International Conference on Advances in Materials, Computing and Communication Technologies
In Association withInternational Journal of Scientific Research in Science and Technology
Volume 9 | Issue 1 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)

Deep Learning Based Visceral Adipose Tissue Segmentation Using Abdominal Images on MRI

Sudha Devi. B[1], Dr. D.S. Misbha[2]

[1]Research Scholar, Department of Computer Science, Nesamony Memorial Christian College Marthandam, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India

[2]Assistant Professor, Nesamony Memorial Christian College, Marthandam, Tamil Nadu, India

ABSTRACT

Obesity is closely related with increased risk of many diseases such as diabetes, hypertension, strokes, cardiovascular diseases and cancer. MRI is an accurate and most prevalent method to determine body fat distribution and quantification that causes obesity. The purpose of this work was to propose a novel approach for segmenting subcutaneous adipose tissue (SAT) and visceral adipose tissue (VAT) using MR images of abdomen. A deep convolutional neural network was used to segment VAT and SAT. The segmentation results were compared to the ground truth produced by base-line CNN. The Dice coefficients, and accuracy shows similarity between two methods in quantifying VAT and SAT. This study demonstrates the feasibility of applying a new deep learning based scheme to automatically segment VAT from MRI data with high accuracy.

Keywords: Visceral adipose tissue (VAT), Subcutaneous adipose tissue (SAT), Convolutional neural network, Deep learning.

I. INTRODUCTION

Obesity generally refers to the excess of fat in the human body. Abdominal adipose tissue is classified into visceral adipose tissue (VAT) and subcutaneous adipose tissue (SAT) [1]. VAT appears between the internal organs like kidney, liver, pancreas intestines and heart [2], [3]. SAT is a layer that lies under the skin. An excess of VAT is known as central obesity or belly fat in which the abdomen protrudes excessively which is the main cause of diabetes and cardiac problems. Apart from the traditional measure, Body Mass Index (BMI) which was used as a measure of obesity, CT and MRI contributes more due to their reliability and accuracy in the measurement of VAT. Accurate segmentation and quantification of VAT is important for clinical evaluation of obesity, cardiac syndromes and prediction of various diseases.

The imaging process of VAT deposited in the abdominal region is effectively done by CT or MRI [4], [5]. MRI is considered to be a safer imaging modality since it is not subjected to ionizing radiations. As VAT manifests mainly in the abdominal region, it is considered as an important marker for evaluating

central obesity. Deep learning models provide powerful tools for addressing the problems of object recognition, image segmentation and classification due to the availability of large amount of well annotated datasets and high speed Graphical Processing Units[6] . The convolutional neural network (CNN) is one of the important classes of deep learning most commonly applied for analyzing visual images. The purpose of this study was to develop a deep convolutional neural network to automatically segment SAT and VAT using abdominal MR images. Various studies are there in the literature to segment and quantify VAT and SAT separately. In this work, a deep convolutional neural network is used to automatically segment VAT and SAT without human intervention.

In the recent years, deep learning models like deep convolutional neural networks promote research interest and remains the state of art in many computer vision applications[7]–[10]. Segmentation of adipose tissue is mostly done using CT and MRI images. Weston et al. [11] used deep learning approach to separate VAT and SAT based on abdominal CT images. Grainger et al. [12] applied deep learning concept using the U-Net architecture for the segmentation of VAT and SAT on CT images the majority of MR-based adipose tissue segmentation algorithms focus on raw features such as intensity, shape and location rather than the intrinsic characteristics of tissues. Image histogram, K-means clustering algorithms and fuzzy c-means algorithms rely on image quality. K-means clustering algorithms [13] and fuzzy c-means algorithms [14] groups the pixels based on intensity features. A three channel U-Net architecture was suggested by Langner et al. [15] for the segmentation of SAT and VAT in the abdominal region using MRI images of humans. Estrada et al. proposed a deep learning pipeline to segment VAT and SAT in the abdominal region using DIXON MRI images [16]. Finally, a 3D CNN was proposed by Kustner et al.[17] for the assessment of VAT and SAT in the whole body MRI images. In the proposed CNN model, segmentation is done by focusing on the intrinsic features of tissues.

