

# ELEN4002: Digital Estimation of Body Mass Index

Darrion Singh (1056673)

**Abstract**—The following reports the findings of a study used to design and implement an application used to estimate Body Mass Index from a photograph.

## 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 inhabitants 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 presents a study of a potential technique to estimate BMI from a photograph. It also reports on the design and implementation of a program that estimates BMI from a photograph using these techniques, which include a mixture of Computer Vision and Machine Learning. The School of Electrical and Information Engineering (EIE) have collaborated with the Perinatal 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 (RO) 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 RO 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 RO, i.e. if the RO 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

#### A. Weight and Diameter Estimation of Apples

Comert et al. aims to estimate the mass of an apple through image processing and machine learning techniques [3]. The process involves performing edge detection on the original image using the Prewitt edge extraction algorithm, which creates a binary image containing the edges in the original image. The binary image then undergoes 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 above operations successfully segmented the image such that the pixels of the apple is white whilst the background is black.

Feature extraction is carried out such that the area and diameter of the apple are extracted. 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].

#### 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.

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. Types of Neural Networks

Mehta discusses the various types of neural networks and what applications in which they prove to be useful [8]. 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.

#### D. 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 [9]. 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 17 layers was used, details of which can be found in Ref. [9]. The results showed the BMI predictions having a correlation of 0.96 with the actual BMI values.

### IV. SYSTEM DESIGN

The proposed solution comprises of two subsystems, namely the Extraction layer and the Machine Learning layer. These subsystems have no dependency on each other and are to be built independently of one another. Figure 1 in Appendix 5 provides a system diagram showing the inputs and outputs of each system, as well as the components of each system. The Extraction layer comprises of the Object Detection, Image Segmentation and Input Extraction components, whilst the Machine Learning layer comprises of the Front View, Side View and Weighted Averaging components.

#### A. Extraction Layer

A front and side view photograph of a person with the aforementioned RO in the top left corner is the input into the system. For the sake of clarity and simplicity, the RO chosen is a standard A3 cardboard, and its chosen colour is black for the sake of high contrast with the assumed light coloured wall in the background. The Object Detection component will isolate the two objects in the photograph, namely the person and the RO, and return the bounding box for each. The width of the bounding box surrounding the RO is measured in pixels. Knowing its physical width allows us to calculate the pixels per meter ratio for that photograph. In other words, all objects that exist in line with

the RO will be represented by approximately the same number of pixels per meter as that object. This technique was inspired by Ref. [10].

For both the front and side view photographs the Image Segmentation component will receive the locations of the RO and the person in the form of their bounding boxes. Using the bounding boxes, both the RO's and person's pixel mask (all pixels belonging to an object) will be extracted for both the front and side photographs.

Using the pixels per meter found earlier the Input Extraction component performs feature extraction on the person's front and side pixel masks. Six features are extracted from each pixel mask:

- Their approximate height in meters.
- The approximate two dimensional area contained within their shape (silhouette) in squared meters.
- The maximum thickness of their thoracic region in meters.
- The maximum thickness of their abdominal region in meters.
- The maximum thickness of their pelvic region in meters.
- The maximum thickness of their upper leg region (thigh) in meters.

Note that from front images, the 'thickness' is hereafter referred to as 'width' whilst from the side images, 'thickness' is hereafter referred to as 'depth'. The four body regions mentioned above have been chosen strategically as to represent common areas of body fat deposit. In other words, this study explores the use of Computer Vision and Machine Learning to relate the common regions of body fat deposit to BMI by extracting information about these regions from photographs. The output of the Extraction layer is two vectors, namely the front view vector (FVV) and side view vector (SVV). Both FVV and SVV will contain the above mentioned features extracted from their respective photograph.

### B. Machine Learning Layer

All the outputs of the Extraction layer mentioned above are provided as inputs into the Machine Learning layer. The FVV and SVV are passed into the Front View and Side View components respectively, both of which are neural network models trained on a dataset of FVV's and SVV's, respectively. Both networks will have been trained to estimate BMI independently until an acceptable accuracy is achieved, or falling short of that, the best possible accuracy is achieved.

The estimated BMI from the Front View and Side View component are then used as inputs into a final neural network, the Error Correction / Compensation model, which performs a weighted averaging between the Front and Side BMI estimations. This network will have been trained to estimate BMI by deciding the contribution that the Front and Side view estimates should make towards a final BMI estimation with improved accuracy.

