Automatic Identification of Somatotype by Digital Images

Antonio Ricardo A. Brasil * **Thales de Oliveira Gonçalves** **
Jorge Washington Brombley Castro *** **Eliane Cunha Gonçalves** ****
Klaus Fabian Côco * **Patrick Marques Ciarelli** *

* *Departamento de Engenharia Elétrica, Universidade Federal do Espírito Santo, Vitória, ES, Brasil.*
** *Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Paulo, Brasil.*
*** *Secretaria de Saúde da Serra, Serra, ES, Brasil.*
**** *Faculdade Estácio de Sá, Vitória, ES, Brasil.*

Abstract: Somatotype is a metric that concerns human body shape and composition. It is important in many applications, especially for the fields of physical education and health. However, obtaining the somatotype is very time-consuming, it requires experts to take several measurements manually on the individual's body and demanding several anthropometric devices.The proposal of this work is to obtain the somatotype using their body images in different positions, based on image processing and machine learning techniques, which is a less expensive alternative to obtain this information. Thus, a dataset of 46 bodybuilders was obtained along the years from 2014 to 2016 in the state of Espírito Santo. The results obtained show that the type of bodybuilders somatotype can be estimated reasonably based only on their images, obtaining the best classification rate equal to 92%.

Keywords: Somatotype, Feature Selection, Artificial Neural Network, Genetic Algorithm.

1. INTRODUCTION

In the last centuries, various researchers have tried to understand the human body and the relationships existing in the physical manifestations of body proportions. However, research of this nature only began to be made with scientific rigor in the early 20th century (Carter and Heath, 1990), when Sheldon et al. (1940) coined the term "somatotype" after performed various measurements from photographs of different individuals. Years later, Heath and Carter (1967) proposed a method in which measurements were made directly on the subject's body and they used mathematical expressions to infer the somatotype.

Over the years, the classification of the individual by his somatotype has become a fundamental tool for the evaluation of the life quality and athletic performance (Thorland et al., 1980; Bolonchuk et al., 2000; Massidda et al., 2013; Gutnik et al., 2015).

Nowadays, the technique proposed by Heath and Carter (1967) is still used in the same way. However, this method has disadvantages, such as the fact that the method is invasive, the time required to obtain the measurements, the need for specific equipment, which are not so portable, and the necessity of qualified professionals. Nevertheless, the use of image processing with machine learning (Klafke, 2018; Santos, 2020) and the arrival of more affordable, higher-quality and cheaper photographic equipment, it is possible to develop faster and less invasive methodologies.

The first work which proposed the automatic classification of somatotype by digital images was (Gonçalves et al., 2016). The authors manually collected measures from digital images and, subsequently, applied machine learning techniques to estimate the somatotype from these measurements. However, the manual collection of the measurements keeps the same problem from the existing methods.

This work aims to perform the somatotype classification by means of measurements obtained automatically from images, using image processing techniques and machine learning. For evaluating the proposed method, a dataset composed by images of 46 bodybuilders will be used, and the results will be compared with those obtained by the method of Heath and Carter (1967).

The main contributions of this work are: proposing a protocol for acquisition of images to estimate somatotype; a methodology for estimating somatotype components; and the indication of the most relevant body measurements to perform this task.

This work has the following structure. In Section 2 is presented the current procedures for calculating the somatotype. In Section 3 the proposed methodology is discussed. In Section 4 the results are shown. Finally, in Section 5 the conclusions and future paths are presented.

⋆ The authors thank the financial support for the research from the project of the Fundação de Amparo à Pesquisa do Espírito Santo (FAPES), number 598/2018.

2. SOMATOTYPE

Sheldon et al. (1940) were the first that coined the term somatotype and defined that it is composed of three embryonic tissues: ectoderm, mesoderm and endoderm. In order to study and conceptualize the somatotype, they used photographs of 4.000 naked students in front, profile, and dorsal positions. Individuals were photographed with the palmar region facing the thigh and their bodies in the images were then divided into five regions for measurements, comprising 16 measurements.

However, the technique proposed by Sheldon et al. (1940) encountered a number of constraints, given the technology of that period: the use of photographs and obtaining anthropometric measurements from them made the method costly, requiring time and an adequate environment, and the possibility of human mistakes (Fernandes Filho, 2003).

For this reason, Carter and Heath (1990) proposed a new method, where ten measurements are obtained directly from the body: stature (H), body mass (M), triceps skinfold (TR), supraspinale skinfold (SE), subscapular skinfold (SB), medial calf skinfold (PA), bi-epicondylar bone diameter of humerus (DU), bi-epicondylar bone diameter of femur (DF), perimeter of flexed arm in maximum contraction (PB) and calf circumference (PP).

To carry out the measurements, five devices are needed: a scale (to measure body mass), a stadiometer with a moving head (equipment that measures height), an adipometer/plicometer (for measuring skinfolds), a small caliper (used to measure bone diameters) and a flexible metric tape made of metal or fiberglass (which is used to measure perimeters) (Fernandes Filho, 2003).