II. MATERIALS AND METHODS

Imaging dataset:

The cohort for this study consists of 250 non-contrast MRI datasets collected for standard CAC scoring. These datasets were collected randomly from the prospective EISNER trial obtained at NIMS Medical Centre. The patient population consisted of asymptomatic subjects with no prior history of coronary artery disease but with cardiovascular risk factors. The population is represented in Table 1. There were no scans excluded because of loss of quality in image or artefacts.. The total number of transverse slices per scan was 55, with 2.5 mm or 3 mm slice spacing. Each axial slice had a dimension of 512×512 pixels of 0.684 mm × 0.684 mm. In this study, a total of 11235 transverse views were used.

Table 1: Patient Characteristics

Characteristic	200
No. of women	63
No. of men	137
Age (years)	59.8±7
Diabetes	11
Hypercholesterolemia	124

Overview of traditional CNN architecture:

The Convolution Neural Network (CNN) is a deep neural network that was developed to analyse images. CNN extracts the high level information from the images without any pre-processing and feature extraction. A CNN is made up of different types of layers: input layer, convolution layer, the pooling layer, fully connected layer and output layer [18], [19]. The general architecture of CNN is shown in figure 1. The convolution layer is the most important layer in a CNN which extracts the key features from the input image. A convolution layer consists of a set of filters or kernel through which the convolution operation of

the input image is done [20]. The size of the kernel changes depending upon the filter size as 3x3, 5x5, and 7x7. The convolution operation with multiple filters is capable of extracting features from the dataset while preserving their spatial details. The feature map is the output of the convolution process.

Pooling is a convolution operation that reduces the dimensionality of feature maps. CNN employs both maximum and average pooling. In max pooling, the highest values are extracted from the feature map. Average pooling calculates the average value in the feature map. Deep convolutional neural network produces state of the art results related to image recognition problems [20].

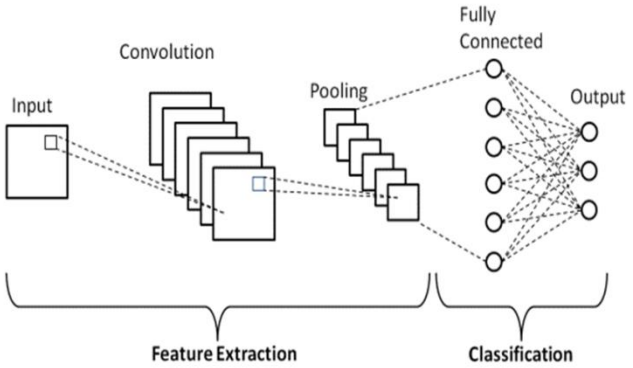

Fig.1. Common architecture of CNN

III. PROPOSED METHODOLOGY

CNN provides accurate segmentation results in most of the medical imaging applications. The sole aim of the proposed work is to classify an adipose tissue pixel as belonging to SAT or VAT depending up on the features. A deep neural network model is used in this study to mitigate the effects of various noises by adopting intrinsic features of different tissues. The internal boundary of SAT is determined using a graph cut algorithm. The flowchart representing the whole process of the proposed system is shown in figure 2. CNN is commonly used to perform pixel-wise classification [21]. In the first step, a deep neural network is used to construct an improved pixel-wise classification algorithm based on high-level representations of pixels.

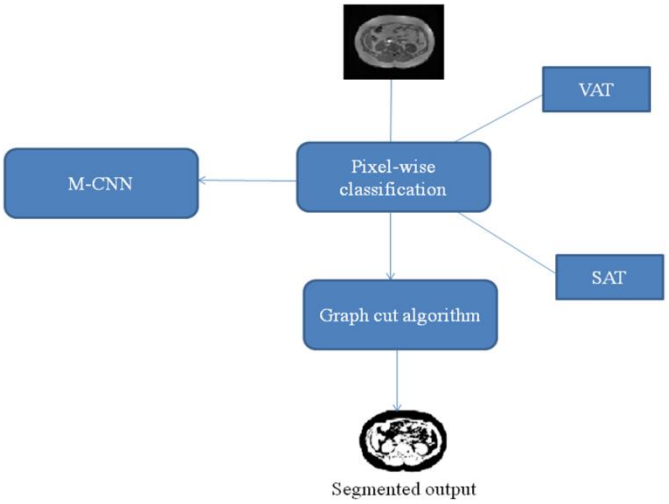

Fig.2. Overview of the complete segmentation process of adipose tissue

Four types of features: shape, location, intensity and context are considered. Polar coordinates are used to represent shape and location features. Intensity and contextual features are represented by multi-scale patches centered at each pixel. In the second stage, to investigate the intrinsic characteristics of tissues, a multi-scale deep neural network model called M-CNN is adopted that isolates the more abstract and high-level representations from these raw features.. In some cases, the segmentation algorithm misclassifies SAT and VAT pixels. Some VAT areas were misinterpreted as SAT and some SAT areas were misinterpreted as VAT.

