BENCHMARKING 3D FACE DE-IDENTIFICATION WITH PRESERVING FACIAL ATTRIBUTES

Kevin H. M. Cheng, Zitong Yu, Haoyu Chen, Guoying Zhao[*]

Center for Machine Vision and Signal Analysis, University of Oulu, Finland

ABSTRACT

Privacy with the use of face images is becoming a major concern in civilians' applications. Recent studies have exploited privacy protection methods by means of facial attributes editing or de-identifying face images. Altering attributes causes loss of information for facial analysis while most de-identification studies did not quantitatively evaluate how well facial attributes are preserved. Moreover, state-of-the-art face analysis utilized 3D information for better performance. Existing face privacy studies only focusing in 2D domain is a key limitation towards the compatibility of more advanced 3D face analysis. This paper presents the *first* study on the possibility of 3D face de-identification with preserving facial attributes. We systematically evaluate the performance of 2D/3D face/facial attribute recognition and develop 2D/3D de-identification methods with preserving facial attributes using Auto Encoder and Generative Adversarial Networks approaches. We present comprehensive and reproducible experimental results using a publicly available 3D face database with facial attribute annotations for benchmarking and further research.
https://github.com/kevinhmcheng/3d-face-de-id

Index Terms— 3D face privacy, 3D face de-identification, 3D facial attributes preservation

1. INTRODUCTION

Human face images embed identity and many other personal information such as gender, age, ethnicity, emotion status, and health conditions [1]. When associating the personal information with the identity, privacy of individuals can be infringed, which can also result in discrimination and unfairness in societies [2]. The General Data Protection Regulation (GDPR) in European Union stated in recital 71 that "the controller should use appropriate mathematical or statistical procedures … prevent, inter alia, discriminatory effects on natural persons" [3], showing the importance of data processing for privacy protection.

To protect the privacy when using face images, recent studies have exploited two major approaches: facial attributes editing [4-7] and de-identification [8-13]. However, altering facial attributes results in the loss of original information (e.g., facial expressions) for facial analysis applications (e.g., psychological diagnosis). Meanwhile, most existing de-identification studies (e.g., [9-13]) only demonstrated the de-identification performance with producing realistic face images, but merely quantitatively evaluate whether the facial attributes in the original images are preserved. Those facial attributes are critical for facial analysis, which should be preserved while de-identifying face images. Moreover, state-of-the-art biometrics research [14-16] utilized 3D information for more accurate and robust face recognition. Challenging tasks like facial macro/micro-expression recognition [17-18] can also be better addressed in 3D domain. Existing face privacy studies only focusing in 2D domain is the key limitation towards the compatibility of more advanced face analysis using 3D information.

There are several related works in the 2D domain. These studies include face swapping [19], region of interest editing [20], template morphing [21], adding mosaic [22], while the adversarial training approach [8-13] has attracted the most attention. Apart from those focusing on face deidentification, other related research focuses more on facial attributes such as gender [6], age and ethnicity [7], and facial expression [8, 23-24]. The investigation of privacy preservation methods can also be extended to enhance the diversity of deidentified images [25]. However, none of the existing work has ever investigated privacy protection for 3D face data.

This paper presents the *first* study on the possibility of 3D face de-identification with preserving facial attributes. This challenging research spans across several topics including face recognition, facial attribute classification, privacy protection, multi-task and adversarial learning, 3D representation, information fusion, and so on. This paper systematically evaluates the performance of 2D/3D face/facial attribute recognition and develops 2D/3D face de-identification methods with preserving facial attributes using both Auto Encoder (AE) and Generative Adversarial Networks (GAN) approaches. Our technique enables de-identifying images with preserving facial attributes. The comprehensive and reproducible experimental results (with *implementation codes* [26]) presented in this paper provide a significant benchmark for much needed further research in this area. Open research problems are also discussed.

[*] corresponding author

2. FACE DE-IDENTIFICATION FRAMEWORK

Figure 1 shows the overall framework for 2D/3D face de-identification with preserving facial attributes. Firstly, we train the classification models for each classification tasks (Section 2.1). Secondly, we train the de-identification models with constraints to preserve facial attributes (Section 2.3). Finally, test images are processed by the de-identification models and evaluated with the classification models.

