Supplementary Material

1 Introduction

This document provides (i) qualitative results illustrating pose and shape diversity, (ii) additional statistics of the ChangeLing18K (gender, weight, hip size distributions ,etc), (iii) additional qualitative results on the dataset generated by Seedream 4.0 [2], and (iv) additional results for qualitative comparison with Flux Kontext [5]. We have explained in detail the dataset generation process in our main paper, so we will not be going into that in this material. We manually curated every prompt to ensure diversity of pose, clothing, backgrounds and extent of transformations in our data.

Original **(a)** **(b)** **Original** **(a)** **(b)**

Figure 1: Odo generates realistic body transformations while preserving identity, clothing, and background details. Examples show transformations from original images to (a) thinner and (b) overweight body types across various poses including sitting and standing positions. All are in-the-wild images.

2 Dataset Statistics

In this section, all the metrics related to the human body in the image are calculated by firstly extracting the SMPL (Skinned Multi-Person Linear) body shape parameters using NLF [8] and then calculating the weights,heights and all the body measurements using SMPL-Anthropemtry [1].

2.1 Training Dataset

Our training dataset includes a total of **18,573** pairs of images. This is split into 3 transformation pairs—thin-fat,fat-muscle and thin-muscle. We have tried to ensure a fair gender ratio in our dataset through meticulous manual pruning resulting in total number of **male pairs** to be **10,377 (55.9%)** and **female pairs** to be **8,196 (44.1%)**. A few sample images showcasing the diversity in transformation and pose of our training set is shown in Fig. 3. The gender wise as well as the transformation type wise statistics for the training dataset are shown in the tables below.

Table 1: Training Set Gender-wise Weight Statistics (in kg) for each Body Type

Body Types	**Female**			**Male**		
	Mean	Min	Max	Mean	Min	Max
Thin	53.4	40.9	82.8	73.5	27.1	103.1
Fat	77.4	44.6	124.6	106.9	64.4	145.4
Muscle	56.2	44.5	95.9	84.3	55.2	121.2

Table 2: Training Set Weight Statistics and Pair-wise deltas (all values in kg).

Body Types	**Absolute values**				**Body Pairs**	**△ b/w body types**		
	Mean	Median	Min	Max		Mean	Min	Max
Thin	64.3	67.1	27.1	103.1	Thin–Fat	29.3	-10.4	91.7
Fat	94.2	94.0	44.6	145.4	Thin–Muscle	7.3	-9.8	41.2
Muscle	71.6	77.3	41.1	121.2	Fat–Muscle	-22.0	-69.0	23.5

As seen from Table 1 andTable 2, the training set covers a very wide weight spectrum, from 27kg to 145kg, while keeping clear separation between body-type categories. On average, males are roughly 20kg heavier than females in every class (e.g., Thin≈74kg vs 53kg; Fat≈107kg vs 77kg; Muscle≈84kg vs 56kg). The transformation pairs are also substantial: Thin→Fat increases weight by about 30 kg on average, Thin→Muscle by 7kg, and Fat→Muscle reduces weight by 22kg.

Some of the other metrics related to various dimensions of the human body parts for both male and female are shown in Fig. 14 and Fig. 15. We also showcase the large pose variety (in Fig. 3) and weight range of our test dataset through some qualitative samples in Fig. 4, for fat transformation, and in Fig. 5 for muscular transformation.

2.2 Test Dataset

We create a custom test using the exact same pipeline we used to create our training data, but using 450 previously unseen faces from the CelebA-HQ dataset [4] as well as 15 license-free faces from and Pixabay [6]. For each face, we generate 5 images per body type (thin, fat, muscular) using PuLID [3], giving 2,325 images per type (6,975 total) and 2325 matched triplets *thin, fat, muscular*. Each triplet yields six ordered body-shape pairs, so the pool contains up to 13,950 candidate transformations. After the same manual pruning used for training (removing pose/clothing mismatches, artifacts, identity drift, minimal shape change) we retain **3600** high-quality pairs for evaluation. The number of male and female pairs in our test dataset is **2244 (62.3%)** and **1356 (37.7%)** respectively. A few sample images showcasing the diversity in transformation and pose of our training set is shown in Fig. 6. The gender wise as well as the transformation type wise statistics for the test set are shown in the tables below.

As seen from Table 3 and Table 4, the test set mirrors the training set's diversity: body weights span ≈42kg to 166kg, again with males about 18–35kg heavier than females in each class (Thin≈71kg vs 53kg, Fat≈117kg vs 82kg, Muscle≈85kg vs 57kg). Transformation magnitudes remain substantial—Thin→Fat gains 34kg on average, Thin→Muscle adds 10kg, and Fat→Muscle sheds 24kg—providing strong supervision for evaluating shape-editing performance. The Table 3 and Table 4 confirm that the held-out data preserves wide weight coverage, clear gender/body-type separation, and meaningful before-after deltas, making it a reliable benchmark for generalization tests.

Table 3: Test Set Gender-wise Weight Statistics (in kg) for each Body Type

Body Types	**Female**			**Male**		
	Mean	Min	Max	Mean	Min	Max
Thin	52.8	41.6	76.1	71.4	59.7	104.1
Fat	82.3	39.7	113.5	117.2	76.7	165.5
Muscle	57.0	45.1	69.0	84.8	73.4	127.9

Table 4: Test Set Weight Statistics and Pair-wise deltas (all values in kg).

Body Types	**Absolute values**				**Body Pairs**	**△ b/w body types**		
	Mean	Median	Min	Max		Mean	Min	Max
Thin	64.9	67.1	41.6	104.1	Thin–Fat	33.8	-5.6	72.9
Fat	99.6	98.3	39.7	165.5	Thin–Muscle	10.2	-2.1	41.4
Muscle	76.2	80.7	45.1	127.9	Fat–Muscle	-24.2	-20.5	24.4

Some of the other metrics related to various dimensions of the human body parts for both male and female are shown Fig. 16 and Fig. 17. We also showcase the large weight range of our test dataset through some qualitative samples in Fig. 7, for fat transformation, and in Fig. 8 for muscular transformation.

3 Qualitative Results

In Fig. 1, we showcase our model's capability to generate realistic body transformations while preserving the identity as well as the complex design of the original garments for in=the-wild images. We show (see Fig. 10) the inference results our model generated on our test set. The input image and the depth map are passed into the model as conditioning. Our model learns to follow the depth map and produces realistic transformations. We also show (see Fig. 11) the inference results on a completely dataset generated using Seedream 4.0 [2]. We also show a qualitative comparison of our model's output with that of Flux Kontext [5] in Fig. 12 and Fig. 13.

