2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Healthcare Robotics for Light-Based Cosmetic Treatments

Anqing Duan[1], Wanli Liuchen[2], Domingo Gomez[2], Muhammad Muddassir[2], David Navarro-Alarcon[2*]

***Abstract*— Healthcare robotics has been gaining traction as a key area of research focused on enhancing human wellness. This paper explores the use of intelligent robots in the beauty industry, specifically within the context of photorejuvenation-based cosmetic dermatology, aimed at improving facial skin aesthetics. The beauty industry, traditionally labor-intensive, is experiencing a critical shortage of skilled beauticians, highlighting the opportunity for robotic technologies to meet this demand. However, integrating robots into cosmetic procedures presents unique challenges, particularly in tasks requiring high precision, such as laser pulse delivery and thermal dose management. This study addresses these challenges by introducing a deep learning approach for trajectory generation in laser path planning and a model-based control strategy for thermal dose regulation. Our empirical results demonstrate that the presented healthcare robots can deliver effective photorejuvenation treatments, suggesting a promising future for increased automation in cosmetic services.**

I. Introduction

The beauty and personal healthcare industry is a rapidly growing field. The global market size is projected to grow from $1,610.5 billion in 2023 to $2,765.8 billion by 2030, with a steady compound annual growth rate of more than 8% over the forecast period. This expansion is coupled with a notable shortage of qualified aesthetic practitioners [1]. To meet the burgeoning demand for aesthetic services, the deployment of robots holds significant promise. Robots are poised to play a crucial role across a wide array of aesthetic applications. Beyond merely responding to market growth, the adoption of robotic technology in beauty and healthcare offers multiple benefits. These include enhanced standardization of treatments, reduced operational costs, and increased scalability of businesses.

While autonomous robots offer a promising solution, deploying them in real-world aesthetic applications presents complex challenges. Developing a robotic beautician involves addressing several technical and practical issues. First, robotic systems for beauty tasks must process sensory inputs in real time to execute appropriate actions. The perception systems in beauty robots need to handle high-dimensional information, such as visual inputs, efficiently. This efficient processing of complex sensory information is crucial in achieving beauty tasks. Besides, real-world beauty tasks typically involve a series of intricate sub-tasks. Robot motion policies may need to vary spatially, i.e., different areas may require distinct approaches. Developing precise motion policies that are tailored to specific regions is essential for the effective learning and execution of these tasks. Furthermore, the automation of beauty tasks demands meticulous structuring, necessitating the choice between data-driven or model-based strategies for robotic decision-making.

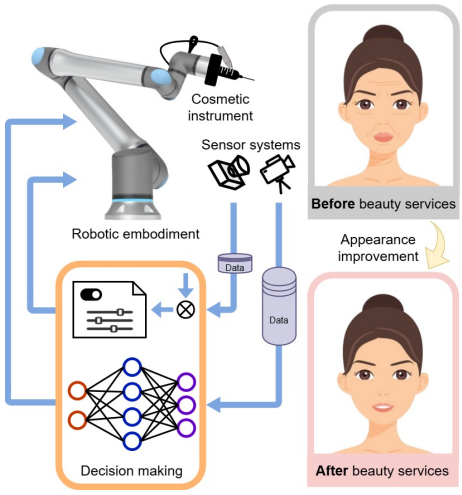

Fig. 1. A robotic system designed for beauty tasks, utilizing model-based approaches for low-dimensional data processing and neural networks for handling high-volume data in the decision-making module.

This paper explores the usage of robots in the beauty industry by addressing the aforementioned challenges with a focus on cosmetic dermatology services, as shown in Fig. 1. Cosmetic dermatology is a major specialized sector in the beauty industry that deals with the appearance of skin conditions [2]. A popular treatment strategy in cosmetic dermatology is *photorejuvenation* [3]. The principle of photorejuvenation-based skin care is to create a controlled wound on the skin by burning with laser pulses. The skin condition is expected to improve after the self-healing procedure. In the procedure of photorejuvenation, a dermatologist carefully manipulates a cosmetic laser to spread the laser pulses over the subject's facial skin.

To advance the development of healthcare robots for cosmetic services, we make the following contributions:

1) We develop a novel deep learning framework, specifically based on *PointNet++*, to efficiently generate cosmetic laser paths from RGBD data;
2) We leverage segmentation techniques for the differentiation of facial regions to facilitate the execution of a primitive motion policy;
3) We address the thermal dose regulation problem by a model-based control strategy, enabling precise control of thermal emissions.

[1]Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, United Arab Emirates.
[2]The Hong Kong Polytechnic University (PolyU), Kowloon, Hong Kong.
*Corresponding author: dnavar@polyu.edu.hk.

The rest of the paper is organized as follows. The relevant literature is reviewed in Sec. II. We present our methodology in Sec. III. Subsequently, the experimental results are reported in Sec.IV. Finally, we conclude the paper in Sec. V.

II. Related Work

Robotic technologies are progressively integrating into the beauty industry, fulfilling various roles. Erdogan et al. [4] developed a novel robotic system that can detect the follicles in pre-op and the placed grafts in post-op. Homma et al. [5] demonstrated a cosmetic robot system to achieve automatic makeup application, utilizing a simplified robotic brush system with motors for directional control and a spring for force regulation. Similarly, an intelligent robotic system is developed to automate the process of eyelash extensions [6]. Compared with the various perspectives on cosmetic applications mentioned above, our work focuses on developing an intelligent robotic system capable of performing skincare tasks. Unlike makeup application or hair-related procedures, skincare requires careful addressing of unique challenges such as laser trace generation and dose regulation.