## V. TRAINING DATA COLLECTION

As mentioned above, the layer of the system that performs the BMI prediction makes extensive use of neural network models. Thus in order to train these models, a dataset containing photographs of people with their corresponding BMI was required.

### A. Ethics Approval

This study was approved by The University of the Witwatersrand Human Research Ethics Committee (Medical), Certificate Number M190539. The committee approved the anonymised collection of the height, mass, age, sex, and a single front and side-facing full-body photograph of 500 fully clothed participants.

### B. Collaboration with the PHRU

The PHRU provided clinical research staff who aided in the collection of the above data. Each person was issued an arbitrary subject number that was attached to each anonymised photograph. The dataset was collected at three sites:

- PHRU MDR Tuberculosis (TB) Research Clinic in Klerksdorp.
- PHRU TB Clinic in Chris Hani Baragwanath Academic Hospital in Soweto.
- The Chamber of Mines building at the University of the Witwatersrand.

Data was collected at the last site by the two principal investigators of the study.

### C. Data Point Composition

As mentioned above, the data collected was as follows:

- A front and side photograph of the subject. Both photographs had the RO in the top left corner of the image.
- The height and mass of each person. These were captured simultaneously and linked to the persons subject number using the *REDCap* application.
- (Optional) Sex of person in each photograph. Whilst this characteristic is not an input into the system, it was recorded for statistical purposes. Furthermore, to achieve the optional criteria discussed in preceding sections, data points must be labelled as male/female to train separate models.
- (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 may 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 was encouraged that the data-set of subjects who volunteer to be photographed be spread in terms of sex, ethnicity and age. The final dataset that was used contained 465 photographs.

## VI. METHODOLOGY

Drawing from the previously discussed literature, the following sections describe methodology employed to achieve each component and subsystem implementation described above. The chosen technologies used for this system is Python, due to its open-source nature, large online community support, and its growing popularity in the Machine Learning field.

### A. Extraction Layer

1) *Reference Object Detection and Segmentation:* The purpose of Reference Object Detection is to extract the pixels per meter metric from the photograph, since this metric is used for all subsequent data extraction in the succeeding components such as the person's height, silhouette area, widths and depths. Two methods have been used to extract the RO from the photograph, both of which make extensive use of the OpenCV-Python library. Method One is the primary operation, while Method Two is activated if Method One fails during runtime. Method One is as follows:

- Knowing that the RO is black, the image was filtered of all pixels whose 8-bit colour fell outside of the gray range of 0 – 150, where 0 is pure black and 255 is white, such that the remaining pixels were considered as the 'black' pixels of the image. The value of 150 was chosen experimentally, as to account for lighting in the photographs causing the RO to appear slightly lighter than it should.
- The filtering produced a binary image which showed the positions of black pixels in the image. Edge detection (using the OpenCV implementation of the Canny algorithm) was performed on the filtered image, returning all contours in the image.
- OpenCV's Polygon approximation was then performed on the contours that were returned. This returned all contours in the photograph that have four vertices (rectangular). The contours were ordered from left to right, and the left most contour was assumed to belong to the RO, and returned it.

This method is precise since it searches for a specific type of contour from a specific subset of the pixels in the image, and was partially inspired by Ref. [11]. Unfortunately, whilst this method works for most images there are instances where the lighting in the room caused the pixels which comprise the RO to be so brightened that it fell outside the range of gray specified above. In this case, no rectangular contours would be returned, triggering Method Two as a fail-safe. Method Two is as follows:

- The image is first converted to grayscale and slightly blurred to reduce noise.
- Morphological closing is performed on the image to fill in gaps that may prevent contour detection.
- Edge detection was performed, and returned all contours in the image, ordered from left to right.
- Smaller contours arising from noise were excluded on the basis of the pixel area contained within.

This method (inspired by Ref. [10]), whilst more reliable in finding the contour of the RO, is far more susceptible to noise in the image. As a result, it is susceptible to returning a distorted RO contour. Furthermore, it is also susceptible to non-rectangular objects appearing to the left of the RO being falsely returned as the RO (due to the left most object being assumed as the RO). Using the known dimensions of the RO, the pixels per meter metric is found and returned.

