ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Silhouette Body Measurement Benchmarks

Song Yan*, Johan Wirta†, Joni-Kristian Kämäräinen*
*Computing Sciences, Tampere University, Finland
† NOMO Technologies Ltd, Espoo Finland

In this work, we provide a number of datasets that will be made publicly available to facilitate fair comparisons.

Our main contributions are:

1) A novel train/test dataset - **BODY-fit** - for benchmarking silhouette based body measurement methods. The dataset is obtained from the local clothing company who have 2,675 female and 1,474 male 3D scans of their customers.

2) A novel testing dataset - **BODY-rgb** - of recently captured RGB images of 86 males and 108 females and tape measured ground truth.

3) A strong **baseline** which achieves good accuracy on the new datasets and the existing data - CAESAR-fit - provided by Pishchulin et.al [1] and for which we define train/test splits files and generated silhouette images.

[1] Pishchulin, Leonid, et al. "Building statistical shape spaces for 3d human modelling." *Pattern Recognition* 67 (2017): 276-286.

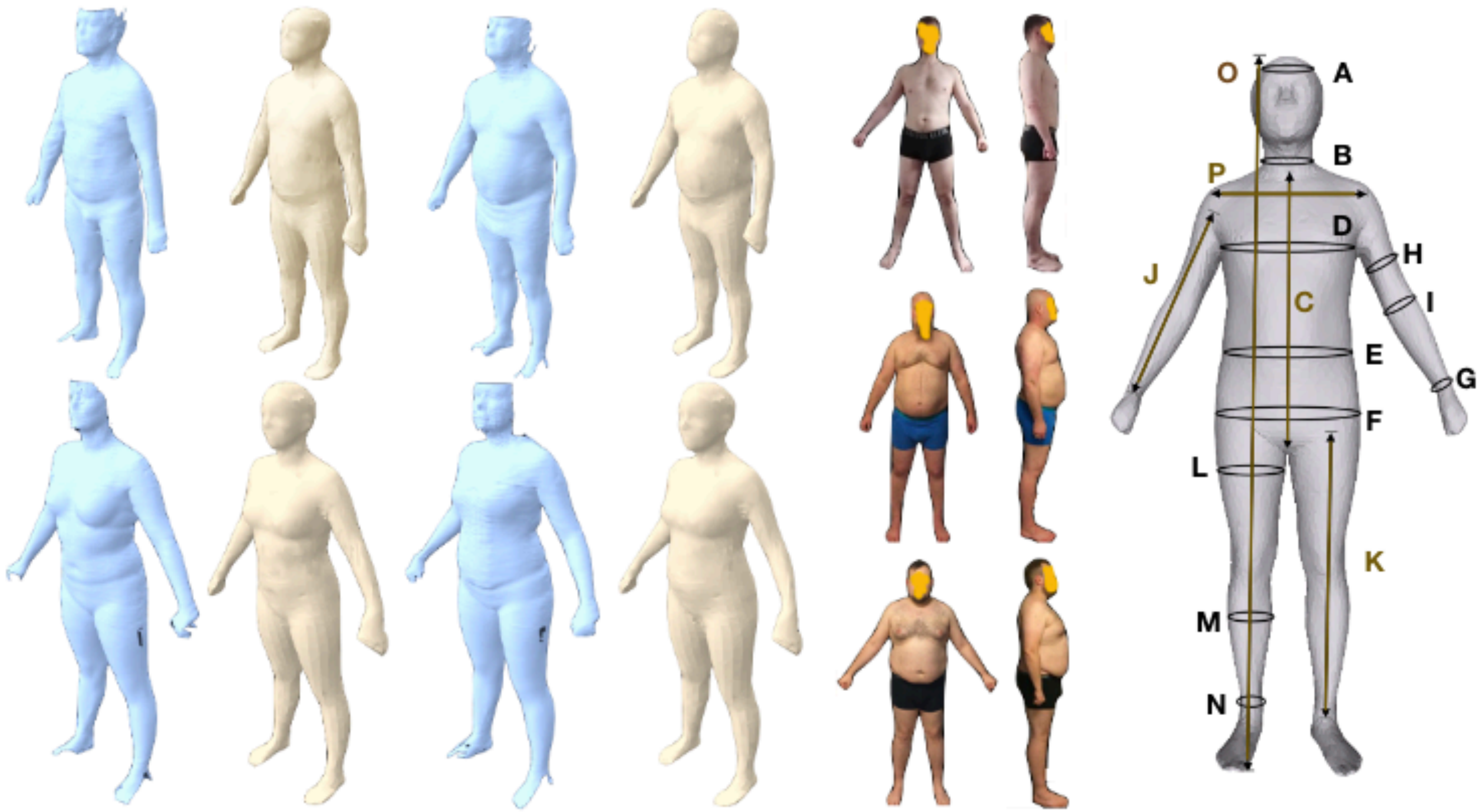

Fig. 1. Examples from the proposed benchmarks, "BODY-fit" and "BODY-rgb", and the 16 body measurements (A-P) used in the method comparison. The blue meshes represent the original BODY scans containing missing points and noise (mainly in the head, feet and hand regions). The yellow meshes result from non-rigid ICP fitting of the mean shape template from the CAESAR fits datasets [24] so that the both datasets now share the same topology. RGB images were captured using Apple iPad.

Silhouette Body Measurement Benchmarks, Song Yan*, Johan Wirta†, Joni-Kristian Kämäräinen*,
*Computing Sciences, Tampere University, Finland, † NOMO Technologies Ltd, Espoo Finland

Related Datasets

Dataset	Description	Available
UMTRI [2]	was collected to find the safest sitting posture of young children in cars.	No
ANSUR 1988 ANSUR 2012	contain 3D scans and 93 tape measured body measurements of US Army Force soldiers. ANSUR 2012 contains 4,082 male and 1,986 female subjects of varying age.	No
CAESAR [3]	contains 3D scans of 2,400 U.S. and Canadian and 2,000 European civilians with tape measured ground truth.	Commercial available
CAESAR-fit	Pishchulin et.al. [1] and Yang et.al. [4] performed mesh registration on the CAESAR scans to bring them in correspondence and use pre-defined geodesic distances as the ground truth.	Yes
Synthetic datasets	to learn a statistical model from CAESAR dataset to synthesize training data.	

[2] Kim, K. Han, et al. *Development of virtual toddler fit models for child safety restraint design*. University of Michigan, Ann Arbor, Transportation Research Institute, 2015.
[3] Robinette, Kathleen M., et al. Civilian american and european surface anthropometry resource (caesar), final report. volume 1. summary. SYTRONICS INC DAYTON OH, 2002.
