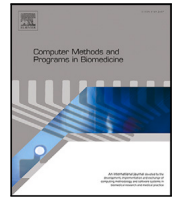




Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: <https://www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine>



BioFace3D: An end-to-end open-source software for automated extraction of potential 3D facial biomarkers from MRI scans

Álvaro Heredia-Lidón^a, Luis M. Echeverry-Quiceno^b, Alejandro González^a,
Noemí Hostalet^{b,c,d}, Edith Pomarol-Clotet^{c,d,1}, Juan Fortea^e, Mar Fatjó-Vilas^{b,c,d},
Neus Martínez-Abadías^b, Xavier Sevillano^a, for the Alzheimer's Disease
Neuroimaging Initiative²

^a Human-Environment Research Group, La Salle - Universitat Ramon Llull, Barcelona, 08022, Spain

^b Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals (BEECA), Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, 08028, Spain

^c FIDMAG, Germanes Hospitalàries Research Foundation, Barcelona, 08035, Spain

^d CIBERSAM (Biomedical Research Network in Mental Health), Instituto de Salud Carlos III, Madrid, 28029, Spain

^e Sant Pau Memory Unit, Hospital de Sant Pau i la Santa Creu, Barcelona, 08025, Spain

ARTICLE INFO

Keywords:

3D facial reconstruction
MRI
Facial biomarkers
3D landmarking
Geometric morphometrics
Software

ABSTRACT

Background and Objectives: Facial dysmorphologies have emerged as potential critical indicators in the diagnosis and prognosis of genetic, psychotic, and rare disorders. While some conditions present with severe dysmorphologies, others exhibit subtler traits that may not be perceivable to the human eye, requiring the use of precise quantitative tools for accurate identification. Manual annotation remains time-consuming and prone to inter- and intra-observer variability. Existing tools provide partial solutions, but no end-to-end automated pipeline integrates the full process of 3D facial biomarker extraction from magnetic resonance imaging.

Methods and Results: We introduce BioFace3D, an open-source pipeline designed to automate the discovery of potential 3D facial biomarkers from magnetic resonance imaging. BioFace3D consists of three automated modules: (i) 3D facial model extraction from magnetic resonance images, (ii) deep learning-based registration of homologous anatomical landmarks, and (iii) computation of geometric morphometric biomarkers from landmark coordinates.

Conclusions: The evaluation of BioFace3D is performed both at a global level and within each individual module, through a series of exhaustive experiments using proprietary and public datasets, demonstrating the robustness and reliability of the results obtained by the tool. Source code, along with trained models, can be found at https://bitbucket.org/cv_her_lasalle/bioface3d.

1. Introduction

The study of facial morphology has gained increasing relevance in biomedical research due to its potential as a diagnostic and prognostic biomarker for genetic, neurodevelopmental, and psychiatric disorders [1–6]. The ability to extract quantitative facial biomarkers provides new opportunities for early diagnosis and prognosis of these conditions. However, current approaches often require using multiple programs as well as manual interventions, which are labor-intensive and susceptible to observer-related variability.

Several computational methods have been developed to address specific steps in the 3D facial analysis workflow [1,7–12]. For instance, MeshMonk [7] enables automated dense surface registration and alignment of anatomical facial landmarks. 3DMedX[®] (3D Lab Radboudumc, Nijmegen, The Netherlands) enables the customization of automated workflows, including 3D volume rendering from magnetic resonance imaging (MRI) and automatic landmark registration. However, as a commercial tool, it does not offer an open-access implementation. 3D Slicer [8], a widely used open-source medical imaging platform, provides plug-ins for 3D volume rendering, segmentation,

* Corresponding author.

E-mail address: alvaro.heredia@salle.url.edu (Á. Heredia-Lidón).

¹ Data used in the preparation of this article were obtained from FIDMAG Germanes Hospitalàries Research Foundation.

² Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

and template-based landmarking with ALPACA [9], among others. Despite its flexibility, Slicer's modules operate independently, requiring significant manual intervention to build coherent workflows.

Recent studies have also proposed multiple approaches for automatic 3D facial landmark detection [13]. However, many of these methods do not publicly release their code or pretrained models. Cliniface [11], an open-access tool built on top of MeshMonk, supports automatic landmark detection and the extraction of standardized facial measurements, such as distances, angles, and asymmetry indices. Nonetheless, it is not compatible with textureless 3D facial reconstructions like those derived from medical imaging such as MRI. Additionally, tools like MorphoJ [12] or the Geomorph R package [14] enable geometric morphometric (GM) analysis for computing potential biomarkers based on the Cartesian coordinates of anatomical landmarks.

In summary, while each of these tools contributes valuable functionality, there is currently no fully integrated pipeline capable of automating the entire process. Therefore, a unified, end-to-end computational tool is needed to support robust and reproducible facial phenotyping from medical images, especially in large-scale or longitudinal biomedical studies.

In this work, we present BioFace3D, an end-to-end open-source software tool for fully automated extraction of 3D facial biomarkers from MRI data. BioFace3D introduces a modular pipeline that performs robust segmentation, facial mesh reconstruction, anatomical landmark detection, and morphometric facial features computation. The system enables the direct application of geometric morphometrics to clinical MRI data without requiring specialized imaging or manual annotation (https://bitbucket.org/cv_her_lasalle/bioface3d).

1.1. Facial biomarkers and their biomedical relevance

Facial dysmorphologies are increasingly recognized indicators of genetic and neurodevelopmental conditions. Many disorders present distinctive craniofacial patterns, often correlated with brain structural abnormalities and sex-specific characteristics [6]. While some conditions, such as Down syndrome, Fragile-X or Morquio syndromes, exhibit severe and easily recognizable dysmorphologies [15–17], others, such as autism spectrum disorder [18] and psychotic disorders [3,19], display subtle facial variations that are challenging to detect without computational tools [20].

Given the high prevalence and diagnostic potential of these dysmorphologies, especially in rare disorders, there is increasing interest in leveraging facial biomarkers for clinical applications [21]. Current diagnostic tools, such as Face2Gene [1] or GestaltMatcher [22], rely on 2D facial images but do not incorporate 3D shape information, which may introduce biases in morphological assessments [2]. Furthermore, these tools often require additional patient data to improve diagnostic accuracy [23], highlighting the need for objective, automated tools capable of analyzing 3D facial structure directly from medical imaging data.

While various 3D imaging techniques – such as stereophotogrammetry [24], laser scanning [25], and structured light cameras [2,26] – enable high-precision facial shape analysis, MRI-based facial biomarker extraction remains an underexplored avenue. The advantage of MRI lies in its ability to provide high-resolution anatomical data, allowing for the extraction of facial surfaces while remaining compatible with brain imaging studies, which is key in neurodevelopmental disorders research [6], since alterations in brain development often manifest in distinct craniofacial phenotypes [27,28].

A well-established methodology for quantifying shape variation, including facial, is geometric morphometrics [2]. GM provides powerful statistical tools to analyze 3D morphological structures with high precision [29,30]. It relies on anatomical landmark coordinates [31,32] to extract global and local shape information, enabling a detailed comparison of shape features between groups. Two commonly used GM

techniques include Generalized Procrustes Analysis (GPA) [33], which aligns individual faces into a common morphospace for global shape comparisons [34], and Euclidean Distance Matrix Analysis (EDMA) [35], which detects localized shape variations by comparing distances between landmarks. Although these techniques are widely used in craniofacial research, no computational pipeline has fully automated their application in MRI-derived facial biomarker analysis.

1.2. BioFace3D: An end-to-end solution for MRI-based facial biomarker extraction

To address this challenge, we introduce BioFace3D, a high-throughput, open-source computational pipeline for the automated extraction and analysis of 3D facial biomarkers from full-head MRI scans. Unlike previous approaches that require manual intervention at multiple stages, BioFace3D fully automates the entire workflow, ensuring efficiency, reproducibility, and scalability.

The innovation of BioFace3D lies in integrating multiple well-established techniques into a single, unified pipeline. This platform could be used by researchers and clinicians to conduct large-scale studies on MRI-derived facial morphology with minimal manual effort, reducing observer-related variability and enhancing diagnostic potential. Furthermore, the potential biomarkers calculated by the tool can be used as feature vectors of clinical value in machine learning-based diagnostic classification tasks.

Specifically, BioFace3D performs the following key functions:

- Automated extraction of 3D facial isosurfaces from MRI scans, enabling precise reconstruction of facial morphology.
- Deep learning based automated anatomical landmark registration, ensuring consistency and reproducibility across large datasets.
- Calculation of potential GM-based facial biomarkers, providing both global and local shape descriptors for statistical analysis.

Another key feature of BioFace3D is its modularity, which allows to integrate new modules and apply the pipeline to any type of 3D facial imaging modality, extending its applicability beyond MRI. By offering an end-to-end, publicly available solution, BioFace3D represents a step forward in standardizing and automating the study of facial biomarkers from MRI data.

As a proof of concept, we demonstrate the diagnostic potential of the facial biomarkers obtained by BioFace3D on a Down syndrome classification case study, achieving competitive classification accuracy only using facial features. These results illustrate and support the applicability of BioFace3D in clinical diagnostic scenarios, as a tool that facilitates translational research in genetic, neurodevelopmental and rare disorders from neuroimaging data.

2. Materials and methods

2.1. Datasets

2.1.1. Training datasets

Training the algorithms in BioFace3D required the manual processing and annotation of a proprietary dataset of head MRI scans. This dataset included 339 high-resolution T1-weighted (T1w) head MRI scans of adult individuals from both sexes provided by: (i) FIDMAG Germanes Hospitalàries Research Foundation (obtained with a 1.5T GE Sigma scanner at Hospital Sant Joan de Déu) and (ii) Sant Pau Memory Unit from Hospital del Mar (Philips 3 Tesla X Series Achieva) and Hospital Clinic (3T Siemens Prisma scanner), all from Barcelona, Spain. Each participant was scanned at only one of these centers and signed an informed consent to use their facial images for research purposes. The study adhered to the guidelines set by the participating institutions and was approved by their respective ethics committees (PR-2023–21,