2) *Person Detection and Segmentation:* The purpose of Person Detection is to return the pixel mask of the person in the image and extract that the aforementioned FVV and SVV. Using the key assumption stated earlier that the person in the photograph is

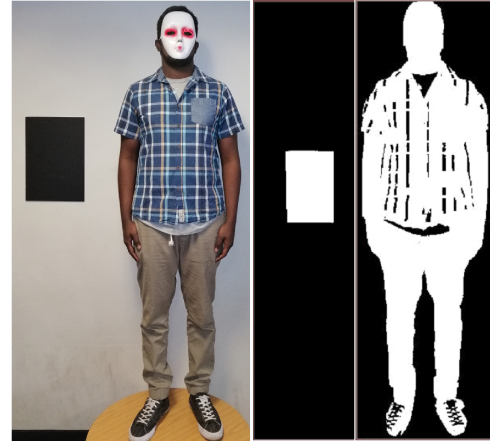


Fig. 1. Example photograph (left), reference object segmentation (middle), person segmentation (right).

standing in front of a uniform background, the following method was developed, partially inspired by Ref. [12]:

- The Mask-RCNN model developed in Ref [13] was used to detect the Region of Interest (ROI) that contained the person – their bounding box. This pre-trained model has been trained on the MS COCO dataset to detect people, among other common objects, and return an approximate (rough) pixel mask of the person [14].
- The bounding box returned from the Mask-RCNN was then used to perform foreground extraction on the corresponding ROI in the original image (using the OpenCV implementation of the GrabCut algorithm), returning the pixel mask of the extracted foreground.
- The foreground-produced mask more accurately defines the silhouette but is susceptible to noise due to nearby objects, whilst the CNN-produced mask provides an approximate shape of the person but is almost unaffected by nearby objects.
- These two pixel masks are combined using the bit-wise AND operation, such that the final pixel mask only contains pixels found in both the foreground-produced and CNN-produced masks.

The returned pixel mask was rid of zero rows as to crop the top and bottom of the mask. The number of rows in the remaining image was taken as the pixel height of the person, and converted using the pixels per meter metric to a physical height. The pixels per meter metric was used to calculate physical area per pixel. The pixels in the mask were summed and then converted to physical area using this metric.

The image was then approximated into eight sections, to represent the eight distinct regions of the body: Head, Neck, Thorax, Abdomen, Pelvis, Upper Leg, Lower Leg, Foot [15]. The row with the largest number of pixels (maximum thickness) of the Thorax, Abdomen, Pelvis and Upper Leg regions were then extracted and converted to physical dimensions. Altogether, these features made the FVV and SVV. This method of data extraction and composition of FVV and SVV were inspired by Ref. [3]. A visual representation of the entire Extraction Layer can be seen in Figure 1.

## B. Machine Learning Layer

1) *Model Choice and Evaluation:* As per consulting the aforementioned Literature Review, it was found that the simplest type of Neural Network that may account for a noisy, small, and potentially linearly inseparable data-set is the Multilayered Perceptron (MLP). The Front, Side and Compensation models were all MLP's. The models were evaluated by a weighted scoring system that took the maximum deviation from the mean as well as the Mean Absolute Error (MAE) into account. MAE (measured in BMI) was considered as the main metric of model loss, since in this particular application, the accuracy of individual results was more important than evaluating the spread of BMI estimations [16].

2) *Training and Final Performance:* Before training, 20% of the original dataset was removed to be used later as unseen data during model evaluation. The remaining data was used with both the Classical Test-Train split and (Stratified) k-Fold cross validation methods to train the models. The Classical method used a 70/30 train/test split for training, whilst the Cross Validation methods used 10 folds.

Different models were trained using different shuffles of data as well as different training parameters such as learning rate, batch size and number of epochs. At the end of each complete training cycle, the models were evaluated by exposing them to the unseen data. The model which had the lowest combination of MAE and maximum deviation when exposed to completely unseen data was saved as having the best performance. The final Model Loss and Training Method for the best performing models were as follows:

- Front: Trained using Stratified Cross Validation using 10 folds, Mean Absolute Error (MAE) — 3.32
- Side: Trained using Cross Validation using 10 folds, MAE — 2.77
- Compensation: Trained using the Classical Test/Train Split, MAE — 2.58

Note that the Front and Side View models were first trained, and the best performing models were used to generate data to train the Compensation model.

The final architecture of the Front model was 4-3-2-1, i.e. four neurons in the input layer, three in the first hidden layer, two in the second hidden layer, and one in the output layer. The final architecture of the Side and Compensation models were \*\*\*\*\* and \*\*\*\*\* respectively.

