Artificial Intelligence Research and Development
T. Alsinet et al. (Eds.)

doi:10.3233/FAIA240438

Landmark Anything: Multi-View Consensus Convolutional Networks Applied to the 3D Landmarking of Anatomical Structures

Álvaro HEREDIA-LIDÓN [a,1], Christian GARCÍA-MASCARELL [a], Luis Miguel ECHEVERRY-QUICENO [b], Daniel HERRERA-ESCARTÍN [b], Juan FORTEA [e], Edith POMAROL-CLOTET [c,d] Mar FATJÓ-VILAS [c,d], Neus MARTÍNEZ-ABADÍAS [b] and Xavier SEVILLANO [a]

[a] *HER - Human-Environment Research Group, La Salle - Universitat Ramon Llull, Barcelona, Spain*

[b] *Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals (BEECA). Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain*

[c] *FIDMAG, Sisters Hospitallers Research Foundation, Barcelona, Spain*

[d] *CIBERSAM (Biomedical Research Network in Mental Health, Instituto de Salud Carlos III), Madrid, Spain*

[e] *Sant Pau Memory Unit, Hospital de Sant Pau i la Santa Creu, Barcelona, Spain*

Abstract. As shape alterations in three-dimensional biological structures are associated to numerous pathological processes, quantitative shape analysis for obtaining phenotypic biomarkers of diagnostic potential has become a prominent research area. In this context, the automatic detection of landmarks on 3D anatomical structures is crucial for developing high-throughput phenotyping tools. This study evaluates the performance of multi-view consensus convolutional networks – originally developed for facial landmarking– in automatically detecting landmarks on three different 3D anatomical structures: the face, the upper respiratory airways and the brain hippocampi. Leveraging magnetic resonance imaging datasets, we trained multiple models and assessed their accuracy against manual annotations, while analyzing the impact of different network hyperparameters on the results.

Keywords. Automatic 3D landmarking, multi-view convolutional networks, face, upper respiratory airways, hippocampus, biomarkers

1. Introduction

Shape variations in three-dimensional biological structures are associated to many physiological and pathological processes [1]. In this context, the advent of advanced imaging technologies, coupled with computational methods for quantitative shape analysis have

[1]Corresponding Author: Álvaro Heredia-Lidón, alvaro.heredia@salle.url.edu

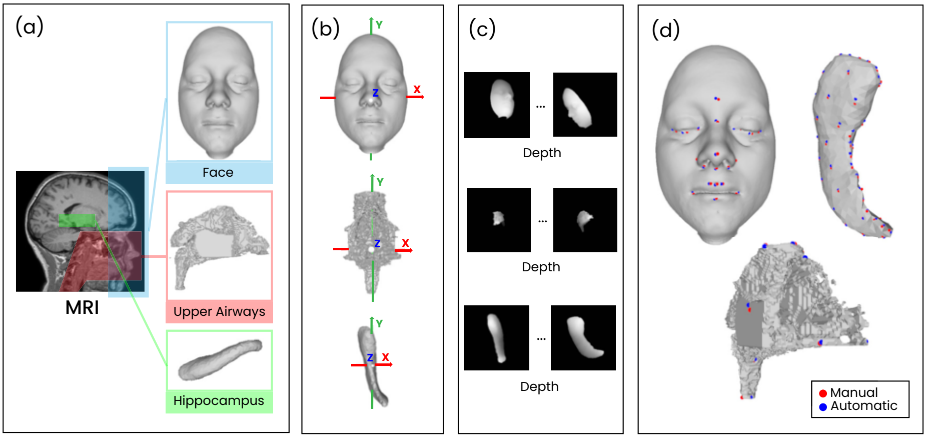

Figure 1. (a) Facial, upper airways and hippocampus 3D models are segmented from a single MRI scan. (b) All anatomical structures are registered to the LSA orientation. (c) Different depth 2D views from the 3D models are extracted for landmarking with MV-CNN. (d) Examples of manual and automatic landmarks placed on the anatomical structures.

driven research to identify phenotypic biomarkers for multiple underlying conditions, like genetic and neurodevelopmental disorders [2,3,4].

