

# ELEN4002: Digital Estimation of Body Mass Index

Darrion Singh (1056673)

*Abstract—*

## I. INTRODUCTION

The Body Mass Index (BMI) is a screening tool used to indicate whether one's body mass is healthy for their corresponding height [1]. Whilst not an accurate diagnostic tool, the BMI provides a healthy body mass range for a particular height as it takes natural body shape variations into account [1]. Unfortunately, the accurate collection of one's height and mass is time-consuming, especially when considering the time required to collect this data for large amounts of people. Furthermore, the distance between rural areas and clinics also inhibit the habitants of the area from regular medical checkups, where a medical practitioner might require to measure and track one's height and mass. A possible solution to both the above mentioned problems would be the creation of a tool that would provide a medical practitioner with this information without the need for the patient to be physically present. The following report contains a design and implementation of a program that estimates BMI from a photograph, using a mixture of Computer Vision and Machine Learning. The School of Electrical and Information Engineering (EIE) have collaborated with the Peri-Natal HIV Research Unit (PHRU) for this project.

## II. BACKGROUND

### A. Project Concept

The aim of this project is to estimate the BMI of a person from their photograph. This is to be achieved by creating a program which accepts a front and side photograph of a person. The program will extract body shape data from the photographs with the use of Computer Vision, and predict one's BMI based on their body shape data using Machine Learning.

The purpose of such a tool has vast applications in the medical field, whether it be used as a time-efficient means of tracking the progress of a disease/disorder that affects one's mass, or for a simple indication of one's own health. Furthermore, a possible extension of this tool would be using it to classify which category of BMI one falls into: Underweight ( $< 18.5$ ), Healthy ( $18.5 - 25$ ), Overweight ( $25 - 30$ ), or Obese ( $> 30$ ).

The success of this project would result in the production of a BMI estimator which precludes the need for a subject to remove their clothes, as well as travel to a clinic for BMI measurement. In large rural clinics, where large patient volumes combined with time taken to undress and redress for accurate height and mass measurements, this digital solution may result in a significant reduction in the time taken for

consultations. Furthermore, this solution would not be affected by measurement errors resulting from incorrect calibration of machinery.

### B. Assumptions & Constraints

- As to provide the least chance of a false human body outline being detected, we assume that any person being photographed is staged in front of a uniform background, such as a wall whose colour significantly contrasts with that of the patient and their clothes. It is further assumed that all photographs used to train the machine learning algorithm should abide by the same assumption.
- In order to provide a measurement reference, the person in each photograph should be standing next to a reference object of known dimensions, such as a piece of standard A4 paper. The colour and dimensions of the piece of paper should be the same throughout the training dataset.
- This reference object is present in the photographs during any live testing or usage of the application.
- All subjects in the photographs stand alongside and in line with the reference object, i.e. if the reference object is stuck to a wall, the subject is standing against the wall.
- The distance between the camera and the subjects being photographed is constant and far enough away to capture the subject's full body photograph.
- Subjects have a maximum height of two meters.
- The time frame for this project is six weeks, with a total budget of R1200.

### C. Success Criteria

- The program should provide an average accuracy of 75% for it to be classified as a potential viable solution.
- To be considered a working prototype, it should provide an average accuracy of 85%.
- To be considered a completed product, it should provide a minimum average accuracy of 95% across all age, gender, and mass groups.
- BMI prediction takes place without the need for one to remove their clothes and thus operate with pictures of clothed people.
- The final solution is a robust desktop application that can be easily used by a medical practitioner.
- The solution should be camera-independent and resolution-independent.

### D. Optional Criteria

- The software and its trained neural network models are exported to either a mobile application or desktop executable. In this way, one could select a photograph

from their cellphone/desktop gallery and the program could perform estimation on the chosen picture.

- The desktop executable can receive pictures from a connected camera.
- A mobile application would access the cellphone's camera and perform a BMI estimation from a photograph taken in-app. The prediction could either be shown to the user or sent to the medical practitioner directly.
- Provided that there is sufficient training data, training two separate models for male and female data points would decrease sex bias in the models. This would be to further improve accuracy by accounting for natural variations of body shape between men and women, with the hope of seeing similar statistical variations in BMI across age and sex as seen in Ref. [2].

### III. LITERATURE REVIEW

Before deciding on the proposed solution described in the succeeding sections, various academic literature regarding image processing and machine learning were consulted. The following provides a summary of findings that were considered when deciding upon the final solution.

#### A. Weight and Diameter Estimation

Comert et al. aims to estimate the mass of an apple through image processing and machine learning techniques [3]. The original apple image is seen as the top left image of Figure 1. The process involves using the Prewitt edge extraction algorithm in order to obtain a binary image, seen as the top right image of Figure 1. The binary image goes through Closing, filtering and Opening.

