Estimation of Body Mass Index using Face Detection.

A Report Submitted

in Partial Fulfillment of the Requirements

for the Degree of

Bachelor of Technology

in

Information Technology

by

Ankit Jain (20178009)

Shakshi Agrawal (20175080)

Brahmchari (20178093)

to the

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

MOTILAL NEHRU NATIONAL INSTITUTE OF TECHNOLOGY

ALLAHABAD, PRAYAGRAJ

10 July, 2020

UNDERTAKING

I declare that the work presented in this report titled "Estimation of Body Mass Index using Face Detection", submitted to the Computer Science and Engineering Department, Motilal Nehru National Institute of Technology Allahabad, Prayagraj, for the award of the Bachelor of Technology degree in Computer Science & Engineering, is my original work. I have not plagiarized or submitted the same work for the award of any other degree. In case this undertaking is found incorrect, I accept that my degree may be unconditionally withdrawn.

July 10, 2020

Allahabad

U.P. (India)

Ankit Jain (20178009)

Shakshi Agrawal (20175080)

Brahmchari (20178093)

CERTIFICATE

Certified that the work contained in the report titled "Estimation of Body Mass Index using Face Detection", by *Ankit Jain, Shakshi Agrawal and Brahmachari*, has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Dr. Rupesh Kumar Dewang
Computer Science and Engineering Dept.
M.N.N.I.T Allahabad, Prayagraj

July 10, 2020

Preface

The main objective of any computer science student is to get as much practical knowledge as possible. Being able to have practical knowledge by developing a project is a lifetime experience. As practical knowledge is as important as theoretical knowledge, we are thankful for having this project.

Through the development of this project, we had a great experience of various strategies that were applied and their pros and cons in judging their applicability. This served as a stepping stone in my career in the computer science field ahead.

We are pleased to present this project. Proper care has been taken while organizing the project in order to make it comprehend and implement various software technologies simultaneously.

Acknowledgement

The satisfaction that accompanies the successful completion of this project would be incomplete without mention of people who made it possible, because without their support, encouragement and guidance, everything would have gone in vain. We consider ourselves highly privileged to express gratitude and respect towards all those who guided us through the completion of the project.

We convey thanks to our project guide ***Dr. Rupesh Kumar Dewang*** of Computer Science and Engineering Department for providing Encouragement constant support, and guidance which served as a great help in completing this project successfully.

Last but not the least, we wish to thank our **parents** for financing our studies in college as well for constantly boosting up our morale. Their personal sacrifice in providing this opportunity to learn Engineering is greatly acknowledged.

Contents

1. Introduction

Together with a person’s gender, age and race, their weight status is a publicly visible signal that can have profound influence on many aspects of their life. Most obviously, it can affect their health as having a larger BMI is linked to an increased risk of both cardio-vascular diseases and diabetes, though not necessarily in a straight-forward manner (Meigs and others 2006). However, other aspects of the burden imposed by obesity come in the form of “fat shaming” and other forms of “sizeism”. For example, obesity is related to a lower income and part of the reason seems to be due to weight-based discrimination.

The human face exhibits information pertaining to identity, a person’s disposition, demeanor, as well as to attributes such as gender, age and ethnicity. From the perspective of biometrics, emphasis has predominantly been placed on facial recognition.

More recently attributes or soft biometrics such as gender, age, height and weight have gained popularity due to their semantic interpretation, i.e., they can provide a description that can be readily understood by humans; for example, the description “young, female, tall”.

1.1. Motivation

Obesity is an important public health concern, and an understanding of the difficulty in reducing obesity in individuals despite the seemingly simple energy balance equation that underlies all weight gain remains elusive. Excessive weight has been associated with obesity, diabetes, and cardiovascular diseases.

Limited attention has been given to the connection between the human face and body characteristics such as body height and weight and even less so on the automatic extraction of such. Estimating body height, weight and the associated BMI is warranted for several reasons.

Firstly, height and weight are attributes frequently used in surveillance, forensics, as well as re-identification applications and image retrieval systems [1].

Secondly, height and weight are primary and obvious attributes used by humans to verbally describe a person often used in police reports, unlike traditional biometrics which may be insufficient, as this was argued, for example, by Klontz and Jain [2] in the case of the 2013 Boston bombings.

Thirdly, body weight and height have been proposed as soft biometric traits in automated biometric systems [3].

