DDF-ISM: Internal Structure Modeling of Human Head Using Probabilistic Directed Distance Field

Zhuoman Liu, Yan Luximon*, Wei Lin Ng, and Eric Chung

***Abstract*—The increasing interest surrounding 3D human heads for digital avatars and simulations has highlighted the need for accurate internal modeling rather than solely focusing on external approximations. Existing approaches rely on traditional optimization techniques applied to explicit 3D representations like point clouds and meshes, leading to computational inefficiencies and challenges in capturing local geometric features. To tackle these problems, we propose a novel modeling method called DDF-ISM. It leverages a probabilistic *Directed Distance Field for Internal Structure Modeling*, facilitating efficient and anatomically accurate deformation of different parts of the human head. DDF-ISM comprises two key components: 1) a probabilistic DDF network for implicit representation of the target model to provide crucial local geometric information, and 2) a conditioned deformation network guided by the local geometry. Additionally, we introduce a large-scale dataset of human heads with internal structures derived from high-quality Computed Tomography (CT) scans, along with well-designed template models encompassing skull, mandible, brain, and head surface. Evaluation on this dataset showcases the superiority of our approach over existing methods, exhibiting superior performance in both modeling quality and efficiency.**

***Index Terms*—Head modeling, anatomical model, mesh deformation, distance field.**

I. Introduction

The burgeoning demand for precise 3D human head modeling, driven by advancements in digital avatars [1, 2, 3] and simulation technologies [4, 5], has underscored the critical need to accurately represent not only the external appearance but also the complex internal anatomies of the human head. Such detailed representations are critical for applications ranging from surgical simulations to anthropometric analyses [6, 7], necessitating models that encompass intricate internal structures like bone tissues and the brain. However, obtaining an anatomically accurate model of the internal structures of human heads poses significant challenges.

The first challenge lies in acquiring data related to internal head structures. While external surfaces can be readily captured through conventional scanning technologies and multi-view reconstructions [8, 9, 10], securing detailed internal structural data necessitates more sophisticated approaches. Medical imaging techniques, such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans, offer a pathway but come with their own problems, including noises and difficulties in aligning these detailed scans with predefined templates. Furthermore, the oversimplification of template models for internal structures [7, 11, 12] are limited to provide enough information for further applications.

The second challenge is that traditional methods [7, 12, 13, 14] primarily employ optimization techniques on explicit 3D representations like meshes and point clouds, aiming to match correspondences and model internal structures. However, these methods often encounter severe limitations, including computational inefficiency and difficulties in capturing fine local geometric details. IFMatch [15] proposes an efficient matching approach based on signed distance fields (SDFs) [16]. An SDF-based method utilizes volumetric representations to encode shape information but may face challenges with symmetric objects [17], where different locations can share the same distance value despite distinct geometric meanings, a situation frequently observed in anatomical structures.

Addressing the challenge of data acquisition, we present a comprehensive dataset featuring detailed internal structures of human heads derived from high-fidelity CT scans. This dataset, covering a demographic spread of 100 adults aged 18 to 65, is designed to provide a wide range of anatomical accuracy. Accompanied by precisely designed template models, including the skull, mandible, brain, and head surface, this resource aims to overcome the prevalent issues of noise and alignment common in medical imaging data. The inclusion of high-resolution imagery and detailed templates makes this dataset a critical source for enhancing the accuracy of internal structure modeling.

In response to the challenge in modeling, we propose a novel approach, DDF-ISM, which leverages a probabilistic directed distance field (DDF) [18] for modeling the anatomical internal structures of human heads. The DDF-ISM framework consists of two primary components: 1) a conditioned deformation network trained to learn deformations on a template model and 2) a probabilistic DDF network that implicitly represents the target model and provides crucial local geometric information to the deformation network. By integrating *directed* distance fields and *conditioning* mechanism, our method enables efficient deformation of various symmetric components of the head while maintaining anatomical accuracy. Furthermore, through rigorous evaluations of the proposed dataset, we demonstrate the superior performance of our method compared to existing point-based and SDF-based modeling methods, both in terms of modeling quality and computational efficiency.

Zhuoman Liu is with the School of Design, The Hong Kong Polytechnic University, Hong Kong, China. E-mail: zhuo-man.liu@connect.polyu.hk

Yan Luximon is with the School of Design, The Hong Kong Polytechnic University, Hong Kong, China, and also with the Laboratory for Artificial Intelligence in Design, Hong Kong, China. E-mail: yan.luximon@polyu.edu.hk

Wei Lin Ng and Eric Chung are with the Department of Biomedical Imaging, Universiti Malaya Medical Centre, University of Malaya Research Imaging Centre (UMRIC), Faculty of Medicine, Universiti Malaya.

* Corresponding author.

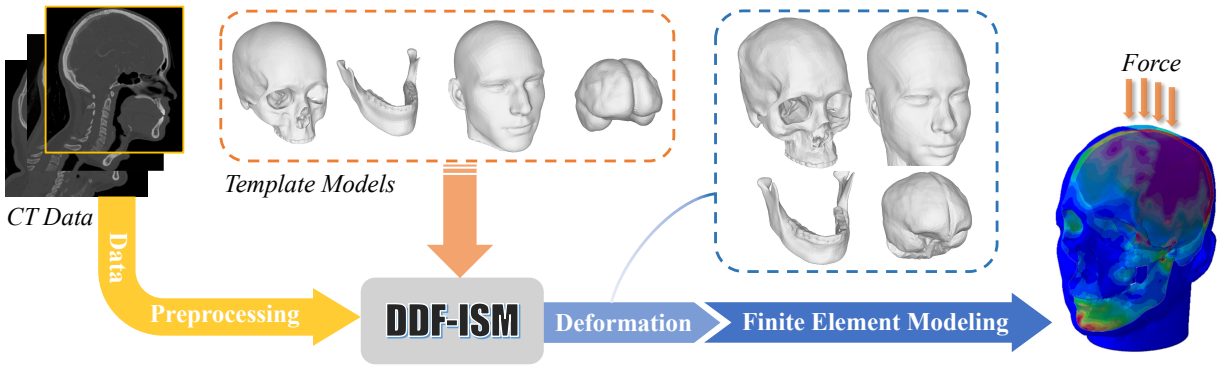

Fig. 1: **Internal structure modeling workflow.** This diagram outlines the key stages from CT data acquisition to deformation and simulation, highlighting the integration of DDF-ISM and finite element modeling for precise biomechanical analysis.

Our main contributions are summarized as follows:

- We propose DDF-ISM, a novel modeling framework for accurately modeling the internal structures of human heads with geometric guidance.
- We integrate the directed distance field (DDF) to capture geometry features that are robust to symmetry, enabling efficient and anatomically accurate deformations.
- We construct a comprehensive dataset of human heads with internal structures derived from CT scans, along with well-designed template models for benchmarking.
- Extensive experiments show the superior performance of DDF-ISM over existing point-based and SDF-based modeling methods. We also conduct the application of finite element modeling based on our deformed head models to show the effectiveness of the proposed method.

II. Related Work

A. Data Sources for 3D Head Modeling

The foundation of 3D human head modeling resides in the meticulous acquisition of precise data. Traditional methodologies have predominantly focused on external detail capture through advanced techniques such as laser scanning, structured light systems, and photogrammetry. The most esteemed human head datasets [19, 20, 21, 22] in the literature, constructed utilizing multi-camera stereo systems, offer comprehensive variations in pose, shape, and expression. However, these approaches often do not adequately address the capture of internal anatomical features essential for a broad range of applications, including detailed medical analyses and the generation of realistic digital human avatars.

Recent developments have prompted a shift toward the integration of medical imaging technologies, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), for the acquisition of intricate internal anatomical details. CT scans, distinguished by their superior spatial resolution, have emerged as invaluable for their precise differentiation between bone and soft tissue, offering remarkable clarity. This advancement in data collection methods represents a notable transition from mere surface capture to the acquisition of detailed internal structures, broadening the utility of models across diverse fields, from medical diagnostics to the animation of digital humans. Notably, [23] introduces a human head dataset integrating full-head CT and facial scans, aimed at advancing craniofacial reconstruction techniques. Besides, SCULPTOR [12] presents the LUCY dataset, a unique collection of shape and skeleton-correlated facial data derived from pre- and post-surgery CT scans, designed to support further facial editing applications.

In addition, template models of human head for parameterization have a profound impact on the functionality of subsequent simulation applications. The design strategy of template models significantly varies, directly influencing the scope and precision of the model's applicability. Specifically, methods focusing on facial features [19, 24] tend to only parameterize the facial area, potentially overlooking the overall representation necessary for comprehensive applications. Moreover, current approaches that consider internal anatomical structures [7, 11, 12] employ overly simplified template models for parameterization, lacking in detail. Such simplifications can limit the effectiveness of the model in simulations, which require high fidelity to real-world anatomical complexities.

B. 3D Morphable Models

3D Morphable Models (3DMMs) have emerged as a key technique in computer vision and graphics, enabling the representation and manipulation of 3D shapes and textures from limited data inputs. Traditional 3DMMs [25] create a statistical model of 3D facial shape and texture, allowing for the generation of novel faces by altering model parameters. The integration of deep learning with 3DMMs has revolutionized the field. [26] proposed a deep convolutional neural network to learn 3DMM parameters directly from images, bypassing the traditional optimization-based fitting methods. This approach, termed Deep 3DMMs, demonstrated superior performance in terms of speed and accuracy, facilitating real-time applications.

Recent works have extended 3DMMs to encompass additional anatomical structures. For instance, FLAME [19] incorporates facial expressions, enhancing the dynamic representation of facial movements. A combined model of the scalp and skull [11] offers a more comprehensive head model, essential for medical and forensic applications. Addressing the variability in scalp shape under hair, a parametric head model considering scalp shape under hair was proposed [24]. SCULPTOR [12] integrates internal skull and mandible structures, ensuring anatomical consistency in generated faces. A

head geometry model accounting for variations in brain structures among different age groups has been introduced [7], beneficial for developmental biology and characterized medicine.

