IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 24, 2019, accepted February 6, 2019, date of publication February 14, 2019, date of current version March 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2899379

Machine Learning Approaches for Prediction of Facial Rejuvenation Using Real and Synthetic Data

SYED AFAQ ALI SHAH[1], MOHAMMED BENNAMOUN[2], AND MICHAEL K. MOLTON[3]
[1]Discipline of Information Technology, Mathematics and Statistics, Murdoch University, Perth, WA 6150, Australia
[2]Department of Computer Science and Software Engineering, The University of Western Australia, Cawley, WA 6009, Australia
[3]Epiclinic Pty, Ltd., Adelaide, SA 5000, Australia

Corresponding author: Syed Afaq Ali Shah (afaq.shah@murdoch.edu.au)

This work has been supported by the Australian Research Council Linkage Project LP130100138. The author would like to thank NVIDIA for their Titan V GPU grant.

ABSTRACT This paper proposes a novel machine learning approaches to predict the outcome of facial rejuvenation prior to a cosmetic procedure. This is achieved by estimating the required amount of dermal filler volume that needs to be applied on the face by learning the underlying structural mapping from the pretreatment and posttreatment 3D face images. We develop and train our proposed deep neural network, called Rejuv3DNet, designed specifically to predict the dermal filler volume. We also propose the kernel regression (KR)-based model to validate and improve our volume estimation results using regression. Our other contributions include the development of the *first* 3D face cosmetic dataset, which consists of real-world pretreatment and posttreatment 3D face images and a novel technique for the generation of synthetic cosmetic treatment 3D face images. Our experimental results show that the proposed Rejuv3DNet and the KR model achieve 62.5% and 66.67%, respectively, on real-world data, while these techniques achieve a prediction accuracy of 75.2% and 89.5%, and 77.2% and 90.1% on our two different synthetic datasets. Our proposed techniques have been found to be computationally efficient, achieving near real-time prediction performance. The reported accuracies are our preliminary results for proof of concept, which can be improved with more data. The proposed approach has the potential for further investigation in the cosmetic surgery domain.

INDEX TERMS Deep learning, deep neural network, facial analysis, regression.

I. INTRODUCTION

During the past decade, the field of computer vision and machine learning has witnessed significant improvements in various face analysis tasks such as face recognition [1]–[3], face detection [4]–[6], facial landmark localization [7]–[9], facial expression analysis [10]–[12] and 3D face reconstruction [13]–[15]. This is primarily attributed to the fact that the research community has made a considerable effort to collect and annotate images [16]–[19] and to the deep learning methodologies which capitalized on the availability of such large amount of data. However, this is not the case in the cosmetic imaging domain. Unlike traditional face recognition, the imaging data related to cosmetic practice is scarce and fraught with legal concerns such as ethical approval and patient privacy. Under these constraints, the prediction and assessment of potential outcomes of cosmetic procedures can be complicated, especially when multiple treatment modalities are proposed [20], [21]. This uncertainty may not only be a concern to the patient, but also to the medical professional performing the procedures. In existing procedures, a pre treatment scan of the patient's face is acquired during their first visit as shown in Fig. 1(a). Cosmetic practitioners then prepare a treatment sheet showing the amount of filler volume required per area (Fig. 1(b)) based on their experience and perform the procedure. A post treatment scan is also acquired to reflect the effect of the cosmetic procedure as shown in Fig. 1(c).

When asked by the patient *how will I look afterwards?*, the cosmetic practitioner presently can only offer descriptive, subjective replies or provide before and after 2D images of previous patients. With experience, and over time, such subjective predictions of the practitioner are likely to improve but remain difficult for the patient to form a predictive picture of appearance after treatment. Subjective dissatisfaction is relatively common in these situations and its incidence is likely to be significantly higher for newcomers to the field who have yet to gain

The associate editor coordinating the review of this manuscript and approving it for publication was Xian Sun.

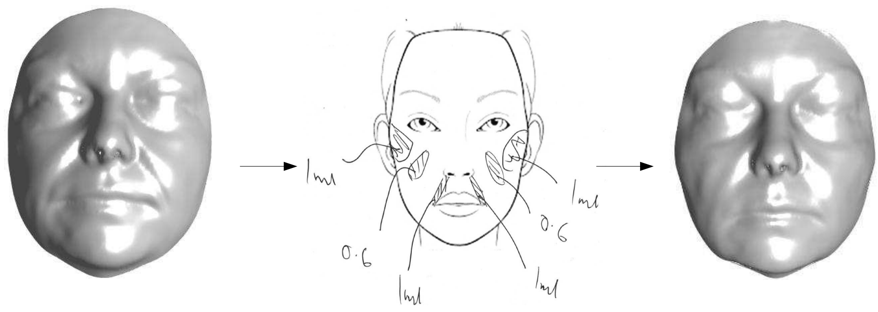

FIGURE 1. 3D Face Scan: (a) Pre Treatment (before cosmetic procedure). (b) Treatment Sheet. (c) Post Treatment (after cosmetic procedure).

enough experience to be acquired from high numbers of treatments.

In this paper, we address these challenges associated with cosmetic procedures. We propose machine learning-based predictive modeling methods to predict the outcome of a cosmetic procedure prior to the actual treatment. This is accomplished by capturing the pre and post treatment 3D face images and learning the underlying structural changes in these images to estimate the required amount of dermal filler volume. The contributions of this paper can be summarized as follows:

- We develop a cosmetic 3D face dataset, which consists of real world pre and post treatment 3D faces of patients. These images have been acquired by an expert in a clinical setting. **To the best of our knowledge, this is the first 3D face cosmetic dataset**.
- To address the data scarcity problem of cosmetic imaging, we also develop a mesh morphing algorithm for the generation of synthetic 3D face images to reflect different cosmetic treatments.
- We design a multi-layered Rejuv3DNet to predict the outcome of facial rejuvenation by estimating the required dermal filler volume in different treatment areas.
- We also design and develop a kernel regression-based KR model to validate and improve our volume estimation results.
- To the best of our knowledge, this is the first ever research effort by computer vision and machine learning researchers in this particular direction.