Another related area lies in the field of service robotics. In the field of service robotics, robots are used for various tasks for the benefit of human wellness, such as dental care [7], hair combing [8], and massage [9]. Compared to general service robots, beauty-focused robots prioritize enhancing personal appearance by providing personalized beauty services tailored to specific individual needs.

Furthermore, the advancement of machine learning technologies has significantly influenced the skincare field [2]. For example, Shen et al. [10] developed a convolutional neural network-based method to automatically detect facial acne vulgaris, which attained a high performance in detecting abnormality in the facial skin. Similarly, Chen et al. [11] proposed a framework based on data width evolution and self-learning to facilitate skin disease recognition. Compared with pure vision-based skin issue diagnosis, our skincare work focuses on the embodiment of the skincare services. As a physical strategy for skin care, a robot platform is developed to spread lotion [12]. While their approach required physical contact with the human back, our work leverages thermal interaction between the subject and the robot. Particularly, as we obtain the subject's face using a depth camera, we leverage PaintNet to generate treatment trajectories, which have the merits of efficiently processing 3D point cloud data [13].

III. Cosmetic Dermatology by Photorejuvenation

This section presents the study of cosmetic dermatology by photorejuvenation, focusing on developing an intelligent robotic system for laser-based facial treatment. To effectively perform this task, we propose a systematic approach consisting of multiple interconnected phases, with facial region segmentation and trajectory planning serving as the core components, as illustrated in Fig. 2.

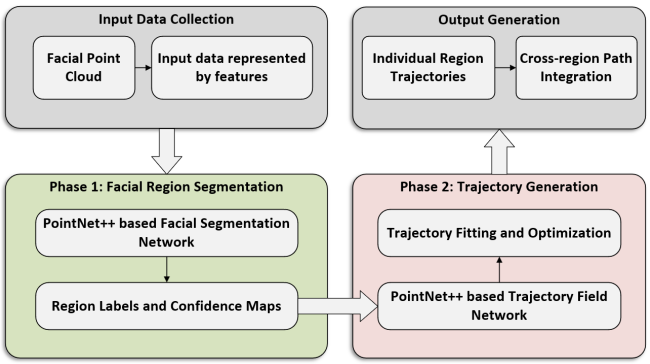

Fig. 2. Core processing pipeline for facial segmentation and trajectory planning.

Our facial segmentation and trajectory planning framework begins with data collection using an RGB-D camera, capturing the face's 3D geometry and surface features. The collected point cloud data includes spatial coordinates (X, Y, Z), surface normals, and geometric features that characterize local surface properties, which are typically challenging to handle [14].

The first processing phase employs a PointNet++ based network to segment the facial point cloud into anatomically meaningful regions, as different facial areas require distinct treatment strategies. The second phase focuses on trajectory generation through a two-stage process: generating a probability field that identifies suitable regions for laser application, then fitting and optimizing trajectories based on this field. This approach ensures that paths are locally smooth and globally optimal for treatment. The output consists of region-specific treatment trajectories tailored to each facial area's characteristics, guiding the robot's end-effector during the procedure.

Beyond this core pipeline, our system incorporates thermal dose regulation that monitors and adjusts treatment parameters to maintain safety while achieving the desired therapeutic effect. This critical aspect will be elaborated in subsequent sections. The following subsections detail each component and its technical implementations.

A. Facial Region Segmentation

In photorejuvenation-based cosmetic dermatology, different facial regions require distinct laser application strategies due to variations in skin thickness, sensitivity, and aesthetic requirements. For example, the forehead area typically requires a different laser path pattern than the cheek area. We first segment the facial point cloud into distinct anatomical regions that align with cosmetic practices to facilitate such region-specific treatment planning.

Our segmentation approach builds upon the hierarchical feature learning capability of PointNet++ [15], which has demonstrated remarkable performance in processing unordered point sets with distance metrics. The key insight of our implementation lies in its ability to progressively capture features at multiple scales through hierarchical feature learning, which naturally aligns with the hierarchical organization