From these measurements, the components of endomorphism (Endo), mesomorphism (Meso) and ectomorphism (Ecto) are estimated using Equations 1 to 3.

$$Endo = -0.7182 + 0.1451XC - 6.8x10^{-3}XC^2 + 1.4x10^{-6}XC^3 \quad (1)$$

$$\text{Meso} = 0.858\text{DU} + 0.601\text{DF} + 0.188\text{PcB} + 0.161\text{PcP} - 0.131\text{H} + 4.5 \quad (2)$$

$$Ecto = \begin{cases} 0.1, & \text{if } IP \leq 38.25 \\ 0.463IP - 17.63, & \text{if } 38.25 \leq IP \leq 40.75 \\ 0.732IP - 28.58, & \text{if } IP \geq 40.75 \end{cases} \quad (3)$$

where Equations 4 to 7 are auxiliary expressions.

$$XC = \frac{170.18}{H}(TR + SB + SE) \quad (4)$$

$$IP = \frac{H}{\sqrt[3]{M}} \quad (5)$$

$$PcB = PB - TR \quad (6)$$

$$PcP = PP - PA \quad (7)$$

This way, the somatotype of an individual is represented by the ordered triple with the values of the corresponding *Endo-Meso-Ecto* measures. However, not all combinations of somatotype values are biologically possible, for example, somatotypes 2-2-2 or 7-8-7 are impossible (Carter, 2002). Generally, somatotypes are high in endomorphy and/or mesomorphy if, and only if, they are low in ectomorphy (Carter, 2002). Components with values from 0.5 to 2.5 are considered low, 3 to 5 are considered moderate, 5.5 to 7 are high and above 7.5 are extreme (Carter and Heath, 1990; Tóth et al., 2014).

The components should be rounded to multiples of 0.1 or 0.5 depending on the application (Carter, 2002). The values are usually rounded to the nearest half unit, to achieve simplicity of communication and summary of the results (Carter, 2002). If any component resulting from the equations is zero or negative, the value of 0.1 is associated with that component, since by definition the components must be positive, however, values smaller than 1 are extremely rare to be observed (Carter, 2002).

Beyond the errors inserted by the anthropometric devices, there is also the possibility of human error in the application of the method and, for this, it is necessary experts during the measurement procedure. It is recommended that the expert performs all measurements three times and computes the average of such values (Carter, 2002). This reduces the interference of human error, but makes the application of the technique more repetitive and even more time-consuming.

2.1 Somatochart

The most usual way to represent the somatotype is through a two-dimensional space called somatochart (Carter and Heath, 1990). The points of the somatochart are determined by Equations 8 and 9.

$$X = Ecto - Endo \quad (8)$$

$$Y = 2Meso - (Endo + Ecto) \quad (9)$$

Note that the shift of the three-dimensional space from the somatotype to the somatochart is associated with an orthogonal linear projection (Figure 1).

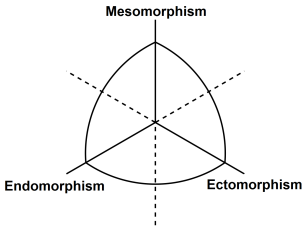


Figure 1. Somatochart. Adapted from (Carter and Heath, 1990).

The origin of the axes of the somatochart locates the individuals with somatotype with the three components of the same value 4-4-4. The approximately circular region around the axes of the somatochart represents the more biologically common somatotypes to be observed.

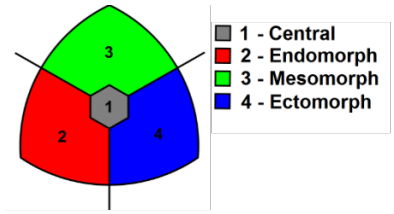
A commonly used metric is the Somatotype Attitudinal Distance (SAD), which is the Euclidean distance between

two somatotypes A and B. This metric is commonly used when an individual presents a given somatotype (A) but wishes had another somatotype (B) through physical exercises and planned diet. Equation 10 shows how to obtain the SAD value between A and B (Carter, 2002).

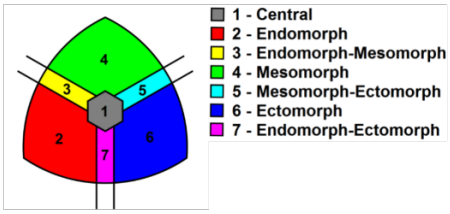
$$SAD_{A,B} = \sqrt{((Endo_A - Endo_B)^2 + (Meso_A - Meso_B)^2 + (Ecto_A - Ecto_B)^2} \quad (10)$$

2.2 Classes of Somatotypes

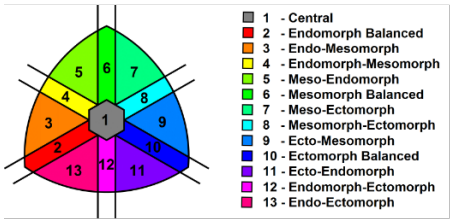
The somatochart can be divided into regions to group individuals. The most commonly used division presents 4 classes (ecto, endo, meso and center), as shown in Figure 2a. Figures 2b and 2c show the somatochart divided into 7 and 13 regions, respectively.



