

Wattendorf Sonja*, Tabatabaei Seyed Amir Hossein, Fischer Patrick, Howaldt Hans-Peter, Wilbrand Martina, Wilbrand Jan-Falco and Sohrabi Keywan

Automatic Detection of Pediatric Craniofacial Deformities using Convolutional Neural Networks

Abstract: The geometric shape of our skull is very important, not only from an esthetic perspective, but also from medical viewpoint. However, the lack of designated medical experts and wrong positioning is leading to an increasing number of abnormal head shapes in newborns and infants. To make screening and therapy monitoring for these abnormal shapes easier, we develop a mobile application to automatically detect and quantify such shapes. By making use of modern machine learning technologies like deep learning and transfer learning, we have developed a convolutional neural network for semantic segmentation of bird's-eye view images of child heads. Using this approach, we have been able to achieve a segmentation accuracy of approximately 99 %, while having sensitivity and specificity of above 98 %. Given these promising results, we will use this basis to calculate medical parameters to quantify the skull shape. In addition, we will integrate the proposed model into a mobile application for further validation and usage in a real-world scenario.

Keywords: Deep Learning, Semantic image segmentation, Biomedical signals, Cranial deformities

<https://doi.org/10.1515/cdbme-2020-3087>

***Corresponding author: Wattendorf Sonja:** Faculty of Health Sciences, University of Applied Sciences Giessen, Wiesenstrasse 14, 35390 Gießen, Germany, e-mail: sonja.wattendorf@ges.thm.de

Tabatabaei Seyed Amir Hossein, Fischer Patrick: Institute of Medical Informatics, Faculty of Medicine, Justus-Liebig-University Giessen, Giessen, Germany

Howaldt Hans-Peter, Wilbrand Martina: Department for Cranio-Maxillofacial Surgery – plastic Surgery-, University Hospital Giessen, Giessen, Germany

Wilbrand Jan-Falco: Department for Cranio-Maxillofacial Surgery – plastic Surgery, Diakonie-Klinikum Jung-Stilling, Siegen, Germany

Sohrabi Keywan: Faculty of Health Sciences, University of Applied Sciences Giessen, Giessen, Germany

1 Introduction

Nowadays visual appearance is becoming more and more important in our society. Especially the human face and the shape of the head are aspired to be flawless. But not only the aesthetic aspect, but also the health plays a role in the human skull shape [1]. Ideally, the skull grows normally, which ensures healthy growth of the brain and face. However, if the skull shape is abnormal, different types of skull deformities can happen in newborns and infants [2]. The objective of our project is to develop a mobile application to detect pediatric cranial deformities by simply taking a bird's eye view image of the child's head. Given the recent advances in machine learning, especially in the field of deep learning, it is possible to leverage convolutional neural networks for semantic segmentation. Once the area of interest is segmented, different clinical parameters can be extracted and compared to the current standard in order to assess the necessity of therapeutic intervention [3]. In this paper, we propose the use of deep and transfer learning to achieve highly accurate segmentation results.

2 Material and Methods

2.1 Dataset

More than 300 images were acquired within the consultation hours for craniofacial abnormalities of the Department of Cranio-Maxillofacial Surgery at the University Hospital of Giessen. In order to measure clinical parameters after segmentation, we used a 50-cent coin as a reference, which was placed on the center of the head. Manual image annotations were performed by trained medical experts using Labelme [4].

2.2 Neural Network Architecture

The desired segmented output of classification model is a 3-label segmented image in which the coin, the head and the background are semantically segmented. So, a semantic segmentation task wherein each pixel is assigned a label will be performed. Our network architecture is based on DeepLab v3+ which is a state-of-the-art convolutional neural network (CNN) structure introduced by Google Research [5]. The basis of DeepLab v3+ is ResNet-18 (residual network) wherein one or more layers will be skipped while training [6, p. 374]. Also, by customizing activation functions in ResNet better performance can be achieved. The whole network is a DAGNetwork (directed acyclic graph) with 100 layers. For image segmentation background, head and coin build the three classes for the neural network. Each class has a different colour to show the own relevant region. The complete dataset was split into 60 % as a training set, 20 % as a test set and 20 % as a validation set. The original images are used as input, while the output of the neural network is compared against the manually annotated ground truth data set. Sensitivity, specificity and accuracy are calculated as the main classification parameters indicating the overall classification accuracy of the model over all three classes.

