# BODY WEIGHT ESTIMATION USING 2D BODY IMAGE

**A Project Report**

***Submitted by***


**DINESH BALAJI [CB.EN.U4CSE17113]**
**SAI TRINADH [CB.EN.U4CSE17129]**
**PRASHANTH S [CB.EN.U4CSE17140]**
**ROHAN SONEJA [CB.EN.U4CSE17246]**


*Under the guidance of*


**Ms. Aarthi R.**
(Asst. Professor, Department of Computer Science & Engineering)


***in partial fulfillment for the award of the degree***

***of***

**BACHELOR OF TECHNOLOGY**

***in***

**COMPUTER SCIENCE & ENGINEERING**


**AMRITA VISHWA VIDYAPEETHAM**

Amrita Nagar PO, Coimbatore - 641 112, Tamilnadu


**Review #3 – November 2020**

# ABSTRACT

Body weight and associated BMI are key human features that are useful in a number of applications including health apps, remotely practising doctors, dynamic luggage allowance in airports and surveillance. The aim of this project is to estimate the body weight of a person given their digital image. This will be done by extracting facial features using computer vision on the images and machine learning techniques on the derived measurements.

Previous work includes the use of 2D face images to estimate the body weight. In addition to that, the scope is extended to include anthropometric measurements obtained from full body images to compare the results.

# TABLE OF CONTENTS

# Chapter 1

# INTRODUCTION

## 1.1 Problem Definition

To estimate the body weight of a person given a two-dimensional image of the person.

## 1.2 Motivation

There are several usecases for weight estimation where it may not be desirable to have body weight measured discretely and in other cases, the hardware to measure weight may not be available.
Observers seem to be able to glean body weight information from frontal views of a face (Schneider et al., 2012) and past research has replicated this using machine learning models. Additionally, it is inevitable that full body image should provide better insight as the anthropometric measurements can be used to further increase accuracy.

## 1.3 Applications

Following are some of the applications of body weight estimation using 2D full body images.

- Airport dynamic luggage allowance – Airports can use it to dynamically change luggage allowance policy based on how much weight passengers are taking up till a given point. Passengers will usually be hesitant to be weighed during check-in since it is a public space and also, hardware based weighing takes a few seconds which will cause propagation of delay in check-in queues. Image based weight determination will help in this case.

- Fitness recommendation on health apps – Health apps can provide better fitness recommendations based on image data provided by user instead of having to input weight manually.

- Surveillance – Extracting BMI is also pertinent in surveillance to add additional features to a photograph which would be useful in short descriptions of people such as "a short stout person", "tall lean person", "average height and fat" that would ease identification of a subject. While some soft biometrics like colour or length of hair can be changed quickly, it is not trivial to change visual body appearance.

- Remote BMI calculation by doctors – Measure weight in rural settings where hardware is not available such that a doctor practising remotely can have access to that metric too.

In the subsequent chapters, we will see how BMI estimation is done using facial and full body image data. We will also see the software requirements and datasets applicable for the project. We will finish off with the modularisation and month-wise plan.

# Chapter 2

# LITERATURE SURVEY

## 2.1 BMI Estimation Using Full Body Image Data

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. Jiang and Guo (2019) used images of entire front body by scraping data from a reddit page subreddit called preogress pics. To estimate BMI they used body contour and skeleton joints detected by CSJ detector and estimated the weight based on the full body landmarking of both the skeleton joints and hence the body contours. Several other features such as shoulder width, shoulder to waist ratio were were determined to aid in the process. Velardo and Dugelay (2010) studied the body weight directly from anthropometric data (body measurements) collected by National Health and Nutrition Examination Survey, CDC. A polynomial regression model was employed to analyze the anthropometric data. Velardo and Dugelay (2010) estimated the weight of a person within 4% error using 2D and 3D data extracted from a low-cost Kinect RGB-D camera output. Pfitzner et al. (2017) also used RGB-D camera data, this demonstrated a body weight estimation by volume extraction from RGB-D data. The presented algorithm provided an accuracy of 79% for a cumulative error of ±10%. The approach was tested with 110 patients from trauma room, focusing on body weight estimation for stroke patients. Due to uncertainties in volume estimation this approach had outliers up to 32%. Compared to a physician's estimation this approach is already more suitable for drug dosing. Also anthropometric features from depth data can be found in related work: Santner et al. (2009) estimated human body volume in clinical environment by eight stereo cameras around a stretcher and bioelectrical impedance analysis. With a 3D reconstruction the volume of a frontal surface can be calculated towards the medical stretcher which has been modeled as a plane. Nguyen et al. (2014) developed a method to predict body weight by a side view feature and a support vector regression model. Separating datasets by gender their approach reached an average error of 4.62 kg for females and 5.59 kg for males. Finally, they compared the body weight estimation by the algorithm in contrast to visual estimation.

