

Non-intrusive Body Composition Assessment from Full-body mmWave Scans

Miriam Senne^{1,2*}[0009–0008–6995–3033],
Benjamin D. Killeen^{1,2*}[0000–0003–2511–7929],
Tony Danjun Wang^{1,2}[0000–0002–0883–1744], and Nassir Navab^{1,2}

¹ Chair for Computer Aided Medical Procedures
Technical University of Munich, Boltzmannstr. 3, 85748 Garching, Germany
{miriam.senne,bd.killeen,tony.wang,nassir.navab}@tum.de

² Munich Center for Machine Learning, Munich, Germany

*Equal contribution

Abstract. Body composition assessment (BCA) provides detailed information about the distribution of different tissue types in the body, enabling more precise characterization of individuals than BMI or weight alone. Consistent and frequent BCA would be valuable for personalized medicine, but the gold standard methods for BCA, such as CT and MRI, are only practical for opportunistic monitoring of patients with clinical indications for imaging and are not suitable for routine use in the general population. Here, we consider an imaging modality which is not currently used in medical applications: millimeter wave (mmWave) radar. Commonly used in security settings, mmWave scans enable fast, non-intrusive, and privacy-preserving reconstruction of full body shape without the need to remove clothing. To demonstrate the feasibility of fast and convenient BCA from mmWave scans, we present a method for BCA value regression using a multi-task learning strategy that leverages synthetic mmWave-like point clouds derived from clinical imaging and parametric human models. We evaluate the model on a pilot cohort of real mmWave scans with bioimpedance-derived body fat measurements, supporting the feasibility of estimating VAT and body fat percentage (BFP) from mmWave data acquired through clothing in a standing posture. We find that the model can predict VAT and BFP with a mean absolute error of 1.0 and 3.2%, respectively, demonstrating the potential of mmWave scanning for routine BCA in a wide range of settings.

Keywords: Point cloud · Personalized care · Longitudinal monitoring.

1 Introduction

Body composition assessment (BCA) is an important component in the emerging landscape of personalized medicine, providing detailed information about the distribution of different tissue types in the body. Quantitative measurements of skeletal muscle volume, abdominal fat volume, or bone mineral density, for example, characterize individuals more precisely than body mass index (BMI)

or weight, and studies have repeatedly demonstrated their clinical utility for risk stratification and prognosis in a wide range of conditions, including cardiometabolic disease [15,14], cancer cachexia (*i.e.*, rapid muscle degeneration), and frailty in elderly patients [2]. Consistent and frequent BCA would expand these and other benefits to the general population. At the same time, with the maturity of fully automated algorithms for medical image analysis [19], and in particular for segmentation of CT and MRI scans, there is growing interest in opportunistic body composition assessment as a byproduct of routine clinical imaging [7]. Recent work has shown that external body shape alone can provide strong signals for adiposity and fat distribution, including models that predict body fat percentage from three dimensional body point clouds [21] and surface meshes [12], as well as models that estimate VAT and related depots from two dimensional silhouettes derived from MRI [8]. Here, we consider an imaging modality which is not currently used in medical applications: millimeter wave (mmWave) radar. Commonly used in security settings, mmWave scans enable fast, non-intrusive, and privacy-preserving reconstruction of full body shape without the need to remove clothing [16]. It can be deployed in a wide range of settings, including primary care clinics and community centers. Our goal is to enable fast and convenient BCA from routine mmWave scans, with a particular focus on the estimation of visceral adipose tissue (VAT) and body fat percentage (BFP).