The rest of this paper is organized as follows. The next section surveys the related work. Section III introduces our proposed methodology for synthetic data generation using mesh morphing and predictive modeling techniques. Our experimental results and evaluation of the proposed techniques are provided in Section IV. Finally, the paper is concluded in Section VI.

II. RELATED WORK

Facial analysis is one of the most researched topics in machine learning and computer vision [22]–[25]. Here, we present the most relevant works to this paper and divide them into 'facial analysis in cosmetic medicine' and 'synthetic data generation'.

A. FACIAL ANALYSIS IN COSMETIC MEDICINE

In cosmetic medicine, many attempts have been made to derive models that can be used to evaluate existing pathophysiology of aging of the facial features [26]. Most have focused on the many facets of classifying wrinkles. Few have been derived to evaluate outcomes, outside of proprietary models used to acquire regulatory approval [27]. Such proprietary models are not available for general use and are also unlikely to be universal. Those that are in general use, involve parameters that, among others, range from the grading of wrinkles, to measuring skin rigidity, defining descriptive differences between 'crinkles', folds and microscopic evaluation of the elastoses. There is therefore a strong need to develop quantitative techniques for facial and surface quantification before and after cosmetic medical interventions.

The Lemperle rating scale [28], exists as one of the earliest methods of wrinkle evaluation. Using a scale of 0-5 and a comparative photographic evaluation, the wrinkles of the face are graded according to anatomical landmarks and severity of appearance. The Lemperle scale is therefore useful in providing useable language to record the physical examination of the patient that can be interpreted by other professionals. Criticism of the Lemperle scale includes the absence of the stipulation of using standardized photographic modeling. Buchner *et al.* [28] describe improved scaling using photonumeric wrinkle assessment, thereby extending the accuracy of the Lemperle scale. Their assessment method assists practitioners to record more physical observations for the individual patient and discuss results among colleagues. However, **the lack of precise quantification continues to produce subjective and arguable results** and provides limited interpretation of useful evidence to the patient indicating the degree of improvement that could be anticipated following treatment [29].

Hatzis [30] advocates the technique known as profilometry which to date has been the most accurate in attempting to quantify the depth and area of skin wrinkles. This method

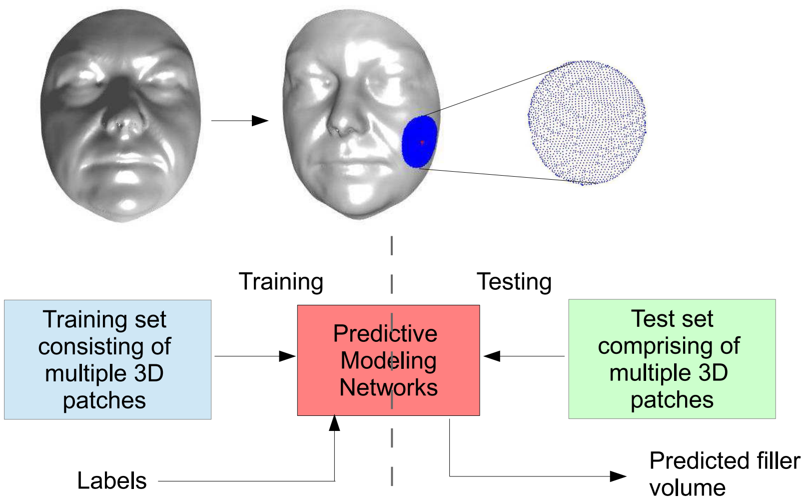

FIGURE 2. Proposed Framework. First row: Landmarks are carefully selected on the patient's face and a fixed size 3D patch is extracted around the landmark. Second Row: Several 3D patches extracted from multiple faces form our training and test datasets. This data is used for the training and testing of our Rejuv3DNet and KR predictive models to estimate the volume of dermal filler for a given test face prior to actual treatment.

requires the application of an impression of the skin not unlike impressions used in dentistry. A replica model is then made from the mold and measurements of actual physical topography of the region are made using a micrometer. From this data, parameters are measured including wrinkle depth, wrinkle area, wrinkle volume, wrinkle tissue reservoir volume (WRTV) and wrinkle to wrinkle distance. WRTV measurement appears particularly interesting as a tool to estimate a volume of dermal filler that might be required to replace lost volume. The limitation of the profilometry model is obviously the amount of effort involved for both patient and doctor in the taking of the impression(s) [31]. Less obvious is the limitation of understanding the dynamic movements of mimetic lines and wrinkles using this technology. In addition, profilometry appears to be more useful in very deep lines and tissue deficit and of only small benefit in evaluating more superficial and shallow deficits.

Carruthers and Carruthers [32] describe a number of subjective rating scales and a single quantitative scale for evaluating general changes in facial skin, none of which are widely used or intimately known. While the development of rating scales of aging of the face provide a useful vehicle for dialogue between practitioners and case presentations, ultimately patient satisfaction is the aim, and any method that can demonstrate quantifiable, as opposed to subjective, evidence of treatment takes the clinician one step closer to that goal.

B. SYNTHETIC DATA GENERATION

Richardson *et al.* [33] and Dou *et al.* [34] used Annotated Face Model (AFM) [35], Basel Face Model (BFM) [36] and 3D Morphable Model (3DMM) [37] techniques to generate thousands of synthetic 3D images for face reconstruction. These methods generate synthetic 3D faces within the linear space of a specific statistical face model. However, the generated faces have a variation of ± 3 standard deviations from the model mean with highly smooth surfaces. Kim *et al.* [38] fitted the BFM to 577 identities of FRGCv2 database [39] and induced 25 expressions in each identity. Minor pose variations between $\pm 10^o$ in roll, pitch and yaw for each original scan were also introduced. To simulate occlusions, eight random occlusion patches were introduced to each depth map to increase the database size to 123,325 scans. This method only increases the intra-person variations without augmenting the number of identities, which in this case remained 577. As can be noted, these synthetic data generation techniques are geared towards traditional facial analysis such as face reconstruction, recognition and expression evaluation. The generation of synthetic data specific to facial cosmetic procedure and prediction of dermal filler volume is a specialized task and is still an open ended research problem.