Despite these advancements, a comprehensive model that includes more internal anatomical structures of the head is still evolving. Our approach aims to address this gap by developing a morphable model that includes more detailed head internal structures. By integrating features of the head surface, skull, mandible, and brain, our model enhances accuracy and utility across various applications.

C. 3D Head Deformation

In the process of parametric modeling, vertices are sampled on meshes for matching with the template model, followed by the deformation of the template. Traditional methods typically involve point-based optimization during modeling. A classic non-rigid deformation method focused on points is the Coherent Point Drift (CPD) [27], which also has a variant guided by landmarks, known as LGCPD [13]. However, when the number of vertices to be modeled is large or the deformation is significant, these methods require substantial computational memory and suffer from reduced processing speed, leading to less accurate deformation results.

To address these challenges, some approaches utilize hierarchical structures to reduce computational memory while simultaneously improving modeling precision. For instance, Neural Deformation Pyramid (NDP) [14] utilizes a pyramid structure with Multi-Layer Perceptrons (MLPs) [28] at each level to process sinusoidally encoded 3D points, enabling incremental non-rigid deformations. Similarly, SCULPTOR [12] utilizes control nodes to accelerate deformation, incorporating landmarks for more precise guidance. Besides, attention mechanisms have been employed to enhance shape deformation accuracy by capturing geometric relationships. [29] is the first to apply an attention module for shape deformation, showcasing its potential in non-rigid transformations.

Despite these advancements, point-based modeling approaches continue to be encumbered by their computational speed. Leveraging advancements in implicit functions, IF-Match [15] introduces a paradigm shift by utilizing Signed Distance Functions (SDFs) [16] for geometrically representing target objects, coupled with the deployment of an MLP-based network for learning deformations, thus significantly enhancing computational efficiency. Nonetheless, the application of SDF-based matching to anatomical head modeling can benefit from enhancements in capturing geometric features. While SDF effectively encodes shape information, incorporating directional information (*i.e.*, DDF [18]) alongside point queries may offer advantages in better representing the intricate geometry of symmetrical structures [17]. This approach could provide a more detailed representation of anatomical features compared to traditional SDF-based methods.

III. Overview

In this work, we introduce DDF-ISM, a novel parametric anatomical head model that integrates comprehensive internal structures to enhance medical imaging simulations and diagnostics. Our approach includes the development of a large-scale anatomical head dataset comprising CT scans, segmented meshes, labeled landmarks, and meticulously designed template models. We propose the utilization of a directed distance field (DDF) to innovatively guide the mesh deformation. This methodology aims to tackle significant challenges in data construction and parametric modeling, particularly those related to anatomical accuracy and computational efficiency. The structure of this paper is organized as follows: we begin with a detailed description of the data acquisition and processing, proceed to elaborate on the network architectures, the training and inference processes, and conclude with experimental results and analyses.

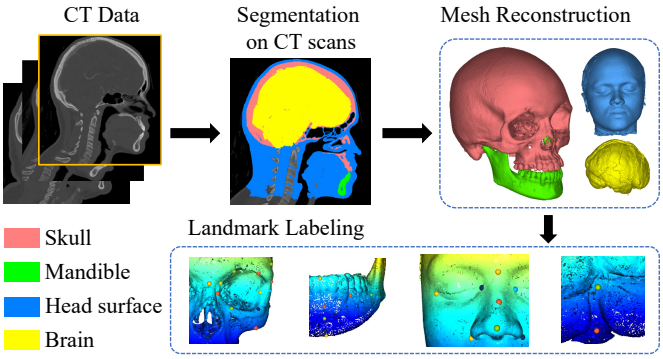

Fig. 2: Overall workflow of data preprocessing of CT scans.

IV. Data Construction

A. Data Acquisition of CT Scans

In collaboration with experts from a Malaysian hospital, we ethically collected CT data following approval from the university's ethics committee (HSEARS20210804001). This dataset encompasses all parts of the human head, clearly capturing the necessary anatomical details that are essential for subsequent 3D model fitting. Specifically, this dataset includes scans from 100 Malaysian adults, whose ages range from 18 to 65. To facilitate robust analysis and minimize biases associated with age and gender, we organized the participants into five age groups (*i.e.*, 18-25, 26-35, 36-45, 46-55, and 56-65), collecting 20 CT scans for each category. Each group maintains an equal distribution of male and female participants, thereby reducing potential analytical discrepancies due to age and gender differences. These scans were captured at a resolution of $0.54 \times 0.54 \times 1\ mm^3$, providing the high-resolution data necessary for precise anatomical modeling and the accurate depiction of complex structures in subsequent parametric modeling. This high level of detail is crucial for the effective segmentation and modeling processes that follow.

B. Data Preprocessing

The preprocessing of the collected CT scans involves two critical steps: segmentation and downsampling. In the segmentation phase, we employ Materialise Mimics[1] to delineate

[1]www.materialise.com/en/healthcare/mimics-innovation-suite/mimics

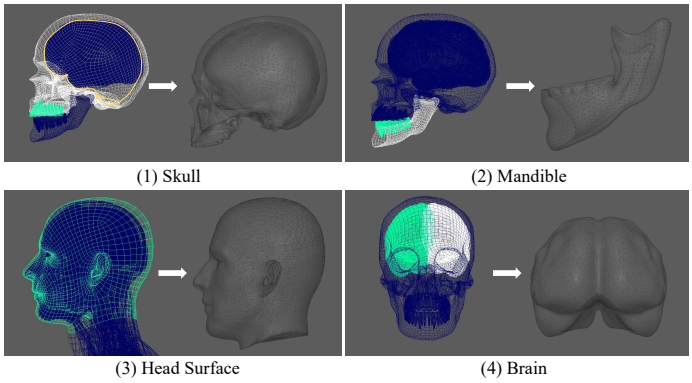

Fig. 3: Template construction from source models.

four key anatomical parts of the human head (*i.e.*, skull, mandible, head surface, and brain) from a series of 2D CT images, subsequently reconstructing individual 3D meshes for each part. Initially, we apply predefined thresholds within the software to automatically extract masks for bone and soft tissue structures. Following this, as illustrated in Fig. 2, the bone is further segmented into two distinct components: the skull (shown in pink) and the mandible (shown in green), through manual labeling of multiple masks across three views, *i.e.*, axial, coronal, and sagittal. In a similar manner, the soft tissue is divided into the head surface (shown in blue) and the brain (shown in yellow). Once segmentation is complete, the software calculates the meshes for these four structures.

Given that the original meshes are overly dense for effective 3D registration, a downsampling process is implemented. This involves using a uniform sampling filter to reduce mesh density according to a specified percentage, denoted by σ. Specifically, we set $\sigma = 0.8$ for the skull and head surface, and $\sigma = 1.5$ for the mandible and brain, corresponding to a downsampling factor of approximately 100. This adjustment ensures that the meshes are optimized for subsequent computational tasks without compromising the structural integrity and detail required for accurate modeling and analysis.

C. Template Model Construction

Prior to exploring the approach for 3D mesh deformation, it is imperative to develop template models for the four key anatomical components of the human head. We acquired a detailed 3D human model, including comprehensive anatomical structures, from TurboSquid. Using AutoDesk Maya, we extracted and refined key components for template construction: the skull with upper teeth, the mandible with bottom teeth, the head surface, and the brain, as shown in Fig. 3. Refinements included merging and cleaning intersections, closing open surfaces, aligning the brain with the skull's inner surface, and resampling and smoothing meshes. These processes ensured the templates retained sufficient detail for comprehensive biomechanical analyses.

The refined template models are characterized by their intricate structures and are depicted in Fig. 4, which provides both a visual representation and detailed geometric data under X-Ray rendering. We denote the template model as $\bar{\mathbf{T}} = \{\bar{\mathbf{T}}_{skl}, \bar{\mathbf{T}}_{mdb}, \bar{\mathbf{T}}_{surf}, \bar{\mathbf{T}}_{brn}\}$, including skull, mandible, head surface, and brain. Specifically, the skull model comprises $N_{skl} = 21,294$ vertices with $42,597$ polygons; the mandible model includes $N_{mdb} = 4,000$ vertices with $8,008$ polygons; the head surface is detailed with $N_{surf} = 20,234$ vertices and $40,464$ polygons; and the brain includes $N_{brn} = 4,838$ vertices with $9,672$ polygons. Additionally, we build an internal structure of the skull closely matching the brain's surface to facilitate realistic simulations.

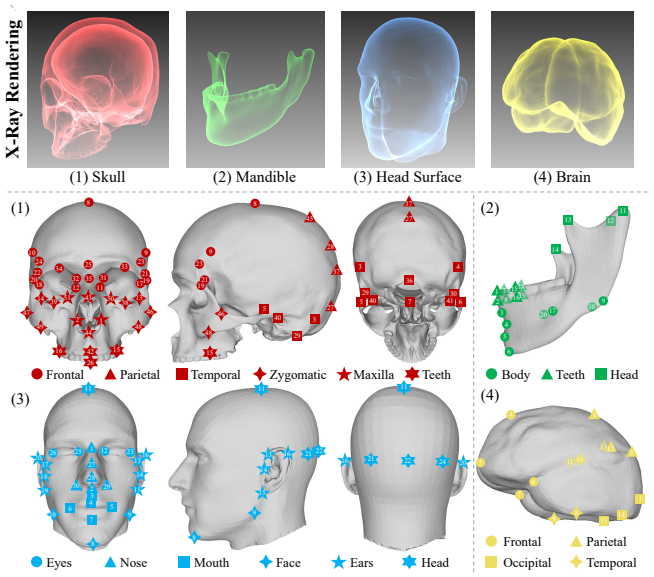

Fig. 4: Anatomical template models with regional landmarks of four key parts of human head.

