Personalized Fitting of Respiratory Mask Using Deep Learning and 3D Facial Modeling

Eya MLIKA [1], Bahe HACHEM [2], Yamen ALHABASH [2], Loïc DEGUELDRE [2], Luc DUONG [1]
[1] École de Technologie Supérieure, Montréal QC, Canada;
[2] Numalogics, Montréal QC, Canada

https://doi.org/10.15221/24.23

Abstract

Respiratory masks are highly used in healthcare and industrial environments to protect individuals from airborne contaminants and infectious pathogens. However, conventional respiratory masks are often one-sized, therefore it becomes challenging to fit diverse facial morphologies. If loosely fit, the mask may not adequately safeguard or be over-tight to achieve protection, which could cause discomfort or even generate pressure wounds over extended wear.

This project aims to propose an artificial intelligence approach for the design and customization of respiratory masks, emphasizing customized products to better fit into the variety of human face morphology. The proposed approach begins with the creation of facial geometries using 3D facial data obtained by using the ARKit framework. ARKit allows to acquire a structured and complete mesh of the subject's face despites incomplete and noisy data. Each facial scan captures 5,023 3D points, providing a detailed map of individual facial features. The resulting dataset, including 60 different facial scans, forms the basis of our machine learning algorithm. This algorithm is designed to improve the customization and fit of respiratory masks, enhancing wearer safety and comfort. From this facial modelling, a deep learning model designed to predict the deformations of a mask when fitted to the face was deployed. The model enables the identification of potential areas of pressure and mask misfits, predicting problems before they become critical. A predictive model was further introduced to simulate the interaction between the facial structure and the mask as closely as possible. Combining scanning technologies and predictive modelling will alleviate the detection of gaps and pressure points, enabling preventive measures to be taken to rectify these defects. Due to an in-depth understanding of these interactions, the newly developed model proposes modifications to the mask design to better match the unique contours of each face, thus improving the mask's seal and comfort.
This research might contribute to improve fitting and address important health protection issues and accelerate safety regulation compliance, thereby lessening health risks related to the long-term use of poorly fitted masks. Adequate fitting might have an important impact on the design of personalized protective equipment.

Keywords: Respiratory masks, Customization, Artificial Intelligence, Face Mask, 3D Face Reconstruction, Mesh generation, non-rigid deformation, ARKit, Deep Learning, Personal Protective Equipment, Safety, Health Care Safety, Industrial Safety, Facial structure interaction, thin plate spline, TPSNet

1. Introduction

Respiratory masks play a crucial role in protecting against hazardous airborne contaminants in both healthcare and industrial environments. These devices are essential for preventing infections caused by pathogens in medical environments, and for protecting workers from chemical pollutants posing various health risks in industrial settings **[1,2]**. Although occupational health and safety regulations require personal protective equipment, including respiratory masks, to follow safety standards, the conventional “one size fits all” design of these masks does not consider the different facial morphologies of users, compromising their effectiveness and comfort, particularly during prolonged wear **[3,4].**
Faced with this problem, our study proposes an innovative approach using artificial intelligence to personalize respiratory masks. By combining 3D facial scanning, finite element analysis and deep learning, we develop customized masks that fit perfectly to the unique contours of each face. Our research aims to set up a method capable of predicting facial deformation under the mask and providing these results in real time when ordering a mask via a mobile application.
This article is structured as follows: the first section presents a review of related work. The second section details our project's method, including the generation and processing of facial mesh data and the development of our deep-learning model. The third section presents the results obtained, highlighting the accuracy of our dataset and the effectiveness of our prediction model. Finally, the conclusion summarizes the main contributions of our work and discusses future implications for the design of personalized protective equipment

2. Related work

In our earlier work, our team developed a personalized mask adapted to the morphology of the user's face. This customization was made possible by a mobile application that scanned the face and created the mask geometry. Their study combined 3D scanning with finite element analysis to assess the mask's performance, while confirming the numerical results through experimental measurements of pressure and spacing **[5]**. Machine learning models were applied to predict pressures at the mask-face interface and adjust mesh points according to the degree of clamping. Among these models, the Random Forest Regressor showed superior accuracy compared with the Decision Tree Regressor and Elastic Net models. However, these approaches met challenges, notably the integration of facial morphology into model training and the management of protection gaps due to incomplete data.