This paper aims to address the existing challenges in facial analysis techniques adopted in cosmetic procedures. To achieve this aim, we propose machine learning-based predictive modeling systems to predict dermal filler volume prior to actual cosmetic treatment. We also develop a real and synthetic 3D face dataset of pre and post treatment images and evaluate our techniques on these datasets to demonstrate its effectiveness. In the following sections, we shall discuss our proposed algorithms and introduce our new datasets.

III. PROPOSED TECHNIQUE

In this section, we discuss our proposed synthetic data generation and predictive modeling techniques for the prediction of dermal filler volume prior to treatment. Fig. 2 shows the block diagram of our proposed method. Given a 3D face,

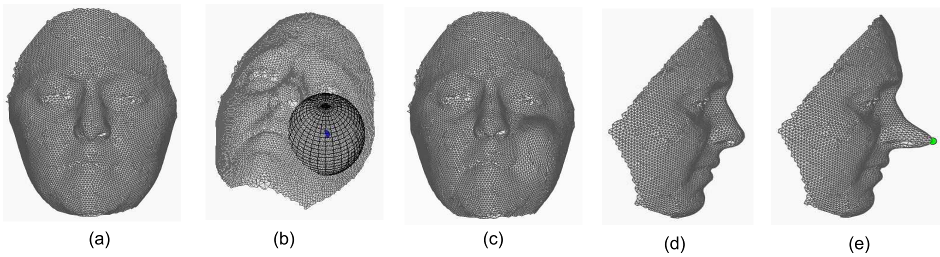

FIGURE 3. Proposed Mesh Morphing Algorithm. Illustration of synthetic image generation from a single real scan image. (a) Original Input Image. (b) Landmark selection with support radius. (c) Outcome of mesh morphing to synthetically generate the effect of a particular dermal filler volume. (e) Side view of original image. (d) Nose stretched to illustrate the outcome of our proposed mesh morphing algorithm.

a landmark at the location of the injection point (used for dermal filler) is first selected. A fixed size 3D patch is then extracted around a selected landmark (first row of Fig. 2). Multiple such patches are extracted from different locations of 3D faces (e.g., cheeks and temples) and from multiple faces to form our training and test datasets. The treatment sheet (Fig. 1) is used as a reference in this case. During training, 3D patches from the training dataset are fed to our proposed predictive modeling networks (second row of Fig. 2). Training is performed in a supervised way. During testing, a 3D patch from the test dataset is fed to the network, which then predicts the dermal filler volume to be injected on test 3D face. We address the problem of dermal filler volume prediction in two different ways. We use classification with Rejuv3DNet and regression for KR model to demonstrate the effectiveness of the proposed approach and for proof of concept. In the following, we shall discuss our synthetic data generation method and introduce Rejuv3DNet and KR predictive models for prediction of dermal filler volume prior to cosmetic treatment.

A. PROPOSED SYNTHETIC DATA GENERATION

Assume we have the 3D mesh (shape) s_h of a face with N vertices as a 3 × N matrix:

$$s_h = [x_i y_i z_i], \quad i = 1, \ldots, N \tag{1}$$

Let L_m denotes the matrix of facial landmarks on the 3D mesh s_h. These landmarks are the injection points on the face where the dermal filler is injected during the cosmetic procedure (treatment sheet Fig. 1(b)). One such landmark is shown in Fig. 3(b). R_i is the support radius (scale) related to the ith landmark ((Fig. 3(b))). R_i defines the influence hull of a given landmark i.e., the 3D region within which any 3D vertex is influenced by the related landmark. To achieve the desired synthetic data generation results, we propose an interactive mesh morphing technique to reconstruct multiple facial shapes, reflecting different cosmetic treatments, from a single image. Mesh morphing is a popular method used in computer graphic applications, as a powerful tool for free-shape designing and modeling. 3DMMs are powerful statistical models of 3D facial shape and texture.

Next, the displacement, $\Delta\vartheta_j$ of any vertex ϑ_j of the 3D model (mesh) is calculated as a weighted mean of all the landmarks as follows:

$$\Delta\vartheta_j = \omega(\vartheta_j).\chi \tag{2}$$

where ω and χ are the weighted and the morphing matrix, respectively.

The ith element of ω depends on the position of the vertex ϑ_j, position of the ith landmark L_{mi} and its radius R_i:

$$\omega_i(\vartheta_j) = f(\vartheta_j, L_{mi}, R_i) \tag{3}$$

where f is the weight function. f is generated by using the de Casteljau algorithm [40] and is given by,

$$f = \begin{Bmatrix} 1, & \left\|\vartheta_j - L_{mi}\right\|_2 = 0 \\ 0, & \left\|\vartheta_j - L_{mi}\right\|_2 > 0 \end{Bmatrix} \tag{4}$$

The morphing matrix χ can be calculated by applying Eq. 2 to all the landmarks as follows:

$$\begin{aligned} \Delta L_{m1} &= \omega(L_{m1}).\chi \\ \Delta L_{m2} &= \omega(L_{m2}).\chi \\ &\ldots \\ \Delta L_{mr} &= \omega(L_{mr}).\chi \end{aligned} \tag{5}$$

where ΔL_m is the matrix of displacement assigned at landmarks. These displacements are calculated when user selected landmark is moved. Then, χ can be calculated by solving a linear system:

$$\chi = \omega(L_m)^{-1}.\Delta L_m \tag{6}$$

Fig. 3(c) illustrates the generation of synthetic 3D images from a single input 3D image to reflect the effect of the dermal filler volume injected at a landmark location. Fig. 3(d) shows the side view, while Fig. 3(e) demonstrates the output of our algorithm for the nose region. As can be noted, our algorithm enables synthetic data generation by interactively modifying facial vertices to achieve the desired results for a given dermal filler volume. These 3D patches, around the landmarks, are extracted and saved in our database.