[4] Yang, Yipin, et al. "Semantic parametric reshaping of human body models." *2014 2nd International Conference on 3D Vision*. Vol. 2. IEEE, 2014.

Related Methods

- Earlier works use engineered features and regression [5], [6], [7], [8], [9], [10], [11].

- Recently deep architectures have become more popular.
  - In [12] the hand-crafted features are extracted from silhouettes and mapped to the shape PCA (Principal Component Analysis) sub-space via the Random Forest regressor, then the body measurements are obtained from the reconstructed meshes.

  - HS-NET [13] learns a global mapping from silhouettes to shape parameters by training CNNs.

  - HKS-Net [14] Dibra et al. firstly construct a rich body shape representation space from the pose invariant Heat Kernel Signature (HKS) descriptors, then learn a mapping from silhouettes to this embedded space.

[5] L. Sigal, A. Balan, and M. J. Black, "Combined discriminative and generative articulated pose and non-rigid shape estimation," in Advances in neural information processing systems, 2008, pp. 1337–1344.
[6] Chen, T.-K. Kim, and R. Cipolla, "Inferring 3d shapes and deformations from single views," in European Conference on Computer Vision.Springer, 2010, pp. 300–313
[7] Chen, D. P. Robertson, and R. Cipolla, "A practical system for modelling body shapes from single view measurements." inBMVC,2011.
[8] Chen, T.-K. Kim, and R. Cipolla, "Silhouette-based object phenotype cognition using 3d shape priors," inComputer Vision (ICCV), 2011IEEE International Conference on. IEEE, 2011, pp. 25–32
[9] J. Boisvert, C. Shu, S. Wuhrer, and P. Xi, "Three-dimensional human shape inference from silhouettes: reconstruction and validation,"Machine vision and applications, vol. 24, no. 1, pp. 145–157, 2013
[10] A. Tsoli, M. Loper, and M. Black, "Model-based anthropometry: Predicting measurements from 3d human scans in multiple poses," inWinterConference on Applications of Computer Vision (WACV), 2014
[11] S. Yan, J. Wirta, and J.-K. Kämäräinen, "Anthropometric clothing measurements from 3d body scans,"Machine Vision and Applications, vol. 31, no. 1, p. 7, 2020.
[12] Dibra, Endri, et al. "Shape from selfies: Human body shape estimation using cca regression forests." *European conference on computer vision*. Springer, Cham, 2016.
[13] Dibra, Endri, et al. "Hs-nets: Estimating human body shape from silhouettes with convolutional neural networks." *2016 fourth international conference on 3D vision (3DV)*. IEEE, 2016.
[14] Dibra, Endri, et al. "Human shape from silhouettes using generative hks descriptors and cross-modal neural networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Silhouette Body Measurement Benchmarks, Song Yan*, Johan Wirta†, Joni-Kristian Kämäräinen*,
*Computing Sciences, Tampere University, Finland, † NOMO Technologies Ltd, Espoo Finland

