Image and Vision Computing 89 (2019) 183–196

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

# On visual BMI analysis from facial images ☆

Min Jiang[a, b], Yuanyuan Shang[a], Guodong Guo[a, b,*]

[a] Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing 100048, China
[b] Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506, USA

ARTICLE INFO

*Article history:*
Received 29 January 2019
Received in revised form 1 July 2019
Accepted 16 July 2019
Available online 1 August 2019

*Keywords:*
Visual BMI estimation
Facial images
Facial representations
FIW-BMI database

ABSTRACT

Automatically assessing body mass index (BMI) from facial images is an interesting and challenging problem in computer vision. Facial feature extraction is an important step for visual BMI estimation. This work studies the visual BMI estimation problem based on the characteristics and performance of different facial representations, which has not been well studied yet. Various facial representations, including geometry based representations and deep learning based, are comprehensively evaluated and analyzed from three perspectives: the overall performance on visual BMI prediction, the redundancy in facial representations and the sensitivity to head pose changes. The experiments are conducted on two databases: a new dataset we collected, called the FIW-BMI and an existing large dataset Morph II. Our studies provide some deep insights into the facial representations for visual BMI analysis: 1) The deep model based methods perform better than geometry based methods. Among them, the VGG-Face and Arcface show more robustness than others in most cases; 2) Removing the redundancy in VGG-Face representation can increase the accuracy and efficiency in BMI estimation; 3) Large head poses lead to low performance for BMI estimation. The Arcface, VGG-Face and PIGF are more robust than the others to head pose variations.



## 1. Introduction

Face images contain much biometric information, such as identity, gender, age, weight, etc. Decoding facial cues has attracted much attention from sociologists and computer scientists. In this work, we study the problem of body mass index (BMI) estimation from facial images, called the visual BMI. The BMI (given by $\frac{weight(lb)}{height(in)^2} \times 703$) is a general body fat indicator widely used in medical research. It can reveal various health and lifestyle issues. Close connections exist between BMI and some diseases, such as cancers [1, 2], unstable angina [3] and type 2 diabetes or cardiovascular disease (CVD) [4], etc. Thus, BMI is important to personal health monitoring and medical research. BMI is traditionally measured in person with special devices. Recent studies show that facial cues are likely to influence facial perception and are important for BMI prediction [5-8]. Automatically assessing BMI from facial images is a great benefit to health condition monitoring and researchers who are interested in studying obesity in large populations.

Fig. 1 shows a typical framework for BMI estimation from two-dimensional (2D) facial images. It consists of four steps: face detection, image alignment, facial representation extraction, and regression. The first and second steps are the preparation for feature extraction. The third step is the most important which dominantly determines the performance for BMI estimation. Thereby, in this work, we study the visual BMI estimation problem from this key aspect: facial feature extraction methods, and explores methods to improve the performance.

We study the visual BMI estimation problem by analyzing two types of facial representation methods. The psychology inspired geometric features (PIGF) is used by Wen and Guo in [7]. Considering the whole facial shape may not be exactly defined by the PIGF, we explore another method for extracting geometric facial representation-pointer feature (PF), which defines the face shape by a series of facial landmarks. In addition, to take advantage of the above two geometric representations, a fusion method is utilized to extract a richer geometric representation, denoted as PIGF + PF. In terms of deep learning, the VGG-Face model has been utilized for BMI prediction in [8]. Considering the very high dimension of the VGG-Face feature, we also explore other deep models to extract the deep representations, e.g. the LightCNN [9], Centerloss models [10] and Arcface [11]. Thus we can get a deep insight into deep learning based facial representations for visual BMI analysis.

☆ This paper has been recommended for acceptance by Sinisa Todorovic.

* Corresponding author at: Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506, USA.
*E-mail addresses:* minjiang.aca@gmail.com (M. Jiang), syy@bao.ac.cn (Y. Shang), guodong.guo@mail.wvu.edu (G. Guo).

https://doi.org/10.1016/j.imavis.2019.07.003
0262-8856/

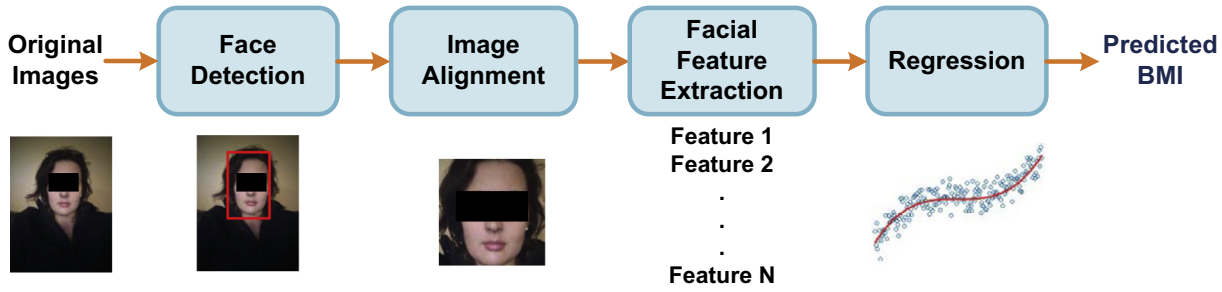


**Fig. 1.** A typical framework for visual BMI estimation from two-dimensional (2D) facial images.

The main contributions of this work include:

1. We study the visual BMI estimation problem systematically by evaluating two representative types of facial representations.
2. To explore the efficient methods to improve the performance, we study the redundancy and pose sensitivity of the facial representations, which are critical for practical applications.
3. A visual BMI analysis face database is collected, called face in the wild for BMI analysis, or FIW-BMI. It contains 7930 wild face images of 4881 individuals along with the labels of gender, height, and weight.

The remainder of the paper is organized as follows. Section 2 covers the existing facial feature extraction methods for visual BMI analysis. The principles and related methods are systematically presented and discussed in Section 3. Section 4 presents two databases used for performance evaluation: the newly collected FIW-BMI and Morph II. In Section 5, we conduct three experiments and provide detailed analysis and discussion. Finally, the conclusions are given in Section 6.

## 2. Related works

In the field of psychology, researchers study the association between facial measures and body weight. Coetzee et al. [12] demonstrated that facial adiposity (the perception of weight in the face) is associated with the perceived health and attractiveness in a curvilinear relationship. Coetzee et al. [13] studied the correlation between BMI and three facial measures—cheek-to-jaw-width ratio (CJWR), face width-to-height ratio (WHR) and face perimeter to area ratio (PAR). They recruited 95 Caucasian and 99 African participants to capture facial images for their study. Experimental results showed that these three facial measures are correlated with BMI values. Pham et al. [14] further studied the correlation between BMI values and four other facial measures—eye size, lower face to face height ratio (LF/FH), face width to lower face height ratio (FW/LFH) and the mean of eyebrow height among the group of young and elder in Korean. Recently, Henderson et al. [15] investigated the effect of multiple facial cues on health judgments from two-dimensional (2D) and three-dimensional (3D) facial images. They found that the skin color cues (including texture, color and color distribution) may provide perceived health information. Mayer et al. [6] assessed the association of BMI values with facial attributes—shape and texture (color pattern) in a female group. The experiment was conducted on 49 standardized images of female participants. They showed that BMI can be better predicted from facial attributes than from traditional facial measures.

Computationally the BMI values can be estimated from 2D facial images by geometric features. Wen and Guo [7] proposed for the first time to automatically predict the BMI values from facial images. The psychology inspired geometric features (PIGF) are computed for BMI representation. The method was evaluated on 14,500 images from Morph II [16]. Lee and Kim [17] examined 15 2D geometric facial characteristics from frontal and lateral face images of 11,347 Korean men and women to identify the strongest predictor of normal and viscerally obese subjects, and assessed the predictive power of the combined characteristics. Barr et al. [18] utilized the method in [7] to identify whether the BMI values can be correctly identified from the 1412 young adults in order to improve data capturing in dissemination and implementation.

Deep learning based approaches have shown promising results in face recognition [19, 20], and other visual tasks [21, 22]. Kocabey et al. [8] employed the pre-trained VGG-Face model [23] to extract facial representations for BMI analysis. The method was evaluated on 4206 images. A comparison between BMI prediction and human perception was presented. It is shown that humans perform better on small BMI differences predictions, and there is no performance difference for larger BMI difference predictions. Recently, Dantcheva et al. [24] explored the possibility of estimating height, weight, and BMI from single-shot facial images by using a regression method based on a 50-layer ResNet architecture. They assembled a dataset which consists of 1026 subjects to facilitate the study.

In the above studies, each BMI estimation method was evaluated on one dataset only. The influence of data distribution on the method can not be fully analyzed on the single dataset. Evaluating the methods on a second dataset with different characteristics is valuable. More important, so far there are no systematic studies to evaluate and compare various face representations for visual BMI analysis, especially the two representative types of facial representations: the geometric features and deep learning based features. The aim of this work is to present a study on visual BMI analysis on two databases with different characteristics to establish a better understanding of various representations.

## 3. A deep insight into the visual BMI representations

As mentioned above, we consider that there are two representative types of facial representations for visual BMI analysis from facial images. We examine the principles of the facial representations systematically and discuss some related issues.

### 3.1. Geometry based representations

The principle of a geometric model is to mathematically describe the facial shapes related to body fat. It shows that facial geometry measures are correlated to body fat or BMI. Inspired by these psychology studies [13, 14], the first computational method PIGF was developed by Wen and Guo [7], using geometric features to estimate seven facial metrics: cheek-to-jaw-width ratio (CJWR), face width-to-height ratio (WHR), face perimeter to area ratio (PAR), eye size (ES), lower face to face height ratio (LF/FH), face width to lower face height ratio (FW/LFH) and the mean of eyebrow height (MEH). Given the geometric features, some statistical methods can be used to map the features to BMI values. The facial landmarks need to be extracted

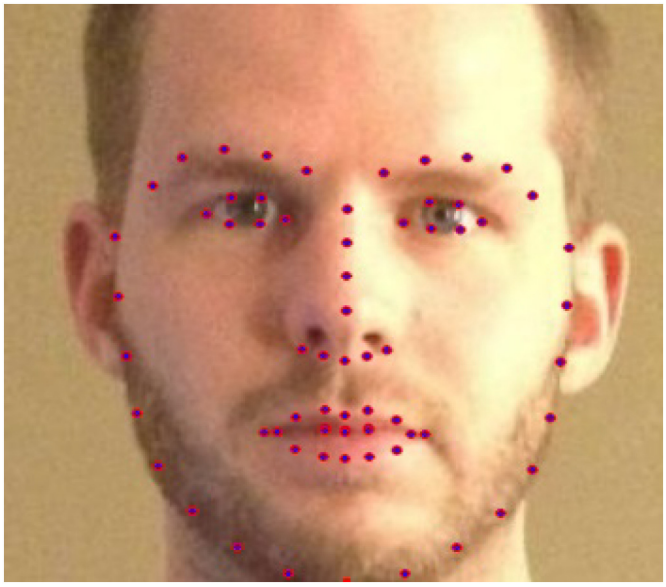

**Fig. 2.** Illustration of pointer feature (PF), which consists of a series of facial landmarks.

prior to the geometric feature extraction. So the performance of BMI estimation is related to the accuracy of facial landmark detection. The experimental results reported in [7] show that the PIGF performs quite well on BMI estimation.

On the other hand, Mayer et al. [6] analyzed the association of facial landmarks with body fat. They found that comparing with WHR, the whole facial shape is also good at reflecting the total fat proportion of the body. The facial shape based points were used to study the correlation between facial shape and body fat. Inspired by this, we want to explore another method for extracting the geometric facial representation, called the pointer feature (PF). It defines the facial shape and features by a series of facial landmarks as shown in Fig. 2. The PF consists of coordinates of M facial landmarks, denoted as $(x_i, y_i), i = 1, \ldots, M$, which can be directly concatenated into a vector. Then the PF representation is a 2M-dimensional vector: $[x_1, y_1, \ldots, x_n, y_n, \ldots, x_M, y_M]^T$. It is obvious that PF relies on the accurate detection of facial landmarks.