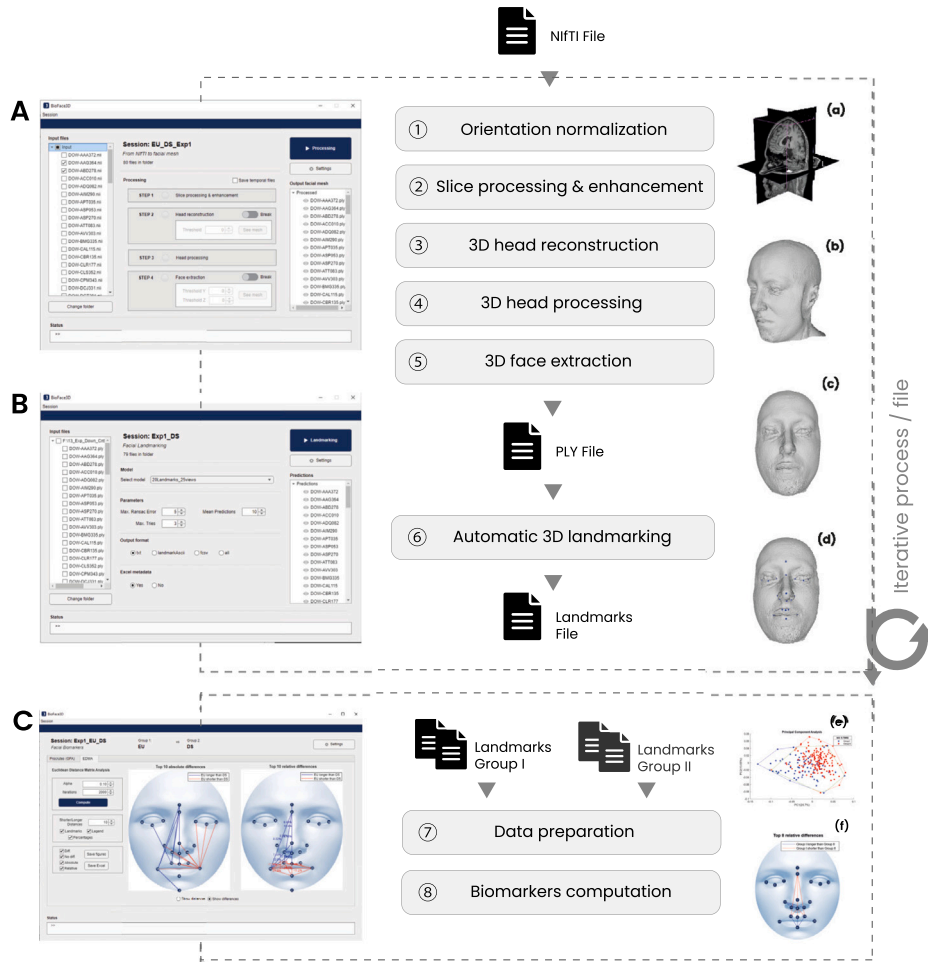


Fig. 1. BioFace3D pipeline. (A) Module for MRI processing and 3D facial mesh extraction. ① From an (a) MRI, the orientation of the head is normalized to a common direction. ② Image processing and enhancement to correct for intensity variations along the MRI 2D slices. ③ Estimation of the optimal facial soft tissue segmentation threshold and reconstruction of the (b) 3D head mesh. ④ Processing of the 3D head for centering, emptying and alignment. ⑤ Extraction and remesh of the (c) 3D facial mesh and storage as a PLY file. (B) Module of 3D facial landmarking. From the PLY file, ⑥ an automatic prediction of the (d) 3D coordinates of the landmarks is performed with multi-view consensus CNNs [10] models. (C) Module for facial biomarkers computation between groups. ⑦ From the landmark files, data are prepared for biomarker extraction. ⑧ Computation of facial biomarkers based on GM techniques. (e) GPA and (f) EDMA biomarkers identify statistically significant global and local differences in facial morphology between groups.

IRB00003099-CER122317 and CER URL_2023_2024_012). All procedures were carried out according to the Declaration of Helsinki. A summary of the dataset composition is presented in the Supplementary Material.

From the NIfTI files derived from each MRI scan, 3D head reconstructions were manually generated using the 3D scientific data analysis software Amira 2019.2 (Thermo Fisher Scientific, Waltham, MA, USA). Head isosurfaces were created by adjusting a gray-scale threshold value that optimized skin segmentation for each participant and facial region. The result of this manual segmentation was used to train the automatic head reconstruction algorithms in BioFace3D.

Additionally, two trained researchers manually registered the 3D coordinates of anatomical facial landmarks on the reconstructed surfaces using the Amira software. This landmark selection included homologous landmarks that could be reliably located on the facial structures of all individuals in the sample, and have been used and validated in previous works [6,17,24] — see the Supplementary Material for landmark definition. These manually annotated landmarks were used to train one of the automatic landmarking deep learning models included in BioFace3D. Prior to training, the manual landmarks reliability was assessed by evaluating intra- and inter-observer variability across multiple manual annotation repetitions. Statistical analysis confirmed that these errors were not significant, ensuring the consistency and accuracy

of the manual annotations used for training the automatic landmarking models of BioFace3D.

2.1.2. Validation datasets

To validate BioFace3D's performance in a completely unbiased manner and to test its generalizability across different MRI scanners, we employed distinct datasets from those used during the training.

We used 20 MRI scans of healthy adult controls from the IXI [36], and the Alzheimer's Disease Neuroimaging Initiative (ADNI) (adni.loni.usc.edu) public datasets (10 from each dataset), denoted as IXI₁₀ and ADNI₁₀. The selection criterion for these data included T1w head MRI acquired with MPRAGE protocol (1.5T and 3T) from individuals with preserved facial morphology, ensuring complete facial reconstruction from the forehead to the chin, and including the tip of the nose. Noisy or artefactual MRIs were not considered in the analysis.

Although we obtained express consent from IXI and ADNI to use their data in our experiments, the approval excluded displaying facial reconstructions in this paper. For this reason, we acquired six MRI scans from our team members, who gave explicit consent to use their facial images in scientific publications. These scans constitute the BF3D dataset, and were obtained through the FIDMAG Germanes Hospitalàries Research Foundation with a 3T Philips Ingenia scanner

at the Hospital Sant Joan de Déu (Barcelona, Spain). Only facial reconstructions from the BF3D proprietary dataset have been used for the elaboration of the figures in this manuscript.

Finally, a proprietary dataset of 234 MRI scans of adult male and female individuals with Down syndrome (80 individuals) and euploid controls (154 individuals) provided by Hospital Sant Pau Memory Unit (Barcelona, Spain) was used to validate the computation of potential facial biomarkers in scenarios with highly diverse facial anatomies.

A summary of the validation datasets composition is presented in the Supplementary Material.

2.2. BioFace3D pipeline

BioFace3D is divided into three major modules (Fig. 1), comprising: (i) facial mesh extraction from MRI (Fig. 1A), (ii) 3D facial landmarking (Fig. 1B) and (iii) facial biomarkers computation (Fig. 1C). To perform an end-to-end facial biomarker discovery study, it is necessary to iteratively apply modules 1 and 2 to all individuals in the sample. Finally, from the 3D coordinates of the landmarks, the biomarker analysis can be completed with module 3.

Each module has its own user interface (UI). The software was written in MATLAB 2023b [37], with Deep Learning, Statistics and Machine Learning, Image Processing, Computer Vision, and Mapping Toolboxes, and with dependencies on Python 3.8 and PyTorch 1.10. It can be run on a Windows personal computer (16 GB RAM with 6-core CPU). As an orientation, the process of converting a NIfTI file to an automatically landmarked 3D facial model takes six minutes in average. The use of GPU can accelerate the execution of some of its components. The code is available in the Bitbucket repository (https://bitbucket.org/cv_her_lasalle/bioface3d).

2.3. Module 1: Face extraction from MRI

This module automatically produces a facial mesh from an MRI scan, taking as input NIfTI file(s) (.nii/.nii.gz) and generating as output 3D facial meshes encoded in PLY file(s) (.ply). Module 1 can process files individually or in batch mode.

Fig. 2A shows the UI of the facial mesh extraction module. It is designed as a multi-step process that can be interrupted by the user to supervise and to validate intermediate results (Fig. 2B and C). Algorithm 1 presents a pseudo-code of the automatic processes performed by the module, as detailed below.

Orientation normalization (① in Fig. 1). The orientation of an anatomical structure – such as the head – in an MRI scan is determined by the patient-based coordinate system defined by the DICOM standard system [38]. These coordinates relate the Cartesian axes to the anatomical coordinate system. Generally, NIfTI or DICOM files define RAS (x-axis pointing to Right, y-axis to Superior, z-axis to Anterior) as the default orientation. This information is stored in the `DataDirection` field of the NIfTI file header, codifying the reference orientation as a 3×3 diagonal matrix.

However, depending on the reconstruction software, technique or post-acquisition alignment procedures used, up to 24 possible orientations can be encountered, assuming the coordinate systems are always orthogonal and comply with the right-hand rule [39]. This information is codified in the `QForm` field of the NIfTI file header as a 3×3 rotation matrix with respect to the orientation determined by `DataDirection`. In this way, we can determine the orientation of all MRI files and apply the corresponding rotation transformations to automatically orient all heads to a common direction. Following [10], the orientation of the heads is normalized in BioFace3D to the LSA orientation (x-axis pointing to Left, y-axis to Superior, z-axis to Anterior).

Slice processing & enhancement (② in Fig. 1). Most head MRI scans are acquired to analyze the brain, so the surrounding regions often present problems of non-uniform illumination and signal attenuation. To reduce this variability, a harmonization process is performed to minimize intensity inhomogeneity, either intra-MRI or inter-MRI [40]. This processing is divided into two stages: histogram matching and bias field correction.

To normalize intensity and to reduce variability across MRI slices, histogram matching is applied using as a reference an MRI scan from the IXI dataset [36] that shows high contrast among tissues and uniform intensity across slices. The SimpleITK [41] histogram matching implementation is applied to any input MRI, increasing intensity uniformity. Next, to reduce non-uniform intensity bias, the N4ITK (N4 Inhomogeneity Correction) method is used [42]. To determine the MRI region where bias correction is applied, a binary mask is created by thresholding the histogram-matched MRI, and subsequently applying the hole-filling and closing mathematical morphology operations. The results of the bias field and logarithmic bias correction processes are portrayed in the Supplementary Material.

3D head reconstruction (③ in Fig. 1). The 3D head models are reconstructed from the MRI scan by automatically determining a grayscale intensity threshold that optimizes skin segmentation and distinguishes between background and foreground voxels.

For this purpose, facial anatomy experts manually selected the optimal threshold value for the segmentation of a subset of 150 MRI scans from our training dataset. We measured a correlation coefficient of $r = 0.81$ between the mean and the maximum intensity level of the MRI and the manually selected optimal threshold. Then, we trained a linear regression model to automatically estimate this segmentation threshold value using the mean and maximum intensity level of the MRI as input variables. Based on this threshold, the Flying Edges algorithm [43] is applied to generate a 3D binary polygonal reconstruction of the head, guided by the linear regression model, which enables accurate segmentation of the facial surface from the MRI scan. The resulting output is a triangular manifold mesh, obtained after removing duplicate points and correcting potential topological errors. However, if the output is not satisfactory, the users can adjust the predicted threshold value manually through the UI and previsualize the resulting 3D head mesh.

Head processing (④ in Fig. 1). This is a denoising and normalization process implemented using PyMeshLab [44]. It involves four steps (Fig. 3): (1) cleaning isolated elements, (2) centering of the head bounding box, (3) fine alignment to a reference, and (4) internal emptying of the head.