2. Related Work

Some of the works related to automated face-based estimation of BMI include a study by Wen and Guo [4], based on the MORPH-II dataset, which obtained mean absolute errors (MAEs) for BMI in the range from 2.65−4.29 for different ethnic categories. The study explored handcrafted features for BMI-estimation and specifically in the method the face was detected, normalized, and an active shape model was fitted, based on which, geometry and ratio features were extracted (cheekbone to jaw width, width to upper facial height ratio, perimeter to area ratio, eye size, lower face to face height ratio, face width to lower face height ratio and mean of eyebrow height), normalized and finally subjected to support vector regression.

In another study by A. Dantcheva, P. Bilinski, F. Bremond [5], they proposed a method for face-based estimation of height, weight and BMI based on the ResNet-50 architecture which resulted in promising correlation accuracies of up to $\rho = 0.78$ for female weight estimation and mean absolute errors of 2.3 for female BMI estimation.

3. Proposed Work

Motivated by the above, we propose here a method for face-based estimation of height, weight and BMI based on various regression algorithms.

Our BMI prediction system is composed of two stages: (i) extraction of face embeddings using a pre-trained face-net architecture and (ii) training regression models and finding out the best one.

3.1. Face Detection

There exists a large no of face detection algorithms, we used python face recognition library which is based on ResNet-34 from the ***Deep Residual Learning for Image Recognition*** paper by He et al [6], but with fewer layers and the number of filters reduced by half. It maps a “face” into a feature vector of 128-d which can comprise various features like:

a. Height of face (cm)
b. Width of face (cm)
c. Average color of face (R, G, B)
d. Width of lips (cm)
e. Height of nose (cm)

These detected faces are provided as input to regression models.

3.2. Training Regression Models

We trained various regression models like simple linear regression model, ridge linear regression model, random forest regressor and kernel ridge regression on 1026 images from the VIP_attribute dataset and calculated the goodness metric for each model.

Goodness metric consists of mean square error, variance score i.e., r2 score, mean absolute error and accuracy of the model. We compared models based on this criterion and used the best one for final prediction.

We approached the problem by observing performing of various loss functions: -

- ### Simple Linear Regression

  Given our simple linear equation y=wx+b, we can calculate MSE as:
  $MSE = 1/N \sum (y_i - (w.x_i + b))^2$

  - N is the total number of observations (data points)
  - $1/N \sum$ is the mean
  - y_i is the actual value of an observation and wx_i+b is our prediction

- ### Ridge Linear Regression

  It learns w, b using the same least squares criterion but adds a penalty for large variations in w parameters.

  $RSS_{RIDGE} = 1/N \sum (yi - (w.x_i + b))^2 + \lambda \sum w_j^2$

- ## Random Forest Regressor

Every decision tree has high variance, but when we combine all of them together in parallel then the resultant variance is low as each decision tree gets perfectly trained on that particular sample data and hence the output doesn't depend on one decision tree but multiple decision trees. In the case of a regression problem, the final output is the mean of all the outputs.

$$\text{Cost} = \text{sum}\ (y - \text{prediction})^2$$

The decision tree will start splitting by considering each feature in training data. The mean of responses of the training data inputs of particular group is considered as prediction for that group. The above function is applied to all data points and cost is calculated for all candidate splits. *Again, the split with lowest cost is chosen*.

We also tried adjusting the following set of hyperparameters:

- n_estimators = number of trees in the forest
- max_features = max number of features considered for splitting a node
- max_depth = max number of levels in each decision tree
- min_samples_split = min number of data points placed in a node before the node is split
- min_samples_leaf = min number of data points allowed in a leaf node
- bootstrap = method for sampling data points (with or without replacement)

The *hyper parameters* of random forest regressor are tuned with randomized search cv with 3-fold cross validation.

- ## Kernel Ridge Regression

Kernel ridge regression combines Ridge regression and classification (linear least squares with l2-norm regularization) with the kernel trick. It thus learns a linear function in the space induced by the respective kernel and the data.

4. Experimental Setup and Results Analysis

We implemented our approach on “Python 3 Google Compute Engine backend (GPU)”.