In addition, a fusion of these two kinds of geometric features could be considered. A simple way is to concatenate the PIGF and PF, denoted as PIGF + PF. Hopefully, it can take the advantages of the two geometric representations.

As a result, we will investigate three different geometric features, in order to get a deep understanding. The computation of geometric features is very fast, and the geometric representations are with very low dimensions.

### 3.2. Deep learning based representations

Recently, deep neural networks have been successfully applied to various applications. Fig. 3 shows a general pipeline of the deep learning approach. The VGG-Face is one of the deep convolutional networks originally proposed for face recognition, which learns a face embedding using a triplet loss function [23]. The network contains 13 convolutional layers, 5 max-pooling layers, 3 fully-connected (fc) layers and a final layer with the soft-max function. VGG-Face model is trained on 2.6 million face images from the web. It takes a face image of size $224 \times 224$ with a constant image with all pixels equal to (94,105,129) subtracted as the input. Having about 144 million parameters indicates that the VGG-Face is a complex model. Kocabey et al. [8] employed the pre-trained VGG-Face models to extract facial features for BMI analysis. The extracted features from layer fc6 of VGG-Face is utilized. The size of the feature vector in VGG-Face is 4096. Thus the dimension of the VGG-Face representation is quite high.

Considering the high computational complexity for VGG-Face model, it is interesting to investigate other deep models with a lower computational cost for visual BMI analysis. Here we explore the LightCNN, Centerloss and Arcface models.

LightCNN is a network with a low computational complexity which learns a compact face embedding on a large-scale dataset with noisy labels [9]. It proposed a Max-Feature-Max (MFM) activation function to suppress a small number of neurons and to make CNN models light and robust. The model is trained on 493,456 face images from CASIA WebFace dataset. The input of the network is a gray-scale face image of size $128 \times 128$. Considering the significant

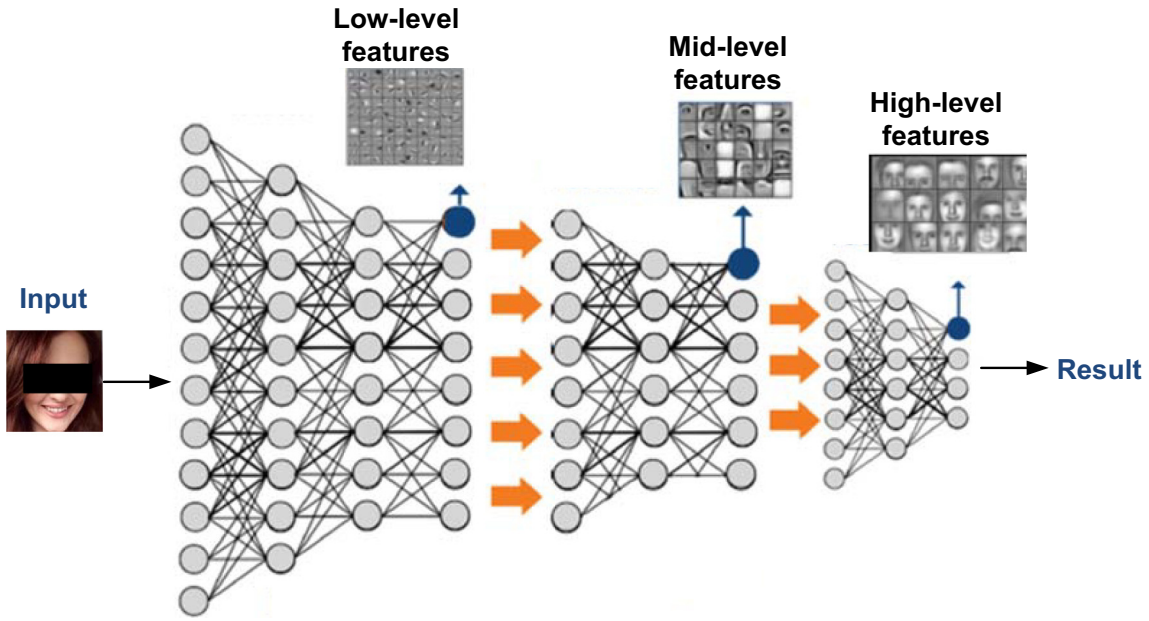


**Fig. 3.** The pipeline of deep learning approach.

performance achieved by this model on face recognition task, the layer fc1 (with size 256 × 1) of LightCNN is used to extract deep facial features for BMI estimation.

The features learned by the deep networks trained under the supervision of softmax loss [25] may not be discriminative enough. In order to improve the discriminative power of the learned features, Wen et al. [10] proposed a center loss function to minimize the intra-class variations while keeping features of different classes separable. The center loss based network takes a face image of size 96 × 112 as the input. This model was developed for face recognition, which we evaluate its use for visual BMI estimation. We extract features for each image and its horizontally flipped one, and concatenate them as a 1024 dimensional feature vector.

Additive Angular Margin Loss (ArcFace) [11] can extract highly discriminative features for face recognition by directly optimizing the geodesic distance margin through the correspondence between the angle and arc in the normalized hypersphere. It outperforms many other deep models for face recognition. The input of Arcface model is face images of size 112 × 112. The final 512-dimension embedding feature of the network is utilized for visual BMI analysis.

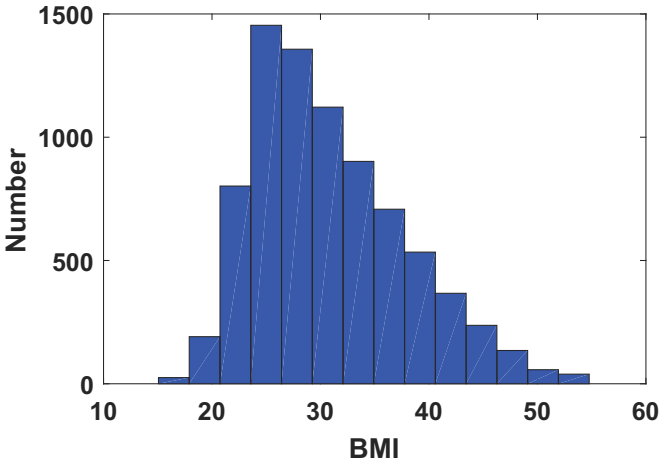


**Fig. 4.** Distribution of BMI values on FIW-BMI database. The BMI values span a wide range with most of the values distribute between 20 to 50.

## 4. Databases

Given various face representations for BMI analysis, we conduct experiments on two databases: a newly collected face database by us and the Morph II. They are different in size and characteristics. The details about the two databases are given below.

### 4.1. FIW-BMI database

We collect a new dataset, called face in the wild for BMI analysis (FIW-BMI)[1]. The facial images were collected from a social website—Reddit posts[2]. We went through the original images by a deep cascaded multi-task based face detector [26]. Given the detected face landmarks (two eyes, nose and mouth), each face image is cropped and normalized to the size of 256 × 256. Considering all the images are from social networks, they are not strictly frontal view face images with clean background. We visually checked all images and discarded the images which are not appropriate for visual BMI analysis, such as large head pose changes or exaggerated facial expressions. Finally, the annotation for each image is manually checked to generate the correct labels.

After all the above procedures, 7930 images from 4881 individuals were kept, along with the corresponding gender, height, and weight labels. Among these individuals, there are 3192 males (5197 images) and 1689 females (2733 images). Fig. 4 shows the BMI distribution of the whole database. Because the Reddit posts is a social network displaying people's progress of weight loss, weight gain, or essentially any type of body changes, the BMI values of these images distribute over a very wide range: 15 to 60. The mean BMI value of the database is 30.8, the standard deviation of the BMI values is 6.97. Among these images, 43 are underweight (BMI $\leq$ 18.5), 1662 are normal (18.5 $<$ BMI $\leq$ 25), 2455 are overweight (25 $<$ BMI $\leq$ 30), 3770 are in obese (BMI $>$ 30). Fig. 5 shows a few examples of the facial images from our FIW-BMI database.

### 4.2. Morph II database

Morph II database contains 55,608 mugshot-style frontal view face images along with the age, gender, and ethnicity labels. Most of them are with height and weight values. The BMI values can be computed from the weight and height. There is an uneven distribution of the ethnicity in the database, eg. about 96% identities are Black and White, while 4% are Hispanic, Asian, Indian, and others. Only images from African American and White are used for this study. Totally 29,033 images of 9693 identities were selected. Details about the selected Morph II dataset are given in Table 1. The images are separated into four groups by gender and ethnicity. The distribution of BMI values on the selected database (includes training and test sets) is shown in Fig. 6. Comparing to Fig. 5, the BMI values of Morph II mainly distribute on a relatively small range: 15 to 35. The mean BMI value of selected Morph II is 24.8, the standard deviation of the BMI values is 4.61. Among these, 893 are underweight, 16,582 are normal, 8237 are overweight and 3321 are obese.

## 5. Experiments and analysis

Experimentally, we evaluate and analyze the two major types of facial representations for BMI estimation. The experiment settings and performance metrics are briefly described first. Then the overall performance of the facial representations on two databases is presented. We further analyze the facial representations from two perspectives: the redundancy of the facial representations, and the sensitivity of facial representations to various head pose changes. Finally, two influence factors for BMI analysis are discussed.

### 5.1. Experiment setting

#### 5.1.1. Database setting

The BMI values are estimated based on separated training in each gender (and ethnicity) group using the SVR model in this work without any other specification. To evaluate the overall performance of the seven facial representations, FIW-BMI is split into 10 subsets for each gender group (10 subsets for the male group and 10 subsets for female). A similar process is applied to Morph II. It is also split into 10 subsets for each gender and ethnicity group. We use a cross-validation for performance measure. 8 subsets are used as the training set and the remaining 2 are the testing set in each round for each gender-ethnicity group. There is no overlap of individuals between the training and test sets in each round. Such a process is repeated 30 rounds for each group (the training and test sets are different for each round). Then 95% confidence intervals are calculated based on the results of these 30 repeated experiments.

To analyze the redundancy of the facial representations and the sensitivity to head pose variations, the training and test sets of the

[1] Please contact the authors for the dataset.
[2] Website: http://www.reddit.com/r/progresspics.

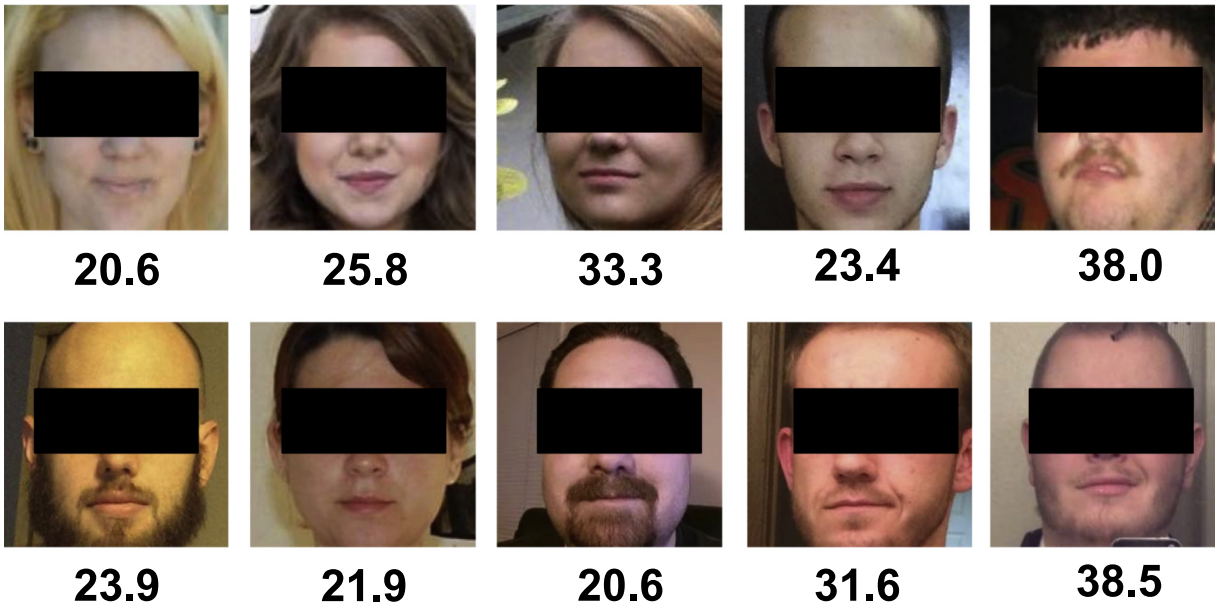


**Fig. 5.** Samples of the cleaned images in FIW-BMI database.

two databases are given in Tables 2 and 3, respectively. The same individual can only appear in either training or test set, but not both.

