

# Body Weight Analysis from Human Body Images

Min Jiang, Guodong Guo\*, *Senior Member, IEEE*,

Human body images encode plenty of useful biometric information, such as pupil color, gender, weight, etc. Among this information, body weight is a good indicator of health conditions. Motivated by the recent health science studies, this work investigates the feasibility of analyzing body weight from 2-dimensional (2D) frontal view human body images. The widely used body mass index (BMI) is employed as a measure of body weight. To investigate the problems at different levels of difficulties, three feasibility problems, from easy to hard, are studied. More specifically, a framework is developed for analyzing body weight from human body images. Computation of five anthropometric features is proposed for body weight characterization. Correlation is analyzed between the extracted anthropometric features and the BMI values, which validates the usability of the selected features. A visual-body-to-BMI dataset is collected and cleaned to facilitate the study, which contains 5900 images of 2950 subjects along with the labels corresponding gender, height, and weight. Some interesting results are obtained, demonstrating the feasibility of analyzing body weight from 2D body images. In addition, the proposed method outperforms two state-of-art facial images based weight analysis approaches in most cases.

**Index Terms**—Body weight analysis, visual analysis of Body mass index (BMI), anthropometric features, visual-body-to-BMI dataset

## I. INTRODUCTION

In modern lives, there are various social networks with different functions, such as image sharing, dating, job hunting, and blogging. With the popularity of digital camera, more and more people record their lives via photos or videos and post the records to social media. Photos from social networks contain lots of hard biometric and soft biometric information, such as pupil color, gender, height, weight, age, etc. Such biometric information can be utilized for individual identification [1]–[6]. Among the soft biometric measures, body weight and fat are good indicators of health conditions.

The purpose of this work is to explore the feasibility of body weight analysis from the visual appearance of human body images. We develop some useful cues to characterize body weight/fat from human body images. The widely used body fat indicator—body mass index ( $BMI = \frac{weight(lb)}{height(in)^2} \times 703$ ) is used as a measure for body weight. The BMI has been employed as a measure in many previous studies [7], [8], which is also a risk factor for many diseases. For example, [9], [10] demonstrated that increased BMI is associated with some cancers (such as breast cancer, colon cancer, thyroid cancer, etc.) for both males and females. Wolk et al. [11] presented that BMI is a

Min Jiang and Guodong Guo are with the Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506, USA. email: minjiang.aca@gmail.com and guodong.guo@mail.wvu.edu. Guodong Guo is the corresponding author.

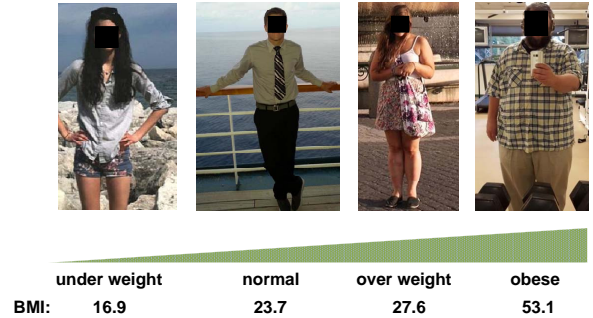


Fig. 1. Some frontal body images with BMI values and corresponding categories. The increase in body adiposity is observed as the BMI value increases.

risk factor for unstable angina and myocardial infarction in patients. Meigs et al. [12] studied the risk of type 2 diabetes and cardiovascular disease (CVD) stratified by BMI. Considering the close connection between BMI and some diseases, BMI is important for personal health monitoring and medical research.

Generally, BMI is measured in person with special devices. For convenient monitoring, this work explores an automatic BMI prediction approach from people’s daily life photos. Our work can be of great benefit to medical researchers to access BMI data from social networks, which may provide lots of sources for health monitoring in large populations.

### A. Motivation

The motivation of this study comes from several aspects. First, from human vision, the body weight/fat can be intuitively observed by humans from 2D images. Some examples of body images with corresponding BMI values are shown in Fig. 1. The increase in body adiposity can be observed by human vision without difficulty. Second, many studies in health science [13]–[16] had shown that some anthropometric measures, such as waist-thigh ratio, waist-hip ratio, waist circumference, etc., are indicators for obesity and are correlated to BMIs. Based on the above intuitive observation (as shown in Fig. 1) and health science studies, we believe that it is worth investigating a computational approach to analyze body weight from human body images.

### B. Related work

There are a few studies working on estimating human body weight or BMI from body related data, such as body measurements, 3-dimensional (3D) body data and RGB-D body images. Velardo et al. [17] studied the body weight directly from anthropometric data (body measurements) collected by NHANES [18]. A polynomial regression model was employed

to analyze the anthropometric data. Cao et al. [19] investigated the use of true measurements of the body (provided by CAESAR 1D database) for the prediction of certain soft biometrics, such as gender and weight. Detailed definitions about many different measurements of the anthropometric feature were included in their work. Velardo et al. [20] studied the weight estimation from 3D human body data by the same anthropometric features as in their previous work [17]. Velardo et al. [21] estimated the weight of a person within 4% error using 2D and 3D data extracted from a low-cost Kinect RGB-D camera output. Nguyen et al. [22] proposed a weight estimator based on single RGB-D images, which utilized the visual color cues, depth, and gender information. Nahavandi et al. [23] presented a skeleton-free Kinect system to estimate BMI of human bodies. Recently, Pfitzner et al. [24] described the estimation of the body weight of a person who is in front of an RGB-D camera with three different poses: lying, standing and walking.

Instead of estimating body weight or BMI from body images, some work analyzed body weight or BMI from face images. Wen and Guo [25] first proposed a computational method for BMI prediction from face images. They also analyzed the correlations between facial features and BMI values. Lee et al. [26] examined 15 2D facial characteristics to identify the strongest predictor of normal and viscerally obese subjects. Later on, Pascali et al. [27] proposed a method for automatic extraction of geometric features, related to weight parameters, from 3D facial data. Kocabey et al. [28] estimated BMIs from face images collected from a social media website. They employed the pre-trained VGG-Net and VGG-Face models to extract features and then utilized a support vector regression model to predict BMI. Recently, Dantcheva et al. [29] explored the possibility of estimating height, weight, and BMI from single-shot facial images by proposing a regression method based on the 50-layer ResNet architecture. All the above approaches require clear frontal view face images as the input.

Some works studied gender or body shape from body images or 3D scanners. Wu et al. [30] explored gender classification from unconstrained and articulated human body images. Cao et al. [31] developed a method based solely on metrological information from facial landmarks of 2D face images for gender prediction and demonstrated that the geometric features achieve comparable performance as appearance features in gender prediction. Gonzalez-Sosa et al. [32] studied gender estimation based on information deduced jointly from face and body and presented two score-level-fusion schemes of the face and body-based features which outperformed the two individual modalities in most cases. Balan et al. [33] studied the markerless human shape and pose capture from multi-camera video sequences using a richly detailed graphics model of 3D human shape. Their approaches required multi-camera video sequences for 3D model reconstruction. Lu et al. [34] collected anthropometric data by 3D whole body scanners, which consist of four sets of laser beams and CCD cameras. [33] and [34] share the limitation of relying on complex 3D data collection devices to generate precise body data.

In contrast to the above works, we propose an approach to analyze body weight just from 2D body images. Neither depth

images nor clear face images are required for this approach. One advantage is that the approach is non-invasive. Moreover, it can work on images with incomplete body parts, even the back view of the body and low-quality images. To the best of our knowledge, this is the first work to explore weight/BMI related information from 2D body images only.

### C. Our contribution

It can be challenging to directly estimate BMI values from 2D human body images. We consider three problems, from easy to hard. By investigating these problems at different levels, one can understand how well the algorithms can address the related problems in real applications. The main contributions of this work are as follows:

- A new visual-body-to-BMI dataset is collected and cleaned, containing 5900 images of 2950 subjects (each contains a pair of images), which is the first dataset of its kind.
- A computational framework is developed for body weight and BMI analysis from 2D human body images, which can process either a single image or a pair of images.
- Five anthropometric features are proposed for body weight analysis from 2D body images. Computational methods are developed to extract these features and map them into weight/BMI values.

The remainder of the paper begins with describing the newly collected and cleaned visual-body-to-BMI dataset in Section II. Section III introduces the three problems we study and presents the framework for body weight analysis. Details about the feature detection and computational method are given in Section IV. Section V describes the employed machine learning models. The calculation of Pearson's correlation coefficient and the metrics used to evaluate the performance are presented in Section VI. In Section VII, we first calculate the correlation between the extracted features and the BMI values; and then provide the detailed experimental results and discussion. Finally, conclusions and future work are given in Section VIII.

