ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/331723086

Deep learning-based detection of anthropometric landmarks in 3D infants head models

Conference Paper · March 2019

DOI: 10.1117/12.2512196

CITATIONS
2

READS
91

10 authors, including:

Helena Torres
University of Minho
10 PUBLICATIONS **41** CITATIONS
SEE PROFILE

Bruno Oliveira
University of Minho
14 PUBLICATIONS **102** CITATIONS
SEE PROFILE

Fernando Veloso
University of Minho
11 PUBLICATIONS **18** CITATIONS
SEE PROFILE

António H. J. Moreira
University of Minho
31 PUBLICATIONS **193** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Guidance of transseptal punctures for left heart interventions using personalized biomechamical models and volumetric ultrasound imaging View project

Clinical Pediatric Surgery View project

All content following this page was uploaded by Fernando Veloso on 29 June 2020.

The user has requested enhancement of the downloaded file.

Deep Learning-based Detection of Anthropometric Landmarks in 3D Infants Head Models

Helena R. Torres[a,b] , Bruno Oliveira[a,b], Fernando Veloso[a,b], Mario Ruediger[c], Wolfram Burkhardt[c], António Moreira[a], Nuno Dias[a], Pedro Morais[a], Jaime C. Fonseca[b], João L. Vilaça[a]

[a]2Ai – Polytechnic Institute of Cávado and Ave, Barcelos, Portugal
[b]Algoritmi Center, School of Engineering, University of Minho, Guimarães, Portugal
[c]Department for Neonatology and Pediatric Intensive Care, Children's Hospital, Medical Faculty of TU Dresden, Germany

ABSTRACT

Deformational plagiocephaly (DP) is a cranial deformity characterized by an asymmetrical distortion of an infant's skull. The diagnosis and evaluation of DP are performed using cranial asymmetry indexes obtained from cranial measurements, which can be estimated using anthropometric landmarks of the infant's head. However, manual labeling of these landmarks is a time-consuming and tedious task, being also prone to observer variability. In this paper, a novel framework to automatically detect anthropometric landmarks of 3D infant's head models is described. The proposed method is divided into two stages: (i) unfolding of the 3D head model surface; and (ii) landmarks' detection through a deep learning strategy. In the first stage, an unfolding strategy is used to transform the 3D mesh of the head model to a flattened 2D version of it. From the flattened mesh, three 2D informational maps are generated using specific head characteristics. In the second stage, a deep learning strategy is used to detect the anthropometric landmarks in a 3-channel image constructed using the combination of informational maps. The proposed framework was validated in fifteen 3D synthetic models of infant's head, being achieved, in average for all landmarks, a mean distance error of 3.5 mm between the automatic detection and a manually constructed ground-truth. Moreover, the estimated cranial measurements were comparable to the ones obtained manually, without statistically significant differences between them for most of the indexes. The obtained results demonstrated the good performance of the proposed method, showing the potential of this framework in clinical practice.

Keywords: anthropometric landmark detection, deep learning, deformational plagiocephaly, mesh unfolding

1. INTRODUCTION

Deformational plagiocephaly (DP) refers to a cranial deformity occurring in infants. This condition is characterized by an occipital flattening of the skull due to external pressures, resulting in cranial asymmetry [1], [2]. Identified causes include restrictive intrauterine environment, unusual birth position, supine sleeping, and bottle-feeding positions [3],[4]. Moreover, preterm birth is also associated to DP. In fact, preterm infants show high probability of developing head deformities, being the prevalence of DP in very preterm infants 38% worldwide [5]. Besides the DP-related cosmetic concerns, infants with this condition can also experience developmental delay [6],[7], thus, prevention or early treatment of this condition is needed. There are different treatment options for DP, such as repositioning, physical therapy, or cranial orthosis. To select the best treatment option for DP, accurate diagnosis of DP and assessment of its severity is required. The diagnosis and evaluation are performed based on cranial asymmetry indexes estimated from anthropometric skull measurements. These measurements are usually extracted manually using calipers [8]. However, a high intra- and inter-observer variability of that method can be expected since accuracy of measurement will vary between investigators. Thus, other techniques have been used, namely techniques based on 3D laser scanning [15]. In these strategies, the estimation of the cranial asymmetry indexes is performed using anthropometric landmarks estimated in a 3D model of the infant's head. Since the manual selection of these landmarks is not straight forward, development of automated solutions to extract them are potential and useful tools to normal clinical practice.