*5.1.2. Images preprocessing*

The face image alignment is applied prior to the extraction of geometric features. The alignment is based on the detected eye coordinates. It basically performs translation, rotation, and scaling of the faces so as to align all face images into the common eye coordinates. The output is a cropped $256 \times 256$ image. The Openface toolkit [27] is employed for detecting 68 face landmarks. The output of the PIGF is a 7 dimensional representation: $[CJWR, WHR, PAR, ES, LF/FH, FW/LFH, MEH]^T$. The PF consists of the coordinates of 68 facial landmarks, denoted as $(x_i, y_i), i = 1, \ldots, 68$, resulting in a 136 dimensional representation: $[x_1, y_1, \ldots, x_n, y_n, \ldots, x_{68}, y_{68}]^T$.

The implemented and pre-trained models of VGG-Face, LightCNN-29 and Centerloss are used from the Caffe deep learning framework [28]. All weights of the fully-connected layer of each deep network are used for feature extraction. These layers are noted as fc6 in VGG-Face, fc1 in LightCNN and fc5 in Centerloss. The VGG-Face model takes a $224 \times 224$ color image with the mean subtracted and outputs a 4096 dimensional feature vector. The LightCNN model provides a 256 dimensional representation extracted from a $128 \times 128$ gray-scale image. The Centerloss model outputs a 1024 dimensional representation with the input $96 \times 112$ color images. The Arcface model takes $112 \times 112$ color image and outputs a 512 dimensional feature vector. The image alignment is required before extracting deep representations. It is done by following the alignment protocol provided by each deep model.

*5.1.3. Implementation details for machine learning*

As shown in Fig. 1, the extracted facial representations are then used to train a regression model. We employ the support vector regression (SVR) [29] model to learn the mapping from the extracted representations to BMI values. The SVR is selected due to its robust generalization behavior. The Gaussian Radial Basis Function (RBF) is utilized as the SVR kernel.

**Table 1**
Details about the selected Morph II database.

| | Black male | Black female | White male | White female |
|---|---|---|---|---|
| #Subject | 6497 | 1096 | 1565 | 535 |
| #Images | 19,290 | 2824 | 4862 | 2057 |

*5.2. Performance metrics*

Mean absolute error (MAE) is employed to measure the performance on BMI estimation. It is defined as the average of the absolute error between the estimated BMI values and the ground truth BMI values, which is computed by: $MAE = \frac{1}{N}\sum_{k=1}^{N}|\hat{p}_k - p_k|$, here $p_k$ is the ground truth BMI value for image $k$, $\hat{p}_k$ is the corresponding estimated BMI value, $N$ is the number of test images. This measure is motivated by its use in age estimation [30].

The second measurement is the accuracy of the predicted BMI category. According to the estimated BMI values, we can compute the corresponding BMI category (underweight, normal, overweight and obese). The accuracy of the predicted category is the proportion

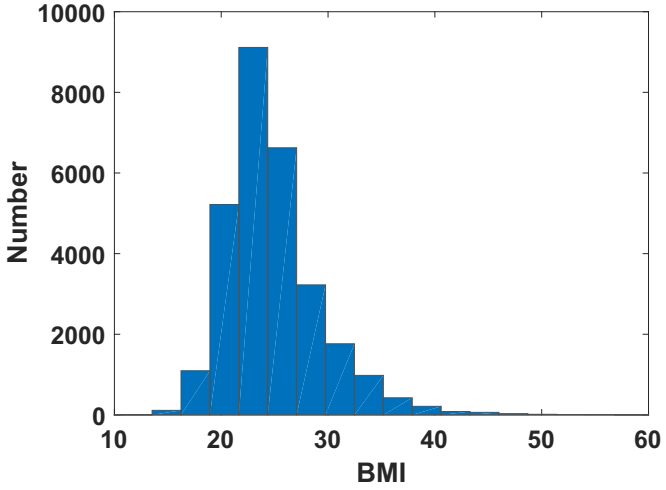


**Fig. 6.** Distribution of BMI values on Morph II. The BMI values mainly distribute between 15 to 35.

**Table 2**
Splitting FIW-BMI by gender.

| | Training set | | Test set | |
|---|---|---|---|---|
| | #Subject | #Images | #Subject | #Images |
| Male | 2551 | 4136 | 641 | 1061 |
| Female | 1329 | 2164 | 360 | 569 |

**Table 3**
Splitting selected Morph II by gender and ethnicity.

| | Training set | | Test set | |
|---|---|---|---|---|
| | #Subject | #Images | #Subject | #Images |
| Black male | 4568 | 13,574 | 1929 | 5716 |
| Black female | 873 | 2218 | 223 | 606 |
| White male | 1245 | 3856 | 320 | 1006 |
| White female | 428 | 1615 | 107 | 442 |

of the total number of predictions that are correct. This measurement is helpful to decide if the errors are acceptable. For example, given an image with ground-truth BMI value 24, the estimated value is 19. Though the absolute error is 5, the predicted category (normal) is correct. On the other hand, the category has a limitation. For example, if the ground-truth BMI of an image is 30 and the estimated value is 30.5, though the absolute error is 0.5, the predicted category (obese) is incorrect.

Mean absolute percentage error (MAPE) is proposed as the third measure. It is a relative error computed as:

$$MAPE = \frac{100}{N}\sum_{k=1}^{N}\left|\frac{\hat{p_k} - p_k}{p_k}\right|. \quad (1)$$

Considering the advantages and limitations of the above three measurements, we use all of them to evaluate the performance.

A 95% confidence interval (CI) is a range of values that it can be 95% certain contains the true mean of the population. We calculate the 95% confidence intervals of the above three metrics based on the results of 30 repeated experiments. It is computed by:

$$CI = \bar{X} \pm Z\frac{s}{\sqrt{n}}, \quad (2)$$

here $n$ is the number of observations, $\bar{X}$ is the mean of observations, and $s$ is the standard deviation. For 95% confidence interval, the Z value is 1.96.

### 5.3. Overall performance comparison

The 95% confidence interval of MAEs for the seven representations for BMI estimation on the two databases are given in Tables 4 and 5. The MAEs are calculated from the whole test set. To better present the details about the performance, we further calculated the MAEs from each BMI category. In addition to the MAEs, the 95% confidence intervals of the accuracy for the predicted BMI category are given in Tables 6 and 7. The 95% confidence interval of MAPEs are given in Tables 8 and 9. Combining MAEs (Tables 4 and 5), the accuracy for category classification (Tables 6 and 7) and MAPEs (Tables 8 and 9) to evaluate and analyze the performances with more specific information, some interesting observations can be obtained.

#### 5.3.1. Performance of the two types of facial representations

The performances of the seven facial representations are different from each other. Overall, the experimental results show that these

**Table 4**
95% confidence interval of MAEs for the seven facial representations for BMI prediction on FIW-BMI.

| | Male | | | | | Female | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 3.78 ± 0.07 | 7.90 ± 0.25 | 3.79 ± 0.08 | 2.52 ± 0.50 | 4.76 ± 0.14 | 4.26 ± 0.08 | 10.37 ± 1.10 | 5.30 ± 0.07 | 2.68 ± 0.05 | 4.68 ± 0.13 |
| PF | 3.76 ± 0.07 | 8.96 ± 0.37 | 3.81 ± 0.10 | 2.51 ± 0.04 | 4.56 ± 0.12 | 4.15 ± 0.08 | 9.43 ± 0.90 | 5.08 ± 0.09 | 2.91 ± 0.07 | 4.60 ± 0.10 |
| PIGF + PF | 3.70 ± 0.07 | 8.16 ± 0.27 | 3.79 ± 0.09 | 2.50 ± 0.04 | 4.49 ± 0.10 | 4.10 ± 0.07 | 9.97 ± 0.87 | 5.02 ± 0.08 | 2.71 ± 0.07 | 4.53 ± 0.12 |
| VGG-Face | 3.26 ± 0.06 | 5.61 ± 0.34 | 2.99 ± 0.10 | 2.50 ± 0.05 | 4.24 ± 0.10 | 3.66 ± 0.08 | 9.79 ± 0.95 | 4.42 ± 0.11 | 2.67 ± 0.09 | 3.81 ± 0.12 |
| LightCNN | 3.44 ± 0.06 | 5.76 ± 0.36 | 3.17 ± 0.09 | 2.50 ± 0.05 | 4.30 ± 0.09 | 3.90 ± 0.03 | 9.94 ± 1.00 | 4.62 ± 0.09 | 2.86 ± 0.07 | 4.12 ± 0.06 |
| Centerloss | 3.40 ± 0.05 | 8.02 ± 0.35 | 3.19 ± 0.08 | 2.54 ± 0.05 | 4.30 ± 0.11 | 3.82 ± 0.11 | 8.56 ± 0.50 | 5.00 ± 0.21 | 2.70 ± 0.11 | 3.97 ± 0.11 |
| ArcFace | **3.15 ± 0.07** | 5.52 ± 0.21 | 3.18 ± 0.04 | 2.25 ± 0.05 | 4.07 ± 0.14 | **3.51 ± 0.09** | 9.76 ± 0.85 | 4.47 ± 0.08 | 2.53 ± 0.08 | 3.62 ± 0.17 |

**Table 5**
95% confidence interval of MAEs for the seven facial representations for BMI prediction on Morph II database.