B. PROPOSED REJUV3DNET

To predict the volume of dermal filler for a given 3D face, we first define our Rejuv3DNet, which is shown in Fig. 4. Rejuv3DNet is a multi-layered neural network, which consists of two hidden layers, sandwiched between an input

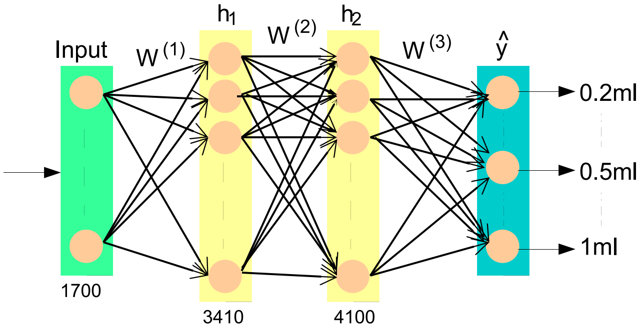

FIGURE 4. Proposed Rejuv3DNet. The network consists of four layers including the input, output and two hidden layers. The input to the network is the 3D patch extracted from the pre treatment face and the output is the volume of the dermal filler (0.2ml to 1ml).

and an output layer. These layers map the input data x to a representation $\hat{y}$ as follows:

$$\begin{aligned} h_1 &= s(w^{(1)}x + b^{(1)}) \\ h_2 &= s(w^{(2)}h_1 + b^{(2)}) \\ \hat{y} &= s(w^{(3)}h_2 + b^{(3)}) \end{aligned} \tag{7}$$

where $w^i \in \mathbb{R}^{d_{i-1}\times d_i}$ is the weight matrix for layer i with d_i nodes, $b^i \in \mathbb{R}^{d_i}$ is the biased vector and $s(.)$ is the non-linear activation function.[1] The network parameters i.e., weights of each layer are initialized with parameters drawn from a Gaussian distribution with zero mean and a standard deviation adjusted using the Xavier's method [41]. Given k training 3D face patches $P_1, P_2, \ldots P_k$ and their corresponding labels y_c, where these labels indicate the quantity of dermal filler volumes, Rejuv3DNet is trained on this training data. During training, we optimize the learning by Stochastic Gradient Descent (SGD) with standard L2 norm over the learned weights.

During classification, given a test 3D face patch P_t, the network predicts the class $\hat{y}_t$ of the test patch.

C. PROPOSED KR MODEL

To address the problem of dermal filler volume estimation using the regression strategy, we propose a kernel regression based model. Let there be c number of distinguished classes with P_i number of training 3D images (patches) from the ith class, $i = 1, 2, \ldots, c$. Let x be the test 3D image and our problem is to classify x as one of the classes $i = 1, 2, \ldots., c$. We next model the joint distribution $p(x, y)$ using a Parzen density estimator given as:

$$p(x, y) = \frac{1}{N}\sum_{n=1}^{N} f(x - x_n, y - y_n) \tag{8}$$

where x_n is the training set with labels y_n and $f(x, y)$ is the component density function. Next, we calculate the regression function $\hat{y}$, corresponding to the conditional average of the target variable i.e., test image conditioned on the input variable:

$$\begin{aligned} \hat{y} &= \varepsilon[y|x] = \int_{-\infty}^{\infty} yp(y|x)dy \\ &= \frac{\int yp(x, y)dy}{p(x, y)dy} \end{aligned} \tag{9}$$

This can easily be expanded to take the following form:

$$\hat{y} = \frac{\sum_n \int yf(x - x_n, y - y_n)}{\sum_n \int f(x - x_m, y - y_m)} \tag{10}$$

For simplicity, we assume that the component density function share zero mean i.e.,

$$\int_{-\infty}^{\infty} f(x, y)ydy = 0 \tag{11}$$

Finally, by using a simple change of variables, we obtain our KR model as follows:

$$\begin{aligned} \hat{y} &= \frac{\sum_n \Re(x - x_n)y_n}{\sum_m \Re(x - x_m)} \\ &= \sum_n \mathbb{K}(x, x_n)y_n \end{aligned} \tag{12}$$

where $m, n = 1, \ldots, K$ and the kernel function $\mathbb{K}(x, x_n)$ is given by:

$$\mathbb{K}(x, x_n) = \frac{\Re(x - x_n)}{\sum_m \Re(x - x_m)} \tag{13}$$

and $\Re(x)$ is defined as:

$$\Re(x) = \int_{-\infty}^{\infty} f(x, y)dy \tag{14}$$

We then calculate the distance measure between the predicted response $\hat{y}$ and the original response y,

$$d_i(y) = \left\| y - \hat{y}_i \right\|_2 \tag{15}$$

and rule in favor of the class with minimum distance.

$$\underbrace{min}_{i}\, d_i(y), i = 1, 2, \ldots, N \tag{16}$$

IV. EXPERIMENTAL RESULTS

We evaluated the performance of our proposed Rejuv3DNet and KR model for the task of dermal filler volume prediction. In the following, we introduce our new real 3D face cosmetic dataset and synthetic 3D face cosmetic dataset, and evaluate our proposed models on these two datasets.

A. REAL 3D FACE COSMETIC DATASET

Our 3D face cosmetic dataset is a real world dataset that contains 80 pre and post surgery 3D images pertaining to 40 subjects from different age groups and ethnicity. The age group of the subjects varies from 40 to 70 years (and includes both males and females) and include different ethnicity including Asian, Australian and African subjects. For each individual, there are two frontal face images with neutral expression. The first 3D image is taken before the cosmetic

[1] In our case we use a sigmoid defined as $s(z) = \frac{1}{1+e^{-z}}$

procedure and the second is taken after the procedure using high resolution Artec 3D scanner. Different dermal filler volumes have different impact on the facial features. Therefore, to diversify our dataset we include a wide variety of cases such as different dermal filler volumes and facial treatment areas i.e., cheeks and temples.