2.3 Implementation

We have utilized DeepLab v3+ [5], the development environment of MATLAB Deep Learning toolbox for the sake of implementation. The DeepLab v3+ is based on the ResNet-18 network architecture [7]. In our model, the ResNet-18 is a pretrained network originated from Deep Learning Toolbox. Beyond the original data set, the training sample size can be augmented using corresponding data augmentation techniques just in case [8]. The network consists of different types of layers. Especially convolutional and addition layers as well as batch normalization have been customized. Rectified Linear Unit function (ReLU) has been mainly utilized as the activation function. The engaged optimizer in our convolutional neural network is a stochastic gradient descent with momentum. The network is converged in six epochs. The regularization term with the factor of 0.005 has been added to reduce possible overfitting event.

2.4 Calculating clinical parameters

First step is to convert the segmented head into a binary image. The row with maximum number of non-zero-value pixels corresponds to the width, the same principle applies to

the length with consideration of the columns. The diagonals at 40° angle are calculated using the cosine theorem. In the second step the ratio of the radius of the segmented coin is calculated and the radius of an original 50ct coin to determine the real clinical parameters.

3 Results

To evaluate the approach, classification parameters including accuracy, specification and sensitivity based on the acquired data have been calculated. According to the results, the accuracy of 99.01 % for the implemented neural network architecture is achieved. The following figure shows the high-level architecture of the implemented convolutional neural network (Figure 1).

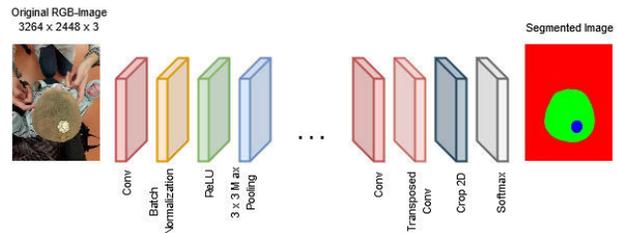


Figure 2: Architecture of the utilized deep network. The input of network is a RGB image with the size of 3264 x 2448 x 3. The network consists of 100 layers. To get a segmented output image the last layer is a softmax, which indicates a pixel classification layer. The output image has the same size as the input image with three labels.

The achieved specificity and sensitivity are 99.46 % and 98.94% respectively. Also, the precision is about 98.94 %. The corresponding 3-class confusion matrix is shown as follows in Figure 2.

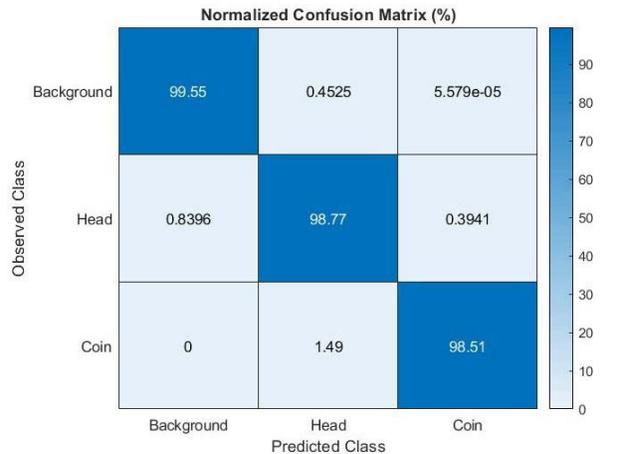


Figure 1: Confusion matrix of the three-class problem. Each pixel is assigned by one of the three classes. The observed and predicted classes are shown on the left side and on the bottom side respectively. Correct classifications are shown on the main diagonal.