Proposed Datasets

BODY-fits

1) A local clothing company NOMO3d provided us a dataset of real 3D body scans of people wearing only tight underwear. The scans were captured using a commercial TC2 device and software[1].
2) The scanner outputs a triangulated mesh structure in the regular OBJ file format. Each mesh contains on average 57,000 vertices and around 113,000 faces.
3) Similar to the original CAESAR scans also our scans are of various qualities and contain holes, in particular, near the feet, hand and head regions. To compensate the missing regions, the scans were converted to watertight meshes by applying the non-rigid ICP algorithm of Amberg et al.[15] and a 3D body template.

BODY-rgb

1) In addition to the 3D scan datasets we collected a small dataset of real RGB images of people in underwear.
2) 8-20 body measurements were measured using a tape measure.
3) The dataset consists of 86 male and 108 female subjects.
4) The approximate capturing distance was 2.4m and the camera height 1.6m.
5) Images of front and side views and their manually segmented silhouettes are included.

[1] https://www.tc2.com

[15] Amberg, Brian, Sami Romdhani, and Thomas Vetter. "Optimal step nonrigid ICP algorithms for surface registration." *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2007.

Silhouette Body Measurement Benchmarks, Song Yan*, Johan Wirta†, Joni-Kristian Kämäräinen*,
*Computing Sciences, Tampere University, Finland, † NOMO Technologies Ltd, Espoo Finland