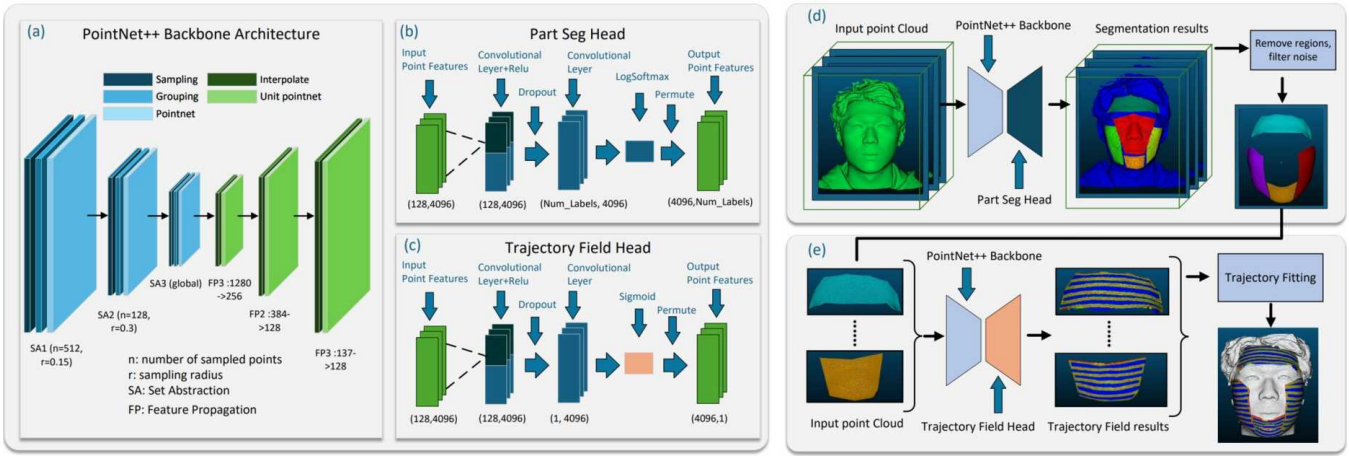

Fig. 3. Overview of our framework architecture. (a) PointNet++ backbone architecture shows the hierarchical processing structure with set abstraction (SA) and feature propagation (FP) layers for point cloud feature learning. (b) Part segmentation head detailing the network structure for facial region segmentation, which processes high-dimensional point features (128,4096) to output region labels. (c) Trajectory field head illustrating the parallel network branch that generates trajectory probability fields from the same point cloud features. (d) The face segmentation pipeline demonstrates the workflow from the input point cloud to the filtered segmentation results. (e) The trajectory generation pipeline shows how segmented regions are processed to create optimized laser treatment paths through probability field computation and trajectory fitting. The complete framework effectively integrates segmentation and path planning into a unified architecture.

of facial features, from local surface details to global facial structures.

The segmentation branch of our framework, as detailed in Fig. 3(a-b), processes the input point cloud through three set abstraction levels (SA1-SA3), followed by feature propagation layers (FP1-FP3) that restore point-level predictions. The part segmentation head processes high-dimensional point features (128,4096) through convolutional layers with ReLU activation and dropout for robust region classification. The effectiveness of our segmentation approach is demonstrated in Fig. 3(d), where the network successfully segments the face into distinct regions (forehead, cheeks, chin, nose, and eye area) with high prediction certainty, particularly in regions with distinctive geometric features.

The segmented regions directly inform the subsequent trajectory generation process by:

- Enabling region-specific path patterns that align with each area's treatment requirements;
- Ensuring clear treatment boundaries between different facial zones;
- Facilitating parallel processing of different regions for efficient path planning.

These segmented regions serve as the basis for our trajectory field generation, which we will discuss in detail in the following section.

B. Laser Trace Generation

Given the segmented facial regions, our next task is to generate appropriate laser traces for each region. The challenge lies in generating paths that ensure uniform coverage while adapting to the geometric characteristics of different facial areas. We propose a two-stage approach to address this challenge: first, generating a probability field highlighting suitable regions for laser application, then fitting smooth trajectories based on the generated field.

1) Trajectory Field Generation: A common approach to trajectory generation is to directly regress trajectory coordinates using deep learning models [16]. However, this strategy faces significant challenges in our context, where the desired outputs are geometrically constrained path segments. Deep neural networks, which inherently produce distributions over their output space, struggle to generate precise trajectories that strictly adhere to geometric constraints, even when guided by carefully designed loss functions. Furthermore, traditional path planning methods that rely on discrete waypoints or binary labels fail to capture the continuous nature of laser paths and are sensitive to point cloud sampling variations [17].

To better align with the probabilistic nature of deep learning and our task requirements, we reformulate the trajectory generation problem as learning a probability field over the facial surface. This probability field indicates regions suitable for laser application, from which geometrically valid trajectories can be extracted. This approach offers several advantages: it leverages the probabilistic output nature of neural networks, handles uncertainties in expert demonstrations, and provides a robust foundation for generating paths that satisfy geometric constraints.

Given an expert-demonstrated trajectory denoted by $\mathcal{T} = \{q_1, ..., q_m\} \subset \mathbb{R}^3$, we create soft labels through a distance-based Gaussian transformation. Specifically, for each point p in the point cloud, its probability label y_p is computed as:

$$y_p = \exp(-d_p^2/2\sigma^2) \tag{1a}$$

$$d_p = \min_{q \in \mathcal{T}} \|p - q\|_2 \tag{1b}$$

where d_p represents the minimum Euclidean distance from point p to the expert trajectory, and σ is a bandwidth parameter that controls the spread of the probability field. This formulation naturally creates a smooth transition region

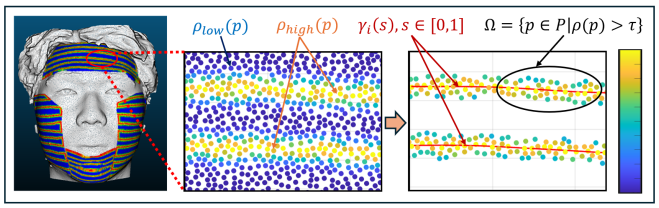

Fig. 4. Visualization of trajectory fitting process. The probability field $\rho(p)$ is computed from point cloud data, with high values (ρ_{high}) indicating suitable regions for laser application. Smooth trajectories $\gamma_i(s)$ are fitted through the high probability regions Ω, naturally following the geometric patterns. Colorbar shows probability values from 0 to 1.

around expert trajectories, transforming discrete demonstrations into a continuous probability field more suitable for deep learning.

For each point, we compute a rich feature set, including spatial coordinates, surface normals, and local geometric descriptors (linearity, planarity, and sphericity) calculated from eigenvalues of the local covariance matrix.