In this paper, we present a method for VAT and BFP regression from mmWave scans. After reconstruction, the output of an mmWave scan is a point cloud representing the front and back surface of the body, with missing data in the regions with a surface normal not aligned with the AP axis. To train a model that estimates internal body composition from this surface geometry, we confront two main challenges: 1) the lack of paired mmWave and body composition data, and 2) the domain gap between supine or prone MRI/CT scans and standing mmWave scans. Using tissue type segmentation [19] of MRI and CT scans is a fast, scalable way to obtain VAT and BPT values, and we introduce a method to synthesize mmWave-like point clouds from these scans based on the body surface geometry. However, because these scans are acquired in the supine or prone position, the observed surface geometry is affected by gravitational deformation and contact with the support surface on the anterior and posterior sides, respectively. Alternative sources of body surface data, like the Skinned Multi-Person Linear Model (SMPL) [10], can provide a complete body surface in arbitrary poses, but lack any information about the internal body composition. The same is true of available datasets of mmWave scans, although other measurements can be automatically derived from body surfaces [16] as well as CT/MRI scans [7]. Thus, we propose a multi-task learning strategy that leverages each data source for the information it provides, with the goal of learning a shared representation that can bridge the domain gap between the different modalities and postures. Specifically, we train a single point cloud encoder with multiple regression heads, regressing volumes of intramuscular adipose tissue (IMVAT), visceral adipose tissue (VAT), subcutaneous adipose tissue (SAT), total body, lean tissue (LT) and

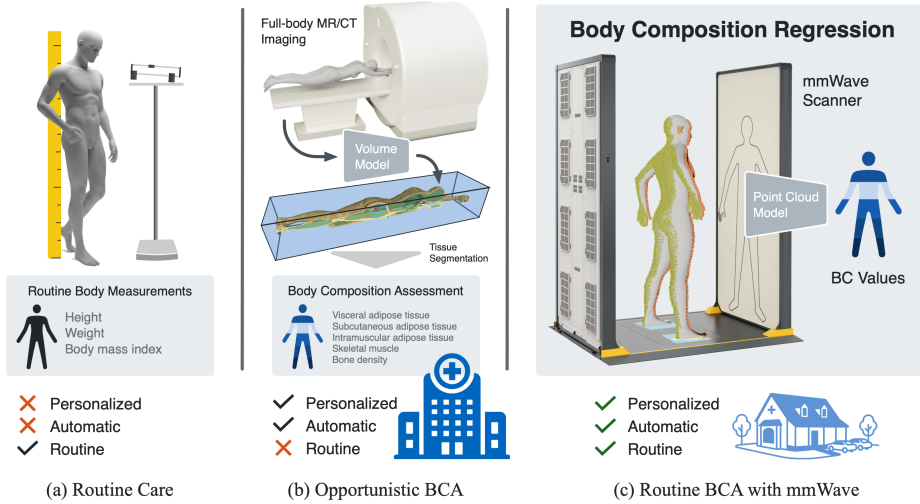


Fig. 1: Routine body measurements like height, weight, and BMI are widely used but provide limited information about body composition (a). In contrast, detailed body composition assessment (BCA) can provide more precise characterization of individuals and has demonstrated clinical utility, but is currently only practical for opportunistic monitoring of patients with clinical indications for imaging (b). Millimeter wave (mmWave) radar is a fast, non-intrusive, and privacy-preserving imaging modality that can reconstruct a partial body shape without the need to remove clothing, enabling BCA value regression (c).

muscle (MV); height; and chest, waist, and hip circumferences, using posterior and anterior surface data from MRI/CT scans, SMPL meshes, and mmWave scans. We demonstrate the feasibility of this approach by evaluating the model on a withheld set of synthetic mmWave-like point clouds derived from MRI/CT scans and SMPL models, the real mmWave dataset including anthropometric measurements, and on a pilot cohort of real mmWave scans with bioimpedance-derived body fat measurements.

2 Methods

2.1 Dense Point Cloud Extraction from CT/MRI Data

Paired millimeter wave point clouds and ground truth body composition labels are expensive to acquire and are therefore scarce. To enable supervised training, publicly available full body MRI and CT datasets with body composition annotations are used. From each volume, surface geometry is extracted and converted into point clouds that simulate mmWave scanner output. Because the scanner provides full body coverage, only full body reference scans are considered. Full body MRI scans are provided in the HIT dataset [6], where subjects