This paper presents a framework to automatically detect anthropometric landmarks of 3D head models of infants. The main goal of this work is to remove the user interaction influence in the landmarks' extraction stage, increasing therefore its reproducibility while reducing the processing time. The detected landmarks can be used to estimate several cranial indexes, allowing not only to diagnose DP easily, but to assess infant's head shape to detect other deformities in the future.

2. METHODS

The proposed automatic method is divided into two main parts as can be seen in Figure 1. The first stage concerns the unfolding of the 3D surface of the head model to create a 2D flattened version of it that is later used to create 2D informational maps representative of the main features of the head. The second stage relies on a deep learning strategy to detect the anthropometric landmarks on the maps created in the first stage. Both stages are described in more detail in the next sub-sections.

2.1 Unfolding of 3D head models using geodesic embedding flattening

To unfold a 3D mesh, a specific parameterization that transforms the 3D space of the head model into a 2D one is required. Nevertheless, while performing this transformation, correct parameterization is required in order to minimize distortions in the original mesh. In this work, a traditional geodesic embedding flattening was used, where the flattened mesh was obtained by minimizing the geodesic distortion between the points in the 3D mesh and the points in the parameterized domain [9].

Using this unfolded space, three 2D informational maps with specific features of the head were created (see Figure 1). The first is denominated depth map where its intensity values represent the distance of each point of the mesh from a viewpoint. In this case, this map contains information on the depth of each point of the 3D head mesh when seen from the front, consisting in the coordinate of each point of the mesh in the sagittal direction. The second map (termed as polar map) represents the distance between each point of the 3D mesh and its centroid. Lastly, the third, named curvature map, comprises information of the principal curvature directions in each face of the head mesh, using for that the normal vector at each point of the mesh. By combining these three maps after normalization, a 3-channel image (defined to have a resolution of 368x368 pixels) that represents the main head features is created and used as input for the deep learning step.

2.2 Deep learning-based landmark detection

After creating the 3-channel image, a deep learning strategy is applied to detect the anthropometric landmarks. This detection is performed using the abovementioned 3-channel image. The architecture of the proposed convolutional network can be seen in Figure 2. As shown in the figure, the convolution network is firstly initialized and fine-tuned by the first layers of the VGG-19 [10], which are used to perform a first analysis of the image. The generated feature maps are later introduced in a set of convolutional layers that produces the confidence maps for each anthropometric landmark as suggested in [11]. These confidence maps represent the belief of a landmark occurs in each pixel of the 2D image, and it is computed by applying a Gaussian-like function where the maximum of the gaussian map represents the ideal landmark position.