(a) Four classes.



(b) Seven classes.



(c) Thirteen classes.

Figure 2. Somatochart with four, seven e thirteen classes.

3. MATERIALS AND METHODS

Figure 3 presents a diagram of the methodology used in this research. It will be discussed in the following sections.

3.1 Dataset

The data were collected with different resolutions, 960×1280, 2448×3264, 2432×4320 and 4000×6000 pixels, with the athletes positioned in front of a white background and with artificial lighting. The cameras were positioned at the same height. The images were captured from bodybuilders in the regional championships of the International Federation of Bodybuilding (IFBB) during the years 2014 to 2016 in the State of Espírito Santo - Brazil.

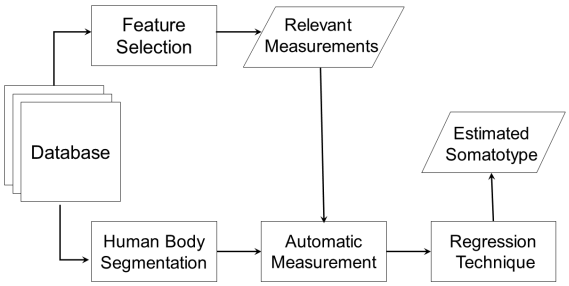


Figure 3. Diagram with the proposed solution.

(a) Frontal. (b) Lateral.

Figure 4. Example of images in frontal and lateral positions.

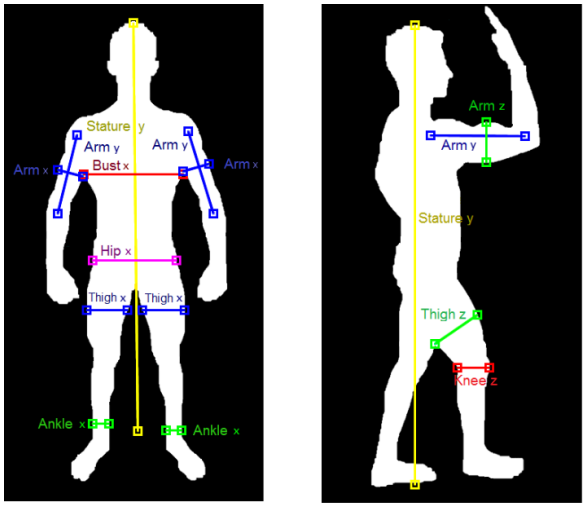
The images were recorded in JPG format, with two images for each athlete, one at each position (Figure 4a and 4b). With these images all the necessary measures for the proposed method are obtained.

For both images, it can be observed that the individual should have legs and arms open, so that it is possible to correctly segment the silhouette. It is also necessary to wear tight bathing suits so that the real measures of the body are not altered. Long hair should be tied to the back, so that it does not hamper the measurements. A mark of known size is also used to estimate the real size of the measurements (rectangular element in the background).

The dataset is composed of images from 46 individuals (15 males and 31 females). The average age of individuals is 29.3 ± 6.9 years.

3.2 Identification of the Most Relevant Body Measurements

In order to automate the measurement extraction, it is important to identify first which measurements are the most relevant. For this, we can imagine the horizontal (x), vertical (y) and depth (z) axes positioned in the human body (frontal image). Considering these axes, 29 measurements of the body were proposed: height (y), neck (x and z), shoulder (x), bust (x), waist (x and z), hip (x and z), arm (x, y and z), elbow (x and z), forearm (x, y and z), wrist (x and z), thigh (x, y and z), knee (x and z), calf (x, y and z) and ankle (x and z). Figures 5a and 5b



(a) Frontal. (b) Lateral.

Figure 5. Manual measurements obtained from images.

show measurement points obtained manually in the image of an athlete silhouette.

From these direct measurements, 17 indirect measurements were proposed by Gonçalves et al. (2016): span; perimeters (neck, waist, hip, arm, elbow, forearm, wrist, thigh, knee, calf, ankle); ratios (arm, forearm, thigh, calf); and the height-to-span ratio. As the span corresponds to the distance between the fingertips of an individual when with arms horizontally open, this was approximated by the sum of the measurements: left forearm (y) + left arm (y) + shoulder (x) + right arm (y) + right forearm (y). To obtain the perimeter, the parts of the body involved are approximated by ellipses of diameters equal to their measurements obtained on the x and z axes. The ratio measurements were obtained by dividing the area of the ellipse by its respective part in the y direction. Finally, the height-to-span ratio is obtained by taking the ratio between height and span.

Additionally, two external information were added: gender and body mass of the bodybuilders. Thus, a total of 48 characteristics are obtained for each individual (29 direct measurements, 17 indirect measurements, gender, and body mass). The proposed measurements were inspired by the measurements taken by Sheldon et al. (1940), and other ones were taken to capture more information from the shape of the silhouettes.