were scanned in the prone position stretching their arms over their heads, leading to detailed information of the back surface while producing a limited and flat representation of the anterior surface. The HIT dataset contains 398 whole body MRI scans with four tissue classes, namely muscle, SAT, IMVAT, and bone and fitted SMPL body models. IMVAT corresponds to a combined segmentation of VAT and intramuscular adipose tissue (IMAT). A binary body volume is constructed by retaining background as label 0 and mapping all other tissue labels with values greater than 0 to 1. Since air cavities such as the lungs may be labeled as background, slice based morphological operations are applied to obtain a closed and solid body volume. A surface mesh is then reconstructed using marching cubes with a threshold of 0.5. To obtain complementary anterior surface information, we employ a subset of full body CT scans from the NMDID [3] database, which contain examinations of diseased subjects that have their arms crossed in front of their chest. The NMDID subset used in this work is filtered to include only cases of natural death with minimal decomposition, resulting in 798 full body CT scans. Tissue labels are generated using TotalSegmentator, and the labels for SAT, VAT, and skeletal muscle are retained. 41 scans are excluded due to extreme BF below 5% or above 70%. At the low end, even minimally decomposed samples can carry anatomical artifacts that corrupt TotalSegmentator-derived BCA labels. At the high end, field-of-view truncation rendered both tissue labels and body surfaces impossible to obtain reliably. Surface meshes are reconstructed from the body label, which covers the trunk and the remaining body regions, using marching cubes. To obtain a high resolution and smooth surface suitable for point cloud synthesis, Laplacian smoothing is applied to all reconstructed meshes, after which 800000 points and corresponding normals are sampled from each surface. Body composition labels are computed as tissue volumes by counting voxels assigned to the corresponding tissue class and multiplying by the voxel volume derived from the image spacing provided with each scan. The resulting values are converted to liters.

2.2 Dense Point Cloud Reconstruction from Parametric Models

To bridge the domain gap between supine/prone clinical imaging and the standing posture required for mmWave scanning, we extend our training dataset with synthetic point clouds derived from the SMPL model [10]. We establish base postures—specifically, an upright standing position with arms slightly abducted (A-pose) and a posture with arms crossed in front of the chest, resembling CT acquisitions (X-pose). We captured informal reference photographs of these poses and fit SMPL parameters to them [13]. For each base pose, we generate a variety of plausible human shapes by sampling the SMPL shape parameters (β) with a standard deviation that ensures a realistic, diverse distribution of body measurements and overall composition.

Achieving high pose diversity alongside this shape variation presents a distinct challenge, as naively applying noise to pose parameters often results in implausible articulations, such as unnaturally bent joints. To overcome this, we perturb the initial poses and project them back onto the closest plausible

configuration on a learned manifold of valid human articulations using Neural Riemannian Distance Fields (NRDF) [4]. Finally, we employ VolumetricSMPL [11] to guarantee that the resulting meshes are physically realistic and free of self-intersections, which we then convert into synthetic mmWave point clouds. Since the HIT dataset provides fitted SMPL models that exclude the flattened regions caused by the scan bed, these meshes are used directly for the prone poses. SMPL meshes in A-pose and X-pose are sampled in the same quantity as the HIT SMPL instances, yielding a total of 1194 SMPL based samples. Resulting meshes are Laplacian smoothed and 800000 surface points with corresponding normals are sampled. Anthropometric labels are derived by measuring chest, waist, and hip circumferences according to ISO 8559 [5]. For each shaped SMPL mesh, the model is first placed in a canonical pose. Axial cross sections are then extracted at fixed keypoint locations, and each circumference is computed as the perimeter of the convex hull of the corresponding cross section.

2.3 Data Pre-Processing and Normal Filtering

Microwave imaging is derived from radar and exploits electromagnetic waves in the 30 to 300 GHz range to penetrate materials and interact with internal structures [1]. The used mmWave sensor comprises two opposing panels, and subjects stand between them in an A pose as shown in Fig. 1. Acquisition requires 25 ms to 30 ms and yields an intensity volume with 1.9 mm by 1.9 mm lateral resolution and 5.5 mm depth resolution, where high values correspond to the skin surface. Due to viewpoint dependent illumination, regions with surface normals oriented away from the panels do not get captured, resulting in missing scanner observation, particularly along the body sides. Since front and back volumes are available separately, point clouds are extracted by locating depthwise intensity maxima and applying thresholding. Synthesized mmWave data from meshes have to exhibit those properties.