2.1. Multi-task Classification

Four classification tasks are addressed in this paper (biometric, expression, gender, ethnicity recognition). To develop de-identification methods with preserving facial attributes, it is important to first select the classifiers for each task so that we can fairly compare the classification performances between the original and de-identified images. Unlike multi-task learning, our goal is to select a classifier for each task as an independent evaluator for the de-identification performances. Therefore, we train the classifiers independently. With the insights from the literature on 3D face [14] and facial expression recognition [17], there is no such a model that works the best for all the tasks. Therefore, we considered multiple models: GoogleNet [27], ResNet50 [28], VGG16 [29], and DenseNet121 [30], which are the most popularly used deep learning models for general classification tasks. We will show in Section 3 that these models generally perform quite well for our classification tasks. During the training, we modify the last classification layer to accommodate the class numbers for each task. Except this layer, we fine-tuned the networks with the respective pre-trained weights to enable transfer learning for these tasks.

2.2. Use of 3D Information

This subsection describes the use of 3D information in addition to only using 2D images. It is an open research on how to utilize 3D information for various classification tasks. While this paper does not focus on advancing such technique but on providing reliable evaluation settings for de-identification, we employ conventional strategies to use 2D/3D information. With point cloud data in 3D, it is convenient to project these points on a plane to form depth images [14, 17]. Similarly, surface normal images also possess 3D information [31], which can be considered as a candidate input. Meanwhile, we also include several fusion strategies for fusing 2D/3D information, ranging from input, feature and score levels. For the input fusion, we include RGBD (depth image as the fourth channel), RGBxD (multiplying the depth with the RGB channels), and GDN (using grayscale, depth, and grayscale surface normal images as the three channels). For the feature fusion, we use the feature vectors resulted from 2D/depth data, and then train an additional layer for the classification. For the score level, we use the classification scores resulted from 2D and depth data, and then apply weighted sum between these scores.

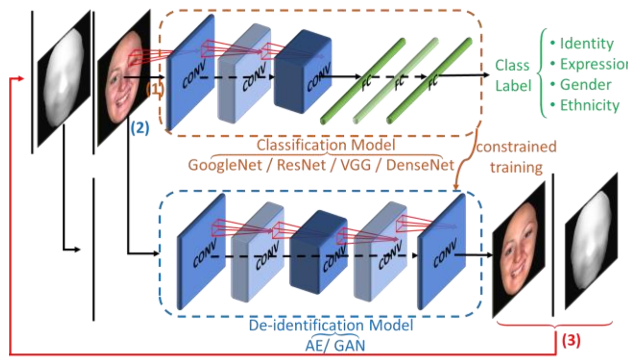

Figure 1. Framework for 2D/3D face de-identification with preserving facial attributes. The face images are extracted from BU3DFE database [33].

2.3. De-identification with Preserving Facial Attributes

Inspired by the research on privacy protection via altering facial attributes [6-7], AE/GAN can be employed to process images with additional constraints. Unlike references [6-7], we aim at de-identifying images with preserving facial attributes. Therefore, we do not alter facial attributes, but preserve them via constraint losses. The backbone of AE is adopted from the semi-adversarial networks [6] while that of GAN is adopted from StarGAN [32] with the use of pre-trained models. Note that the backbone of GAN in PrivacyNet [7] is also adopted from StarGAN [32].

Let $\boldsymbol{X}$ and $\boldsymbol{X}'$ be an input face image and the de-identified image, m, n, c be the dimension and channels, $\boldsymbol{y}_t$ and f_t be the label vector and the classifier for task t, g be the de-identification model such that $\boldsymbol{X}' = g(\boldsymbol{X})$. For the AE based approach, the reconstruction loss function J_{rec} for training the de-identification model g can be defined as:

$$J_{rec}(\boldsymbol{X}, \boldsymbol{X}') = \frac{1}{m \times n \times c} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{c} |X_{i,j,k} - X'_{i,j,k}| \quad (1)$$