Libraries Used: -

- Numpy
- Pandas
- Face-Recognition
- Chart-studio
- Matplotlib
- Seaborn
- Sklearn
- Principal Component Analysis (PCA)
- Ipywidgets

4.1. Dataset

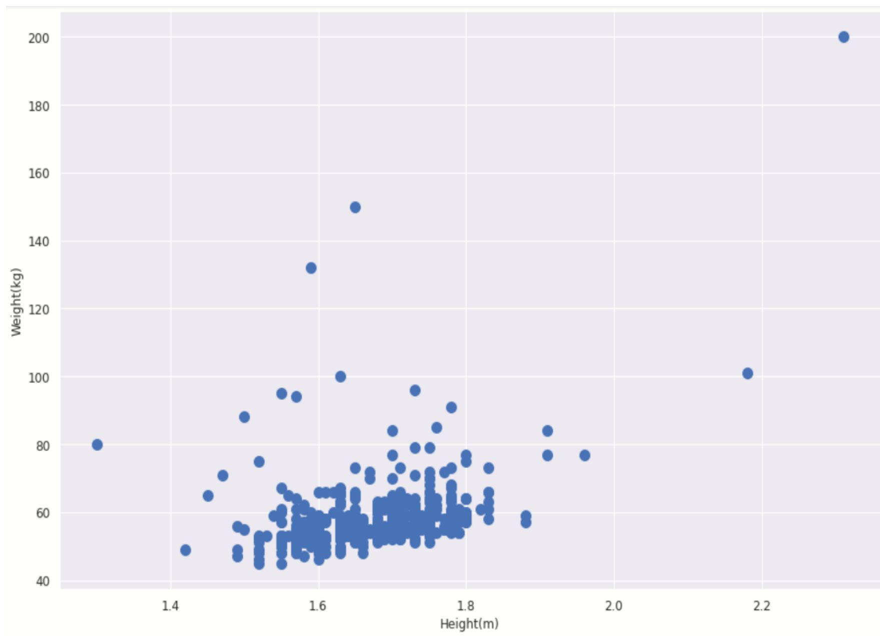
For this study, we used the ***VIP_attribute dataset***, consisting of 1026 subjects, specifically 513 female and 513 male celebrities (mainly actors, singers and athletes) collected from the WWW. The images are mainly frontal images. Covariates include illumination, expression, image quality and resolution.

Further challenges are beautification (photoshop) of the images, as well as the presence of makeup, plastic surgery, beard and mustache. We obtained annotations related to the subjects’ body weight and height available on websites such as www.celebheights.com, www.howtallis.org and celebsize.com, and proceeded to calculate the associated BMI.