| | Black male | | | | | Black female | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 2.70 ± 0.04 | 6.61 ± 0.22 | 1.82 ± 0.02 | 2.36 ± 0.01 | 7.38 ± 0.10 | 3.77 ± 0.08 | 6.25 ± 0.16 | 2.26 ± 0.05 | 2.95 ± 0.09 | 8.38 ± 0.19 |
| PF | 2.67 ± 0.04 | 6.69 ± 0.24 | 1.84 ± 0.02 | 2.29 ± 0.02 | 7.22 ± 0.11 | 3.76 ± 0.08 | 6.53 ± 0.15 | 2.35 ± 0.05 | 2.43 ± 0.09 | 8.33 ± 0.19 |
| PIGF + PF | 2.63 ± 0.04 | 6.61 ± 0.24 | 1.84 ± 0.03 | 2.28 ± 0.02 | 7.14 ± 0.11 | 3.68 ± 0.07 | 6.37 ± 0.14 | 2.39 ± 0.06 | 2.60 ± 0.08 | 8.11 ± 0.15 |
| VGG-Face | 2.45 ± 0.05 | 6.01 ± 0.21 | 1.87 ± 0.03 | 2.10 ± 0.03 | 5.73 ± 0.11 | 3.48 ± 0.04 | 5.21 ± 0.14 | 2.25 ± 0.05 | 2.61 ± 0.08 | 7.25 ± 0.16 |
| LightCNN | 2.42 ± 0.10 | 5.81 ± 0.16 | 1.78 ± 0.02 | 2.16 ± 0.02 | 5.84 ± 0.10 | 3.55 ± 0.05 | 5.40 ± 0.17 | 2.38 ± 0.05 | 2.53 ± 0.05 | 7.82 ± 0.23 |
| Centerloss | 2.50 ± 0.04 | 6.23 ± 0.28 | 1.84 ± 0.03 | 2.12 ± 0.04 | 6.37 ± 0.12 | 3.63 ± 0.06 | 5.41 ± 0.16 | 2.42 ± 0.06 | 2.71 ± 0.10 | 7.94 ± 0.16 |
| ArcFace | **2.40 ± 0.03** | 6.24 ± 0.28 | 1.85 ± 0.03 | 2.01 ± 0.03 | 5.65 ± 0.08 | **3.51 ± 0.07** | 5.12 ± 0.13 | 2.43 ± 0.05 | 2.57 ± 0.07 | 7.26 ± 0.17 |
| | White male | | | | | White female | | | | |
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 2.67 ± 0.03 | 5.73 ± 0.36 | 1.94 ± 0.03 | 2.35 ± 0.06 | 7.30 ± 0.15 | 2.96 ± 0.04 | 4.27 ± 0.21 | 1.73 ± 0.03 | 4.74 ± 0.09 | 8.82 ± 0.38 |
| PF | 2.57 ± 0.03 | 5.85 ± 0.43 | 1.86 ± 0.03 | 2.28 ± 0.04 | 7.17 ± 0.15 | 3.13 ± 0.07 | 4.68 ± 0.18 | 1.88 ± 0.03 | 4.41 ± 0.10 | 8.88 ± 0.38 |
| PIGF + PF | 2.49 ± 0.03 | 5.69 ± 0.37 | 1.84 ± 0.03 | 2.22 ± 0.05 | 7.15 ± 0.17 | 3.01 ± 0.05 | 4.42 ± 0.17 | 1.78 ± 0.04 | 4.35 ± 0.09 | 8.80 ± 0.41 |
| VGG-Face | **2.30 ± 0.03** | 4.83 ± 0.31 | 1.82 ± 0.03 | 2.08 ± 0.04 | 5.35 ± 0.21 | 2.96 ± 0.06 | 4.11 ± 0.15 | 1.77 ± 0.08 | 3.99 ± 0.09 | 8.72 ± 0.32 |
| LightCNN | 2.35 ± 0.04 | 4.73 ± 0.25 | 1.82 ± 0.04 | 1.98 ± 0.03 | 6.15 ± 0.15 | **2.87 ± 0.07** | 4.08 ± 0.17 | 1.72 ± 0.05 | 3.89 ± 0.07 | 7.80 ± 0.61 |
| Centerloss | 2.41 ± 0.03 | 5.21 ± 0.47 | 1.87 ± 0.07 | 2.09 ± 0.05 | 6.03 ± 0.25 | 2.94 ± 0.09 | 4.22 ± 0.22 | 1.79 ± 0.07 | 3.89 ± 0.15 | 8.94 ± 0.52 |
| ArcFace | 2.32 ± 0.02 | 5.45 ± 0.30 | 1.77 ± 0.03 | 1.96 ± 0.04 | 6.27 ± 0.18 | 2.90 ± 0.05 | 4.02 ± 0.16 | 1.76 ± 0.03 | 3.42 ± 0.10 | 8.63 ± 0.36 |

**Table 6**
95% confidence interval of BMI category prediction accuracy (%) for the seven facial representations on FIW-BMI.

| | Male | | | | | Female | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 72.1 ± 0.6 | 3.8 ± 2.1 | 45.8 ± 1.0 | 79.0 ± 0.7 | 80.6 ± 0.7 | 68.6 ± 0.7 | 3.5 ± 2.1 | 18.0 ± 1.5 | 78.9 ± 1.0 | 82.5 ± 1.2 |
| PF | 73.9 ± 0.6 | 4.0 ± 2.7 | 46.0 ± 1.3 | 76.7 ± 0.8 | 81.7 ± 0.7 | 69.4 ± 0.7 | 3.0 ± 1.9 | 29.8 ± 1.1 | 70.1 ± 4.4 | 83.2 ± 1.0 |
| PIGF + PF | 74.1 ± 0.5 | 4.1 ± 2.5 | 46.2 ± 1.1 | 78.5 ± 0.7 | 82.3 ± 0.6 | 70.1 ± 0.6 | 3.2 ± 1.8 | 28.8 ± 1.3 | 75.4 ± 2.3 | 83.5 ± 1.1 |
| VGG-Face | 78.0 ± 0.6 | 15.5 ± 5.8 | 62.5 ± 1.3 | 79.8 ± 1.2 | 85.5 ± 0.8 | 74.5 ± 0.7 | 4.9 ± 1.9 | 35.9 ± 2.3 | 73.5 ± 1.6 | 89.9 ± 0.8 |
| LightCNN | 76.3 ± 0.5 | 3.3 ± 2.2 | 60.0 ± 2.6 | 74.3 ± 0.8 | 86.7 ± 0.5 | 71.9 ± 0.6 | 4.5 ± 2.0 | 36.8 ± 1.6 | 68.2 ± 1.3 | 87.8 ± 0.7 |
| Centerloss | 75.4 ± 0.5 | 4.1 ± 3.0 | 60.3 ± 1.6 | 72.9 ± 1.0 | 85.8 ± 0.6 | 73.3 ± 1.3 | 4.7 ± 2.1 | 33.1 ± 3.7 | 73.8 ± 1.7 | 88.2 ± 0.8 |
| ArcFace | **79.1 ± 0.3** | 5.7 ± 2.7 | 62.4 ± 0.9 | 77.5 ± 0.9 | 90.3 ± 0.6 | **75.7 ± 0.8** | 5.2 ± 2.3 | 39.5 ± 1.8 | 72.0 ± 1.2 | 91.4 ± 1.0 |

two types of facial representations both are effective for addressing BMI estimation. And the deep model based methods (VGG-Face, LightCNN, Centerloss and Arcface) perform better than the geometry based methods (PIGF, PF, and PIGF + PF). Among them, measuring with MAEs, the VGG-Face and Arcface show more robustness than the others in most cases.

For the white female group, the deep learning based representations do not show clear advantages over the geometric representations as in other groups. From Table 1, we can see this group has the least number of images for training and testing. Since the training time of SVR models for deep representations is much longer than the geometric representations. The geometric representations are more suitable for small datasets. The deep representations perform better on a large dataset with much more time cost.

From Tables 4 and 5, it can be seen that the confidence intervals of PIGF + PF are smaller than both PIGF and PF for most groups. To decide whether a significant performance difference exists between the fused geometric feature (PIGF + PF) and the individual feature (PIGF, PF), we apply a hypothesis testing with a statistical significance measure. The null hypothesis is: there is no performance difference between the two features. We can make a decision by:

- If the p-value is smaller than the significance level $\alpha$, it can reject the null hypothesis;
- If the p-value is larger than the significance level $\alpha$, it fails to reject the null hypothesis.

Here the significance level $\alpha$ is set to 0.01. The p-value is computed from the MAEs of the two features obtained from the repeated (30 times) experiments on each group of the two databases. According to the calculation, the range of p-value is from 2.7e−3 to 1.3e−06. This result reveals the significant differences between the fused geometric feature (PIGF + PF) and the individual feature (PIGF, PF).

### 5.3.2. Performance on the four BMI categories

All seven representations have different performances in the four categories. As shown in Tables 4, 5, 8 and 9, the MAEs and MAPEs

**Table 7**
95% of BMI category prediction accuracy (%) confidence interval for the seven facial representations on Morph II.

| | Black male | | | | | Black female | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 71.7 ± 0.7 | 4.0 ± 0.3 | 94.7 ± 0.2 | 51.0 ± 0.9 | 19.5 ± 0.9 | 65.0 ± 1.0 | 2.5 ± 0.8 | 93.9 ± 0.6 | 48.5 ± 1.9 | 29.3 ± 3.1 |
| PF | 73.0 ± 0.6 | 3.8 ± 0.3 | 94.5 ± 0.3 | 55.2 ± 1.0 | 24.7 ± 1.2 | 67.7 ± 1.2 | 0.6 ± 0.5 | 91.4 ± 1.0 | 58.5 ± 1.9 | 33.2 ± 2.5 |
| PIGF + PF | 73.0 ± 0.6 | 3.7 ± 0.5 | 94.7 ± 0.2 | 55.9 ± 1.0 | 21.8 ± 1.2 | 67.9 ± 1.1 | 0.9 ± 0.3 | 92.4 ± 1.3 | 59.7 ± 1.0 | 31.2 ± 2.6 |
| VGG-Face | 73.5 ± 0.3 | 14.1 ± 2.5 | 89.0 ± 0.7 | 57.1 ± 2.1 | 35.4 ± 1.7 | 65.9 ± 1.0 | 10.1 ± 1.7 | 87.1 ± 1.4 | 60.7 ± 1.3 | 30.0 ± 2.3 |
| LightCNN | 77.1 ± 0.2 | 17.3 ± 3.5 | 90.7 ± 0.2 | 62.7 ± 0.3 | 48.3 ± 1.1 | 70.3 ± 0.8 | 14.2 ± 5.5 | 88.9 ± 0.7 | 64.6 ± 2.0 | 48.7 ± 2.1 |
| Centerloss | 75.6 ± 0.4 | 16.9 ± 4.1 | 93.1 ± 0.6 | 61.7 ± 1.9 | 33.5 ± 1.7 | 66.4 ± 1.2 | 9.3 ± 1.6 | 89.1 ± 1.1 | 59.4 ± 1.7 | 30.7 ± 3.6 |
| ArcFace | **78.4 ± 0.4** | 18.7 ± 3.5 | 93.0 ± 0.7 | 66.4 ± 1.6 | 47.3 ± 1.3 | **69.2 ± 1.0** | 7.0 ± 1.2 | 88.2 ± 1.0 | 64.7 ± 1.1 | 45.0 ± 3.2 |
| | White male | | | | | White female | | | | |
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 71.9 ± 1.0 | 5.5 ± 1.1 | 95.9 ± 0.4 | 53.2 ± 1.9 | 4.6 ± 0.6 | 70.2 ± 1.0 | 5.5 ± 1.0 | 99.1 ± 0.5 | 19.2 ± 2.2 | 25.3 ± 3.9 |
| PF | 72.5 ± 0.9 | 4.5 ± 0.9 | 95.7 ± 0.3 | 54.4 ± 1.6 | 7.7 ± 1.4 | 70.6 ± 1.2 | 13.9 ± 3.6 | 99.6 ± 0.2 | 22.5 ± 1.9 | 24.7 ± 1.4 |
| PIGF + PF | 74.5 ± 1.0 | 4.7 ± 1.0 | 91.7 ± 0.4 | 61.5 ± 2.0 | 28.9 ± 1.5 | 69.1 ± 1.1 | 9.5 ± 1.5 | 99.7 ± 0.2 | 14.1 ± 3.4 | 15.1 ± 4.9 |
| VGG-Face | 75.2 ± 0.7 | 11.1 ± 2.4 | 90.5 ± 0.9 | 60.5 ± 1.1 | 33.0 ± 1.7 | **73.2 ± 0.8** | 30.4 ± 1.9 | 98.7 ± 0.2 | 27.9 ± 2.1 | 15.9 ± 3.1 |
| LightCNN | **76.7 ± 0.6** | 9.2 ± 1.9 | 91.5 ± 0.4 | 64.1 ± 1.4 | 36.7 ± 2.5 | 72.9 ± 0.6 | 27.4 ± 2.5 | 97.6 ± 0.2 | 30.2 ± 2.3 | 26.5 ± 3.6 |
| Centerloss | 75.8 ± 0.7 | 10.7 ± 2.5 | 91.6 ± 0.9 | 62.8 ± 1.0 | 30.6 ± 2.6 | 70.1 ± 1.0 | 17.0 ± 1.8 | 99.4 ± 0.2 | 27.8 ± 2.7 | 19.4 ± 2.7 |
| ArcFace | 75.8 ± 0.5 | 14.5 ± 3.7 | 94.0 ± 1.5 | 63.8 ± 1.1 | 17.8 ± 3.7 | 70.6 ± 0.7 | 17.5 ± 1.4 | 98.7 ± 0.3 | 34.5 ± 1.5 | 14.6 ± 2.6 |