## 2.2 BMI Estimation Using Facial Image Data

Instead of estimating body weight or BMI from body images, some work analyzed body weight or BMI from face images. Wen and Guo (2013) first proposed a computational method for BMI prediction from face images 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. They also analyzed the correlations between facial features and BMI values. An Active Shape Model is used to extract facial features which are used to predict BMI using various regression techniques. In Barr et al. (2018) Facial landmarking was used to figure out adiposity (facial fattness) which positively correlates to the weight of the person. This method is less acurate in exteme underweight and obese cases though. An SVR regression model was used. Windhager et al. (2017) showed that Shape of the face has direct correlation with several body characteristics such as height and weight and can be determined by facial landmarking and spatial scaling, A total of 71 landmarks and semi landmarks were digitized to capture facial shape. Regression and Geometric morphometric toolkit tool used to estimate facial fatness . Additionally as a feature set height measurement using anthropometer and saliva sample testing done apart from facial front photograph. Tai and Lin (2015) used where estimation of BMI is done using facial data on a regression model.

## 2.3 Summary

Most of the research in the past years has been to estimate BMI using facial images and in some cases body weight. While that is shown to have a high correlation to the BMI and body weight, it is prone to high error. To reduce the error, body image has been incorporated in the newer research although in some cases, at the cost of expensive hardware (thermal camera/depth sensor). Jiang and Guo (2019) showed that using a 2D body image dataset is viable to estimate BMI with reasonable accuracy.

With our project, we intend to combine both – facial and full body – image data and develop a model to estimate the body weight with improved accuracy.

Additional Papers: Haritosh, Gupta, Chahal, Misra, and Chandra (2019); Dantcheva, Bremond, and Bilinski (2018); Kocabey, Camurcu, Ofli, Aytar, Marín, Torralba, and Weber (2017); Bolukbaş, Başaran, and Kamaşak (2019); Nahavandi, Abobakr, Haggag, Hossny, Nahavandi, and Filippidis (2017)

## 2.4 Datasets

| S. N. | Name | Source | Size | Features | Remarks |
|---|---|---|---|---|---|
| 1. | Visual-Body-to-BMI Dataset | Received on request, Min Jiang | 2950 (x2) | body image pair (before and after weight change), weight, height, sex | full body image pairs scraped from ProgressPics subreddit |
| 2. | VIP-Attribute Dataset | Received on request, Antitza Dantcheva | 1026 | facial image, sex, height, weight | privately generated dataset containing facial images of celebrities |
| 3. | Illinois Department of Corrections (IDOC) Mugshots | Downloaded from Kaggle | 68492 | image pair (front and side mugshot), date of birth, weight, sex, height | front-facing and side-facing images of unique prisoners from the USA |
| 4. | Reddit-HWBMI Dataset | Received on request, Ankur Haritosh | 982 | facial image, age, sex, weight, height | facial images scraped from ProgressPics subreddit |