D. Landmark Labeling

In conjunction with the development of templates and subsequent preprocessing of meshes, we define and label anatomical landmarks to facilitate semantic information for modeling. As illustrated in Fig. 4, we define $K_{skl} = 51$ landmarks for the skull, $K_{mdb} = 14$ for the mandible, $K_{surf} = 30$ for the head surface, and $K_{brn} = 15$ for the brain. The definitions of these landmarks are guided by the anatomy of each part, drawing from established standards in the field. Specifically, the landmark definitions for the skull and mandible models are adapted from [30]. Specifically, we separate the skull into six anatomical regions, including frontal, parietal, temporal, zygomatic, maxilla, and teeth, while mandible is divided into three regions, *i.e.*, body, teeth, and head. Similarly, we separate the head surface into six regions, *i.e.*, eyes, nose, mouth, face, ears, and head, while brain is divided into four regions, including frontal, parietal, occipital, and temporal. We split out these anatomical regions to further analyze the modeling errors of each region. For the landmarks annotated for templates, we denote them as $\bar{\mathbf{x}}_{lmk} = \{\bar{\mathbf{x}}^{skl}_{lmk}, \bar{\mathbf{x}}^{mdb}_{lmk}, \bar{\mathbf{x}}^{surf}_{lmk}, \bar{\mathbf{x}}^{brn}_{lmk}\}$. For the landmarks labeled for the target meshes, we denote them as $\mathbf{x}_{lmk} = \{\mathbf{x}^{skl}_{lmk}, \mathbf{x}^{mdb}_{lmk}, \mathbf{x}^{surf}_{lmk}, \mathbf{x}^{brn}_{lmk}\}$.

Besides, before internal structure modeling, we apply rigid alignment using iterative closest point (ICP) algorithm [31] to coarsely align the preprocessed target meshes to the templates.

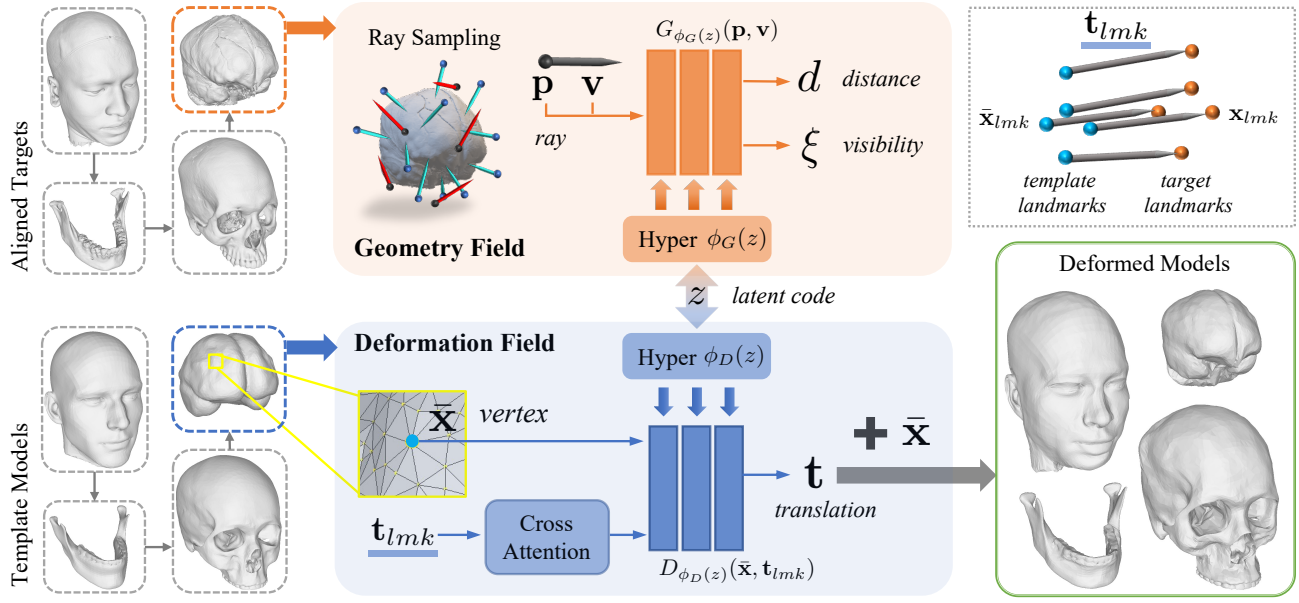

Fig. 5: The overall pipeline of DDF-ISM. Starting with the alignment of target and template models, the pipeline proceeds through the stages of ray sampling, distance computation, and visibility assessment. It then leverages the geometry and deformation fields, guided by latent codes, to achieve precise deformation of the template model. The cross-attention mechanism ensures the accurate translation of vertices, resulting in a deformed model that closely matches the target's anatomical features.

We select 5 landmarks from $\bar{\mathbf{x}}_{lmk}^{surf}$ and $\mathbf{x}_{lmk}^{surf}$ (label 8, 11, 22, 23, 26 in Fig. 4 (3)), respectively. With the paired landmarks, we compute the coarse transformation and scale and perform them on the four anatomical parts of target meshes to obtain aligned targets $\mathbf{M} = \{\mathbf{M}_{skl}, \mathbf{M}_{mdb}, \mathbf{M}_{surf}, \mathbf{M}_{brn}\}$. With the template models $\bar{\mathbf{T}}$ and aligned targets $\mathbf{M}$, we discuss how to model these internal structures of human head in next section.

V. Internal Structure Modeling

The process of internal structure modeling is crucial for creating an accurate and comprehensive 3D representation of the human head. Our approach meticulously integrates various anatomical details, including the skull, mandible, head surface, and brain, to enhance both the accuracy and utility of the model. We develop a **geometry field** to capture local geometric features of each anatomical part and propose a **deformation field** to learn the deformation on vertices from templates to targets. A *latent code* is then introduced to transfer geometric information from the geometry field and as the *condition* for the deformation field. The overall pipeline of our method is shown in Fig. 5.

A. Geometry Field

The geometry field G is a probabilistic directed distance field (DDF) [18], which takes a ray originated at point $\mathbf{p}$ with a view direction $\mathbf{v}$ as input and predicts the distance value d and binary visibility ξ. The visibility $\xi = 1$ indicates the surface is visible from the ray viewpoint. When the surface is visible, we have distance value $d \geq 0$, which is the distance from $\mathbf{p}$ to the surface along $\mathbf{v}$. To transmit the geometric information to the subsequent deformation field, we introduce a latent code $z \in \mathbb{R}^{C_z}$ and a hypernetwork [32] $\phi_G(z)$. Since the DDF requires only one query to obtain the surface distance, it is efficient to provide geometric information. Meanwhile, the query here is directional, so the geometric information provided by the DDF is not ambiguous for the approximately symmetric internal structure of the human head.

Ray sampling. We sample multiple rays in free space according to some criteria to learn the whole geometry of the target mesh. Following [18], we apply three data types for ray sampling, *i.e.*, *Uniform (U)*, *Bounding (B)*, and *Surface (S)*. We visualize the sampled rays in Fig. 5, with rays that successfully hit the surface colored in green ($\xi = 1$) and rays that miss the surface colored in red ($\xi = 0$). More details of ray sampling are in Sec. VII-A.

Loss terms. During ray sampling, it is intuitive to acquire the ground-truth distance d and visibility ξ. Therefore, we have distance loss $\mathcal{L}_d$ using ℓ_1 loss and visibility loss $\mathcal{L}_\xi$ using binary cross entropy as follows:

$$\mathcal{L}_d = \xi|\hat{d} - d|, \text{where } \hat{d} = \hat{d}_{i^*},\ i^* = \arg\max_i w_i \quad (1)$$

$$\mathcal{L}_\xi = BCE(\hat{\xi}, \xi) \quad (2)$$

In the paper, we set $K = 2$, which means the geometry network outputs two distance values with weights w_0, w_1 and $\sum_i w_i = 1$ to handle the discontinuity of vanilla DDF. Besides, we adopt the proposed weight field regularizations in [18], *i.e.*, weight variance loss $\mathcal{L}_V = \prod_i w_i$ and weight transition loss $\mathcal{L}_T = \max(0, \varepsilon_T - |\nabla_{\mathbf{p}} w_0 \hat{\mathbf{n}}|)$. Specifically, $\hat{\mathbf{n}}$ is the surface normals computed from the derivative of predicted distance $\hat{d}$ w.r.t. the ray view point $\mathbf{p}$, *i.e.*, $\hat{\mathbf{n}} = \tau \nabla_{\mathbf{p}} \hat{d} / \|\nabla_{\mathbf{p}} \hat{d}\|^2$, where $\tau \in \{-1, 1\}$ is the sign to make sure $\hat{\mathbf{n}}^T \mathbf{v} < 0$.

Moreover, according to the geometric properties of DDF [18], we have constraints on distance, visibility, and

surface normals. We adopt the directed Eikonal regularization $\mathcal{L}_{DE}$ and first-order normals loss $\mathcal{L}_{\mathbf{n}}$ from [18] as follows:

$$\mathcal{L}_{DE} = \lambda_{E,d} \sum_{i=0}^{K} \xi(\nabla_{\mathbf{p}} \hat{d}_i \mathbf{v} + 1)^2 + \lambda_{E,\xi}(\nabla_{\mathbf{p}} \hat{\xi} \mathbf{v}) \quad (3)$$

$$\mathcal{L}_{\mathbf{n}} = -\xi |\mathbf{n}^T \hat{\mathbf{n}}_{i^*}| \quad (4)$$

In summary, we have a geometry loss $\mathcal{L}_G = \lambda_d \mathcal{L}_d + \lambda_\xi \mathcal{L}_\xi + \lambda_V \mathcal{L}_V + \lambda_T \mathcal{L}_T + \lambda_{DE} \mathcal{L}_{DE} + \lambda_{\mathbf{n}} \mathcal{L}_{\mathbf{n}}$.

