surface muscles by a 3D U-Net was performed using training VPs composed of 56 slices selected at equal intervals in each case from whole-body CT images for the training algorithm.

First, we compared the results of this study with a conventional method [13]. In the conventional method [13], surface muscle segmentation was performed in a 2D U-Net using all slices of the whole body, with an average DC of 0.749. The average DC was 0.900 using the VP composed of 56 slices, and a high segmentation rate was obtained with the proposed method. In the VP-based method,
the 3D U-Net could extract relative surface muscle features between slices even with a limited number of training slices.

**Figure 4.** Slice example of the segmentation results in each section [DC (dice coefficient) in case, DC in slice]. The left column is a good result and the right column is a bad result. Yellow, red, and green show the overlapped area between ground truth and the result, the over-extracted, and unextracted regions, respectively. In (a,e), the segmentation results were accurate for areas not adjacent to the limbs and abdominal organs. Over-extraction occurred in the mammary gland area (b) and the area where the surface muscle was adjacent to the abdominal organ (c). Un-extraction occurred in the surface muscles of the abdomen where the recognition target was small (d). DC, dice coefficient.

The point of this study was to show that for algorithm training, 56 images selected at equal intervals, in each case, from the whole-body CT images obtained a higher average DC value without using all slices. Despite the random selection of training and validation data in three experiments,
the mean DC was 0.900 and (SD: 0.022). This can be considered to be a very stable result. In this study, 56 slices were selected from each case at equal intervals for the VP creation. Furthermore, the number of slices and the imaging range of test images are different in each case. Therefore, by selecting slices at equal intervals, we considered that each VP is composed of surface muscles at various anatomical positions, trained with many surface muscle variations, and obtained high generalization performance for the test VP. In other words, since the number of slices in the image database is 1305–1726, when 56 slices are selected from each case, it can be said that surface muscle segmentation of the whole body is performed with the limited selection of slices of (3.2–4.3% of the whole). In this way, the method of selecting slices at equal intervals during VP construction is also an effective method for the generation of annotation images in the automatic segmentation of surface muscles.

There are some limitations to our study. First, the number of images when creating a VP was fixed at 56. For the creation of the VP, we selected 56 because it was the maximum number of slices that could be learned by the GPU in this experimental environment. Thus, we did not consider another number of slices. Therefore, for the segmentation of the surface muscles using VP, determination of the number of slices composing the VP remains to be studied. However, it is necessary to consider that changing the number of slices requires the creation of new ground truth labels. In this study, surface muscles were defined as skeletal muscles outside of the body cavity, and whole-body CT images taken with a standard protocol were used. The training increases as the number of slices increases, so improving the DC rate is possible. However, ground truth segmentation of skeletal muscle on non-contrast CT is not easy. It is difficult to increase the ground truth easily because it is hard to strictly distinguish tissues such as microvessels and tendons. On the other hand, if the training VP is composed of less than 56 slices, the number of features that can be extracted and the segmentation result will decrease accordingly. However, by reducing the number of images that make up the training VP, there is an advantage of a reduced cost of creating ground truth labels, and the number of learning cases can be increased. It is necessary to examine the change of the segmentation result and determine the number of slices constituting the training VP required for the surface muscle segmentation when taken all together. Furthermore, even when selecting the same 56 slices, a different slice selection method, based on the anatomical position in each case, could be considered (instead of selecting equal intervals). The reason is that slices required for training are not included in the VP slice from the high-density mammary gland in the over-extracted region or abdominal region that were not extracted (Figure 4). Therefore, it is necessary to consider the slice selection method that constitutes the training VP. However, when selecting slices based on the anatomical position, the variation of the surface muscles in each training VP is reduced, so the generalization performance for surface muscle segmentation in regions other than the selected slice may decrease. Therefore, it is necessary to consider the enhancement of training by augmentation.

Second, all the CT data used in this study were acquired with the same scanner from Gifu University Hospital. It would be interesting to apply our trained model to CT images from other scanners in order to test the inter-scanner robustness. Considering the fact that unlike MR image values, CT values are correlated with tissue attenuation coefficients, we hypothesize that we can directly apply our trained model to CT data acquired from other scanners. Such a hypothesis needs to be verified in our future work. In addition, although there were no artifacts that affected the segmentation results in the present data, there is a paper that mentions the improvement of segmentation accuracy by reducing metal artifacts in thigh muscle segmentation [21], so there is room for further investigation.

Last but not least, the present method was evaluated on CT data collected with a standard clinical protocol. Whether it will work or not on heterogeneous data acquired in clinical routine needs to be further checked in the future. Similarly, since this study was applied only to patients with atrophic muscle diseases, including ALS, it is necessary to verify its applicability to muscle diseases that do not involve atrophy.
Although the above issues remain, the proposed method is useful because the surface muscles can be segmented with high accuracy (approximately 90%) from a training VP composed of only 56 slices selected at equal intervals.

5. Conclusions

In this study, we defined a VP and proposed a surface muscle segmentation method using whole-body CT images. Segmentation with a 3D U-Net was performed using a training VP composed of 56 slices selected at equal intervals from all slices. As a result, the DC showed a segmentation agreement rate between the result and the correct answer of 0.900 on average (mean). Furthermore, we found that the surface muscle segmentation could be efficiently obtained with a limited number of training images with a VP. Future research is needed to examine the number of slices and the slice selection method when creating VPs for erroneous extraction in regions where CT values are similar to those of the surface muscle and where the target region is small.