In this context, Geometric Morphometrics (GM) has revolutionized biological shape quantification [5]. To use GM methods, the 3D Cartesian coordinates of anatomical landmarks, that correspond to homologous points across individuals, must be registered to encode morphology. However, manual 3D landmarking is a burdensome, error-prone task that requires anatomical expertise and is subject to inter/intra-observer variability. For this reason, there is a need for automatic 3D landmarking algorithms that *i)* perform as accurately as expert anatomists, and *ii)* can be applied to any anatomical structure.

While numerous 3D automatic landmarking methods exist in the literature [6], most are designed and optimized for a single anatomical structure, typically the face [7], with deep learning-based methods demonstrating the highest accuracy [8].

In this work, we explore the generalization ability of multi-view consensus convolutional neural networks (MV-CNN) [8], a deep learning state-of-the-art method devised for automatic 3D facial landmarking. The novelty of this work is the use of MV-CNN to register landmarks on three different anatomical structures: the face, the upper respiratory airways, and the brain hippocampus. To that end, we trained specific MV-CNN models for each anatomic structure, leveraging manually landmarked datasets of 3D models segmented and reconstructed from magnetic resonance imaging (MRI) scans (see Figure 1a). By tuning the training hyperparameters appropriately, we have successfully applied this method to all these anatomical structures, achieving average landmarking errors below 2 mm in all of them, comparable to the accuracy of expert morphologists [9].

2. Multi-view Convolutional Neural Networks Architecture and Training

This automatic landmarking method is based on creating 2D views of the 3D anatomical models from different viewpoints. Then, the hourglass model is used to predict candidate

Anatomical structure	**Number of landmarks**	**Training hyperparameters**		**Landmarking error (mm)**
		XYZ rotation angles	**Views**	
Face	20	[±40°,±80°,±20°]	25	**1.78 ± 0.38**
			100	1.83 ± 0.40
Upper airways	15	[±40°,±80°,±20°]	25	**1.84± 0.75**
			100	1.87 ± 0.76
Hippocampus	60	[±40°,±180°,±20°]	25	0.48 ± 0.08
			100	**0.44 ± 0.08**

Table 1. Data, models and training hyperparameters of the MV-CNN, and landmarking error in mm.

landmarks on 2D heat maps, and the final position is determined by ray-projecting onto the surface of the 3D mesh using least squares combined with random sample consensus (RANSAC) [8]. This 2D multi-view approach enables 3D landmark registration with high precision while maintaining a reduced computational cost.

For model training, we used three datasets of anatomical structures that were manually segmented from MRI scans[2] and landmarked by expert morphologists. The datasets comprised: *i)* 538 facial meshes (20 landmarks), *ii)* 250 upper airways (15 landmarks), and *iii)* 356 hippocampi (60 landmarks and semilandmarks). All 3D models were preprocessed (including mesh cleaning, centering to the origin, and resizing) and registered to the Left-Superior-Anterior (LSA) orientation with PyMeshlab (see Figure 1b).

As for the training hyperparameters, they were optimized heuristically after several experiments for each anatomical structure. These parameters included the number of views (i.e. 2D projections of the 3D model, see Figure 1c) and XYZ camera rotation angles range for projections creation (see Table 1). Texture and color information were not considered in this study.

3. Experiments and results

Table 1 presents the MV-CNN hyperparameter configurations that yielded the highest landmarking accuracy for each model. All experiments[3] used a 90%-10% training/test split. At inference time, the landmarks coordinates were obtained as the average of 10 predictions with a maximum acceptable RANSAC value of 5. The metric used to evaluate landmarking accuracy is the average Euclidean distance between manually and automatically registered landmarks. The rightmost column of Table 1 shows the landmarking error (in mm) for each anatomical structure (average value ± standard deviation).

From a quantitative standpoint, both the facial and the upper airways models attain average landmarking error lower than 2mm, an accuracy below the intra-observer variation accepted in craniometric measurements [9]. The slightly lower accuracy of the airways model could be due to the poorer quality of the meshes obtained from the manual segmentation from MRIs. In both models, most landmarks are located frontally or

[2]High-resolution T1w head MRI scans provided by Hospital Sant Joan de Déu (1.5 GE Sigma scanner), Fundació Pasqual Maragall (Siemens Magnetom PRISMA 3T), and Hospital Sant Pau Memory Unit (Philips 3 Tesla X Series Achieva), all from Barcelona (Spain). All the participants signed an informed consent approved by the Bioethics Committes of the participant institutions.

