Scattering Transform for Matching Surgically Altered Face Images

Ishita Gupta[1], Ikshu Bhalla[1], Richa Singh[2], and Mayank Vatsa[2]
[1]Google Inc. USA and [2]IIIT-Delhi, India
Email: {ishitagupta45, ikshubhalla}@gmail.com, {rsingh, mayank}@iiitd.ac.in

***Abstract*—The use of face as a biometric feature has been widely accepted and used in security and surveillance systems. Recent studies have made significant advancements to address various challenges in face recognition such as illumination, age, pose and disguise. Another important covariate is recognizing faces with pre-and-post facial plastic surgery. Facial plastic surgeries alter the geometry and texture of facial regions, the extent of which is dependent on both the number, and the type of surgeries performed. The increasing reach of plastic surgery and its expanding user base present an indispensable challenge that must be dealt with while devising robust face recognition systems. In this paper, we present Invariant Scattering transform based feature extraction to compute translation invariant representation at local and global levels that is stable against plastic surgery variations. The identification accuracy achieved by the proposed algorithm is over 97% at rank-10 on the IIITD plastic surgery face database.**

I. Introduction

Daily Mail reported in March 2016 that one of Britain's most wanted criminals evaded law enforcement agencies for four years after fleeing the country on a false passport. He underwent plastic surgery to change his appearance and steer clear of police capture [1]. Daily Times reported in December 2014 that four terror suspects had undergone plastic surgery in an attempt to escape from Lahore International Airport [2]. The widely infamous White Widow terror suspect was reported to constantly change her appearance using plastic surgery to evade capture. As the Mirror reported in early 2016 [3], she is accused of 400 murders and is one of the world's most wanted terrorists who has not been apprehended. These incidences and previous research [4] demonstrate the misuse of plastic surgery by individuals to elude law enforcement agencies or conceal own one's identity. Even though plastic surgery is primarily being used for cosmetic and treatment purposes such as burns and tumors, it's misuse is also widely reported. As shown in Fig. 1, plastic surgery procedures can significantly alter facial biometric features, which may lead an automated system to misclassify before surgery and after surgery faces of an individual as two different subjects. Therefore, it is important to build a robust face recognition system that can recognize face images altered with plastic surgery.

The effect of variations in pose, illumination, age, and disguise have been considered while developing robust face recognition systems, however, plastic surgery has been perceived as a difficult challenge. Singh et al. [4], [6] and Bhatt et al. [12] analyzed the effects of plastic surgery on face recognition algorithms and showed that surgical alterations affect the performance of face recognition algorithms. They also collected the one and only publicly available face database with plastic surgery variations [4]. Thereafter, several algorithms have been proposed to address this covariate and a summary of existing literature is presented in Table I. Existing literature of face recognition with plastic surgery variations has focused on extracting handcrafted features and utilizing a classifier to identify the probe image. Although advancements in deep learning have seen significant improvement in face recognition performance, existing algorithms with facial plastic surgery as covariate have not seen much development. The primary reason is limited training data in the problem domain.

Fig. 1: Plastic surgery procedures may produce such drastic effects that may make it difficult for even humans to match post-surgery images (right) with pre-surgery ones (left) [5].

This paper proposes a face recognition algorithm using Scattering Network Transform [13] for facial feature extraction. Such an architecture avoids information loss as wavelet coefficients at a given layer recover the information lost in the preceding layer. Filters are not learned from data but are predefined wavelets. These wavelets can build invariance to translation, scaling and rotation by performing convolution along the respective variables. It is our assertion that the use of an algorithm using Scattering Transform can mitigate the problems described above efficiently. We further demonstrate that on the only publicly available database [4], the proposed algorithm yields very promising results in identification scenario and can provide an efficacious contribution to building robust face recognition systems, less affected by variations introduced due to plastic surgery procedures.

II. Proposed ScatNet based face recognition algorithm

The proposed algorithm is divided into four major steps: preprocessing, feature extraction and dimensionality reduction, matching, score fusion, and classification. Fig. 2 shows the steps involved in the proposed algorithm and the following subsections describe the individual steps in detail.