Prior to training, each point cloud is registered to the millimeter wave scanner coordinate space. Augmentation is applied using a random rotation about the vertical axis, with angles sampled from a normal distribution and clipped to the range from -10 to 10 degrees. Positional jitter is added with a maximum magnitude of 2 mm to reflect the lateral scanner resolution. A normal based filtering step is then used to simulate scanner illumination. For each point, the squared dot product between the point normal and a panel normal vector is computed, and points below a threshold are removed. The threshold is sampled from a normal distribution with mean 0.7 and is treated as an additional augmentation parameter, as well as a subsequent random down-sampling to 100000 points.

2.4 Network Architecture

The proposed multi-headed regression model, shown in Fig. 2, predicts ten body related measurements from three dimensional point clouds. The backbone is a PointTransformerV3 encoder [20] initialized from large-scale pretraining on

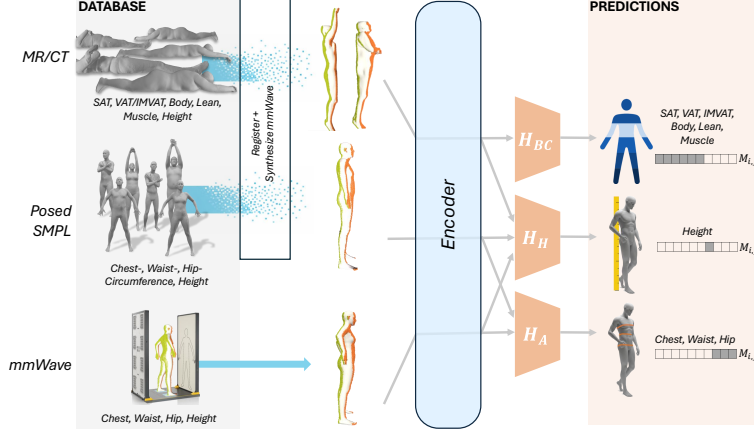


Fig. 2: Overview of the proposed pipeline, including mmWave specific point cloud synthesis from CT, MR, and SMPL surfaces, for a n -d multi-headed model with a shared encoder and three task-specific regression heads.

diverse indoor three dimensional datasets, since pretrained weights for human point clouds are not available. Since the pretrained configuration expects six input feature channels, the three dimensional coordinates are duplicated to form an N by 6 feature tensor. The encoder extracts a shared 512-dimensional feature representation from three dimensional point clouds, which is passed to three task specific MLP heads. $Head_H$ predicts body height and outputs one value, while $Head_A$ predicts three anthropometric measurements and $Head_{BC}$ predicts six body composition metrics. Each head uses the same MLP design. The input layer receives the 512-dimensional encoder feature, followed by two hidden layers with 256 and 128 units. Each hidden layer is followed by layer normalization, GELU activation, and dropout. A final linear layer maps the 128 dimensional representation to the required output dimension.

All target labels are Z-score normalized using a precomputed mean and standard deviation for each target. In multi task learning, manual selection of loss weights is challenging because tasks differ in scale and convergence behavior. This work therefore uses a masked and dynamically weighted objective that supports partial labeling and stabilizes optimization. First, we define a localized Masked Sample Loss ($\ell_{i,h}$) to ensure gradient isolation:

$$\ell_{i,h} = \frac{\sum_{j \in \text{Head}_h} \mathcal{L}(\hat{y}_{i,j}, y_{i,j}) \cdot M_{i,j}}{\left(\sum_{j \in \text{Head}_h} M_{i,j} \right) + \epsilon} \quad (2)$$

The binary mask $M_{i,j}$ zeros out missing labels, while the denominator normalizes the gradients. This allows the shared encoder to learn universal geometric features from disjoint datasets. The per target loss $\mathcal{L}$ is the Huber loss, which is robust to outliers via a quadratic to linear transition. The small constant ϵ

Table 1: MAE and standard deviation on withheld test sets, reported per dataset for the pretrained multi-headed model. *BFP in % is included based on dataset specific fat definition.