We introduce a constraint loss to preserve facial attributes:

$$J_{cons}^{t}(\boldsymbol{X}', f_t, \boldsymbol{y}_t) = -\sum_{l=1}^{L_t} y_t^l \log(f_t(\boldsymbol{X}')^l) \quad (2)$$

where L_t is the number of classes for task t. Note that the classifier f_t can be obtained from the best model for each task (section 2.1). We can preserve a specific facial attribute for task t with the following loss function J_1:

$$J_1 = \lambda_{rec} J_{rec} + \lambda_{cons} J_{cons}^{t} \quad (3)$$

where λ_{rec} and λ_{cons} are the weights for J_{rec} and J_{cons}^{t}. We can also preserve N multiple facial attributes (e.g., facial expressions, gender, and ethnicity) by loss function J_2:

$$J_2 = \lambda_{rec} J_{rec} + \lambda_{cons} \sum_{t=1}^{N} J_{cons}^{t} \quad (4)$$

For simplicity, the same weight λ_{cons} can be applied to all J_{cons}^{t}. We can also de-identify an image with preserving N multiple facial attributes by loss function J_{de}:

$$J_{de} = \lambda_{rec} J_{rec} - \lambda_{id} J_{cons}^{id} + \lambda_{cons} \sum_{t=1}^{N} J_{cons}^{t} \quad (5)$$

where λ_{id} enables the control of the de-identification degree.

For the GAN based approach, the training of the discriminator is almost the same as StarGAN except the classification loss is disabled because classification models have been pre-trained (section 2.1). When training the generator, similar to the AE based approach, we disable the options to alter facial attributes but to introduce the constraint loss function (Equation 2). The adversarial loss function J_{adv} for training the generator g can be defined as follows:

$$J_{adv} = \frac{1}{N_d} \sum_{i=1}^{N_d} \log(1 - D(X')_i) \quad (6)$$

where D is the discriminator network and N_d is the number of scalar values for the output of D. The final loss function J_{de} for training the generator g can be defined as follows:

$$J_{de} = \lambda_{adv} J_{adv} + \lambda_{rec} J_{rec} - \lambda_{id} J_{cons}^{id} + \lambda_{cons} \sum_{t=1}^{N} J_{cons}^{t} \quad (7)$$

In this way, it is possible to de-identify an image with preserving facial attributes into some extent.

3. EXPERIMENTS AND RESULTS

Dataset and Evaluation Protocol. Among existing 3D face databases, BU3DFE database [33] provides annotations of identity, expressions, gender, and ethnicity. It is widely used for 3D facial analysis (e.g. [14, 17]) and is highly suitable for this paper. This database provides 2500 2D/3D face images/point cloud data acquired from 100 subjects. There are six expressions per subject, four intensity levels per expression, and a neutral sample. As a research frontier on the new problem addressed in this paper, it is reasonable to first consider a closed set biometric recognition scenario [34] so that same identity and facial attributes classes are seen during the training. We define the training set as all the 100 subjects and six expressions, with intensity level 1 to 3; the test set as all the 100 subjects and six expressions, but with intensity level 4, i.e., 1800 images for the training and 600 images for the testing. This evaluation protocol enables fair evaluation on the performance of biometric recognition and facial attributes classification, thus the performance of de-identification with preserved facial attributes. For the evaluation metrics, we present identification accuracies (ACC), equal error rates (EER). Due to space limitations, we only present receiver operating characteristics (ROC) curves for the final results using both 2D and depth information. The implementation details and parameters such as λ_{id} can be found in the implementation codes [26].

Comparative Performance Evaluation. Table 1 shows the experimental results of four tasks (biometric, expression, gender, ethnicity recognition) using four models (GoogleNet, ResNet50, VGG16, DenseNet121), different input data and fusion strategies (input level, feature level, score level), which justifies the choices of the classification models (section 2.1) and the use of 3D data (section 2.2). The first basic finding can be referred to the first row where 2D texture images are the inputs. The recognition performance of all four tasks is quite high, indicating that this experimental setting is effective for the evaluation of de-identification because there is plenty of room to de-identify the images (lowering the biometric recognition performance) while preserving other attributes (maintaining other recognition performance). Comparing the performance of different classification models, we can select the best model for each input and task. For example, for biometric recognition, ResNet is selected when 2D data is used while DenseNet is selected when depth data is used. Comparing the first three rows, we find that using 2D images performs better than depth images which performs better than using surface normal images. Therefore, 2D and depth images with their respective best models are selected for feature and score level fusion. Comparing the last five rows regarding the fusion strategies, we find that score level fusion performs the best.