## II. DATASET WITH CLEANING

The human body images are downloaded from the website Reddit posts<sup>1</sup>. In total there are 47,574 images of 16,483 individuals. Each individual has at least one "previous" and one "current" images (or a collage which was made by sticking several images). As shown in Fig. 2, all the images under the same individual folder have the same annotations (except the image number). The format of the original annotation is "ID\_image number\_previous weight\_current weight\_height\_gender". Thereby, all the images under the same individual folder share the same information about weights ("previous" and "current") and height, the weight for each image cannot be automatically distinguished by algorithms. It needs manual processing (visually check) to correct the weight for the individuals.

We processed and cleaned the dataset with automatic and manual steps which are described below. First, we went

<sup>1</sup>Website: <http://www.reddit.com/r/progresspics>



Fig. 2. The illustration of cleaning images with automatic and manual steps. Two cases are given. The first case (left panel) shows the individual just contains one collage (an image made by sticking several images). The second case (right panel) shows the individual contains 3 images, among them there are 2 group photos (more than one person shown on the image). The blue arrow represents the process of cropping every single body from a composite image based on automated body detection. The orange arrow represents the manual process of correcting annotations. The annotations for the “previous” and “current” images are visually distinguished by body size and shape.

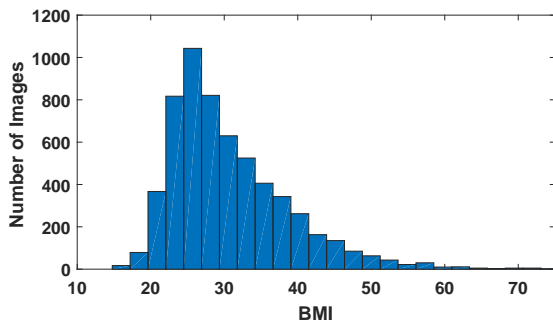


Fig. 3. The distribution of BMI values in the body-to-BMI dataset. The BMI distribution is in a wide range from 15 to 75.

through the original images by a body detector, using a method similar to [35]. Then, given the detected bodies, each single body image was cropped from the original images. During the process, we kept the cropped body images containing both head and frontal body (with required joints detected). If there are greater than or equal to 2 cropped images kept for an individual, the algorithm keeps the individual folder. Now the left (cropped) images under the same individual folder still share the same annotation (“ID\_image number\_crop number\_previous weight\_current weight\_height\_gender”). The next step was to visually distinguish which image has the “previous” weight and which has the “current” weight. Since the annotations only have the “previous” and “current” body weights for each individual, just one “previous” image and one “current” image were kept for each individual. Finally, we manually corrected the annotations for these images.

Fig. 2 shows the procedure of processing the images with automatic and manual steps. Two cases are introduced: the first case shows the individual just contains one collage (an image made by sticking several images); the second case shows the individual contains 3 images, among them there are 2 group

photos (more than one person shown in the image). The blue arrow represents the automatic process of cropping every single body from the images. Each cropped body in the image is labeled by a red boundary box. The orange arrow represents the manual process of distinguishing and correcting annotations. The annotations for the “previous” and “current” body images are visually distinguished by body size and shape. A pair of images mentioned throughout this work is one “previous” and one “current” body images from the same individual.

After these procedures, there are 2950 subjects (individuals) left, each contains two images: one “previous” and one “current”. This leads to a total of 5900 images with corresponding labels of gender, height, and weight. The set of images is noted as visual-body-to-BMI dataset<sup>2</sup> containing 966 females and 1984 males. The ground truth of BMI can be calculated. The BMIs distribution of the body-to-BMI dataset is shown in Fig. 3. The BMIs distribution is in a wide range from 15 to 75. Specifically, 46 body images are in the underweight range ( $BMI \leq 18.5$ ), 1416 are normal ( $18.5 < BMI \leq 25$ ), 1863 are overweight ( $25 < BMI \leq 30$ ) and 2575 are obese ( $BMI > 30$ ). By comparing the weight of the “previous” and “current” images, we conclude that 1246 subjects show an increase in weight, 1233 subjects show a decrease in weight, and the rest 481 subjects have the same weight in both images. The height of each subject remains the same in a pair of images. The subjects are natural with various clothing styles.

### III. PROBLEMS TO STUDY

The studies in health science [13]–[15] show evidence on the relation between some anthropometric measures and obesity. Considering that BMI is a widely used body weight/fat indicator, we employ this index as the measure for the body weight. This work explores the relationship between the BMI values and visual appearances of the human body. The correlation

<sup>2</sup>Please contact the authors for the dataset.



Fig. 4. Three kinds of problems explored for body weight analysis. For the pairwise images, the change is from the left one to right one.

between BMI values and the computed anthropometric features is studied first. Given the correlation, we analyze the body weight issue from 2D human body images at different levels of difficulties (from easy to hard, based on human perception).

Fig. 4 shows the three problems studied for body weight analysis. First, we recognize the weight difference from a pair of frontal view body images. This is defined as a three-class classification problem. The output is a triple classification result which decides whether the weight of the subject is increased, decreased or keeping the same. In our dataset, the height of each subject remains the same height in the corresponding pairwise images, thereby the weight difference is equivalent to BMI difference. Then, we go further and estimate how big the weight or BMI difference between the pairwise images is. The above two problems are studied based on the pairwise images from the same individual. The key of these two problems is to measure whether the change between the two images can be computed or not. A more challenging task is to directly estimate the BMI value from a single body image.

Fig. 5 depicts the framework of the body weight analysis approach, which consists of three steps:

- 1) Body contour and skeleton joints detection.
- 2) Anthropometric feature computation from the body images.
- 3) Apply statistical models to map the features to the weight differences or the BMI values.

As shown in Fig. 5, the approach can classify the weight difference from the pairwise images. The classification is an output of three different results  $\{0, 1, -1\}$ :  $0$  indicates no weight change,  $1$  indicates weight increased, and  $-1$  indicates weight decreased. The order of the images in the pair does matter, and the change is from “previous” (left) to “current” (right), as indicated in Fig. 4. Note that the prediction of the BMI differences and the BMI values are solved by two different regression models. The details about feature extraction and mapping will be given in Sections IV and V, respectively.

#### IV. FEATURE EXTRACTION

In this section, we present the details about feature extraction for the proposed approach. Body contour and skeleton joints

(CSJ) detection is the first step for feature extraction. The output of the detection is used for anthropometric feature computation.

##### A. Contour and skeleton joints detection

Body contour and skeleton joints (CSJ) detection are based on deep networks for contour and skeleton joints detection. Fig. 6 shows the body contour and skeleton joints detected by the CSJ detector. The brick red area represents the detected body part. The asterisks represent the detected skeleton joints.

In order to detect the body contour from an image, pixel-level image segmentation is applied to it. We use the conditional random fields as recurrent neural networks (CRF-RNN) method [36] for body detection. The mean-field CRF inference is reformulated as an RNN, then the CRF-RNN layer (iterative mean-field layer) is plugged into a fully convolutional neural network (FCN). By applying the CRF-RNN method to the image, the body regions are labeled out, while all other regions in the image are labeled as the background. This leads to a set  $B$  contains all pixels which are labeled as the human body, and a set  $G$  contains all pixels which are labeled as the background.  $l_{x,y}$  represents the label assigned to the pixel locates at  $(x, y)$ , where  $(x, y)$  denotes the horizontal and vertical coordinates on the image.  $l_{x,y}$  is from a pre-defined set of labels  $L = \{b, g\}$ . Here  $b$  is the label for the human body and  $g$  is for the background. Then  $B = \{(x, y) : l_{x,y} = b\}$  and  $G = \{(x, y) : l_{x,y} = g\}$ . We will use this in Section IV-B for computing anthropometric features.

With the locations of skeleton joints in an image, the key parts (waist, hip, etc.) are located for extracting the anthropometric features. In this work, the convolutional pose machine (CPM) [37] is employed to detect the skeleton joints from body images. CPM consists of a series of convolutional neural networks (CNN) that repeatedly produce 2D belief maps for the location of each body part. The belief maps produced by the previous CNN are used as the input of the next CNN. By using the CPM, a list of coordinates of the key skeleton joints can be obtained, such as left hip, right hip, left shoulder, and right shoulder, etc. The coordinates of skeleton joints will be used for computing anthropometric features.