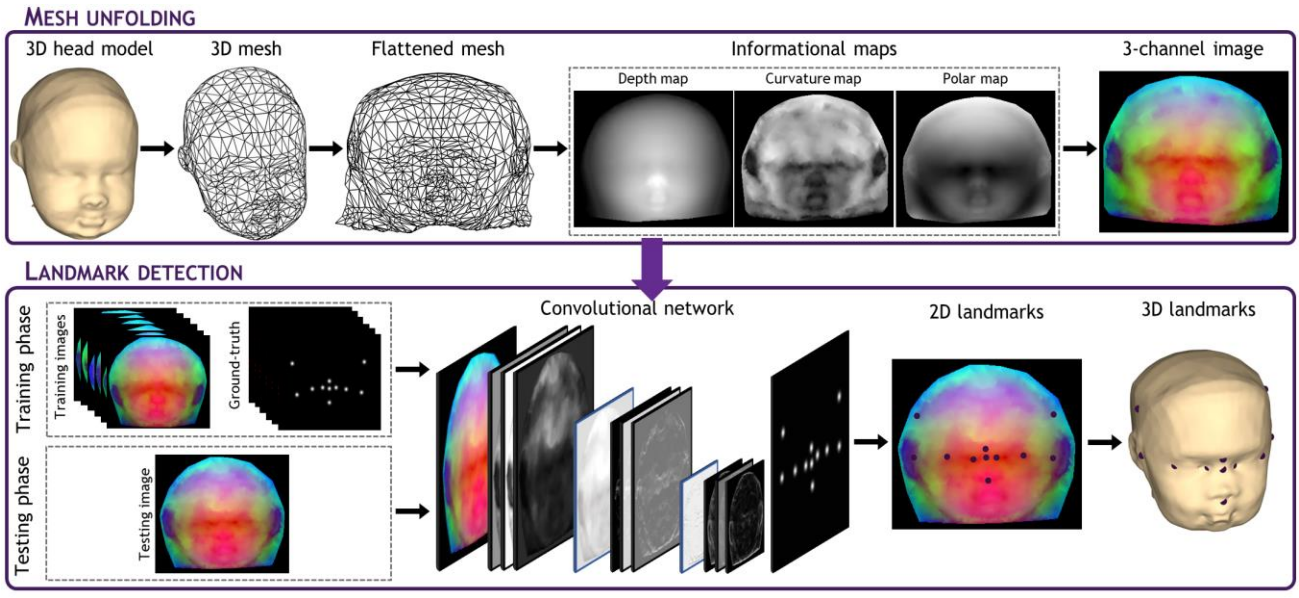

Figure 1 – Overview of the proposed method for the detection of anthropometric landmarks.

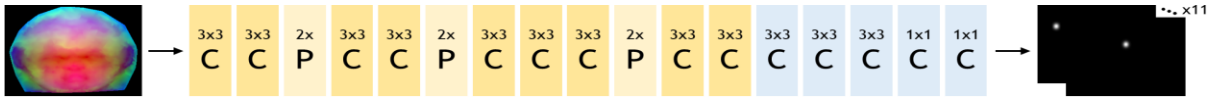

Figure 2 – Architecture of the convolutional network used in the proposed framework. The VGG-19 is represented in yellow colors. The output of the network is confidence maps for each anthropometric landmark.

In the training phase, where the method is trained to predict the confidence maps, a loss function f was applied in the end of the convolutional network to estimate the difference between the predictions and the ideal confidence map, which was constructed by applying a Gaussian-like function around each landmark of the ground-truth (constructed after projection of a manual 3D labeling to the 2D image). This loss function f is given by:

$$f = \sum_{j=1}^{J} ||S_j - S_j^*||, \quad (1)$$

where J represents the number of landmarks, and S_j and S_j^* are the prediction and ground truth maps for landmarks j, respectively.

In the test phase, each landmark (in the unlabeled case) is detected by estimating the respective confidence maps, being the optimal position defined as the maximum of the respective confidence map, *i.e.* its peak, after a non-maximum suppression processing. The obtained landmarks are then transferred to the 3D world, by reverting the parameterization used in the unfolding stage.

3. RESULTS

3.1 Dataset creation

To train and validate the proposed framework, a synthetic dataset of 65 infants' head was created using the *MakeHuman* software [12]. Different full-body infants' models were initially generated using random characteristics, including anatomical features and facial expressions. Afterwards, these models were post-processed by removing several parts while keeping the head only. For each head model, a ground truth was created by manually labeling the anthropometric landmarks. The database was then divided into two subsets, namely training and testing datasets, with 50 and 15 head models, respectively. To increase the size of the training dataset, a data augmentation strategy was applied by scaling the head models, allowing the models to grow/shrink in different axis randomly until 2 cm (Figure 3). After the data augmentation strategy, a training database of 250 head models was available, mimicking, therefore, the required variability in terms of shape, size, and anatomical features.

3.2 Anthropometric landmarks and cranial asymmetry indexes

To compute the asymmetry indexes to evaluate DP, eleven landmarks were used, namely: glabella (GL), sellion (SL), and subnasal (SN) and right and left exocanthions (EX), endocanthions (ED), tragions (TR), and eurions (EU) (Figure 4A) [13]. By defining these landmarks, a head's coordinate system is possible to be estimated (Figure 4B) with [14]:

- the origin being defined as the midpoint between the tragions;
- the X-axis and the Y-axis being defined as the vector from the origin through the left tragion and sellion, respectively;
- the Z-axis being defined as the vector from the origin through the top of the head, such that it is perpendicular to the X- and Y-axis.

Once defined the coordinate system, different measurement planes, parallel to the XY plane, are generated. In this work, the measurement plan that passes through the glabella point was defined.

Figure 3 – Example of one synthetic head model (first model showed in the image) and respective augmented models used in the experiments.