Author Contributions: Conceptualization, N.K.; methodology, N.K. and A.O.; software, A.O.; validation, N.K. and A.O.; resources, X.Z. and T.H.; data curation, H.K., T.M. and M.M.; project administration, N.K. and H.F.; funding acquisition, H.F. and N.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by a JSPS Grant-in-Aid for Scientific Research on Innovative Areas (Multidisciplinary Computational Anatomy, #26108005 and #17H05301), JAPAN.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kamiya, N.; Ieda, K.; Zhou, X.; Yamada, M.; Kato, H.; Muramatsu, C.; Hara, T.; Miyoshi, T.; Inuzuka, T.; Matsuo, M.; et al. Automated analysis of whole skeletal muscle for muscular atrophy detection of ALS in whole-body CT images: preliminary study. In Proceedings of the SPIE Medical Imaging 2017, Orlando, FL, USA, 11–16 February 2017; pp. 1–6. [CrossRef]

2. National Institute of Neurological Disorders and Stroke. Amyotrophic Lateral Sclerosis (ALS) Fact Sheet. Available online: https://www.ninds.nih.gov/disorders/Patient-Caregiver-Education/Fact-Sheets/Amyotrophic-Lateral-Sclerosis-ALS-Fact-Sheet (accessed on 7 March 2020).

3. Shelhamer, E.; Long, J.; Darrell, T. Fully convolutional networks for semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 640–651. [CrossRef] [PubMed]

4. Zhou, X.; Takayama, R.; Wang, S.; Hara, T.; Fujita, H. Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. Med. Phys. 2017, 44, 5221–5233. [CrossRef] [PubMed]

5. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention 2015, Munich, Germany, 5–9 October 2015; pp. 234–241. [CrossRef]

6. Klein, A.; Warszawski, J.; Hillengaß, J.; Maier-Hein, K.H. Automatic bone segmentation in whole-body CT images. Int. J. Comput. Assist. Radiol. Surg. 2019, 14, 21–29. [CrossRef] [PubMed]

7. Kamiya, N.; Zhou, X.; Chen, H.; Muramatsu, C.; Hara, T.; Yokoyama, R.; Kanematsu, M.; Hoshi, H.; Fujita, H. Automated segmentation of psoas major muscle in X-ray CT images by use of a shape model: Preliminary study. Radiol. Phys. Technol. 2012, 5, 5–14. [CrossRef] [PubMed]

8. Hiasa, Y.; Otake, Y.; Takao, M.; Ogawa, T.; Sugano, N.; Sato, Y. Automated Muscle Segmentation from Clinical CT using Bayesian U-Net for Personalized Musculoskeletal Modeling. IEEE Trans. Med. Imaging 2020, 39, 1030–1040. [CrossRef] [PubMed]

9. Sakamoto, M.; Hiasa, Y.; Otake, Y.; Takao, M.; Suzuki, Y.; Sugano, N.; Sato, Y. Automated segmentation of hip and thigh muscles in metal artifact contaminated CT using CNN. In Proceedings of the International Forum on Medical Imaging in Asia, Singapore, 6–9 January 2019; Volume 11050, pp. 1–6. [CrossRef]

10. Hashimoto, F.; Kakimoto, A.; Ota, N.; Ito, S.; Nishizawa, S. Automated segmentation of 2D low-dose CT images of the psoas-major muscle using deep convolutional neural networks. Radiol. Phys. Technol. 2019, 12, 210–215. [CrossRef] [PubMed]
11. Lee, H.; Troschel, F.M.; Tajmir, S.; Fuchs, G.; Mario, J.; Fintelmann, J.F.; Do, S. Pixel-level deep segmentation: Artificial intelligence quantifies muscle on computed tomography for body morphometric analysis. J. Digit. Imaging 2017, 30, 487–498. [CrossRef] [PubMed]
12. Kamiya, N.; Kume, M.; Zheng, G.; Zhou, X.; Kato, H.; Chen, H.; Muramatsu, C.; Hará, T.; Miyoshi, T.; Fujita, H. Automated recognition of erector spinae muscles and their skeletal attachment region via deep learning in torso CT images. Comput. Methods Clin. Appl. Musculoskeletal Imaging 2019, 1–10. [CrossRef]
13. Oshima, A.; Kamiya, N.; Zhou, X.; Hará, T.; Fujita, H. Automated Segmentation of Surface Muscle in Whole-Body CT Images using 2D U-Net: Preliminary Study. In Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biological Society, Berlin, Germany, 23–27 July 2019; p. 71.
14. Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention 2016, Athens, Greece, 17–21 October 2016; pp. 424–432. [CrossRef]
15. Keras Documentation. Keras: The Python Deep Learning Library. Available online: https://keras.io/ (accessed on 10 July 2018).
16. Zhou, X.; Yamada, K.; Kojima, T.; Takayama, R.; Wang, S.; Zhou, X.; Hará, T.; Fujita, H. Performance evaluation of 2D and 3D deep learning approaches for automatic segmentation of multiple organs on CT images. In Proceedings of the SPIE Med Imaging 2018, Houston, TX, USA, 10–15 February 2018; pp. 1–6. [CrossRef]
17. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv 2015, arXiv:1502.03167v3.
18. Kingma, P.D.; Ba, J.; Adam. A method for stochastic optimization. arXiv 2017, arXiv:1412.6980v9.
19. Nimura, Y.; Deguchi, D.; Kitasaka, T.; Mori, K.; Suenaga, Y. Pluto: A common platform for computer-aided diagnosis. Med. Imaging Technol. 2008, 26, 187–191.
20. Dice, R.L. Measures of the amount of ecologic association between species. Ecology 1945, 26, 297–302. [CrossRef]
21. Sakamoto, M.; Hiasa, Y.; Otake, Y.; Takao, M.; Suzuki, Y.; Sugano, N.; Sato, Y. Bayesian segmentation of hip and thigh muscles in metal artifact-contaminated CT using convolutional neural network-enhanced normalized metal artifact reduction. J. Signal Process. Syst. 2020, 92, 335–344. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).