[3]Run on a PC with i9-10980XE CPU @ 3.00GHz × 36 cores and NVIDIA GeForce RTX 2080 Ti GPU.

slightly laterally, allowing for a reduced XYZ angle range configuration and a limited number of views (25), as additional views did not yield significant improvements.

In contrast, the hippocampus model exhibited an error below 0.5 mm. Considering that this organ is around 5 cm long, this accuracy is impressive despite its convex-shaped geometry lacking prominent reference points and the number of semilandmarks distributed over its entire surface. However, this model required an increased range of rotation angles (±180° in the Y-axis) and a higher number of views (100) for the MV-CNN to successfully place landmarks all over the hippocampus. For a qualitative evaluation, Figure 1d compares automatic and manual landmarks on the three anatomical structures.

4. Conclusions

This work has evaluated the performance of a deep learning-based method originally devised for 3D facial landmarking to automatically register landmarks on three fairly different anatomical structures, two of which had never been tested before. Thanks to the flexibility for tuning hyperparameters during its training, we have demonstrated the generic nature and high accuracy of multi-view convolutional neural networks for 3D automatic landmarking, validating its applicability in a wide variety of phenotyping pipelines.

Acknowledgements The research in this paper was supported by the Joan Oró grant (2023 FI-2 00160) from the Recerca i Universitats Departament (DRU) of the Generalitat de Catalunya with grant 2023 FI-2 00160 and the European Social Fund, by Ministerio de Ciencia, Innovación y Universidades with grant PID2020-113609RB-C21, and by Fondation Jerome Lejeune with grant 2020b cycle-Project No.2001. The authors would also like to thank the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) of the Generalitat de Catalunya (2021 SGR01396, 2021 SGR00706, 2021 SGR1475).

References

[1] Dryden IL, Mardia KV. Statistical shape analysis: with applications in R (Vol. 995). John Wiley & Sons. 2016

[2] Hallgrímsson B, et al. Automated syndrome diagnosis by three-dimensional facial imaging. Genet Med. 2020;22(1010):1682–1693. doi:10.1038/s41436-020-0845-y.

[3] Heuzé Y, Caballol Bachs E, Echeverry LM, Salgado Pineda P, Fortea J, Martínez-Abadías N. Could hippocampus shape be a reliable non-invasive biomarker of Alzheimer disease in Down syndrome. Poster presented at: Anatomy Connected 2023; 2023; Washington D.C., USA.

[4] Martínez-Abadías N, Echeverry LM, Sevillano X, Giménez S, Fortea J, Heuzé Y. Shape and volume features of the upper airways and the tongue: useful biomarkers to improve sleeping apnea diagnosis in Down syndrome? Poster presented at: Anatomy Connected 2023; 2023; Washington D.C., USA.

[5] Mitteroecker P, Schaefer K. Thirty years of geometric morphometrics: Achievements, challenges, and the ongoing quest for biological meaningfulness. Am J Biol Anthropol. 2022;178:181–210.

[6] Porto A, Rolfe S, Maga AM. ALPACA: A fast and accurate computer vision approach for automated landmarking of three-dimensional biological structures. Methods Ecol Evol. 2021;12(11):2129-2144.

[7] Berends B, Bielevelt F, Schreurs R, Vinayahalingam S, Maal T, de Jong G. Fully automated landmarking and facial segmentation on 3D photographs. Sci Rep. 2024 Mar;14(1):6463.

[8] Paulsen RR, et al. Multi-view consensus CNN for 3D facial landmark placement. In: Lecture Notes in Computer Science. Springer International Publishing; 2019. p. 706–719.

[9] Stull KE, Tise ML, Ali Z, Fowler DR. Accuracy and reliability of measurements obtained from computed tomography 3D volume rendered images. Forensic Sci Int. 2014;238:133–140.