**Table 8**
95% confidence interval of MAPEs (%) for the seven facial representations for BMI prediction on FIW-BMI.

| | Male | | | | | Female | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 13.7 ± 0.3 | 31.6 ± 0.7 | 14.0 ± 0.3 | 8.7 ± 0.2 | 17.5 ± 0.5 | 15.5 ± 0.3 | 38.3 ± 2.7 | 19.4 ± 0.2 | 8.7 ± 0.1 | 17.6 ± 0.5 |
| PF | 13.4 ± 0.3 | 33.0 ± 1.0 | 13.9 ± 0.4 | 8.9 ± 0.2 | 16.7 ± 0.4 | 15.0 ± 0.5 | 34.4 ± 2.3 | 18.8 ± 0.3 | 9.5 ± 0.2 | 16.3 ± 0.6 |
| PIGF + PF | 13.3 ± 0.3 | 32.1 ± 0.8 | 14.1 ± 0.3 | 8.8 ± 0.2 | 16.2 ± 0.4 | 14.9 ± 0.3 | 36.7 ± 2.5 | 18.2 ± 0.2 | 8.8 ± 0.1 | 16.3 ± 0.5 |
| VGG-Face | 11.4 ± 0.2 | 23.9 ± 1.5 | 12.0 ± 0.3 | 8.5 ± 0.1 | 13.0 ± 0.3 | 12.3 ± 0.2 | 35.5 ± 3.5 | 17.1 ± 0.4 | 9.0 ± 0.2 | 12.5 ± 0.3 |
| LightCNN | 11.6 ± 0.2 | 24.2 ± 1.2 | 11.9 ± 0.3 | 8.5 ± 0.2 | 13.5 ± 0.2 | 12.7 ± 0.1 | 34.8 ± 2.7 | 16.2 ± 0.3 | 9.1 ± 0.2 | 13.4 ± 0.3 |
| Centerloss | 11.9 ± 0.1 | 30.0 ± 1.0 | 12.3 ± 0.2 | 9.1 ± 0.2 | 13.7 ± 0.3 | 12.6 ± 0.3 | 35.6 ± 1.9 | 17.5 ± 0.7 | 8.8 ± 0.3 | 12.7 ± 0.4 |
| ArcFace | **11.0 ± 0.2** | 24.0 ± 0.7 | 11.8 ± 0.1 | 9.1 ± 0.1 | 12.3 ± 0.3 | **12.0 ± 0.2** | 35.9 ± 2.2 | 16.7 ± 0.3 | 9.3 ± 0.2 | 11.8 ± 0.4 |

**Table 9**
95% confidence interval of MAPEs (%) for the seven facial representations for BMI prediction on Morph II.

| | Black male | | | | | Black female | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 10.9 ± 0.2 | 27.9 ± 0.9 | 7.5 ± 0.1 | 9.7 ± 0.1 | 28.5 ± 0.4 | 15.3 ± 0.3 | 26.2 ± 0.6 | 9.3 ± 0.2 | 12.6 ± 0.4 | 32.4 ± 1.0 |
| PF | 10.7 ± 0.1 | 28.2 ± 0.9 | 7.4 ± 0.1 | 9.4 ± 0.1 | 27.7 ± 0.4 | 15.1 ± 0.3 | 26.0 ± 0.5 | 9.6 ± 0.2 | 11.8 ± 0.4 | 31.7 ± 0.9 |
| PIGF + PF | 10.7 ± 0.1 | 27.9 ± 0.9 | 7.5 ± 0.1 | 9.3 ± 0.1 | 27.7 ± 0.4 | 15.1 ± 0.2 | 26.1 ± 0.5 | 9.0 ± .2 | 12.0 ± 0.3 | 33.5 ± 0.9 |
| VGG-Face | 9.5 ± 0.1 | 25.9 ± 0.8 | 7.8 ± 0.1 | 8.7 ± 0.3 | 21.3 ± 0.4 | **12.4 ± 0.2** | 21.9 ± 0.5 | 9.2 ± 0.1 | 9.7 ± 0.4 | 24.3 ± 0.6 |
| LightCNN | 9.3 ± 0.1 | 24.7 ± 0.6 | 7.0 ± 0.1 | 8.3 ± 0.1 | 20.9 ± 0.2 | 12.9 ± 0.2 | 21.1 ± 0.7 | 9.5 ± 0.2 | 9.8 ± 0.3 | 24.5 ± 0.8 |
| Centerloss | 10.0 ± 0.1 | 26.4 ± 1.0 | 7.5 ± 0.1 | 8.7 ± 0.2 | 23.7 ± 0.5 | 14.5 ± 0.2 | 23.0 ± 0.6 | 10.0 ± 0.2 | 11.2 ± 0.3 | 29.8 ± 0.7 |
| ArcFace | **9.1 ± 0.1** | 25.5 ± 1.1 | 7.4 ± 0.1 | 8.1 ± 0.1 | 20.2 ± 0.3 | 12.8 ± 0.3 | 22.1 ± 0.5 | 10.0 ± 0.2 | 10.5 ± 0.3 | 24.3 ± 0.7 |
| | White male | | | | | White female | | | | |
| | All | Underweight | Normal | Overweight | Obese | All | Underweight | Normal | Overweight | Obese |
| PIGF | 10.5 ± 0.1 | 24.6 ± 1.5 | 7.6 ± 0.1 | 9.5 ± 0.2 | 28.5 ± 0.6 | 13.8 ± 0.2 | 19.6 ± 0.8 | 7.9 ± 0.1 | 22.0 ± 0.4 | 36.8 ± 1.9 |
| PF | 10.5 ± 0.1 | 24.9 ± 1.6 | 7.7 ± 0.3 | 9.5 ± 0.2 | 28.1 ± 0.7 | 10.5 ± 0.1 | 24.9 ± 1.6 | 7.7 ± 0.1 | 12.5 ± 0.2 | 28.1 ± 0.7 |
| PIGF + PF | 10.1 ± 0.1 | 24.7 ± 1.3 | 7.3 ± 0.3 | 9.6 ± 0.2 | 28.0 ± 0.5 | 11.9 ± 0.2 | 22.5 ± 1.1 | 7.7 ± 0.1 | 16.5 ± 0.3 | 32.1 ± 1.1 |
| VGG-Face | 8.6 ± 0.1 | 22.1 ± 1.1 | 7.0 ± 0.2 | 7.5 ± 0.1 | 19.3 ± 0.7 | 13.9 ± 0.5 | 17.3 ± 0.9 | 8.7 ± 0.2 | 19.9 ± 0.5 | 37.1 ± 2.3 |
| LightCNN | **8.5 ± 0.1** | 20.3 ± 0.9 | 6.8 ± 0.2 | 7.7 ± 0.1 | 18.9 ± 0.6 | **11.3 ± 0.3** | 16.6 ± 0.6 | 7.4 ± 0.2 | 13.6 ± 0.3 | 28.7 ± 3.1 |
| Centerloss | 9.7 ± 0.1 | 24.1 ± 1.8 | 7.7 ± 0.2 | 8.5 ± 0.2 | 22.6 ± 1.0 | 13.4 ± 0.4 | 19.2 ± 0.8 | 8.7 ± 0.2 | 17.4 ± 0.7 | 36.0 ± 2.6 |
| ArcFace | 9.6 ± 0.1 | 24.0 ± 1.2 | 7.7 ± 0.1 | 8.4 ± 0.1 | 23.0 ± 0.6 | 12.8 ± 0.2 | 18.7 ± 0.6 | 8.2 ± 0.1 | 15.3 ± 0.1 | 34.4 ± 1.8 |

are higher for underweight category on FIW-BMI dataset; MAEs and MAPEs are high for underweight and obese categories on selected Morph II dataset. This is caused by the BMI distributions of the datasets. From Figs. 5 and 6, it can be seen that most images in the FIW-BMI database are in the categories of overweight and obese, while most images in Morph II are in the normal and overweight categories. The performance of the facial representations is influenced by the number of training images. Less training images in the specific category leads to a worse performance for the corresponding category.

#### 5.3.3. Performance on the two databases

Comparing to Morph II, all facial representations show less robustness on FIW-BMI database. This is caused by the wild data collection of FIW-BMI. Slight head pose changes exist on this database, and the BMI values distribute in a larger range on FIW-BMI (20–55) than the Morph II (15–35). To further analyze the performances of these facial representations on the two databases, we do another experiment which uses Morph II for training and FIW-BMI for testing, and vice versa. The experimental results are given in Tables 10 and 11. Comparing with the results in Tables 4 and 5, one can see that the performances of all seven facial representations drop significantly. This may be caused by the quite different BMI distributions of the two databases and the different "domains" of the images (Morph II has mugshot face images, while FIW-BMI is with daily life face images).

### 5.4. Redundancy in facial representations

According to the overall performance of these facial representations on BMI estimation, it is shown that the deep representations perform better on a large dataset. Since the number of training samples is limited, we try to eliminate the negative influence caused by

**Table 10**
Performance of the seven facial representations where using Morph II for training and FIW-BMI for testing.

| | Male | | | Female | | |
|---|---|---|---|---|---|---|
| | MAE | Accuracy (%) | MAPE (%) | MAE | Accuracy (%) | MAPE (%) |
| PIGF | 5.41 | 53.5 | 20.6 | 6.73 | 39.4 | 25.2 |
| PF | 5.35 | 54.3 | 21.9 | 6.86 | 40.7 | 26.8 |
| PIGF + PF | 5.30 | 54.2 | 21.2 | 6.88 | 40.9 | 26.9 |
| VGG-Face | 4.32 | 64.5 | 15.3 | 5.79 | 51.9 | 20.5 |
| LightCNN | 4.45 | 63.2 | 15.6 | 5.91 | 51.3 | 20.6 |
| Centerloss | 4.61 | 60.8 | 16.5 | 6.22 | 49.5 | 22.6 |
| ArcFace | 4.21 | 64.1 | 15.1 | 5.73 | 52.2 | 20.1 |

**Table 11**
Performance of the seven facial representations where using FIW-BMI for training and Morph II for testing.

| | Black male | | | Black female | | | White male | | | White female | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MAE | Accuracy (%) | MAPE (%) | MAE | Accuracy (%) | MAPE (%) | MAE | Accuracy (%) | MAPE (%) | MAE | Accuracy (%) | MAPE (%) |
| PIGF | 4.82 | 50.1 | 15.3 | 5.80 | 44.7 | 18.4 | 5.58 | 40.2 | 18.3 | 6.43 | 37.7 | 22.4 |
| PF | 4.95 | 46.7 | 16.7 | 5.95 | 42.6 | 20.0 | 5.82 | 36.8 | 19.0 | 6.58 | 37.0 | 24.5 |
| PIGF + PF | 4.72 | 48.5 | 16.1 | 5.86 | 43.6 | 19.8 | 5.78 | 37.2 | 18.9 | 6.40 | 38.6 | 23.9 |
| VGG-Face | 4.00 | 56.4 | 14.1 | 5.52 | 44.5 | 18.8 | 3.69 | 63.6 | 13.1 | 5.00 | 46.7 | 18.3 |
| LightCNN | 3.59 | 63.5 | 12.8 | 5.52 | 44.0 | 18.7 | 3.51 | 64.9 | 12.5 | 4.99 | 49.1 | 18.0 |
| Centerloss | 3.92 | 59.8 | 13.8 | 6.08 | 40.5 | 20.2 | 4.08 | 60.0 | 14.1 | 5.23 | 48.5 | 18.5 |
| ArcFace | 3.59 | 62.7 | 13.2 | 4.90 | 53.0 | 17.4 | 3.32 | 67.4 | 11.1 | 5.03 | 48.6 | 18.1 |

**Table 12**
Performance (MAEs) of applying PCA to the five facial representations for BMI prediction. A downward arrow (↓) denotes the MAE decreases, comparing with the method without PCA. And an upward arrow (↑) denotes the MAE increases.

| Method | FIW-BMI | | Morph II | | | |
|---|---|---|---|---|---|---|
| | Male | Female | Black male | Black female | White male | White female |
| PF + PCA | 3.82 ↑ | 4.14 ↑ | 2.67 ↑ | 3.75 ↑ | 2.73 ↑ | 3.14 ↑ |
| VGGFace + PCA | **3.15** ↓ | **3.57** ↓ | **2.41** ↓ | **3.56** ↓ | **2.41** ↓ | 2.91 ↓ |
| LightCNN + PCA | 3.41 ↑ | 3.86 ↑ | 2.45 ↑ | 3.71 ↑ | 2.49 ↑ | 3.08 ↑ |
| Centerloss + PCA | 3.31 ↓ | 3.77 ↑ | 2.50 ↑ | 3.73 ↑ | 2.50 ↑ | **2.86** ↓ |
| ArcFace + PCA | 3.19 ↑ | 3.62 ↑ | 2.38 ↑ | 3.51 ↑ | 2.57 ↑ | 2.94 ↓ |

the small number of training samples. Thereby it is essential to analyze the redundancy in facial representations and explore efficient methods to improve their performance.