The 3D head models obtained from the reconstruction usually present small isolated elements corresponding to the noise outside the head (Fig. 3A). To remove this noise and to keep only the head mesh, a 3D filter is applied to remove isolated elements with a diameter less than 90% of the maximum object size, which corresponds to the head. This allows to define the 3D bounding box of the head and normalize its position in terms of translation. To do this, we calculate the centroid of the 3D bounding box and use this value to translate the entire mesh to the origin of the Cartesian coordinate system (0,0,0) (Fig. 3B).

To complete the fine alignment of the 3D head, a 3D rigid transformation is applied to align all heads to exactly the same reference (see Supplementary Material). To that end, we use a reference head mesh that has been manually aligned in the LSA orientation, and registered with 5 anatomical control points corresponding to the inner corners of the eyes, the tip of the nose and the corners of the mouth. We refer to the coordinates of these control points as $CP_{ref} \in \mathbb{R}^{5 \times 3}$. Subsequently, we automatically register these 5 control points on the input head mesh to be aligned ($CP_{in} \in \mathbb{R}^{5 \times 3}$), leveraging a 5-landmark selection of the Multi-view Consensus Convolutional Neural Network (MVCNN) model [10] pre-trained on the DTU-3D dataset [45]. Next, we perform an inter-point match that estimates the similarity transformation that

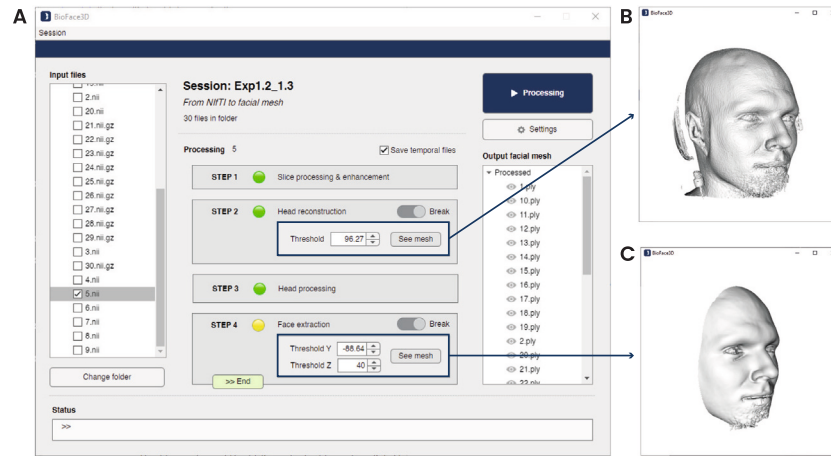


Fig. 2. Module 1: Face extraction from MRI. (A) UI of the module. The left side panel lists the files in the input folder (.nii/.nii.gz) that can be selected to process, and the right side panel lists the already processed (.ply) files. The central panel shows the different steps that are automatically completed to obtain a 3D facial mesh from the MRI. (B) Intermediate preview of the 3D head reconstruction and manual adjustment of the segmentation threshold. (C) Intermediate preview of the resulting 3D facial mesh and manual adjustment of the cutting thresholds.

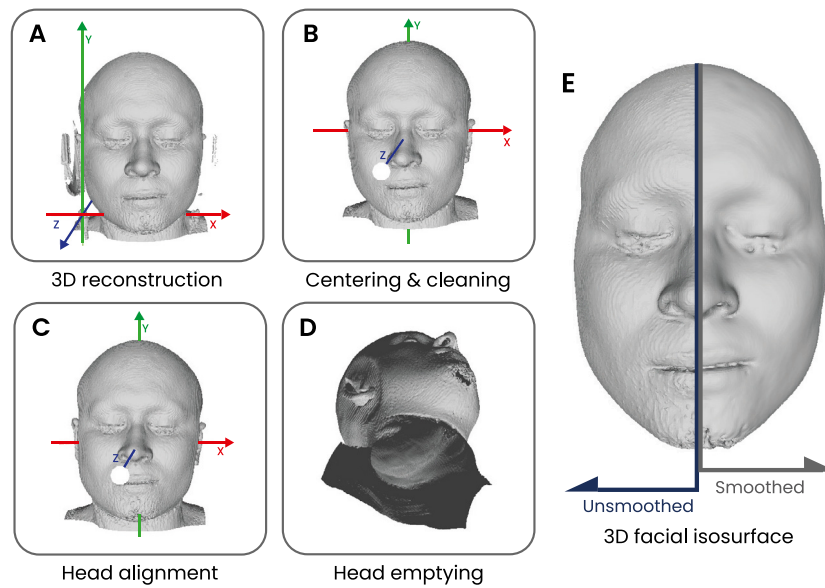


Fig. 3. 3D head processing. (A) 3D head model reconstructed from MRI. (B) 3D head model after cleaning and centering of the bounding box at the origin of coordinates. (C) 3D head model perfectly aligned. (D) 3D head model internally emptied. (E) 3D facial isosurface extracted from the head before applying a remeshing and smoothing process (right) and after applying it (left).

maps the position of CP_{in} to that of CP_{ref} . And then, the corresponding 3D rigid geometric transformation is applied to the input head mesh to perfectly align it to the reference (Fig. 3C). It is important to note that this step can be disabled from the BioFace3D UI, since it may be a time-consuming process.

Finally, to empty the head, internal vertices are automatically selected based on their quality and subsequently removed. To this end, an ambient occlusion filter is applied and vertices are categorized in terms of their quality in the 0 to 1 range. The darkest (or deepest) vertices – corresponding to the most internal areas of the head – will have values close to 0. In contrast, the vertices of the most external facial surface will have quality values close to 1. Finally, vertices with a quality between 0 and a heuristically adjusted threshold of 0.027 are automatically selected and removed (Fig. 3D).

3D facial mesh extraction (5 in Fig. 1). Considering that the head is already aligned to a common reference space, we can apply fixed thresholds to cut the 3D mesh along the Z (*thZ*) and Y (*thY*) axes and to extract only the facial surface. These threshold values are displayed

in the BioFace3D UI and can be modified by the user, offering also a preview of the result, as shown in Fig. 2C. In addition, all vertices and faces that ended up isolated due to the breaking of the original mesh are also removed.

Optionally, to further improve the visual quality of the resulting 3D facial mesh, a Taubin low-pass filter [46] can be used to smooth the facial surface and to eliminate the high-frequency noise resulting from the 3D scanning and reconstruction. Moreover, to close possible holes left on the facial surface, the screened Poisson reconstruction method [47] can be applied, which generates a new isosurface that preserves the original morphology of the face. This filter is based on Neumann boundary conditions, so the newly generated mesh extends beyond the boundaries of the original face mesh. Therefore, the same cropping strategy explained above is applied on the mesh resulting from the screened Poisson reconstruction. It is important to note that the Taubin low-pass filtering and screened Poisson reconstruction are optional to the user, since the new meshes may present slight shape variations with respect to the original mesh (Fig. 3E). These differences are discussed in the results section.

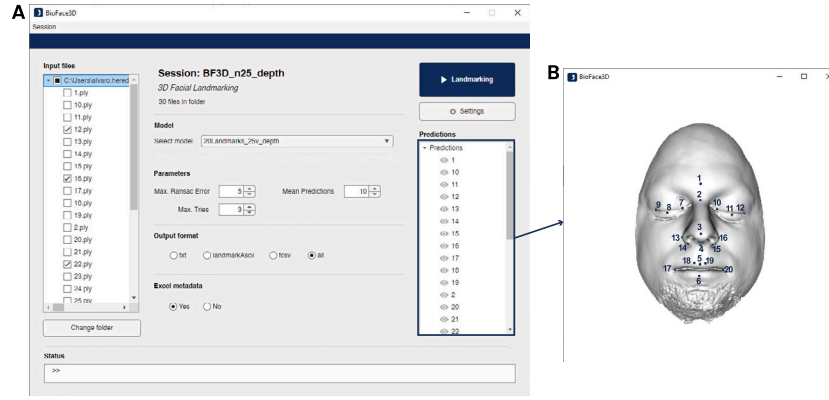


Fig. 4. Module 2: 3D landmarking. (A) UI of the module. The left side panel lists the files in the input folder (.ply) that can be selected to landmark, and the right side panel lists the already landmarked files. In the central panel, it is possible to select the MVCNN model to perform landmarking, modify prediction hyperparameters and adjust the output formats. (B) Previsualization of landmark prediction from UI using BioFace3D-20 landmark model.

Algorithm 1: 3D Facial Mesh Extraction from MRI

Input: Head MRI M
Output: Manifold facial mesh $F \in \mathbb{R}^{n \times 3, m}$ (n : vertices, m : faces)

```

 $M \leftarrow \text{TRANSFORMORIENTATION}(M, 'LSA');$ 
 $M \leftarrow \text{HISTOGRAMMATCHING}(M, \text{Template});$ 
 $\text{mask} \leftarrow \text{GETMASK}(M);$ 
 $M \leftarrow \text{BIASCORRECTION}(M, \text{mask});$ 
 $[MaxI, MeanI] \leftarrow \text{GETINTENSITY}(M);$ 
 $th \leftarrow \text{PREDICTTHRESHOLD}(MaxI, MeanI);$ 
if  $th$  is valid then
     $H \leftarrow \text{FLYINGEDGES3D}(M, th);$ 
     $H \leftarrow \text{REMOVEISOLATEDPIECES}(H, 90\%);$ 
     $H \leftarrow \text{CENTREMESH}(H);$ 
    if  $\text{fine\_alignment}$  then
         $CP \leftarrow \text{MVCNN}(H);$ 
         $H \leftarrow \text{SIMILARITYTRANSFORMATION}(CP, CP_{ref});$ 
     $H \leftarrow \text{AMBIENTOCCLUSION}(H);$ 
     $v \leftarrow \text{SELECTVERTEXQUALITY}(0, 0.027);$ 
     $H \leftarrow \text{REMOVESSELECTEDVERTEX}(H, v);$ 
     $T \leftarrow \text{EXTRACTFACIALMESH}(H, thY, thZ);$ 
    if  $\text{smooth\_remesh}$  then
         $T \leftarrow \text{TAUBINFILTER}(T);$ 
         $F \leftarrow \text{SCREENEDPOISSON}(T);$ 
    else
         $F \leftarrow T;$ 
    if  $\text{save\_snapshot}$  then
         $\text{SAVESNAPSHOT}(F);$ 
else
     $\text{SETMANUALTHRESHOLD}(th);$  restart from thresholding;
return  $F$ 

```

2.4. Module 2: 3D landmarking

The 3D landmarking module (6 in Fig. 1) takes as input data the 3D facial meshes in PLY (.ply) format from the previous module and generates output data file(s) in various formats with the automatically predicted 3D facial landmarks coordinates. Fig. 4A presents the UI of the module and Fig. 4B displays the output of the landmarking process.