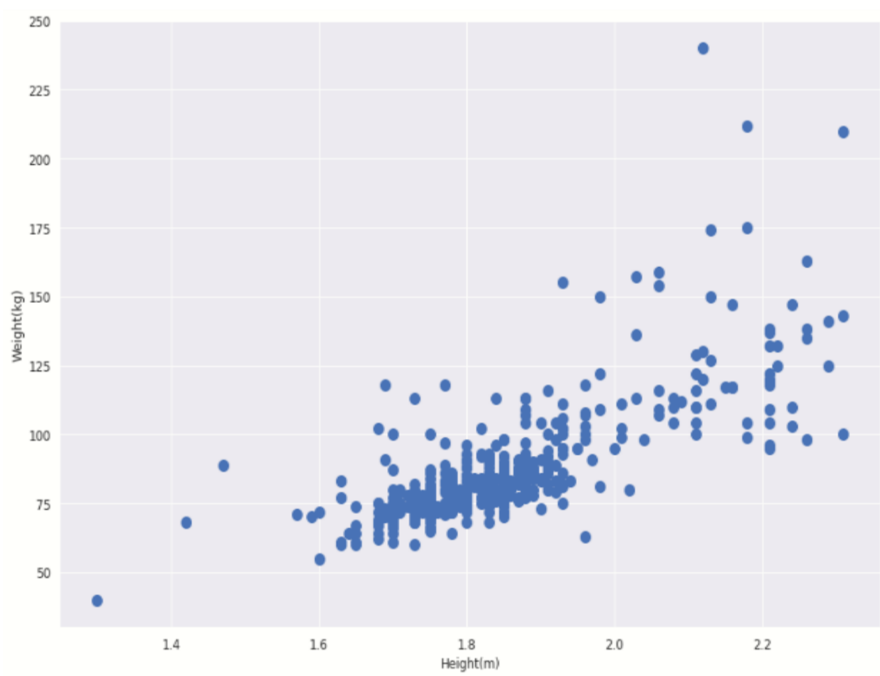
4.2. Relationship between parameters

- Weight vs Height

Females



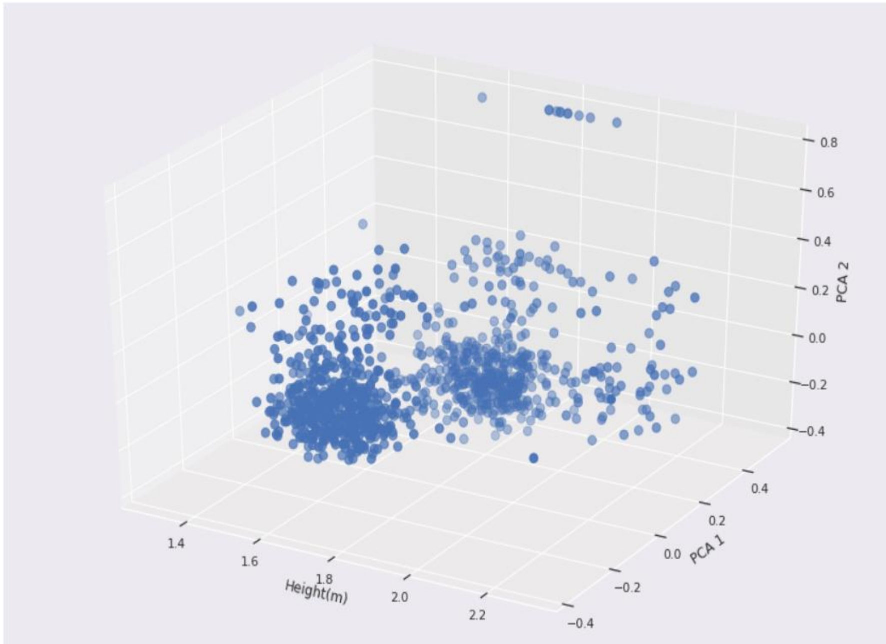
Males



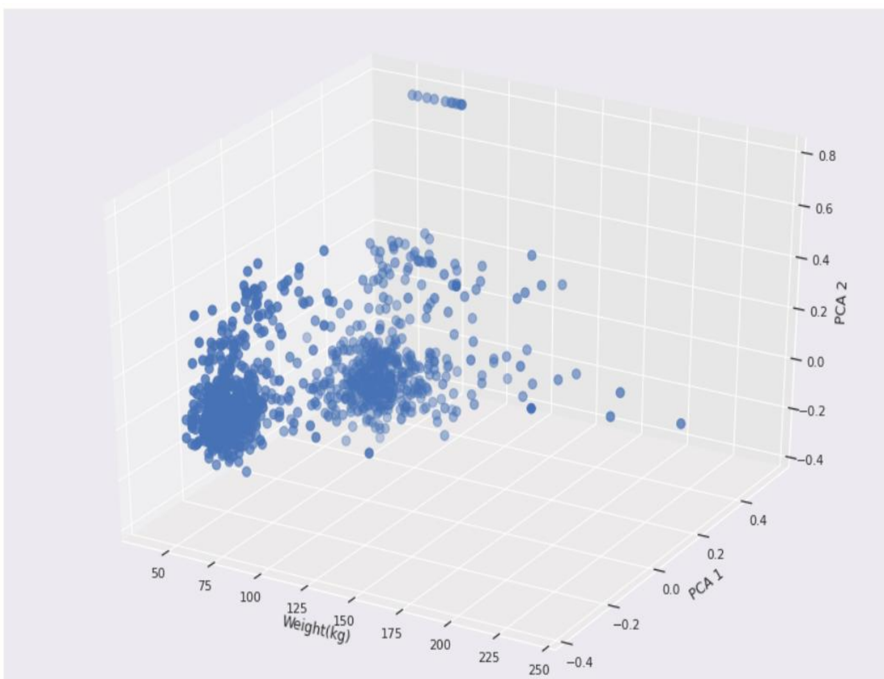
- Visualization of Face Embeddings

PCA was used for dimension reduction of 128-d feature vector to 2-d.