Table 2 shows the experimental results of the de-identification approaches (Section 2.3) with different constraint losses. The row “Original” refers to the baseline classification performance when the images are not processed; the rows between “Expression” to “Ethnicity” refer to constraining each attribute using Equation 3 so that such attributes can be better preserved; the row “Exp.+Gen.+Eth.” refers to constraining multiple attributes together using Equation 4; the row “Ours” refers to de-identifying images with preserving facial attributes using Equation 5 or 7. The weight λ_{id} controls the strength to de-identify the images, which is selected by achieving the accuracies of biometric recognition to be 1%.

We also compare with two existing de-identification methods: VGAN [8] aims at face de-identification with preserving facial expressions while CIAGAN [9] aims at de-identification with preserving visual quality, which are of similar objectives to the problem addressed in this paper. We implement these methods based on the partially available source codes and try our best to improve the performance by adjusting the hyper-parameters. Same dataset and evaluation protocol are used in the evaluation. For the facial expressions preservation in VGAN, similar to our AE/GAN approach, we replace the classification loss with our pre-trained classification models. It can be observed from the last two rows in Table 2 that these two methods are effective for de-identification, but do not perform well in preserving facial attributes.

Figure 2 shows the experimental results of de-identification performance using both 2D and 3D information. These results are obtained by using score fusion between 2D and depth images. It can be observed that using the AE/GAN based approach can enable strong de-identification (heavy drops in the biometric recognition performance), while maintaining the recognition performance of other attribute classification tasks to some extent. The GAN based approach generally outperforms the AE approach in terms of maintaining better recognition performance of the attribute recognition tasks.

4. CONCLUSIONS

This paper presents the *first* study on the possibility of 3D face de-identification with preserving facial attributes.

Incorporating constraint loss functions in AE/GAN based approach is effective to de-identify both 2D/3D data with preserving facial attributes. The current approach relies on specific classification models for de-identification, other classification models may still be able to classify the original identity in some cases. It is also important to de-identify face images with stronger preservation of facial attributes. Furthermore, how to represent 3D face data for biometric and attribute recognition is an open problem to be addressed. For instance, PointNet++ [35] can be an interesting alternative instead of projecting the 3D point clouds as depth maps.

5. ACKNOWLEDGEMENTS

This work was supported by the Academy of Finland for Academy Professor project EmotionAI (grants 336116, 345122) and Infotech Oulu, as well as the CSC-IT Center for Science, Finland, for computational resources.

Table 1. Experimental results of tasks using different classification models, input data and fusion strategies. The numerical results are ACC (in %) and EER (in % and in brackets). Finally selected options are in bold.