In BioFace3D, automatic 3D facial landmarking is performed using state-of-the-art Multi-view Consensus Convolutional Neural Networks (MV-CNN) [10]. This method is based on performing a 2D projection of the 3D head models from different viewpoints, using the hourglass

model to predict candidate landmarks on 2D heat maps and ray-projecting the final position onto a real vertex of the 3D mesh by using least squares (LSQ) combined with random sample consensus (RANSAC). This 2D multi-view approach allows 3D landmark registration at a reduced computational cost but with high precision. The main training hyperparameters affecting the model such as the number of views (i.e. the number of 2D projections with randomly taken camera positions) and image channels have been chosen following [48].

For flexibility, BioFace3D includes two MVCNN models for facial landmarking, so the user chooses the best-suited model.

- **BioFace3D-20 model:** 20 landmarks, trained from scratch on 339 3D facial models segmented from MRI scans. Each sample was encoded using 25 multi-view projections combining depth and geometry channels, with random rotations applied in the XYZ-axes ranges $[\pm 40^\circ, \pm 80^\circ, \pm 20^\circ]$. The model was trained over 100 epochs (early stopping enabled), using a batch size of 4, Adam optimizer (initial learning rate 0.001), and a feature extractor with 256 channels and 0.2 dropout rate. The average training time was approximately 9.8 h on an NVIDIA RTX 2080 Ti GPU. Full landmark definitions and training configuration details are available in the Supplementary Material.
- **DTU-3D model:** 73 landmarks (24 landmarks + 49 semi-landmarks), pre-trained on 601 3D facial models obtained with a Canfield Vectra M3 surface scanner. See [10,45] for landmarks definition and training hyperparameter details.

2.5. Module 3: Facial biomarkers

This last module allows the statistical comparison between two study groups of individuals (e.g., case and control groups) through the automatic discovery of potential biomarkers. These biomarkers, obtained by GM from the coordinates of homologous facial landmarks, characterize the diversity of faces in feature vectors that encode their morphology. Fig. 5 shows the different sections of the UI module.

Data preparation (7 in Fig. 1). From the landmark files obtained from the previous module, BioFace3D generates an Excel spreadsheet for each study group with the facial landmark coordinates of all the individuals in the sample.

Biomarkers computation (8 in Fig. 1). Two GM-based biomarker calculation techniques have been implemented in BioFace3D: Generalized Procrustes Analysis (GPA) and Euclidean Distance Matrix Analysis (EDMA).

To obtain GPA-based facial biomarkers, it is necessary to remove the influence of scale and position from the landmarks configurations, so

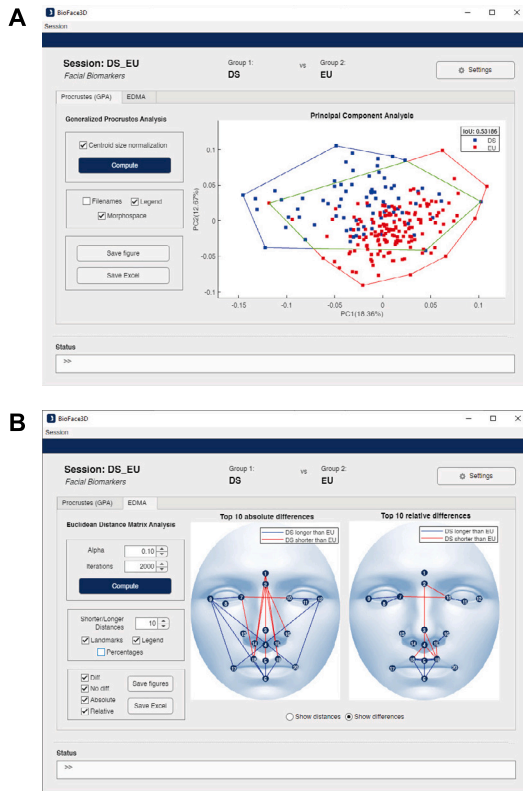


Fig. 5. Module 3: Facial biomarkers. UI of the module is divided into two sections: (A) GPA analysis and (B) EDMA analysis. Each section allows the configuration of analysis parameters and the saving of figures and metadata files.

that they are in the same morphospace [32]. For this, the coordinates of the landmarks of each subject are normalized by their centroid size (defined as the square root of the sum of squared distances of all the landmarks to their centroid), translated to the same origin of coordinates, and rotated until achieving a least-squares fit between homologous landmarks, minimizing the Procrustes distance. Then, the Procrustes coordinates of each landmark configuration can be interpreted as a single n -dimensional vector, where each dimension is equivalent to a spatial coordinate of one of the landmarks. Considering that we work with l 3D landmarks, vectors will have a dimension of $n = 3 \cdot l$.

This set of vectors is then decomposed using Principal Component Analysis (PCA) into its principal components, i.e. linear combinations of the original dimensions that simplify the description of the variation between individuals in the sample and allow the visualization of natural groupings of individuals in spaces of reduced dimension. In this way, it is easier to analyze which groups of individuals have a more similar facial shape. GPA-based facial biomarkers feature vectors are the first vector components obtained after the dimension reduction of PCA. Therefore, different facial biomarkers can be built based on GPA retaining a greater or lesser number of principal components. As user feedback, the UI of this module displays the morphospaces generated from the first two principal components, with the corresponding percentage of total morphological variance explained by each component (Fig. 5A). In addition, the convex hull of each group on the morphospace can be displayed, and the intersection over union (IoU) metric between convex hulls can be computed as a measure of the separability of the two groups.

On the other hand, EDMA biomarkers are based on the calculation of the matrix of linear distances between all possible pairs of landmarks of each individual. This enables to statistically determine which

linear measurements differ between the two study groups by pairwise comparisons based on confidence interval ($\alpha = 0.10$) tests [35].

This analysis is performed under the statistical hypothesis that for the two groups and given a specific inter-landmark distance, this has the same expected mean value for both groups. If the hypothesis is not rejected, the inter-landmark distance is not significantly different between the two groups. After all the statistical tests have been carried out on each of the inter-landmark distances, those that are found to be statistically different between both groups will be considered the distinctive features of one group with respect to the other. These characteristic distances are ordered according to the degree of difference, from greatest to least. Therefore, we can construct different facial biomarker feature vectors based on EDMA using a greater or lesser number of inter-landmark distances among the most significant differences between both groups.

In the UI (Fig. 5B), the number of significantly different inter-landmark distances used for the construction of the EDMA facial biomarkers can be selected from 2 to the number of significantly different distances obtained in each experiment. In addition, the percentage of significantly different distances is indicated. These distances are iteratively displayed on a simulated facial figure with the landmarks.

2.6. Evaluation methodology

2.6.1. Face reconstruction

Facial isosurface extraction from MRI was evaluated both quantitatively and qualitatively by comparing the automatically generated meshes from BioFace3D with expert-guided manual reconstructions.

To that end, an expert morphologist manually selected the segmentation threshold value for the reconstruction of the 3D head model from the MRI with Amira. Then, the 3D model was internally and externally cleaned, translated to the center of the coordinate axis, and rotated to LSA orientation with Meshlab. Finally, the 3D facial surface was extracted.

For the quantitative evaluation, three complementary geometric metrics were computed between the automatic (source, A) and manual (target, B) facial meshes:

- **Unidirectional Chamfer Distance (uCD):** Measures the average squared Euclidean distance from each vertex of the automatic mesh A to its nearest neighbor in the manual mesh B :

$$\text{uCD}(A \rightarrow B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} \|a - b\|_2^2 \quad (1)$$

- **Unidirectional Hausdorff Distance (uHD):** Measures the maximum of the minimum distances from points in A to their nearest neighbors in B :

$$\text{uHD}(A \rightarrow B) = \max_{a \in A} \min_{b \in B} \|a - b\|_2 \quad (2)$$

- **Centroid Shift Magnitude (CSM):** Computes the Euclidean distance between the centroids of both meshes:

$$\text{CSM}(A, B) = \left\| \frac{1}{|A|} \sum_{a \in A} a - \frac{1}{|B|} \sum_{b \in B} b \right\|_2 \quad (3)$$

While uCD evaluates the overall average geometric fidelity between meshes, uHD is more sensitive to localized deviations and helps identify outliers, and CSM captures potential non-uniform vertex distributions or misalignment.

Prior to metric computation, both reconstructions were first cropped using the same thresholds and manually fine-aligned using control points and the iterative closest point (ICP) algorithm. For efficient nearest-neighbor search between mesh vertices, a kD-tree-based algorithm was used.

Additionally, for qualitative evaluation, representative examples of manual and automatic facial reconstructions are presented. To visualize local discrepancies, the per-vertex uCD was encoded as a color map using a blue-to-yellow gradient, highlighting the facial regions in the automatic reconstruction that exhibit the largest deviations from the manual ground truth.

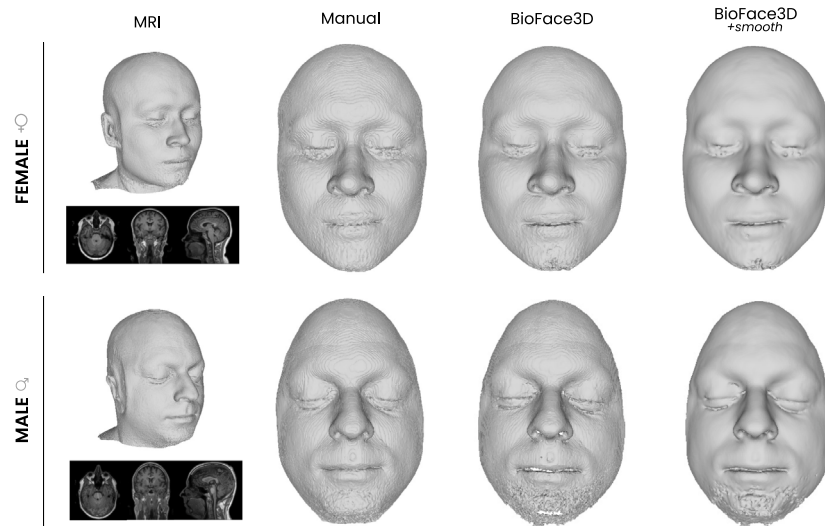


Fig. 6. Qualitative comparison of facial reconstructions from MRI. Each row shows the reconstruction of an individual obtained manually (second column), automatically with BioFace3D, without and with smoothing (third and fourth columns).

2.6.2. Landmarking accuracy

Automatic landmarking accuracy was evaluated against manual landmarking in terms of Euclidean distance metrics and morphometric variables.

A single expert morphologist manually registered 20 anatomical landmarks on the non-smoothed facial meshes obtained automatically by BioFace3D. Regarding automatic landmarking, we used the BioFace3D-20 trained model. The MVCNN hyperparameters were set to 25 views, depth + geometry channels and rotation angles ($\theta_x = \pm 40^\circ$, $\theta_y = \pm 80^\circ$, $\theta_z = \pm 20^\circ$), following [48].