TABLE I: Summary of research directions related to facial plastic surgery using the IIITD facial plastic surgery database.

Reference	Algorithm	Maximum Rank-1 Accuracy
Singh et al. [6]	PCA, FDA, GF, LFA, LBP, GNN, Sum Rule Fusion with min-max normalization	Local Surgery: 49.2%, Global Surgery: 15%
Singh et al. [4]	PCA, FDA, LFA, Circular Local Binary Pattern (CLBP), Speeded Up Robust Features(SURF) GNN	53.7%
Jillela and Ross [7]	Combining information from face and ocular regions at score level	87.4% (removing non-detected and low resolution samples
Aggarwal et al. [8]	Part-wise and Sparse representation	77.9%
Bhatt et al. [9]	Multilevel non-disjoint face granules assimilated using a multiobjective genetic approach to optimize feature extractor from each granule, weighted χ^2 matching	87.32%
Marsico et al. [10]	Region based approaches PCA, FDA, LBP, FARO (FAce Recognition against Occlusions and Expression Variations) FACE (Face Analysis for Commercial Entities)	85.4% (disregarding non-detected samples)
Moeini et al. [11]	3D face reconstruction and sparse and collaborative representations	94.98% (External data for 3D model is used; therefore, direct comparison with other algorithms is not feasible)

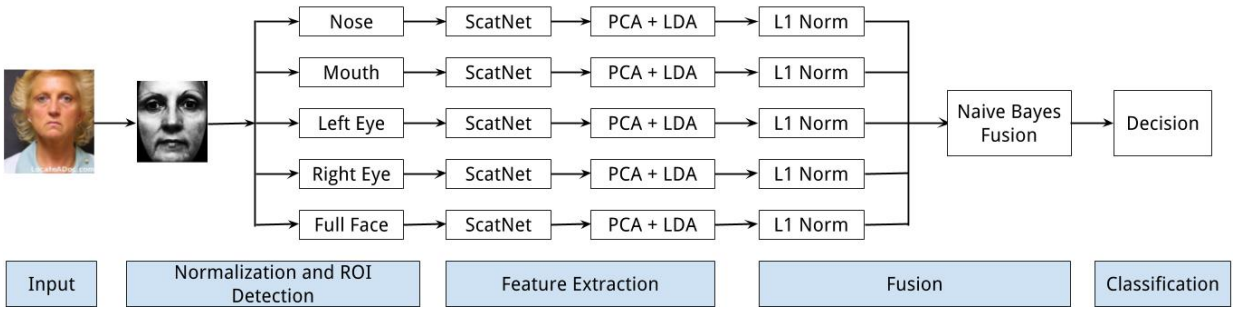

Fig. 2: Complete pipeline of the proposed ScatNet based face recognition algorithms.

A. Face Detection

The face images are detected and normalized using Haar face detector and region of interest is extracted. The ROI is resized to 196×224 pixels and converted to grayscale format. Using The CSU Face Identification Evaluation System, Version 5.1 [14], the images are geometrically normalized via two point eye-coordinates and face only region is obtained with the help of an elliptical mask. Histogram equalization is the final step to enhance the contrast in the images. The first row of Fig. 3 shows the final normalized image obtained after detection and pre-processing.

From the normalized images, four regions of interest (right eye, left eye, nose, and mouth) are detected using a bounding box method with golden ratio template [15]. Conventionally these regions are chosen as they are believed to be important for human physiognomy and have been shown to possess significant discriminatory facial information [9], [16]. In the feature extraction process, as shown in Fig. 3, we have utilized five components, i.e., complete face, right eye, left eye, nose, and mouth.

B. Feature Extraction using ScatNet

Bruna and Mallat [17] proposed Scattering Convolutional Networks as an invariant feature representation technique for generic texture images. In this research, we propose to use Scattering transform [13] for feature extraction from face images with pre-post-facial plastic surgery. We first briefly explain the theory of Scattering transform followed by the details of feature extraction and facial feature matching.