##### B. Anthropometric feature computation

Several anthropometric indicators suggested in health science [13]–[16] are used as measures for the obesity. Some listed indicators include waist-thigh ratio, waist-hip ratio, abdominal sagittal diameter, waist circumference, and hip circumference. Taking into account these indicators, we have five anthropometric features automatically detected and computed from body images, including waist width to thigh width ratio ( $WTR$ ), waist width to hip width ratio ( $WHpR$ ), waist width to head width ratio ( $WHdR$ ), hip width to head width ratio ( $HpHdR$ ), and body area between waist and hip ( $Area$ ). Among these features,  $Area$  is inspired by our human perception.

The measurement of the waist circumference and the hip circumference cannot be directly obtained from 2D images. We consider the particular body part as a cylinder. Then we use the width of the body part (on a 2D image) to approximately represent the circumference of a particular body part. Similar



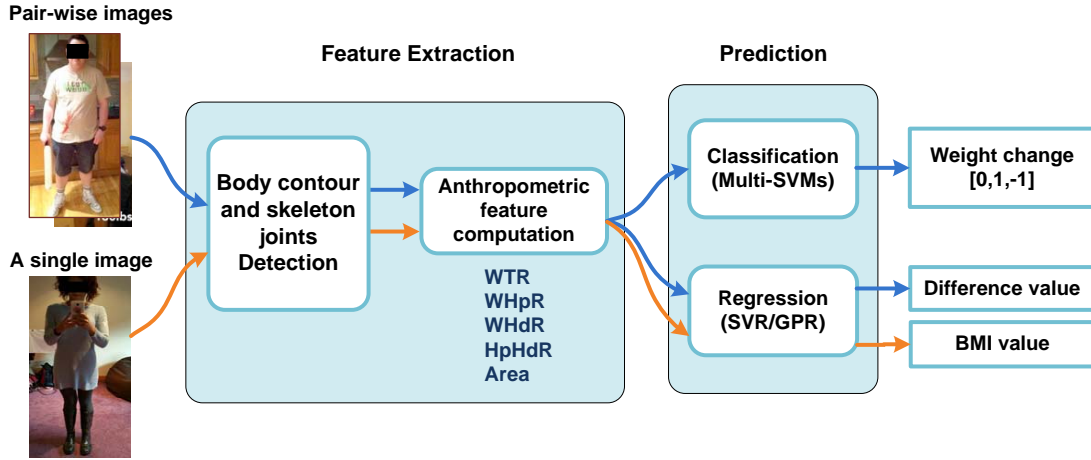


Fig. 5. The framework of our proposed weight analysis approach. The approach can take either pairwise body images or a single image as input. It classifies and predicts the BMI difference from pairwise images, or estimate the exact BMI value from a single image.



Fig. 6. The body contour and skeleton joints detected by the CSJ detector. The brick red area represent the detected body part. The asterisks represents the skeleton joints.

TABLE I  
ABBREVIATIONS OF BODY PARTS FOR FEATURE COMPUTATION.

Body part	Abbrev.	Body part	Abbrev.
Nose	n	Hip	h
Left ear	le	Left hip	lh
Right ear	re	Right hip	rh
Center shoulder	cs	Left hip boundary	lhb
Waist	w	Right hip boundary	rhb
Left waist	lw	Thigh	t
Right waist	rw	Left thigh boundary	ltb
Left waist boundary	lwb	Right thigh boundary	rtb
Right waist boundary	rwb	Knee	k
Left knee	lk	Right knee	rk

approximation has been utilized and verified in [17]. They used the width of the upper arm, leg, waist, and calf to test a polynomial regression model, which was trained by the real circumferences of the body parts. Since the absolute measures of the waist width and hip width cannot be obtained from 2D images without metric/scale information, thereby we compute the ratio to characterize the relative measures.

Fig. 7 illustrates the anthropometric features visually. There are 18 detected skeleton joints shown in the figure labeled with asterisks. In the following, we use the coordinates of 8 detected skeleton joints for computing anthropometric features. These 8 skeleton joints are the nose, left ear, right ear, center shoulder,

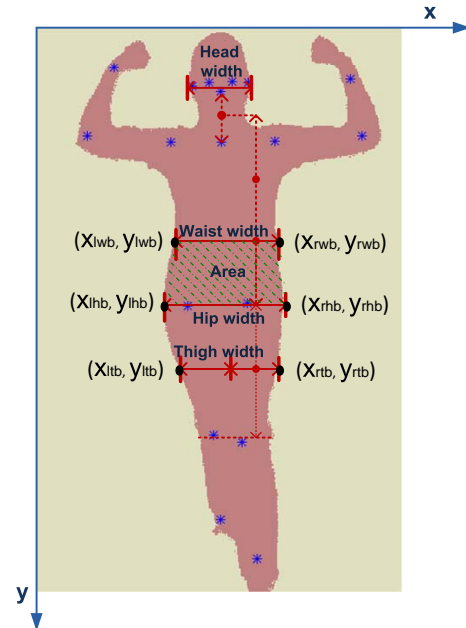


Fig. 7. The anthropometric features computed for body weight analysis. The 18 skeleton joints (labeled by asterisks) are nose, left eye, right eye, left ear, right ear, center shoulder, left shoulder, right shoulder, left elbow, left hand, right elbow, right hand, left hip, right hip, left knee, right knee, left ankle and right ankle. The area filled with green dash dots denotes the feature *Area*.

left hip, right hip, left knee and right knee. The abbreviations of skeleton joints or boundaries involved for feature computation are given in Table I. The abbreviation of a body part is used as an index which denotes the location of the pixel. For example, the pixel of left hip is denoted as  $p_{lh}$ , and its coordinate is denoted as  $(x_{lh}, y_{lh})$ . The size of the input image is  $M \times N$ . The methods for computing the five anthropometric features are described below:

1) *WTR*: the ratio of waist width to thigh width. A general knowledge about human body proportions [38] is used to initially estimate the location of waist and thigh based on the detected locations of hip and head. As shown in Fig. 7, the vertical location of the waist is computed by:  $y_w = \frac{2}{3}y_h +$

$\frac{1}{6}(y_n + y_{cs})$ , where  $y_h = \frac{1}{2}(y_{lh} + y_{rh})$ . Similarly, the vertical location of the thigh  $y_t = \frac{1}{2}(y_k + y_h)$ , where  $y_k = \frac{1}{2}(y_{lk} + y_{rk})$ . With the vertical locations of waist and thigh, the next step is to estimate the waist width and thigh width. Taking waist width as an example, this calculation can be considered as fixing  $y = y_w$ , and searching for the x-axis coordinates of the left and right waist boundaries  $x_{lwb}$  and  $x_{rwb}$  from the contour image. The x-axis coordinate of left waist boundary  $x_{lwb}$  can be computed by:

$$\begin{aligned} x_{lwb} &= \underset{x}{\operatorname{argmin}} |x - x_{cw}|, \\ \text{s.t. } x &\in [0, x_{cw}], (x, y_w) \in G. \end{aligned} \quad (1)$$

Here  $x_{cw}$  is x-axis coordinate of the center waist, which can be approximated by x-axis coordinate of the center shoulder  $x_{cs}$ .  $y_{lw}$  and  $y_{rw}$  both are equal to  $y_w$ .  $G$  is a set contains all pixels labeled as the background. Similarly,  $x_{rwb}$  is given by:

$$\begin{aligned} x_{rwb} &= \underset{x}{\operatorname{argmin}} |x - x_{cw}|, \\ \text{s.t. } x &\in [x_{cw}, M], (x, y_{rw}) \in G. \end{aligned} \quad (2)$$

Here  $M$  is the width of the image.  $y_{lwb}$  and  $y_{rwb}$  both are equal to  $y_w$ . The thigh boundary along the x-axis is determined by  $x_{ltb}$  and  $x_{rtb}$ , which can be calculated in the same way as Eqns. (1) and (2). With the coordinates of these boundaries, the waist width is the Euclidean distance between  $p_{lwb}$  and  $p_{rwb}$ . Thigh width is half of the Euclidean distance between  $p_{ltb}$  and  $p_{rtb}$ . So  $WTR$  is computed by:

$$WTR = \frac{d(p_{lwb}, p_{rwb})}{0.5 \cdot d(p_{ltb}, p_{rtb})}, \quad (3)$$

where  $d(\cdot)$  denotes the Euclidean distance between the two pixels.