One of the problems with high-dimensional features is that, in many cases, not all the measured features are relevant or important for understanding the underlying phenomena of interest. It is, therefore, interesting to analyze the redundancy in the representations. To figure out the issue, we first apply dimension reduction to the five facial representations (VGG-Face, LightCNN, Centerloss, Arcface and PF), then evaluate the performance of the reduced dimensions. As one of the typical dimension reduction methods, Principal Component Analysis (PCA) is selected. Note that the PCA projection is only learned with the training set. The dimensions of the four analyzed facial representations are as follows: PF is 128-dimension, VGG-Face is 4096-dimension, LightCNN is 256-dimension, Centerloss is 1024-dimension and Arcface is 512-dimension. Because the dimension of PIGF is seven and each dimension has its physical meaning (as mentioned in Section 3), PIGF and PIGF + PF are not involved in this investigation.

The percentage of explained variance is an index of the goodness of fit when applying PCA. It can be easily computed as the eigenvalues of corresponding components divided by the total variance. Here the total variance is the sum of all eigenvalues. Because the percentage of explained variance is a key factor to influence the performance of dimension reduced representations, different percentages (99%, 98%, 95%, 90%, 85%, 80%, and 75%) of explained variance for PCA are analyzed.

This experiment is conducted on FIW-BMI and selected Morph II dataset, respectively. And the details about the training and test sets are given in Tables 2 and 3. Table 12 presents the MAEs of applying PCA to the five facial representations for BMI estimation. Here the reported MAE is the best performance of each representation among the different percentages of explained variance. Comparing the MAEs of facial representations without applying PCA as given in Tables 4 and 5, we mark each MAE with a sign indicating the positive or negative effect of applying PCA to the representation. More specifically, a downward arrow (↓) denotes the MAE decreases (positive effect), and an upward arrow (↑) denotes the MAE increases (negative effect). It can be seen that VGG-Face + PCA performs better than VGG-Face in all groups on both databases. Centerloss + PCA achieves lower MAEs than Centerloss in the male group (BMI analysis database) and the white female group (Morph II). Arcface + PCA achieves lower MAE only in white female group (Morph II). While applying PCA to LightCNN and PF representations does not bring any positive effect. Such different changes observed in the five facial representations caused by the different feature redundancy. Thereby, it is concluded that removing the redundancy in VGG-Face representation can increase the accuracy and efficiency in BMI estimation.

More details about BMI estimation performance (MAEs) obtained by applying different percentages of explained variance are shown in Fig. 7 (FIW-BMI database) and Fig. 8 (Morph II). The horizontal axis denotes the percentages of explained variance. We conduct the experiment by seven different percentages: 99%, 98%, 95%, 90%, 85%, 80% and 75%, respectively. Here 100% denotes the facial representation without applying PCA. It can be seen that the curve denoted VGG-Face + PCA drops obviously after a short rise, while most of the

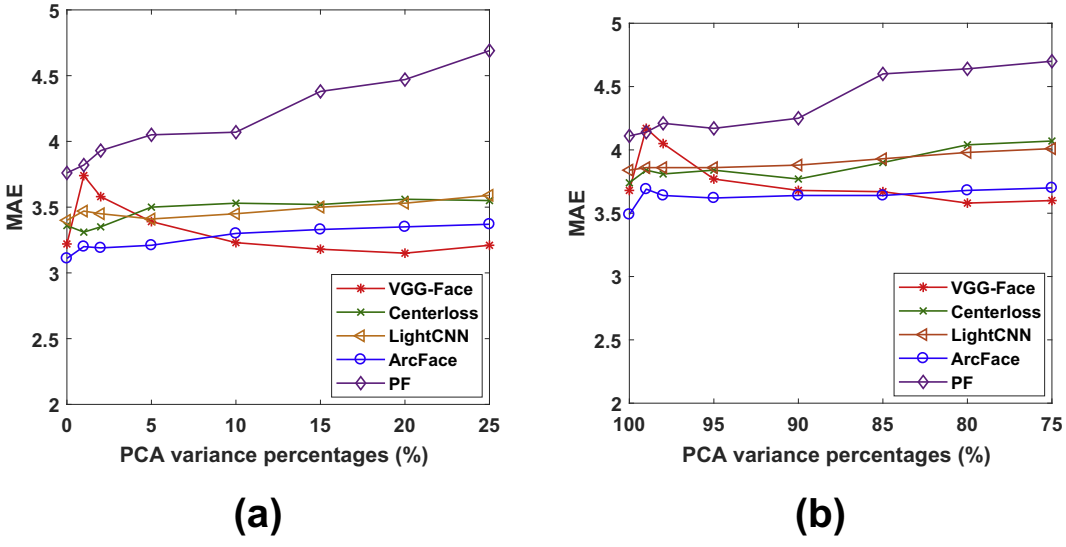


**Fig. 7.** BMI estimation error (measured by MAEs) of applying PCA to facial representations by different percentages of explained variance on FIW-BMI database. (a) is the results of the male group, and (b) is the female group.

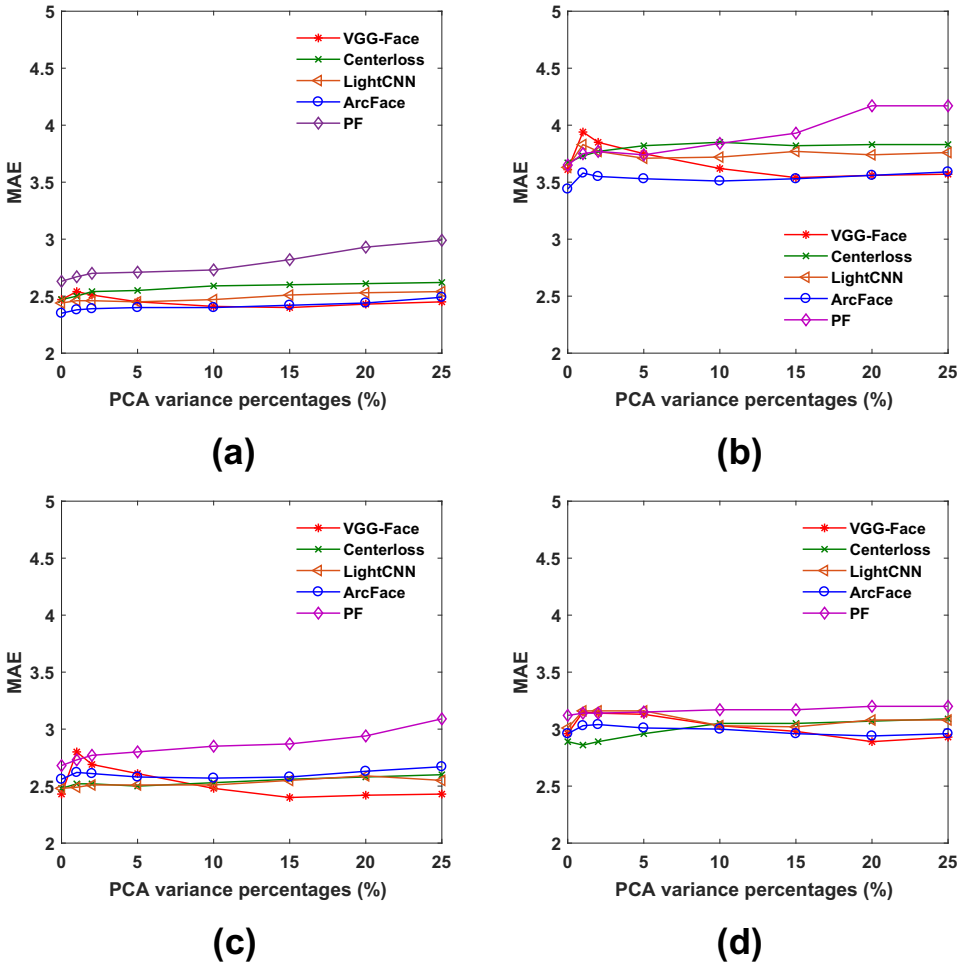


**Fig. 8.** BMI estimation error (measured by MAEs) of applying PCA to facial representations by different percentages of explained variance on Morph II. Each sub-figure shows the result of the different gender-ethnicity group: (a) black male, (b) black female, (c) white male, and (d) white female.

other curves are in a gradual uptrend. The best performance of VGG-Face is obtained at 80%–85% of the explained variance. Tables 13 and 14 report the kept dimensions of the five facial representations after applying PCA based on different percentages of explained variance on the two databases, respectively.

### 5.5. Sensitivity to head pose variations

The performance of face recognition is related to head pose changes. However, the influence of head pose on visual BMI estimation has not been studied yet. BMI estimation influenced by various head poses is conducted on the FIW-BMI dataset. As mentioned in Section 4.1, head pose variations exist in this database. To benchmark the robustness of the seven facial representations against pose variations, we group the face images by head pose angles.