## VII. SOURCES OF ERROR

As expected when performing feature extraction on a photographic dataset, the system design either had to account for or circumvent various errors in the dataset. The following errors were circumvented by creating a Method One as described in Section VI-A1:

- Poorly lit photographs, or non-uniform low-contrasting backgrounds that cause the RO to be confused with the background.
- The person being too close to or making contact with the RO.

- Another object being more to the left of the photograph than the RO.

The following errors were circumvented by combining the masks produced by foreground extraction and the Mask-RCNN:

- The person being in contact with another object that would affect normal edge detection.
- The photograph containing other large objects that were in the foreground with the person.

Unfortunately there are some errors that could not be accounted for or circumvented, and must be prevented in any future data collection:

- Taking front and side photographs from an angle such that the RO is seen as behind the person instead of alongside them.
- A more pronounced error of parallax in side photographs due to the RO being a person-width behind the side surface of the person being photographed, resulting in the side dimension extractions being slightly exaggerated.

## VIII. FINDINGS & ANALYSIS

### IX. ADDITIONAL FEATURES

The program allows for height and mass models to be trained, as to provide another means of BMI estimation. The training and evaluation is the same as described above, producing the highest performing model.

All models are automatically indexed by date, time, model type (BMI/Height/Mass), Network architecture. The training also produces an output file listing the parameters used for the cycle, as well as model performance per training type (Classical/Cross Validation).

Finally, the program also provides the user with the option to visualise the data being extracted from the image for the purpose of error detection. This can be extended in future versions of the program to interactively choose a reference object in the photograph.

## X. RECOMMENDATIONS

In accordance with Section VII, as to improve the quality of the training data as well as model performance, the following minor changes must be implemented when creating a dataset for this system in future:

- There must be at least half a meter of space to the left of the RO in the photograph, as well as to the right of the person in the photograph. Having other objects vertically below the RO or between the persons head and shoulder in the picture puts the data point at risk of being discarded by the program due to its built-in data sanitisation.
- Consider making the RO a bright, distinctly coloured object worn on the person. This would improve the system's ability to locate the RO by first detecting the person, and using the same ROI to find the RO.

With regards to the speed of execution and resource usage of the program, the following changes should also be implemented in future:

- The program should be parallelised as far as possible. Processes such as RO detection and person detection are not required to be sequential.
- Instead of using the pre-trained Mask-RCNN model that is trained on 81 classes, train a new model on the same MS COCO dataset but only on two classes: person and background. This would decrease the size of the model, as well increase the speed of person detection.

## XI. CONCLUSION

## XII. ACKNOWLEDGEMENTS

The author would like to extend his gratitude to the following people:

- Sachin Govender (1036148) for his contribution, particularly to the Machine Learning layer of the system. His findings can be seen in Ref [17].
- Prof. Scott Hazelhurst from the University of the Witwatersrand, who supervised this study.
- Dr. Neil Martinson from the Peri-Natal HIV Research Unit. The PHRU have been invaluable to this study due to their contribution to data collection.

## 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] A. Mehta, "A complete guide to types of neural networks," Jan 2019. [Online]. Available: <https://www.digitalvidya.com/blog/types-of-neural-networks/>
- [9] 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," *arXiv.org*, Aug 2019. [Online]. Available: <http://arxiv.org/abs/1908.11694>
- [10] A. Rosebrock, "Measuring size of objects in an image with opencv," Mar 2016. [Online]. Available: <https://www.pyimagesearch.com/2016/03/28/measuring-size-of-objects-in-an-image-with-opencv/>
- [11] A. Rosebrock, "Opencv shape detection," Feb 2016. [Online]. Available: <https://www.pyimagesearch.com/2016/02/08/opencv-shape-detection/>
- [12] A. Rosebrock, "Keras mask r-cnn," Jun 2019. [Online]. Available: <https://www.pyimagesearch.com/2019/06/10/keras-mask-r-cnn/>
- [13] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [14] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.
- [15] "Human figure proportions." [Online]. Available: <https://anatomyclass.com/anatomy-video-lessons/human-figure-proportions>
- [16] L. Simmons, "Estimating body mass index from face images using keras and transfer learning," Feb 2019. [Online]. Available: <https://medium.com/@leosimmons/estimating-body-mass-index-from-face-images-using-keras-and-transfer-learning-de25e1bc0212>
- [17] S. Govender, "Digital estimation of body mass index," School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, 4th Year Project Report 19P44, Sep 2019.