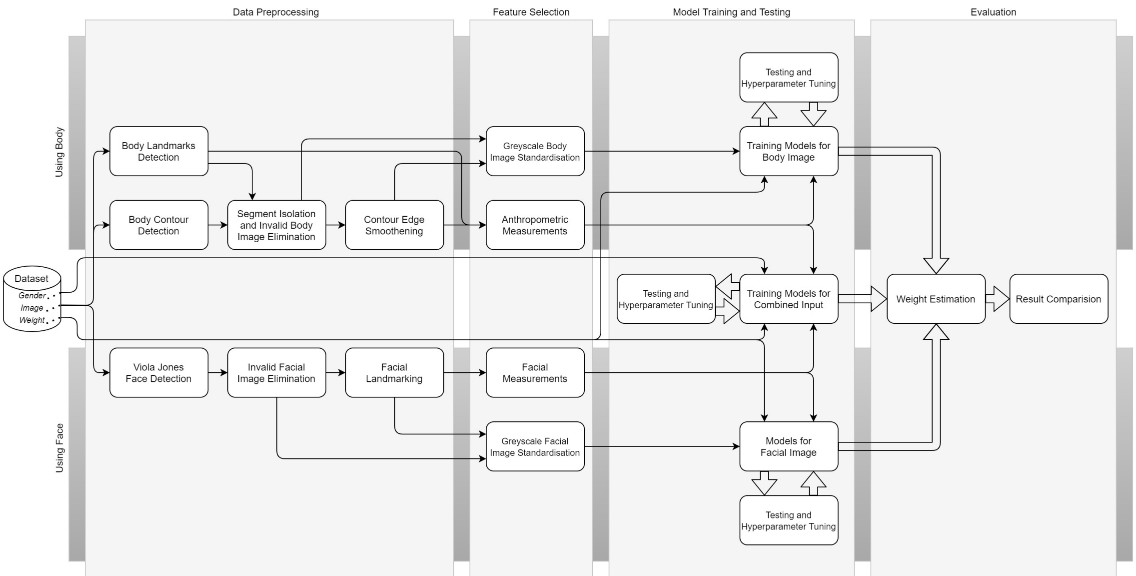
## 2.5 Software Requirements

- Python – language of choice for implementing core part
- OpenCV – image pre-processing and feature extraction
- Dlib – feature extraction
- MediaPipe – body landmarks extraction
- TensorFlow – implementing ML model
- Azure – deploying ML model API
- ReactJS – front-end for web-app

# Chapter 3

## PROPOSED SYSTEM

## 3.1 Architecture Diagram




(zoom in to view diagram)

## 3.2 Description of Modules

### 3.2.1 Dataset

Here, we use a combination of four diverse datasets which increase the robustness of the model. They include, a full body image dataset that has been scraped from reddit website under the subreddit r/ProgressPics that consist of publicly available images of the entire front of the body of a person, the Reddit-HWBMI Dataset consisting of only facial images, a facial image dataset consisting of celebrity photographs from popular websites and a novel dataset from kaggle consisting of facial photographs of prison inmates. All the images have the associated features of Gender and Body Weight which will be used to train a model to predict the body weight of a person given the photograph of the person.

Subsequently, the description will be split into two for each module that deal with the model that uses facial images and the model that uses body image separately.

### 3.2.2 Data Pre-processing

This module consists of all the stages that deal with the cleaning and preprocessing of the dataset that will aid in the stage of feature selection.

#### Using Body Image

This module consists of three submodules namely Body Contour detection, Invalid Body Image Elimination and Contour Edge Smoothening. This is done to first detect the contour of the body using an existing model that extracts the exact shape of the body within the image and hence isolates it from the environment thus aiding in measurements done on the body. Then certain images that haven't been properly diagnosed and have a fair bit of noise are eliminated. Then to make the body contour even sharper it is smoothened.

- Body Contour Detection
  The extraction of human contours is an essential part of the development of vision-based non-contact human body measurements and modelling systems. Here we colour code the human body shape alone from the image thus enabling ease in segregation from the background environment to aid measurement of various features that will serve as input features to the CNN model.

- Contour Edge Smoothening
  Due to several factors such as noise in the image, posture, background discrepancy and inaccuracy of the contour detection algorithm itself, this becomes a crucial step in ensuring the measurements derived from the image in the feature selection stage are as accurate as possible. This will in-turn positively affect the accuracy of the overall model.

#### Using Facial Image

This module consists of three submodules namely Viola-Jones Face detection, Invalid Facial Image Elimination and Facial Landmarking. First Viola-Jones Face detection is utilised to extract the face of the subject alone from the image, any imperfections or images without any face detected are then removed and the remaining extracted facial images are then landmarked with spots that highlight various facial landmarks such as eyebrow position, cheek position, eye position, jaw contour etc.