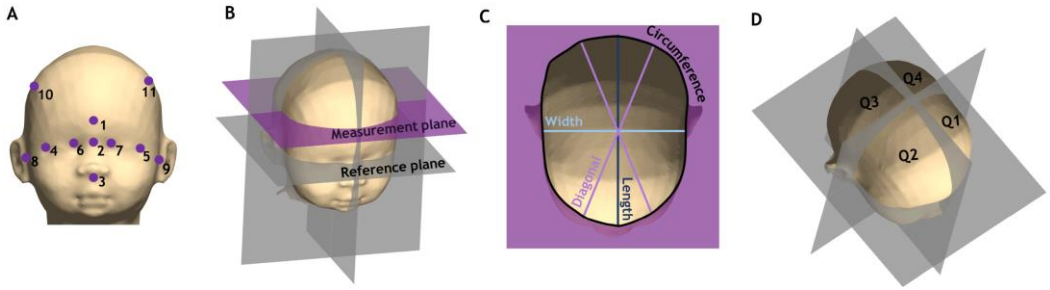

Figure 4 – Head mapping. (A) Infant's head model with anthropometric landmarks: 1-glabella (G), 2-sellion (SL), 3-subnasal (SN), 4-right exocanthion (R. EX), 5-left exocanthion (L. EX), 6-right endocanthion (R. ED), 7-left endocanthion (L. ED), 8-right tragion (R. TR), 9-left tragion (L. TR), 10-right eurion (R. EU), 11-left eurion (L. EU); (B) Head's coordinate system; (C) Cranial measurements computed in the measurement plane; (D) Quadrants of the head model.

The coordinate system and measurement plane were then used to compute the relevant anthropometric measurements (Figure 4C), such as cranial circumference, cranial width, cranial length, oblique transcranial diagonals (measured with an angle of 30º and -30º), and volume of each head quadrant (Figure 4D). Later, these measurements are used to estimate the cranial asymmetry indexes. In this work, four indexes were evaluated, namely the cephalic index (ratio between the cranial width and length), transcranial diagonal difference (difference between transcranial diagonals), anterior cranial asymmetry (ratio of the difference between the largest and smaller anterior quadrant volumes to the smaller anterior quadrant volume), and posterior cranial asymmetry (ratio of the difference between the largest and smaller posterior quadrant volumes to the smaller posterior quadrant volume) [14].

3.3 Framework performance evaluation

To evaluate the proposed method, two experiments were performed and presented in this work: 1) assessment of the accuracy of the automatic landmark estimation approach by comparing the detected landmark position with the ground-truth; 2) a clinical validation by comparing the estimated cranial indexes with the manual ones.

3.3.1 Accuracy assessment

Table 1 summarizes the performance of the proposed detection method in terms of mean distance error between the automatic results for each detected landmark position and the manual ground-truth. Moreover, to compare the detection results with the inter-observer variability, a second observer also performed a manual labeling of the relevant landmarks. Overall, mean errors (for all landmarks) of 3.5 mm were achieved. Distances errors lower than 3 mm were achieved for the majority of landmarks. Distances errors higher than 6 mm were obtained for both eurions only. In fact, considering a distance threshold of 5 mm, an average accuracy of 94% was found for the remaining landmarks, excepting the eurions that obtained an accuracy of 50%. Figure 5 presents example results of the proposed method.

Table 1- Performance of the proposed method to detect the different anthropometric landmarks in terms of distance errors (in mm)

	GL	SL	SN	R. EX	L. EX	R. ED	L. ED	R. TR	L. TR	R. EU	L. EU	Mean
Detection method	2.7±1.3	2.3±1.4	2.1±0.7	3.6±1.8	3.2±1.4	2.9±1.7	2.7±1.9	3.2±1.2	2.6±2.0	6.6±3.1	6.7±4.5	3.5±2.6
Inter-observer	4.5±2.5	2.8±1.3	2.4±1.0	7.0±1.5	5.9±3.2	2.1±0.7	2.4±1.1	2.5±1.5	2.3±1.2	10.5±3.0	8.3±3.2	4.6±1.8

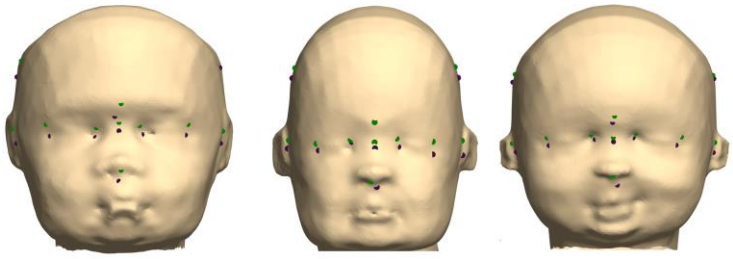

Figure 5- Landmark detection result using the automatic method (dark violet) and ground-truth (green).