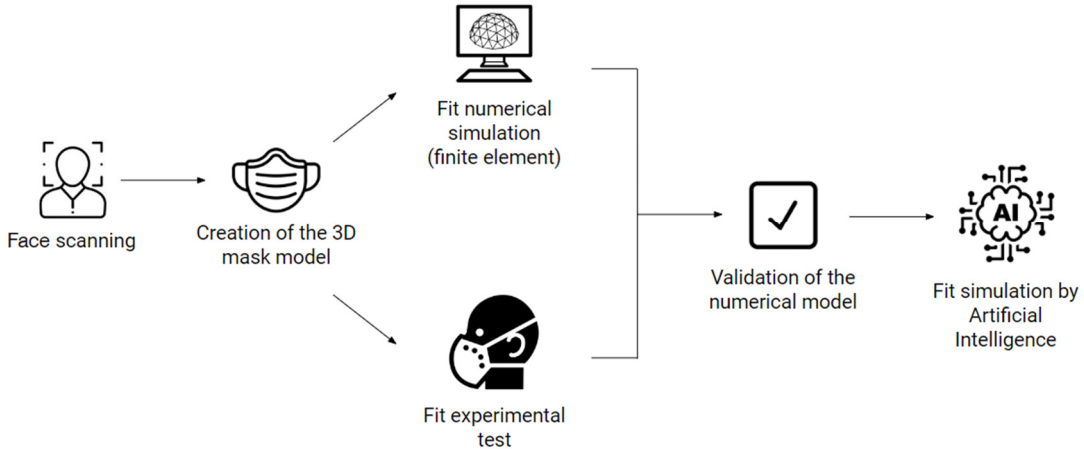


Fig. 1. Protocol for Mask Fit Simulation

The advancement in transforming unstructured and incomplete 3D point clouds into detailed meshes is significantly driven by the development of 3D morphable models (3DMMs) **[6,7,8,9]**. These models allow for the creation of precise and adaptable shapes. By employing adjustable and morphable reference models, researchers can keep consistency and enhance the integration of scanning technologies with advanced mesh generation techniques. 3DMMs facilitate the filling of gaps in scanned data, helping to generate detailed and exact three-dimensional representations.

A main drawback from earlier work is that most studies assure a rigid fit and very little work focus on non-rigid deformation in shape transformation learning. Graph Convolution Networks (GCN) **[10]** and CDPNet have been proposed for learning non-rigid shape transformations **[11]**. GCN processes relational data within a graph structure, ideal for analysing non-rigid deformations. CDPNet predicts post-deformation transformations of 3D objects, maintaining structural integrity under various conditions. P2P-NET **[12]** offers a generalized approach to learning various shape transformation tasks, while its derivative, P2MAT-Net **[13]**, transforms point clouds into sets of medial spheres, although it requires further refinement to understand global shape properties.

3. Methodology

3.1. Creation of Structured and Complete Facial Mesh Databases from ARKit scan

The ARKit framework has significantly advanced the capability to capture 3D scans of human faces. However, despite efforts to minimize noise and address incomplete data, challenges remain in creating structured and complete facial meshes **[14].** To address this issue, we propose a method that integrates the average face from the FLAME database as a reference to enhance the ARKit scan data **[15].** The FLAME database is a statistical 3D model of the human face that captures a wide range of facial shapes, expressions, and poses. It is constructed by learning from a large dataset of 3D scans and is designed to stand for facial geometry using a low-dimensional parameter space. Integrating the average face from the FLAME database as a reference in this method provides a robust and consistent template that can significantly improve the quality of ARKit scan data. This method involves the extraction of facial landmarks, mesh alignment, and transformation techniques to generate a structured and complete facial mesh

3.1.1. Landmarks extraction

The first step involves extracting 100 facial landmarks from the ARKit scan. We start by manually selecting 68 reference landmarks of the face. However, these first 68 landmarks do not accurately capture the full facial morphology. To address this, we employ the Geodesic Ensemble Surface Sampling Algorithm (GESSA) to add more landmarks. GESSA begins by randomly sampling points across the facial surface and then improves their distribution to achieve uniformity. This optimization process maximizes the entropy in the landmark distribution, ensuring a detailed capture of the morphological variability **[16].**