The Euclidean distance-based evaluation was carried out using (i) the Mean Radial Error (MRE), defined as the average Euclidean distance (in mm) between the manually and automatically placed landmarks, and (ii) the Successful Detection Rate (SDR), which measures the percentage of predicted landmarks located within a given radius from the reference (gold standard) landmarks. The SDR was calculated for a series of increasing distances (from 1 to 4 mm), where higher SDR values indicated greater accuracy in landmark detection.

Morphometric evaluation was conducted using Centroid Size (CS) and Pairwise Procrustes Distances (PPD) derived from the anatomical coordinates of the manual and automatic landmarks. For CS, we calculated the square root of the summed squared distances between each landmark coordinate and the centroid of the landmark set. Regarding PPD, we performed GPA on the automatically and manually landmarked data separately and computed the Procrustes distances between each subject and the rest of the sample. For both CS and PPD, we then computed the correlation between the values obtained from the automatic and manual landmarking methods, computing the p -value to test for the significance of the correlation coefficient between methods.

2.6.3. Facial biomarker computation

A comparison of facial biomarkers between euploid (EU) and Down syndrome (DS) groups of individuals was performed using the fully automatic BioFace3D pipeline – including the BioFace3D-20 automatic landmarking model – and a fully manual process. This comparison was based on GPA and EDMA.

In the GPA analysis, we compared the intersection over union (IoU) between the convex hulls of the EU and DS groups in the PCA morphospace. In addition, we calculated the Procrustes Distance (PD) to assess the overall difference in facial shape between groups [49]. This PD was calculated based on the mean shapes of each group,

with 10,000 permutations computed for both the manual and automatic studies. This analysis aimed to evaluate whether both methods identified significant differences in facial shapes between the groups.

For EDMA, each individual was represented as a matrix of 190 linear distances between all possible pairs of anatomical landmarks. Based on a confidence interval ($\alpha = 0.10$) [35], we calculated the Facial Difference Score (FDS) as the percentage of significantly different facial distances among the total number of inter-landmark distances [24].

To validate whether the FDS percentage is statistically significant or a product of the normal morphological variability expected in a general population, we performed a series of EDMA simulations, as in [6,17,24]. To that end, we applied iterative bootstrapping to generate random samples of EU and DS individuals. We first created subsamples of 70 randomly chosen EU subjects. Then, using a subsampling approach, we automatically generated random pseudo-subsamples containing an increasing number of DS individuals in each iteration. This procedure resulted in a series of staggered pseudo-subsamples that contained from 0 to 70 DS subjects. A total of 100 simulations were run in each round, and in each of these simulations, we computed an EDMA analysis and an FDS score.

The results from each round of random groups were separately represented in histograms. The first round of simulations contained no DS subjects and only included EU individuals, representing facial differences that can be found randomly in the general EU population. To assess whether the FDS value obtained using the complete DS sample was significantly different or similar to the FDS resulting from a random sample, we compared the distribution of FDS random values with the FDS observed in the whole sample. The p -value assessing the statistical significance of the comparison was computed as the ratio between the number of simulations containing no patients that provided a higher FDS than the observed FDS divided by the total number of cumulative simulations. P -values below 0.05 indicated that the FDS obtained using the real dataset was higher than the FDS obtained randomly in a sample of EU subjects.

Finally, to reveal the unique morphological facial pattern of each group, we computed the top-10 significant absolute and relative largest and shortest distances between the DS and EU groups.

3. Results and discussion

3.1. High similarity between manual and automatically extracted facial meshes

Fig. 6 presents the facial reconstructions obtained manually and automatically with BioFace3D on a male and a female subjects of the

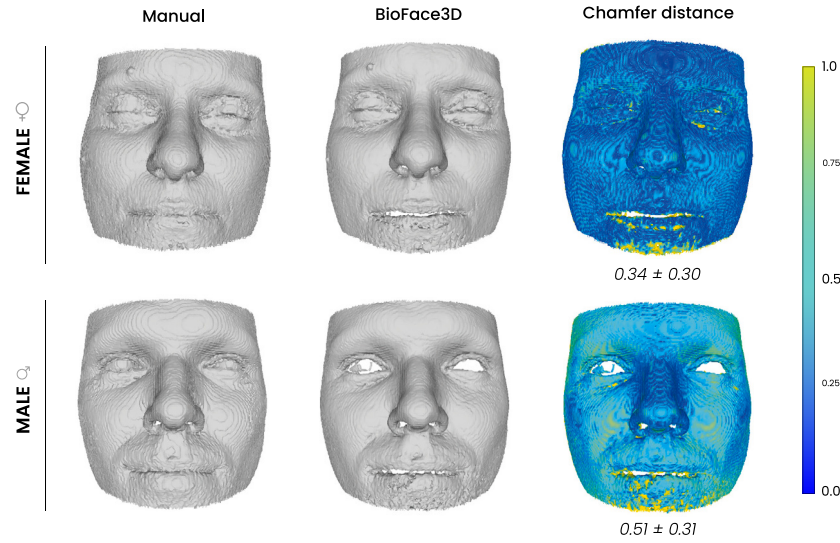


Fig. 7. Chamfer distances between automatic and manual face reconstruction. Each row represents an BF3D individual reconstructed manually (first column) and automatically with BioFace3D without smooth (second column). The third column shows the surface error between both reconstructions color-coded in millimeters and the average Chamfer distance ($\pm$ standard deviation).

Table 1

Geometric distances (in mm) in terms of unidirectional Chamfer distance (uCD), unidirectional Hausdorff distance (uHD) and centroid shift magnitude (CSM) between manual and automatic BioFace3D facial reconstruction of test individuals.

	IXI ₁₀	ADNI ₁₀	BF3D	Average
uCD	0.20 ± 0.21	0.32 ± 0.27	0.72 ± 0.51	0.37 ± 0.30
uHD	3.52	4.24	6.21	4.42
CSM	0.83	1.09	2.33	1.28

BF3D proprietary dataset. The manual and automatic reconstructions of the remaining subjects of the BF3D dataset are portrayed in the Supplementary Material.

From a qualitative point of view, we observed a high similarity between the manual and automatic facial models, especially in the central area of the face where the biomarkers are calculated. Some differences could be noted in the chin, corners of the eyes and ends of the mouth, areas highly sensitive to the selected segmentation threshold. However, if necessary, this automatic threshold value can be modified from the UI.

Although manual reconstruction still proved to be the best method, the automatic reconstruction preserved the main features of facial morphology. Moreover, the automatic process halved the average time required compared to manual processing.

Concerning smoothed reconstructions, high-frequency noise was eliminated and missing parts of the facial surface were reconstructed. However, details of the facial morphology were lost, especially in the eyes and mouth. For this reason, non-smoothed reconstructions were used in the following experiments of this article.

Quantitatively, Table 1 reports the geometric distances between the manual and automatic reconstructions across all validation datasets. On average across all datasets, the automatic reconstructions achieved a Chamfer distance of 0.37 ± 0.30 mm, indicating high global geometric fidelity between meshes with submillimeter accuracy, comparable to state-of-the-art methods [50]. The Hausdorff distance reached 4.42 mm, reflecting the presence of a few isolated outlier points; however, these were mainly located at the outer facial boundaries or in superficial artifacts beyond the coverage area of the landmarking system. The centroid shift was 1.28 mm, suggesting a uniform vertex distribution

Table 2

Mean radial error (in mm) between manual and automatic BioFace3D landmarks of test individuals.

	IXI ₁₀	ADNI ₁₀	BF3D	Average
MRE	1.61 ± 0.72	1.81 ± 0.90	1.63 ± 0.65	1.69 ± 0.83

in the automatic meshes and acceptable global alignment concerning the manual references.

As observed, performance was slightly better for the public cohorts (IXI₁₀ and ADNI₁₀), likely due to higher MRI resolution, standardized acquisition protocols, and more homogeneous intensity distributions. These results confirm the robustness and geometric accuracy of the BioFace3D pipeline for the automatic reconstruction of 3D facial surfaces from different sources of MRI data.

To complement the numerical evaluation, qualitative inspection was performed by mapping per-vertex uCD values onto the automatic mesh surface (Fig. 7). As expected, the most prominent deviations were located in high-curvature regions such as the mouth, nose, and periocular area, which are typically more susceptible to segmentation variability. Full-color maps for all individuals in the BF3D dataset are provided as Supplementary Material.

3.2. Automatic landmarking accurately models facial shape

From a Euclidean distance standpoint, Fig. 8 shows the MRE box-plots for each landmark. Overall, we observed a higher accuracy in midline landmarks as compared to lateral landmarks, such as the corners of the mouth and eyes, which is consistent with the literature [51]. Moreover, we observed that the MRE of most landmarks was below 2 mm, an accuracy level that is below the intra-observer variation commonly accepted in craniometric measurements [52]. In average terms, the MRE for all landmarks of test individuals was 1.69 ± 0.83 mm (Table 2).

Furthermore, when landmarking accuracy was evaluated in terms of SDR, we found that more than 68% of the landmarks were below the 2 mm critical threshold, as shown in Table 3. This value is comparable to state-of-the-art automatic facial landmarking methods [48,53].

In terms of morphometric variables, Fig. 9 displays the correlation between the CS and PPD computed from the automatically and

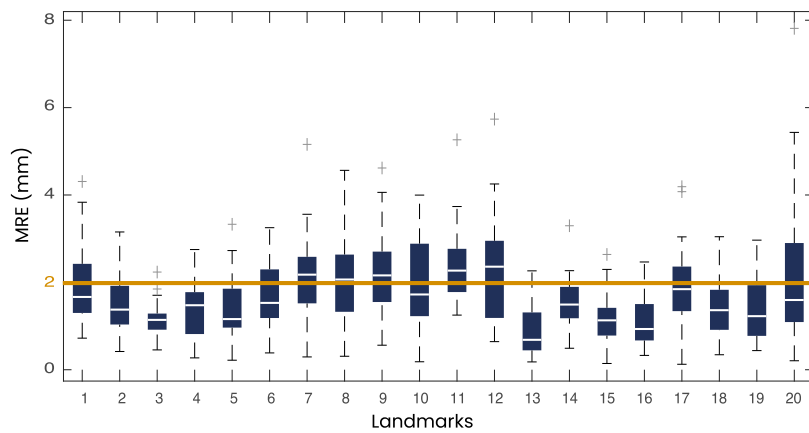


Fig. 8. Mean radial error (MRE) boxplots. Comparison of Euclidean distance (in mm) per landmark between automatic and manual landmarks. The superimposed line of 2 mm marks the commonly accepted error in craniometric measurements.