### 3.2.3 Feature Selection

This is a vital stage between the raw data preprocessing and the model training stage that involves using the cleaned and labelled image from the data preprocessing stage to extract features to be used in model training which are unique for the two models being built. For the purely facial image model, facial measurements are extracted from the facial landmarking performed in the preprocessing stage and for the full-body image-based model anthropometric measurements are derived from the labelled contour detected on the body image. This stage consists of another sub-module involving the Grayscale Image standardisation to be used in both the models on the respective datasets.

#### Using Body Image (Anthropometric Measurements)

From the labelled image with body contour and smoothening done key anthropometric measurements are extracted such as ratios of shoulder-length, height to waist length ratio etc. These become features that will be used for training the CNN model along with the image from with these features were extracted.

#### Using Facial Image (Facial Measurements)

From the labelled image with facial landmarking done several vital features such as 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 etc. These become features that will be used for training the CNN model along with the image from with these features were extracted.

### 3.2.4 Model Training and Testing

This module consists of all the stages involved in building the CNN model using the extracted features from the feature selection module along with the image used to extract those features. This module involves the building of two CNN models on each Face and Full body image dataset. It also consists of the building of an additional DL model using features used by both the CNN models. The models will be subjected to continuous testing and tuning of various hyperparameters to have the best fit on the dataset.

### 3.2.5 Evaluation

#### Weight Estimation

Using the prediction output from the models, weight of the person is estimated from the image given as input.

#### Result Comparision

Here, we compare the results obtained from the three models built:

- CNN model for facial image
- CNN model for full-body image
- DL model that combines both the datasets

and compare them using various metrics of supervised models and come to a conclusion on which model best suits the use case of Estimating the weight of a person using only the image of the person which has proven to be a useful feature in several fields of study.

## 3.3 Conclusion

- Through the research, we intend to use full body images as well as facial images to predict weight of person.
- We achieve this using various novel stages of preprocessing and feature selection that optimise the dataset thus enabling extraction of latent features of pertinence to help improve the decision making of the chosen model built.
- We then intend to build three models for making this prediction.
- Finally, we will compare the results by each of the models and see if we can combine them to improve overall accuracy.

# Chapter 4

# IMPLEMENTATION

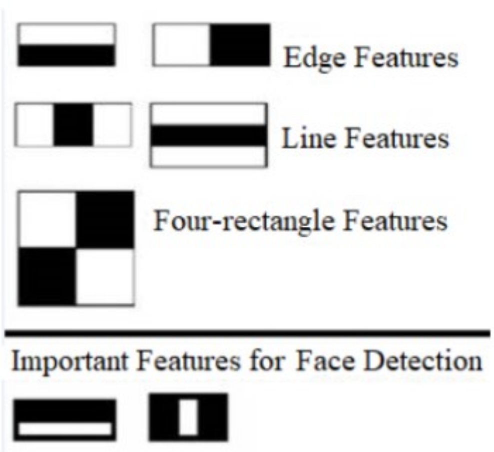
## 4.1 Pre-processing

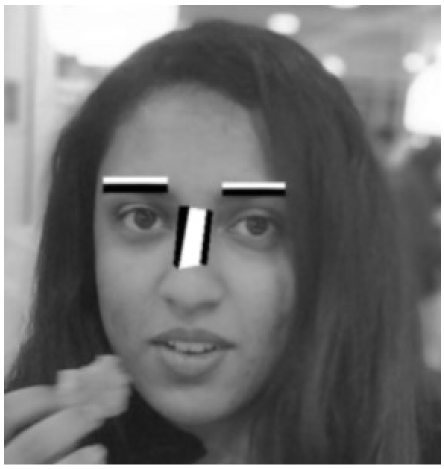
Since weight is to be estimated only from the images, all features from except the weight are dropped from the feature set. For our convinience, the gender has not been dropped as it is a realtively simple task to classify gender for real world images.

## 4.2 Feature Extraction

### 4.2.1 Viola-Jones Face Detection

This algorithm was used to detect the face in the image. It works by segregating the white and dark lines in a face. To do this, the face is first grey scaled. Then, the features below show a box with a light side and a dark side, which is how it is determined what the feature is. Sometimes one side will be lighter than the other, as in an edge of an eyebrow. Sometimes the middle portion may be shinier than the surrounding boxes, which can be interpreted as a nose.