In the next stage, graph cut algorithm is used to locate the internal boundary of SAT that improves the accuracy of segmentation[22] .Graph cut algorithms construct the graph depending up on pixel intensity and location information by using minimum-cut algorithms and find the internal boundary of SAT. Figure 3 demonstrates the separated internal boundary of SAT obtained as a result of minimum-cut algorithm. In this paper, exploring the internal distribution of adipose tissues in the abdomen with location information using deep neural network is focused to segregate SAT and VAT.

IV. PIXEL-WISE CLASSIFICATION

It is important to focus on the volumes of SAT and VAT in abdominal images for adipose tissue segmentation, and this is determined by the number of pixels that belong to VAT and SAT. So, a pixel-wise classification is done to segment the adipose tissue and all the pixels of the abdominal images are categorized into two groups: SAT and VAT. Three features are observed in T1-weighted abdominal MR images (i) adipose tissues are generally brighter than other tissues (ii) location distributions of SAT and VAT are different (iii) the ellipse shaped appearance of SAT. The following sections deals with the extraction of intrinsic features of tissues from the above-said raw features.

Selection of features

Step I – Location and shape features: Location and shape features are very important to differentiate SAT and VAT. A polar transform has been applied to these raw location and shape features, and polar coordinates are used to define the location and shape features in the classification algorithm. Since the internal and external boundaries of SAT are both ellipse shaped, the SAT boundaries will be much easier to identify if we make a polar transform of the original abdominal image. Following the polar transformation, the SAT boundaries can be seen at the bottom sections of the image. The polar coordinates (the radial coordinates) r and (polar angle) θ for each pixel are defined in terms of Cartesian coordinates by

$$r = \sqrt{(x - x_1)^2 + (y - y_1)^2} \qquad (1)$$

$$\theta = \tan^{-1}\left(\frac{y - y_1}{x - x_1}\right) \qquad (2)$$

Where (x1, y1) is the actual point, and (x, y) is Cartesian coordinate of pixel. When the location of different tissues is considered at the polar angle θ, the radial coordinates of SAT pixels are larger than other tissues. To preserve this property, all the radial coordinates of pixels along the given polar angle is normalized to [0,1]. First, the polar angles are divided into k regions. After that, the radial coordinates for each area is normalized to [0, 1]. The radial coordinate distributions of SAT pixels were found to be larger than those of VAT pixels.

Step II – Intensity and Contextual Features: In general, classes between pixels and their neighbours are closely related, and contextual features like multi-scale patches are used to capture these relationships.. The grey intensity information of different tissues varies in a single scale patch, which is inadequate for separating SAT and VAT. Larger patches provide more contextual information, but they lose local intensity discriminative features and increase the computational complexity. A multi-scale input is investigated to fully capture the local intensity discriminative features and global contextual information. An average pooling operator is used to compress larger sized patches in order to reduce the storage memory and computational complexity. On the other hand, the raw intensity patches suffer from inhomogeneous image intensities which are typical in MR images. A multi-scale deep neural network M-CNN is proposed to solve this problem by learning the discriminative and intrinsic representations from these features.

V. MULTI-SCALE DEEP NEURAL NETWORK (M-CNN)

A multi-scale deep neural network M-CNN is used to derive more abstract and intrinsic features from multi-scale patches. To learn the hidden representations from each scale patches, a separate two-layer Deep Belief Network (DBN) is used and in order to learn the correlations between high-level contextual features, location features and intensity features two additional hidden layers are added . The architecture of M-CNN with two-scale patches is shown in Fig. 3.