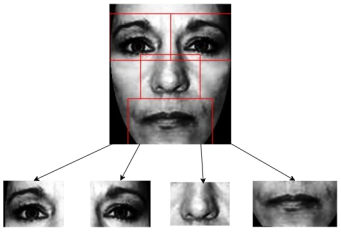

Fig. 3: Regions of interest extracted from a face image.

1) Scattering transform: Scattering transform or ScatNet [13] provides local image descriptors computed with a cascade of wavelet transform and modulus operators. It corresponds to a convolutional network where filter coefficients are given by a wavelet operator. As mentioned by Bruna and Mallat [17], in Scattering transforms, signal information is de-localized into scattering decomposition paths via non-linear transforms to build invariant, stable and informative representations. As shown in Fig. 4, Scattering transform scatters the signal information along multiple paths. Scattering transform generates coefficients encompassing varying degree of compromise between discrimination (the ability of being reactive to larger variations in the pattern) and invariance (immune to smaller variations). The representation is generated by concatenating these coefficients having the ability to produce an appropriate classification boundary.

Let $I(z)$ be an image in $\mathbb{R}^2$ and ϕ_J be an averaging window. This window carries the low frequencies of I above the scale

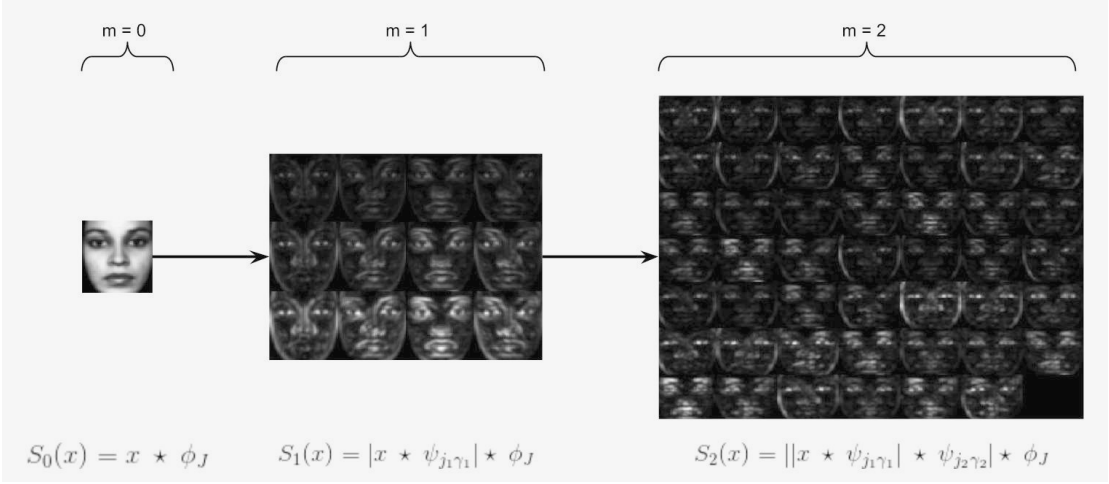

Fig. 4: Visualization of Scattering coefficients in layer by layer approach using an input face image x.

2^J (a Gaussian low-pass filter):

$$\phi_J(z) = 2^{-2J}\phi(2^{-J}z)$$

The 0^{th} layer of 2D Scattering transform of an image I is defined as,

$$S_0 I(z) = I \star \phi_J(z) \tag{1}$$

These coefficients are translation and deformation invariant to those deformations that are small relative to 2^J. The averaging with ϕ_J leads to loss of high-frequency information that can be recovered by convolution with wavelets while computing the coefficients for next layers.