3.1 Comparison with BR-5K Dataset

BR-5K dataset, introduced by Ren et al. [7], consists of 5000 pairs of high resolution images with the faces blurred out. They construct the pairs with the help of professionals who edit the source image to get the transformed counterpart. However, their transformations are very minimal making the transformed image hard to distinguish from the source image. They also do not provide any annotations on what type of transformation is in a particular pair. Their dataset also has a imbalanced gender ratio with female and male pairs being **4,621 (92.4%)** and **381 (7.6%)** respectively. We show the qualitative comparison of BR-5K [7] with our training dataset images in Fig. 9.

Using SMPL-Anthropometry [1] and NLF [8], we fit the SMPL model on the BR-5K dataset and measure the weight of the model for all of their images. We simply show through Fig. 2 that our dataset has a larger, more visible change in the the weights of people in thin and fat body types. We compare only thin and fat because, BR-5K [7] does not annotate the dataset with transformation type, hence through visual inspection of BR-5K [7] we chose to put them in thin and fat categories. Our dataset exhibits quite a symmetric as well as broad distribtutions for weight deltas (fat-thin) as compared to BR-5K [7], portraying that the transformation extent is greater in our dataset.

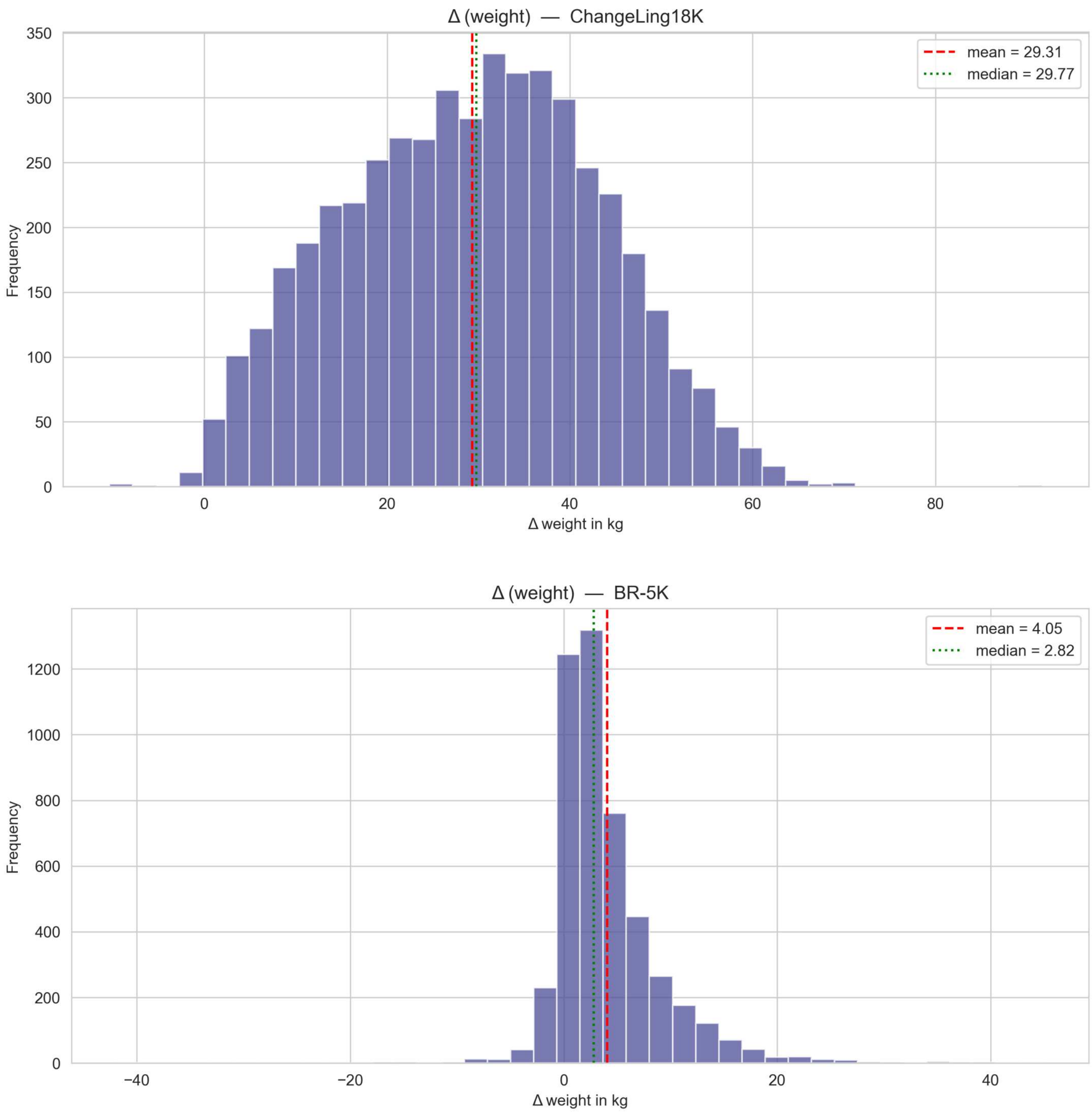

Figure 2: Distribution of weight Δ(fat$-$thin). Top: our ChangeLing18K training set. Bottom: BR-5K [7]. Our transformations showcase a wider range of weight difference when going from thin to fat body type.

Figure 3: Sample images from our training set showing variety of poses and transformations

Figure 4: Wide weight range for the fat body type in our training set.

Figure 5: Wide weight range for the muscle body type in our training set.

Figure 6: Sample images from our test set showing variety of poses and transformations.