Closing is the process of performing Dilation followed by Erosion on an image, whilst Opening is the process of performing Erosion followed by Dilation on an image [4], [5]. Dilation is a morphological operation whereby the value (intensity) of a pixel is increased to that of the pixel in its neighbourhood with the highest intensity [6]. Erosion is a morphological operation whereby the value of a pixel is decreased to that of the pixel in its neighbourhood with the lowest intensity [6].

This is done in order to improve the image quality for feature extraction, since Closing an image improves the prevalence of contours, whilst Opening an image improves the uniformity of the spaces between contours. The bottom image of Figure 1 shows the binary image after the above operations were performed on it, successfully segmenting the pixels of the apple as white whilst the background is black.

Feature extraction is carried out such that the area and diameter of the apple are extracted seen in Figure 2. This feature extraction would then be performed on a set of apple photographs whose masses were known. Thereafter, the diameter and area of the apples were related to their recorded masses using a linear regression model. This method reported a 96.5% accuracy of estimation [3].

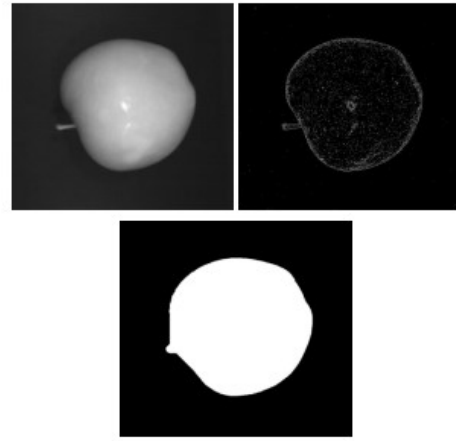


Fig. 1. Segmentation techniques to isolate the apple in the image [3].

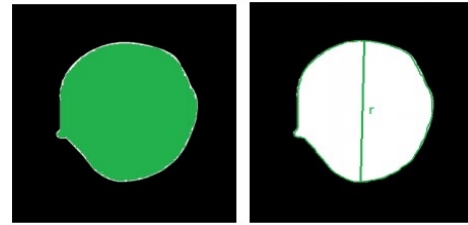


Fig. 2. Apple Area and Diameter Extraction [3].

#### B. Weight Estimation Using Image Analysis and Statistical Modelling

Similar to the requirements of the BMI system, Kollis et al. discusses a robust image analysis based mass estimation of pigs in large scale livestock farming [7]. The software of choice for image processing is MATLAB in which the Image Processing Toolbox V4.1 is used. Object detection occurred through comparison between the camera environment without the pig and with the pig. This results in a binary image that undergoes segmentation, filtering and opening to get a clear image for feature extraction.

Feature extraction in this paper is particularly relevant to this report as the body shape of the pigs is more complex than the features of apples described by Ref. [3] in the previous section. The feature extraction method is comprised of the skeletonisation function in MATLAB which generates a skeletal structure of the object in question.

Figure 3 shows the skeletonisation process for the image of the pig.

Using this process the spine (main branch) is used to estimate the length of the pig by counting the pixels which comprise the spine. Thereafter, the estimated spine length and recorded pig masses were related using a linear regression model.

#### C. Human Weight from a RGB-D Image

Nguyen et al. explores human mass prediction through a single RGB-D image [8]. The camera in question uses a depth

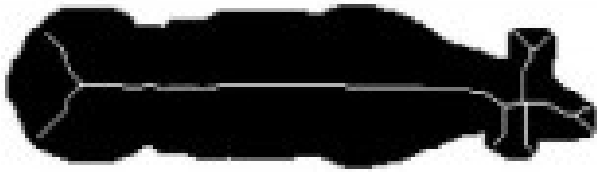


Fig. 3. Skeletonisation function implemented on pig binary image [7].

sensor to create a separate depth image. This extra image proves to be useful in object detection due to depth allowing an accurate trace of the human body shape with the aid of the 2D RGB image.

This silhouette is passed through similar image processing methods used in the aforementioned papers. However, mass prediction is different in this scenario as the author used various mathematical relationships relating body width, height and area with Support Vector Regression to predict the mass of the person through the extracted features.

#### D. Types of Neural Networks

Mehta discusses the various types of neural networks and what applications in which they prove to be useful [9]. The summary is as follows: Feed-Forward neural networks were useful in situations with minimal data-sets, as well as data-sets that have a substantial amount of noise. The multi-layer perceptron is used when the training data-set is not linearly separable. Networks that make use of back propagation need large data-sets and is generally used in when the network needs to recall or retain information. Convolutional neural networks are used in applications such as object detection and image classification.

#### E. Estimation of BMI from Photographs Using Deep Convolutional Neural Networks

Pantanowitz et al. have conducted a similar undertaking to that described in this report. The paper discusses the estimation of BMI from a photograph using a Deep Convolutional Neural Network [10]. The analysis was performed on a dataset of 161 photographs that were used to create a set of silhouette images via manual marking as well as automated image processing techniques. The silhouette images were standardised to a particular pixel width without changing aspect ratio, and were thereafter padded appropriately and cropped to a square bounding box. The silhouette images were augmented via minimal rotations and width shifts, with the author stating that the model could only train sufficiently with augmentation. A rather complicated model consisting of seventeen layers was used, details of which can be found in Ref. [10]. The results showed the BMI predictions having a correlation of 0.96 with the actual BMI values.