B. Deformation Field

Mesh deformation is a critical component of fitting the morphable model to the target data. The deformation field $D_{\phi_D(z)}(\bar{\mathbf{x}}, \mathbf{t}_{lmk})$ operates on the vertex $\bar{\mathbf{x}}$ from template model with the landmark translations $\mathbf{t}_{lmk} = \mathbf{x}_{lmk} - \bar{\mathbf{x}}_{lmk}$, then learn the translation $\mathbf{t}$ from the template to target. By applying the predicted translation on the template vertex, we can obtain the deformed mesh with new vertex $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \hat{\mathbf{t}}$. It is noted that $\mathbf{t}_{lmk}$ is fed into a cross-attention module [33] to extract the potential geometric relationship between the template and targets. Then the landmark translation features are concatenated with $\bar{\mathbf{x}}$ as inputs of D. Similar to the geometry field, we introduce a hypernetwork $\phi_D(z)$ for the deformation field. The latent code z passes geometric features through ϕ_D to the deformation network D. Therefore, this field allows for the precise translation of vertices, ensuring that the deformed models accurately represent the target anatomical structures. The cross-attention mechanism facilitates the refinement of these deformations by considering the relationship between the template and target landmarks.

Loss terms. After deformation, we have a new set of vertices of each part, denoted as $\hat{\mathbf{X}}_{part} \in \mathbb{R}^{N_{part} \times 3}$, where $part \in \{skl, mdb, surf, brn\}$. With the vertices $\mathbf{X}_{part} \in \mathbb{R}^{N \times 3}$ from target $\mathbf{M}_{part}$ (N is the number of vertices of target mesh), we minimize ℓ_2 Chamfer-Distance [34] for shape fitting.

$$\mathcal{L}_{CD} = CD(\hat{\mathbf{X}}_{part}, \mathbf{X}_{part}) \quad (5)$$

In addition, we minimize the ℓ_2 distances between deformed landmarks and target landmarks using a landmark loss $\mathcal{L}_{lmk}$. To preserve the smoothness of deformed surface, we add a smoothness regularization $\mathcal{L}_{smooth}$ by controlling the gradients of deformation field.

$$\mathcal{L}_{lmk} = \|\hat{\mathbf{x}}_{lmk} - \mathbf{x}_{lmk}\|_2 \quad (6)$$

$$\mathcal{L}_{smooth} = \|\nabla_{\bar{\mathbf{x}}} t\|_2 \quad (7)$$

In summary, we have a deformation loss $\mathcal{L}_D = \lambda_{CD} \mathcal{L}_{CD} + \lambda_{lmk} \mathcal{L}_{lmk} + \lambda_{smooth} \mathcal{L}_{smooth}$.

C. Training

When training both the geometry field and deformation field, we introduce a regularization $\mathcal{L}_z = \|z\|_2$ on latent code. Therefore, when learning the morphable model of one of the anatomical parts, we minimize the total loss $\mathcal{L} = \mathcal{L}_G + \mathcal{L}_D + \mathcal{L}_z$. Furthermore, as shown in Fig. 5, to reduce collisions among the four anatomical parts, we learn the morphable models in the order of brain, skull, mandible, then head surface. For collision detection, we convert the DDF values to signed distances [18] and eliminate intersections by pushing vertices out of the volume.

D. Inferencing

Our method employs an optimization-based strategy, leveraging hypernetworks to dynamically generate parameters for shape deformation during fitting. To infer deformed models of new subjects, we first load the pre-trained geometry field G and minimize $\mathcal{L}_G$ to obtain initial latent code $\hat{z}$. Subsequently, we load the pre-trained deformation field D and minimize $\mathcal{L}_{CD}$ and $\mathcal{L}_{lmk}$ to obtain the optimized latent code $\tilde{z}$ (excluding the smoothness term to reduce gradient computation time). With $\tilde{z}$, we can infer the deformation field D to obtain translations and then get the morphable model.

VI. Application

Finite Element Analysis (FEA) is a critical application of our anatomical head model. By providing accurate representations of the skull, mandible, head surface, and brain, it is straightforward to use our model for precise simulation of physical forces and their effects on human head. This is particularly valuable in medical fields such as neurosurgery, orthodontics, and craniofacial surgery, where understanding the impact of surgical interventions on the entire head anatomy is crucial. In this study, we use a software called ABAQUS for FEA, and the pipeline of FEA is as follows:

1) **Mesh Generation:** We first randomly select one participant from our dataset, then by applying our DDF-ISM on the CT data, we can obtain the morphed models of skull, mandible, head surface, and brain. Before performing FEA, we add ligaments between skull and mandible (label 40 of $\mathbf{x}^{skl}_{lmk}$ and label 11 of $\mathbf{x}^{mdb}_{lmk}$, label 41 of $\mathbf{x}^{skl}_{lmk}$ and label 12 of $\mathbf{x}^{mdb}_{lmk}$) as constraints. Besides, a boolean operation is conducted on the bone (including both skull and mandible) and head surface to obtain the mesh of soft tissue. Then, the skull, mandible, soft tissue, and brain can be assembled to form a Parasolid file of the human head as a finite element model.
2) **Material Properties:** After loading the above finite element model in ABAQUS, we set appropriate material properties [35] for different head regions as shown in Table I, where ρ is density, E is elastic modulus, and ν is poison ratio. The material properties are assigned to corresponding parts respectively.
3) **Impact Force and Boundary Condition:** To simulate the impacts of an external force on different parts of the human head, we apply an impact force of 100N directed along the negative Y-axis at the head vertex, over an area of 50×50 mm^2, as shown in Fig. 6 (a). The bottom of the soft tissue (neck) is constrained as a fixed boundary condition throughout the simulation process.
4) **Simulation and Analysis:** In Fig. 6 (b) and (c), we show the simulation results at two scales after applying the impact force on the head vertex. For visualization, the maximum von Mises stress is 0.2 MPa for the soft tissue scale and 50 MPa for bone scale. The top of soft tissue,

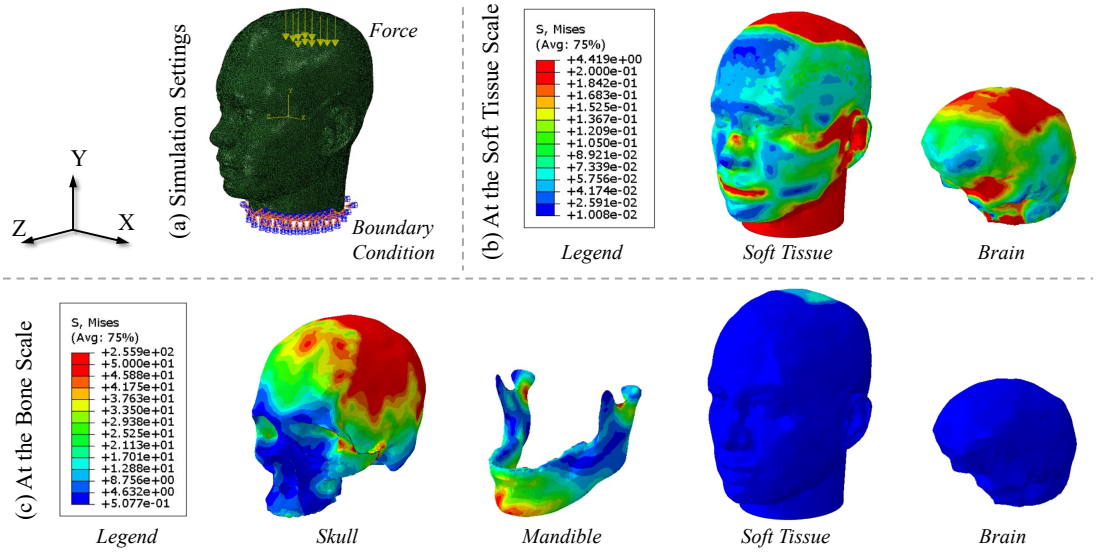

Fig. 6: Illustration of simulation settings and FEA results of different parts. This illustration presents the Von Mises stress distribution across various components within a biomechanical model. The legend provides a guide to the stress values, indicating the structural integrity and potential areas of concern within the model.

TABLE I: Material properties for different anatomical parts in a biomechanical model.

Part	$\rho \times 10^{-8}$ (kg/mm^3)	E (MPa)	ν	Matreial Type	Element Type
Bone and ligament	193.5	4096.8	0.22	Elastic	C3D4
Soft tissue	140.0	0.7	0.45	Elastic	C3D4
Brain	104	1.58	0.5	Viscoelastic	C3D4H

brain, and skull are in red, indicating the highest stress. It is noted that the bottom of soft tissue, *i.e.*, the neck part, and the bottom of mandible are also in red, it is because the constraints when applying impact force are set at the neck, therefore, the bottom of soft tissue takes counterforce and squeezes the bottom of mandible. The results show accurate stress of the four anatomical parts according to the force, demonstrating that our DDF-ISM is feasible to work for precise simulations.

VII. Experiments

A. Implementation Details

Our implementation leverages a multi-stage process for training the geometry and deformation fields. The training process begins with the acquisition of CT scans, followed by segmentation and mesh reconstruction of the skull, mandible, head surface, and brain. These segmented meshes serve as the basis for training our DDF-ISM.

For the **geometry field**, we utilize a combination of *Uniform (U)*, *Bounding (B)*, and *Surface (S)* ray sampling strategies [18] to train, ensuring comprehensive coverage of the target mesh's geometry. In the Uniform way, a point is sampled randomly from within a specified boundary, and a direction is chosen randomly from the surface of a sphere. In the Bounding way, the point is selected randomly from the boundary itself, and the direction is also chosen from the surface of the sphere, but with the restriction that it points toward the interior of the boundary. In the Surface way, both the point and the direction are selected randomly from the surface of a sphere. To optimize the geometry field, the geometry loss includes distance loss $\mathcal{L}_d$, visibility loss $\mathcal{L}_\xi$, weight variance loss $\mathcal{L}_V$, and weight transition loss $\mathcal{L}_T$, with specific regularizations (*i.e.*, $\mathcal{L}_{DE}$ and $\mathcal{L}_{\mathbf{n}}$) to handle the symmetry and discontinuities of the human head's internal structures.