Directional wavelet $\psi_{j,\gamma}$ are constructed by rotating a single wavelet ψ along the angles $\gamma \in \Gamma$, where Γ is a group of rotations, and scaled by 2^J resulting in

$$\psi_{j,\gamma} = 2^{-2j}\psi(2^{-j}R_\gamma z) \tag{2}$$

where, R is the planar rotation operator. The high-frequency information eliminated by Equation (1) can be recovered by convolution with the wavelet filter defined above. The first order scattering coefficients are obtained by taking the modulus of wavelet transform with $\psi_{j,\gamma}$ and then regularizing by averaging their amplitude with ϕ_J. The modulus removes the complex phase, reducing the variability as the coefficients are now independent of local translation. Further,

$$S_1 I(z) = |I \star \psi_{j_1,\gamma_1}| \star \phi_J(z) \tag{3}$$

The convolution with ϕ_J generates coefficients that are locally translation invariant but removes high frequencies. To restore high frequencies, wavelet coefficients at finer scale are regularized by averaging their amplitude with ϕ_J. This process is applied q times repeatedly to generate the q_{th} order scattering coefficients given by

$$S_q I(z) = \left(\left| \left| I \star \psi_{j_1,\gamma_1}| \star \cdots \right| \star \psi_{j_q,\gamma_q} \right| \star \phi_J(z) \right) \tag{4}$$

such that $\{j_1 < \cdots < j_q < J\}$ and $\{\gamma_1, \cdots, \gamma_q) \; \epsilon \; \Gamma\}$. Scattering transform applies decomposition of the the input signal iteratively using wavelet coefficients of different orientations and finer scales. It is generally recommended to take Scattering transform up to the second layer (m = 0, 1, 2). Scattering wavelet decomposition produces a set of coefficients at every layer and the visualizations of scattering transform coefficients at three different levels ($m = 0, 1, 2$) are shown in Fig. 4. This example helps visualize the output of each layer of ScatNet when applied on face images. It shows that applying ScatNet at different scattering orders (m) encode varying information; however, this also leads to increase in the size of feature vector and may also have redundancy in features.

2) ScatNet based feature extraction: In the proposed algorithm for face recognition with variations caused by plastic surgery, as shown in Fig. 2, the image is divided into four facial patches pertaining to that of the nose, mouth, left eye, and right eye. Thus, including the holistic (full) face image, there are in-total five facial regions. The samples of the four local patches for a subject are shown in Fig. 3. The ScatNet transform is applied with fine-tuned parameters such as filter-banks and associated parameters. Empirically, we observe that Morlet filter bank, which has close relationship to human perception in vision, yields the optimal results for this application. ScatNet transform involves four other parameters:

- M**:** For an input face x, this parameter defines the maximum scattering order, that is, the order up to which the scattering coefficients will be computed, i.e. $S_0 x, S_1 x, \cdots, S_M x$.
- L**:** This parameter tunes the number of wavelet orientations. The angular selectivity of the wavelet filters increases by increasing the value of L.
- J**:** This parameter is the value of scale in the filter bank $\psi_{j,\gamma}$ (Equation 2). Increasing the value of J increases the translational invariance.
- σ_ϕ**:** A 2D Morlet filter bank consists of a Gaussian window which tunes the spread of ϕ.:

$$\phi_J(u) = 2^{-2j/Q}\phi\left(2^{-j/Q(u,v)}\right)$$

We extract the ScatNet features for all five facial regions; the best set of parameters may vary for every facial region. Therefore, all four parameters are optimized using grid search over a range of values.

C. Facial feature matching

As described in Section II-B1, the size of ScatNet coefficients can be large and may possibly contain redundant information. Therefore, the next step involves reducing the feature dimensionality while preserving the unique and invariant ScatNet coefficients. Dimensionality reduction is achieved by first applying Principal Component Analysis (PCA) [18], and then Linear Discriminant Analysis (LDA) [19] on the coefficients of each of the local facial regions, independently [20]. From all the principal coefficients, set of coefficients that preserve a specific percentage of Eigen energy, E_i, are retained, i.e,

$$E_i = \sum_{j=1}^{i} \lambda_j / \sum_{j=1}^{n} \lambda_j \tag{5}$$

Here, $\lambda_1, \lambda_2, \cdots, \lambda_n$ is the set of Eigen values. On the reduced feature set obtained after PCA, we perform LDA to improve class separability by minimizing the intra-class variance and maximizing the inter-class variance. To match the gallery and probe features, L_1 distance is applied on the output of LDA from individual regions, i.e.