To identify the most relevant measurements for each of the three components of the somatotype, a Multiobjective Optimization with Genetic Algorithm (GA) was performed (Deb, 2001), with settings: population of 200 individuals, maximum of 100 generations, uniform mating and mutation, and selection by tournament size 4 (2% of population), each individual corresponds to a binary vector of 48 positions indicating whether the i-th characteristic is selected.

The optimization has two objectives to be minimized: the amount of measurements to be extracted from the image and the mean square error for Endo, Meso and Ecto obtained by a Multilayer Perceptron (MLP) with 10 hidden neurons. The Leave-One-Out method was used to split the dataset (Witten and Frank, 2005).

The sets selected by the GA to estimate each of the three components of the somatotype has 5 measurements each, comprising 11 different measurements. Table 1 shows these measurements. The measurements that appear more than once are highlighted in bold. It is interesting to note that the hip measurement appears to be relevant for the three components of the somatotype.

Table 1. Selected measures by GA.

Endomorphism	Mesomorphism	Ectomorphism
Gender	**Stature** (y)	**Gender**
Stature (y)	**Hip** (x)	**Hip** (x)
Hip (x)	Ankle (x)	Arm (x)
Bust (x)	Thigh (z)	Thigh (x)
Knee (z)	Arm (ratio)	Arm (y)

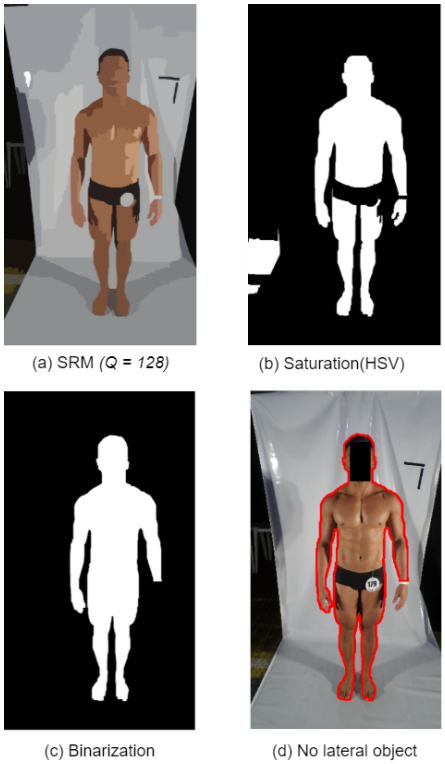
After selecting the 11 most relevant measurements, an approach was developed for automatic extraction of these measurements through image processing. The approach is divided into two phases: segmentation of the human body and automatic measurement extraction. The next sections will describe these procedures.

3.3 Image Segmentation of Human Body

In order to extract body measurements from an individual, it is necessary that his body be segmented. This step is divided in two parts: resizing the images to the height of 1280 pixels preserving proportions; and application of Gaussian low pass filter 3×3 with standard deviation of 0.5 to blur the image.

Given the pre-processed image, the segmentation of the human body image follows another five steps:

(1) Separation of the image of the individual in disjoint regions using the Statistical Region Merging (SRM) technique (Nock and Nielsen, 2004) with parameter $Q = 128$ defined by trial and error (Figure 6a).
(2) Then binarization using the Otsu method (Gonzalez and Woods, 2011) on the saturation channel in the HSV color space (Figure 6b) is applied.
(3) For removal the lateral objects, the number of pixels in each column of the binarized image is counted and a graph of pixel count is generated. The region with the largest area of the graph is chosen and the others regions is discarded.
(4) Low-intensity regions in general have low saturation and therefore are not segmented by binarization. Some regions of low intensity can be fundamental parts of the composition of the individual's silhouette, such as shadow regions caused by lighting, clothing, or hair. To reconstruct, Otsu's limiar is calculated over the SRM image. Regions that have an intensity below otsu's limiar and that are adjacent to some already segmented regions are candidates for regions for reconstruction. Then, three tests are applied: the first checks if the centroid of the candidate region is greater than 25% of the image height. If pass, then the second test is applied, which checks the quantity of image connected components. If when reconstructing the region the amount of components decreases, then



(a) SRM (*Q = 128*) (b) Saturation(HSV)

(c) Binarization (d) No lateral object

Figure 6. Steps realized for body segmentation.

the region is reconstructed; otherwise, a third test is applied, which checks if the perimeter/area ratio decreases, it is understood that the candidate causes a change to a smoother region and its reconstructed. Otherwise, the region is discarded.