2)  $WHpR$ : the ratio of waist width to hip width. Given the left and right hip skeleton joints  $p_{lh}$  and  $p_{rh}$ , the left hip boundary  $p_{lhb}$  and right hip boundary  $p_{rhb}$  are calculated following the rules in Eqns. (1) and (2). Then hip width is the Euclidean distance between  $p_{lhb}$  and  $p_{rhb}$ . The  $WHpR$  is computed by:

$$WHpR = \frac{d(p_{lwb}, p_{rwb})}{d(p_{lhb}, p_{rhb})}. \quad (4)$$

3)  $WHdR$ : the ratio of waist width to head width. Since the images have different scales, the waist widths computed from the images cannot directly represent the measured waist width. According to the anthropometry study on adult head circumferences [39], there are tiny differences on the width of adult heads. Thereby,  $WHdR$  is computed to represent the waist width. Here head width is the Euclidean distance between left ear  $p_{le}$  and right ear  $p_{re}$ . Then  $WHdR$  is given by:

$$WHdR = \frac{d(p_{lwb}, p_{rwb})}{d(p_{le}, p_{re})}. \quad (5)$$

4)  $HpHdR$ : the ratio of hip width to head width. As described above, we use this ratio to represent the hip width in each body image. The  $HpHdR$  is computed by:

$$HpHdR = \frac{d(p_{lhb}, p_{rhb})}{d(p_{le}, p_{re})}. \quad (6)$$

5)  $Area$ : the area between waist and hip. Because of the unknown scale information for each image,  $Area$  is expressed as the number of pixels per unit area between waist and hip. The number of pixels between waist and hip is given by:

$$\#pixels = \sum_{\substack{x \in [0, M] \\ y \in [y_w, y_h]}} \mathbb{1}[l_{x,y} = b], \quad (7)$$

here  $\mathbb{1}[\cdot]$  is an indicator function.  $l_{x,y}$  represents the label (obtained from CSJ detection) assigned to the pixel that locates at  $(x, y)$ . Then the  $Area$  is calculated by:

$$Area = \frac{\#pixels}{(y_h - y_w) \cdot 0.5 \cdot [d(p_{lwb}, p_{rwb}) + d(p_{lhb}, p_{rhb})]} \quad (8)$$

As shown in Fig. 5, the prediction approach can take the different input: either a pair of images or a single image. The BMI difference can be classified and estimated from a pair of input images. On the other hand, the BMI value can be estimated from a single body image. Five anthropometric features are extracted from each body image, resulting in a feature vector  $\mathbf{f} = [WTR, WHpR, WHdR, HpHdR, Area]^T$ . For a single image,  $\mathbf{f}$  is the feature used for estimation. For pairwise images, the following transformation is applied to the features  $\mathbf{f}_1$  and  $\mathbf{f}_2$  for generating the transformed feature:

$$\mathbf{f}_t = \log \mathbf{f}_1 - \log \mathbf{f}_2, \quad (9)$$

where  $\mathbf{f}_1$  and  $\mathbf{f}_2$  are features extracted from the ‘‘previous’’ and ‘‘current’’ images, respectively.  $\log(\cdot)$  denotes applying logarithmic operation to each element in the vector.

After extracting the features from a pair of images or a single image, we apply a normalization to the features by:

$$\mathbf{m}' = \frac{\mathbf{m} - \mu}{\sigma}, \quad (10)$$

where  $\mathbf{m}$  is the extracted feature (denoted as  $\mathbf{f}$  or  $\mathbf{f}_t$  above).  $\mu$  is the mean value and  $\sigma$  is the standard deviation, both are calculated from the training data along each feature dimension (there are 5 feature dimensions). The normalization is essential in order to obtain a robust estimation.

## V. LEARNING THE MAPPING

The weight/BMI analysis is to map the anthropometric features to BMI values. The training process is to learn the mapping function. In the estimation, the learned function is used to estimate the BMI values from extracted features. We study the problem in different settings. Since the problem is relatively new and challenging, we explore how well we can achieve at different levels of difficulties:

- Recognize the weight difference (increase, decrease or the same)  $\hat{t}_c$ , given a pair of images.
- Predict how big the weight or BMI difference  $\hat{t}_d$  is between a given pair of images.
- Estimate the BMI value  $\hat{t}_v$  from a single body image.

Weight difference recognition is a three-class classification problem. The pairwise feature  $\mathbf{f}_t$  used for training and testing in this problem is obtained from Eqn. (9). The ground-truth label  $t_c$  is generated based on the weight change on the pairwise images (suppose the height of each subject remain the same in

a pair of images).  $t_c \in [0, 1, -1]$ :  $0$  denotes keeping the same weight,  $1$  denotes weight increased, and  $-1$  denotes weight decreased.

The level of BMI differences is considered as a regression problem. The pairwise-features  $\mathbf{f}_t$  are also used for training and testing in this problem. The ground-truth label  $t_d$  is the BMI difference of the pair images which may be positive or negative.

BMI value estimation is also defined as a regression problem. The feature vector extracted from each single image  $\mathbf{f} = [WTR, WHpR, WHdR, HpHdR, Area]^T$  is used for this problem. The ground-truth label  $t_v$  is the BMI value.

We employ the multi-class support vector machines (multi-SVMs) [40] for classification, and the support vector regression (SVR) [41] and Gaussian process regression (GPR) [42] for BMI difference prediction and BMI estimation.

### A. Support vector machine

Support vector machines (SVMs) are supervised learning algorithms that analyze data for classification or regression. There are two main categories for SVMs: support vector classification (SVC) and support vector regression (SVR). They have been widely utilized in many problems [43], [44]. The SVM can do nonlinear classification using kernel functions. Gaussian radial basis Function (RBF) kernel is one of the most popular kernels. In this work, the RBF kernel achieves a better performance in classification and regression than other kernels.

The SVC is a binary classifier. To get multi-class classification, a set of binary classifiers are constructed with each trained to separate one class from another. For  $n$  classes, this results in  $\frac{(n-1)n}{2}$  binary classifiers. Since our classification on BMI difference has three classes  $\{0, 1, -1\}$  for a pair of images, 3 binary classifiers are trained accordingly. The SVR uses the same principle, similar to the SVC, but with differences in the optimization.

### B. Gaussian processing regression

A Gaussian process (GP) is a collection of random variables and a finite number of variables which have a joint Gaussian distribution [45]. GPR means Gaussian process regression. The prior mean and covariance of the GP need to be specified. The prior mean is assigned constantly with zero, or the mean of the training data. The prior covariance is specified by passing a kernel object. The hyper-parameters of the kernel are optimized by maximizing the log-marginal-likelihood. A rational quadratic kernel is employed for GPR. Given a set of training examples  $(a_1, b_1), \dots, (a_n, b_n)$ , the rational quadratic kernel is defined as:

$$k(a_i, a_j) = \left( 1 + \frac{D(a_i, a_j)^2}{2\alpha\iota^2} \right)^{-\alpha}, \quad (11)$$

here  $\iota$  is a length-scale parameter,  $\alpha$  is a scale mixture parameter, and  $D(\cdot)$  denotes the distance between two sample points.

## VI. PERFORMANCE MEASURES

It is critical to measure the correlation between the extracted anthropometric features and body weights or BMIs. Pearson's correlation coefficient ( $PCC$ ) is employed for measuring the correlation. It is a measure of the linear correlation between two variables. It was developed by Karl Pearson in 1895 from a related idea introduced by Francis Galton [46]. Given two sets of data  $\{a_1, \dots, a_n\}$  and  $\{b_1, \dots, b_n\}$ , the formula for  $PCC$  is:

$$PCC = \frac{\sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2} \sqrt{\sum_{i=1}^n (b_i - \bar{b})^2}}, \quad (12)$$

here  $PCC$  is a scalar value between  $-1$  and  $1$ . If  $PCC < 0$ , it shows a negative correlation between the two sets. If  $PCC > 0$ , it shows a positive correlation. When  $PCC = 0$ , it indicates that there is no correlation between the two sets. When  $PCC$  is close to  $-1$  or  $1$ , there is a very strong correlation.

We apply a hypothesis testing with a statistical significance measure [47]. The p-value is utilized to decide whether a significant correlation exists between the two sets of data. We can make a decision by:

- If the p-value is smaller than the significance level  $\alpha$ , it can reject the null hypothesis (there is no correlation between the two sets).
- If the p-value is larger than the significance level  $\alpha$ , it fails to reject the null hypothesis.

The significance level  $\alpha$  can be set to, e.g., 0.001, 0.01 or 0.05. If the p-value is equal to or smaller than the threshold, it indicates a significant correlation between the two data sets.

In addition to correlation, we measure the performance of the proposed approach for weight or BMI estimation. The recall is used to evaluate the classification. And mean absolute error (MAE), mean absolute percentage error (MAPE) and absolute percentage error (APE) are used to measure the regression results:

- Recall: it is a performance measure that quantifies the ability of the classifier to correctly classify the positive training instances (also true positive rate, sensitivity). It is computed as the number of corrected classification divided by the number of samples that should have been classified as this class.
- MAE: it is defined as the average of absolute error between the estimated value and the ground truth:

$$MAE = \frac{1}{N} \sum_{j=1}^N |\hat{r}_j - r_j|, \quad (13)$$

here  $\hat{r}_j$  is the estimated value for  $j$ -th sample,  $r_j$  is the ground truth for  $j$ -th sample, and  $N$  is the total number of test samples.