$$d(I_1, I_2) = \sum_{i=1}^{m} |I_{1-i} - I_{2-i}| \tag{6}$$

Here, I_1 and I_2 represent the coefficients obtained after applying LDA. For each gallery-probe image pair, L_1 norm is calculated independently for each of the components mentioned above. The L_1 distance obtained by matching individual components of the gallery and probe images are concatenated to get a score vector, which is further classified using Naive Bayes classification algorithm. The genuine and impostor scores pertaining to the images in the training database are modeled with two Gaussian distributions. For every test sample, posterior probability for each class is computed and a decision is taken depending on the class (genuine or impostor) with higher posterior probability.

III. Experimental Details

Database and Protocol: The plastic surgery face database [4], is a publicly available database consisting of 1800 real world pre-surgery and post-surgery images corresponding to 900 subjects. The effectiveness of proposed algorithm is shown on the original protocol with 10 times repeated cross-validation [4] and [9]. For each fold, the dataset is partitioned into train and test sets, each having 40% and 60% unseen subjects from the dataset, respectively. Images taken before plastic surgery are fixed as gallery whereas, images taken post plastic surgery form the probe. Identification accuracies at different ranks are used to report the results, along with cumulative match characteristic curves (CMC).

Algorithms used for comparison: The following algorithms and commercial systems are used for performance comparing with the proposed algorithm: (i) Three Patch Local Binary Patterns (TPLBP) [21], (ii) Four Patch Local Binary Pattern (FPLBP) [21], (iii) Verilook - a commercial off-the-shelf algorithm [22], (iv) CNN based commercial face recognition algorithm [23], and (v) Face reconstruction using 3D models [11]. Since the code of [11] is not available and the results are not reported on pre-defined protocol, the results of [11] are computed based on our implementation using standard protocol for the dataset and an existing 3D face database (USF Human-ID) is used to learn 3D model. Other than these, the results are also compared with reported results of Multi-objective evolutionary algorithm [9] and Region based approach [10].

IV. Experimental Results and Analysis

The result section is divided into two parts: we first present the experiments performed for optimizing the parameters involved in the proposed algorithm. In the second part, we focus on evaluating the effectiveness of each component of the proposed approach and comparison with state-of-the-art algorithms.

A. Parameter Optimization

As mentioned previously, ScatNet transform has four parameters ($M, L, J,$ and σ_ϕ) and it has been observed that for each facial region, different set of combinations provides optimal accuracies. In our experiments, total 144 parameter combinations, i.e. $3 \times 2 \times 4 \times 6$, are explored for each of the facial regions. Experimental results over different cross-validation folds demonstrate that with varying σ_ϕ, the best parameter values of ScatNet transform are **L** = 8, **M** = 3, and **J** = 5. The optimal values for individual facial regions are summarized in Table II.

These results correspond to the natural expectation because a higher value of **L** increases the angular selectivity of the wavelet filters. A higher value of **M** means more layers in the scattering transform, which avoids losing crucial information from the signal it operates on. Average pooling loses information, so a multilayer network structure is important as it recovers wavelet coefficients at the next layer. It should, however, be noted that increasing the value of **M** comes at a great computation cost and thus can not be set to an arbitrarily large value. The improvisation in classification results will be much less compared to the computational cost. A higher value of **J** makes the analysis of the signal robust to translational invariance.

TABLE II: Optimal values of the parameters involved in scattering transform for individual regions.