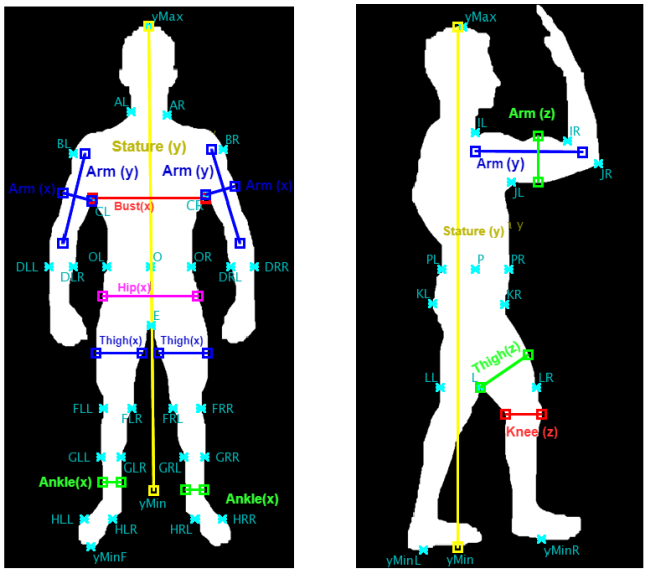
(5) Morphological closing of empty regions (holes) resulting in the image of the silhouette in Figure 6c. Initially, a morphological closing operation is performed on the image using a 5x5 mask of ones as a structuring element. Then, the holes of the binarized image, candidates to be closed, are obtained. To decide whether a hole should be closed, the height/width ratio is calculated. If the height/width is lower than 2.1, the hole is closed; otherwise, the region is not closed. This limiar (2.1) was chosen after an extensive parameter change.

Figure 6d shows the final contour of the segmentation over the original image.

3.4 Automatic Measurement Extraction

For each silhouette (frontal and lateral), key points are identified and used to extract the measurements. Each key point is described with a label or a sequence of letters. Figure 7 presents all key points and the details are described below. For simplicity, when it is mentioned silhouette, we are referring to the silhouette contour. To walk on the silhouette, we used the Freeman chain code (Gonzalez and Woods, 2011).

For the frontal image, the following key points are identified: The centroid of the individual (O) and, at the same height, two points on the hip (OL and OR) and two points on each arm (DLL and DLR, on the left arm, and DRL and DRR, on the right arm). The centroid of the highest



(a) Frontal. (b) Lateral.

Figure 7. Key points on human body.

points is used to represent the highest point ($yMax$) on the silhouette. In order to compute the stature (y), it is computed the centroid from the region that is 80% of distance from O to the lowest point on the silhouette ($yMinF$). This centroid is projected ($YMin$) on the line that passes on O and $yMax$. The difference between $yMax$ and $YMin$ is the stature (y) in pixels.

For the bust (x), the points CL and CR are identified. CL is the highest point on the silhouette in the path from DLR to OL (in clockwise direction). The same procedure is used to identify CR, but using the path from OR to DRL. The difference between CL and CR is the bust (x) in pixels.

For the ankle (x), a horizontal line is drawn in the middle distance between $yMin$ and $yMinF$. So, four points are identified on the line that touches the silhouette: from left to right, HLL, HLR, HRL and HRR. A key point E is defined as the highest one in the path from HLR to HRL. After then, a horizontal line at 20% of distance from $yMin$ to E is used to locate the points GLL, GLR, GRL, and GRR. The left ankle (x) is obtained by looking for the smallest horizontal distance in pixels between XL and XR (which corresponds to the smallest width of the leg among the points HLL, GLL, GLR and HLR). Similarly, the right ankle (x) is obtained.

The hip (x) is the distance between the points on the silhouette that touches the horizontal line that passes in the half distance between O and E.

A 10th order polynomial regression (order defined by trial and error) is applied to smooth the path between GRR and CR and between GLL and CL. Figure 8 shows these paths in red. Green asterisks correspond to a local minimum and blue the local maximum points of the polynomial regressions. With $j1$ (local minimum) and $j2$ (local maximum) being these points of the knee, (Figure 8), a horizontal line is drawn in the middle of the heights of $j1$ and $j2$. This line cuts the individual's edge into four points, defined from left to right as FLL, FLR, FRL, and

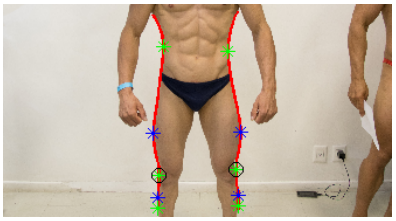


Figure 8. Polynomial regression between GRR and CR.

FRR. Then, in the half distance from hip x to FLL and FLR is the left thigh (x), and in the half distance from hip x to FRL and FRR is the right thigh (x). Figure 7 helps to visualize this procedure.

For the arm, initially, the clockwise direction on the silhouette from DLL to DRR is coursed. Four points are selected from the gradient angle of this path: pll, plr, prl and prr. The first and last represent the most horizontal gradients (probably, shoulder points), while the middle ones represent head points (commonly on the ears). Two straight line segments are created, the one on the left (sl) joins pll and plr, while the one on the right (sr) joins prl and prr. The key point AL is defined as the point between pll and plr that is as far away as possible from sl. Similarly, AR is the point of X between prl and prr that has the greatest possible distance from sr. Defined AL and AR, then, BL is defined as the first point on silhouette before AL whose gradient angle is greater than 45°. Similarly, BR is the first point after AR whose gradient angle is less than -45°.