Figure 7: Wide weight range for the fat body type in our test set

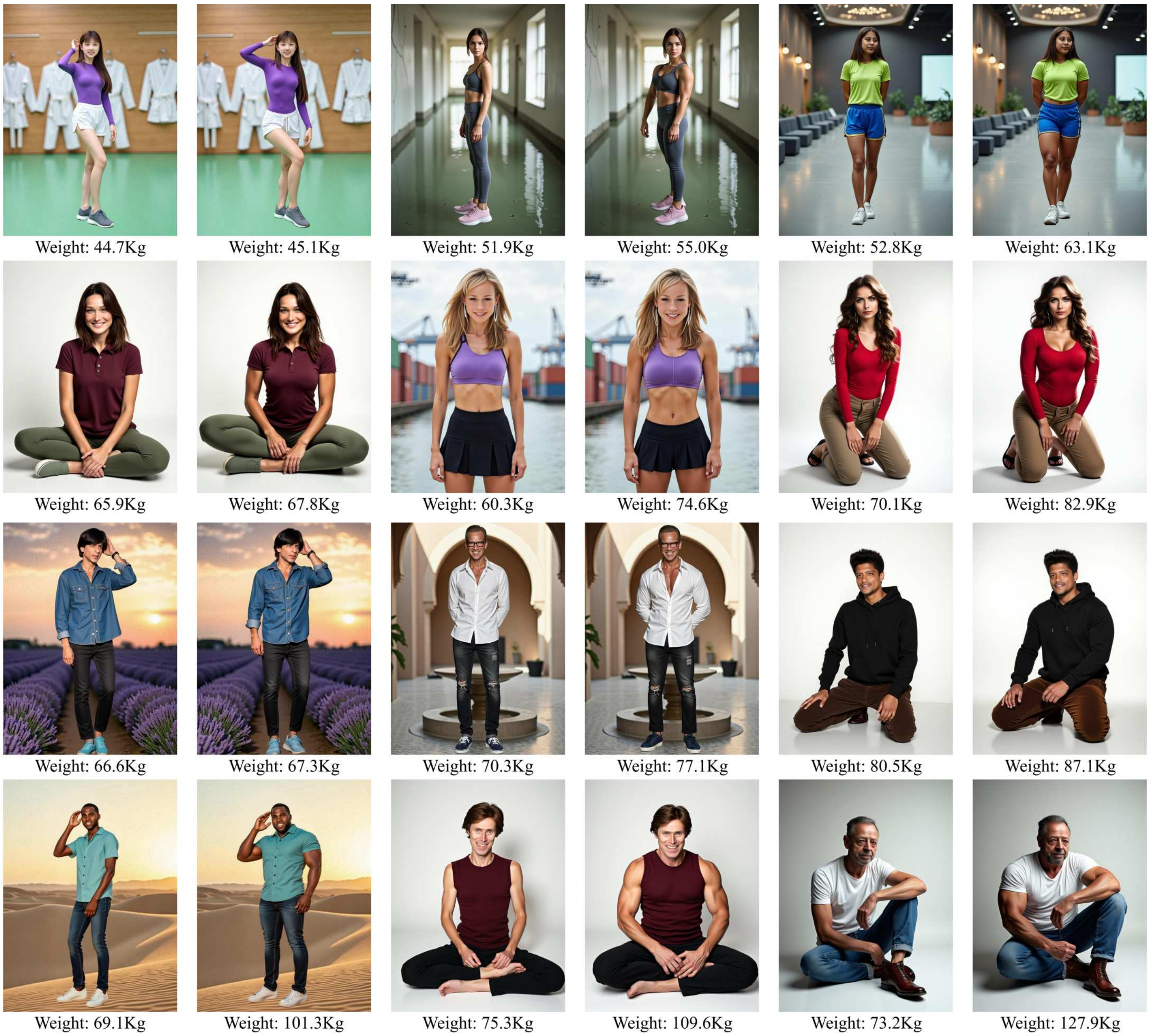

Figure 8: Wide weight range for the muscle body type in our test set.