For the **deformation field**, we implement the cross-attention module following [33]. The cross-attention module captures the geometric relationship between the template and target models by processing the landmark translations. This helps to enhance the accuracy of the deformation process by allowing the model to better understand how the template should be adjusted to match the target. The deformation loss incorporates Chamfer Distance loss $\mathcal{L}_{CD}$, landmark loss $\mathcal{L}_{lmk}$, and smoothness regularization $\mathcal{L}_{smooth}$, balancing shape-fitting accuracy with smooth surface transitions.

To train the two fields for each internal structure of the human head, we use SIREN (MLPs using Sine activation) [36]. The geometry field $G_{\phi_G(z)}$ is modeled using three hidden layers with 256 hidden units, while the deformation field $D_{\phi_D(z)}$ employs five hidden layers to incorporate the cross-attention features, with 512 hidden units. The latent dimension is set to 512. DDF-ISM is implemented by PyTorch and the network is optimized using the Adam optimizer [37] with a learning rate of 0.0001 over 50 iterations on a single NVIDIA RTX 3090 GPU.

The inference phase involves loading the pre-trained geometry and deformation fields and optimizing the latent codes to fit new subject data with the same settings as the above training phase. The operation of latent code optimization in the inference phase ensures that the deformed models accurately represent the anatomical structures of new subjects while maintaining computational efficiency.

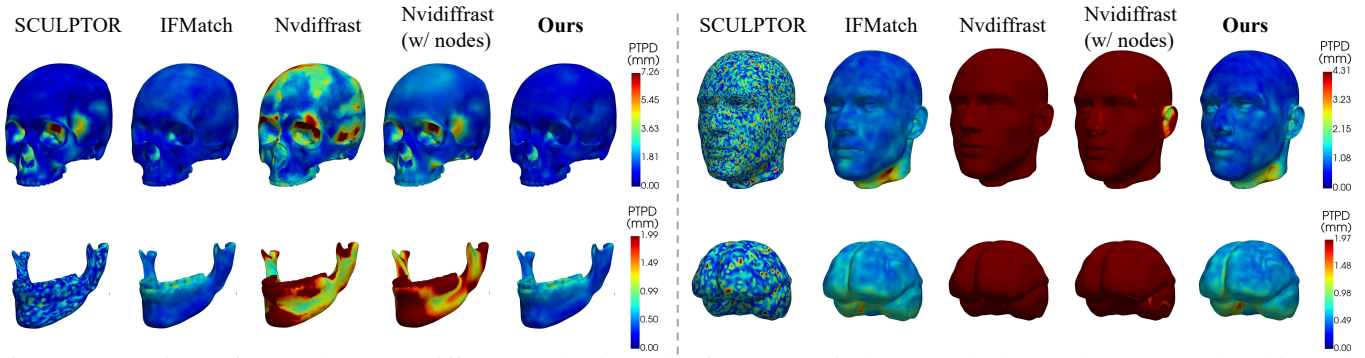

Fig. 7: Comparison of PTPD between different methods on the four anatomical parts. The lower the PTPD value (blue), the closer the deformed mesh is to the target mesh, indicating better performance.

TABLE II: **PTPD** (mm) ↓ and **RMSE**$_{lmk}$ (mm) ↓ (mean ± std. dev.) for different methods on four anatomical parts.

	Head Surface		Brain		Skull		Mandible		Avg.	
	PTPD	**RMSE**$_{lmk}$	**PTPD**	**RMSE**$_{lmk}$	**PTPD**	**RMSE**$_{lmk}$	**PTPD**	**RMSE**$_{lmk}$	**PTPD**	**RMSE**$_{lmk}$
SCULPTOR [12]	1.41	6.46 ± 2.78	**0.61**	6.73 ± 3.68	1.50	6.21 ± 3.63	**0.42**	2.86 ± 1.72	0.99	5.57 ± 2.95
IFMatch [15]	1.16	8.47 ± 4.11	0.68	23.02 ± 10.63	1.27	7.42 ± 4.87	0.51	4.40 ± 2.60	0.91	10.83 ± 5.55
Ours	**0.95**	**4.33 ± 2.34**	0.66	**3.78 ± 2.15**	**1.06**	**3.92 ± 2.48**	0.46	**2.39 ± 1.31**	**0.78**	**3.61 ± 2.07**
Nvdiffrast [38]	14.21	23.83 ± 10.14	7.64	29.62 ± 10.71	3.19	13.68 ± 7.76	2.33	10.86 ± 5.56	6.84	19.50 ± 8.54
Nvdiffrast [38] (w/ nodes)	6.48	13.28 ± 18.79	13.08	28.63 ± 16.63	2.10	6.77 ± 4.37	2.11	7.06 ± 3.27	5.94	13.94 ± 10.77
Ours (w/o CA, dim=512)	0.97	7.27 ± 5.39	0.70	17.37 ± 8.52	1.11	5.94 ± 4.27	0.48	4.13 ± 2.27	0.82	8.67 ± 5.11
Ours (w/o CA, dim=256)	1.05	7.41 ± 5.47	0.71	17.83 ± 8.75	1.15	6.07 ± 4.28	0.51	4.29 ± 2.37	0.86	8.90 ± 5.22

TABLE III: Region-specific errors **RMSE**$_{reg}$ (mm) ↓ (mean ± std. dev.) for different methods on *head surface*.

	Eyes	Nose	Mouth	Face	Ears	Head	**Avg.**
SCULPTOR [12]	5.43 ± 2.36	4.26 ± 1.55	4.61 ± 2.12	15.31 ± 6.50	6.80 ± 2.98	**5.77 ± 2.66**	6.46 ± 2.78
IFMatch [15]	6.19 ± 3.11	4.06 ± 2.07	7.13 ± 3.87	11.59 ± 5.50	9.18 ± 4.69	15.29 ± 6.27	8.47 ± 4.11
Ours	**3.85 ± 1.90**	**2.97 ± 1.51**	**3.41 ± 1.68**	**5.85 ± 3.07**	**4.70 ± 2.86**	6.11 ± 3.22	**4.33 ± 2.34**

TABLE IV: Region-specific errors **RMSE**$_{reg}$ (mm) ↓ (mean ± std. dev.) for different methods on *brain*.

	Frontal	Parietal	Occipital	Temporal	**Avg.**
SCULPTOR [12]	5.16 ± 2.19	5.02 ± 1.90	7.17 ± 2.27	9.71 ± 2.31	6.73 ± 3.68
IFMatch [15]	16.35 ± 6.37	24.53 ± 8.19	22.26 ± 5.15	28.76 ± 5.80	23.02 ± 10.63
Ours	**3.57 ± 2.10**	**4.01 ± 2.22**	**3.55 ± 2.22**	**3.94 ± 2.05**	**3.78 ± 2.15**

B. Evaluations

Comparing methods. In our experiments, we compare the performance of our DDF-ISM with existing state-of-the-art methods SCULPTOR [12] and IFMatch [15], as well as differentiable rendering-based shape fitting method (*i.e.*, Nvdiffrast [38]). SCULPTOR models not only the head surface but also the internal structures of the skull and mandible. However, SCULPTOR is developed on explicit mesh geometric representation, making it significantly slower. IFMatch introduces the use of implicit geometric representation combined with deformation to enhance speed. This approach contrasts with SCULPTOR and aligns more closely with our method's goals of efficiency and effectiveness. To evaluate Nvdiffrast with multi-view L_2 loss on mesh normals, we optimize vertex deformations over 500 random viewpoints for 200 iterations. Additionally, we implement a version with control nodes (*w/ control nodes*) based on SCULPTOR's strategy. For a fair comparison, all baselines are supervised with the same sampled rays and points as ours, validated through 5-fold cross-validation. The 100 scans were randomized and divided into five groups (20 scans each), with one group used as the test set and the remaining four as the training set, repeated across five iterations.

Evaluation metrics. The evaluation metrics include: 1) point-to-point distance (**PTPD**) [39, 40] between deformed mesh and target mesh, 2) root mean square error (**RMSE**$_{lmk}$) for landmarks, and 3) root mean square error for specific regions (**RMSE**$_{reg}$). All the above metrics are in millimeters (mm) averaged across testing meshes.

Our results, as detailed in Table II, highlight the superior performance of our DDF-ISM over conventional point-based and Signed Distance Function (SDF)-based methods in terms of both PTPD and RMSE$_{lmk}$. The directed distance field approach facilitates an accurate representation of internal structures, while the cross-attention mechanism ensures precise alignment of template and target landmarks.