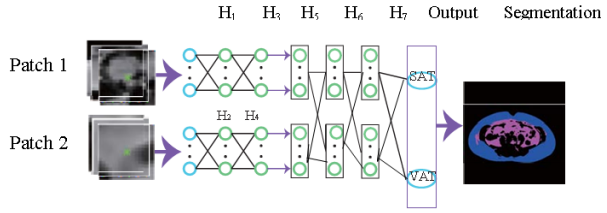

Fig.3. M-CNN architecture with patch 1 and patch 2 as inputs. Here two DBNs are used for learning the high level representations of patches.

For extracting a feature using multi-scale deep neural network two-scale patches P1 and P2 are considered. The weights of the deep learning architecture are considered as {W1, W2, W3, W4, W5, and W6}. The features used for the i-th pixel are (P1, P2, I, r, θ), where P1 and P2 indicate two-scale patches, I indicates the intensity, and r and θ indicate the polar coordinates. The hidden representations are initially learned from two different sized patches. The representations for the first patch P1 are learned using the following equations:

$$H_1=\sigma(W_1P_1 + b_1)\ ;\ H_3=\sigma(W_3P_1 + b_3) \qquad (3)$$

Where σ(x) =1/ (1 + exp (–x)) is the active function, and bj, 1, (j = 1, 2), indicates the bias of the j-th hidden layer that is connected to patch P1. In the same way, the hidden layers representations H2, H4 for patch P2 are learned. Then, all the representations of the i-th pixel (H3, H4 I, r, θ) are concatenated to a big vector H5 which is the input for the third hidden layer. Next, the hidden layer representations H6 and H7 for the i-th pixel are learned as,

$$H_6=\sigma(W_6H_5 + b_6)\ ;\ H_7=\sigma(W_7H_6 + b_4) \qquad (4)$$

Where bj, (j = 3, 4) is also the bias.

The classification results are obtained by softmax classifier. Given the classification parameters Wclass and bclass, the predicted label for the i-th pixel is calculated by the following normalized probability distribution.

$$\text{Prob}(y_i=c)= \frac{exp\left(W_{class,c}H_7+b_{class,c}\right)}{\sum_{k=1}^{2} exp\left(W_{class,q}H_7+b_{class,k}\right)} \qquad (5)$$

Where c ∈ {1, 2} indicates the label, where SAT and VAT are defined as class 1, class 2 respectively. Wclass,k represents the k-th column of Wclass, and bclass,k represents the j-th element in bclass,k

VI. REFINED SEGMENTATION

The resulting segmentation is not so accurate, as in some cases the algorithm misclassifies VAT pixels as SAT pixels or vice versa. To overcome this situation, the internal boundary of SAT is extracted using graph cuts algorithm. Based on the intensity of pixel and location information graph cut algorithms construct the graph and find the internal boundary of SAT by minimum cut algorithm. The image is represented as a weighted graph G = (V,E,W) where V represent the vertices (voxels in image), E represent the edges between the vertices and W represent the weights of the edges. The graph cut method requires the initial selection of foreground (FG) and background (BG) seeds. Here, SAT is supposed to be the region of interest and VAT is the region to be excluded to create a SAT boundary, because it is easier to define seeds on SAT when compared to VAT.

A weight is assigned to every edge of the graph using the distance-transformed values of the voxel

$$W_{ij}= \max_{u_iu_j\in FG} \left(D(u_i), D\left(u_j\right)\right) \qquad (6)$$

Where D(ui) and D(uj) are the distance transform values of the vertices ui and uj from the nearest tissue boundaries. The edges between fat and non-fat voxels have weights of 1 and the weights between non-fat voxels have weights of infinite value. The minimum cut algorithm finds the cut in the graph with the smallest sum of weights. Since the weights at the connections between SAT and VAT are smaller, the

cut happens that divides SAT and VAT. The process of refined segmentation is shown in Figure 4.

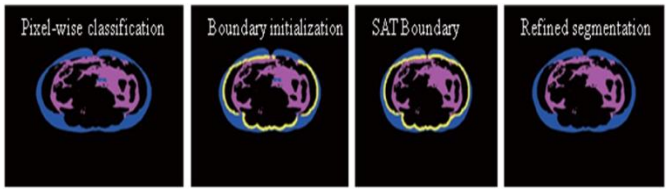

Fig.4: Results of Refined segmentation. SAT and VAT regions with a clear internal SAT boundary is shown. In the segmentation refinement subfigure, magenta coloured portion indicates VAT and blue coloured portion indicates SAT respectively.