1) 3D DATA NORMALIZATION

The 3D images in our dataset are acquired from shoulder level up. The size and resolution of the meshes vary from an image to image. We perform an efficient fully automatic 3D face normalization to obtain uniformity in terms of the size and resolution of all meshes. The input to our algorithm is a raw 3D mesh. The algorithm removes outlier points, detects the nose tip and crops the Region Of Interest (ROI) around the detected nose tip. Figure 5 illustrates the different steps of the 3D face normalization. The details of each block are given below.

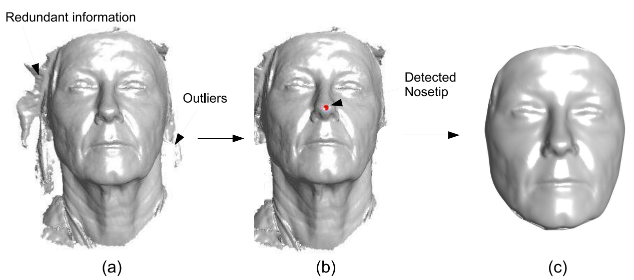

FIGURE 5. 3D Face Normalization. (a) Raw 3D mesh with outliers and redundant information. (b) Nose tip detection. (c) Crop region of interest around the nose tip.

We represent a 3D face by a matrix F ($N \times 3$) of point clouds, where N is the total number of points and each row of F corresponds to x, y, z coordinates of a point (vertex). A raw 3D face contains outlier points and redundant information as shown in Fig. 5(a). To filter the outliers, we find the mean μ and the standard deviation σ of the depths (z_i) of all points as follows:

$$\mu = \frac{1}{m}\sum_{i=1}^{m} z_i$$
$$\sigma = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(z_i - \mu)} \tag{17}$$

Any point whose depth is outside the $\mu \pm 3\sigma$ limit is considered an outlier and is removed. Figure 5(b) shows a 3D face after the removal of outlier points. To reduce dimensionality and computational complexity, we fit a surface of the form $z(x, y)$ to the 3D poincloud using the *gridfit* algorithm [42]. The resulting uniformly sampled pointcloud is input to the nose tip detection algorithm [43].

After a successful detection of the nose tip, a sphere of radius (r = 98mm) centered at the nose tip is used to crop the face. Fig. 5(c) shows the cropped 3D face.

Facial landmarks are then selected and 3D patches are extracted from pre treatment faces (as discussed in Sec. III) and their corresponding labels i.e., quantity of filler volume is obtained from the treatment sheet. We have five different class labels each indicating the volume of the dermal filler i.e., 0.2, 0.5, 0.6, 0.8 and 1ml. These patches along with their labels are stored in the database.

2) EVALUATION

To evaluate our proposed Rejuv3DNet and KR model, we randomly select 80% of the 3D images (patches) for training and 20% for testing. Next, ten runs of experiments are performed for different random selections of the training and test sets. Table 1 summarizes our average prediction results achieved by Rejuv3DNet and KR model for this small dataset.

TABLE 1. Dermal filler volume prediction results achieved by Rejuv3DNet and KR model on real 3D face cosmetic dataset.

Proposed Techniques	Prediction Results
3D-Div [44]	51.5
3D-Vor [25]	55.8
Rejuv3DNet	62.5 ± 2.4
KR Model	66.67 ± 3.2

As can be noted, our proposed Rejuv3DNet and KR model achieves 62.5% and 66.67% prediction accuracy, respectively. This performance is expected due to the smaller size of the training dataset which makes it hard for the proposed models to learn the underlying structural patterns. Another reason for this prediction performance is the close structural similarity due to the small difference in dermal filler volume. For instance, there is a slight difference in the facial surface for 0.5, 0.6 and 0.8ml, and our proposed algorithm could not easily distinguish between these classes, while making a prediction. We believe, a diverse real dataset with a very large number of 3D images will significantly improve the performance of the proposed predictive modeling systems. A large dataset will also enable training and testing of deep learning based prediction models, which will further improve the prediction performance.

To the best of our knowledge, this is the first ever reported result in the area of facial rejuvenation prediction. We therefore cannot find any baseline methods for comparison in this paper. However, we design some baseline methods for comparison purposes. These methods include 3D-Div [44], [45] and 3D-Vor [25] local features, and feature matching algorithms for comparison in this work. As can be noted the proposed techniques outperform the baseline methods by significant margin. A detailed analysis about the performance of Rejuv3DNet and KR model is provided in Sec. V.

B. SYNTHETIC 3D FACE COSMETIC DATASET

To generate our synthetic dataset, we use 3D face images from two different 3D face datasets including BU-3DEF and GavabDB 3D face dataset. A brief description of these datasets is provided below.

1) BU-3DEF DATASET

This dataset collects 3D face models and face images from 100 subjects of different genders, races and ages. Each subject has 25 face images, one neutral image and six expressions (happiness, disgust, fear, angry, surprise and sadness), each of which is at four levels of intensity. Figure 6 shows sample images from BU-3DEF dataset.

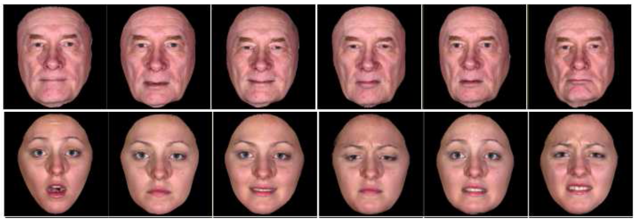

FIGURE 6. Sample images from BU-3DEF dataset.