Dataset	Height (cm)	Circumference (cm)			Volume (L)				BFP* (%)
		Chest	Waist	Hip	SAT	IMVAT	VAT	Body	
<i>Synthetic</i>									
NMDID	1.9 $\pm$ 1.6	—	—	—	2.5 $\pm$ 2.2	—	1.0 $\pm$ 0.9	2.0 $\pm$ 1.5	4.07 $\pm$ 3.30
HIT	1.6 $\pm$ 1.3	—	—	—	2.2 $\pm$ 1.5	0.7 $\pm$ 0.7	—	1.6 $\pm$ 1.1	2.36 $\pm$ 1.54
SMPL Σ	1.2 $\pm$ 0.9	1.2 $\pm$ 1.0	1.5 $\pm$ 1.2	1.1 $\pm$ 0.8	—	—	—	—	—
<i>Real</i>									
mmWave _A	1.1 $\pm$ 0.4	2.8 $\pm$ 1.3	1.8 $\pm$ 1.3	1.5 $\pm$ 0.6	—	—	—	—	—
mmWave _{BC}	1.4 $\pm$ 1.0	—	—	—	—	—	—	—	—

ensures stability when no labels are present for a given head. Second, we employ Dynamic Weight Averaging (DWA) to adaptively balance the H heads:

$$\mathcal{L}_{total} = \sum_{h=1}^H w_h(t) \cdot \left[\frac{1}{N_h} \sum_{i=1}^N \ell_{i,h} \right] \quad (3)$$

By monitoring relative descent rates, DWA [9] automatically prioritizes lagging tasks. The constraint $\sum w_h = H$ stabilizes the total gradient energy, ensuring robust optimization without manual hyperparameter tuning.

Training uses AdamW with separate learning rates for the encoder (0.00005) and heads (0.0001), a cosine annealing scheduler, and a 10-epoch warmup for stable optimization. Weight decay and gradient clipping are applied for regularization and stability. On a single Nvidia RTX 6000 Ada GPU, training for 150 epochs with a batch size of six takes about six hours.

3 Evaluation

Evaluation is conducted on synthetic point clouds derived from CT, MR, and SMPL data, as well as on real millimeter wave scans. For the HIT dataset, the provided train/val/test split of 8/1/1 is used. All remaining datasets are split across subjects and gender with the same ratio. In addition to the synthetic data, we use a real mmWave anthropometry dataset [16], denoted *mmWave_A*, with 54 scans from 27 subjects. Table 1 shows the performance of the model on withheld test sets for each dataset, using mean absolute error (MAE), where SMPL Σ denotes aggregated results across all SMPL subsets. The results show that across synthetic datasets, anthropometric errors are typically within 1 to 2 cm, indicating that the model compensates for the mmWave illumination gaps. SAT and VAT errors should be interpreted alongside label variability across datasets, since boundary errors from abdominal region segmentation and inconsistent annotation protocols contribute to the reported MAE values beyond model prediction error alone.

To assess transfer to real mmWave scans for BCA, a pilot cohort dataset, denoted *mmWave_{BC}*, was collected and comprises scans of 14 subjects (8 men, 6

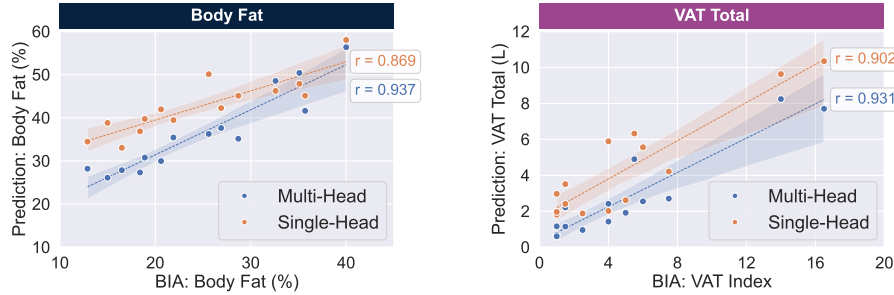


Fig. 3: Correlation of BFP (left) and VAT (right) with a 95% confidence interval for the mmWave_{BC} data, for multi-head and single-head architectures. Pearson score r rates the correlation of predicted and target values.