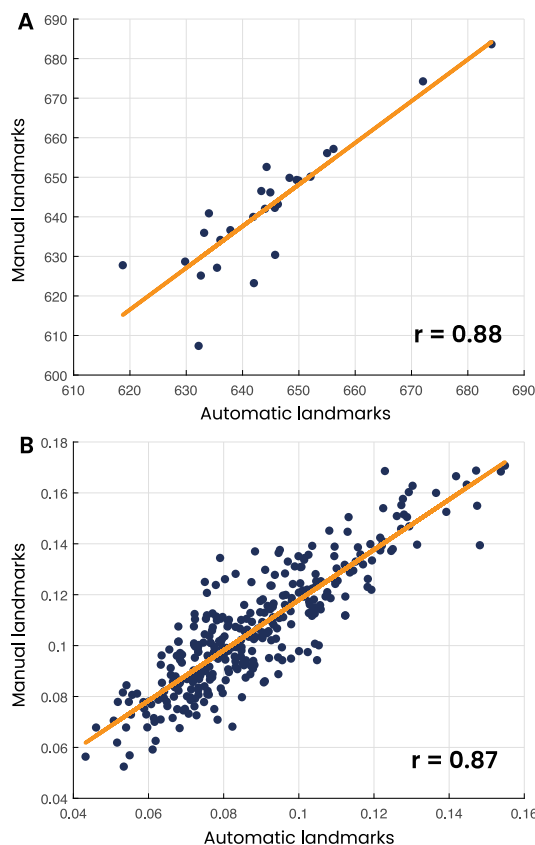


Fig. 9. Automatic landmarking evaluation in terms of morphometric variables. Correlation between automatic and manual landmarking in terms of (A) Centroid Size and (B) Pairwise Procrustes Distances.

Table 3
Successful detection rate (in %) between manual and automatic BioFace3D landmarks within a given radius (from 1 to 4 mm).

	1 mm	2 mm	3 mm	4 mm
SDR	25.34	68.65	90.58	96.92

manually registered landmarks on the validation datasets, portraying the correlation coefficient r for each comparison. In all cases, the correlations were significant with a p -value < 0.0001 and presented

values higher than 0.87, which validated that automatic landmarking with MVCNN was able to code facial shape as accurately as manual landmarks, despite the variability found among the individuals in the three validation datasets.

Finally, we also measured the time savings gained by automatic landmarking with respect to manual registration. Manual landmarking took 8 min on average for each individual, while the automatic method took 3 min to average 10 model predictions on a conventional Windows PC (16 GB RAM with 6-core CPU) without GPU.

3.3. Down syndrome: A case-study

3.3.1. Potential facial biomarkers

Down syndrome (DS) condition is associated by a characteristic facial phenotype including hypertelorism, slanting eyes, facial flattening, and short and up-turned noses. These facial morphological differences compared to euploid individuals (EU) were automatically encoded by potential GPA and EDMA biomarkers with BioFace3D.

Regarding the GPA biomarkers, Fig. 10 presents the 2D PCA morphospaces obtained by fully manual and automatic processing, respectively. We observed high similarity between the distributions of the manually and automatically processed individuals in both cases, as suggested by the convex hulls – traced in blue for DS and in red for EU – that enclosed the individuals of each group. The shape of the convex hulls was similar in both cases (up to rotation and reflection caused by the eigenvector sign ambiguity inherent to PCA), and the corners of the convex hulls were delimited by the same individuals in both cases. Moreover, the Intersection over Union (IoU) of the convex hulls was 0.53 for manual processing and 0.54 for fully automatic analysis with BioFace3D.

In terms of Procrustes Distance (PD) between groups, both manually (PD = 0.0569; p -value < 0.0001) and automatically with BioFace3D (PD = 0.0483; p -value < 0.0001) reported very close PD values and significant differences in global facial shape between EU and DS individuals.

Regarding EDMA biomarkers, the FDS percentage revealed similar values for both manual (75.8%) and automatic (74.6%) processes, for the 190 total facial distances between DS - EU. To validate the statistical significance of these values we used EDMA simulations, as shown in Fig. 11. In both simulations, with manual and automatic processing, we observed an increase in the different distances as the number of DS in the mixed group increases, however, this percentage of distances was lower than the FDS value obtained with the full sample of EU and DS (dashed lines) until the last histogram was reached. Furthermore, the p -value < 0.05 in both cases validates the statistical significance of

Table 4

Classification results (in %) for each classifier–biomarker combination, along with the optimal feature vector dimensionality d (in parentheses). Results are reported for the diagnostic class (DS).

Classifier	Biomarkers	Precision	Recall	F1-score	Accuracy
LinSVM	GPA (50)	88.2	92.3	90.2	93.2
	EDMA (96)	88.4	93.8	91.0	93.7
SubDis	GPA (50)	89.7	93.8	91.7	94.2
	EDMA (96)	88.4	93.8	91.0	93.7
MLP50	GPA (40)	88.4	93.8	91.0	93.7
	EDMA (96)	85.7	92.3	88.9	92.1

the results and the use of the automatic method to code local facial differences using EDMA biomarkers.

Finally, Fig. 12 showcases the top-10 significantly shorter and longer distances between EU and DS individuals using EDMA biomarkers computation based on BioFace3D automatic processing. The detected significant distances were consistent with the facial phenotype of individuals with DS, such as greater distance between the eyes, due to hypertelorism, and a shortened midfacial height caused by midface hypoplasia. This DS facial phenotype pattern was in accordance with the literature [17,24].

These results highlight the capacity of BioFace3D to successfully analyze facial data with diverse anatomical and phenotypic traits in a fully automatic manner.

3.3.2. Biomarker-driven automatic diagnosis

As an additional experiment, we explored the potential of GM-based biomarkers for automatic diagnostic classification based solely on 3D facial information using machine learning models. To this end, we trained three classifiers with feature vectors derived from GPA and EDMA, as computed by BioFace3D. Specifically, we used: (i) a support vector machine with a linear kernel (LinSVM), (ii) a random subspace ensemble classifier based on discriminant analysis (SubDis), and (iii) a multilayer perceptron with a single hidden layer of 50 fully connected neurons (MLP50).

Each feature vector consisted of a given number d of (i) principal component values in the Procrustes morphospace (GPA), and (ii) significantly different inter-landmark distances (EDMA). Table 4 reports the performance of the classifiers at their optimal operating point, determined by Youden's Index, based on the optimal dimensionality (d) of each feature vector.

Overall, all classifier–biomarker combinations yielded strong performance, with F1-scores around 90%. The best-performing pair was SubDis+GPA, achieving an accuracy of 94.2%. Regarding dimensionality, all experiments showed improved performance with increased dimensionality, highlighting the richness of the features extracted by BioFace3D and their diagnostic potential.

4. Conclusions and future work

This work presents BioFace3D, a fully automated, end-to-end computational pipeline for extracting and analyzing geometric morphometric (GM)-based facial biomarkers from 3D facial models reconstructed from structural head MRI scans. The tool enables high-throughput processing, starting from NIfTI MRI scan files and culminating in quantitative GM-based comparisons between diagnostic groups. While the pipeline operates fully automatically, BioFace3D provides flexibility for user intervention at key processing stages, allowing for manual adjustments and refinements to optimize accuracy when needed.

The robustness and reliability of BioFace3D have been evaluated against manual processing results, demonstrating its effectiveness in

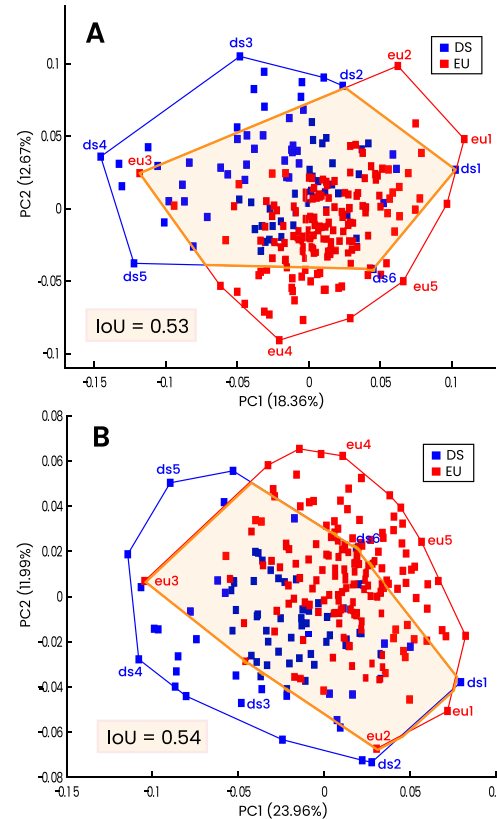


Fig. 10. BioFace3D GPA facial biomarkers computation for Down syndrome. (A) Generalized Procrustes Analysis (GPA) morphospaces between manual and (B) automatic biomarkers.

extracting reproducible and clinically relevant facial biomarkers, accounting for variations in facial anatomy related to different conditions. Furthermore, its clinical applicability has been validated through a diagnostic classification case study in Down syndrome using machine learning models, achieving results comparable to those of other state-of-the-art methods. By providing an open-source, scalable, and reproducible framework, BioFace3D facilitates large-scale studies on facial morphology and enhances biomedical and clinical research applications.

Despite its promising results, BioFace3D presents several limitations that warrant further investigation. First, its performance is influenced by the quality and resolution of the input MRI data, which can vary across clinical sites. Second, the landmarking module may be less accurate in subjects with severe craniofacial anomalies or in the presence of MRI artifacts.

Future work will focus on expanding BioFace3D's capabilities and addressing its limitations in the following directions:

- Improving facial reconstruction methods by enhancing MRI slice quality and implementing advanced deep learning segmentation techniques.
- Extending Module 1 to support additional 3D facial imaging modalities, such as stereophotogrammetry-derived models, increasing the tool's versatility.
- Integrating machine learning models trained on extracted facial biomarker feature vectors to enable automated diagnostic classification based on facial shape data.