Face pose distortion based sample pose (*SP*) index proposed by Marsico et al. [31] is utilized for measuring the head pose angles. SP index is given by the linear combination of three components, which are inversely proportional to *roll*, *yaw*, and *pitch*, respectively:

$$SP = \alpha(1 - roll) + \beta(1 - yaw) + \gamma(1 - pitch), \tag{3}$$

with $\alpha = 0.1$, $\beta = 0.6$ and $\gamma = 0.3$. See details about the calculation for *roll*, *yaw*, and *pitch* in [31], whose ranges are from 0

**Table 13**
The number of kept dimensions corresponding to different percentages of explained variance on FIW-BMI database.

| | Male | | | | | Female | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PF | VGG | LightCNN | Centerloss | ArcFace | PF | VGG | LightCNN | Centerloss | ArcFace |
| 99% | 19 | 1763 | 124 | 203 | 228 | 18 | 1326 | 121 | 199 | 218 |
| 98% | 15 | 1458 | 101 | 178 | 215 | 14 | 1101 | 99 | 174 | 203 |
| 95% | 10 | 988 | 67 | 138 | 189 | 10 | 755 | 67 | 130 | 174 |
| 90% | 7 | 616 | 50 | 104 | 160 | 7 | 479 | 47 | 94 | 145 |
| 85% | 6 | 414 | 41 | 85 | 138 | 6 | 325 | 38 | 74 | 124 |
| 80% | 5 | 285 | 34 | 71 | 121 | 5 | 224 | 32 | 61 | 107 |
| 75% | 4 | 196 | 29 | 60 | 106 | 4 | 153 | 27 | 51 | 93 |

**Table 14**
The number of kept dimensions corresponding to different percentages of explained variance on Morph II.

| | Black male | | | | | Black female | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PF | VGG | LightCNN | Centerloss | ArcFace | PF | VGG | LightCNN | Centerloss | ArcFace |
| 99% | 18 | 1675 | 118 | 177 | 216 | 16 | 1140 | 115 | 181 | 208 |
| 98% | 14 | 1352 | 93 | 144 | 199 | 13 | 905 | 90 | 154 | 188 |
| 95% | 9 | 825 | 59 | 97 | 167 | 9 | 568 | 58 | 112 | 153 |
| 90% | 6 | 441 | 41 | 66 | 135 | 6 | 325 | 40 | 78 | 119 |
| 85% | 5 | 265 | 32 | 50 | 114 | 5 | 204 | 31 | 59 | 98 |
| 80% | 4 | 166 | 26 | 41 | 97 | 4 | 131 | 25 | 46 | 82 |
| 75% | 3 | 104 | 22 | 34 | 84 | 4 | 84 | 20 | 38 | 69 |
| | White male | | | | | White female | | | | |
| | PF | VGG | LightCNN | Centerloss | ArcFace | PF | VGG | LightCNN | Centerloss | ArcFace |
| 99% | 18 | 1625 | 115 | 184 | 226 | 18 | 1076 | 115 | 191 | 213 |
| 98% | 13 | 1311 | 91 | 158 | 213 | 14 | 885 | 91 | 167 | 196 |
| 95% | 9 | 845 | 59 | 120 | 186 | 8 | 591 | 60 | 127 | 165 |
| 90% | 6 | 502 | 44 | 91 | 156 | 6 | 363 | 44 | 94 | 134 |
| 85% | 5 | 328 | 36 | 74 | 134 | 5 | 242 | 35 | 74 | 112 |
| 80% | 4 | 222 | 30 | 61 | 116 | 4 | 166 | 30 | 61 | 95 |
| 75% | 3 | 151 | 26 | 52 | 101 | 4 | 114 | 25 | 51 | 82 |

to 1, where 0 means almost no distortion and 1 means the worst distortion. Thereby, large *SP* represents small head pose and vice versa.

This experiment is conducted on FIW-BMI database. The dataset is divided as shown in Table 2. The number of images of the test set for each range of SP values in given in Table 15. The obtained MAEs of the seven facial representations for BMI estimation with various head poses on FIW-BMI database are shown in Fig. 9. The values of SP index are divided into four intervals: $SP \geq 0.9$, $0.9 > SP \geq 0.8$, $0.8 > SP \geq 0.7$ and $SP < 0.7$. It can be seen that when the SP decreases (head pose increases), the MAEs of the seven facial representation all increases, except the VGG-Face and Arcface representations on the male group in the interval $0.8 > SP \geq 0.7$. This experimental result demonstrates that large head pose changes lead to low performance for both geometric based and deep learning based representations. Thus the visual BMI estimation can be further improved by employing efficient pose normalization approaches.

It is interesting to observe that the VGG-Face and Arcface perform better on the range from 0.7 to 0.8 than on higher SP ranges in the male group. While such a phenomenon does not exist in the performance of the other two deep features (Centerloss and LightCNN). This may be caused by the different architectures of the four deep models and the different properties of the training sets. The VGG-Face and Arcface were trained on larger datasets which contain more pose conditions. In addition, the VGG-Face and Arcface have more sophisticated architectures which may lead to richer representations.

Among the seven facial representations, the Arcface, VGG-Face and PIGF show greater robustness than other representations w.r.t head pose variations, since the MAEs increase much less than the others. LightCNN, PF and PIGF + PF show lower robustness to head pose variations, since their performances drop significantly with the decrease of *SP* value, especially when the *SP* values are smaller than 0.7.

**Table 15**
The number of images for each range of SP values in the test set of FIW-BMI database.

| | Male | Female |
|---|---|---|
| $SP \geq 0.9$ | 464 | 307 |
| $0.9 > SP \geq 0.8$ | 468 | 57 |
| $0.8 > SP \geq 0.7$ | 109 | 202 |
| $SP < 0.7$ | 20 | 3 |

### 5.6. Discussion

We discuss the two influence factors on the performance of facial representations. One is the BMI distribution on the dataset. Another is the accuracy of landmark detection.

#### 5.6.1. Influence of BMI distribution on the estimation

As shown in Figs. 4 and 6, there are unbalanced BMI distribution over the two datasets. Very few samples distribute on the underweight category (BMI $\leq$ 18.5), while most samples distribute on the normal and overweight categories. This phenomenon also exists in real life. Most people are in normal and overweight ranges. To analyze the influence of unbalanced data on the estimated BMIs, we conduct an experiment on a balanced dataset. 5556 images were selected from the Morph II database. Among the selected images, 893 are underweight, 1788 are normal, 1505 are overweight and 1370 are obese. Fig. 10 shows the BMI distribution over the selected dataset. The Comparing with the BMI distribution in Fig. 6, Fig. 10 shows a relatively balanced distribution (with a relatively higher portion for underweight and obese). Then the images are randomly split into training and test sets. The training set contains 4319 images, and the test set contains 1237 images. There is non-overlap of individuals between the training and test sets. Considering the size of the training set is small, we use mixed training without separating the four gender and ethnicity groups. The experimental results on this balanced dataset are given in Table 16. Comparing with the results shown in Table 5, the performance on the balanced test set becomes worse. The experimental results indicate that the performance of BMI estimation depends on the prior distribution of the training set and the specific properties of the test set.

#### 5.6.2. Influence on accuracy by landmark detection

The three geometric facial representations are computed from the detected landmarks. Though the recently proposed facial landmark detection methods [27, 32] achieve a quite high accuracy with good resists for low resolution, blur and noisy images, an evaluation of the influence on the accuracy of BMI analysis is still necessary. To report a fair evaluation, we generate three sets of data. First, we randomly select 100 images from the Morph II dataset (there are no head pose variations in this dataset), and manually label all the needed landmarks (68 landmarks) for each image. These manually labeled landmarks are used as the ground truth. Then we apply an automatic landmark detection by Openface toolkit [27] to the selected 100 images, with

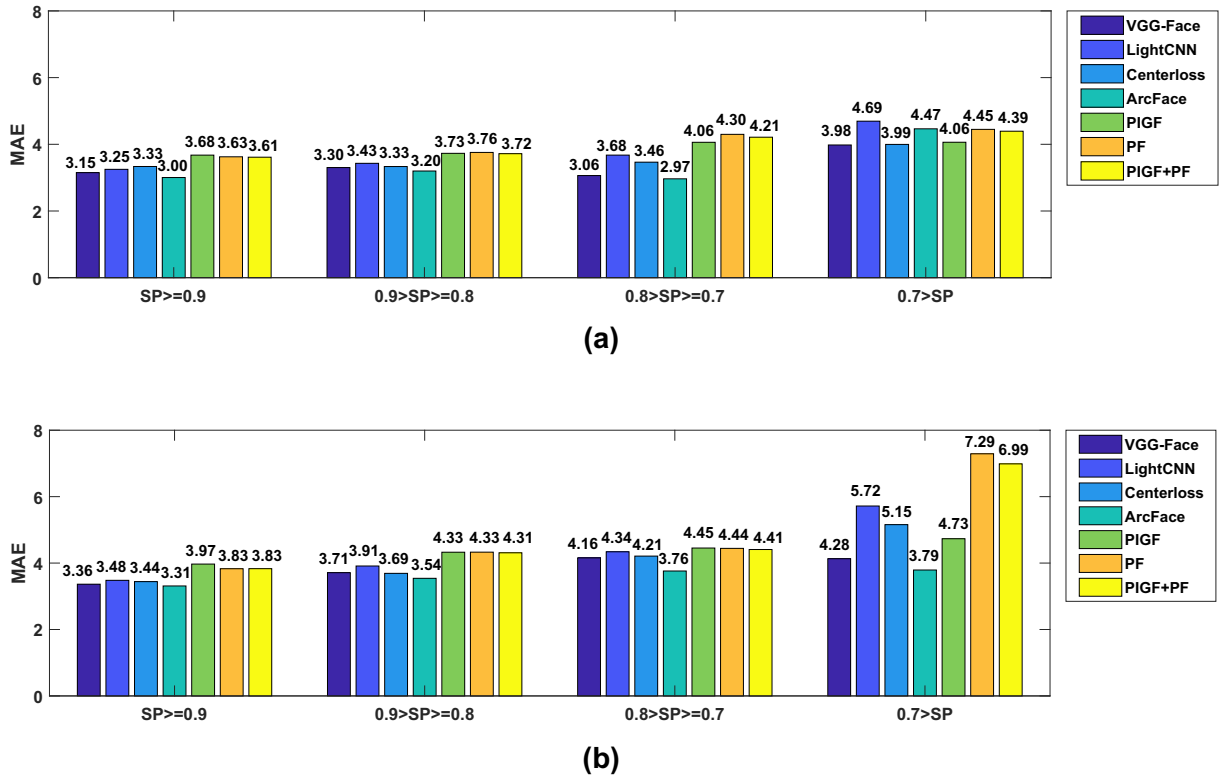


**Fig. 9.** The sensitivity of facial representations to invariant head pose. (a) shows the performance on the male group, and (b) shows the performance on the female group.

68 landmarks detected for each image. Finally, we generated noisy landmarks by adding white Gaussian noise to the ground truth landmarks with the mean set to 3 pixels, variance set to 2 pixels. The BMIs are estimated from these three sets of data by the three geometric representations. The experimental results are presented in Fig. 11. One can see that the difference between the MAEs from manually labeled landmarks and the automatically detected are very small. While the performance degrades significantly on the noisy set. Among the three representations, PIGF shows relatively more robust to inaccurate and noisy landmarks. These results justify that the accuracy of landmark detection methods has a limited influence to geometric facial representations.

Finally, we study another interesting problem. Whether more landmarks could bring an improvement to BMI estimation? The performance of PF feature with 119 landmarks [6] are analyzed on Morph II dataset, and compared with 68 landmarks. Fig. 12 shows an example of the detected 119 landmarks on a face image. The extended landmarks are around the neck, ears, forehead and around the vertex to the ears. The training and test sets are the same as shown in Table 3. Table 17 shows the comparison on the performance between 119 landmarks and 68 landmarks. As it can be observed, PF with 68 landmarks providing more promising results than 119 landmarks on each set. This reveals that the facial points around the neck, ears, forehead and the vertex to the ears are not as important as those round the face for estimating BMI values.