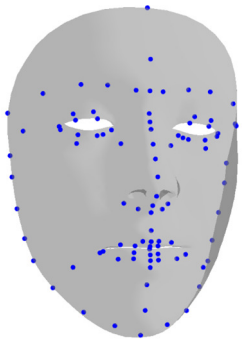


Fig. 2 Landmarks from ARKit scan

3.1.2. Mesh Alignment

Once the landmarks are identified, the next step is aligning the ARKit landmarks with the reference face using the Iterative Closest Point (ICP) algorithm. This process begins with proving initial landmark and reference face vertices correspondences. The ICP algorithm then iteratively refines the alignment by repeatedly matching corresponding points between the two meshes, estimating the best transformation (rotation and translation) to minimize the distance between these points, and applying this transformation to the ARKit mesh. This procedure iterates until the algorithm converges by reaching the best transformation **[17].**

3.1.3. Mesh Generation

Following the alignment, the landmarks on the reference face are selected using the nearest neighbour algorithm. The transformation between the ARKit landmarks and the reference landmarks is computed using the Thin Plate Spline (TPS) algorithm. The TPS algorithm is particularly suited for capturing non rigid deformation, since it works by modeling the transformation needed to bend a thin plate spline from one shape into another, adhering to certain fixed points. It minimizes the energy needed to achieve this bending, ensuring a smooth and continuous transformation **[18].** This computed transformation is then applied to the entire reference face, resulting in a complete and structured mesh that accurately stands for the morphology of the scanned face. This transformation ensures that the final mesh captures the unique facial features of each user, which is particularly crucial for applications such as the customization of respiratory masks. The proposed method yields precise and structured facial meshes that will be used to train a deep learning model to predict the deformation of the face after waring the mask.

3.2. Deep learning model for Respiratory Mask Fitting

TPS-Net was developed to optimize the customization and fitting of respiratory masks by accurately predicting facial structure deformations upon mask application. This network is specifically designed to manage non-rigid transformations between 3D facial meshes, both pre- and post-mask application. By learning smooth and continuous geometric transformations, TPS-Net utilizes a neural network to precisely map facial points to their new positions after mask wear. The model is trained using a combination of Chamfer distance, which evaluates alignment accuracy, and a regularization term that ensures smooth transformations. This approach is particularly effective for accurately aligning source and target points, essential for fitting masks to diverse facial morphologies. To validate TPS-Net's performance, various evaluation metrics are employed, including Chamfer Distance, Mean Squared Error (MSE), and Mean Absolute Error (MAE). These metrics ensure that the model achieves precise and smooth transformations, with lower Chamfer distances showing better alignment and MSE, MAE providing additional measures of alignment accuracy.

4. Results and discussions

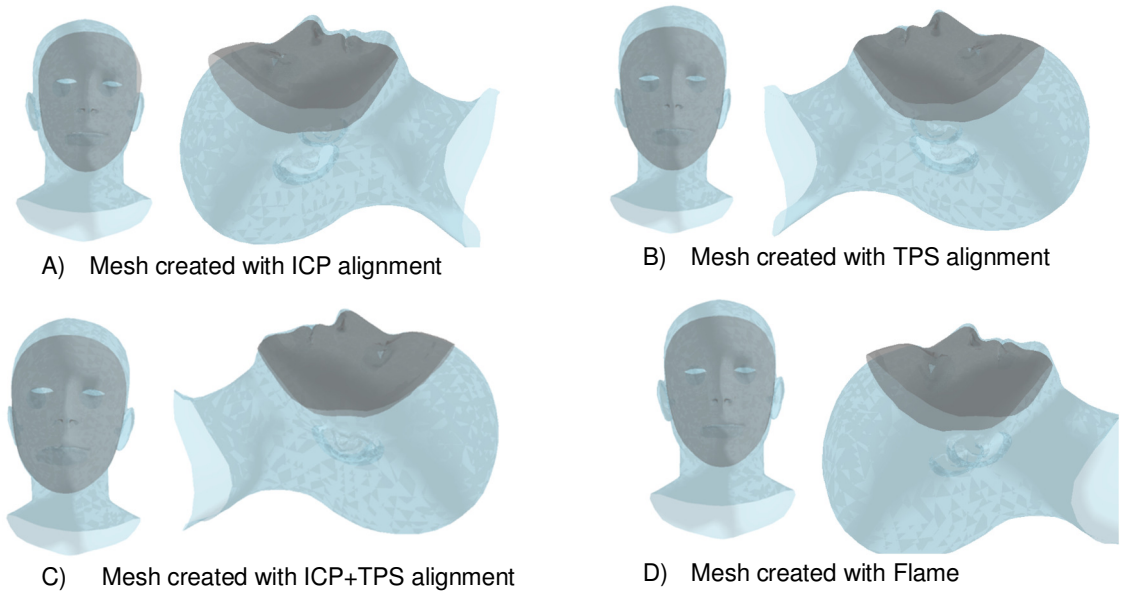
4.1. Creation of Structured and Complete Facial Mesh Databases from ARKit scan