Region	σ_ϕ	L	J	M
Full Face	0.7	8	5	3
Left Eye	0.5	8	5	3
Right Eye	0.8	8	5	3
Nose	0.8	8	5	3
Mouth	0.8	8	5	3

TABLE III: Identification accuracies (%) for individual regions of interest of the proposed algorithm. These accuracies are the average of 10-fold cross validation.

Region of interest	Rank 1	Rank 10
Full Face	67.00	84.62
Nose	58.31	76.24
Left eye	54.44	74.77
Right eye	53.83	73.87
Mouth	52.03	71.07

Dimensionality of ScatNet features is quite high and to address this, we have utilized PCA followed by LDA for dimensionality reduction and improve class separability. To elucidate, let us consider the case of feature vector size of an image at $M = 3$, which is the highest order of scattering coefficients. The size of the feature vector generated by ScatNet is $2,000,376$. To determine the best value of Eigen energy to be retained, we compute the results by varying value of E_i to $0.95, 0.97, 0.98, 0.99,$ and 1.0. We experimentally observed that reducing the retained Eigen energy reduces the rank-1 accuracy of individual facial regions by up to 15%. Therefore, in order to maintain a trade-off between accuracy and feature dimensionality, the results are demonstrated with 99.9% Eigen energy at which the dimensionality is 352. It is interesting to note that this corresponds to 99.98% decrease in feature vector size, thereby providing a significant reduction in the computational time during matching.

B. Region Analysis and Comparison with Existing Algorithms

Both cosmetic and disease correcting plastic surgeries are generally local in nature. Local surgeries can be performed on individual facial features and in such cases, the remaining parts of the face appear to remain unchanged. We hypothesize that in such cases, it is difficult to match faces based on the altered regions. However, other facial regions can be efficiently and accurately used for face recognition. CMC curves in Fig. 5 summarize the results of the five local regions used in the proposed algorithm. Table III summarizes the rank-1 and rank-10 accuracies of the individual regions of interest. As expected, the best results are obtained by full face, which yields the rank-1 identification accuracy of 67% followed by nose region with 58.31% accuracy. The accuracy of both eye regions are lower than nose and full face, with mouth providing the lowest accuracies. Variations in mouth region are not only due to lip augmentation but it can also be significantly affected due to face lift and other global surgical procedures. Moreover, even though the database contains frontal face images, there are minor expression variations in some images which may affect the appearance of eyes and mouth regions. Therefore, another possible reason for high accuracy of nose region is that there are minimal variations possible in nose region with changes in expression. As explained in the algorithm section, the final classification output is obtained by combining the results of the five facial regions, which improves the accuracy significantly and addresses these variations.

TABLE IV: Comparing the identification accuracies (%) of the proposed and existing face recognition algorithms. The results of the algorithms marked with * are taken from the respective papers while the remaining results are computed by the authors.

Algorithms	Rank 1	Rank 5	Rank 10
TPLBP [21]	70.33	85.33	88.70
FPLBP [21]	69.31	83.59	86.77
Bhatt et al. [9]*	**87.32**	92.05	97.26
Marsico et al. [10]*	84.80	-	-
Moeini et al. [11]	60.41	68.53	73.32
Verilook	84.66	89.82	94.02
Face++	84.92	90.40	95.34
Proposed	85.43	**95.91**	**97.61**

Table IV summarizes the results of the proposed algorithm at three different ranks (1, 5, and 10) and compares the performance with several existing face recognition algorithms. Fig. 5 (Right Plot) illustrates the results in terms of CMC curves. The results show that the accuracies of TPLBP and FPLBP are significantly lower while the proposed algorithm yields highest rank-10 accuracy of 97.61%. These results showcase the efficacy of our algorithm for matching faces that have undergone plastic surgery, even without the need of large training database. The only training based components involved in the proposed algorithm are PCA and LDA. [4] computed the results of PCA and LDA on input images, and reported the rank-1 accuracy of 27.2% and 37.8% respectively. These values are 50-60% lower than the identification accuracies of the proposed algorithm - with the difference being ScatNet transform and local facial regions. From the results, it can be inferred that ScatNet transform is very effective in encoding texture information and dividing the image into local regions also helps to segregate the effect of plastic surgery. The standard deviation across different cross-validation trials is only 1.14% which showcases the stability of the algorithm with variations in training data.