As illustrated in Fig. 3(c) and (e), our trajectory generation approach utilizes a specialized network branch parallel to the segmentation head. The trajectory field head shares the same backbone features but processes them through a different sequence of convolutional layers, ultimately producing a continuous probability field instead of discrete region labels. This design enables the network to learn complex spatial relationships while maintaining geometric consistency with the segmented regions. Fig. 3(e) demonstrates how the generated trajectory field is transformed into practical laser paths through our fitting algorithm, with each facial region receiving customized treatment patterns based on its geometric and anatomical characteristics.

To maintain spatial continuity in the predicted probability field, we incorporate a smoothness constraint in the loss function that penalizes sharp probability variations between neighboring points, balanced with the primary binary cross-entropy loss for probability prediction. The resulting probability field effectively captures both the geometric constraints of the facial surface.

2) Trajectory Fitting: Given the probability field $\rho : \mathbb{R}^3 \to [0,1]$ generated by our deep learning model from the point cloud data $P \in \mathbb{R}^3$, where $\rho_{\text{high}}(p)$ indicates regions suitable for laser application and $\rho_{\text{low}}(p)$ represents unsuitable regions (as visualized in Fig. 4), we aim to fit smooth trajectories that naturally follow these high-probability regions. Let $\gamma_i : [0,1] \to \mathbb{R}^3$ denote the i-th trajectory function, where $i \in \{1, \ldots, N\}$ and N is the total number of trajectories. The trajectory fitting process is formulated as follows:

$$\Omega = \{p \in P \mid \rho(p) > \tau\} \tag{2a}$$

$$C_i = \{p \in \Omega \mid \exists q \in C_i : |p - q|_2 \leq \epsilon\} \tag{2b}$$

$$\min_{\gamma_i} \sum_{p \in C_i} |\gamma_i(s_p) - p|^2 + \lambda \int_0^1 |\gamma_i''(s)|^2 \, ds \tag{2c}$$

where $\tau \in (0,1)$ is the probability threshold that identifies high-probability regions in (2a). Equation (2b) represents the process of density-based spatial clustering of applications with noise, where the distance threshold $\epsilon > 0$ segments the high-probability points into distinct trajectory groups C_i. The smooth trajectory fitting optimization is defined in (2c), where $s_p \in [0,1]$ is the curve parameter corresponding to point p, and $\lambda > 0$ controls the trajectory smoothness.

This approach ensures that the trajectories naturally follow the geometric patterns in the high-probability regions while maintaining sufficient smoothness for robotic execution. The data-driven nature of our method means that path characteristics emerge from the underlying probability distribution rather than being imposed through geometric constraints, resulting in trajectories that naturally adapt to the specific requirements of each facial region.

C. Thermal Dose Regulation

This section considers the regulation of thermal doses, which has wide applications in robotics [18]. Particularly, thermal dose regulation plays an important role in the effectiveness of photorejuvenation treatment, as the treatment performance relies on the self-healing of the skin cells for a better appearance. Both under- and over-emission of laser pulse energy will cause issues. Under-emission of the laser thermal energy will lead to insufficient treatment results, as the skin cells may not be thermally stimulated well. Over-emission of laser thermal energy will cause a more severe critical safety issue, as laser pulses could cause irreversible thermal damage.

Given the important role of thermal dose regulation in the success of photorejuvenation, our goal is to precisely control the thermal doses on the skin area of interest. Furthermore, we opted for a model-based control strategy over a learning-based policy for several reasons. First, control policies derived from demonstration data may pose safety concerns, resulting in laser intensities exceeding permissible limits. Additionally, the dimensionality of thermal sensing is relatively low, and the thermal dynamics on the skin surface are well understood and readily accessible. From a practical perspective, developing a model-based laser controller is more feasible and efficient. Consequently, we have employed a model-based control strategy to address the thermal control requirements associated with laser dosing.

Formally, assuming a homogeneous tissue, the thermal dynamics model on the facial skin is given by the Pennes' bioheat equation [19]:

$$\rho C \frac{\partial T}{\partial t} = -k \nabla^2 T + \omega_b c_b (T_b - T) + Q_e \tag{3}$$

where T denotes the temperature of the skin tissue, ρ is the density of the tissue, C is the specific heat capacity of the tissue, k denotes the thermal conductivity of the tissue, ω_b is the blood perfusion rate, c_b is the specific heat capacity of the blood T_b is the arterial blood temperature, and finally Q_e represents the volumetric heat source due to the laser-skin photothermal interaction.

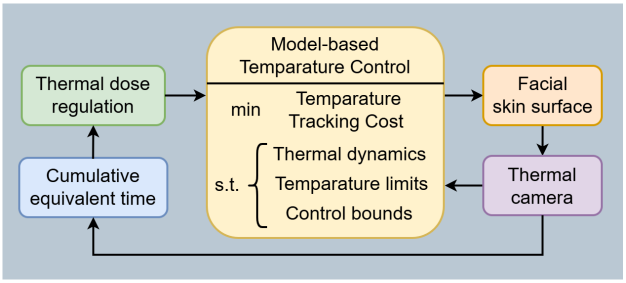

Fig. 5. Schematic illustration of the architecture for thermal dose control.