The generation of the facial mesh from ARKit scan landmarks with the reference face from the FLAME database was performed using various algorithms to figure out the most effective approach for this task. In this study, the ARKit face scan, captured with high precision, served as the reference comparator against which the performance of each method was evaluated. The performance of each method was evaluated using metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Standard Deviation of Errors (STD), and Hausdorff Distance of Errors (HDE) -a measure of the maximum distance between points in one set and the nearest points in another set, reflecting the maximum possible error. The results, summarized in the table below, prove the effectiveness of each method:

Table 1: Comparison of the four facial mesh generation algorithms.

Method	RMSE (mm)	MAE (mm)	STD (mm)	HDE (mm)
ICP	0.037	0.022	0.030	0.183
TPS	0.037	0.021	0.030	0.184
ICP+TPS	0.033	0.019	0.027	0.165
Flame	1.610	1.510	0.565	0.181

Figure 5 provides visual comparisons of the generation results for ICP, TPS, and the combined ICP+TPS approach. The images clearly illustrate that TPS and ICP+TPS methods provide more exact generations compared to ICP and Flame. The TPS method shows superior performance in aligning the landmarks closely with minimal gaps. With the ICP, TPS and Flame, we notice the significant gaps between the points in certain regions, showing poor representation of morphological variations. But with ICP+TPS, the points are more closely aligned, showing a better fit and more exact representation of the facial morphology. The TPS method captures non-rigid deformations effectively, which is crucial for accurately representing facial morphology. In contrast, the ICP method, being a rigid transformation, does not adequately account for morphological variations, leading to larger gaps between point sets in certain regions. The FLAME model, however, produced poor results due to suboptimal parameter optimization and its limited use of only 100 landmarks, which is insufficient to accurately capture the detailed features and variations in facial morphology. The high RMSE and MAE values observed for the FLAME model reflect this limitation.



A) Mesh created with ICP alignment

B) Mesh created with TPS alignment

C) Mesh created with ICP+TPS alignment

D) Mesh created with Flame

Fig. 5 Comparative Analysis of Mesh Generation technique

4.2. Thin Plate Spline Networks TPS-NET for Respiratory Mask Fitting

In our study, the dataset was distributed with 80% used for training, 10% for validation, and the remaining 10% dedicated to testing. This distribution aligns with established practices commonly referenced in literature. To provide a comprehensive evaluation of our model's performance, we conducted a comparative analysis with Coherence Point Drift Network (CPD-Net), a well-recognized point-to-point network.

Table 2 presents the averaged results for all models tested. We report the root mean square error (RSME), mean absolute error (MAE), and Chamfer distance—a metric that quantifies the similarity between two-point clouds by averaging the nearest point distances between sets. The TPS-NET model achieved the lowest error across all metrics, showing its superior performance in predicting facial deformations. The Point-to-point network also performed well but showed higher errors compared to TPS-NET. The CPDNET model had the highest errors, suggesting less effectiveness for this specific application. Adding a learning rate decay strategy provided more flexibility to the networks by reducing the learning rate by 50% after every 50 epochs, enhancing the convergence characteristics in the later stages of training. The rigorous training and testing regimen ensured robust evaluation of each model. Comparing the results, TPS-NET significantly outperformed the other models, achieving an RSME of 0.0005 mm, an MAE of 0.0002 mm, and a Chamfer distance of 0.0225 mm. These values show the model's ability to accurately predict facial deformations, crucial for applications where precise geometric adjustments are necessary.

Table 2: Performance Comparison of Facial Deformation Prediction Models in millimeters (mm)

Model	RSME (mm)	MAE (mm)	Chamfer distance (mm)
TPS-NET	0.0005	0.0002	0.022
Point-to-point network	0.0017	0.0003	0.033
CPDNET	0.0706	0.0582	0.046

5. Conclusion and Future Work

The goal of this study was to propose and to evaluate a new fitting algorithm, designed for respiratory masks. This has important implications for health and safety in both medical and industrial settings, potentially reducing the risks associated with prolonged mask usage and ensuring better compliance with safety regulations. Future work will focus on further refining the models and expanding the dataset to include a wider variety of facial morphologies, thereby enhancing the generalizability and robustness of the proposed solutions.