2) GAVABDB 3D FACE DATASET

GavabDB is a 3D face dataset, which contains 549 3D face images of 61 different individuals (45 male and 16 female). There are 9 images for each subject i.e., 2 frontal views with neutral expression, 2 x-rotated views with neutral expression, 2 y-rotated views with neutral expression and 3 frontal gesture images. The subjects are Caucasian and their age ranges between 18 and 40 years old. Figure 7 shows sample images from Gavab dataset.

FIGURE 7. Sample images from GavabDB 3D face dataset.

For our synthetic dataset, we used the frontal neutral images from both datasets to reflect the clinical settings and morphed those images using our synthetic image generation algorithm (Sec. III-A) to generate different facial rejuvenation scenarios, such as injection of 0.2 to 1ml of dermal filler volume on the left and right side, for a given 3D face as shown in Fig. 8. Next, 3D patches were extracted from these 3D faces as discussed in Sec. III. This allowed us to generate around 200 synthetic 3D image patches in total for BU3DEF dataset and 244 patches for Gavab 3D face dataset. These patches were then used for the training and testing of our proposed techniques on these datasets.

EVALUATION

We evaluate the proposed Rejuv3DNet and KR model on our synthetic 3D face cosmetic dataset. For our experiments, we select 80% of the images for training and the remaining 20% for testing. The experiments are repeated 10 times for different random selections of the training and testing data. Our experimental results are reported in Table 2. The proposed models achieve a prediction accuracy of 75.2% and 89.5%, respectively, on the first synthetic 3D face dataset (i.e., images generated from BU-3DEF) and 77.2% and 90.1% on our second synthetic 3D face dataset i.e., GavabDB 3D face dataset. These results seem very promising and we believe that a large synthetic data generation has contributed to this improvement by enabling proper training of our proposed Rejuv3DNet and KR model. This implies that a large real (non-synthetic) images will help in the proper training of Rejuv3DNet and will further improve the prediction performance.

TABLE 2. Dermal filler volume prediction results achieved by Rejuv3DNet and KR model on synthetic 3D face cosmetic dataset .

Techniques	Prediction Results (BU-3DEF)	Prediction Results (Gavab3D)
3D-Div [44]	62.7	54.3
3D-Vor [25]	68.5	58.2
Rejuv3DNet	75.2 ± 3.1	77.2 ± 5.3
KR Model	89.5 ± 5.5	90.1 ± 2.2

C. COMPUTATIONAL TIME ANALYSIS

A comparison of the computational complexity of our proposed methods on the real 3D face cosmetic dataset using a Corei7 3.40GHz CPU with 8GB RAM is presented in Table 3. The proposed Rejuv3DNet requires comparatively more time for training than KR model. However, the training is performed offline. The proposed methods require only 0.0043 and 0.0005seconds, respectively to identify an image. The efficient testing of Rejuv3DNet is attributed to the fact that the proposed deep learning method only requires a few matrix multiplications (using Eqs. 7) during testing. KR model takes very less training and testing time which makes this model suitable for real time facial rejuvenation predictions.

TABLE 3. Computational time analysis on real data.

Proposed methods	Training (sec./image)	Testing (sec./image)
Rejuv3DNet	40.86	0.0043
KR Model	0.0052	0.0005

V. DISCUSSION AND ANALYSIS

In Table 1 and 2, we reported the prediction performance of our proposed Rejuv3DNet and KR model on real and synthetic data. As can be noted, the performance of Rejuv3DNet (neural network) is lower compared to the regression-based KR model. There are few reasons behind this low performance which can be explained intuitively. The **first** issue is the lack of diversity in data i.e., treatment cases. Our real world 3D cosmetic dataset has more cases where the patients were injected with 1ml and 0.5ml dermal filler volume, and very few 0.2ml and 0.6ml cases. We believe this resulted into

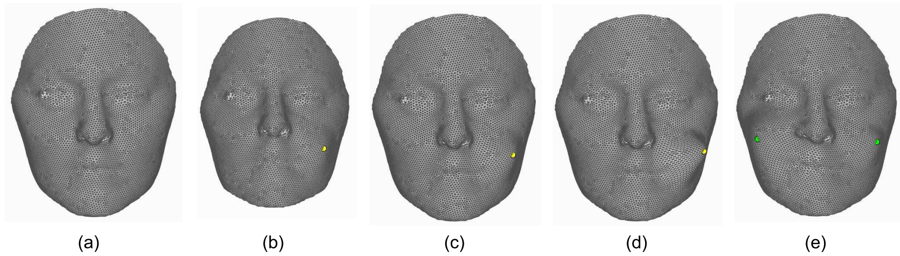

FIGURE 8. Proposed Synthetic 3D Face Cosmetic Dataset. (a) Original Image from BU-3DEF dataset. (b) - (d) Synthetically generated affect of different filler volumes. (e) Different filler volumes and support radius R_i to generate diverse 3D images.

bias of our neural network during training, as it was learning from a non-uniformly distributed data samples. The network incorrectly predicted 0.5ml for 0.2ml and 0.6ml volumes.

The **second** issue is with the size of the dataset i.e., training examples. Our datasets are smaller in size with only 40 subjects in the real 3D face dataset and 100 in the synthetic dataset. Neural networks are data driven models and have been shown to produce a high performance in data rich domains. For instance, in the case of natural image classification, neural networks have achieved considerably lower error rates. This is because of the fact that the dataset, ImageNet LSVRC, used for training and testing of the network consists of about 1.2 million training images belonging to 1000 different classes, while the test set has around 150,000 images.

We believe that the achieved accuracies can be increased by collecting a more diverse real world cosmetic treatment data with very large training and test dataset of 3D cosmetic faces. This will also enable the training of very deep neural networks, which will significantly reduce the dermal filler volume prediction error. To achieve this, we are still collecting more real world data and our dataset will grow over a period of time.