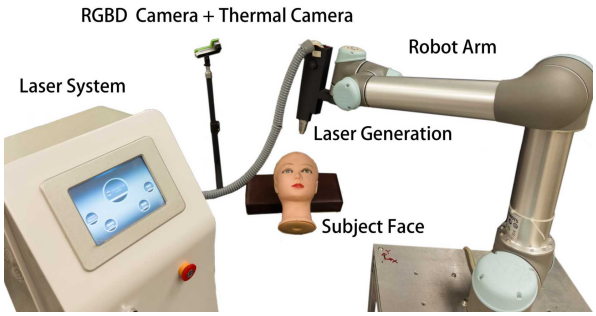

Fig. 6. Experimental setup for performing robotic cosmetic dermatology using photorejuvenation.

In our application, the external heat source in (3) is from the cosmetic laser pulses. The heat source resulting from the laser-skin photothermal interaction is given by $Q_e = \mu_a I S(t)$, where μ_a represents the absorption coefficient of the tissue, $I = I_0 I_r I_z$ denotes the laser intensity with I_0 the peak intensity, I_r the radial spread of light following a Gaussian profile, and I_z accounting for the exponential attenuation of intensity due to absorption and scattering. Besides, $S(t)$ is a switch function modeling the activation and deactivation of the laser power source. Its value takes one during the pulse width period τ_w and equals zero otherwise.

Since the control signal $S(t)$ that induces heat to the tissue is a discontinuous function, designing a continuous control signal is preferable to simplify the control design. As the frequency of the laser pulses can usually be directly chosen, we would like to use the frequency of the laser pulses as our control input. By observing that $\int S(t)\,dt = \int \tau_w u(t)\,dt$, we specify the control signal as $S(t) = \tau_w u(t)$.

Treatments that involve heating biological tissues have mainly three categories, namely hyperthermia, coagulation, and vaporization, depending on irradiation exposure time and the elevated tissue temperature [19]. Hyperthermia refers to the overheating of tissue, coagulation involves the denaturation of proteins to solidify tissue, and vaporization is the conversion into gas through intense heat. In cosmetic dermatology, skin photorejuvenation-based thermal treatment falls into hyperthermia. In hyperthermia, one important notion is the thermal dose, which is quantified as the cumulative equivalent time at a treatment temperature T_0. The thermal dose D is calculated as [20]:

$$D = \int_0^t R^{T_0 - T(\tau)}\,d\tau \tag{4}$$

Here $T_0 = 43\,^\circ\mathrm{C}$ is the thermal isoeffective temperature, indicating the reference point for calculating equivalent thermal dose. R is a temperature-dependent damage rate constant with $R = 0.25$ when $T < 43\,^\circ\mathrm{C}$ and $R = 0.5$ when $T \geq 43\,^\circ\mathrm{C}$.

In clinical practice, the desired thermal dose is typically the primary parameter of interest, as it directly determines the effectiveness and precision of thermal treatments for the skin. To this end, our controller first takes the reference thermal dose and then computes the required temperature value based on (4). Subsequently, to track the desired temperature value, we formulate a model predictive temperature controller as follows:

$$\min \sum_{i=0}^{N-1} \|T_{k+i} - T_{\mathtt{ref}}\|^2 + \lambda \|u_{k+i}\|^2 \tag{5a}$$

$$T_{k+1} = AT_k + Bu_k + d_k, \tag{5b}$$

$$\text{s.t. } T_{\mathtt{min}} \leq T_{k+i} \leq T_{\mathtt{max}}, \tag{5c}$$

$$u_{\mathtt{min}} \leq u_{k+i} \leq u_{\mathtt{max}}, \tag{5d}$$

where (5a) represents the control objective, which consists of tracking the desired temperature value and penalizing the control strength with λ being a regularization parameter. The discretized thermal dynamics is specified by (5b), where A is the state transition matrix, B is the control input matrix, and d is the thermal disturbances, as shown by [21]. The temperature limits and the control limits are specified by (5c) and (5d), respectively.

An overall schematic illustration of the thermal dose control framework during the skin photorejuvenation procedure is shown in Fig. 5.

IV. Evaluation Results

A. Experimental Setup

The experimental setup for photorejuvenation-based cosmetic dermatology is illustrated in Fig. 6. An Nd:YAG laser serves as the laser source, operating at a wavelength of $1064\,\mathrm{nm}$. Since this wavelength falls within the near-infrared (NIR) spectrum and is invisible to the human eye, the laser generator also emits a low-energy flash of visible light to aid in visualizing the laser pulses. The laser generator is mounted at the end link of a UR5 robotic arm from Universal Robots.

An RGB-D camera from RealSense is positioned above the face to capture the subject's facial skin. In addition, a Boson thermal camera from FLIR is employed to monitor the skin's surface temperature.

We utilize the FaceVerse dataset [22], which contains 2,688 high-quality 3D face scans encompassing 21 expressions from 128 distinct identities. We carefully selected 110 representative fitted models that comprehensively cover diverse facial characteristics and features for ground truth annotation. The preprocessing pipeline consists of the following steps:

TABLE I

ABLATION STUDY ON LABEL SOFTENING STRATEGIES

Method	Training Stability	Path Error (mm)	Coverage (%)
Hard Labels	Failed	> 10.0	< 40.0
Soft Labels	**High**	**2.12**	**94.7**

- Point cloud extraction from fitted models while preserving geometric fidelity
- Uniform sampling to maintain consistent point density
- Normal vector computation for each point to capture local surface characteristics

The processed dataset is split into training (70%), validation (15%), and test (15%) sets, ensuring a balanced distribution of identities and expressions across splits.