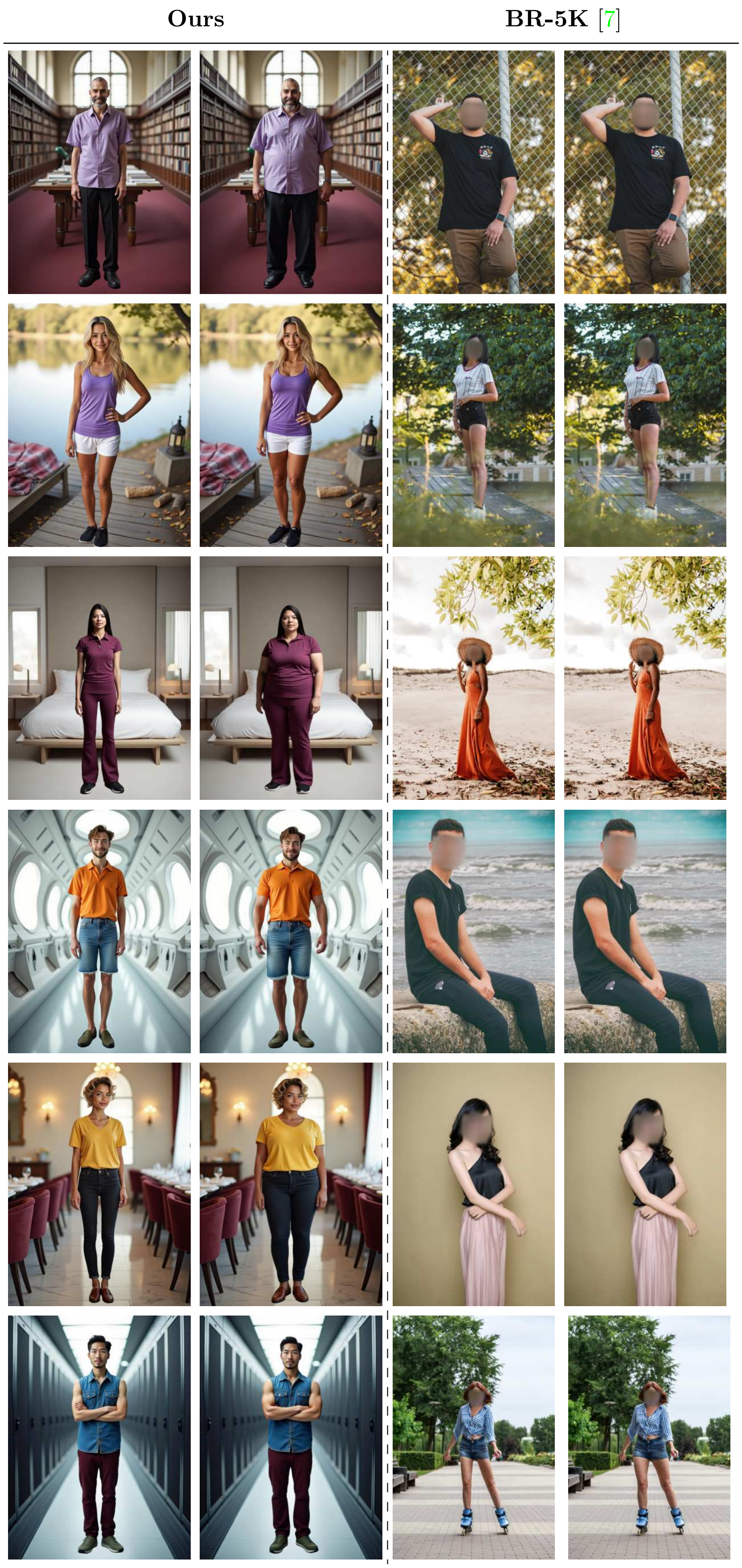

Figure 9: Qualitative comparison. Each row shows a distinct identity. On the left, our dataset exhibits large, identity-preserving body-shape edits. While on the other side, BR-5K dataset [7] pairs show only minor or ambiguous changes and the faces are blurred.

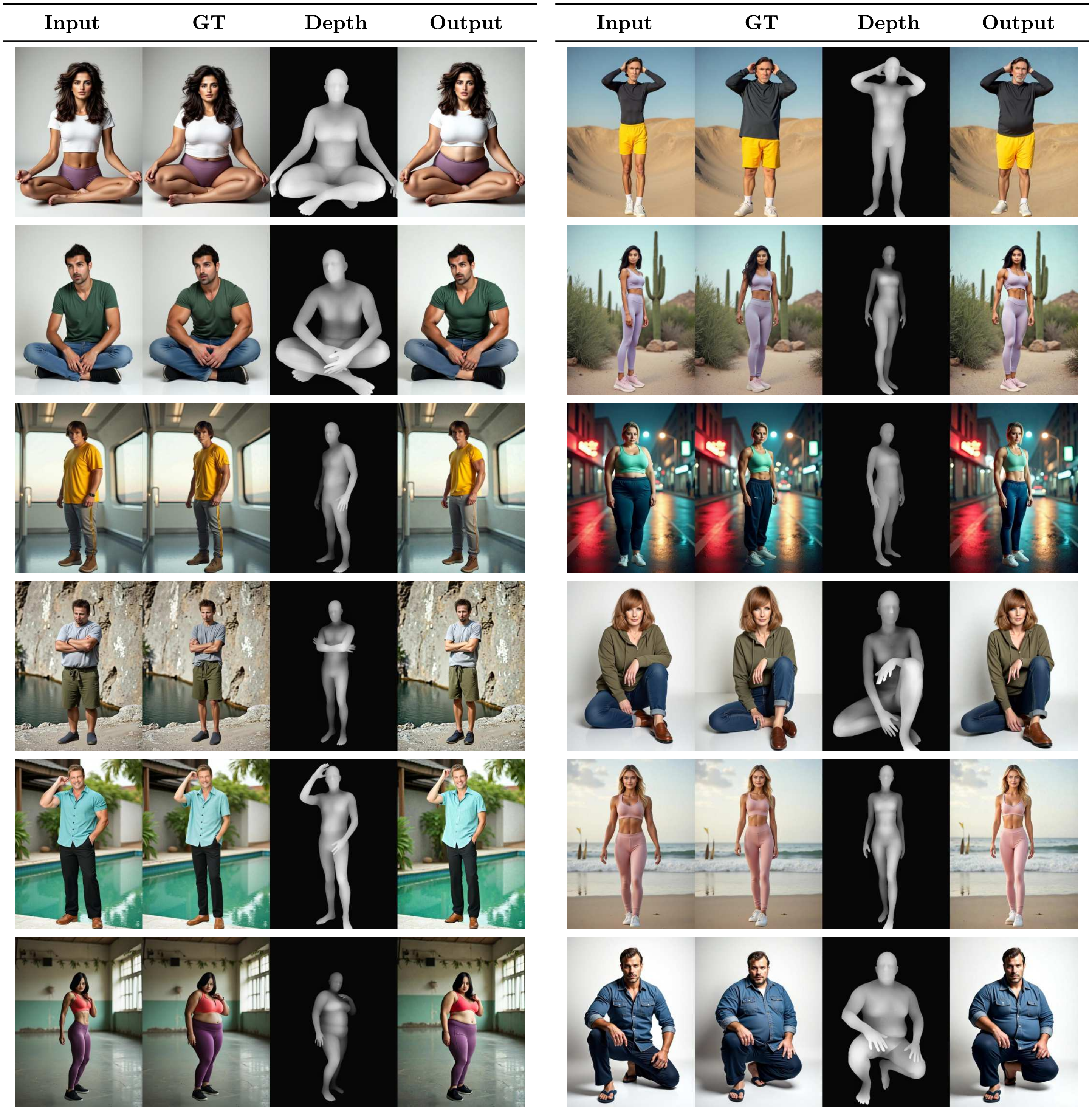

Figure 10: Inference results. The rows from top-to-bottom are the transformations- thin→fat, thin→muscle, fat→muscle, fat→thin, muscle→thin and muscle→fat.