3.3.2 Clinical validation

Figure 6 presents the Bland-Altman plot analysis of the automatic method against manual analysis for the selected cranial indexes. Moreover, the difference between observers is also presented. Overall, no statistically significant biases (red dashed line) were obtained by the automatic value for all the measurements ($p>0.05$ in a two-tailed t-test against 0). When assessing the agreement between manual observers and automatic measurements, a similar performance was found for the cephalic index, the transcranial diagonal difference, and the posterior cranial asymmetry index ($p>0.05$ in a two-tailed F-Test against the intra-observer).

4. DISCUSSION

This paper presents a landmark detection approach based on a deep learning strategy to estimate the location of anthropometric landmarks in 3D models of infants' head. Analyzing Table 1, it is possible to verify that the results corroborates the high accuracy of the proposed method, being achieved a mean distance error of 3.5 mm. In fact, the distance errors for each landmark were below 3.5 mm for the majority of the landmarks, except for both eurions. Moreover, considering a distance threshold of 5 mm, an average accuracy of 94% was found for the majority of the joints. It is important to notice that these results were obtained in testing head models that presented cranial widths and lengths ranging between 94 and 112 mm and 117 and 144 mm, respectively. However, it is expected that these errors do not change considerably if testing in head models with different head sizes, once the proposed method was training using heads with variable sizes. The higher errors found for the eurions can be explained by the difficulty to detect specific and unique anatomical features of this landmark, owing to its location in the central area of the parietal side of the head. Indeed, manual labeling of this specific landmark is challenging, which ultimately influence the training stage of the proposed method and the accuracy of the developed strategy. This also hampers the process of manual labeling for these landmarks, which can contribute for a less effective training of the network to predict the eurions and consequently a less robust detection result for these landmarks. Nevertheless, the results clearly prove that the proposed method overperform the inter-observer variability in seven of the eleven landmarks detected, showing a comparable result for the remaining ones. These accurate segmentation results demonstrate the robustness of both stages of the method. On one hand, it is possible to conclude that the information maps created in the first stage of the method correctly represent the main features of the head, allowing to correctly differentiate specific characteristics for the anthropometric landmarks. This is essential to perform a correct prediction using a deep-learning strategy, where specific features for each landmark are searched. On the other hand, the results obtained also suggest that the convolutional network configuration proposed is suitable for prediction of the location of the landmarks.