Proposed baseline

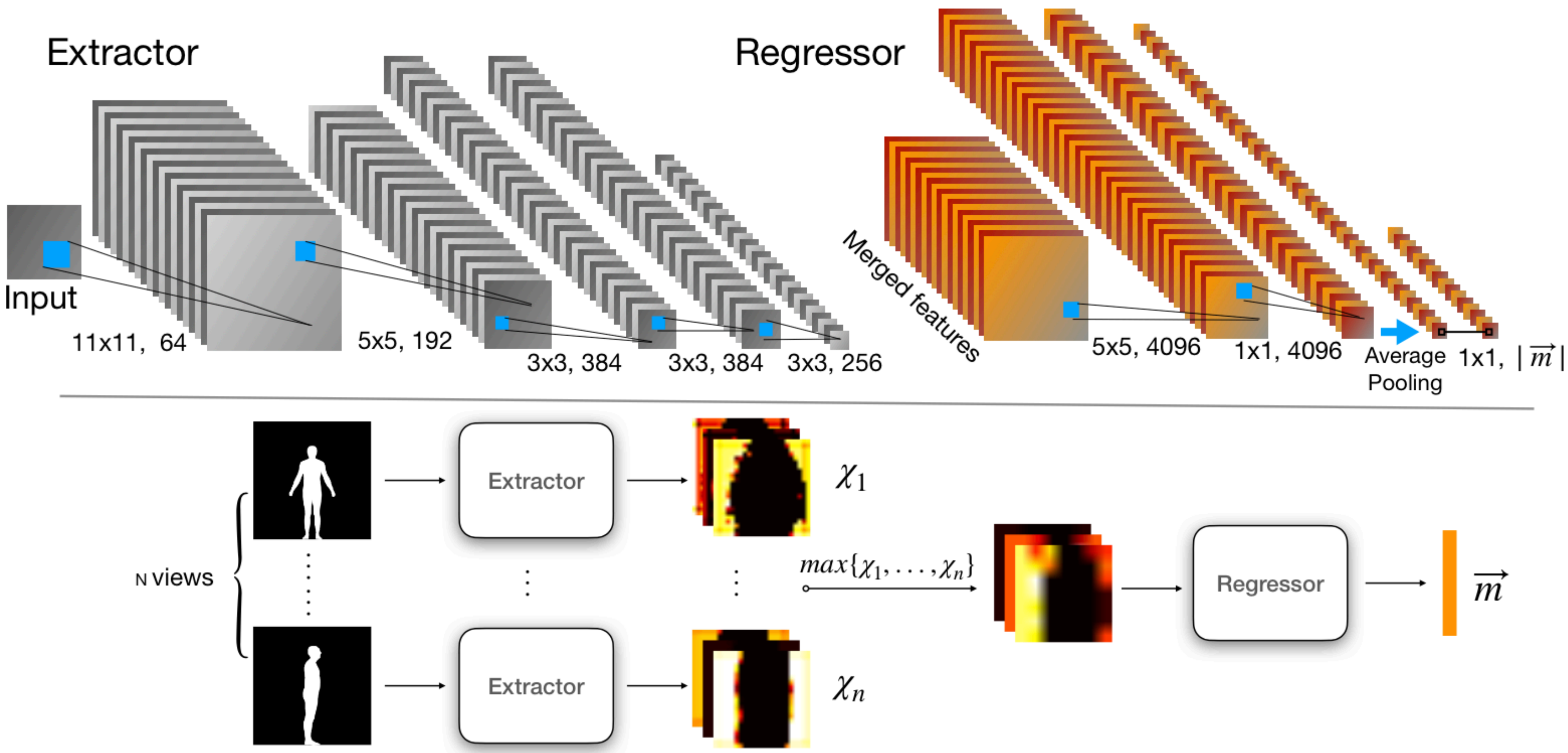

Fig. 2. The overall deep architecture for the baseline method. Blue blocks denote conv kernels, and grey and yellow blocks denote the feature maps. Kernel sizes and the number of output feature maps are shown as $< k \times k, C >$.

Results

TABLE I
COMPARISON OF THE PROPOSED METHOD (OUR) TO THE PRIOR ART (UF-US-2 [14]). UF-US-2 CODE WAS OBTAINED FROM THE ORIGINAL AUTHORS. METHODS WERE TESTED USING THE SAME TRAIN/TEST SPLITS AND ALL UNITS ARE MILLIMETERS (MM).

	CAESAR-fit				BODY-fit			
	Male		Female		Male		Female	
Measure	UF-US-2 [14]	Our	UF-US-2 [14]	Our	UF-US-2 [14]	Our	UF-US-2 [14]	Our
A. Head circ.	10.6	**8.6**	18.1	**15.9**	26.0	**17.2**	13.9	**9.2**
B. Neck circ.	11.6	**9.3**	**11.6**	15.5	13.4	**11.8**	**14.5**	14.6
C. Shoulder-b/c len.	9.9	**5.4**	**10.7**	16.3	12.3	**11.2**	9.4	**7.7**
D. Chest. circ.	27.4	**18.2**	32.3	**24.8**	32.1	**23.0**	26.2	**21.7**
E. Waist circ.	27.6	**17.0**	32.0	**22.9**	42.5	**16.5**	22.3	**17.1**
F. Pelvis circ.	**22.9**	30.6	29.0	**24.0**	24.8	**13.3**	20.6	**14.7**
G. Wrist circ.	**9.5**	10.7	**12.2**	13.3	4.2	**4.1**	**4.8**	5.2
H. Bicep circ.	14.9	**12.5**	16.6	**11.5**	13.8	**11.4**	11.9	**9.3**
I. Forearm circ.	12.4	**7.9**	13.5	**10.7**	8.7	**7.2**	8.6	**8.5**
J. Arm len.	8.9	**4.2**	**8.9**	13.1	9.2	**7.6**	7.4	**6.4**
K. Inside leg len.	**9.8**	13.5	**13.3**	14.8	11.9	**9.2**	10.0	**6.5**
L. Thigh circ.	21.9	**16.5**	28.2	**16.4**	**16.9**	17.8	14.8	**11.6**
M. Calf circ.	12.5	**7.2**	16.0	**10.3**	11.0	**8.8**	13.6	**9.2**
N. Ankle circ.	9.2	**4.6**	10.6	**6.1**	6.4	**5.4**	7.2	**6.1**
O. Overall height	**14.8**	15.1	**20.2**	34.7	25.8	**9.9**	17.1	**8.6**
P. Shoulder breadth	9.0	**5.6**	**9.8**	10.9	12.0	**9.2**	9.3	**7.6**