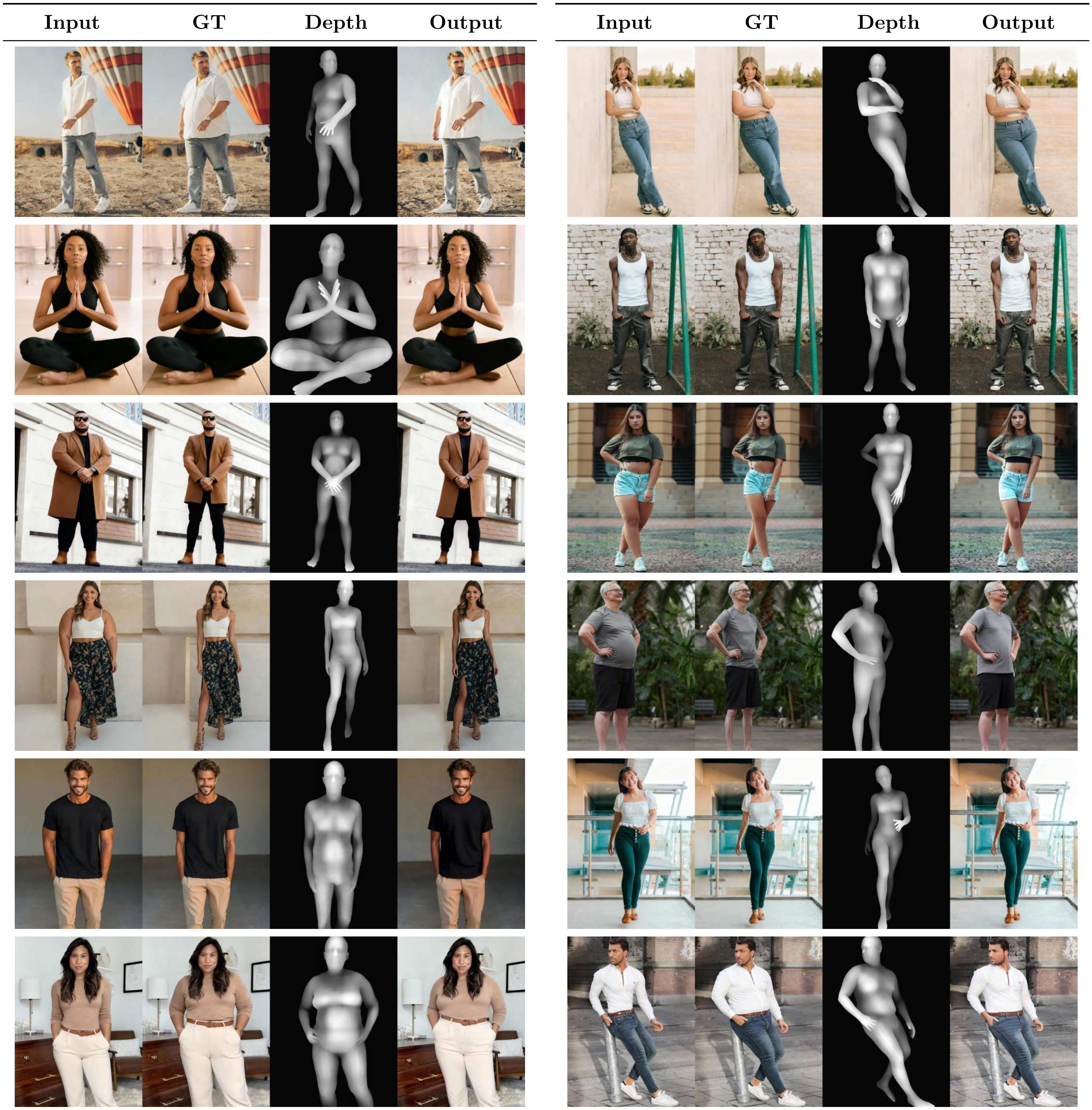

Figure 11: Inference results on dataset generated by Seedream 4.0 [2]. The rows from top-to-bottom are the transformations- thin→fat, thin→muscle, fat→muscle, fat→thin, muscle→thin and muscle→fat.

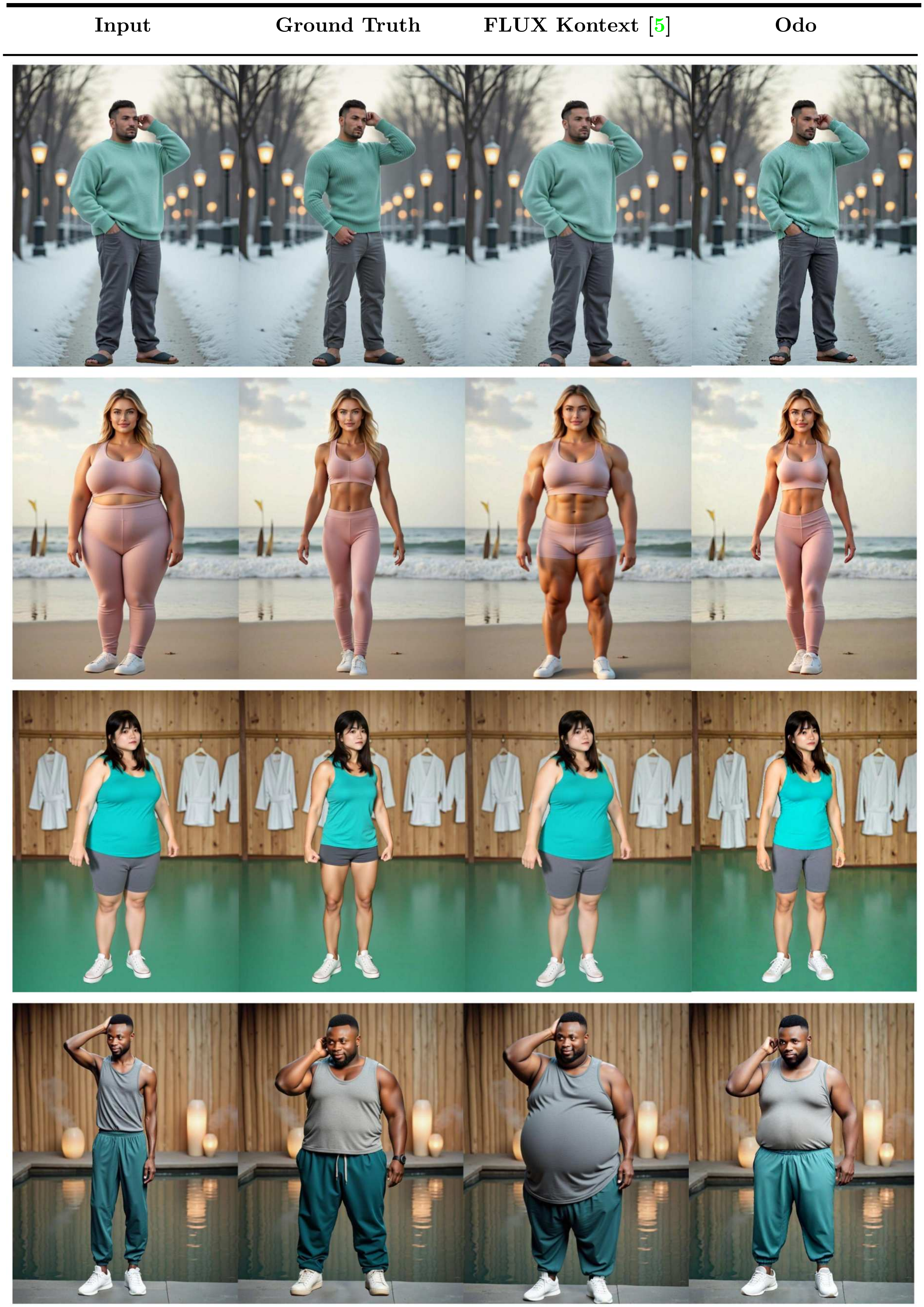

Figure 12: Comparison of FLUX Kontext [5] with Odo.