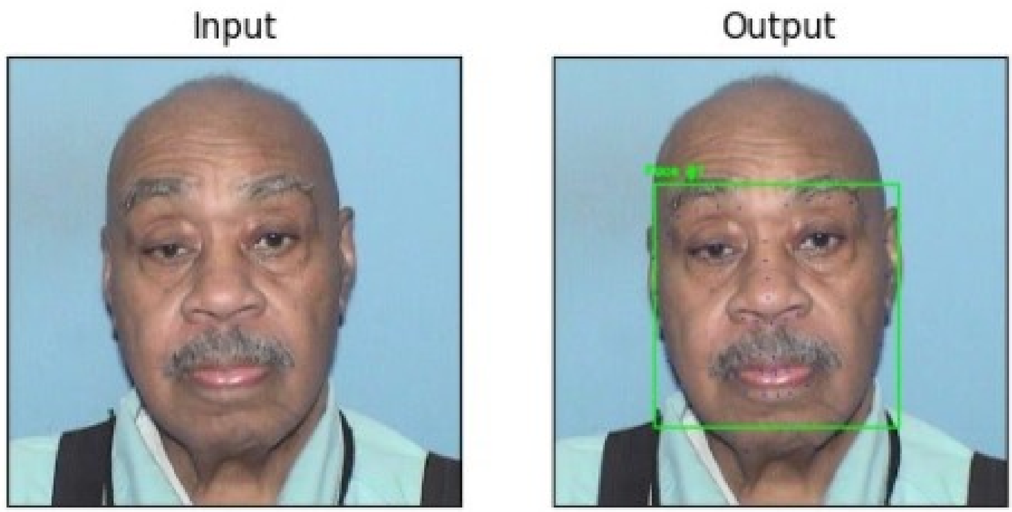


Then Adaptive boosting (Adaboost) algorithm is used to identify the face in the image. A bounding box is drawn around the image and this cropped image is the output of the algorithm.

### 4.2.2 Detecting Facial Landmarks Using dlib

Dlib is a landmark's facial detector with pre-trained models. It is used to estimate the location of 68 coordinates (x, y) that map the facial points on a person's face like image below.


### 4.2.3 Calculating Facial Ratios

From the landmarks on the face, we are able to calculate the length of various features in a face such as nose width, nose length, eye width, jaw width, face height, outer and inner lip widths, etc.

Since the length may not be the same depending on the distance from the subject ratio of the various extracted features to the jaw width was taken to remove the bias induced by the depth of the subject from the camera in the calculation of the length of the facial features.

### 4.2.4 Detecting Body Contour

As implemented by Zheng et al. (2015), we use a custom Python library based on Conditional Random Fields as Recurrent Neural Networks to isolate the body of the person from the image. Since this model is pre-trained to isolate several kinds of objects, the contour that is detected for the human is quite noisy and has been a limiting factor for the usability of anthropometric measurements.
Following image shows how the segmented output looks like for a sample image.

### 4.2.5 Anthropometric Measurements

We can use the body segment returned by the body contour detector in conjunction with body landmarks to make measurements related to body dimensions by using pixel counting operations. Following are some examples.

- Arm Width
  Mid-point of shoulder joint and elbow joint can be taken and number of pixels on either side can be calculated. Since orientation of arm may differ for different images, the count is multiplied by the sine of the angle made by the connecting line.

- Neck Width
  Mid-point of mid-shoulders-point and nose can be taken and number of pixels on either side can be calculated.