B. Ablation Study

1) Effect of Label Softening: To investigate the effectiveness of our probabilistic label transformation, we compared binary labels with our proposed soft label approach:

- **Hard Labels**: Binary labels indicating trajectory/non-trajectory points (0 or 1);
- **Soft Labels**: Our proposed Gaussian transformation with bandwidth σ.

The results demonstrate that hard binary labels are unsuitable for our trajectory learning task. The network fails to converge to meaningful solutions with hard labels, showing poor learning stability and severe generalization issues. This is primarily due to the discrete nature of binary labels, which creates sharp discontinuities in the supervision signal and provides no meaningful gradients in regions between demonstrated trajectories.

The transition to soft labels brings dramatic improvements. Our probabilistic transformation enables stable training and effective performance, reducing path error by over 75% compared to hard labels. These findings strongly support our decision to reformulate trajectory learning as a probability field prediction task rather than a binary classification problem. The significant performance gap between hard and soft labels indicates that the probabilistic representation is an implementation choice and a crucial design decision that fundamentally enables effective trajectory learning in our framework.

2) Impact of Geometric Features: To evaluate the impact of geometric features in our framework, we conducted an ablation study comparing different feature combinations:

- **Basic**: Point contains only Cartesian coordinates (x, y, z);
- **Basic+Normal**: Cartesian coordinates and surface normals (x, y, z, n_x, n_y, n_z);
- **Full**: Cartesian coordinates, surface normals, and geometric features, i.e., linearity, planarity, sphericity (x, y, z, nx, ny, nz, $\mathcal{L}$, $\mathcal{P}$, $\mathcal{S}$).

We analyzed the convergence speed and final performance metrics for each variant. As described in Section III-A, the geometric features were computed using local neighborhoods of $k = 16$ points.

TABLE II

ABLATION STUDY ON GEOMETRIC FEATURES

Method	Conv. Time (epochs)	Path Error (mm)	Smoothness
Basic	> 1800	4.42	0.56
Basic+Normal	1150	2.85	0.83
Full	**680**	**2.12**	**0.91**

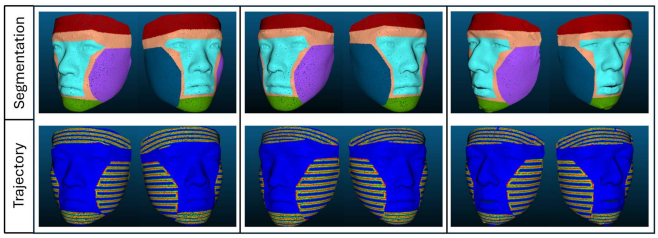

Fig. 7. Facial region segmentation and trajectory generation results within the segmented regions for training samples.

Table II shows that incorporating geometric features significantly improves convergence speed and final performance. The full model with geometric features achieves convergence in approximately half the number of epochs compared to the basic model. This acceleration in training can be attributed to the rich structural information encoded by the geometric features, particularly the local shape descriptors (linearity, planarity, and sphericity).

Moreover, we observe that the geometric features contribute to better path planning quality, reducing path generation error by 38% compared to using only basic coordinates. The improvement in smoothness metrics (from 0.56 to 0.91) suggests that geometric features help the network better understand local surface characteristics, leading to more natural and continuous laser paths. This is particularly important in regions with high curvature or complex geometric structures, where basic coordinates alone may not capture sufficient local context for optimal path planning.

Adding surface normal information (Basic+Normal) provides an intermediate improvement, but the full geometric feature set shows superior performance across all metrics. This demonstrates that our carefully designed geometric features capture complementary aspects of local surface geometry that are particularly relevant for laser path planning tasks.

C. Deep Learning Model Performance

We evaluated the performance of our proposed deep learning framework in three key aspects: (1) the model's segmentation accuracy and trajectory generation quality across different facial regions; (2) comparative analysis against existing methods such as PaintNet; and (3) generalization capability on unseen faces and practical implementation.

1) Segmentation and Trajectory Generation Quality: Our model achieved an average segmentation accuracy of 94.2% across distinct facial regions on the test set. The central facial area (shown in cyan) demonstrated the highest accuracy at 96.8%, while the peripheral regions (including the red,

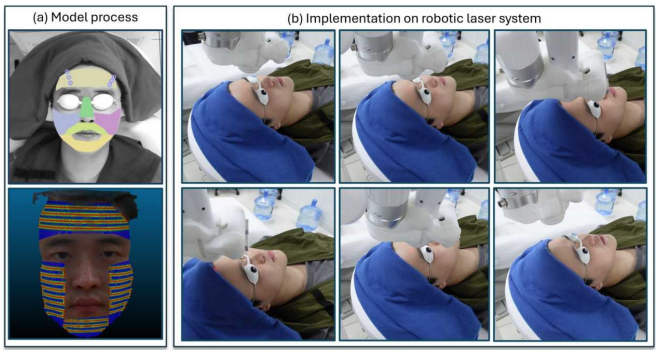

Fig. 8. (a) Visualization of automated region segmentation and trajectory planning results on an unseen test subject. (b) Sequential frames showing the automated treatment process execution on the robotic laser platform.

TABLE III

COMPARISON OF TRAJECTORY GENERATION METHODS

Metric	Ours	PaintNet	Manual
Data complexity	Low	High	N/A
Training samples	∼120	>2,000	N/A
Process time (s)	1.3	8-10	1200-2400
Uniformity (%)	90.5	87.3	70-85
Coverage (%)	94.1	91.8	85-92

blue, purple, and green segments) showed relatively lower accuracy at 91.3%. This variation in accuracy is consistent with the inherent challenges in defining precise boundaries for transitional facial zones.