Input	Ground Truth	FLUX Kontext [5]	Odo

Figure 13: Comparison of FLUX Kontext [5] with Odo.

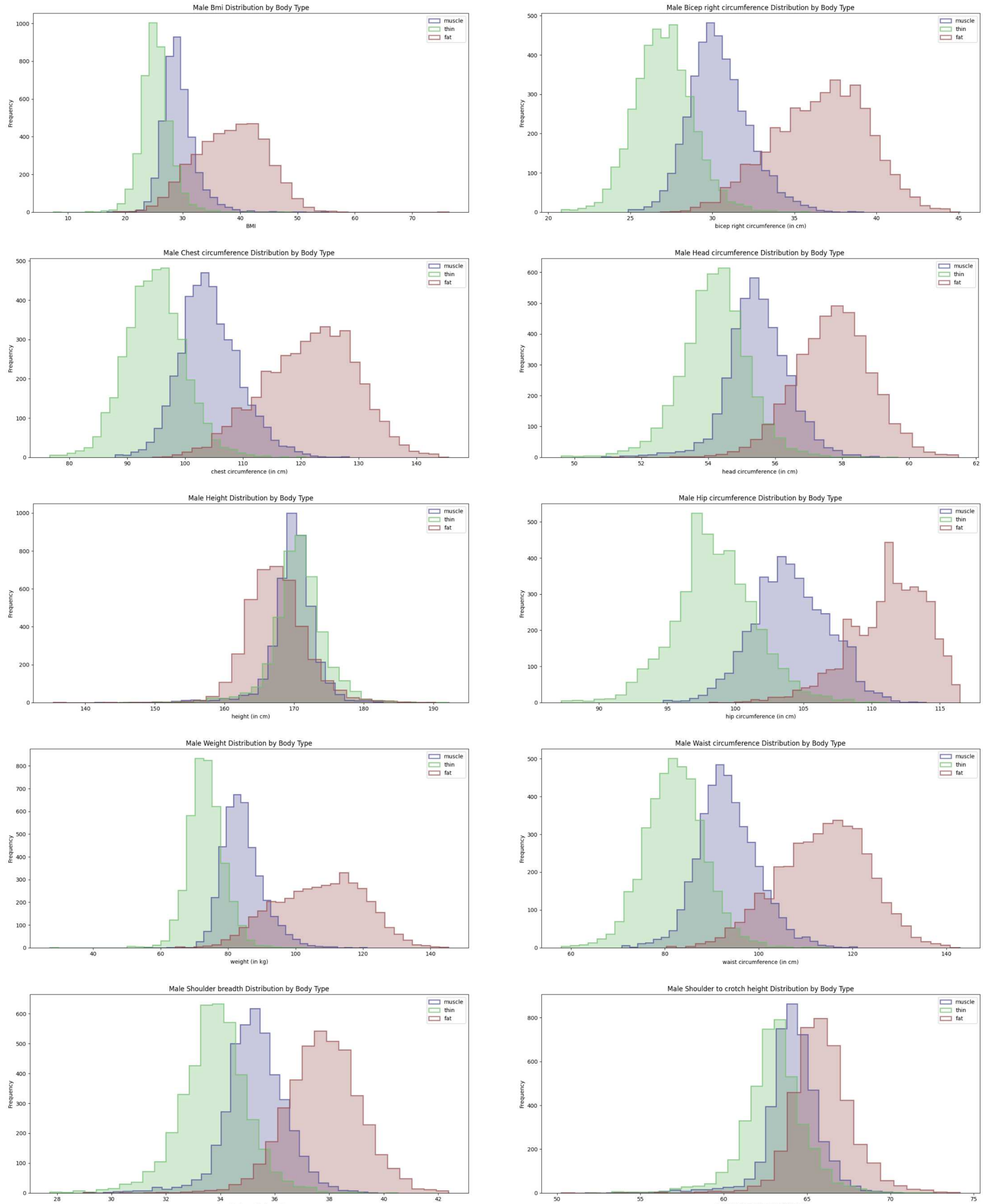

Figure 14: SMPL Anthropometric [1]-measurement distributions for males in our train set. The grpahs from top to bottom and left to right are : BMI, Bicep right circumference (in cm), Chest circumference, Head circumference (in cm), Height distribution (in cm), Hip circumference (in cm), Weight distribution (in kg), Waist circumference (in cm), Shoulder breadth (in cm) and Shoulder-to-crotch height (in cm).