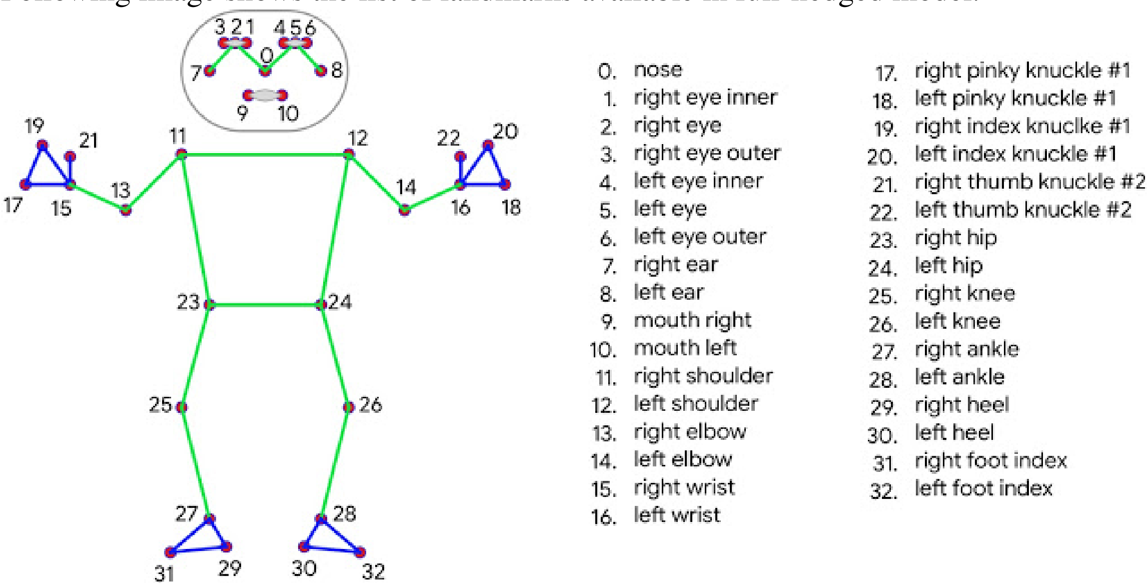
- Hip Width
  Mid-point of hip points can be taken and number of pixels on either side can be calculated.
- Waist Width
  Mid-point of hip points and shoulder points can be taken and number of pixels on either side can be calculated.

Due to the unreliability of the segmentation algorithm and the huge variance in the poses of bodies, at this stage, these features are not included for training the models.

### 4.2.6 Body Landmark Detection

As implemented by Bazarevsky et al. (2020), we use the pose detection framework available through Google's MediaPipe library to return coordinates as well as their respective confidence values for a given body image. Key coordinates include shoulder joints, hip joints, elbow joints, etc.

Following image shows the list of landmarks available in full-fledged model.




The API in its current iteration does not provide any landmarks below the hip joints. Crucially, BlazePose model is optimised for on-board mobile computation. Therefore, the features calculated from it are extremely fast to obtain and they can be used to make a rough estimation of weight if allowed by the usecase. Furthermore, the landmarks are used to assist the anthropometric measurement algorithm make required measurements on the body segment.

### 4.2.7 Final List of Features

Following facial distances (scaled down by jaw width) are considered.

- Left-eye width
- Right-eye width
- Face height
- Outer-lip width
- Inner-lip width
- Nose width
- Nose length
- Left-eyebrow width
- Right-eyebrow width

Following body distances (scaled down by inter-shoulder-joint distance) are considered.

- Left-shoulder to left-hip
- Right-shoulder to right-hip
- Left-hip to right-hip
- Left-shoulder to right-hip
- Right-shoulder to left-hip
- Left-shoulder to left-elbow
- Left-elbow to left-wrist
- Right-shoulder to right-elbow
- Right-elbow to right-wrist

Additionally, as mentioned earlier, gender is also incorporated in the feature list.

## 4.3 Models

A deep learning model and an XGboost regressor model was built for selected datasets.

### 4.3.1 Deep Learning

In prediction problems involving unstructured data (eg. images, text, etc.) sequential neural networks tend to outperform all other algorithms or frameworks. Hence, we have used this deep learning model. The library of choice is Keras which is built upon TensorFlow.

### 4.3.2 XGBoost Regressor

Since the problem of finding the weight, a continuous variable, requires regression also, XGBoost is a decision-tree-based ensemble machine learning algorithm that uses a gradient boosting framework. When it comes to small-to-medium structured/tabular data, decision tree-based algorithms perform at a much better level. It also has a perfect combination of software and hardware optimization techniques to yield superior results using less computing resources in the shortest amount of time.

# Chapter 5

# RESULTS AND DISCUSSION

To measure the performance of the models, a common metric of mean absolute error (MAE) was used, taken from the sklearn library.