Task	Biometric Recognition				Expression Recognition				Gender Recognition				Ethnicity Recognition			
Input\Model	Gnet	Rnet	Vnet	Dnet	Gnet	Rnet	Vnet	Dnet	Gnet	Rnet	Vnet	Dnet	Gnet	Rnet	Vnet	Dnet
2D	99.7 (0.2)	**100 (0.0)**	99.2 (1.4)	99.8 (0.2)	90.0 (5.1)	**94.0 (4.1)**	89.8 (6.1)	93.3 (4.5)	99.5 (0.5)	**100 (0.0)**	99.3 (0.7)	99.8 (0.2)	99.7 (0.2)	**100 (0.0)**	99.8 (0.8)	99.8 (0.2)
Depth	98.3 (1.2)	99.3 (0.3)	88.3 (3.9)	**99.3 (0.2)**	88.7 (5.2)	91.2 (4.7)	88.3 (8.2)	**92.7 (4.8)**	99.2 (0.8)	**99.3 (0.7)**	99.0 (1.0)	99.3 (0.7)	98.5 (1.0)	**99.5 (0.3)**	99.2 (1.4)	99.3 (0.5)
Normal	95.5 (2.0)	98.2 (0.8)	92.8 (3.5)	98.8 (0.8)	88.2 (6.3)	91.2 (5.0)	89.7 (6.2)	91.7 (4.8)	98.0 (2.0)	99.0 (1.0)	97.7 (2.3)	98.7 (1.5)	98.0 (1.2)	98.5 (0.7)	99.0 (0.7)	99.2 (1.2)
RGBD (Input Fusion)	98.5 (0.8)	99.3 (0.5)	93.2 (3.6)	99.7 (0.3)	88.8 (6.1)	90.2 (5.6)	88.8 (8.4)	92.0 (5.1)	99.5 (0.5)	99.5 (0.5)	99.3 (0.7)	100 (0.0)	99.3 (0.3)	99.8 (0.2)	99.3 (1.2)	99.7 (0.2)
RGBxD (Input Fusion)	99.8 (0.3)	99.7 (0.2)	97.2 (1.9)	99.7 (0.3)	90.3 (5.6)	92.0 (5.0)	92.0 (5.8)	91.8 (4.4)	99.7 (0.3)	100 (0.0)	99.7 (0.3)	99.8 (0.3)	99.7 (0.4)	99.8 (0.0)	99.8 (0.5)	99.8 (0.3)
GDN (Input Fusion)	98.3 (1.0)	99.3 (0.8)	95.3 (2.3)	98.3 (0.6)	90.2 (5.0)	91.2 (5.3)	90.5 (6.7)	91.7 (5.1)	99.7 (0.3)	99.3 (0.7)	99.3 (0.7)	99.7 (0.3)	98.8 (0.7)	99.7 (0.5)	99.7 (0.8)	99.5 (0.5)
Approach\Model	2D(Rnet)+Depth(Dnet)				2D(Rnet)+Depth(Dnet)				2D(Rnet)+Depth(Rnet)				2D(Rnet)+Depth(Rnet)			
Feature Fusion	100 (0.1)				92.7 (3.3)				100 (0.0)				99.8 (0.3)			
Score Fusion	100 (0.0)				95.3 (3.5)				100 (0.0)				100.0 (0.0)			

Table 2. Experimental results of the de-identification methods with different constraint objectives. The numerical results are ACC (in %) and EER (in % and in brackets). Finally selected options are in bold.

Task	Biometric Recognition				Expression Recognition				Gender Recognition				Ethnicity Recognition			
Input	2D		Depth		2D		Depth		2D		Depth		2D		Depth	
Original	100 (0.0)		99.3 (0.2)		94.0 (4.1)		92.7 (4.8)		100 (0.0)		99.3 (0.7)		100 (0.0)		99.5 (0.3)	
De-id Model Constraints \	AE	GAN	AE	GAN	AE	GAN	AE	GAN	AE	GAN	AE	GAN	AE	GAN	AE	GAN
Expression	53.3 (10.8)	73.8 (5.2)	95.8 (1.4)	91.3 (2.0)	91.3 (6.2)	89.3 (6.0)	90.8 (5.0)	91.0 (5.5)	96.0 (4.1)	98.8 (1.2)	98.8 (1.2)	98.3 (1.7)	92.0 (7.0)	97.0 (1.5)	99.0 (1.0)	97.7 (1.8)
Gender	63.3 (7.7)	83.7 (3.3)	91.5 (2.2)	78.5 (5.3)	78.0 (11.3)	86.7 (8.0)	91.0 (5.4)	82.7 (9.3)	99.8 (0.2)	99.7 (0.3)	99.0 (1.0)	97.8 (2.2)	94.3 (5.0)	97.8 (1.8)	98.0 (1.5)	95.7 (3.1)
Ethnicity	64.2 (6.7)	83.5 (4.5)	92.7 (1.7)	93.2 (1.8)	83.8 (9.8)	86.3 (7.8)	91.2 (5.3)	87.8 (6.9)	98.2 (1.8)	98.3 (1.7)	98.7 (1.3)	98.3 (1.7)	99.3 (0.8)	99.7 (0.5)	99.2 (0.5)	99.0 (1.0)
Exp. + Gen. + Eth.	59.5 (7.8)	84.0 (2.8)	97.2 (1.2)	95.5 (1.3)	87.3 (6.8)	90.2 (5.5)	90.0 (5.1)	89.2 (6.3)	98.3 (1.7)	99.7 (0.3)	99.3 (0.7)	99.0 (1.0)	97.3 (1.8)	99.8 (0.5)	99.7 (0.7)	99.5 (0.7)
Ours	1.0 (37.6)	1.0 (45.4)	1.0 (37.8)	1.0 (41.4)	86.0 (8.0)	86.5 (8.3)	85.8 (8.7)	86.7 (7.7)	97.7 (2.3)	98.8 (1.1)	94.7 (5.2)	96.7 (3.3)	93.8 (4.6)	96.7 (3.5)	93.0 (5.1)	94.7 (3.7)
VGAN [8]	3.7 (43.4)		1.7 (46.6)		53.5 (25.3)		66.5 (20.0)		49.5 (50.5)		55.0 (45.0)		46.0 (30.8)		41.2 (32.4)	
CIAGAN [9]	4.0 (38.3)		1.0 (47.5)		25.8 (41.7)		18.2 (49.5)		72.8 (27.1)		46.3 (53.0)		40.5 (25.8)		41.8 (31.4)	