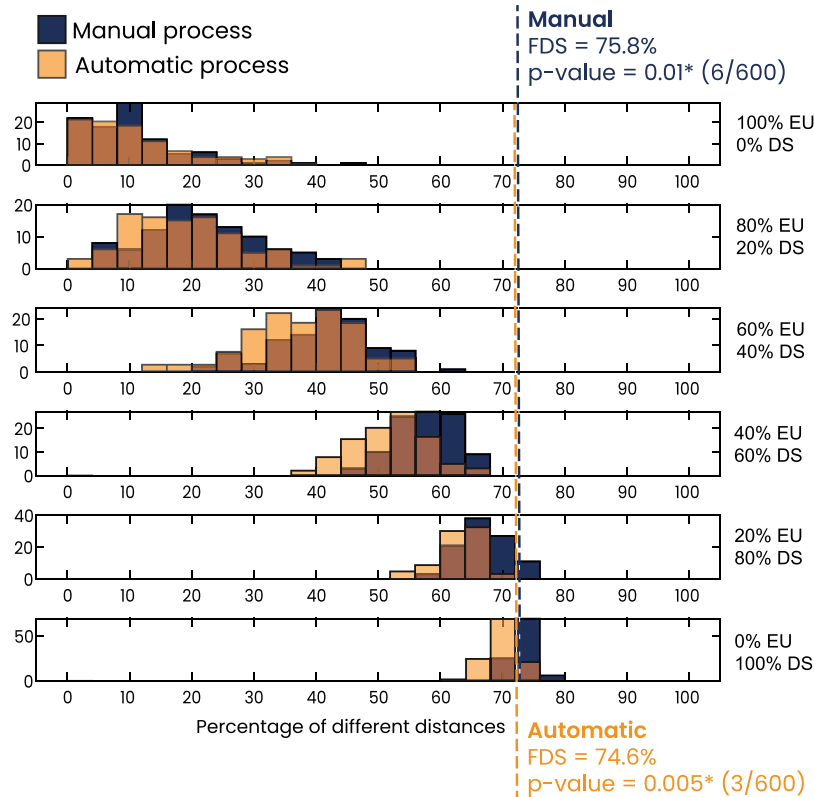


Fig. 11. EDMA simulations based on manual and automatic processes. The histograms represent the percentage of different distances for each simulation. The dashed line shows the FDS obtained with the original DS and EU sample. *Statistically significant cumulative p -value.

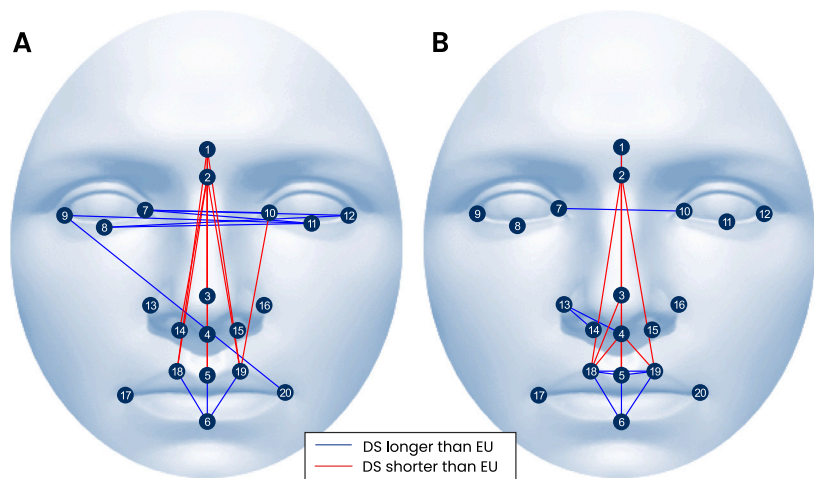


Fig. 12. BioFace3D EDMA facial biomarkers computation for Down syndrome. (A) Top-10 absolute and (B) relative Euclidean Distance Matrix Analysis (EDMA) biomarkers.

CRediT authorship contribution statement

Álvaro Heredia-Lidón: Writing – review & editing, Writing – original draft, Validation, Software, Resources, Methodology, Investigation, Conceptualization. **Luis M. Echeverry-Quiceno:** Validation, Resources, Methodology, Data curation. **Alejandro González:** Software, Resources, Methodology, Conceptualization. **Noemí Hostalet:** Writing – review & editing, Methodology, Investigation. **Edith Pomarol-Clotet:** Writing – review & editing, Data curation. **Juan Fortea:** Data curation. **Mar Fatjó-Vilas:** Writing – review & editing, Supervision, Resources, Methodology. **Neus Martínez-Abadías:** Writing – review & editing,

Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Xavier Sevillano:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Ethical statement for solid state ionics

- (1) This material is the authors' own original work, which has not been previously published elsewhere.
- (2) The paper is not currently being considered for publication elsewhere.

(3) The paper reflects the authors' own research and analysis in a truthful and complete manner.

(4) The paper properly credits the meaningful contributions of co-authors and co-researchers.

(5) The results are appropriately placed in the context of prior and existing research.

(6) All sources used are properly disclosed (correct citation). Literally copying of text must be indicated as such by using quotation marks and giving proper reference.

(7) All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

Funding

The research in this paper was supported by the Joan Oró grant (2024 FI-3 00160) from the Recerca i Universitats Departament (DRU) of the Generalitat de Catalunya, Spain and the European Social Fund, by the Fundación Álvaro Entrecanales-Lejeune predoctoral grant, by Agencia Española de Investigación, Spain (PID2020-113609RB-C21/AEI/10.13039/501100011033), by Instituto de Salud Carlos III (ISCIII), Spain through the contracts FI21/00093 and CP20/00072 (co-funded by European Regional Development Fund (ERDF)/European Social Fund "Investing in your future"), 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).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Alvaro Heredia Lidón reports financial support was provided by Generalitat de Catalunya Ministry of Research and Universities. Neus Martinez-Abadías reports financial support was provided by Government of Spain Ministry of Universities. Luis Miguel Echeverry-Quiceno reports financial support was provided by Jérôme Lejeune Foundation. Noemi Hostalet reports financial support was provided by Carlos III Health Institute. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge the participation of all volunteers supporting this study. We also acknowledge the Alzheimer's Disease Neuroimaging Initiative (ADNI) for the data shared. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private

sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.cmpb.2025.109010>.

References

- [1] Y. Gurovich, Y. Hanani, O. Bar, G. Nadav, N. Fleischer, D. Gelbman, et al., Identifying facial phenotypes of genetic disorders using deep learning, *Nat. Med.* 25 (2019) 60–64, <http://dx.doi.org/10.1038/s41591-018-0279-0>.
- [2] B. Hallgrímsson, J.D. Aponte, D.C. Katz, J.J. Bannister, S.L. Riccardi, N. Mahasuwan, et al., Automated syndrome diagnosis by three-dimensional facial imaging, *Genet. Med.* 22 (2020) 1682–1693, <http://dx.doi.org/10.1038/s41436-020-0845-y>.
- [3] R.J. Hennessy, P.A. Baldwin, D.J. Browne, A. Kinsella, J.L. Waddington, Frontonasal dysmorphism in bipolar disorder by 3D laser surface imaging and geometric morphometrics: Comparisons with schizophrenia, *Schizophr. Res.* 122 (2010) 63–71, <http://dx.doi.org/10.1016/j.schres.2010.05.001>.
- [4] S. Katina, B.D. Kelly, M.A. Rojas, F.M. Sukno, A. McDermott, R.J. Hennessy, et al., Refining the resolution of craniofacial dysmorphism in bipolar disorder as an index of brain dysmorphogenesis, *Psychiatry Res.* 291 (2020) 113243, <http://dx.doi.org/10.1016/j.psychres.2020.113243>.
- [5] A. Schmidt, M. Danyel, K. Grundmann, T. Brunet, H. Klinkhammer, T.-C. Hsieh, et al., Next-generation phenotyping integrated in a national framework for patients with ultrarare disorders improves genetic diagnostics and yields new molecular findings, *Nat. Genet.* 56 (2024) 1644–1653, <http://dx.doi.org/10.1038/s41588-024-01836-1>.
- [6] N. Hostalet, A. González, P. Salgado-Pineda, et al., Face-brain correlates as potential sex-specific biomarkers for schizophrenia and bipolar disorder, *Psychiatry Res.* 339 (2024) 116027, <http://dx.doi.org/10.1016/j.psychres.2024.116027>.
- [7] J.D. White, A. Ortega-Castrillón, H. Matthews, A.A. Zaidi, O. Ekrami, J. Snyder, Y. Fan, T. Penington, S. Van Dongen, M.D. Shriver, P. Claes, MeshMonk: Open-source large-scale intensive 3D phenotyping, *Sci. Rep.* 9 (11) (2019) 6085, <http://dx.doi.org/10.1038/s41598-019-42533-y>.
- [8] R. Kikinis, S.D. Pieper, K.G. Vossburgh, 3D slicer: A platform for subject-specific image analysis, visualization, and clinical support, in: F.A. Jolesz (Ed.), *Intraoperative Imaging and Image-Guided Therapy*, Springer, New York, NY, USA, 2014, pp. 277–289, http://dx.doi.org/10.1007/978-1-4614-7657-3_19.
- [9] A. Porto, S. Rolfe, A.M. Maga, ALPACA: A fast and accurate computer vision approach for automated landmarking of three-dimensional biological structures, *Methods Ecol. Evol.* 12 (11) (2021) 2129–2144, <http://dx.doi.org/10.1111/2041-210X.13689>.
- [10] R.R. Paulsen, K.A. Juhl, T.M. Haspang, T. Hansen, M. Ganz, G. Einarsson, Multi-view consensus CNN for 3D facial landmark placement, in: *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Vol. 11361, 2019, pp. 706–719, http://dx.doi.org/10.1007/978-3-030-20887-5_44.
- [11] R.L. Palmer, P. Helmholtz, G. Baynam, Cliniface: phenotypic visualisation and analysis using non-rigid registration of 3d facial images, *Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. XLIII-B2-2020* (2020) 301–308, <http://dx.doi.org/10.5194/isprs-archives-XLIII-B2-2020-301-2020>.
- [12] C.P. Klingenberg, MorphoJ: an integrated software package for geometric morphometrics, *Mol. Ecol. Resour.* 11 (2) (2011) 353–357, <http://dx.doi.org/10.1111/j.1755-0998.2010.02924.x>.
- [13] B. Al-baker, A. Alkalaly, A. Ayoub, X. Ju, P. Mossey, Accuracy and reliability of automated three-dimensional facial landmarking in medical and biological studies: a systematic review, *Eur. J. Orthod.* 45 (4) (2023) 382–395, <http://dx.doi.org/10.1093/ejo/cjac077>.
- [14] D.C. Adams, M.L. Collyer, A. Kalliontzopoulou, E. Sherratt, Geomorph: Software for geometric morphometric analyses, 2023, R package, URL <https://cran.r-project.org/package=geomorph>.
- [15] N. Kaczorowska, K. Kaczorowski, J. Laskowska, M. Mikulewicz, Down syndrome as a cause of abnormalities in the craniofacial region: A systematic literature review, *Adv. Clin. Exp. Med.* 28 (2019) 1587–1592, <http://dx.doi.org/10.17219/acem/112785>.
- [16] A.G. McKechnie, S.M. Lawrie, A.C. Stanfield, The never-changing face of fragile X? *Lancet Psychiatry* 7 (2020) e26, [http://dx.doi.org/10.1016/S2215-0366\(20\)30127-9](http://dx.doi.org/10.1016/S2215-0366(20)30127-9).
- [17] L.M. Echeverry-Quiceno, E. Candelo, E. Gómez, P. Solís, D. Ramírez, D. Ortiz, et al., Population-specific facial traits and diagnosis accuracy of genetic and rare diseases in an admixed Colombian population, *Sci. Rep.* 13 (2023) 6869, <http://dx.doi.org/10.1038/s41598-023-33374-x>.