## IV. TRAINING DATA COLLECTION

### A. Ethics Approval

### B. Collaboration with the PHRU

The PHRU have provided clinical research staff

### C. Data Types

The data to be collected are as follows:

- A front and side photograph of the subject. Both photographs should have the reference object in the top left corner of the image. As mentioned above, the background of the image should be a uniform surface such as a wall. This is to ensure that the image pre-processing techniques described in the succeeding sections of this report can be carried out without error.
- The height and mass of each person. These must be captured simultaneously and attached to the front and side photographs. Data survey applications such as *REDCap* may be useful in achieving this.
- (Optional) Sex of person in each photograph. Whilst this characteristic is not an input into the system, it should be recorded for statistical purposes for later quantification of algorithmic bias. Furthermore, to achieve the optional criteria discussed above, data points must be labelled as male/female to train separate models.
- (Optional) Race of person in each photograph. This characteristic could also be recorded for statistical purposes. Furthermore, it opens an avenue for further research into how ethnicity can be used to narrow BMI distribution and hence provide more accurate BMI estimations.
- (Optional) Age of person in each photograph. This characteristic, despite only being recorded for statistical purposes, is by far the largest contributor to the distribution of BMI, as seen in Ref. [2]. Knowing the distribution of the age of the subjects will provide a range of ages where the program is likely to provide the most accurate estimation.

With the ultimate aim in mind for this program to be a universal BMI estimator, it is encouraged that the data-set of subjects who volunteer to be photographed be spread in terms of sex, ethnicity and age.

### D. Obtaining Data

It is possible to find such images of people on the internet, however it is unlikely for these images to have the corresponding height and mass on record. The chosen method for obtaining the data is to illicit the help of a medical research group who have access to large amounts of people, along with properly calibrated height and mass measuring equipment.

It should be noted that obtaining data in this manner requires prior clearance from a recognised ethics committee, as well as the permission of those in charge of the proposed sites such as research clinics. As with most neural network related projects, it is encouraged for a large data-set to be collected, and it was decided that a reasonable minimum number of subjects would be 100.

## V. SYSTEM DESIGN

## VI. METHODOLOGY

## VII. SOURCES OF ERROR

## VIII. FINDINGS & ANALYSIS

## IX. ADDITIONAL FEATURES

## X. RECOMMENDATIONS

## XI. CONCLUSION

## XII. ACKNOWLEDGEMENTS

## REFERENCES

- [1] N. H. S. UK, "What is the body mass index (bmi)?" Jul 2016. [Online]. Available: <https://www.nhs.uk/common-health-questions/lifestyle/what-is-the-body-mass-index-bmi/>
- [2] A. P. De Silva, S. H. P. De Silva, R. Haniffa, I. K. Liyanage, K. S. A. Jayasinghe, P. Katulanda, C. N. Wijeratne, S. Wijeratne, and L. C. Rajapakse, "A cross sectional survey on social, cultural and economic determinants of obesity in a low middle income setting," *International journal for equity in health*, vol. 14, no. 1, p. 6, 2015.
- [3] O. CÖMERT, M. HEKİM, and K. ADEM, "Weight and diameter estimation using image processing and machine learning techniques on apple images," *International Journal of Engineering Research and Development*, vol. 9, p. 3, 2017.
- [4] R. Fisher, S. Perkins, A. Walker, and E. Wolfart, "Morphology - opening," 2003. [Online]. Available: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/open.htm>
- [5] R. Fisher, S. Perkins, A. Walker, and E. Wolfart, "Morphology - closing," 2003. [Online]. Available: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/close.htm>
- [6] I. The MathWorks, "Understanding dilation and erosion :: Morphological operations (image processing toolbox user's guide)," 2005. [Online]. Available: <http://matlab.izmiran.ru/help/toolbox/images/morph3.html>
- [7] K. Kollis, C. S. Phang, T. M. Banhazi, and S. J. Searle, "Weight estimation using image analysis and statistical modelling: A preliminary study," *Applied engineering in agriculture*, vol. 23, no. 1, pp. 91–96, 2007.
- [8] T. V. Nguyen, J. Feng, and S. Yan, "Seeing human weight from a single rgb-d image," *Journal of Computer Science and Technology*, vol. 29, no. 5, pp. 777–784, 2014.
- [9] A. Mehta, "A complete guide to types of neural networks," Jan 2019. [Online]. Available: <https://www.digitalvidya.com/blog/types-of-neural-networks/>
- [10] A. Pantanowitz, E. Cohen, P. Gradidge, N. Crowther, V. Aharonson, B. Rosman, and D. M. Rubin, "Estimation of body mass index from photographs using deep convolutional neural networks," Aug 2019. [Online]. Available: <http://arxiv.org/abs/1908.11694>