- MAPE: it is the mean absolute percentage error, computed as:

$$MAPE = \frac{100}{N} \sum_{j=1}^N \left| \frac{\hat{r}_j - r_j}{r_j} \right|, \quad (14)$$

where all variables in Eqn. (14) have the same meaning as in Eqn. (13). Considering the large range of BMIs (15 to 75) in the visual-body-to-BMI dataset, the absolute

TABLE II  
PEARSON'S CORRELATION BETWEEN THE EXTRACTED FEATURES AND THE BMI VALUES IN DIFFERENT GENDER GROUPS.

	Male		Female		Overall	
	n = 1334		n = 666		n = 2000	
	p-value	correlation	p-value	correlation	p-value	correlation
WTR	0.0000	0.1774	0.0078	0.1033	0.0000	0.1320
WHpR	0.0000	0.1771	0.0018	0.1301	0.0000	0.1371
WHdR	0.0000	0.3317	0.0000	0.2992	0.0000	0.3038
HpHdR	0.0000	0.2791	0.0000	0.2769	0.0000	0.2785
Area	0.0000	0.4082	0.0000	0.3219	0.0000	0.3873

percentage error can be another useful measure for the performance of BMI prediction from single images. For example, two individuals with the same height, one's BMI is 20 and another is 40. If they both have their BMI increase by 2, such a change is more obvious on the individual with BMI = 20. MAPE measures the error by taking the BMI as the base. APE is calculated by a single estimated value and ground-truth. It is a relative error.

## VII. EXPERIMENTS

In this section, we explore the feasibility of analyzing body weight from 2D body images. We first examine the correlation between the extracted anthropometric features and the BMI values and then perform three estimation experiments using the extracted features.

The visual-body-to-BMI dataset is randomly split into training and test sets. The training set contains 2000 subjects (4000 images) of 1334 males and 666 females. The test set contains 950 subjects (1900 images): 650 males and 300 females. There is no overlap of subjects between the training and test sets.

### A. Correlations between body features and BMI values

According to the hypothesis test, we can measure whether the extracted features and BMI values are correlated. Here we assume the correlation with p-value < 0.01 is a significant correlation, and vice versa.

We measure the correlation between the extracted anthropometric features and BMI values on the training set. The results are shown in Table II. p-value = 0.0000 indicates that the value is smaller than 0.0001. From Table II we can see that the feature *Area* shows a higher correlation with BMI than other features. The correlation is a little lower in the female group than the male group, which may be caused by the different clothing styles or body fat distribution between females and males. The correlation coefficients of WTR and WHpR are lower than the other three features. Velardo et. al [16] reported an average correlation coefficient of 0.27 for BMI and waist to thigh ratio, and Vazquez et. al [14] reported a correlation coefficient of 0.34 for BMI and waist to hip ratio. Considering that various clothes styles exist in the dataset which may bring negative influences to feature calculation, the correlation coefficients of these two features given in Table II are slightly low but still acceptable. According to the above analysis, the conclusion can be drawn that the extracted features

TABLE III  
RECALL OF THREE-CLASS CLASSIFICATION FROM THE PAIR-WISE IMAGES.

Class	Recall (%)		
	Male	Female	Overall
0	63.6	40.0	56.3
1	81.0	89.2	83.6
-1	77.3	88.0	81.1

**Overall Accuracy: 81.3%**

Predicted Class	Target Class		
	-1	0	1
-1	81.1% 231	18.8% 3	14.5% 30
0	7.4% 21	56.3% 9	1.9% 4
1	11.6% 33	25.0% 4	83.6% 173

Fig. 8. Confusion matrix of weight difference classification results. The diagonal cells show the number and percentage of correct classifications by the method.

are correlated with the corresponding BMI values. Thereby, it is reasonable to estimate BMI values using the extracted features.

### B. Recognize weight difference from a pair of images

The proposed approach takes either a pair of body images or a single body image as the input. For the pairwise images, the approach performs a three-class classification which decides the subject in the pairwise images as weight increased, decreased or keeping the same. Furthermore, we estimate how much the BMI difference between the pairwise images is.

#### 1) Weight change or not?

The approach can process a three-class classification  $\{0, 1, -1\}$  for a pair of images from the same subjects. We use the features calculated by Eqn. (9) in Section IV-B to train a multi-SVMs which contains 3 binary classifiers. The RBF is utilized for the SVM kernel which achieves a better performance on classification than other kernels.

The recall of weight difference classification is given in Table III. Taking into account the different body fat rate between



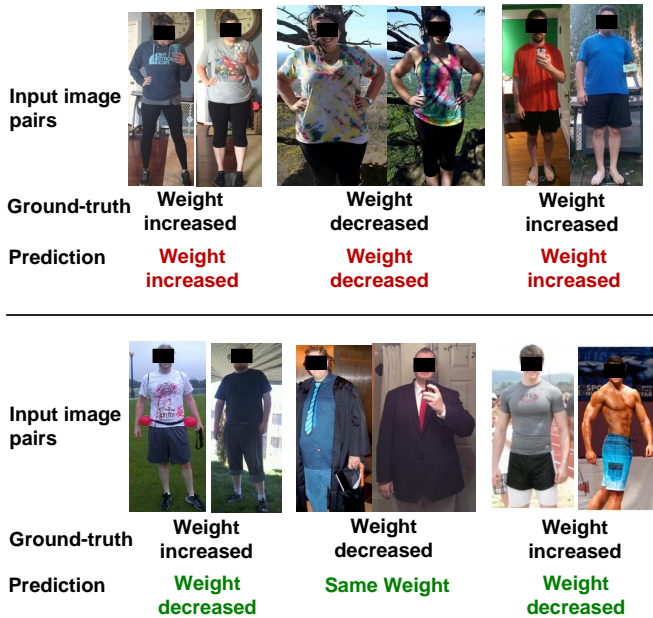


Fig. 9. Some results of the weight difference classification. The upper panel shows good cases, and the lower panel shows failure cases. The BMI difference is from the left one to the right one.

males and females, the recall is measured for each gender group. It is seen that the accuracy for class 0 (keep the same weight) is much lower than the other two classes: 1 (weight increase) and -1 (weight decrease). The reason may be that the number of subjects in class 0 (481) is much less than the other two classes (1246 + 1223). There is an uneven distribution among the three classes. Fig. 8 shows the confusion matrix of weight change classification. The accuracy of weight increased pairs is 83.6%, and the accuracy of weight decreased pairs is 81.1%. They are both within the acceptable range. Fig. 9 shows some examples of the classification. The upper panel shows some good cases, while the lower panel shows some failure cases. Failure cases are observed due to the interference, occlusion of the body, or large body pose, etc.

## 2) How big is the weight change?

Further exploration is to discover how big the weight or BMI change between pairwise images is. The features computed by Eqn. (9) are used to train the regression model. Here we employ the SVR (with the RBF kernel) and GPR models. Table IV shows the MAEs and standard deviations of the estimated BMI differences by the two regression models. We can see that the GPR model performs slightly better than the SVR model. Fig. 10 depicts the comparison of MAEs between the SVR and the GPR broken down by the absolute BMI differences. The difference between SVR and GPR for all ranges are less than 1 except for the range of 0.5 – 5.5 (approximately 1.16). The MAEs in the absolute BMI difference range > 15.5 are relatively higher than other ranges. This may be caused by the small number of subjects (about 7.90%) with BMI differences larger than 15.5. The distribution of BMI differences in the visual-body-to-BMI dataset is given as 492 subjects are in the range of BMI difference < 0.5, 921 are between 0.5 and 5.5, 866 are between 5.5 and 10.5, 438 are between 10.5 and

TABLE IV  
THE MAES AND STANDARD DEVIATIONS OF THE ESTIMATED BMI DIFFERENCES USING SVR AND GPR MODELS.

Model	MAE			Std		
	Male	Female	Overall	Male	Female	Overall
SVR	3.6	4.1	3.8	3.6	3.5	3.6
GPR	3.6	4.0	3.7	3.4	3.5	3.5

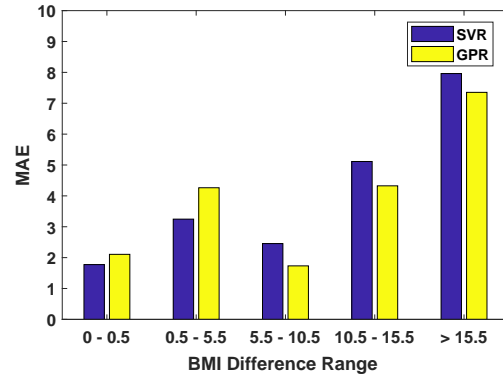


Fig. 10. Comparison of MAEs between SVR and GPR broken down by the absolute BMI differences.