Comparison with two state-of-the-art algorithms in face recognition designed for plastic surgery variations ([9] and [10]) are also documented in Table IV. While at rank-1, the proposed algorithm yields second best results, it outperforms [9] at all other ranks. Additionally, we have also performed experiments with CNN architecture based [23] as well as commercial matcher Verilook (both come with pre-trained models/APIs with very large training database) and the results show that the proposed algorithm outperforms these two face recognition approaches at all ranks. This experiment also showcases that, unlike other deep learning approaches, the proposed algorithm does not require large training data.

In literature, the best reported results are by [11]: Rank-1, 5, and 10 accuracies of 94.98%, 95.43% and 96.88%, respectively. However, it must be noted that the approach uses an external 3D dataset to train the 3D model while the predefined protocol of [4] recommends a fixed training data and therefore a direct comparison is not feasible. Further, the algorithm is based on *dictionary learning* paradigm where discriminative

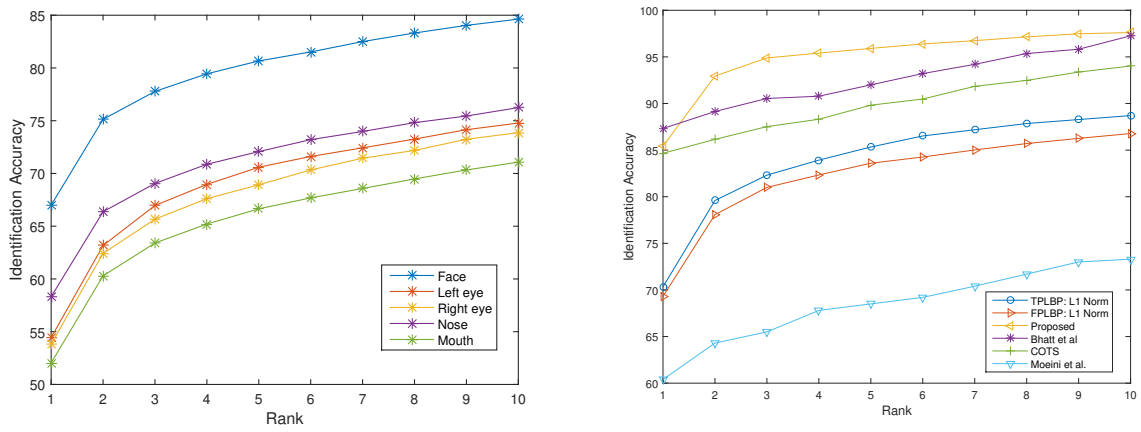

Fig. 5: Left: CMC curves for of the proposed algorithm with individual facial regions of interest. Right: CMCs for various popular algorithms on the Plastic Surgery Database. External data used in [11].

feature learning is proposed. Since the plastic surgery database has only two images per subject (one image as gallery and one as probe), it is challenging to learn discriminative features and encode both interclass and intraclass variations. Based on our understanding of discriminative dictionary learning and 3D face reconstruction (model generated using USF Human-ID database), we have implemented the algorithm proposed by [11] and achieve the Rank-1 accuracy of 60.4% only.

V. Conclusion

Plastic surgery has long been used for both cosmetic and disease corrective surgeries. Given the advancements and state-of-the-art of surgical procedures, this procedure can also be used for impersonating someone else's identity or evading one's own identity. With an ever increasing use of face recognition systems in practical world, such as for border control and law enforcement, this variant has the potential to become a major covariate in building reliable face recognition systems. This research presents a novel algorithm using an efficient ScatNet transform based feature extraction technique coupled with PCA and LDA to reduce dimensionality and increase separation between intra-class and inter-class variations. The proposed algorithm yields over 97% rank-10 identification accuracy on the IIITD plastic surgery face database. In future, we plan to extend the algorithm to address other covariates, including video based face recognition [24], [25].