The arm (y) starting point is given by the center of mass of the arm/forearm of each side (left and right). Thus, the individual's arm/forearm is divided into two parts: above this starting point (arm) and below (forearm). To determine the measurement direction of the arm (y) from its starting point, the Principal Component Analysis (PCA) (Gonzalez and Woods, 2011) is used. Let $v1$ and $v2$ the eigenvectors of the first and second greatest eigenvalues, respectively. So, the ending point of measurement of the arm y is the intersection of lines $r1$ and $r2$, where $r1$ and $r2$ have the directions of $v1$ and $v2$, respectively, and $r1$ pass through the starting point of arm (y) and $r2$ passes through BL (left arm) or BR (right arm). Finally, the arm (x) is the greatest width of the arm at the $v2$ direction.

For the lateral image, the key points P (center of mass of the individual), and PL, PR are defined, the last two are the edge points respectively to the left and right of P, at the same height. Then, the key points $yMax$, $yMinL$, and $yMinR$ are defined, where the first one is the highest point of the individual to the left of PR. The points $yMinL$ and $yMinR$ are taken as the smallest points that are, respectively, to the left of PL and to the right of PR. The point $yMin$ is defined as the intersection of the lines $r1$ and $r2$, where $r1$ is the vertical line that passes through the point $yMax$ and $r2$ is the line that passes between $yMinL$ and $yMinR$. The distance between $yMax$ and $yMin$ is the stature (y).

For the knee (z), the internal path from $yMinL$ to $yMinR$ is obtained. The key point L is the highest point on the path. Thus, the points LL and LR are located on the silhouette and at the same height of L. So, the gradient angles of the path from L and $yMinR$ are obtained, and the local minimum of these angles represents the moment of contour change of the hind thigh to the calf. This is the starting point of the knee (z). The end measurement point is the point on the silhouette that is horizontally at the right of the starting point.

For the arm (x), the path from PR to $yMax$ is coursed in the anticlockwise orientation. With the passage from the chest to the arm, the gradient angle takes the value of 0°, and this point is identified as JL. When the path goes from the arm (nearly horizontal) to the forearm (approximately vertical), the gradient angle takes the value of 45°, and this key point is set to JR, and is found in the subject's elbow. With the path going around the individual's hand, the key point IR is defined, which assumes an angle value of 225°. With the shift from the arm to the individual's neck/head, the gradient angle decreases and assumes a value of 135°, setting the key point IL.

Similar to the measurement of the frontal arm (x), the PCA is applied, obtaining the vector $v1$ and $v2$, considering the region delimited by JR, IR, IL, and JL. The arm start point (y) is the midpoint of IR and JR and the end point is the intersection of lines $r3$ and $r4$, where $r3$ is the line that passes through the arm start point (y) and has $v1$ direction; and $r4$ is a vertical line that passes through IL. Finally, the arm (z) is obtained by taking the longest width of the arm in the $v2$ direction which is to the right of JL and to the left of IR.

For the thigh (z), a path from the key point LL and PL is created. From this path, the local minimum is extracted and defined as the key point KL (Figure 7). Similarly, with a path from the key point LR to PR, another local minimum is extracted, defined as key point KR. Then, the thigh is extracted from a region defined by for lines: $r1$ is the line that connects the key point KL to KR; $r2$ is the vertical line that passes on the final knee (z). $r3$ is the horizontal line at knee height (z), and $r4$ is the line that connects KL to the initial knee point (z). By applying PCA, the eigenvectors $v1$ and $v2$ are obtained, where $v1$ is the most representative. The initial point of the thigh (z) starts at L and the ending point is the border obtained of L extending in the direction of $v2$. So, thigh (z) is the distance between these two points, in pixels.

All measurements are made in pixels. For measures taken more than once, we computed the average value. In order to convert to the metric system, it is used a reference (as shown in Figure 5), whose real dimensions are known and, by proportion, the real human body measurements are estimated.

4. EXPERIMENTS

4.1 Experimental methodology

Due to the limitation of the dataset, the Leave-One-Out method was used to maximize the amount of data (samples) of individuals for training. In this method, $n-1$ samples are used to train and 1 sample to test, where n is the total number of samples. The method is repeated n times, so that each sample is used once as a test. At the end, the result is averaged.

An MLP with one hidden layer was used to estimate the somatotype and the measurements. The number of neurons in the hidden layer was adjusted between even values from 2 to 20.

Five different metrics were used to evaluate the results: Mean Error (ME), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Equations 11 to 14 show the calculation of each metric, where n is the number of samples, $\hat{y}_i$ is the given output by the estimators and y_i is the desired output (target). For each metric, the closer to zero, the better the performance of the model is.

$$ME = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \tag{11}$$

$$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right|\right) \times 100\% \tag{12}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{13}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{14}$$

4.2 Results

Table 2 shows the value of the metrics for each measurement extracted from the image. The values obtained automatically were compared with the measurements that were taken manually in the images.