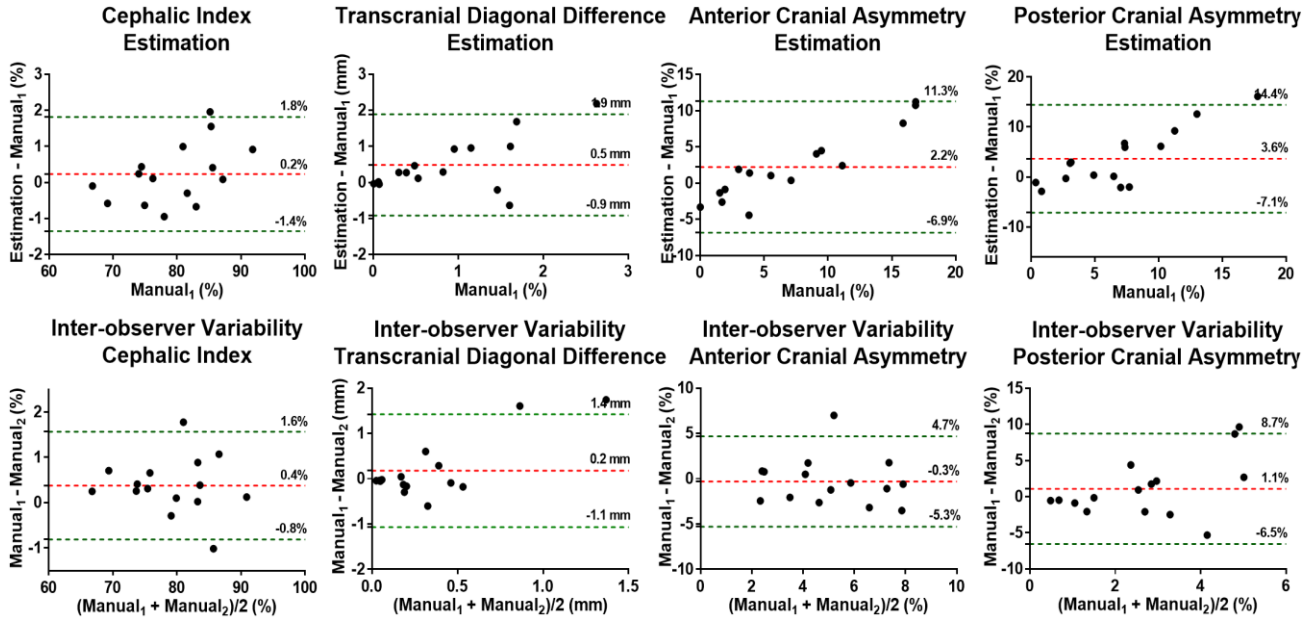

Figure 6- Bland-Altman analysis for the estimated cranial indexes.

Besides the assessment of the accuracy of the automatic landmark estimation, a clinical validation on the defined cranial asymmetry indexes was also performed. To evaluate possible statistically significant biases between the automatic results for the testing set, a two-tailed t-test against 0 was performed, being achieved a $p\text{-}value$ lower than 0.05 for all the measurements. No statistically significant biases (red dashed line in Figure 6) were obtained by the automatic method for all the measurements. Moreover, when assessing the agreement between manual observers and automatic measurements, a similar performance was achieved for most of the cranial indexes, namely the transcranial diagonal difference, and the posterior cranial asymmetry index ($p>0.05$ in a two-tailed F-Test against the intra-observer). These results proved the added-value of the method and its potential to be used in clinical practice. In fact, a high reproducibility in the estimation of the cranial measurements is of high importance in clinical practice. Manual estimation of the cranial measurements is prone to inter-observer variability which hampers diagnosis certainty. In this sense, new methods that rely on labelling landmarks on 3D models of the head (generated using laser scanners) to estimate the measurements have been used. These methods proved to be useful to overpass the lack of reproducibility of the traditional manual estimation of cranial measurements, providing reproducibility data for these indexes [15]. However, the use of these methods only offers a semi-automatic estimation of the measurements, once it is always needed to label manually the landmarks in the 3D model. In these sense, full automatization of the entire process by developing methods to automatically extract the landmarks can present a huge impact in clinical practice. Moreover, based on the current data an entire new field of application opens. The methods does not only allow an accurate description of cranial symmetry but can be also used to quantify mid-face problems, which can be often found in preterm infants after application of respiratory support systems.
Finally, another important aspect to evaluate is the computational burden of the framework. Regarding the first stage of the method, the mesh unfolding process, the creation of the informational maps, and the creation of the 3-channels images take only 20 seconds with a nonoptimized algorithm MATLAB code. Concerning the deep-learning approach, a fast detection was obtained in the experiments (average time of 80 milliseconds per image). Overall, the low computational burden of method shows the clear advantages of the proposed framework against the tedious and time consuming manual estimation of the cranial indexes and manual landmark labeling.

5. CONCLUSIONS

In this paper, a novel deep learning-based technique for detecting anthropometric landmarks in 3D infants head models is proposed. This method consisted in a two-stage approach where a set of informational maps representative of the main features of the head are firstly generated, being afterwards used to create 3D-channel images that are used as input for a convolutional network. The detection method proved its high accuracy in 15 3D head models, with a mean error of 3.5 mm. Moreover, the relevant clinical cranial measurements showed a performance comparable to the inter-observer ones. As a future work, it is expected to increase the size and variability of the training dataset, namely by using head models from different age groups, in order to increase the accuracy of the method. Moreover, it is expected to evaluate the method's performance in an exhaustive clinical database, ultimately corroborating the added-value of this methodology.