vs Height



vs Weight



4.3. Performance of Models

5. Data/Models	Height Model	Weight Model	BMI Model
Simple Linear Regression	Mean square error = 0.01 Variance score = 0.31 Average error = 0.0509 Accuracy = 90.22%	Mean square error = 13.57 Variance score = -193.80 Average error = 3.6763 Accuracy = 13.32%	Mean square error = 6.58 Variance score = -240.37 Average error = 2.5612 Accuracy = 18.02%
Ridge Regression	Mean square error = 0.00 Variance score = 0.36 Average error = 0.0479 Accuracy = 90.83%	Mean square error = 0.03 Variance score = 0.60 Average error = 0.1130 Accuracy = 97.39%	Mean square error = 0.01 Variance score = 0.47 Average error = 0.0839 Accuracy = 97.36%
Random Forest	Mean square error = 0.01 Variance score = 0.30 Average error = 0.0483 Accuracy = 90.78%	Mean square error = 0.04 Variance score = 0.48 Average error = 0.1306 Accuracy = 96.98%	Mean square error = 0.02 Variance score = 0.31 Average error = 0.0950 Accuracy = 97.01%

Data/Models	Height Model	Weight Model	BMI Model
Random Forest with tuned hyper parameter	Mean square error = 0.00 Variance score = 0.34 Average error = 0.0472 Accuracy = 90.96%	Mean square error = 0.03 Variance score = 0.59 Average error = 0.1124 Accuracy = 97.41%	Mean square error = 0.02 Variance score = 0.43 Average error = 0.0830 Accuracy = 97.40%
Kernel Ridge Regression	Mean square error = 0.01 Variance score = 0.29 Average error = 0.0531 Accuracy = 90.18%	Mean square error = 0.03 Variance score = 0.54 Average error = 0.1243 Accuracy = 97.10%	Mean square error = 0.02 Variance score = 0.32 Average error = 0.1042 Accuracy = 97.60%

It was found that Random Forest Regressor with tuned hyper parameters outperformed all the models in terms of the mean squared error and explained variance.

BMI can be experimentally verified from the predicted height and weight as: -

$$\text{BMI} = \text{weight [kg]} / \text{height}^2\ [\text{m}^2]$$

4.4 Some Examples

Height: 1.86m
Weight: 76.28kg
Bmi: 25.55kg/m^2

Height: 1.77m
Weight: 76.82kg
Bmi: 24.39kg/m^2

Height: 1.84m
Weight: 71.34kg
Bmi: 24.71kg/m^2

Height: 1.84m
Weight: 89.07kg
Bmi: 21.76kg/m^2

5. Conclusion and Future Work

This project presented a novel approach for estimating height, weight and BMI from single-shot facial images, based on regression models. Experiments conducted on the dataset, resulted in **absolute mean error** of 0.083 and **variance score** of 0.43. We did not observe a significant gender-bias in estimating height, weight and BMI.

However, more work is necessary in this regard. Future work will involve the additional study of age and ethnicity in order to improve utilization of facial appearance for height, weight and BMI estimation.

The height, weight and BMI estimator were motivated by the current need for self-diagnostic tools for remote healthcare, as well as for soft biometrics categorization in security applications.

References

1. J. Bekios-Calfa, J. M. Buenaposada, and L. Baumela, “Robust gender recognition by exploiting facial attributes dependencies,” PRL, 2013.

2. J. C. Klontz and A. K. Jain, “A case study on unconstrained facial recognition using the boston marathon bombings suspects,” Technical Report, no. MSU-CSE-13-4, 2013

3. A. Dantcheva, C. Velardo, A. D’angelo, and J.-L. Dugelay, “Bag of soft biometrics for person identification,” Multimedia Tools and Applications, vol. 51, no. 2, pp. 739–777, 2011.

4. L. Wen and G.-D. Guo, “A computational approach to body mass index prediction from face images,” Image and Vision Computing, vol. 31, no. 5, pp. 392–400, 2013.

5. A. Dantcheva, P. Bilinski, F. Bremond, "Show me your face and I will tell you your height, weight and body mass index," Proc. of 24th IAPR International Conference on Pattern Recognition (ICPR), (Beijing, China), August 2018.

6. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

7. “Using Computer Vision to Infer Body Mass Index on Social Media”, by Enes Kocabey, Mustafa Camurcu, Ferda Ofli, Yusuf Aytar, Javier Marin, Antonio Torralba, Ingmar Weber at ICWSM'17.