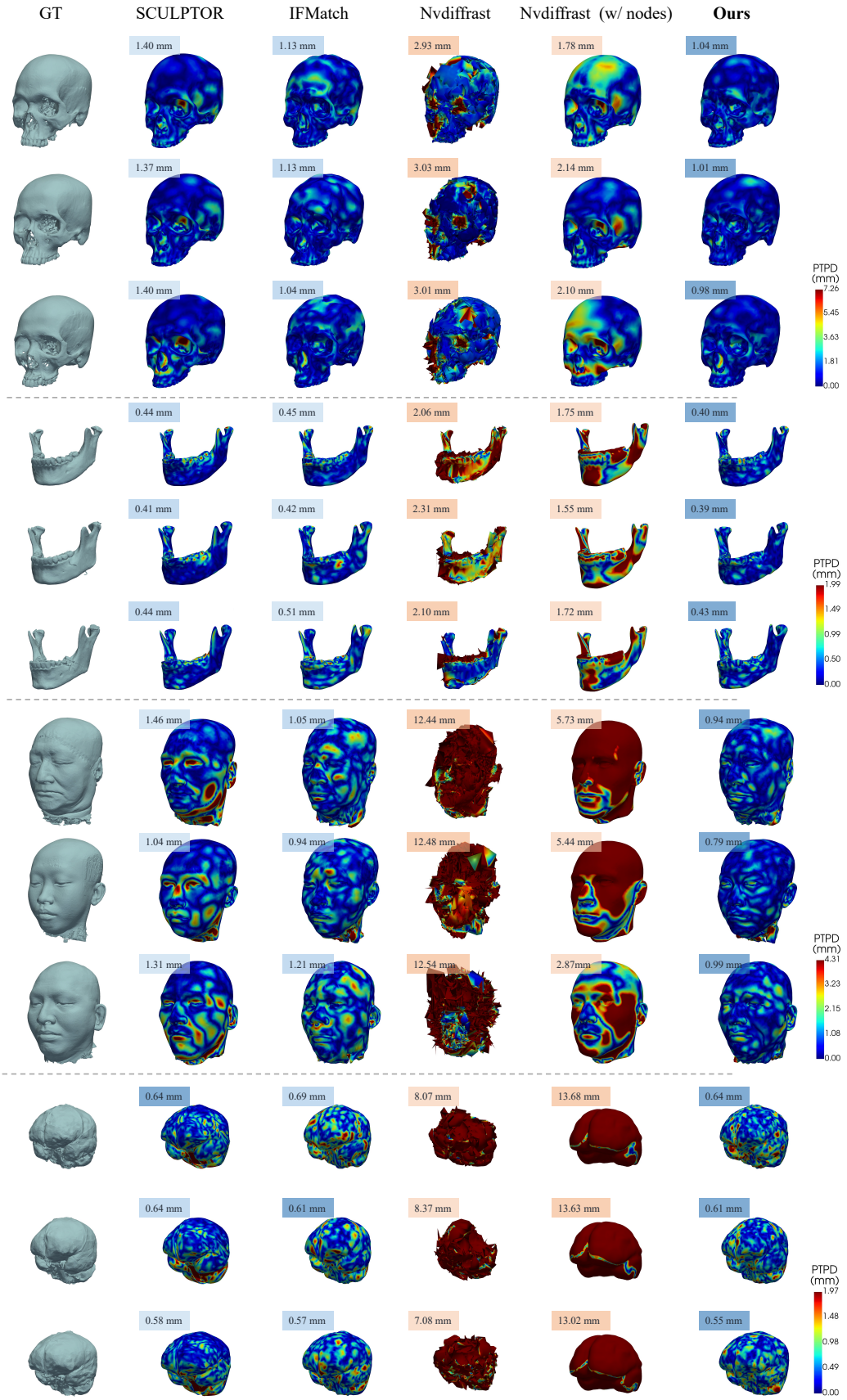

Fig. 8: Deformation result comparisons of our method and baselines. The visual representation of the deformations highlights the differences in accuracy and alignment between the methods.

TABLE V: Region-specific errors $\mathbf{RMSE}_{reg}$ (mm) ↓ (mean ± std. dev.) for different methods on *skull*.

	Frontal	Parietal	Temporal	Zygomatic	Maxilla	Teeth	**Avg.**
SCULPTOR [12]	5.09 ± 2.28	8.52 ± 3.52	6.22 ± 2.58	7.16 ± 2.56	6.87 ± 1.87	5.33 ± 3.18	6.21 ± 3.63
IFMatch [15]	7.70 ± 3.50	12.99 ± 5.70	6.99 ± 3.48	5.90 ± 3.16	7.10 ± 2.77	5.39 ± 3.71	7.42 ± 4.87
Ours	**3.95 ± 2.00**	**5.48 ± 2.67**	**3.92 ± 2.23**	**3.37 ± 2.07**	**3.65 ± 1.65**	**3.93 ± 2.30**	**3.92 ± 2.48**

TABLE VI: Region-specific errors $\mathbf{RMSE}_{reg}$ (mm) ↓ (mean ± std. dev.) for different methods on *mandible*.

	Body	Teeth	Head	**Avg.**
SCULPTOR [12]	2.46 ± 1.31	3.54 ± 1.38	2.58 ± 2.47	2.86 ± 1.72
IFMatch [15]	4.12 ± 2.21	5.30 ± 2.24	3.78 ± 3.35	4.40 ± 2.60
Ours	**2.41 ± 1.29**	**2.45 ± 1.34**	**2.31 ± 1.30**	**2.39 ± 1.31**

TABLE VII: Runtime comparisons (in minutes).

	Head	Surface	Brain	Skull	Mandible	**Total**
SCULPTOR [12]	0.99	1.36	4.07	0.66	7.08	
IFMatch [15]	**0.26**	**0.09**	**0.33**	**0.09**	**0.77**	
Ours	0.31	0.12	0.48	0.12	1.03	
Nvdiffrast [38]	2.89	2.66	2.98	2.66	2.80	
Nvdiffrast [38] (w/ nodes)	6.11	7.07	9.01	5.64	6.96	
Ours (w/o CA, dim=512)	0.30	0.10	0.38	0.10	0.88	
Ours (w/o CA, dim=256)	**0.26**	**0.09**	**0.33**	**0.09**	**0.77**	

Furthermore, we conduct a detailed comparison of errors in specific anatomical regions, including the eyes, nose, mouth, and various parts of the brain, skull, and mandible. To quantify the deformation error corresponding to each anatomical region, we select landmarks within these regions and calculate their respective errors. The experiment results are presented in Tables III, IV, V, and VI.

Beyond quantitative assessments, we also perform qualitative experiments to visually evaluate the anatomical fidelity of the models. Fig. 7 and 8 provide a comprehensive evaluation of the deformation accuracy of baselines and our DDF-ISM on anatomical meshes. The PTPD values in Fig. 7 reveal that our method consistently achieves lower distances, indicating a closer match to the target mesh across all anatomical parts. This suggests a higher precision in modeling complex anatomical structures. The visual comparison in Fig. 8 corroborates these findings, showcasing the superior alignment and structural fidelity of our method's deformation results. Notably, our method outperforms the baselines, particularly in regions critical for biomechanical relevance, such as the eyes, nose, and mouth. Additionally, differentiable rendering-based methods yield unsmooth results (without control nodes) or minimal deformation (with control nodes), failing to match target shapes effectively compared to direct 3D geometry methods (*e.g.*, the baseline methods and ours).

The lower PTPD values and enhanced visual accuracy in our method can be attributed to the effective cross-attention mechanism and the directed distance field approach, which facilitates precise landmark alignment and internal structure modeling. In summary, our experimental results underscore the efficacy of our method in delivering high-fidelity deformations, essential for reliable biomechanical analysis.

C. Computational Efficiency

We report the runtime comparisons in Table VII, highlighting the efficiency of different methods in processing anatomical meshes. SCULPTOR shows the longest total runtime at 7.08 minutes, indicating it may require more computational resources or time for processing. Conversely, IFMatch has the shortest total runtime of 0.59 minutes, suggesting it is highly optimized for speed, which is beneficial for time-sensitive applications. The total runtime of our method is much lower than that of SCULPTOR while is closer to that of IFMatch, potentially balancing computational efficiency with the quality of deformation results.

Compared to SCULPTOR, our method processes N vertices directly through the network layers to compute deformations, ensuring computational efficiency. In contrast, SCULPTOR samples K control nodes from N vertices based on predefined intervals σ, followed by vertex deformation optimization. In Table VIII, SCULPTOR evaluates intervals of decreasing size (A: $\sigma_{surf} = 512, \sigma_{brn} = 16, \sigma_{skl} = 128, \sigma_{mdb} = 16$; B: $\sigma_{surf} = 256, \sigma_{brn} = 8, \sigma_{skl} = 64, \sigma_{mdb} = 8$; C: $\sigma_{surf} = 128, \sigma_{brn} = 4, \sigma_{skl} = 32, \sigma_{mdb} = 4$; D: $\sigma_{surf} = 64, \sigma_{brn} = 2, \sigma_{skl} = 32, \sigma_{mdb} = 2$). Smaller intervals increase runtime but improve performance, especially when sampling iteratively from large to small intervals (A to D) in a coarse-to-fine manner, compared to single-interval sampling. In contrast, our method eliminates iterative sampling, offering a more efficient and effective solution for deformation.

Compared to IFMatch, our design introduces two key improvements. First, we integrate a cross-attention (CA) module to capture landmark translation features, requiring five hidden layers in the deformation field, compared to IFMatch's three layers without CA. Second, we increase the hidden units in both the deformation field and latent codes from 256 to 512, enhancing representational capacity. As shown in Table VII, our method without CA and with 256 hidden units matches IFMatch's runtime. However, the full model, incorporating CA and larger hidden dimensions, outperforms IFMatch and the simplified variants of our method (Table II), demonstrating the benefits of these design choices.

D. Ablation Studies

We conducted ablation studies to understand the contribution of different components in DDF-ISM on both original and noisy data. Noise is introduced by randomly adding small outlier structures to the original mesh, simulating a low-quality scan. We train on the noisy data and test on the original data.