For the validation of the calculation of the clinical parameters, we used eight images which were not trained or tested with the neural network. The following figure shows the parameters drawn in (Figure 3).

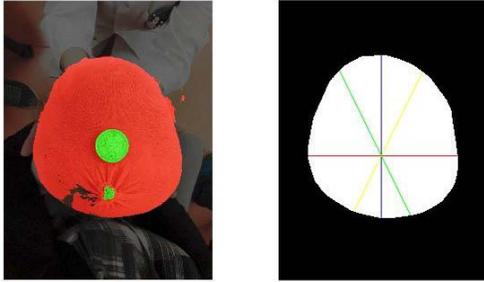


Figure 3: On the left a colored segmented image is shown. On the right side the binary image of the head is shown. The blue line describes the length, the red one the width. The green line is the diagonal A 40° and the yellow one is the diagonal B 40°.

Statistically, the following results are obtained for the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) of the clinical parameters.

Table 1: Overview of the Mean Absolute Error and the Root Mean Square Error

Clinical parameters	Mean Absolute Error (cm)	Root Mean Square Error (cm)
Length	0.64	2.499
Width	0.328	2.14
Diagonal A	0.828	1.995
Diagonal B	0.474	2.867

4 Discussion

In this paper, we proposed the use of deep and transfer learning to tackle a real-world semantic segmentation problem. Based on well-established deep models (ResNet-18 and DeepLab v3+), we have been able to achieve a highly accurate segmentation result. Despite the promising results, additional steps are necessary to provide the usability of the application for physicians and users (patients and care givers). Future work will include the measurement of clinical parameters regarding the shape of skull as well as fine tuning of the whole network and the algorithm for calculating the

clinical parameters. An extensive validation study for the sake of usability of the proposed classification model in the real-world situations will be performed as well.

Author Statement

Research funding: This Project is funded by Dr. Hans-Messer-Stiftung, Bad Soden, Germany.

Conflict of interest: Authors state no conflict of interest.

Informed consent: Informed consent was obtained from legal guardians in all cases.

Ethical approval: The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration, and has been approved by the authors' institutional review board or equivalent committee (AZ 143/16, ethics committee of Justus-Liebig-University Giessen).

References

- [1] D. Filter and J. Reich, Eds., “Bei mir bist du schön...”: kritische Reflexionen über Konzepte von Schönheit und Körperlichkeit. Freiburg im Breisgau: Centaurus-Verl, 2012.
- [2] J. Zöllner and A. Grandoch, “Schädeldeformitäten – Kraniosynostosen und lagebedingte Wachstumsstörungen,” *Pädiatrie up2date*, vol. 12, no. 04, pp. 355–370, Dec. 2017, doi: 10.1055/s-0043-115275.
- [3] J.-F. Wilbrand *et al.*, “Normal Head Shape Parameters in the First 2 Years of Life and Effect of Helmet Therapy,” *Neuropediatrics*, vol. 48, no. 06, pp. 432–441, Dec. 2017, doi: 10.1055/s-0037-1604482.
- [4] Kentaro Wada, “Labelme,” *labelme*, 05-Mar-2020. [Online]. Available: <https://github.com/wkentaro/labelme>.
- [5] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation,” in *ECCV*, 2018.
- [6] A. Géron, *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*, First edition. Beijing; Boston: O’Reilly Media, 2017.
- [7] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *arXiv:1512.03385 [cs]*, Dec. 2015.
- [8] The MathWorks, Inc., “deeplabv3plusLayers,” *deeplabv3plusLayers*, 05-Mar-2020. [Online]. Available: <https://de.mathworks.com/help/vision/ref/deeplabv3pluslayers.html>.