## 5.1 Compiled Results

| **S. N.** | **Dataset** | **Model** | **MAE (kg)** |
|---|---|---|---|
| 1. | VIP-Attribute | Deep Learning (Sequential)<br>$\langle 8, 16, 8\rangle \times relu + \langle 1\rangle \times linear$ | 5.2 |
| 2. | VIP-Attribute | XGBoost Regressor<br>$nestimators = 40$ | 5.1 |
| 3. | IDOC Mugshots | Deep Learning (Sequential)<br>$\langle 8\rangle \times relu + \langle 0.2\rangle \times dropout + \langle 1\rangle \times linear$ | 13.2 |
| 4. | IDOC Mugshots | XGBoost Regressor<br>$nestimators = 50$ | 13.2 |
| 5. | Visual Body to BMI (w/o face) | Deep Learning (Sequential)<br>$\langle 64, 128, 256, 256, 64\rangle \times relu + \langle 1\rangle \times linear$ | 18.4 |
| 6. | Visual Body to BMI (w/o face) | XGBoost Regressor<br>$nestimators = 8$ | 18.5 |
| 7. | Visual Body to BMI | Deep Learning (Sequential)<br>$\langle 64, 128, 128, 64\rangle \times relu + \langle 1\rangle \times linear$ | 19.2 |
| 8. | Visual Body to BMI | XGBoost Regressor<br>$nestimators = 6$ | 17.9 |

## 5.2 Dataset-Wise Analysis

To get a deep understanding of the results, we need to analyse the measured metrics based on the datasets that they were trained on, as the distribution of them differs significantly.

### 5.2.1 VIP-Attribute Dataset

This was a very even dataset with respect to the weights of the celebrities whose images and body weights were used to train the model. Because of this, the MAE for body

weight in this model was about about 5 kg.
Because this dataset was quite standaradised, the results obtained were well within tolerable limits for error in weight.
XGBoost gave similar results, but faster.

### 5.2.2 IDOC Mugshot Dataset

This dataset containing face photographs of prison inmates was less standard and had more diversity in the ethnicity of the subject. On training the model on this dataset and tuning the parameters of the DL model the best MAE obtained was around around 13 kg.
This error value though larger than the VIP-Attribute dataset gives a more robust model due to the variety in the race and ethnicity of the subjects. This is also still a considerably better accurate model than any of the papers referred previously.
Again, XGBoost gave similar results, but faster.

### 5.2.3 Visual Body to BMI Dataset

This data set was extremely unstandardised with images vastly varying in body sizes, body poses, perspectives, distances from camera, lighting and many other factors, due to the nature of its source. Hence, understanstandably, this dataset performs the worst among all. That being said, at its current stage, the features that are provided to the model are quite primitive and once the full set of anthropometric measurements are incorporated into the model, it is expected to perform better.
In both the following cases, the XGBoost model, as usual, ran faster than the deep learning model.

#### Without Face

Without considering facial features, the deep learning model has performed slightly better at a MAE of about 18.4 kg as opposed to 18.5 for the XGBoost model.

### Including Face

Once facial features are also considered, the XGBoost model performs much better than the deep learning model with a MAE of about 17.9 kg as opposed to 19.2 kg for the latter. While this is still bad, it is crucial to understand that the facial features that are detected are heavily approximated due to the fact that they are so small compared to the rest of the body, and with limitations of resolution, error is introduced. Nonetheless, if higher resolution images are available for the model, it should surely perform better.

# Chapter 6

# CONCLUSION

By obtaining a mean absolute error of about 5 kg in finding the weight of a subject from face image and best case of MAE as 18 kg from full body image, it can be safely said that this method of weight estimation is a viable alternative to find the weight of a person in situations where the actual weight of the person cannot be physically measured. Therefore, it can hence be useful to have a rough estimate of a person's weight in several use cases such as in airports to detect the approximate total weight of all the passengers in a flight, forensics to identify people by adding more features than just an image such as a fat man, thin man, etc. and also for remotely practicising physicians to measure such a key human metric.

To summarise:

- Weight of a person can be estimated with a very small error range given a well lit standard facial image of the person.

- In cases of reduced quality of facial image, error increases dramatically and based on the situational tolerance, the predicted weight may be used.

- In case whole body image is used to predict weight, initial estimates suggest that the weights predicted have a rather large error range, but the correlation between image and body weight is clearly defined.