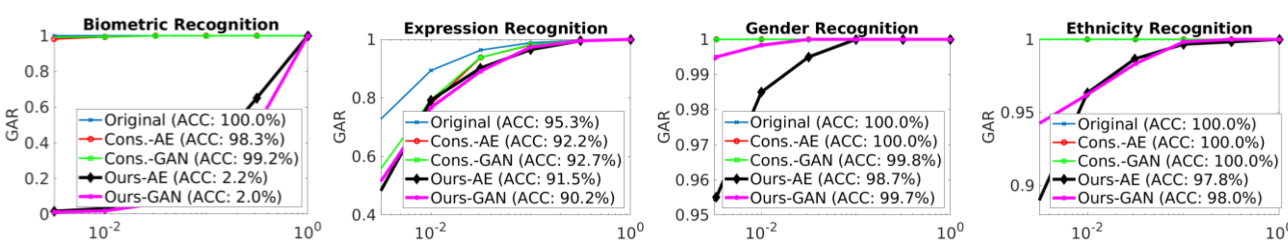

Figure 2. Experimental results of de-identification performance using both 2D and depth information. “Cons.” refers to preserving the three attributes by Equation 4. “Ours” refers to de-identification with preserving facial attributes by Equation 5 or 7.

6. REFERENCES

[1] X. Zheng, Y. Guo, H. Huang, Y. Li and R. He. A survey of deep facial attribute analysis. *IJCV*, 128(8), pp.2002-2034, 2020.

[2] A. Morales, J. Fierrez, R. Vera-Rodriguez and R. Tolosana. SensitiveNets: Learning agnostic representations with application to face images. *IEEE TPAMI*, 43(6), 2158-64, 2020.

[3] The Regulation (EU) 2016/679 (General Data Protection Regulation). *Available online*: https://gdpr-info.eu/

[4] G. Yang, N. Fei, M. Ding, G. Liu, Z. Lu and T. Xiang. L2M-GAN: Learning To Manipulate Latent Space Semantics for Facial Attribute Editing. In *IEEE (CVPR)*, 2021.

[5] Z. He, W. Zuo, M. Kan, S. Shan and X. Chen. Attgan: Facial attribute editing by only changing what you want. *IEEE Transactions on Image Processing*, 28(11), 5464-5478, 2019.

[6] V. Mirjalili, S. Raschka, A. Namboodiri and A. Ross. Semi-adversarial networks: Convolutional autoencoders for imparting privacy to face images. In *ICB*, 2018.

[7] V. Mirjalili, S. Raschka and A. Ross. PrivacyNet: semi-adversarial networks for multi-attribute face privacy. *IEEE Transactions on Image Processing*, 29, pp.9400-9412, 2020.

[8] J. Chen, J. Konrad and P. Ishwar. Vgan-based image representation learning for privacy-preserving facial expression recognition. In *CVPRW*, 2018.

[9] M. Maximov, I. Elezi and L. Leal-Taixé. Ciagan: Conditional identity anonymization generative adversarial networks. In *IEEE CVPR*, 2020.