women; age 21–31 years; weight 51–108 kg) with paired height and BCA values acquired using the Tanita RD 545HR bioelectrical impedance scale [17]. Participants signed an informed consent and ethics waiver was granted. To reduce variability in impedance based estimates, measurements were performed in the morning and participants were instructed to avoid food and drink beforehand. Each participant underwent three consecutive scans in a standard A-pose, standing on fixed floor markers and clothed in everyday streetwear to emulate casual acquisitions. Model outputs are averaged across the three scans and compared against Tanita VAT and BFP measurements. The model predicts VAT volume in liters, whereas Tanita reports a VAT index in the range 0 to 59, where values from 1 to 12 indicate a healthy level and values above 12 indicate an elevated level. BFP is used as a second metric, and a comparable value is computed as the sum of IMVAT, or VAT when IMVAT is unavailable, and SAT divided by the total body volume. For completeness, this derived BFP MAE is also reported in Table 1 for the synthetic CT/MR data. Because the VAT index has no physical unit conversion and the derived BFP may be affected by error propagation from multiple predicted volumes, agreement is assessed using correlation. Associations are visualized in Fig. 3 and quantified using the Pearson correlation coefficient r . We also evaluate a single head H_{BC} , only trained on CT and MR body composition data, to evaluate the effect of multi task learning. The multi-headed model shows consistently stronger association than the single head baseline, with r increasing from 0.868 to 0.937 for BFP and from 0.902 to 0.931 for VAT, indicating that jointly learning related biomarkers improves transfer to real mmWave scans.

4 Discussion and Conclusion

Routine body composition assessment could improve risk stratification and preventive care, but current methods are limited by cost, accessibility, and patient

burden. mmWave scanning offers a rapid, non-invasive, clothing-penetrating alternative performed standing and fully clothed. In this paper, we demonstrated the feasibility of mmWave-based BCA using a multi-task regression model with a shared encoder, trained on synthetic point clouds derived from CT, MR, and SMPL surfaces alongside real mmWave scans. The model successfully predicted anthropometric and body composition metrics, showing strong correlation with reference standards in the pilot cohort. Further work should validate performance against paired mmWave–MRI data rather than bio-impedance devices, which, despite their convenience, are prone to systematic error and cannot capture detailed variations in fatty tissue [18]. A key next step is the collection of large-scale, longitudinal datasets with gold-standard CT or MRI ground truth, enabling rigorous evaluation of clinical utility and translation to routine population-level monitoring.

Acknowledgment This work has been supported by Rohde & Schwarz GmbH & Co. KG (R&S). Additionally, we thank the TUM Ethics Committee for their advice regarding data collection and Armin Keshmiri for supporting this work with his medical expertise. All participants provided informed consent.

Disclosure of Interest M. Senne receives funding from R&S. mmWave device access was provided by R&S. The other authors have no competing interests to declare.

References

1. Ahmed, S.S., Schiessl, A., Gumbmann, F., Tiebout, M., Methfessel, S., Schmidt, L.P.: Advanced microwave imaging. *IEEE microwave magazine* **13**(6), 26–43 (2012)
2. Bates, D.D.B., Pickhardt, P.J., Bates, D.D.B., Pickhardt, P.J.: CT-Derived Body Composition Assessment as a Prognostic Tool in Oncologic Patients: From Opportunistic Research to Artificial Intelligence–Based Clinical Implementation. *Am. J. Roentgenol.* (Jun 2022). <https://doi.org/10.2214/AJR.22.27749>
3. Edgar, H., Daneshvari Berry, S., Moes, E., Adolphi, N.L., Bridges, P., Nolte, K.B.: New mexico decedent image database. Office of the Medical Investigator (2020)
4. He, Y., Tiwari, G., Birdal, T., Lenssen, J.E., Pons-Moll, G.: Nrdf: Neural riemannian distance fields for learning articulated pose priors. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 1661–1671 (2024)
5. Part 1: Anthropometric definitions for body measurement. Standard, International Organization for Standardization, Geneva, CH (Mar 2017)
6. Keller, M., Arora, V., Dakri, A., Chandhok, S., Machann, J., Fritsche, A., Black, M.J., Pujades, S.: Hit: Estimating internal human implicit tissues from the body surface. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 3480–3490 (2024)
7. Killeen, B.D., Wan, B., Kulkarni, A.V., Drenkow, N., Oberst, M., Yi, P.H., Unberath, M.: Towards Virtual Clinical Trials of Radiology AI with Conditional Generative Modeling. *arXiv* (Feb 2025). <https://doi.org/10.48550/arXiv.2502.09688>
8. Klarqvist, M.D., Agrawal, S., Diamant, N., Ellinor, P.T., Philippakis, A., Ng, K., Batra, P., Khera, A.V.: Silhouette images enable estimation of body fat distribution and associated cardiometabolic risk. *NPJ digital medicine* **5**(1), 105 (2022)