For trajectory field prediction, our model generated high-quality probability distribution maps from which the extracted trajectories exhibited superior smoothness and adaptation to facial contours. As shown in Fig. 7, the generated trajectories maintain consistent coverage density across different facial shapes while naturally conforming to variations in facial curvature.

To validate the generalization capability of our method, we conducted experimental evaluations on subjects outside our training dataset. As shown in Fig. 8(a), our segmentation algorithm successfully adapts to new facial structures, automatically dividing the face into five distinct regions. The trajectory planning module demonstrates robust adaptation, generating treatment paths that maintain consistent coverage while conforming to individual facial geometry.

The practical efficacy of our method was further validated through implementation on a robotic laser platform, as illustrated in Fig. 8(b). The sequential frames show the automated execution process, where the system maintains precise positioning and smooth trajectory following. These results confirm that our approach can effectively generalize to new subjects while maintaining treatment precision and safety requirements.

2) Comparison with Existing Methods: Table III compares our method and existing techniques. Compared to PaintNet, our approach significantly reduced data preparation requirements, needing only approximately 120 region-labeled facial point clouds versus PaintNet's 2,000+ trajectory samples.

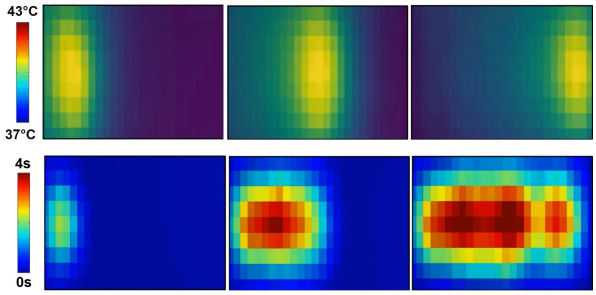

Fig. 9. Evolution trace of temperature (*upper row*) and thermal dose (*lower row*) on the facial skin at different time instances.

Notably, our method demonstrated significant advantages in processing time. The entire pipeline from facial point cloud acquisition to final trajectory generation took an average of only 1.3 seconds, compared to 8-10 seconds for PaintNet and 20-40 minutes for manual planning. This efficiency improvement is crucial for practical clinical applications, as confirmed by our real-world implementation tests shown in Fig. 8.

D. Thermal Regulation

The performance of regulating the thermal doses in our system is verified to showcase the treatment capability in photorejuvenation. We first test the performance on a realistic tissue phantom manufactured of gelatin. Our goal is to administer the desired value of thermal doses along the path on the surface of the tissue. To this end, the robot arm is set to move at a constant speed to manipulate the laser generator over the tissue. During the experiment, the diameter of the laser beam is set to be $2\,\mathrm{mm}$ and the laser energy is $0.75\,\mathrm{J}$

For the MPC controller, the reference thermal dose is specified by $D_{\mathtt{ref}} = 4$ CES The control signal is the PWM frequency constrained between 0 and $10\,\mathrm{Hz}$. Besides, the temperature limit is $43\,^{\circ}\mathrm{C}$. The time step of the controller $0.1\,\mathrm{s}$ and the horizon of the MPC controller is $1\,\mathrm{s}$, which corresponds to ten prediction steps. The initial temperature of the tissue is $36\,^{\circ}\mathrm{C}$, and the tissue starts with a zero thermal dose in the beginning.

The results are illustrated using the heat map shown in Fig. 9, where the surface temperature is shown in the upper row, and the administered thermal dose is shown in the lower row. It can be seen that the temperature is controlled without violating the specified constraints, and the thermal dose is administered to the desired values.

V. CONCLUSION

In this paper, we present the development of healthcare robots in the beauty industry. Particularly, we investigate photorejuvenation-based cosmetic dermatology, which is a typical aesthetic service in the beauty industry. At the core of the proposed methodology lies the deep learning-based trajectory planning for the cosmetic laser, where the learning model takes the point cloud as input and outputs the cosmetic laser path. Additionally, we develop a model-based method

for thermal dose regulation. Our experimental findings suggest that deploying robotics for photorejuvenation treatments holds considerable promise.

There are certainly a few limitations associated with our current achievements. The current robotic system lacks the capability to respond to language instructions, which can affect the precision and safety of the treatment. Moving forward, we aim to address these limitations by enhancing the responsiveness of the robot [23]. Also, we seek to improve the safety of the robot during its operation [24]. Furthermore, we plan to conduct qualitative studies with recipients of the service to gather feedback and improve market acceptance of the robotic system.

REFERENCES

[1] J. M. Marks, E. L. Mackenzie, K. R. Rose, E. C. Shaffrey, J. D. Larson, J. W. Siebert, and A. M. Afifi, "Preparing for a Crowded Cosmetic Market: A Resident Training Model for Minimally Invasive Cosmetic Treatments," *Plastic & Reconstructive Surgery*, vol. 150, no. 2, pp. 317–324, Aug. 2022. [Online]. Available: https://journals.lww.com/10.1097/PRS.0000000000009357

[2] A. Elder, C. Ring, K. Heitmiller, Z. Gabriel, and N. Saedi, "The role of artificial intelligence in cosmetic dermatology—Current, upcoming, and future trends," *Journal of Cosmetic Dermatology*, vol. 20, no. 1, pp. 48–52, 2021.