VII. RESULTS

To implement this proposed model, a workstation with Dual Intel Xeon Processor E5-2630 v2 (Six-Core HT, 2.6GHz Turbo, 15 MB), 4 GB NVIDIA Quadro K5000, 1x8GB DDR3, 1TB 7200RPM SATA, and Windows 10 Pro workstation is used as a platform. Python 3.7 and open source machine learning libraries such as TensorFlow 1.13.1, Keras 2.2.4, and Scikit-learn 0.20.3 have been used to develop its working model. In this work, pixels were considered as samples in the segmentation of visceral adipose tissue. For each subject about 20,000 patches were generated centered at each pixel . The training samples were taken from these patches. Pixel-wise prediction accuracy and Dice coefficients (DC) of SAT and VAT were calculated for each image. Dice coefficient is a similarity measurement index that is commonly used to evaluate the performance of image segmentation algorithm against ground truth. A Dice coefficient of 1.00 is considered to be a perfect similarity. Let P and Q denote the ground truth segmentations and the proposed CNN based segmentations then, DC is defined as follows.

$$DC = \frac{2|P \cap Q|}{|P| + |Q|} \quad (7)$$

The proposed CNN based segmentation method indicates a good agreement with ground truth segmentation. Table 1 displays the evaluation results of the new CNN and baseline CNN. It is observed from the results that the proposed CNN method produced high pixel-wise prediction accuracy (SAT - 0.9798 and VAT – 0.9240) compared to baseline CNN which requires human expertise. The Dice coefficients of SAT and VAT are 0.9798 and 0.9240 which is high compared to baseline CNN. Figure 5 shows the comparison chart of the performance evaluation results.

Table 1: Tabulation of performance evaluation of baseline CNN and proposed CNN

Method of evaluation	Pixel-wise prediction accuracy	SAT Dice coefficient	VAT Dice coefficient
Base-line CNN	0.9432	0.9676	0.8780
Proposed CNN	0.9692	0.9798	0.9240

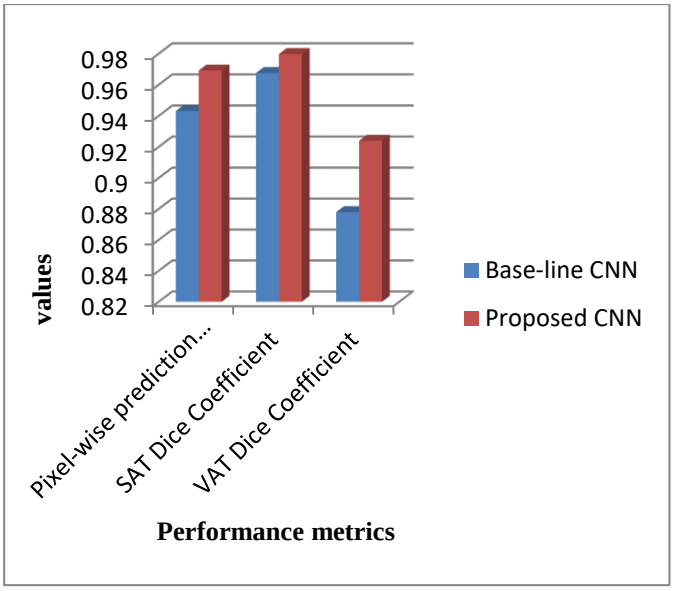

Fig. 5. Segmentation performance of Baseline CNN and proposed CNN

VIII. CONCLUSION

This work presents a method for fully automated segmentation of SAT and VAT on MR images of abdomen. This is done by using a deep neural network with input features like intensity, context,

shape and location. The outputs are the tissue label of all pixels. Development of an automated segmentation algorithm using appropriate feature selection and deep learning techniques with high accuracy could help radiologist to save the excess time that doctors spend on visceral adipose tissue segmentation. Over all, the proposed CNN provides accurate segmentation results and can be extended to handle large data with limited training time.

IX. REFERENCES

[1]. S. Hussein et al., "Automatic Segmentation and Quantification of White and Brown Adipose Tissues from PET/CT Scans," IEEE Trans. Med. Imaging, vol. 36, no. 3, pp. 734–744, Mar. 2017, doi: 10.1109/TMI.2016.2636188.