9. Liu, S., Johns, E., Davison, A.J.: End-to-end multi-task learning with attention. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1871–1880 (2019)
10. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: SMPL: a skinned multi-person linear model. *ACM Trans. Graph.* **34**(6), 1–16 (Aug 2023). <https://doi.org/10.1145/2816795.2818013>
11. Mihajlovic, M., Zhang, S., Li, G., Zhao, K., Muller, L., Tang, S.: Volumetricmpl: A neural volumetric body model for efficient interactions, contacts, and collisions. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5060–5070 (2025)
12. Mueller, T.T., Zhou, S., Starck, S., Jungmann, F., Ziller, A., Aksoy, O., Movchan, D., Braren, R., Kaissis, G., Rueckert, D.: Body fat estimation from surface meshes using graph neural networks. In: International workshop on shape in medical imaging. pp. 105–117. Springer (2023)
13. Pavlakos, G., Choutas, V., Ghorbani, N., Bolkart, T., Osman, A.A.A., Tzionas, D., Black, M.J.: Expressive body capture: 3d hands, face, and body from a single image. In: Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (2019)
14. Pickhardt, P.J., Graffy, P.M., Perez, A.A., Lubner, M.G., Elton, D.C., Summers, R.M.: Opportunistic Screening at Abdominal CT: Use of Automated Body Composition Biomarkers for Added Cardiometabolic Value. *Radiographics* (Mar 2021), <https://pubs.rsna.org/doi/abs/10.1148/rg.2021200056>
15. Quirino-Vela, L., Mayoral-Chavez, M., Pérez-Cervera, Y., Ildefonso-García, O., Cruz-Altamirano, E., Ruiz-García, M., Alpuche, J.: Cardiometabolic risk assessment by anthropometric and biochemical indices in mexican population. *Front. Endocrinol.* **16**, 1588469 (Jul 2025). <https://doi.org/10.3389/fendo.2025.1588469>
16. Senne, M., Killeen, B.D., Baur, C., Navab, N., Farshad, A.: Millimeter-wave imaging for anthropometric measurement (2026), <https://arxiv.org/abs/2605.23064>
17. Tanita Europe B.V.: Tanita RD-545HR. Tanita Europe B.V., Amsterdam, Netherlands (2021), <https://tanita.de/rd-545hr>, dual-frequency segmental bioelectrical impedance analyzer with heart rate monitoring
18. Tewari, N., Awad, S., Macdonald, I.A., Lobo, D.N.: A comparison of three methods to assess body composition. *Nutrition* **47**, 1–5 (Mar 2018). <https://doi.org/10.1016/j.nut.2017.09.005>
19. Wasserthal, J., Breit, H.C., Meyer, M.T., Pradella, M., Hinck, D., Sauter, A.W., Heye, T., Boll, D.T., Cyriac, J., Yang, S., Bach, M., Segeroth, M.: TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiology: Artificial Intelligence* (Jul 2023), <https://pubs.rsna.org/doi/10.1148/ryai.230024>
20. Wu, X., Jiang, L., Wang, P.S., Liu, Z., Liu, X., Qiao, Y., Ouyang, W., He, T., Zhao, H.: Point transformer v3: Simpler faster stronger. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4840–4851 (2024)
21. Zheng, Y., Long, Z., Feng, B., Cheng, R., Vaziri, K., Hahn, J.K.: D3bt: Dynamic 3d body transformer for body fat percentage assessment. *IEEE Journal of Biomedical and Health Informatics* **29**(2), 848–856 (2024)