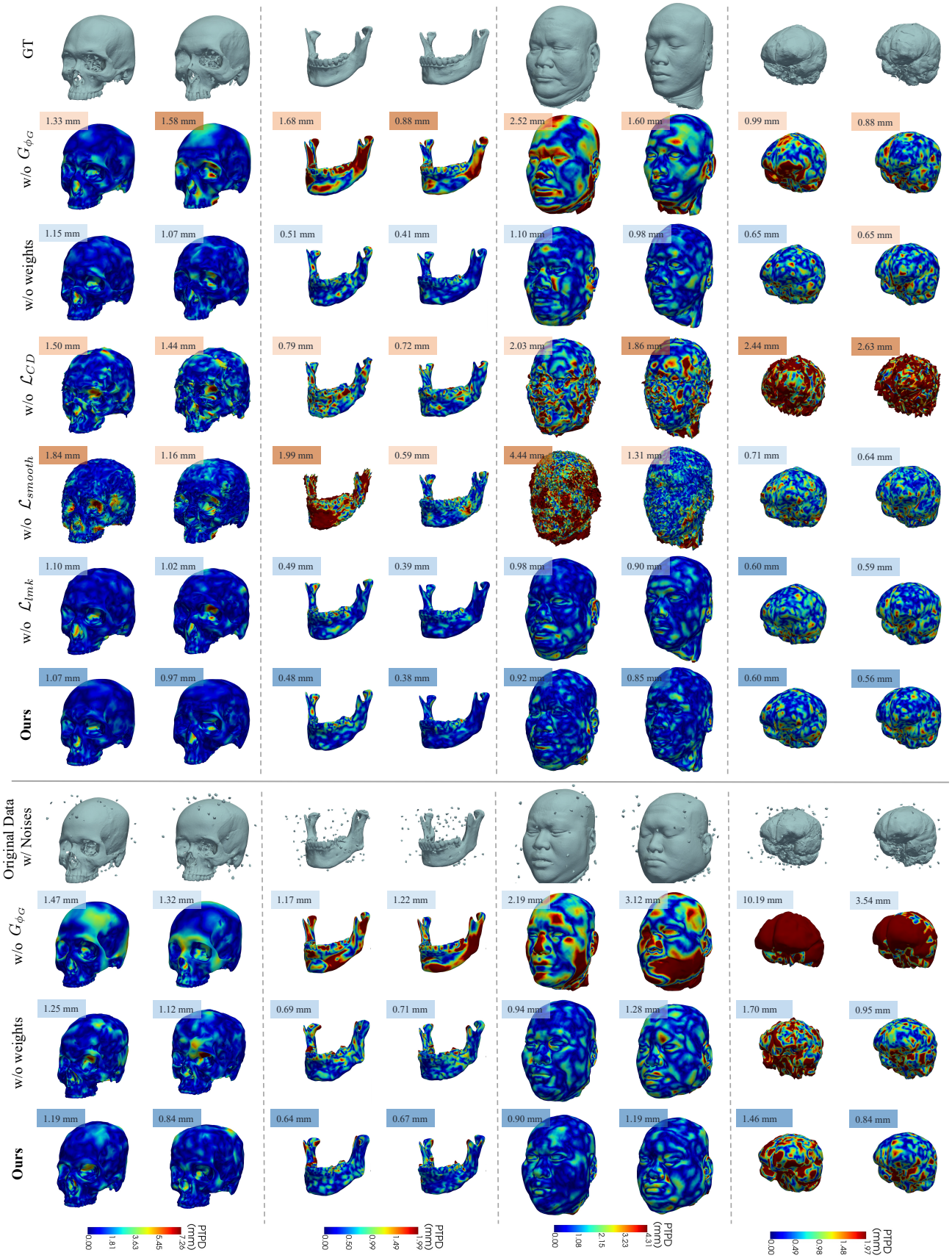

Fig. 9: Deformation result comparisons for ablation study. The values represent the magnitude of deformation, with lower values indicating closer alignment to the target mesh.

TABLE VIII: **PTPD** (mm) ↓, **RMSE**$_{lmk}$ (mm) ↓ (mean ± std. dev.), and **Time** (mins) ↓ for SCULPTOR.

	Head Surface			Brain			Skull			Mandible			Avg.		Sum
	PTPD	**RMSE**$_{lmk}$	**Time**	**PTPD**	**RMSE**$_{lmk}$	**Time**	**PTPD**	**RMSE**$_{lmk}$	**Time**	**PTPD**	**RMSE**$_{lmk}$	**Time**	**PTPD**	**RMSE**$_{lmk}$	**Time**
SCULPTOR	**1.41**	**6.46 ± 2.78**	0.99	**0.61**	**6.73 ± 3.68**	1.36	**1.50**	**6.21 ± 3.63**	4.07	**0.42**	**2.86 ± 1.72**	0.66	**0.99**	**5.57 ± 2.95**	7.08
SCULPTOR-A	1.80	7.73 ± 5.26	**0.08**	0.72	9.44 ± 6.88	**0.10**	1.55	6.67 ± 4.07	**0.35**	0.53	3.35 ± 1.90	**0.09**	1.15	6.80 ± 4.53	**0.62**
SCULPTOR-B	1.68	7.33 ± 5.16	0.14	0.72	9.36 ± 7.01	0.18	1.56	6.72 ± 4.07	0.59	0.50	3.23 ± 1.83	0.11	1.12	6.66 ± 4.52	1.02
SCULPTOR-C	1.70	7.43 ± 5.11	0.24	0.84	11.05 ± 9.67	0.38	1.58	6.80 ± 4.16	1.05	0.49	3.25 ± 1.90	0.17	1.15	7.13 ± 5.21	1.84
SCULPTOR-D	1.68	7.43 ± 5.16	0.53	1.15	12.99 ± 11.45	0.70	1.62	7.01 ± 4.27	2.08	0.52	3.33 ± 1.98	0.29	1.24	7.69 ± 5.72	3.60

TABLE IX: **PTPD** (mm) ↓ and **RMSE**$_{lmk}$ (mm) ↓ (mean ± std. dev.) for ablation studies on four anatomical parts.

	Head Surface		Brain		Skull		Mandible		Avg.	
	PTPD	**RMSE**$_{lmk}$	**PTPD**	**RMSE**$_{lmk}$	**PTPD**	**RMSE**$_{lmk}$	**PTPD**	**RMSE**$_{lmk}$	**PTPD**	**RMSE**$_{lmk}$
Original Data										
w/o G_{ϕ_G}	2.01	6.82 ± 4.40	1.18	7.78 ± 4.95	1.39	4.56 ± 3.17	1.56	4.75 ± 2.48	1.54	5.98 ± 3.75
w/o weights	1.02	5.46 ± 3.91	0.70	6.56 ± 3.71	1.10	4.78 ± 3.33	0.51	2.97 ± 1.63	0.83	4.61 ± 2.97
w/o $\mathcal{L}_{CD}$	1.77	7.27 ± 3.75	2.44	9.06 ± 4.41	1.40	5.49 ± 2.77	0.86	3.88 ± 2.10	1.62	6.43 ± 3.26
w/o $\mathcal{L}_{smooth}$	2.15	6.61 ± 5.71	0.77	3.95 ± 2.10	1.63	5.74 ± 4.05	1.37	4.41 ± 3.11	1.48	5.18 ± 3.74
w/o $\mathcal{L}_{lmk}$	**0.93**	7.67 ± 5.83	**0.61**	21.98 ± 10.46	1.08	6.42 ± 4.71	**0.45**	4.48 ± 2.68	**0.77**	10.14 ± 5.97
Ours	0.95	**4.33 ± 2.34**	0.66	**3.78 ± 2.15**	**1.06**	**3.92 ± 2.48**	0.46	**2.39 ± 1.31**	0.78	**3.61 ± 2.07**
Original Data with Noises										
w/o G_{ϕ_G}	2.64	10.24 ± 6.67	7.06	24.95 ± 11.90	1.51	7.09 ± 4.50	1.45	5.62 ± 2.74	3.17	11.98 ± 6.45
w/o weights	1.12	8.07 ± 5.93	**1.21**	6.77 ± 3.73	1.14	6.12 ± 4.25	0.66	4.28 ± 2.30	1.03	6.31 ± 4.05
Ours	**1.06**	**7.63 ± 5.74**	**1.21**	**4.91 ± 2.70**	**1.13**	**6.04 ± 4.22**	**0.63**	**4.12 ± 2.27**	**1.01**	**5.68 ± 3.73**

Fig. 9 presents deformation magnitudes in millimeters, with lower values indicating better alignment with the target. The variation in values across the board underscores the impact of different factors on the deformation process. Table IX complements this visual assessment with quantitative data, reporting both PTPD and RMSE$_{lmk}$ for four anatomical parts under different conditions.

For our DDF-ISM without $\mathcal{L}_{G_{\phi_G}}$ or without $\mathcal{L}_{CD}$, the increase in PTPD and RMSE$_{lmk}$ values suggests that these components are necessary for enhancing the accuracy of the deformations. The notable increase in error metrics for DDF-ISM without $\mathcal{L}_{smooth}$ or without $\mathcal{L}_{lmk}$ highlights the importance of landmarks in guiding the deformation process to achieve higher fidelity. A comparison of our DDF-ISM without $\mathcal{L}_{G_{\phi_G}}$ and weights reveals that excluding either the DDF or its probabilistic component markedly decreases shape-matching accuracy, especially for noisy scans. Therefore, each part contributes to the overall accuracy, and their combined use yields the best results, as shown by the lowest metric values.

VIII. Conclusion

In conclusion, the proposed DDF-ISM is a novel approach for modeling the internal structures of the human head with high anatomical accuracy and computational efficiency. Our DDF-ISM leverages probabilistic directed distance fields to overcome the limitations of traditional mesh and point cloud-based models, providing superior performance in both quality and speed. The results demonstrate its potential for enhancing applications in biomechanical simulations, medical imaging, and surgical planning. Future work will explore adapting DDF-ISM to model other anatomical regions and integrating it into wearable product design, broadening its impact in various areas including clinical practice and application research.

Acknowledgments

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. GRF/PolyU 15606321).

References

[1] B. Jiang, J. Zhang, J. Cai, and J. Zheng, "Disentangled human body embedding based on deep hierarchical neural network," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 8, pp. 2560–2575, 2020.

[2] L. Bao, X. Lin, Y. Chen, H. Zhang, S. Wang, X. Zhe, D. Kang, H. Huang, X. Jiang, J. Wang *et al.*, "High-fidelity 3d digital human head creation from rgb-d selfies," *ACM Transactions on Graphics*, vol. 41, no. 1, pp. 1–21, 2021.