VI. CONCLUSION

This paper addresses a challenging problem of predicting the outcome of facial rejuvenation by estimating the volume of dermal filler prior to the actual treatment. To achieve this, the paper introduces the first real world 3D Face Cosmetic dataset and synthetic 3D Face Cosmetic datasets. The paper also proposes Rejuv3DNet and KR model, machine learning based approaches, to predict the volume of dermal filler for a given 3D face. Our experimental results show that Rejuv3DNet and KR model achieve very good prediction accuracy on real-world and synthetic 3D face datasets. The computational time analysis shows that the proposed techniques are very efficient which makes them suitable for real time applications.

ACKNOWLEDGMENT

The authors would like to thank NVIDIA for their Titan V GPU donation.

REFERENCES

[1] L. Lin, K. Wang, D. Meng, W. Zuo, and L. Zhang, "Active self-paced learning for cost-effective and progressive face identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 1, pp. 7–19, Jan. 2018.

[2] S. A. A. Shah, M. Bennamoun, and F. Boussaid, "Iterative deep learning for image set based face and object recognition," *Neurocomputing*, vol. 174, pp. 866–874, Jan. 2016.

[3] S. A. A. Shah, U. Nadeem, M. Bennamoun, F. Sohel, and R. Togneri. (2017). "Efficient image set classification using linear regression based image reconstruction." [Online]. Available: https://arxiv.org/abs/1701.02485

[4] S. Zafeiriou, C. Zhang, and Z. Zhang, "A survey on face detection in the wild: Past, present and future," *Comput. Vis. Image Understand.*, vol. 138, pp. 1–24, Sep. 2015.

[5] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, "S^3FD: Single shot scale-invariant face detector," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 192–201.

[6] H. Wu, K. Zhang, and G. Tian, "Simultaneous face detection and pose estimation using convolutional neural network cascade," *IEEE Access*, vol. 6, pp. 49563–49575, 2018.

[7] G. Azzopardi, A. Greco, A. Saggese, and M. Vento, "Fusion of domain-specific and trainable features for gender recognition from face images," *IEEE Access*, vol. 6, pp. 24171–24183, 2018.

[8] S. A. A. Shah, M. Bennamoun, and M. Molton, "A training-free mesh upsampling and morphing technique for 3D face rejuvenation," in *Proc. Int. Conf. Image Vis. Comput.*, New Zealand, Nov. 2018, pp. 1–6.

[9] S. A. A. Shah, M. Bennamoun, and F. Boussaid, "Automatic 3D face landmark localization based on 3D vector field analysis," in *Proc. Int. Conf. Image Vis. Comput.*, New Zealand, Nov. 2015, pp. 1–6.

[10] M. Z. Uddin, W. Khaksar, and J. Torresen, "Facial expression recognition using salient features and convolutional neural network," *IEEE Access*, vol. 5, pp. 26146–26161, 2017.

[11] I. Abbasnejad, S. Sridharan, D. Nguyen, S. Denman, C. Fookes, and S. Lucey, "Using synthetic data to improve facial expression analysis with 3D convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Oct. 2017, pp. 1609–1618.

[12] Y. Ding, Q. Zhao, B. Li, and X. Yuan, "Facial expression recognition from image sequence based on LBP and Taylor expansion," *IEEE Access*, vol. 5, pp. 19409–19419, 2017.

[13] T. Lu, Z. Xiong, Y. Zhang, B. Wang, and T. Lu, "Robust face super-resolution via locality-constrained low-rank representation," *IEEE Access*, vol. 5, pp. 13103–13117, 2017.

[14] P. Huang, G. Gao, C. Qian, G. Yang, and Z. Yang, "Fuzzy linear regression discriminant projection for face recognition," *IEEE Access*, vol. 5, pp. 4340–4349, 2017.

[15] S. A. A. Shah, M. Bennamoun, and M. Molton, "A fully automatic framework for prediction of 3D facial rejuvenation," in *Proc. Int. Conf. Image Vis. Comput.*, New Zealand, Nov. 2018, pp. 1–6.

[16] J. Shin *et al.*, "Open access dataset for EEG+NIRS single-trial classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1735–1745, Oct. 2017.

[17] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar, "Localizing parts of faces using a consensus of exemplars," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2930–2940, Dec. 2013.

[18] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: Database and results," *Image Vis. Comput.*, vol. 47, pp. 3–18, Mar. 2016.

[19] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: The first facial landmark localization challenge," in *Proc. Int. Conf. Comput. Vis. Workshops (ICCVW)*, Dec. 2013, pp. 397–403.

[20] M. Molton, S. A. A. Shah, and M. Bennamoun, "Improving the face of cosmetic medicine: An automatic three-dimensional analysis system for facial rejuvenation," *J. Aesthetic Reconstructive Surgery*, vol. 2, no. 2, 2016. [Online]. Available: http://aesthetic-reconstructive-surgery.imedpub.com/improving-the-face-of-cosmetic-medicine-an-automatic-threedimensional-analysis-system-for-facial-rejuvenation.pdf

[21] S. A. A. Shah, M. Bennamoun, and F. Boussaid, "Performance evaluation of 3D local surface descriptors for low and high resolution range image registration," in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov. 2014, pp. 1–7.

[22] H. Patil, A. Kothari, and K. Bhurchandi, "3-D face recognition: Features, databases, algorithms and challenges," *Artif. Intell. Rev.*, vol. 44, no. 3, pp. 393–441, 2015.

[23] S. Soltanpour, B. Boufama, and Q. J. Wu, "A survey of local feature methods for 3D face recognition," *Pattern Recognit.*, vol. 72, pp. 391–406, Dec. 2017.

[24] H. Zhou, A. Mian, L. Wei, D. Creighton, M. Hossny, and S. Nahavandi, "Recent advances on singlemodal and multimodal face recognition: A survey," *IEEE Trans. Human-Mach. Syst.*, vol. 44, no. 6, pp. 701–716, Dec. 2014.

[25] S. A. A. Shah, M. Bennamoun, and F. Boussaid, "A novel algorithm for efficient depth segmentation using low resolution (Kinect) images," in *Proc. IEEE 10th Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2015, pp. 603–607.