15.5 and 233 are in the range of BMI difference > 15.5. The proposed approach shows effectiveness in predicting how big the weight or BMI change is from a pair of body images.

## C. Estimate BMI from a single image

Now we study the BMI estimation from single images by using the SVR and GPR models. Different from the previous two experiments, we use the anthropometric features  $\mathbf{f} = [WTR, WHpR, WHdR, HpHdR, Area]^T$  extracted from the single image for BMI estimation.

The MAEs and MAPEs of the estimated BMI values by two regression models are given in Tables V and VI, respectively. The overall MAEs of the predicted BMI values are between 3 and 4, the range of BMIs in the dataset is from 15 to 75, as shown in Fig. 3. The error of estimated BMIs is relatively small compared to the large range of BMIs in the dataset. Fig. 11 shows the MAEs and MAPEs between SVR and GPR in different BMI categories: underweight, normal, overweight and obese. We can see that the two regression methods perform well in the normal category. Though the MAEs in obese category is between 5 and 6.5, taking into account the large range of BMI distribution in the obese category (30 to 75), the MAPEs of this category are acceptable. To compare the ground truth BMIs with the estimations, a scatter plot based on the SVR results is shown in Fig. 12. The red dash-dot line shows where the two values are the same. The two green lines show where the absolute differences between the two values are 5. It is shown that points mainly distribute around the red line. Most outliers have the ground truth BMI values larger than 55. It can be seen that the proposed method tends to have a bias with an overestimation for low BMIs (BMI values between 20 and 30) and have an underestimation of high BMIs (BMI values larger than 35).

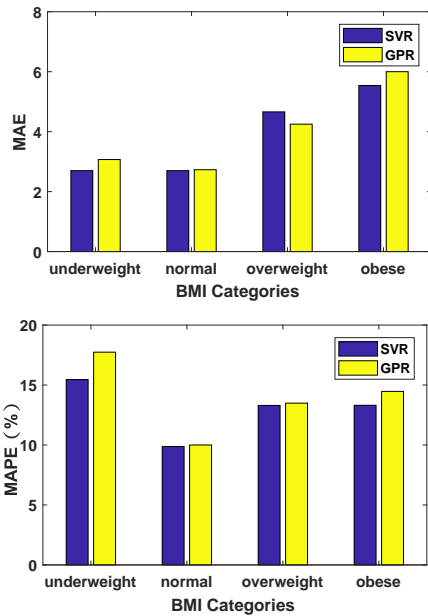


Fig. 11. Comparison of MAEs and MAPEs between SVR and GPR in different BMI categories: underweight (BMI  $\leq 18.5$ ), normal ( $18.5 < \text{BMI} \leq 25$ ), overweight ( $25 < \text{BMI} \leq 30$ ) and obese (BMI  $> 30$ ).

TABLE V  
THE MAES AND STANDARD DEVIATIONS OF PREDICTED BMI IN DIFFERENT GENDER GROUPS USING SVR AND GPR MODELS.

Model	MAE			Std		
	Male	Female	Overall	Male	Female	Overall
SVR	3.4	4.5	3.8	3.3	4.8	3.6
GPR	3.5	4.4	3.9	3.5	4.0	3.7

TABLE VI  
MAPEs OF PREDICTED BMI IN DIFFERENT GENDER GROUPS USING SVR AND GPR MODELS.

Model	Male	Female	Overall
SVR	11.3%	15.0%	12.5%
GPR	12.1%	15.2%	13.1%

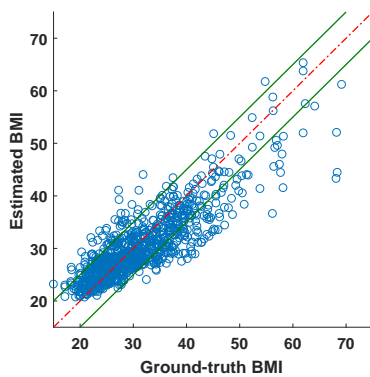


Fig. 12. Scatter plot of the ground-truth BMIs over the estimated BMIs by SVR.

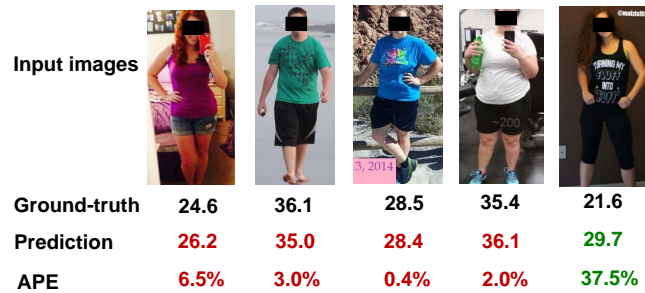


Fig. 13. Examples of estimating BMI from a single body image.

TABLE VII  
COMPARISON OF BMI ESTIMATION BETWEEN OUR METHOD AND OTHER METHODS.

Method	MAE			MAPE (%)		
	Male	Female	Overall	Male	Female	Overall
PIGF	4.61	4.58	4.60	15.5	15.6	15.5
VGG feature	3.72	4.48	3.94	12.7	16.0	13.7
Ours	3.74	4.16	3.86	12.7	14.2	13.1

percentage error ( $APE = \left| \frac{\hat{r}_j - r_j}{r_j} \right|$ ) is calculated for each case. Some failure cases are caused by ambiguous boundaries between the foreground and background, image blur, or large body pose. A detailed discussion about estimation errors and failure cases will be given in Section VII-E.

#### D. Comparison with other methods

To the best of our knowledge, there is no previous approach that can estimate the BMI values from 2D body images only. Thereby, we compare with two methods which predict BMI values from face images. One is a geometric feature based method [25] and another is a VGG-face feature based method [28]. They are denoted as PIGF (psychology inspired geometric feature) and VGG feature, respectively. These two methods both require clear frontal face images as the input, while some images in visual-body-to-BMI dataset do not meet this requirement. For a fair comparison, we select 2000 images which contain the clear frontal view face and then crop the face images. The 2000 images are split into training and test sets, which contains 1500 and 500 images, respectively. The input of our approach is a single body image, and the input of the other two methods is a face image cropped from the same body image. The comparison of the results is shown in Table VII. It can be seen that the proposed method outperforms the PIGF and VGG-face feature based methods in most cases, except on the male set. Moreover, the proposed method does not require a clear frontal view face image as input, which is useful for more general applications.

Furthermore, considering the features learned in deep neural networks (DNN) are demonstrated to be transferable and effective when used in other visual recognition tasks [48], we compare our anthropometric features with that the deep features. In this experiment, we employ the VGG-Net [49] model which is pre-trained on ImageNet database [50] to extract the deep feature. Then an SVR model is trained based on the extracted

Figure 13 shows some examples of prediction. The absolute

TABLE VIII  
RESULTS OF BMI ESTIMATION FROM OUR ANTHROPOMETRIC FEATURES  
AND THE VGG-NET FEATURE.

Feature	MAE			MAPE (%)		
	Male	Female	Overall	Male	Female	Overall
VGG-Net	4.65	5.55	4.94	15.6	17.8	16.3
Ours	3.41	4.52	3.76	11.3	15.0	12.5

TABLE IX  
THE MEAN RELATIVE ERRORS OF THE EXTRACTED FEATURES.

	Head	Waist	Hip	Thigh
Error	2.1%	5.4%	4.7%	9.7%

deep feature. The feature from the  $fc6$  layer is extracted for each body image in the training and test sets. VGG-Net takes an image of size  $224 \times 224$  with the average image subtracted as the input. To normalize the images in visual-body-to-BMI dataset to a common size, we apply zero padding to the images, and then resize them to  $224 \times 224$ . The training and test sets in this experiment are the same as the experiment in Section VII-C. Table VIII presents the results obtained based on the two features. It can be seen that our anthropometric features outperform the VGG-Net feature significantly.

### E. Discussion

In this section, we first analyze the errors generated in feature extraction. Then the statistical analysis will be given, discussing whether the errors are acceptable for the application of BMI estimation from a single image. Finally, we analyze the influencing factors for the proposed method and possible reasons for the failure cases.