[3] A. Duan, W. Liuchen, J. Wu, R. Camoriano, L. Rosasco, and D. Navarro-Alarcon, "Learning rhythmic trajectories with geometric constraints for laser-based skincare procedures," *IEEE Transactions on Robotics*, vol. 41, pp. 1956–1973, 2025.

[4] K. Erdoğan, O. Acun, A. Küçükmanísa, R. Duvar, A. Bayramoğlu, and O. Urhan, "KEBOT: an artificial intelligence based comprehensive analysis system for fue based hair transplantation," *IEEE Access*, vol. 8, pp. 200 461–200 476, 2020.

[5] Y. Homma and K. Suzuki, "A robotic brush with surface tracing motion applied to the face," in *Social Robotics: 10th International Conference, ICSR 2018, Qingdao, China, November 28-30, 2018, Proceedings 10*. Springer, 2018, pp. 513–522.

[6] D. Shira, "Lash extensions from an AI robot: Is this the future of beauty?" https://www.forbes.com/sites/dahvishira/2024/05/28/lash-extensions-from-an-ai-robot-is-this-the-future-of-beauty/.

[7] Z. Xia, F. Ahmad, H. Deng, L. Jiang, W. Qin, Q. Zhao, and J. Xiong, "Robotics application in dentistry: A review," *IEEE Transactions on Medical Robotics and Bionics*, 2024.

[8] N. Dennler, E. Shin, M. Matarić, and S. Nikolaidis, "Design and evaluation of a hair combing system using a general-purpose robotic arm," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 3739–3746.

[9] M. Khoramshahi, G. Henriks, A. Naef, S. S. M. Salehian, J. Kim, and A. Billard, "Arm-hand motion-force coordination for physical interactions with non-flat surfaces using dynamical systems: Toward compliant robotic massage," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 4724–4730.

[10] X. Shen, J. Zhang, C. Yan, and H. Zhou, "An automatic diagnosis method of facial acne vulgaris based on convolutional neural network," *Scientific reports*, vol. 8, no. 1, p. 5839, 2018.

[11] M. Chen, P. Zhou, D. Wu, L. Hu, M. M. Hassan, and A. Alamri, "AI-Skin: Skin disease recognition based on self-learning and wide data collection through a closed-loop framework," *Information Fusion*, vol. 54, pp. 1–9, 2020.

[12] Y. Tsumaki, T. Kon, A. Suginuma, K. Imada, A. Sekiguchi, D. N. Nenchev, H. Nakano, and K. Hanada, "Development of a skincare robot," in *2008 IEEE International Conference on Robotics and Automation*. IEEE, 2008, pp. 2963–2968.

[13] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.

[14] A. Duan, I. Batzianoulis, R. Camoriano, L. Rosasco, D. Pucci, and A. Billard, "A structured prediction approach for robot imitation learning," *The International Journal of Robotics Research*, vol. 43, no. 2, pp. 113–133, 2024.

[15] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space," *Advances in neural information processing systems*, vol. 30, 2017.

[16] G. Tiboni, R. Camoriano, and T. Tommasi, "PaintNet: Unstructured multi-path learning from 3D point clouds for robotic spray painting," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 3857–3864.

[17] H. Chen, S. Huo, M. Muddassir, H.-Y. Lee, Y. Liu, J. Li, A. Duan, P. Zheng, and D. Navarro-Alarcon, "PSO-based optimal coverage path planning for surface defect inspection of 3c components with a robotic line scanner," *IEEE Transactions on Instrumentation and Measurement*, 2025.

[18] L. Hu, A. Duan, M. Li, A. Cherubini, L. Li, and D. Navarro-Alarcon, "Paint with the sun: A thermal-vision guided robot to harness solar energy for heliography," *IEEE Sensors Journal*, vol. 22, no. 18, pp. 18 130–18 142, 2022.

[19] F. Xu, T. J. Lu, K. A. Seffen, and E. Y. K. Ng, "Mathematical modeling of skin bioheat transfer," *Applied Mechanics Reviews*, vol. 62, no. 5, p. 050801, 07 2009. [Online]. Available: https://doi.org/10.1115/1.3124646

[20] S. A. Sapareto and W. C. Dewey, "Thermal dose determination in cancer therapy," *International Journal of Radiation Oncology-Biology-Physics*, vol. 10, no. 6, pp. 787–800, 1984.

[21] M. Muddassir, G. Limbert, B. Zhang, A. Duan, J.-J. Tan, and D. Navarro-Alarcon, "Model predictive thermal dose control of a robotic laser system to automate skin photorejuvenation," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 2, pp. 737–747, 2022.

[22] L. Wang, Z. Chen, T. Yu, C. Ma, L. Li, and Y. Liu, "FaceVerse: a fine-grained and detail-controllable 3D face morphable model from a hybrid dataset," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR2022)*, June 2022.

[23] F. Wang, S. Lyu, P. Zhou, A. Duan, G. Guo, and D. Navarro-Alarcon, "Instruction-augmented long-horizon planning: Embedding grounding mechanisms in embodied mobile manipulation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 14, 2025, pp. 14 690–14 698.

[24] H. Ding, A. Duan, Z. Sun, L. Rozo, N. Jaquier, D. Song, and Y. Nakamura, "Towards safe imitation learning via potential field-guided flow matching," in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2025, pp. 11 693–11 700.