Table 2. Errors of Automatic Measurements.

Measurement	ME cm	MAPE %	MAE cm	RMSE cm
Stature (y)	**5.00**	3.48	**6.17**	**8.66**
Arm (x)	-0.23	7.75	3.10	3.85
Hip (x)	**2.85**	**9.04**	3.34	3.96
Arm (x)	0.19	6.15	0.66	0.88
Thigh (x)	1.01	6.97	1.31	1.57
Ankle (x)	0.00	6.12	0.38	0.46
Thigh (z)	**-2.01**	**10.97**	2.67	3.63
Knee (z)	-0.24	5.11	0.70	0.93
Arm (y)	-0.62	11.37	**4.00**	4.83
Arm (ratio)	0.19	**14.00**	0.51	0.65

Most of the measurements present relatively low MAPE values (below 10%). The measurement with the highest error is Arm (ratio), because it is a function of other measured variables that also already contain approximation errors. The measures Thigh (z) and Arm (y) also presented MAPE above 10% due to the difficulty of precisely determining the starting point of such measurement. The Stature (y) was the measurement that presented the greatest error from the point of view of ME, MAE and RMSE, this is also due to the difficulty of accurately obtaining the start and end points of this measurement.

Another measure that had a polarized estimate was Hip (x), with the second highest ME and, moreover, almost 10% of MAPE. In some cases of the dataset, this measurement was being taken at a less than adequate height at the beginning of the thigh and, therefore, the estimation of this measurement is overestimated, reflected in the relatively high value of ME.

For modeling the classifier, 11 features (10 measurements and gender) were used to infer the somatotype according to Table 3. The entries were normalized to have zero mean and unit variance, the target values were taken by professionals and kept without rounded values, rounded to multiples of 0.1 and 0.5.

Table 3. Errors of Somatotype Estimating.

Component		ME cm	MAPE %	MAE cm	RMSE cm
Endo	Floating Value	-0.04	10.98	0.17	**0.26**
	Multiple 0.1	-0.04	11.87	0.17	0.28
	Multiple 0.5	**-0.02**	**8.13**	**0.13**	0.27
Meso	Floating Value	**-0.02**	21.97	0.37	0.57
	Multiple 0.1	**0.02**	18.35	**0.33**	**0.52**
	Multiple 0.5	-0.06	**11.62**	**0.33**	0.55
Ecto	Floating Value	0.08	15.01	0.15	**0.24**
	Multiple 0.1	-0.06	19.42	0.16	0.26
	Multiple 0.5	**-0.01**	**8.38**	**0.12**	0.26

MAPE had a better result with a multiple rounding of 0.5 in all three classes. For the MAE, the results show that the MLP estimates, for example, that endomorphism and ectomorphism of the dataset individuals, on average, does not distance more than 0.13 in the representation of multiples of 0.5. In general, the use of a multiple rounding of 0.5 presents a satisfactory result for the somatotype classification.

The distances between the three and two-dimensional somatotypes, that is endo-meso-ecto and somatochart components, are shown in Table 4.

Table 4. SAD and Distance in the Average Somatochart.

Representation	SAD Average	Average Distance in Somatochart
Floating Value	0.53	0.52
Multiple 0.1	0.52	0.51
Multiple 0.5	0.50	0.50

Figure 9 shows the real somatotypes (in blue) and the ones estimated (in red) by MLP when using a multiple of 0.5. The straight segments (in black) connects the estimated points with their respective real values. The green points represent error-free matching between the real and the estimated value. The radius around of the black point within the somatochart indicates the region where an estimate of an arbitrary somatotype is considered a true value if located inside and, as can be observed, its area is considerably reduced when compared to the total area. This suggests that the somatopoint location on the somatochart presents reasonable accuracy when estimated by the proposed system.

Table 5. Accuracy in the Classification.

Classification Type	13 Classes	7 Classes	4 Classes
Accuracy	72.14%	89.29%	92.86%

Table 5 shows the results obtained for the somatotype classification considering the three subdivisions of classes presented in Section 2.2. The best result was for the classification into 4 classes, with an accuracy of 92.86%.

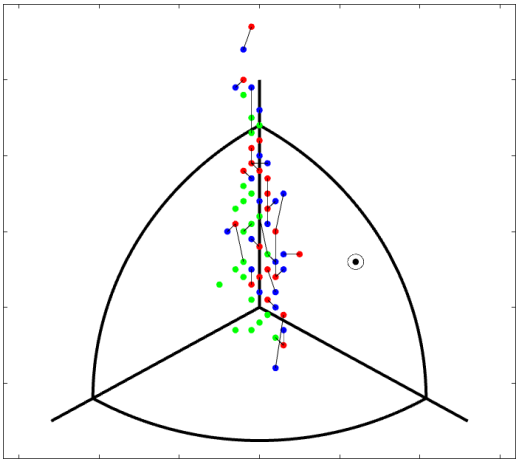


Figure 9. Real somatotypes (in blue) and estimated ones (in red). Green points represent estimates without significant errors.