For feature extraction and regression, the widths of head, waist, hip and thigh are estimated from the 2D body images, and used to calculate the four anthropometric features ( $WTR$ ,  $WHpR$ ,  $WHdR$ ,  $HphdR$ ). To analyze the error, we randomly selected 300 images from the dataset and manually labeled the widths of head, waist, hip, and thigh for each image. Then the labeled widths are used as the ground truth values ( $v$ ) to calculate the relative error ( $\varepsilon$ ) of the estimated values ( $\hat{v}$ ) by:  $\varepsilon = \frac{|v - \hat{v}|}{v}$ . The mean relative errors of the extracted widths are shown in Table IX. The four errors are within a relatively low range. Since it is hard to label the area between waist and hip, where the relative error of estimated *Area* is not given.

To demonstrate whether the errors are acceptable for BMI estimation from a single body image, we further calculate the accuracy of the predicted category. According to the estimated BMI values, we can classify the body belong to which BMI category (underweight, normal, overweight and obese). The accuracy of the predicted category is the proportion of the total number of predictions that are correct. This measurement is helpful to decide if the errors are acceptable. For example, given a body image with ground truth BMI value of 24.5, the estimated value is 20. Though the absolute error is 4.5 which is larger than the MAE (3.8), the predicted category (normal) is correct. On the other hand, this measurement has a limitation.

TABLE X  
THE ACCURACY OF PREDICTED CATEGORY.

	Underweight	Normal	Overweight	Obese
Accuracy	11.1%	78.3%	64.2%	81.0%

<b>Large pose</b>			
<b>Ground-truth</b>	57.0	51.0	35.4
<b>Prediction</b>	44.4	41.9	44.1
<b>Occlusion and loose clothes</b>			
<b>Ground-truth</b>	18.8	35.3	34.5
<b>Prediction</b>	28.5	30.2	25.8
<b>Incorrect segmentation</b>			
<b>Ground-truth</b>	25.4	20.4	27.5
<b>Prediction</b>	33.3	32.7	16.1

Fig. 14. Examples of failure cases with the corresponding detected body contour and skeleton joints. The upper panel shows the cases with the large pose. The middle panel shows the cases with body occlusion or loose clothes. The lower panel shows the incorrect body contour cases.

For example, if the ground truth BMI is 25 and the estimated value is 25.5, though the absolute error is 0.5, the predicted category (overweight) is not correct. Considering the advantage and limitation of this measure, we combine the accuracies of predicted category (as shown in Table X) with the MAEs of predicted BMIs (shown in Fig. 11) to evaluate the performance. All predicted results shown in Table X are based on the SVR method. As it can be observed from Table X and Fig. 11, the prediction accuracy and MAE of the obese category are 81% and 5.5, respectively. Taking into account the large range of BMI on the obese category (30 to 75), the error of the obese category is reasonable. The prediction accuracy and MAE of the overweight category are 64.2% and 4.6, respectively. The performance of the overweight category is a little lower than the obese. The prediction accuracy of the underweight category is the lowest since there are only 46 body images in the database belong to the underweight category, among them, 9 are in the test set and 37 are in training set. The lack of enough underweight body images in the training set could be the reason for this lower performance.

To analyze influencing factors (such as pose, occlusion, loose clothes, and scale) for the proposed method, and the reasons (such as incorrect body contour) of failure, Fig. 14 shows some failure cases with the detected body contour and skeleton joints. Most images in the dataset are frontal

view body images with limited pose changes. Since there is no annotation about body pose, it is difficult to conduct an experiment to evaluate the performance with regard to pose changes. Theoretically, the extracted anthropometric features can tolerate small pose changes. The estimation may be significantly influenced if the input is a profile view image or with the large pose. The upper panel of Fig. 14 shows three failure cases with different poses. Though the detected body contour and skeleton joints are correct, the absolute errors are large. The occlusion always brings negative influences to the method, decreasing the accuracy of body contour detection. Loose clothes is another negative factor to influence the real body shape. Three cases with large body occlusions and loose clothes are shown in the middle panel of Fig. 14. Because all the extracted anthropometric features are relative values (see details in Section IV-B), the scale changes in the image will not impact the method. The lower panel of Fig. 14 shows three failure cases caused by inaccurate contour. The incorrect or inaccurate body contour directly influences the accuracy of the extracted features. The failure of contour detection could be caused by image blurs, ambiguous boundaries between the foreground and background, etc. The proposed method can be further improved by employing more accurate body contour detection algorithms.

### VIII. CONCLUSION

In this work, we investigate the relation between body weight and visual body appearance and estimate the BMI values from 2D body images. Correlation is analyzed between the extracted anthropometric features and BMI values, which validates the usability of the selected features. More specifically, body weight analysis is studied at three different levels of difficulties: the weight change classification is first investigated from a pair of body images of the same subjects; further investigation is conducted to estimate how big the weight change between the pairwise images is; the last is to predict the BMI value from a single body image. To address the visual body weight analysis problem, the computational method of five anthropometric features is developed. And a new visual-body-to-BMI image dataset has been collected and cleaned to facilitate this study. The errors of the three estimation tasks evaluated by several measurements are within acceptable ranges. Comparing with the facial images analysis approaches, the proposed method performs better in most cases. Furthermore, our anthropometric features significantly outperform the VGG-Net feature on BMI estimation. Based on all experimental results, it is promising to analyze body weight or BMI from the 2D body images visually. In the future, we will combine body images with face images to improve the BMI prediction, and will explore the DNN-based method to address this visual body weight analysis problem.

### ACKNOWLEDGMENT

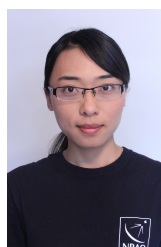
The work is partly supported by an NSF grant IIS 1450620, the Center for Identification and Technology and Research (CITeR), and a WV HEPC grant. The authors would like to thank the editor and anonymous reviewers for comments and suggestions to improve the manuscript.

### REFERENCES

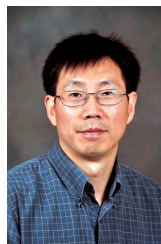
- [1] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: a tool for information security," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 125–143, 2006.
- [2] Y. Zhu, Y. Li, G. Mu, S. Shan, and G. Guo, "Still-to-video face matching using multiple geodesic flows," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 12, pp. 2866–2875, 2016.
- [3] A. Nagar, K. Nandakumar, and A. K. Jain, "Multibiometric cryptosystems based on feature-level fusion," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 255–268, 2012.
- [4] A. Dantcheva, C. Velardo, A. Dangelo, and J.-L. Dugelay, "Bag of soft biometrics for person identification," *Multimedia Tools and Applications*, vol. 51, no. 2, pp. 739–777, 2011.
- [5] M. Günther, P. Hu, C. Herrmann, C. H. Chan, M. Jiang, S. Yang, A. R. Dhamija, D. Ramanan, J. Beyerer, J. Kittler *et al.*, "Unconstrained face detection and open-set face recognition challenge," in *IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2017, pp. 697–706.
- [6] Q. Wang, G. Guo, and M. I. Nouyed, "Learning channel interdependencies at multiple scales on dense networks for face recognition," *arXiv preprint arXiv:1711.10103*, 2017.
- [7] A. J. Henderson, I. J. Holzleitner, S. N. Talamas, and D. I. Perrett, "Perception of health from facial cues," *Phil. Trans. R. Soc. B*, vol. 371, no. 1693, p. 20150380, 2016.
- [8] C. Mayer, S. Windhager, K. Schaefer, and P. Mitteroecker, "Bmi and whr are reflected in female facial shape and texture: a geometric morphometric image analysis," *PLoS one*, vol. 12, no. 1, p. e0169336, 2017.
- [9] M. Arnold, M. Leitzmann, H. Freisling, F. Bray, I. Romieu, A. Renehan, and I. Soerjomataram, "Obesity and cancer: an update of the global impact," *Cancer Epidemiology*, vol. 41, pp. 8–15, 2016.
- [10] A. G. Renehan, M. Tyson, M. Egger, R. F. Heller, and M. Zwahlen, "Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies," *The Lancet*, vol. 371, no. 9612, pp. 569–578, 2008.
- [11] R. Wolk, P. Berger, R. J. Lennon, E. S. Brilakis, and V. K. Somers, "Body mass index," *Circulation*, vol. 108, no. 18, pp. 2206–2211, 2003.
- [12] J. B. Meigs, P. W. Wilson, C. S. Fox, R. S. Vasan, D. M. Nathan, L. M. Sullivan, and R. B. Dagostino, "Body mass index, metabolic syndrome, and risk of type 2 diabetes or cardiovascular disease," *The Journal of Clinical Endocrinology & Metabolism*, vol. 91, no. 8, pp. 2906–2912, 2006.
- [13] A. Molarius and J. Seidell, "Selection of anthropometric indicators for classification of abdominal fatness: a critical review," *International Journal of Obesity and Related Metabolic Disorders*, vol. 22, no. 8, p. 719, 1998.
- [14] G. Vazquez, S. Duval, D. R. Jacobs Jr, and K. Silventoinen, "Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis," *Epidemiologic Reviews*, vol. 29, no. 1, pp. 115–128, 2007.
- [15] M. Ashwell, S. Chinn, S. Stalley, and J. Garrow, "Female fat distribution—a simple classification based on two circumference measurements," *International Journal of Obesity*, vol. 6, no. 2, pp. 143–152, 1982.
- [16] J. C. Seidell, A. Oosterlee, M. Thijssen, J. Burema, P. Deurenberg, J. Hautvast, and J. Ruijs, "Assessment of intra-abdominal and subcutaneous abdominal fat: relation between anthropometry and computed tomography," *The American journal of clinical nutrition*, vol. 45, no. 1, pp. 7–13, 1987.
- [17] C. Velardo and J.-L. Dugelay, "Weight estimation from visual body appearance," in *Proceedings of the IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, 2010, pp. 1–6.
- [18] "National health and nutrition examination survey," *Centers for Disease Control and Prevention*, 1999–2005.
- [19] D. Cao, C. Chen, D. Adjeroh, and A. Ross, "Predicting gender and weight from human metrology using a copula model," in *IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. IEEE, 2012, pp. 162–169.
- [20] C. Velardo and J.-L. Dugelay, "What can computer vision tell you about your weight?" in *Proceedings of the IEEE European Signal Processing Conference (EUSIPCO)*, 2012, pp. 1980–1984.
- [21] C. Velardo, J.-L. Dugelay, M. Palcari, and P. Ariano, "Building the space scale or how to weigh a person with no gravity," in *IEEE International Conference on Emerging Signal Processing Applications (ESPA)*. IEEE, 2012, pp. 67–70.
- [22] T. V. Nguyen, J. Feng, and S. Yan, "Seeing human weight from a single rgb-d image," *Journal of Computer Science and Technology*, vol. 29, no. 5, pp. 777–784, 2014.