- [18] G. Tripi, S. Roux, D. Matrangola, L. Maniscalco, P. Glorioso, F. Bonnet-Brilhault, et al., Cranio-facial characteristics in children with autism spectrum disorders (ASD), *J. Clin. Med.* 8 (2019) 641, <http://dx.doi.org/10.3390/jcm8050641>.
- [19] V.H. Akabaliev, S.T. Sivkov, M.Y. Mantarkov, Minor physical anomalies in schizophrenia and bipolar I disorder and the neurodevelopmental continuum of psychosis, *Bipolar Disord.* 16 (2014) 633–641, <http://dx.doi.org/10.1111/bdi.12211>.
- [20] T. Xu, R.C.K. Chan, M.T. Compton, Minor physical anomalies in patients with schizophrenia, unaffected first-degree relatives, and healthy controls: A meta-analysis, *PLoS One* 6 (9) (2011) e24129, <http://dx.doi.org/10.1371/journal.pone.0024129>.
- [21] J.J. Bannister, S.R. Crites, J.D. Aponte, D.C. Katz, M. Wilms, O.D. Klein, et al., Fully automatic landmarking of syndromic 3D facial surface scans using 2D images, *Sensors* 20 (2020) 3171, <http://dx.doi.org/10.3390/s20113171>.
- [22] T.-C. Hsieh, A. Bar-Haim, S. Moosa, N. Ehmke, K.W. Gripp, J.T. Pantel, M. Danyel, M.A. Mensah, D. Horn, S. Rosnev, N. Fleischer, G. Bonini, A. Hustinx, A. Schmid, A. Knaus, B. Javanmardi, H. Klinkhammer, H. Lesmann, S. Sivalingam, T. Kamphans, W. Meiswinkel, F. Ebstein, E. Krüger, S. Küry, S. Béziau, A. Schmidt, S. Peters, H. Engels, E. Mangold, M. Kreiß, K. Cremer, C. Perne, R.C. Betz, T. Bender, K. Grundmann-Hauser, T.B. Haack, M. Wagner, T. Brunet, H.B. Bentzen, L. Averdunk, K.C. Coetzer, G.J. Lyon, M. Spielmann, C.P. Schaaf, S. Mundlos, M.M. Nöthen, P.M. Krawitz, GestaltMatcher facilitates rare disease matching using facial phenotype descriptors, *Nature Genet.* 54 (3) (2022) 349–357, <http://dx.doi.org/10.1038/s41588-021-01010-x>.
- [23] J.S. Kim, H. Ko, H. Woo, W.S. Kim, Lessons learned from the point-of-care use of a facial analysis technology, *Ann. Child Neurol.* 31 (2023) 271–275, <http://dx.doi.org/10.26815/acn.2023.00227>.
- [24] J.M. Starbuck, S. Llambrich, R. González, J. Albaigès, A. Sarlé, J. Wouters, et al., Green tea extracts containing epigallocatechin-3-gallate modulate facial development in down syndrome, *Sci. Rep.* 11 (2021) 4715, <http://dx.doi.org/10.1038/s41598-021-83757-1>.
- [25] K.C. Koban, P. Perko, L. Etzel, Z. Li, T.L. Schenck, R.E. Giunta, Validation of two handheld devices against a non-portable three-dimensional surface scanner and assessment of potential use for intraoperative facial imaging, *J. Plast. Reconstr. Aesthet. Surg.* 73 (2020) 141–148, <http://dx.doi.org/10.1016/j.bjps.2019.07.008>.
- [26] M. Li, J.B. Cole, M. Manyama, J.R. Larson, D.K. Liberton, S.L. Riccardi, et al., Rapid automated landmarking for morphometric analysis of three-dimensional facial scans, *J. Anat.* 230 (2017) 607–618, <http://dx.doi.org/10.1111/joa.12576>.
- [27] R. Marcucio, B. Hallgrímsson, N.M. Young, Facial morphogenesis: Physical and molecular interactions between the brain and the face, *Curr. Top. Dev. Biol.* 115 (2015) 299–320, <http://dx.doi.org/10.1016/bs.ctdb.2015.09.001>.
- [28] O. Hamadelseed, M.K. Chan, T. Skutella, Distinct neuroanatomical and neuropsychological features of down syndrome compared to related neurodevelopmental disorders: A systematic review, *Front. Neurosci.* 17 (2023) 1225228, <http://dx.doi.org/10.3389/fnins.2023.1225228>.
- [29] C.P. Klingenberg, Evolution and development of shape: integrating quantitative approaches, *Nat. Rev. Genet.* 11 (2010) 623–635, <http://dx.doi.org/10.1038/nrg2829>.
- [30] I.L. Dryden, K.V. Mardia, Statistical Shape Analysis, with Applications in R, Wiley, Chichester, U.K., 2016, URL <https://onlinelibrary.wiley.com/doi/book/10.1002/9781119072492>.
- [31] F.L. Bookstein, Morphometric Tools for Landmark Data: Geometry and Biology, Cambridge Univ. Press, Cambridge, U.K., 1992, <http://dx.doi.org/10.1017/CBO9780511573064>.
- [32] M. Zelditch, D.L. Swiderski, H.D. Sheets, Geometric Morphometrics for Biologists: A Primer, Academic Press, London, U.K., 2012.
- [33] F.J. Rohlf, D. Slice, Extensions of the procrustes method for the optimal superimposition of landmarks, *Syst. Zool.* 39 (1990) 40–59, <http://dx.doi.org/10.2307/2992207>.
- [34] F.J. Rohlf, Shape statistics: Procrustes superimpositions and tangent spaces, *J. Classification* 16 (1999) 197–223, <http://dx.doi.org/10.1007/s003579900054>.
- [35] S. Lele, J.T. Richtsmeier, Euclidean distance matrix analysis: A coordinate-free approach for comparing biological shapes using landmark data, *Am. J. Phys. Anthropol.* 86 (1991) 415–427, <http://dx.doi.org/10.1002/ajpa.1330860307>.
- [36] U.C.L. Imperial College London, U. of Edinburgh, Information extraction from images (IXI) dataset, 2024, Online, [Online]. Available: <https://brain-development.org/ixi-dataset/>. (Accessed 21 May 2024).
- [37] MATLAB, Version R2023b, The MathWorks Inc., Natick, MA, USA, 2023.
- [38] Digital imaging and communications in medicine (DICOM) standard, version 20.0, 2020, [Online]. Available: <http://medical.nema.org/>.
- [39] A. Heredia-Lidón, A. González, C. Guerrero-Mosquera, R. González-Colom, L. M., N. Hostalet, et al., Automated orientation detection of 3D head reconstructions from smri using multiview orthographic projections: An image classification-based approach, in: Proc. Iberian Conf. Pattern Recognit. Image Anal., 2023, pp. 603–614.
- [40] E. Stamoulou, C. Spanakis, G.C. Manikis, G. Karanasiou, G. Grigoriadis, T. Foukakis, et al., Harmonization strategies in multicenter MRI-based radiomics, *J. Imaging* 8 (2022) 303, <http://dx.doi.org/10.3390/jimaging8110303>.
- [41] B.C. Lowekamp, D.T. Chen, L. Ibanez, D. Blezek, The design of SimpleITK, *Front. Neuroinformatics* 7 (2013) 45, <http://dx.doi.org/10.3389/fninf.2013.00045>.
- [42] N.J. Tustison, B.B. Avants, P.A. Cook, Y. Zheng, A. Egan, P.A. Yushkevich, et al., N4ITK: Improved N3 bias correction, *IEEE Trans. Med. Imaging* 29 (2010) 1310–1320, <http://dx.doi.org/10.1109/TMI.2010.2046908>.
- [43] W. Schroeder, R. Maynard, B. Geveci, Flying edges: A high-performance scalable isocontouring algorithm, in: Proc. IEEE Symp. Large Data Anal. Visualization, LDAV, 2015, pp. 33–40, <http://dx.doi.org/10.1109/LDAV.2015.7348069>.
- [44] A. Muntoni, P. Cignoni, Pymeshlab, Zenodo 1 (2021) 2, <http://dx.doi.org/10.5281/zenodo.4903692>.
- [45] J. Fagerlun, S. Harder, A. Rosengren, C. Moeller, T. Werge, R.R. Paulsen, et al., 3D facial landmarks: Inter-operator variability of manual annotation, *BMC Med. Imaging* 14 (2014) 35, <http://dx.doi.org/10.1186/1471-2342-14-35>.
- [46] G. Taubin, A signal processing approach to fair surface design, in: Proc. 22nd Annu. Conf. Comput. Graph. Interact. Techn., ACM, New York, NY, USA, 1995, pp. 351–358, <http://dx.doi.org/10.1145/218380.218473>.
- [47] M. Kazhdan, H. Hoppe, Screened Poisson surface reconstruction, *ACM Trans. Graph.* 32 (3) (2013) 1–13.
- [48] A. Heredia-Lidón, C. García-Mascarell, L.M. Echeverry-Quiceno, N. Hostalet, D. Herrera-Escartín, A. González, et al., A critical comparison between template-based and architecture-reused deep learning methods for generic 3D landmarking of anatomical structures, in: C. Wachinger, B. Paniagua, S. Elhabian, G. Luijten, J. Egger (Eds.), Shape in Medical Imaging, Springer Nature, Cham, Switzerland, 2025, pp. 97–111, http://dx.doi.org/10.1007/978-3-031-75291-9_8.
- [49] P. Mitteroecker, P. Gunz, Advances in geometric morphometrics, *Evol. Biol.* 36 (2) (2009) 235–247, <http://dx.doi.org/10.1007/s11692-009-9055-x>.
- [50] E. Ramon, G. Triginer, J. Escur, A. Pumarola, J. García, X.G. i Nieto, et al., H3D-Net: Few-shot high-fidelity 3D head reconstruction, in: Proc. IEEE/CVF Int. Conf. Comput. Vis., ICCV, 2021, pp. 5600–5609, <http://dx.doi.org/10.1109/ICCV48922.2021.00557>.
- [51] S. Baksi, S. Freezer, T. Matsumoto, C. Dreyer, Accuracy of an automated method of 3D soft tissue landmark detection, *Eur. J. Orthod.* 43 (6) (2021) 622–630, <http://dx.doi.org/10.1093/ejo/cjaa069>.
- [52] K.E. Stull, M.L. Tise, Z. Ali, D.R. Fowler, Accuracy and reliability of measurements obtained from computed tomography 3D volume rendered images, *Forensic Sci. Int.* 238 (2014) 133–140, <http://dx.doi.org/10.1016/j.forsciint.2014.03.005>.
- [53] Z. He, et al., FST-Net: Facial soft tissue landmark localization on 3dMD scans using feature fusion and local coordinate regression, in: 2024 IEEE International Symposium on Biomedical Imaging, ISBI, pp. 1–5, <http://dx.doi.org/10.1109/ISBI56570.2024.10635114>.