[26] A. T. Nguyen, J. Ahmad, S. Fagien, and R. J. Rohrich, "Cosmetic medicine: Facial resurfacing and injectables," *Plastic Reconstructive Surgery*, vol. 129, no. 1, pp. 142e–153e, 2012.

[27] S. Buetow and K. Wallis, "The beauty in perfect imperfection," *J. Med. Humanities*, pp. 1–6, Dec. 2017. [Online]. Available: https://link.springer.com/article/10.1007%2Fs10912-017-9500-2

[28] L. Buchner, G. Vamvakias, and D. Rom, "Validation of a photonumeric wrinkle assessment scale for assessing nasolabial fold wrinkles," *Plastic Reconstructive Surgery*, vol. 126, no. 2, pp. 596–601, 2010.

[29] J. Ayer *et al.*, "A photonumeric scale for the assessment of atrophic facial photodamage," *Brit. J. Dermatol.*, vol. 178, no. 5, pp. e369–e369, 2018.

[30] J. Hatzis, "The wrinkle and its measurement—A skin surface Profilometric method," *Micron*, vol. 35, no. 3, pp. 201–219, 2004.

[31] V. Mazzarello, M. Ferrari, A. Bulla, G. Piu, and A. Montella, "Antiwrinkle effect of topical adhesive pads on crow's feet: How long does the effect last for?" *J. Cosmetic Dermatol.*, vol. 17, no. 3, pp. 484–490, 2018.

[32] A. Carruthers and J. Carruthers, "A validated facial grading scale: The future of facial ageing measurement tools?" *J. Cosmetic Laser Therapy*, vol. 12, no. 5, pp. 235–241, 2010.

[33] E. Richardson, M. Sela, and R. Kimmel, "3D face reconstruction by learning from synthetic data," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 460–469.

[34] P. Dou, S. K. Shah, and I. A. Kakadiaris, "End-to-end 3D face reconstruction with deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, Hawaii, vol. 5, Jun. 2017, pp. 21–26.

[35] I. A. Kakadiaris *et al.*, "Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 640–649, Apr. 2007.

[36] P. Paysan, R. Knothe, B. Amberg, S. Romdhani, and T. Vetter, "A 3D face model for pose and illumination invariant face recognition," in *Proc. 6th IEEE Int. Conf. Adv. Video Signal Based Surveill.*, Sep. 2009, pp. 296–301.

[37] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," in *Proc. 26th Annu. Conf. Comput. Graph. Interact. Techn.* Reading, MA, USA: Addison-Wesley, 1999, pp. 187–194.

[38] D. Kim, M. Hernandez, J. Choi, and G. Medioni. (2017). "Deep 3D face identification." [Online]. Available: https://arxiv.org/abs/1703.10714

[39] P. J. Phillips *et al.*, "Overview of the face recognition grand challenge," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2005, pp. 947–954.

[40] V. B. Anand, *Computer Graphics and Geometric Modeling for Engineers*. Hoboken, NJ, USA: Wiley, 1996.

[41] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.

[42] J. D'Errico, "Surface fitting using gridfit," MATLAB Central File Exchange, USA, Tech. Rep., 2008.

[43] X. Peng, M. Bennamoun, and A. S. Mian, "A training-free nose tip detection method from face range images," *Pattern Recognit.*, vol. 44, no. 3, pp. 544–558, 2011.

[44] S. A. A. Shah, M. Bennamoun, F. Boussaid, and A. A. El-Sallam, "A novel local surface description for automatic 3D object recognition in low resolution cluttered scenes," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Dec. 2013, pp. 638–643.

[45] S. A. A. Shah, M. Bennamoun, F. Boussaid, and A. A. El-Sallam, "3D-Div: A novel local surface descriptor for feature matching and pairwise range image registration," in *Proc. 20th IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2013, pp. 2934–2938.

SYED AFAQ ALI SHAH received the Ph.D. degree in computer vision and machine learning from The University of Western Australia. He is currently a Lecturer with the School of Engineering and Information Technology, Murdoch University, Perth, Australia. He is the co-author of the book *A Guide to Convolutional Neural Networks for Computer Vision*. His research interests include deep learning, 3D face/object recognition, face analysis, and 3D modeling. He received the Start Something Prize for Research Impact through Enterprise for 3D Facial Analysis Project funded by the Australian Research Council.

MOHAMMED BENNAMOUN received the M.Sc. degree in control theory from Queen's University, Canada, and the Ph.D. degree in computer vision from QUT, Brisbane, Australia. He is currently a Winthrop Professor with The University of Western Australia, Australia. He is the co-author of the book *Object Re cognition: Fundamentals and Case Studies* (Springer-Verlag, 2001). He has published more than 300 journal and conference publications. His interests include control theory, robotics, object recognition, artificial neural networks, and computer vision. He was selected to give tutorials at the European Conference on Computer Vision (ECCV 02), the 2003 International Conference on Acoustics Speech and Signal Processing, INTERSPEECH 2014, and CVPR 2016. He has served as the Guest Editor for a couple of special issues in international journals, such as the *International Journal of Pattern Recognition and Artificial Intelligence*.

MICHAEL K. MOLTON received the M.B.B.S. and Diploma degrees in cosmetics and dermatology laser. He is the President of FCPCA. He is a Cosmetic Medical Practitioner for 23 years. He is Currently a Formative Member of the Cosmetic Physicians Society of Western Australia, which ultimately has now become the National Body Cosmetic Physicians College of Australasia, where he is also the President. He is a pioneer in 3D quantitative assessment of the outcomes of facial changes as a result of cosmetic medical procedures that address the appearance of aging. He is the co-author of *Correlates of Dysmorphic Concern in People Seeking Cosmetic Enhancement* (Aust N Z.J Psychiatry, 2004), which, to date, has received 89 citations in other papers. He has authored papers on several other cosmetic medical subjects, including *Postinflammatory Hyperpigmentation* (Australia: Safety of Liposuction) and a published letter to the *Editor Medical Journal of Australia on Deep Vein Thrombosis*.