References

[1] L.K.P. Suen, et al., "Comparing mask fit and usability of traditional and nanofibre N95 filtering facepiece respirators before and after nursing procedures," Journal of Hospital Infection, vol. 104, no. 3, pp. 336–343, 2020, https://doi.org/10.1016/j.jhin.2019.09.014.

[2] J. R. Lewis, S. S. Michael, and T. B. Mary, "N95 Respirators vs Medical Masks for Preventing Influenza Among Health Care Personnel," JAMA Network, vol. 322, no. 9, pp. 824–833, 2019, https://doi.org/10.1001/jama.2019.11645.

[3] L. Shiya, W. Usman, B. Mohanad, and W. Louis Zihao, "A scalable mass customization design process for 3D-printed respirator mask to combat COVID-19," Rapid Prototyping Journal, vol. 27, no. 7, pp. 1302–1317, 2021, https://doi.org/10.1108/RPJ-10-2020-0231.

[4] C. Alison, G. Linsey, Y. Minji, and D. William, "Design considerations for protective mask development: A remote mask usability evaluation," PMC PubMed Central, https://doi.org/10.1016/j.apergo.2022.103751.

[5] T. Hugo, et al., "Personalized Fitting of Respiratory Mask Using 3D Numerical Simulation and Finite Element Analysis," in 3DBODY.TECH 2022, 2022, https://doi.org/10.15221/22.48.

[6] L. Alexandros, M. Stylianos, G. Baris, and P. Stylianos, "AvatarMe: Realistically Renderable 3D Facial Reconstruction 'in-the-wild'," https://doi.org/10.48550/arXiv.2003.13845.

[7] M. Stylianos, P. Stylianos, Z. Stefanos, N. Mihalis, and P. Athanasios, "3DFaceGAN: Adversarial Nets for 3D Face Representation, Generation, and Translation," https://doi.org/10.48550/arXiv.1905.00307.

[8] R. C. Eric, Z. L. Connor, A. C. Matthew, and N. Koki, "Efficient Geometry-aware 3D Generative Adversarial Networks,", https://doi.org/10.48550/arXiv.2112.07945.

[9] T. Li, T. Bolkart, M. J. Black, H. Li, and P. Javier, "Learning a model of facial shape and expression from 4D scans,", https://ps.is.mpg.de/uploads_file/attachment/attachment/400/paper.pdf

[10] W. Zijie and C. Huiyou, "3D Mesh Deformation Using Graph Convolution Network," in Proc. IEEE CCOMS 2019, https://doi.org/10.1109/CCOMS.2019.8821790.

[11] K. Maryam, D. Philippe, L. Hubert, P. Stefan, and C. Farida, "Prediction of postoperative 3D spine shape using Controlled Point Deformation Network (CPDNet)," in Proc. SPIE Medical Imaging 2023, https://doi.org/10.1117/12.2654224.

[12] Y. Kangxue, H. Hui, C.-O. Daniel, and Z. Hao, "P2P-NET: bidirectional point displacement net for shape transform," ACM Transactions on Graphics, vol. 37, no. 4, pp. 1–13, 2018, https://doi.org/10.1145/3197517.3201288.

[13] B. Yang, et al., "P2MAT-NET: Learning medial axis transform from sparse point clouds," Computer Aided Geometric Design, vol. 80, 2020, https://doi.org/10.1016/j.cagd.2020.101874..

[14] J. Borduas, A. C.-H., P. L., S.-P. V., and D. B., "Reliability of Mobile 3D Scanning Technologies for the Customization of Respiratory Face Masks," in Proc. 3DBODY.TECH 2020, 2020, https://doi.org/10.15221/20.34.

[15] Learning a model of facial shape and expression from 4D scans, https://flame.is.tue.mpg.de/, accessed 2024.

[16] T. Dimosthenis, P. Hugo, M. Giovanni, and S. Tim, "Heritability maps of human face morphology through large-scale automated three-dimensional phenotyping," Scientific Reports, 2017, https://doi.org/10.1038/srep45885.

[17] Z. Juyong, Y. Yuxin, and D. Bailin, "Fast and Robust Iterative Closest Point," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 7, 2017, https://doi.org/10.1109/TPAMI.2021.3054619.

[18] L. Jungwoo and Y. Ming-Hsuan, "A direct method for modeling non-rigid motion with thin plate spline," in Proc. IEEE CVPR 2005, https://doi.org/10.1109/CVPR.2005.24.