References

[1] [Online]. Available: https://tinyurl.com/ychv2aq5

[2] [Online]. Available: https://tinyurl.com/ycpb3gjh

[3] [Online]. Available: https://tinyurl.com/y7gz8sqw

[4] R. Singh, M. Vatsa, H. S. Bhatt, S. Bharadwaj, A. Noore, and S. S. Nooreyezdan, "Plastic surgery: A new dimension to face recognition," *IEEE T-IFS*, vol. 5, no. 3, pp. 441–448, 2010.

[5] [Online]. Available: http://www.plasticsurgery.org/news/2015/plastic-surgery-statistics-show-new-consumer-trends.html

[6] R. Singh, M. Vatsa, and A. Noore, "Effect of plastic surgery on face recognition: A preliminary study," in *Proceedings of CVPR Workshops*, 2009, pp. 72–77.

[7] R. Jillela and A. Ross, "Mitigating effects of plastic surgery: Fusing face and ocular biometrics," in *Proceedings of IEEE BTAS*, 2012, pp. 402–411.

[8] G. Aggarwal, S. Biswas, P. J. Flynn, and K. W. Bowyer, "A sparse representation approach to face matching across plastic surgery," in *Proceedings of IEEE WACV*, 2012, pp. 113–119.

[9] H. S. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa, "Recognizing surgically altered face images using multiobjective evolutionary algorithm," *IEEE T-IFS*, vol. 8, no. 1, pp. 89–100, 2013.

[10] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Robust face recognition after plastic surgery using region-based approaches," *Pattern Recogn.*, vol. 48, no. 4, pp. 1261–1276, Apr. 2015.

[11] A. Moeini, K. Faez, H. Moeini, and A. M. Safai, "Open-set face recognition across look-alike faces in real-world scenarios," *Image and Vision Computing*, vol. 57, pp. 1–14, 2017.

[12] H. S. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa, "Face recognition and plastic surgery: social, ethical and engineering challenges," in *Ethics and Policy of Biometrics*. Springer, 2010, pp. 70–75.

[13] S. Mallat, "Group invariant scattering," *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.

[14] D. S. Bolme, J. R. Beveridge, M. Teixeira, and B. A. Draper, "The csu face identification evaluation system: its purpose, features, and structure," in *Proceedings of ICCVS*, 2003, pp. 304–313.

[15] K. Anderson and P. W. McOwan, "Robust real-time face tracker for cluttered environments," *CVIU*, vol. 95, no. 2, pp. 184–200, 2004.

[16] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Robust face recognition after plastic surgery using local region analysis," in *Proceedings of ICIAR*. Springer, 2011, pp. 191–200.

[17] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE T-PAMI*, vol. 35, no. 8, pp. 1872–1886, Aug 2013.

[18] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[19] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K.-R. Mullers, "Fisher discriminant analysis with kernels," in *Proceedings of IEEE SPS Workshop on NNSP IX*, 1999, pp. 41–48.

[20] W. Zhao, A. Krishnaswamy, R. Chellappa, D. L. Swets, and J. Weng, "Discriminant analysis of principal components for face recognition," in *Face Recognition*. Springer, 1998, pp. 73–85.

[21] L. Wolf, T. Hassner, and Y. Taigman, "Descriptor based methods in the wild," in *Real-Life Images Workshop at ECCV*, October 2008.

[22] Verilook. [Online]. Available: http://www.neurotechnology.com/verilook.html

[23] [Online]. Available: https://www.faceplusplus.com/

[24] G. Goswami, R. Bhardwaj, R. Singh, and M. Vatsa, "MDLFace: Memorability augmented deep learning for video face recognition," in *IEEE International Joint Conference on Biometrics*, 2014, pp. 1–7.

[25] G. Goswami, M. Vatsa, and R. Singh, "Face verification via learned representation on feature-rich video frames," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1686–1698, 2017.