[2]. "Yan and Jiao - 2015 - Capsule Robot for Obesity Treatment With Wireless .pdf." .

[3]. "Corbeil et al. - 2001 - Increased risk for falling associated with obesity.pdf." .

[4]. J. Kullberg, "Whole-body MRI for analysis of body composition," in 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Rotterdam, Netherlands, 2010, pp. 1065–1066, doi: 10.1109/ISBI.2010.5490175.

[5]. S. M. Ibrahim, M. S. Ibrahim, M. Usman, I. Naseem, and M. Moinuddin, "A Study on Heart Segmentation Using Deep Learning Algorithm for MRI Scans," in 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), Karachi, Pakistan, Dec. 2019, pp. 1–5, doi: 10.1109/MACS48846.2019.9024793.

[6]. N. Jones, "Using massive amounts of data to recognize photos and speech, deep-learning computers are taking a big step towards true artificial intelligence.," p. 3.

[7]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[8]. S. Lawrence, C. L. Giles, Ah Chung Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," IEEE Trans. Neural Netw., vol. 8, no. 1, pp. 98–113, Jan. 1997, doi: 10.1109/72.554195.

[9]. Y. Lecun, "Gradient-Based Learning Applied to Document Recognition," PROCEEDINGS OF THE IEEE, vol. 86, no. 11, p. 47, 1998.

[10]. P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings., Edinburgh, UK, 2003, vol. 1, pp. 958–963, doi: 10.1109/ICDAR.2003.1227801.

[11]. A. D. Weston et al., "Automated Abdominal Segmentation of CT Scans for Body Composition Analysis Using Deep Learning," Radiology, vol. 290, no. 3, pp. 669–679, Mar. 2019, doi: 10.1148/radiol.2018181432.

[12]. A. T. Grainger et al., "Deep Learning-based Quantification of Abdominal Subcutaneous and Visceral Fat Volume on CT Images," Academic Radiology, p. S1076633220304268, Aug. 2020, doi: 10.1016/j.acra.2020.07.010.

[13]. G. Thörmer et al., "Software for automated MRI-based quantification of abdominal fat and preliminary evaluation in morbidly obese patients," J. Magn. Reson. Imaging, vol. 37, no. 5, pp. 1144–1150, May 2013, doi: 10.1002/jmri.23890.

[14]. A. Zhou, H. Murillo, and Q. Peng, "Novel segmentation method for abdominal fat quantification by MRI," J. Magn. Reson. Imaging, vol. 34, no. 4, pp. 852–860, Oct. 2011, doi: 10.1002/jmri.22673.

[15]. T. Langner et al., "Fully convolutional networks for automated segmentation of abdominal adipose tissue depots in multicenter water–fat

MRI," Magn. Reson. Med., vol. 81, no. 4, pp. 2736–2745, Apr. 2019, doi: 10.1002/mrm.27550.

[16]. S. Estrada et al., "FatSegNet: A fully automated deep learning pipeline for adipose tissue segmentation on abdominal dixon MRI," Magn Reson Med, vol. 83, no. 4, pp. 1471–1483, Apr. 2020, doi: 10.1002/mrm.28022.

[17]. T. Küstner et al., "Fully Automated and Standardized Segmentation of Adipose Tissue Compartments by Deep Learning in Three-dimensional Whole-body MRI of Epidemiological Cohort Studies," p. 30, 2020.

[18]. F. Jiang et al., "Abdominal adipose tissues extraction using multi-scale deep neural network," Neurocomputing, vol. 229, pp. 23–33, Mar. 2017, doi: 10.1016/j.neucom.2016.07.059.

[19]. C. Cao et al., "An Improved Faster R-CNN for Small Object Detection," IEEE Access, vol. 7, pp. 106838–106846, 2019, doi: 10.1109/ACCESS.2019.2932731.

[20]. S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," p. 9.

[21]. Z. Zhang, "Chapter 8 - Multimodal medical volumes translation and segmentation with generative adversarial network," p. 22.

[22]. S. A. Sadananthan et al., "Automated segmentation of visceral and subcutaneous (deep and superficial) adipose tissues in normal and overweight men: Automated Segmentation of Adipose Tissue," J. Magn. Reson. Imaging, vol. 41, no. 4, pp. 924–934, Apr. 2015, doi: 10.1002/jmri.24655.