# Chapter 7

## CHALLENGES AND FUTURE DIRECTION

- Dataset quality
  Since the Visual Body to BMI dataset is scraped from internet, images are extremely varying and lot of them have low resolution, resulting in relatively bad MAE.

- Body contour detection algorithm
  The contour that is made is not reliable for aforementioned dataset hence making anthropometric measurements extremely challenging.

- CNN Model
  Intend to develop CNN model once the segmentation algorithm is potentially improved or if better dataset is available.

- BlazePose
  Currently, API provided by MediaPipe only provides landmarks for upper body. MAE might improve when it is updated to provide other landmarks.

- Anthropometric features
  Once the BCD and BlazePose are improved, full set of anthropometric measurements can be extracted better.

# REFERENCES


1. Barr, M., Guo, G., Colby, S., and Olfert, M. (2018). “Detecting body mass index from a facial photograph in lifestyle intervention.” *Technologies*, 6(3), 83.

2. Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., and Grundmann, M. (2020). “Blazepose: On-device real-time body pose tracking.” *arXiv preprint arXiv:2006.10204*.

3. Bolukbaş, G., Başaran, E., and Kamaşak, M. E. (2019). “Bmi prediction from face images.” *2019 27th Signal Processing and Communications Applications Conference (SIU)*. 1–4.

4. Dantcheva, A., Bremond, F., and Bilinski, P. (2018). “Show me your face and i will tell you your height, weight and body mass index.” *2018 24th International Conference on Pattern Recognition (ICPR)*. 3555–3560.

5. Haritosh, A., Gupta, A., Chahal, E. S., Misra, A., and Chandra, S. (2019). “A novel method to estimate height, weight and body mass index from face images.” *2019 Twelfth International Conference on Contemporary Computing (IC3)*. 1–6.

6. Jiang, M. and Guo, G. (2019). “Body weight analysis from human body images.” *IEEE Transactions on Information Forensics and Security*, 14(10), 2676–2688.

7. Kocabey, E., Camurcu, M., Ofli, F., Aytar, Y., Marín, J., Torralba, A., and Weber, I. (2017). “Face-to-bmi: Using computer vision to infer body mass index on social media.

8. Nahavandi, D., Abobakr, A., Haggag, H., Hossny, M., Nahavandi, S., and Filippidis, D. (2017). “A skeleton-free kinect system for body mass index assessment using deep neural networks.” *2017 IEEE International Systems Engineering Symposium (ISSE)*. 1–6.

9. Nguyen, T. V., Feng, J., and Yan, S. (2014). “Seeing human weight from a single rgb-d image.” *Journal of Computer Science and Technology*, 29(5), 777–784.

10. Pfitzner, C., May, S., and Nüchter, A. (2017). “Evaluation of features from rgb-d data for human body weight estimation.” *IFAC-PapersOnLine*, 50(1), 10148 – 10153. 20th IFAC World Congress.

11. Santner, K., Rüther, M., Bischof, H., Skrabal, F., and Pichler, G. (2009). “Human body volume estimation in a clinical environment.

12. Schneider, T. M., Hecht, H., and Carbon, C.-C. (2012). “Judging body weight from faces: The height—weight illusion.” *Perception*, 41(1), 121–124. PMID: 22611670.

13. Tai, C. and Lin, D. (2015). “A framework for healthcare everywhere: Bmi prediction using kinect and data mining techniques on mobiles.” *2015 16th IEEE International Conference on Mobile Data Management*, Vol. 2. 126–129.

14. Velardo, C. and Dugelay, J.-L. (2010). "Weight estimation from visual body appearance. 1 – 6.

15. Wen, L. and Guo, G. (2013). "A computational approach to body mass index prediction from face images." *Image and Vision Computing*, 31(5), 392 – 400.

16. Windhager, S., Bookstein, F. L., Millesi, E., Wallner, B., and Schaefer, K. (2017). "Patterns of correlation of facial shape with physiological measurements are more integrated than patterns of correlation with ratings." *Scientific Reports*, 7(1), 45340.

17. Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., Huang, C., and Torr, P. H. S. (2015). "Conditional random fields as recurrent neural networks." *International Conference on Computer Vision (ICCV).*