[14] Dibra, Endri, et al. "Human shape from silhouettes using generative hks descriptors and cross-modal neural networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Results

TABLE II
RESULTS FOR BODY-RGB (5-FOLD-CROSS-VALIDATION) WITH TAPE MEASURED GROUND TRUTH (A, C AND I WERE NOT AVAILABLE). "1ST-STAT" USES A TRAINING SET MEAN AS THE PREDICTION TO ALL TEST SAMPLES.

Measure	BODY-rgb Male 1st-stat	BODY-rgb Male Our	BODY-rgb Female 1-stat	BODY-rgb Female Our
A. Head circ.	-		-	
B. Neck circ.	19.8	14.3	20.3	13.8
C. Shoulder-b/c len.	-	-	-	
D. Chest. circ.	76.1	36.1	101.1	31.7
E. Waist circ.	97.6	35.3	121.9	42.7
F. Pelvis circ.	62.2	35.5	90.4	35.5
G. Wrist circ.	8.5	6.6	8.7	6.9
H. Bicep circ.	27.1	20.9	36.3	19.9
I. Forearm circ.	-	-	-	-
J. Arm len.	27.9	22.5	25.1	18.6
K. Inside leg len.	46.7	31.4	37.1	23.7
L. Thigh circ.	43.8	42.8	62.7	44.3
M. Calf circ.	23.1	12.8	29.6	16.7
N. Ankle circ.	12.3	8.5	17.1	13.8
O. Overall height	59.8	14.3	51.5	19.4
P. Shoulder breadth	21.8	15.8	22.0	19.6

Silhouette Body Measurement Benchmarks, Song Yan*, Johan Wirta†, Joni-Kristian Kämäräinen*,
*Computing Sciences, Tampere University, Finland, † NOMO Technologies Ltd, Espoo Finland