[10] E. Chatzikyriakidis, C. Papaioannidis and I. Pitas. Adversarial face de-identification. *IEEE ICIP*, 2019.

[11] T. Li and L. Lin. Anonymousnet: Natural face de-identification with measurable privacy. In *IEEE CVPRW*, 2019.

[12] H. Hukkelås, R. Mester and F. Lindseth. Deepprivacy: A generative adversarial network for face anonymization. In *International Symposium on Visual Computing*, 2019.

[13] X. Yang, Y. Dong, T. Pang, H. Su, J. Zhu, Y. Chen and H. Xue. Towards face encryption by generating adversarial identity masks. In *IEEE ICCV*, 2021.

[14] S.Z. Gilani and A. Mian. Learning from millions of 3D scans for large-scale 3D face recognition. *IEEE CVPR*, 2018.

[15] K.H.M. Cheng, A. Kumar. Contactless biometric identification using 3D finger knuckle patterns. *IEEE TPAMI*, 2020.

[16] I.I. Ganapathi and S. Prakash. 3D ear recognition using global and local features. *IET Biometrics*, 7(3), pp.232-241, 2018.

[17] M. Behzad, N. Vo, X. Li and G. Zhao. Towards reading beyond faces for sparsity-aware 3d/4d affect recognition. *Neurocomputing*, 458, pp.297-307. 2021.

[18] Y. Li, X. Huang and G. Zhao. Micro-expression action unit detection with spatial and channel attention. *Neurocomputing*, 436, pp.221-231, 2021.

[19] S. Mahajan, L.J. Chen and T.C. Tsai. Swapitup: A face swap application for privacy protection. *IEEE AINA*, 2017.

[20] F. Peng, G.L. Ping and S.K. Ge. A face privacy protection scheme using cnn based roi editing. In *HPCC/SC/DSS*, 2019.

[21] M. Long, S. Long, G.L. Ping and F. Peng. A Facial Privacy Protection Framework Based on Component Difference and Template Morphing. *Int. ICCCN*, 2020.

[22] Z. You, S. Li, Z. Qian and X. Zhang. Reversible Privacy-Preserving Recognition. In *IEEE ICME*, 2021.

[23] W. Zhang, X. Ji, K. Chen, Y. Ding and C. Fan. Learning a Facial Expression Embedding Disentangled From Identity. In *IEEE CVPR*, 2021.

[24] J.R. Chang, Y.S. Chen, W.C. Chiu. Learning Facial Representations from the Cycle-consistency of Face. *ICCV*, 2021.

[25] S. Guo, S. Feng, Y. Li, S. An and H. Dong. Integrating Diversity into Neural-Network-Based Face Deidentification. In *CCC*, 2018.

[26] Weblink for Downloading the Implementation Codes: https://github.com/kevinhmcheng/3d-face-de-id

[27] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich. Going deeper with convolutions. In *IEEE CVPR*, 2015.

[28] K. He, X. Zhang, S. Ren and J. Sun. Deep Residual Learning for Image Recognition. In *IEEE CVPR*, 2016.

[29] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *ICLR*, 2015.

[30] G. Huang, Z. Liu, L. Van Der Maaten and K.Q. Weinberger. Densely Connected Convolutional Networks. In *CVPR*, 2017.

[31] K.H.M. Cheng and A. Kumar. Distinctive Feature Representation for Contactless 3D Hand Biometrics using Surface Normal Directions. In *IJCB*, 2020.

[32] Y. Choi, M. Choi, M. Kim, J.W. Ha, S. Kim and J. Choo. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *IEEE CVPR*, 2018.

[33] L. Yin, X. Wei, Y. Sun, J. Wang and M.J. Rosato. A 3D facial expression database for facial behavior research. In FG, 2006.

[34] K.H.M. Cheng, A. Kumar. Deep Feature Collaboration for Challenging 3D Finger Knuckle Identification. *IEEE TIFS*, 2021.

[35] R. Charles, Y. Li, S. Hao and J.P. Leonidas. Deep hierarchical feature learning on point sets in a metric space. In *NIPS*, 2017.