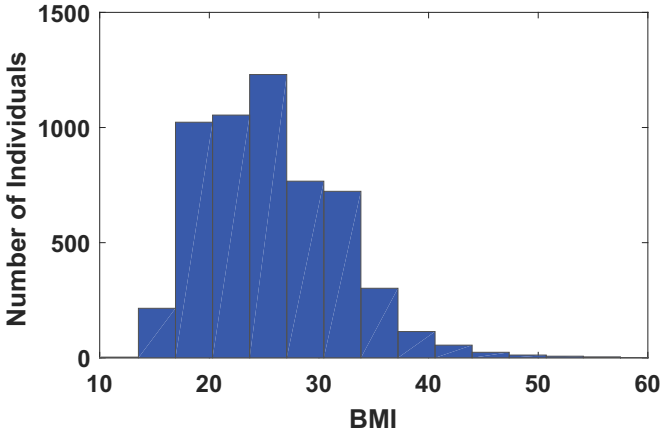


**Fig. 10.** The BMI distribution of the balanced dataset.

## 6. Conclusion

We have studied the visual BMI estimation problem systematically based on the facial representation or feature extraction. According to the inherent properties of representations, they are grouped into two types: geometric based and deep learning based. In addition to the two existing approaches (VGG-Face and PIGF), five other facial approaches: PF, PIGF + PF, LightCNN, Centerloss and Arcface are explored for the first time for BMI analysis. The performance and characteristics of the two types of facial representations have been comprehensively evaluated and analyzed from three perspectives: the overall performance on visual BMI prediction, the redundancy in

**Table 16**
MAE of estimated BMIs on the balanced dataset of selected Morph II.

| | Black male | Black female | White male | White female |
|---|---|---|---|---|
| PIGF | 3.91 | 4.87 | 3.53 | 3.92 |
| PF | 3.87 | 4.81 | 3.47 | 4.03 |
| PIGF + PF | 3.85 | 4.75 | 3.47 | 3.95 |
| VGG-Face | 3.43 | 4.22 | 2.90 | 3.54 |
| LightCNN | 3.50 | 4.38 | 3.43 | 3.99 |
| Centerloss | 3.43 | 4.46 | 2.91 | 3.43 |
| Arcface | 3.44 | 4.13 | 2.93 | 3.51 |

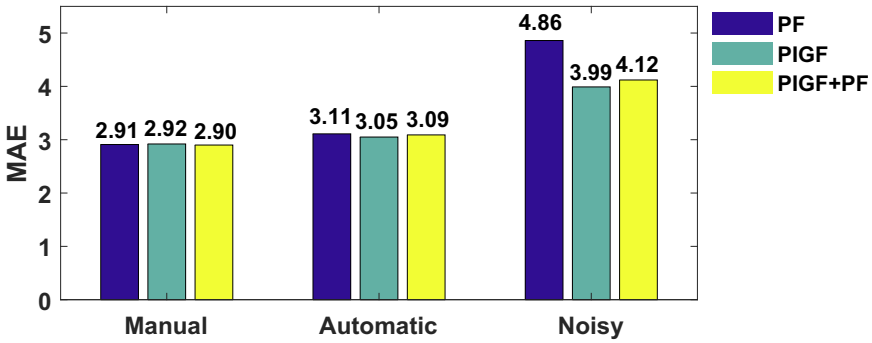


**Fig. 11.** Influence of accuracy of landmark detection to geometric facial representations.

representations and the sensitivity to head pose changes. The experiments are conducted on two databases: FIW-BMI and Morph II, exploring the capability of these approaches, which are summarized as below.

Experimentally we have found that the geometric representations are more suitable for the small dataset while the deep representations could perform better on large datasets with a much higher computation time cost. Among the seven representations, the VGG-Face and Arcface perform better than the others in most cases. For geometric features, more advantages can be achieved by the fused representation, PIGF + PF. The performance of the representations could be influenced by the training images and the BMI distribution.

Considering the limited number of training samples and high dimensions of some facial representations, we explored the efficient methods to improve the performance. We have analyzed the redundancy of the five facial representations (VGG-Face, LightCNN, Centerloss, Arcface and PF) by investigating the effect of applying PCA to the representations. Experimental results have shown that applying PCA to VGG-Face representation leads to better performance on BMI prediction with 80%–85% explained variance. Removing the redundancy in VGG-Face representation can increase the accuracy and efficiency in BMI estimation.

The sensitivity of facial representations to head pose variations for BMI estimation has been investigated as well. Experimental results have shown that large head pose changes lead to a low performance. Among the seven representations, The Arcface, VGG-Face and PIGF show better robustness than the others to head pose variations. The performance of LightCNN, PF and PIGF + PF drop significantly with the increase of head pose angles.

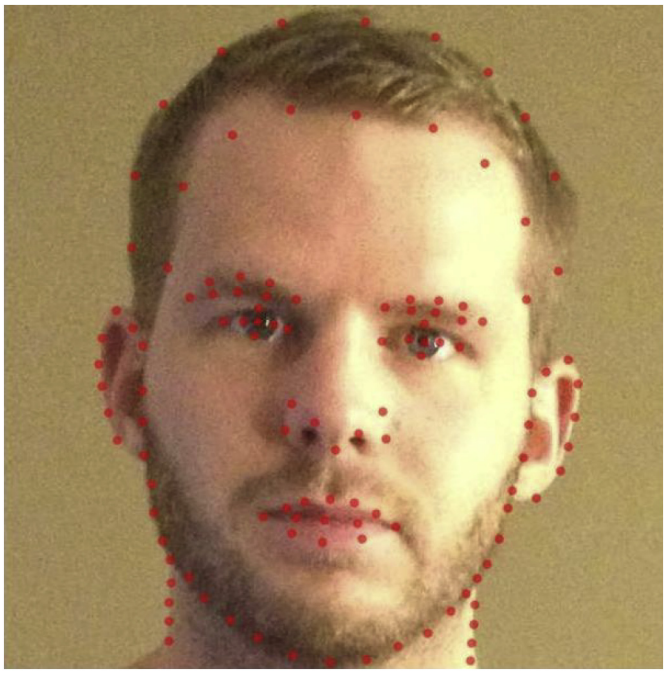

**Fig. 12.** An example of the detected 119 landmarks on a face image.

**Table 17**
MAE of estimated BMIs from 119 landmarks and 68 landmarks.

| | Black male | Black female | White male | White female |
|---|---|---|---|---|
| 119 landmarks | 2.85 | 3.97 | 2.82 | 3.15 |
| 68 landmarks | 2.63 | 3.65 | 2.68 | 3.12 |

## Declaration of competing interest

We have no conflict of interest to declare.

## Acknowledgment

The work was partly supported by an NSF grant IIS-1450620, an NSF-CITeR grant, and an WV-HEPC grant. The authors would also like to thank the editors and anonymous reviewers for their suggestions to improved the manuscript.

## References

[1] M. Arnold, M. Leitzmann, H. Freisling, F. Bray, I. Romieu, A. Renehan, I. Soerjomataram, Obesity and cancer: an update of the global impact, Cancer Epidemiol. 41 (2016) 8–15.
[2] A.G. Renehan, M. Tyson, M. Egger, R.F. Heller, M. Zwahlen, Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies, Lancet 371 (9612) (2008) 569–578.
[3] R. Wolk, P. Berger, R.J. Lennon, E.S. Brilakis, V.K. Somers, Body mass index, Circulation 108 (18) (2003) 2206–2211.
[4] J.B. Meigs, P.W. Wilson, C.S. Fox, R.S. Vasan, D.M. Nathan, L.M. Sullivan, R.B. Dagostino, Body mass index, metabolic syndrome, and risk of type 2 diabetes or cardiovascular disease, J. Clin. Endocrinol. Metab. 91 (8) (2006) 2906–2912.
[5] K. Wolffhechel, A.C. Hahn, H. Jarmer, C.I. Fisher, B.C. Jones, L.M. DeBruine, Testing the utility of a data-driven approach for assessing BMI from face images, PLoS One 10 (10) (2015) e0140347.
[6] C. Mayer, S. Windhager, K. Schaefer, P. Mitteroecker, BMI and WHR are reflected in female facial shape and texture: a geometric morphometric image analysis, PloS One 12 (1) (2017) e0169336.
[7] L. Wen, G. Guo, A computational approach to body mass index prediction from face images, Image Vis. Comput. 31 (5) (2013) 392–400.
[8] E. Kocabey, M. Camurcu, F. Ofli, Y. Aytar, J. Marin, A. Torralba, I. Weber, Face-to-BMI: using computer vision to infer body mass index on social media, Eleventh International AAAI Conference on Web and Social Media, 2017.
[9] X. Wu, R. He, Z. Sun, T. Tan, A light CNN for deep face representation with noisy labels, IEEE Trans. Inf. Forensics Secur. 13 (11) (2018) 2884–2896.
[10] Y. Wen, K. Zhang, Z. Li, Y. Qiao, A discriminative feature learning approach for deep face recognition, European Conference on Computer Vision, 2016. pp. 499–515.

[11] J. Deng, J. Guo, N. Xue, S. Zafeiriou, Arcface: additive angular margin loss for deep face recognition,arXiv preprint arXiv:1801.07698. 2018.
[12] V. Coetzee, D.I. Perrett, I.D. Stephen, Facial adiposity: a cue to health? Perception 38 (11) (2009) 1700–1711.
[13] V. Coetzee, J. Chen, D.I. Perrett, I.D. Stephen, Deciphering faces: quantifiable visual cues to weight, Perception 39 (1) (2010) 51–61.
[14] D.D. Pham, J.-H. Do, B. Ku, H.J. Lee, H. Kim, J.Y. Kim, Body mass index and facial cues in Sasang typology for young and elderly persons, Evid. Based Complement. Alternat. Med. 2011 (2011).
[15] A.J. Henderson, I.J. Holzleitner, S.N. Talamas, D.I. Perrett, Perception of health from facial cues, Phil. Trans. R. Soc. B 371 (1693). (2016)20150380.
[16] K. Ricanek, T. Tesafaye, Morph: a longitudinal image database of normal adult age-progression, Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, IEEE. 2006, pp. 341–345.
[17] B.J. Lee, J.Y. Kim, Predicting visceral obesity based on facial characteristics, BMC Complement. Altern. Med. 14 (1) (2014) 248.
[18] M. Barr, G. Guo, S. Colby, M. Olfert, Detecting body mass index from a facial photograph in lifestyle intervention, Technologies 6 (3) (2018) 83.
[19] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, Deepface: closing the gap to human-level performance in face verification, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014. pp. 1701–1708.
[20] F. Schroff, D. Kalenichenko, J. Philbin, Facenet: a unified embedding for face recognition and clustering, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015. pp. 815–823.
[21] A. Babenko, A. Slesarev, A. Chigorin, V. Lempitsky, Neural codes for image retrieval, European Conference on Computer Vision, Springer. 2014, pp. 584–599.
[22] Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, Realtime multi-person 2D pose estimation using part affinity fields, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1, 2017. pp. 7.
[23] O.M. Parkhi, A. Vedaldi, A. Zisserman, et al. Deep face recognition, British Machine Vision Conference, 1, 2015. pp. 6.
[24] A. Dantcheva, F. Bremond, P. Bilinski, Show me your face and I will tell you your height, weight and body mass index, International Conference on Pattern Recognition, 2018.
[25] N.M. Nasrabadi, Pattern recognition and machine learning, J. Electron. Imaging 16 (4) (2007) 049901.
[26] K. Zhang, Z. Zhang, Z. Li, Y. Qiao, Joint face detection and alignment using multitask cascaded convolutional networks, IEEE Signal Process Lett. 23 (10) (2016) 1499–1503.
[27] B. Amos, B. Ludwiczuk, M. Satyanarayanan, Openface: a general-purpose face recognition library with mobile applications, 2016.CMU School of Computer Science.
[28] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, Caffe convolutional architecture for fast feature embedding, Proceedings of the 22nd ACM International Conference on Multimedia, 2014. pp. 675–678.
[29] H. Drucker, C.J. Burges, L. Kaufman, A.J. Smola, V. Vapnik, Support vector regression machines, Advances in Neural Information Processing Systems, 1997. pp. 155–161.
[30] G. Guo, G. Mu, Y. Fu, T.S. Huang, Human age estimation using bio-inspired features, IEEE Conference on Computer Vision and Pattern Recognition, IEEE. 2009, pp. 112–119.
[31] M. De Marsico, M. Nappi, D. Riccio, Measuring measures for face sample quality, Proceedings of the 3rd International ACM Workshop on Multimedia in Forensics and Intelligence, ACM. 2011, pp. 7–12.
[32] A. Bulat, G. Tzimiropoulos, How far are we from solving the 2D and 3D face alignment problem? (and a dataset of 230,000 3D facial landmarks), Proceedings of the IEEE International Conference on Computer Vision, 2017. pp. 1021–1030.