5. CONCLUSION

In this work, we proposed a method to automatically estimate the somatotype of the individuals by digital images. The proposed methodology includes: 1) protocol for image acquisition; 2) identification of the most relevant body measurements; 3) automatic segmentation and extraction of measurements from the body; 4) estimate of somatotype using the extracted measurements.

The experiments have shown the approach is effective, mainly when the values are represented in multiples of 0.5, and for this situation many estimates were without errors. Regarding classification, the accuracy was higher than 92% for the case of 4 classes.

Thus, the research provided the development of a system that estimates with acceptable precision the somatotype of bodybuilders based on only two images, being a faster option, without needing physical contact with the individual, equipment, and specialists, costing much less than the current method. The main drawback is the necessity of a controlled environment to take photos. Future work can be done to improve the work by (i) apply the method in a larger and more diversified dataset; (ii) use of other methods, such as deep learning, for segmentation of the human body in a non controlled environment.

6. ACKNOWLEDGEMENT

The authors thank the International Federation of Bodybuilding (IFBB) of the State of Espírito Santo - Brazil in the person of its president Andreia Borges for allowing the research with the measurements and use of the images. We also thank the financial support for the research from the project of the Fundação de Amparo à Pesquisa do Espírito Santo (FAPES), number 598/2018.

REFERENCES

Bolonchuk, W.W., Siders, W.A., Lykken, G.I., and Lukaski, H.C. (2000). Association of dominant somatotype of men with body structure, function during exercise, and nutritional assessment. *American Journal of Human Biology*, 12(2), 167–180. doi:10.1002/(SICI)1520-6300(200003/04)12:2<167::AID-AJHB2>3.0.CO;2-3.

Carter, J.L. and Heath, B.H. (1990). *Somatotyping: development and applications*, volume 5. Cambridge University Press.

Carter, J. (2002). The heath-carter anthropometric somatotype - instruction manual. Technical report, San Diego State University.

Deb, K. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, Inc., New York, NY, USA.

Fernandes Filho, J. (2003). *A prática da avaliação física: testes, medidas, avaliação física em escolares, atletas e academias de ginástica*, volume 2. Shape.

Gonçalves, T.d.O., Ciarelli, P.M., Coco, K.F., Gonçalves, E.C., and Castro, J.W.B. (2016). Uma análise da viabilidade de estimar somatotipos de fisiculturistas através de medições em imagens. *Simpósio Brasileiro de Automação Inteligente (SBAI)*.

Gonzalez, R. and Woods, R. (2011). *Processamento Digital De Imagens*. ADDISON WESLEY BRA. URL `https://books.google.com.br/books?id=r5f0RgAACAAJ`.

Gutnik, B., Zuoza, A., Zuoziene, I., Alekrinskis, A., Nash, D., and Scherbina, S. (2015). Body physique and dominant somatotype in elite and low-profile athletes with different specializations. *Medicina (Lithuania)*, 51(4), 247–252. doi:10.1016/j.medici.2015.07.003.

Heath, B.H. and Carter, J.E.L. (1967). A modified somatotype method. *American Journal of Physical Anthropology*, 27(1), 57–74. doi:10.1002/ajpa.1330270108. URL `http:https://dx.doi.org/10.1002/ajpa.1330270108`.

Klafke, B.P. (2018). Sistema para detecção de estradas e obstáculos baseado em imagens rgb e nuvem de pontos para equipamentos de mineração. *Simpósio Brasileiro de Automação Inteligente (SBAI)*.

Massidda, M., Toselli, S., Brasili, P., and Calò, C.M. (2013). Somatotype of elite Italian gymnasts. *Collegium Antropologicum*, 37(3), 853–857.

Nock, R. and Nielsen, F. (2004). Statistical region merging. *IEEE Transactions on pattern analysis and machine intelligence*, 26(11), 1452–1458.

Santos, A.A. (2020). Sistema automático para a inspeção visual de transportadores de correia por meio de redes neurais convolucionais. *Simpósio Brasileiro de Automação Inteligente (SBAI)*.

Sheldon, W.H., Stevens, S.S., and Tucker, W.B. (1940). The varieties of human physique: An introduction to constitutional psychology.

Thorland, W.G., Johnson, G.O., Fagot, T.G., Tharp, G.D., and Hammer, R.W. (1980). Body composition and somatotype characteristics of junior olympic athletes. *Medicine and Science in Sports and Exercise*, 13(5), 332–338.

Tóth, T., Michalíková, M., Bednarčíková, L., Živčák, J., and Kneppo, P. (2014). Somatotypes in sport. *Acta Mechanica et Automatica*, 8(1), 27–32. doi:10.2478/ama-2014-0005.

Witten, I.H. and Frank, E. (2005). *Data mining: practical machine learning tools and techniques*. Elsevier Inc., San Francisco, CA, USA, 2nd edition.