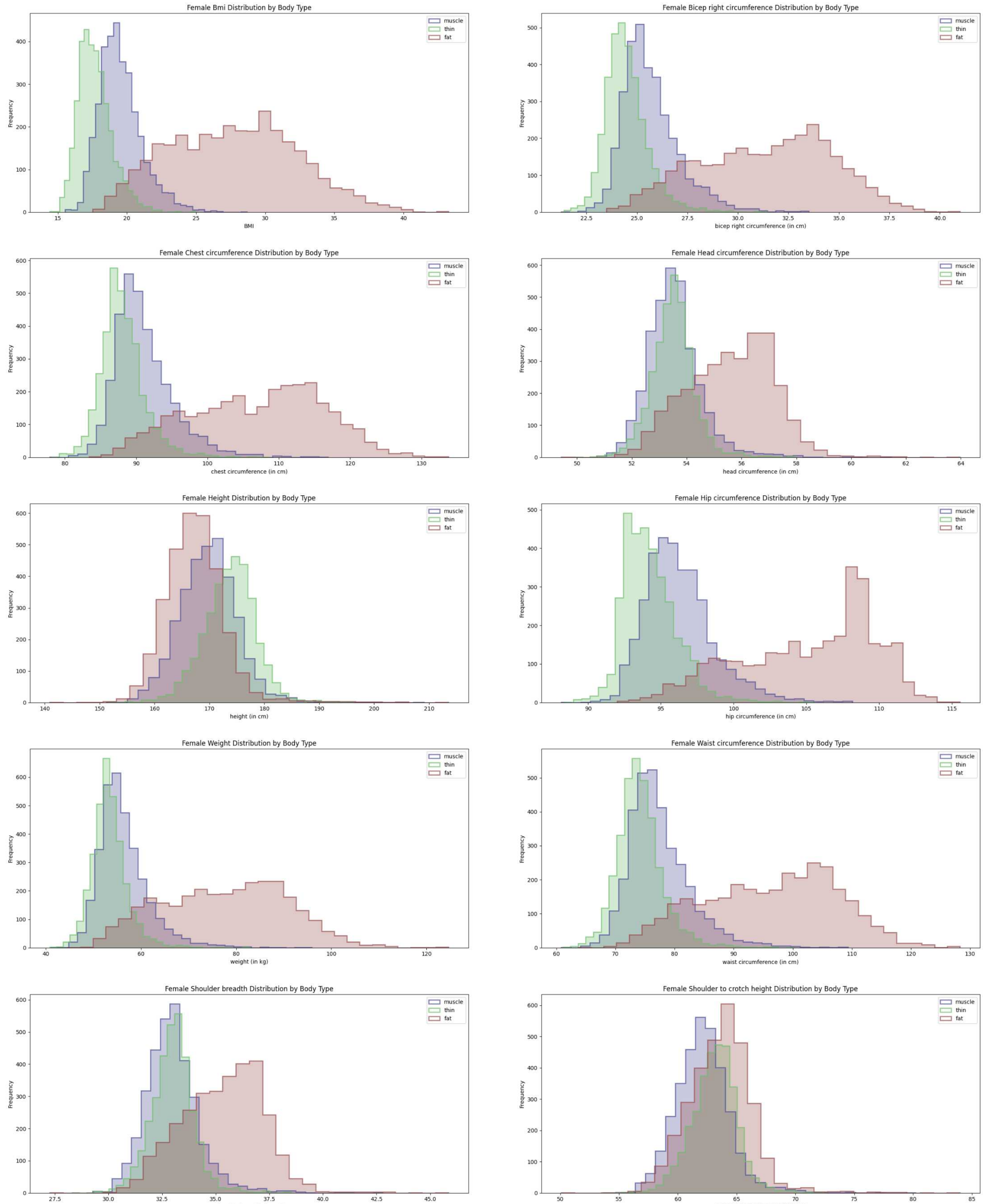

Figure 15: SMPL Anthropometric [1]-measurement distributions for females in our train set. The graphs from top to bottom and left to right are : BMI, Bicep right circumference (in cm), Chest circumference, Head circumference (in cm), Height distribution (in cm), Hip circumference (in cm), Weight distribution (in kg), Waist circumference (in cm), Shoulder breadth (in cm) and Shoulder-to-crotch height (in cm).

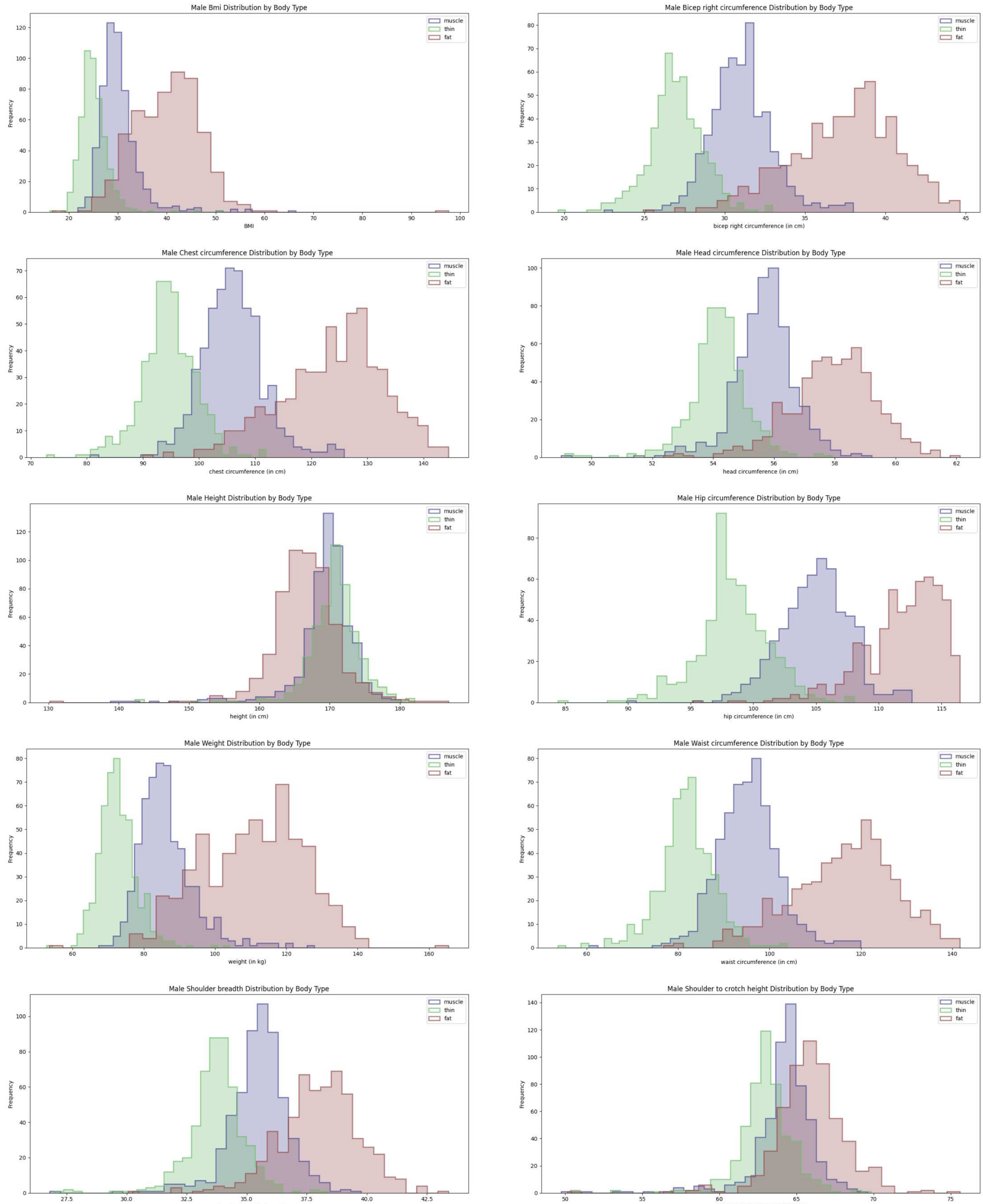

Figure 16: SMPL Anthropometric [1]-measurement distributions for males in our test set. The graphs from top to bottom and left to right are : BMI, Bicep right circumference (in cm), Chest circumference, Head circumference (in cm), Height distribution (in cm), Hip circumference (in cm), Weight distribution (in kg), Waist circumference (in cm), Shoulder breadth (in cm) and Shoulder-to-crotch height (in cm).