- [23] D. Nahavandi, A. Abobakr, H. Haggag, M. Hossny, S. Nahavandi, and D. Filippidis, "A skeleton-free kinect system for body mass index assessment using deep neural networks," in *IEEE International Systems Engineering Symposium (ISSE)*, 2017, pp. 1–6.
- [24] C. Pfitzner, S. May, and A. Nüchter, "Body weight estimation for dose-finding and health monitoring of lying, standing and walking patients based on rgb-d data," *Sensors (Basel, Switzerland)*, vol. 18, no. 5, 2018.
- [25] L. Wen and G. Guo, "A computational approach to body mass index prediction from face images," *Image and Vision Computing*, vol. 31, no. 5, pp. 392–400, 2013.
- [26] B. J. Lee and J. Y. Kim, "Predicting visceral obesity based on facial characteristics," *BMC Complementary and Alternative Medicine*, vol. 14, no. 1, p. 248, 2014.
- [27] M. A. Pascali, D. Giorgi, L. Bastiani, E. Buzzigoli, P. Henríquez, B. J. Matuszewski, M.-A. Morales, and S. Colantonio, "Face morphology: Can it tell us something about body weight and fat?" *Computers in Biology and Medicine*, vol. 76, pp. 238–249, 2016.
- [28] E. Kocabey, M. Camurcu, F. Offli, Y. Aytar, J. Marin, A. Torralba, and I. Weber, "Face-to-bmi: Using computer vision to infer body mass index on social media," *arXiv preprint arXiv:1703.03156*, 2017.
- [29] A. Dantcheva, F. Brémond, and P. Bilinski, "Show me your face and i will tell you your height, weight and body mass index," in *International Conference on Pattern Recognition (ICPR)*, 2018.
- [30] Q. Wu and G. Guo, "Gender recognition from unconstrained and articulated human body," *The Scientific World Journal*, vol. 2014, 2014.
- [31] D. Cao, C. Chen, M. Piccirilli, D. Adjeroh, T. Bourlai, and A. Ross, "Can facial metrology predict gender?" in *International Joint Conference on Biometrics (IJCB)*. IEEE, 2011, pp. 1–8.
- [32] E. Gonzalez-Sosa, A. Dantcheva, R. Vera-Rodriguez, J.-L. Dugelay, F. Brémond, and J. Fierrez, "Image-based gender estimation from body and face across distances," in *International Conference on Pattern Recognition (ICPR)*, 2016, pp. 3061–3066.
- [33] A. O. Balan, L. Sigal, M. J. Black, J. E. Davis, and H. W. Haussecker, "Detailed human shape and pose from images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.
- [34] J.-M. Lu and M.-J. J. Wang, "Automated anthropometric data collection using 3d whole body scanners," *Expert Systems with Applications*, vol. 35, no. 1, pp. 407–414, 2008.
- [35] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," *arXiv preprint arXiv:1611.08050*, 2016.
- [36] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P. H. Torr, "Conditional random fields as recurrent neural networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1529–1537.
- [37] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4724–4732.
- [38] R. Drillis, R. Contini, and M. Bluestein, "Body segment parameters: a survey of measurement techniques," *Artificial Limbs*, vol. 8, pp. 44–66, 1963.
- [39] K. Bushby, T. Cole, J. Matthews, and J. Goodship, "Centiles for adult head circumference," *Archives of Disease in Childhood*, vol. 67, no. 10, pp. 1286–1287, 1992.
- [40] J. Weston and C. Watkins, "Multi-class support vector machines," Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, May, Tech. Rep., 1998.
- [41] H. Drucker, C. J. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems*, 1997, pp. 155–161.
- [42] C. K. Williams and C. E. Rasmussen, "Gaussian processes for regression," in *Advances in Neural Information Processing Systems*, 1996, pp. 514–520.
- [43] G. Guo, S. Z. Li, and K. L. Chan, "Support vector machines for face recognition," *Image and Vision Computing*, vol. 19, no. 9, pp. 631–638, 2001.
- [44] A. Statnikov, D. Hardin, and C. Aliferis, "Using svm weight-based methods to identify causally relevant and non-causally relevant variables," *Sign*, vol. 1, no. 4, 2006.
- [45] C. E. Rasmussen and C. K. Williams, *Gaussian processes for machine learning*. MIT press Cambridge, 2006, vol. 1.
- [46] K. Pearson, "Note on regression and inheritance in the case of two parents," *Proceedings of the Royal Society of London*, vol. 58, pp. 240–242, 1895.
- [47] W. C. Navidi, *Statistics for engineers and scientists*. McGraw-Hill New York, 2006, vol. 2.
- [48] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Advances in Neural Information Processing Systems*, 2014, pp. 3320–3328.
- [49] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [50] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 248–255.



**Min Jiang** received the B.S. degree and M.S. degree in electrical engineering from China University of Mining & Technology, Beijing, China. She is currently pursuing the Ph.D. degree with the Lane Department of Computer Science and Electrical Engineering, West Virginia University. Her research area includes computer vision, machine learning and signal processing, in particular, BMI analysis from human visual appearance and astronomical signal denoising.



**Guodong Guo** (M'07-SM'07) received the B.E. degree in automation from Tsinghua University, Beijing, China, the Ph.D. degree in pattern recognition and intelligent control from Chinese Academy of Sciences, Beijing, China, and the Ph.D. degree in computer science from University of Wisconsin-Madison, Madison, WI, USA. He is an Associate Professor with the Department of Computer Science and Electrical Engineering, West Virginia University (WVU), Morgantown, WV, USA. In the past, he visited and worked in several places, including

INRIA, Sophia Antipolis, France; Ritsumeikan University, Kyoto, Japan; Microsoft Research, Beijing, China; and North Carolina Central University. He authored a book, *Face, Expression, and Iris Recognition Using Learning-based Approaches* (2008), co-edited two books, *Support Vector Machines Applications* (2014), *Mobile Biometrics* (2017), and published more than 100 technical papers. His research interests include computer vision, machine learning, and multimedia. He received the North Carolina State Award for Excellence in Innovation in 2008, Outstanding Researcher (2017-2018, 2013-2014) at CEMR, WVU, and New Researcher of the Year (2010-2011) at CEMR, WVU. He was selected the "People's Hero of the Week" by BSJB under Minority Media and Telecommunications Council (MMTC) on July 29, 2013. Two of his papers were selected as "The Best of FG'13" and "The Best of FG'15", respectively.