Results

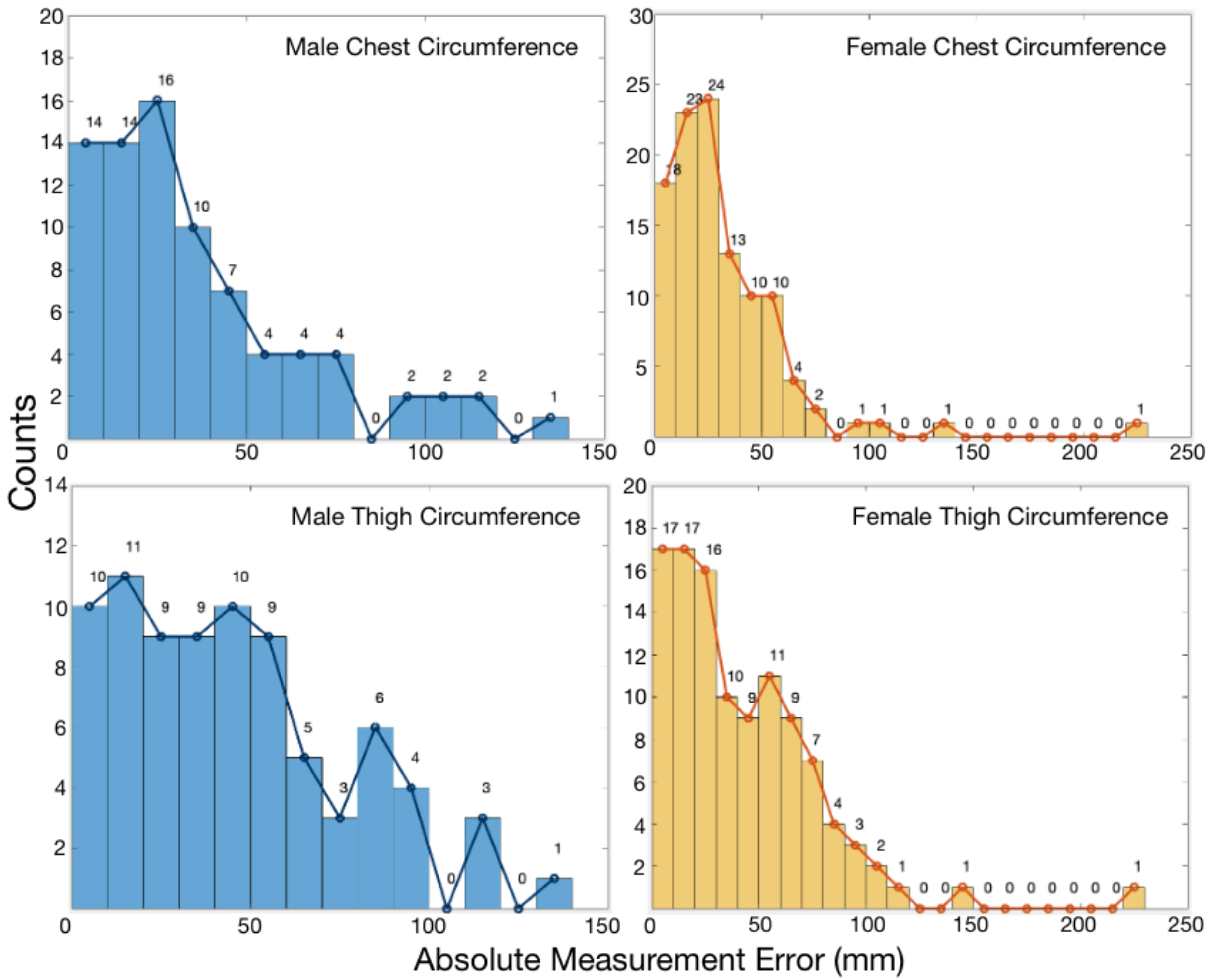

Fig. 3. The error histograms of male/female chest and thigh measurements in our realistic BODY-rgb dataset.

Results

TABLE IV
TEST SET ACCURACY FOR SPECIFIC NETWORKS ($w = 1.0$: SINGLE NETWORK FOR ALL MEASUREMENTS; $w = 10.0$ TARGET MEASUREMENT WEIGHT IS 10.0 AND OTHER MEASUREMENTS 1.0; $w = \infty$: ONLY THE TARGET MEASUREMENT USED).

	CAESAR-fit		
Measure	$w_i = 1.0$	$w_i = 10.0$	$w_i = \infty$
		Male	
A. Head circ.	8.6	**8.2**	11.9
B. Neck circ.	9.3	**9.2**	10.1
C. Shoulder-b/c len.	**5.4**	8.1	19.9
D. Chest. circ.	**18.2**	35.9	34.6
E. Waist circ.	**17.0**	23.5	24.6
F. Pelvis circ.	30.6	30.2	**23.0**
G. Wrist circ.	10.7	**6.8**	8.6
H. Bicep circ.	12.5	14.3	**10.6**
I. Forearm circ.	7.9	**6.9**	9.5
J. Arm len.	**4.2**	15.7	13.1
K. Inside leg len.	13.5	**9.6**	26.8
L. Thigh circ.	16.5	23.4	**15.5**
M. Calf circ.	**7.2**	8.5	10.9
N. Ankle circ.	4.6	**4.5**	6.5
O. Overall height	15.1	**8.3**	22.1
P. Shoulder breadth	**5.6**	5.7	7.6

Silhouette Body Measurement Benchmarks, Song Yan*, Johan Wirta†, Joni-Kristian Kämäräinen*,
*Computing Sciences, Tampere University, Finland, † NOMO Technologies Ltd, Espoo Finland

Conclusion

We introduced new benchmark datasets to boost research on methods that can estimate anthropometric body measurements from image data. The first dataset, BODY-fit, includes 2,675 female and 1,474 male 3D meshes constructed from the scans of real subjects. Similar to previous works, a number of geodesic distance paths on the meshes were measured to provide body measurement ground truth and silhouette images were generated.

We provide the same measurements, similarly generated silhouettes and train/test splits for the existing 1,531 female and 1,517 male CAESAR fitted meshes. Our meshes share the same topology to CAESAR-fit and therefore allows further 3D and 2D cross-dataset comparisons between them. We introduce another realistic dataset of 86 male and 108 female RGB images and corresponding manually made tape measured ground truth (BODY-rgb). As a baseline for these datasets we propose a simple yet effective deep CNN architecture that obtains competitive accuracy on all three datasets.