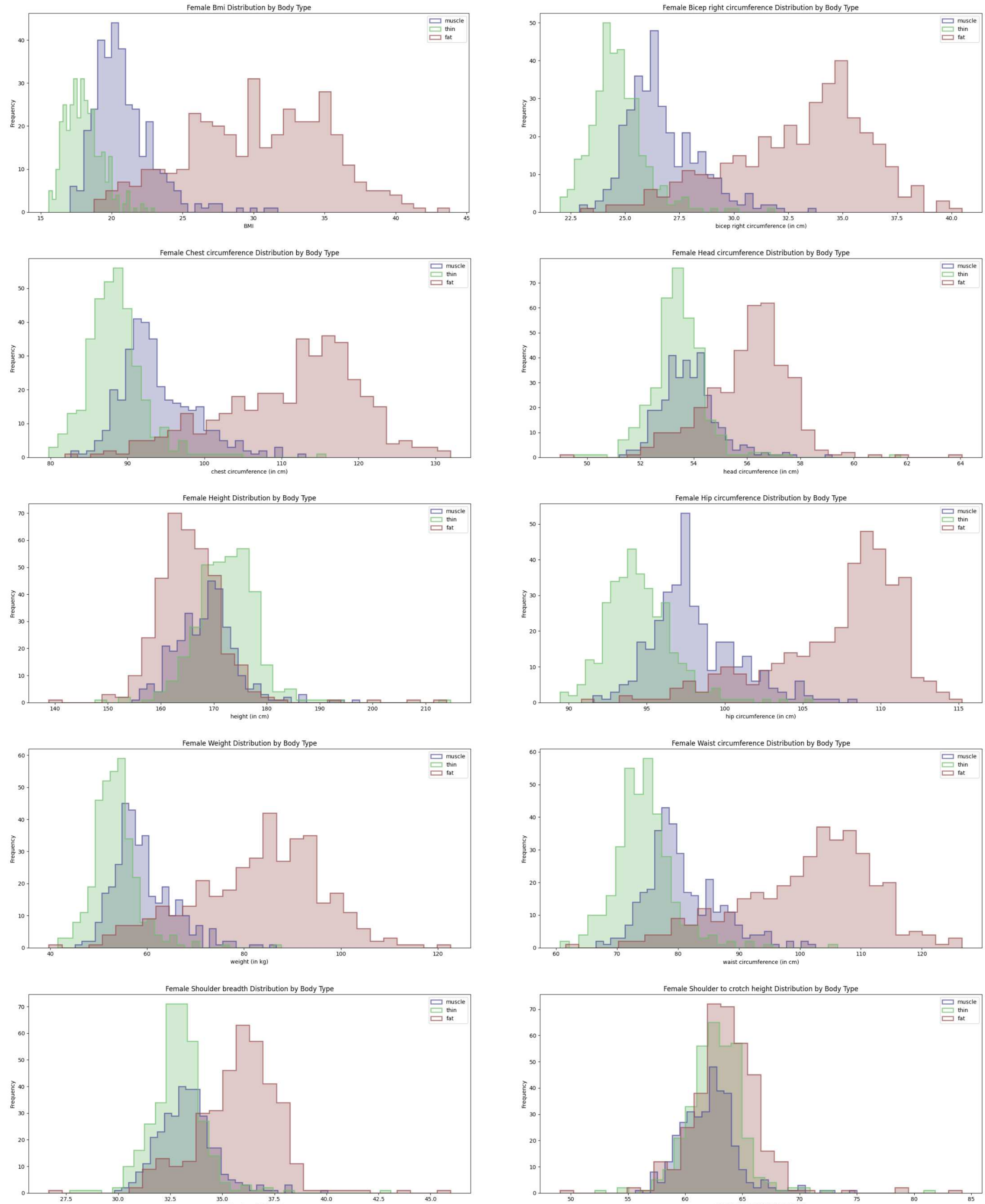

Figure 17: SMPL Anthropometric [1]-measurement distributions for females in our test set. The graphs from top to bottom and left to right are : BMI, Bicep right circumference (in cm), Chest circumference, Head circumference (in cm), Height distribution (in cm), Hip circumference (in cm), Weight distribution (in kg), Waist circumference (in cm), Shoulder breadth (in cm) and Shoulder-to-crotch height (in cm).

References

[1] D. Bojanić. Smpl-anthropometry. https://github.com/DavidBoja/SMPL-Anthropometry, 2024.

[2] ByteDance Seed. Seedream 4.0. https://seed.bytedance.com/en/seedream4_0, 2025.

[3] Zinan Guo, Yanze Wu, Zhuowei Chen, Lang Chen, Peng Zhang, and Qian He. Pulid: Pure and lightning id customization via contrastive alignment. In *Advances in Neural Information Processing Systems*, 2024.

[4] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. 2018.

[5] Black Forest Labs, Stephen Batifol, Andreas Blattmann, Frederic Boesel, Saksham Consul, Cyril Diagne, Tim Dockhorn, Jack English, Zion English, Patrick Esser, Sumith Kulal, Kyle Lacey, Yam Levi, Cheng Li, Dominik Lorenz, Jonas Müller, Dustin Podell, Robin Rombach, Harry Saini, Axel Sauer, and Luke Smith. Flux.1 kontext: Flow matching for in-context image generation and editing in latent space, 2025.

[6] Pixabay Contributors. Pixabay — free images, videos & music. https://pixabay.com, 2025. Accessed: 2025-07-11.

[7] Jianqiang Ren, Yuan Yao, Biwen Lei, Miaomiao Cui, and Xuansong Xie. Structure-aware flow generation for human body reshaping. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7754–7763, 2022.

[8] István Sárándi and Gerard Pons-Moll. Neural localizer fields for continuous 3d human pose and shape estimation. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.