[3] P.-W. Grassal, M. Prinzler, T. Leistner, C. Rother, M. Nießner, and J. Thies, "Neural head avatars from monocular rgb videos," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 18 653–18 664.

[4] T. Wu, A. Hung, and K. Mithraratne, "Generating facial expressions using an anatomically accurate biomechanical model," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 11, pp. 1519–1529, 2014.

[5] D. Du, X. Han, H. Fu, F. Wu, Y. Yu, S. Cui, and L. Liu, "Sanihead: Sketching animal-like 3d character heads using a view-surface collaborative mesh generative network," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 6, pp. 2415–2429, 2020.

[6] M. Seiler, J. Spillmann, and M. Harders, "Data-driven simulation of detailed surface deformations for surgery

training simulators," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 10, pp. 1379–1391, 2014.

[7] A. Wei, J. Wang, J. Liu, M. L. Jones, and J. Hu, "A parametric head geometry model accounting for variation among adolescent and young adult populations," *Computer Methods and Programs in Biomedicine*, p. 106805, 2022.

[8] S. Yang, B. Li, M. Liu, Y.-K. Lai, L. Kobbelt, and S.-M. Hu, "Heterofusion: Dense scene reconstruction integrating multi-sensors," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 11, pp. 3217–3230, 2019.

[9] L. Xu, W. Cheng, K. Guo, L. Han, Y. Liu, and L. Fang, "Flyfusion: Realtime dynamic scene reconstruction using a flying depth camera," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 68–82, 2019.

[10] S.-S. Huang, H. Chen, J. Huang, H. Fu, and S.-M. Hu, "Real-time globally consistent 3d reconstruction with semantic priors," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 4, pp. 1977–1991, 2021.

[11] F. Danckaers, D. Lacko, S. Verwulgen, G. De Bruyne, T. Huysmans, and J. Sijbers, "A combined statistical shape model of the scalp and skull of the human head," *Advances in Human Factors in Simulation and Modeling: Proceedings of the AHFE 2017 International Conference on Human Factors in Simulation and Modeling*, pp. 538–548, 2018.

[12] Z. Qiu, Y. Li, D. He, Q. Zhang, L. Zhang, Y. Zhang, J. Wang, L. Xu, X. Wang, Y. Zhang *et al.*, "Sculptor: Skeleton-consistent face creation using a learned parametric generator," *ACM Transactions on Graphics*, vol. 41, no. 6, pp. 1–17, 2022.

[13] Y. Hu, E.-J. Rijkhorst, R. Manber, D. Hawkes, and D. Barratt, "Deformable vessel-based registration using landmark-guided coherent point drift," in *Medical Imaging and Augmented Reality: 5th International Workshop*, 2010, pp. 60–69.

[14] Y. Li and T. Harada, "Non-rigid point cloud registration with neural deformation pyramid," *Advances in Neural Information Processing Systems*, vol. 35, pp. 27 757–27 768, 2022.

[15] R. Sundararaman, G. Pai, and M. Ovsjanikov, "Implicit field supervision for robust non-rigid shape matching," in *European Conference on Computer Vision*, 2022, pp. 344–362.

[16] J. J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove, "Deepsdf: Learning continuous signed distance functions for shape representation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 165–174.

[17] C. Zhang, Y. Di, R. Zhang, G. Zhai, F. Manhardt, F. Tombari, and X. Ji, "Ddf-ho: Hand-held object reconstruction via conditional directed distance field," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[18] T. Aumentado-Armstrong, S. Tsogkas, S. Dickinson, and A. D. Jepson, "Representing 3d shapes with probabilistic directed distance fields," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2022, pp. 19 343–19 354.

[19] T. Li, T. Bolkart, M. J. Black, H. Li, and J. Romero, "Learning a model of facial shape and expression from 4d scans," *ACM Transactions on Graphics*, p. 1–17, 2017.

[20] A. Ranjan, T. Bolkart, S. Sanyal, and M. J. Black, "Generating 3d faces using convolutional mesh autoencoders," in *European Conference on Computer Vision*, 2018, pp. 725–741.

[21] S. Ploumpis, E. Ververas, E. O'Sullivan, S. Moschoglou, H. Wang, N. Pears, W. A. Smith, B. Gecer, and S. Zafeiriou, "Towards a complete 3d morphable model of the human head," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 11, pp. 4142–4160, 2020.

[22] H. Zhu, H. Yang, L. Guo, Y. Zhang, Y. Wang, M. Huang, M. Wu, Q. Shen, R. Yang, and X. Cao, "Facescape: 3d facial dataset and benchmark for single-view 3d face reconstruction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[23] A. Kustár, L. Forró, I. Kalina, F. Fazekas, S. Honti, S. Makra, and M. Friess, "Face-r—a 3d database of 400 living individuals' full head ct-and face scans and preliminary gmm analysis for craniofacial reconstruction," *Journal of Forensic Sciences*, vol. 58, no. 6, pp. 1420–1428, 2013.

[24] B.-K. D. Park, B. D. Corner, J. A. Hudson, J. Whitestone, C. R. Mullenger, and M. P. Reed, "A three-dimensional parametric adult head model with representation of scalp shape variability under hair," *Applied Ergonomics*, p. 103239, 2021.

[25] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3d faces," in *SIGGRAPH*, 1999, pp. 187–194.

[26] A. T. Tran, T. Hassner, I. Masi, and G. Medioni, "Regressing robust and discriminative 3d morphable models with a very deep neural network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1493–1502.

[27] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[28] S. Haykin, *Neural networks: A comprehensive foundation*. Prentice Hall PTR, 1998.

[29] G. Trappolini, L. Cosmo, L. Moschella, R. Marin, S. Melzi, and E. Rodolà, "Shape registration in the time of transformers," *Advances in Neural Information Processing Systems*, vol. 34, pp. 5731–5744, 2021.

[30] E. Bermejo, K. Taniguchi, Y. Ogawa, R. Martos, A. Valsecchi, P. Mesejo, O. Ibáñez, and K. Imaizumi, "Automatic landmark annotation in 3d surface scans of skulls: Methodological proposal and reliability study," *Computer Methods and Programs in Biomedicine*, vol. 210, p. 106380, 2021.

[31] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 698–700, 1987.
[32] D. Ha, A. M. Dai, and Q. V. Le, "Hypernetworks," in *International Conference on Learning Representations*, 2017.
[33] A. Vaswani, N. M. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017.
[34] H. Fan, H. Su, and L. J. Guibas, "A point set generation network for 3d object reconstruction from a single image," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2463–2471, 2016.
[35] P. Shah, T. Yarlagada, A. Luximon, D. An, and Y. Luximon, "An evaluation of material customized modular protective helmets," *Creativity, Innovation and Entrepreneurship*, vol. 31, p. 48, 2022.
[36] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," *Advances in neural information processing systems*, vol. 33, pp. 7462–7473, 2020.
[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
[38] S. Laine, J. Hellsten, T. Karras, Y. Seol, J. Lehtinen, and T. Aila, "Modular primitives for high-performance differentiable rendering," *ACM Transactions on Graphics*, vol. 39, no. 6, 2020.
[39] K.-L. Low, "Linear least-squares optimization for point-to-plane icp surface registration," *Chapel Hill, University of North Carolina*, vol. 4, no. 10, pp. 1–3, 2004.
[40] Z. Liu, Y. Luximon, W. L. Ng, E. Chung, and J. Zhang, "Anatomical landmark-guided deformation methods for cranial modeling," *Creativity, Innovation and Entrepreneurship*, p. 46, 2023.

Zhuoman Liu is currently a Ph.D. candidate in School of Design at The Hong Kong Polytechnic University. She received the BEng degree in Software Engineering from South China University of Technology. She has published top-tier research articles on NeruIPS, CVPR, IEEE Transactions on Multimedia, *etc*.. Her current research interests include 3D reconstruction, neural rendering, parametric modeling, and generative models.

Prof. Yan Luximon is a full professor in School of Design at The Hong Kong Polytechnic University. She also serves as Chair of School Research Committee, Lab Leader for Asian Ergonomics Design Lab and Deputy Discipline Leader for BA Product Design. She is also President of Hong Kong Ergonomics Society (HKES). She has published over 100 peer-reviewed journal papers, book chapters, patents and international conference papers. Her research interests include computer graphics, 3D digital human modeling and CAD, AI in design and visualization, deep learning, 3D head and face reconstruction, ergonomic design, head and face related products, human computer interaction, virtual reality, and cultural difference.

Wei Lin Ng joined as an academic staff in the Department of Biomedical Imaging, University Malaya in 2018. She obtained her medical degree (MBBS) and master of radiology (MRad) from University of Malaya. Previously she has served in Hospital Taiping, Perak as a radiologist from 2016-2017. She finished her fellowship training in interventional radiology in Siriraj Hospital, Mahidol University, Bangkok Thailand (July 2021 - June 2022). Her fellowship training mainly focused on hepatocellular carcinoma treatment and also included ablation techniques including RFA, MWA, cryoablation, and IRE. Currently, she is the head of unit of Interventional Radiology in University Malaya Medical Center. Her past research includes shear wave ultrasound in breast lesions. She currently has a Fundamental Research Grant Scheme (FRGS) on the study of "Breomics (breast Radiomics): Deciphering Radiomics Signatures Of Breast Cancer".

Eric Chung is a diagnostic and interventional radiologist. He holds a Medical Degree from Kursk State Medical University and has furthered his education with a Master's in Radiology from University Malaya, a Bachelor's in Computer Science with a specialty in Machine Learning and Artificial Intelligence from the University of London, and is currently pursuing an MBA at Sunway University. Dr. Chung has training in pediatric liver transplant imaging and intervention, interventional radiology, and is currently completing his fellowship in Interventional Radiology at Queen Mary Hospital in Hong Kong. In addition to his clinical and administrative roles, Dr. Chung is actively involved in various professional societies and committees. He also contributes to quality improvement, medical records management, and research committees at University Malaya Medical Centre.