ACKNOWLEDGEMENTS

The present submission corresponds to original research work of the authors and has never been submitted elsewhere. Moreover, this work was funded by the project NORTE-01-0145-FEDER-024300, supported by Northern Portugal Regional Operational Programme (Norte2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (FEDER). Moreover, this work has been also supported by FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2019. Furthermore, the authors acknowledge FCT, Portugal, and the European Social Found, European Union, for funding support through the "Programa Operacional Capital Humano" (POCH) in the scope of the PhD grants SFRH/BD/136670/2018 (Helena R. Torres), SFRH/BD/136721/2018 (Bruno Oliveira), and SFRH/BD/131545/2017 (Fernando Veloso).

REFERENCES

[1] S. Robinson and M. Proctor, "Diagnosis and management of deformational plagiocephaly," *J. Neurosurg. Pediatr.*, vol. 3, no. 4, pp. 284–295, 2009.

[2] Kanlaya Ditthakasem and J. C. K. D, "Deformational Plagiocephaly: A Review," *Pedriatric Nurs.*, vol. 43, no. 2,

pp. 59–65, 2017.
[3] L. A. van Vlimmeren, Y. van der Graaf, M. M. Boere-Boonekamp, M. P. L'Hoir, P. J. M. Helders, and R. H. H. Engelbert, "Risk Factors for Deformational Plagiocephaly at Birth and at 7 Weeks of Age: A Prospective Cohort Study," *Pediatrics*, vol. 119, no. 2, pp. e408–e418, 2007.
[4] F. De Bock, V. Braun, and H. Renz-Polster, "Deformational plagiocephaly in normal infants: A systematic review of causes and hypotheses," *Arch. Dis. Child.*, vol. 102, no. 6, pp. 535–542, 2017.
[5] S. Ifflaender, M. Rüdiger, D. Konstantelos, K. Wahls, and W. Burkhardt, "Prevalence of head deformities in preterm infants at term equivalent age," *Early Hum. Dev.*, vol. 89, no. 12, pp. 1041–1047, 2013.
[6] A. L. C. Martiniuk, C. Vujovich-Dunn, M. Park, W. Yu, and B. R. Lucas, "Plagiocephaly and Developmental Delay," *J. Dev. Behav. Pediatr.*, vol. 38, no. 1, pp. 67–78, 2017.
[7] R. I. Miller and S. K. Clarren, "Long-Term Developmental Outcomes in Patients With Deformational Plagiocephaly," *Pediatrics*, vol. 105, no. 2, pp. e26–e26, 2000.
[8] S. Nahles, M. Klein, A. Yacoub, and J. Neyer, "Evaluation of positional plagiocephaly: Conventional anthropometric measurement versus laser scanning method," *J. Cranio-Maxillofacial Surg.*, vol. 46, no. 1, pp. 11–21, 2017.
[9] G. Peyr and L. Cohen, "Geodesic Computations for Fast and Accurate Surface Flattening," *Prog. Nonlinear Differ. Equations Their Appl.*, vol. 63, pp. 157–171, 2005.
[10] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *Computer Vision and Pattern Recognition*, 2014, pp. 1–14.
[11] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," Nov. 2016.
[12] M. Bastioni, S. Re, and S. Misra, "Ideas and methods for modeling 3D human figures: The principal algorithms used by MakeHuman and their implementation in a new approach to parametric modeling," in *Proceedings of the 1st Bangalore Annual Compute Conference*, 2008, pp. 1–6.
[13] J. F. Teichgraeber *et al.*, "Deformational posterior plagiocephaly: Diagnosis and treatment," *Cleft Palate-Craniofacial J.*, vol. 39, no. 6, pp. 582–586, 2002.
[14] P. Meyer-Marcotty *et al.*, "Head orthesis therapy in infants with unilateral positional plagiocephaly: an interdisciplinary approach to broadening the range of orthodontic treatment," *J. Orofac. Orthop. / Fortschritte der Kieferorthopädie*, vol. 73, no. 2, pp. 151–165, 2012.
[15] S. Ifflaender, M. Ru, A. Koch, and W. Burkhardt, "Three-Dimensional Digital Capture of Head Size in Neonates – A Method Evaluation," *PLoS One*, vol. 8, no. 4, pp. 1–6, 2013.

View publication stats