A. Brömme, N. Damer, M. Gomez-Barrero, K. Raja, C. Rathgeb,
A. Sequeira, M. Todisco, and A. Uhl (Eds.): BIOSIG 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022

Protecting the Privacy of Face by De-Identification Pipeline Based on Deep Learning

Anubha Parashar[1] Apoorva Parashar[2] Imad Rida[3] Vidyadhar Aski[4]

Abstract: This paper proposes a reversible face de-identification pipeline that modifies face geometry and texture. Fourteen parameters for geometrical modification are used. For texture modification fixed face texture template is used. We have investigated the impact of various geometrical and surface alterations of face components like eyes, eyebrows, nose, and lips on the ability of humans and machines to recognize faces. The crowdsourcing and machine face recognition experiments were performed on images of famous people collected from the Internet. The obtained results in both types of experiments showed that face texture has a stronger impact on a level of privacy protection than face geometry (shape) modifications.

Keywords: Face de-identification, Face recognition, Affine transformation, Face geometry modification, Face texture modification, Psychological experiments

1 Introduction

One of the main physiological biometric identifies which is used for person identification is face [1]. Face-based identification is used in various application scenarios - from identification of a person based on a still image in a passport or an identity card, to identification based on face images captured by surveillance systems in public and semi public places, without the cooperation of a person and covert person recognition [2]. Obviously, in the situations of covert person recognition, there is need for face de identification for privacy protection. Face de-identification is performed in a de identification pipeline which consists of the following stages [2]: face detection and localization, facial feature points localization, and face de-identification by masking, face swapping [3], wrapping, morphing or by scrambling. A perquisite for an efficient face de-identification is face detection. Face detection in unconstrained conditions is very difficult task and presence of false negative and false positive face detections is common

Privacy protection of false negative (i.e. undetected) faces is impossible. Note that, if there is even only one undetected face in a single frame of a video sequence, the privacy is compromised. Other significant problem in face de-identification are false positive face

[1] Faculty of Engineering, Computer Science and Engineering, Manipal University Jaipur, Jaipur, India, anubhaparashar1025@gmail.com
[2] Consultant, Emerging technology, Mahindra Integrated Business Solutions, Mumbai, India, apoorvaparashar0000@gmail.com
[3] Laboratoire Biomcanique et Bioingnierie UMR 7338, Centre de Recherches de Royallieu, Université de Technologie de Compiègne, Compiègne, France, imad.rida@utc.fr
[4] Manager, Emerging technologies at Mahindra and Mahindra, Mumbai, India, vidyadharstjit@gmail.com

detections. They do not have any impact on privacy protection, but they impede on the naturalness of de-identified images. To achieve naturalness of de-identified faces the common approach is to detect the facial feature points (facial landmarks) which would be the starting points for de-identification based on morphing, face swapping and/or their combination.

In general, face-de identification methods can be classified as naive and complex [3,2]. The naive methods are very fast and have simple implementation. They require that only an approximate bounding box of a face is defined. Utility and naturalness of de identified data are low (i.e. ability to recognize gender, expression, race, age, action, facial features). The methods are irreversible, but privacy can be easily compromised [2]. False positive face detections degrade naturalness of the de-identified data because the original texture is replaced with modified one The complex methods have following characteristics slow, complex implementation, they require that many face components (i.e. eyes, eyebrows, nose, lips) are precisely localized, have very large number of degrees of freedom, their utility and naturalness are high, they are also irreversible, privacy protection is high, and false positive face detections also degrade naturalness because a texture is modified [3], geometry of face components are changed.

The main contributions of the paper are:

1. Reversible face de-identification pipeline that combines the good properties of naive and complex de-identification methods;
2. Results of five psychological experiments that evaluate the impact of face texture and/or face geometry modifications to level of privacy protection;
3. Results of five experiments performed by automatic face recognition of de-identified faces,
4. Comparison of obtained results for both human and automatic face recognition.

2 Face de-identification pipeline

The proposed face de-identification pipeline has four stages: face detection, facial feature points localization, face region decomposition, and face de-identification with modification of face geometry and texture. Faces present in videos or still images are detected at the first stage. Facial feature points are obtained at the second stage. Face region defined by facial feature points is decomposed with Delaunay triangulation at the third stage. Reversible face image de-identification based on piecewise affine transformation of Delaunay triangles and texture modification is performed at the fourth stage. The first three stages are briefly described in this section, while the fourth stage is described in detail in Section 3. One of the main goals of this preliminary research is to evaluate the impacts of face geometry and/or texture modifications on a level of privacy protection when recognition is performed by humans and/or machine. The results of both types of experiments are given in Section 4.

At the first stage (Face detection) we use of-the-shelf NPD face detector [4]. It uses normalized pixel difference features (that are scale invariant, bounded, and able to reconstruct the original image) and deep quadratic trees to learn the optimal subset of features and their combinations, so manifolds. The detector is suitable for fast and accurate unconstrained face detection of faces with arbitrary pose variations and occlusions. It has both low false negative and relatively low false positive face detections owing to the advanced boosting-based method.

Facial feature points are localized by the fast method proposed in [5] at the second stage. It uses an ensemble of regression trees to estimate 68 facial landmark positions from a sparse subset of pixel intensities. The ensemble of regression trees that optimizes the sum of square error loss is learned by gradient boosting The initial assumption about face pose is that both eyes are visible, and a face is in approximately upright position. For faces that are in full profile pose and for false positive detections, the method always returns 68 feature points even though all feature points are not present or visible.

At the third stage the locations of the face feature points localized at the second stage are used to generate constrained Delaunay triangulation [6] that groups certain required segments into the triangulation. These segments are determined by indexes of facial feature points that are positioned on the border between two different face components. The obtained results are sets of Delaunay's triangles that represent face components: eyes, eyebrows, nose, lips, forehead and lower face. This face region partitioning enables both piecewise affine transformation used in the process of de-identification and to preserve naturalness of de-identified face.

3 Fourth stage - face de-identification

The input in the fourth stage is constrained Delaunay triangulation of the face region. The following face components are subject to modification eyes, eyebrows, nose, and lips.

3.1 Face component modification parameters

For each face component a set of modifications is defined taking in consideration morphological characteristics of a human face is very difficult task to define what face components characteristics constitute the core concept that define face identity Face identity is subjective by nature for humans. Only certain combinations of face components modifications can be performed, if we want to preserve naturalness of de-identified faces.

To preserve naturalness of a face, the very first step is to determine following global morphological characteristics of a face distance between eyes, head width and height, and nose tip position. Based on these global morphological characteristics, we introduced a range of allowable face modification parameters expressed in the interval from -1 to 1, of each face modification parameter is (pseudo) randomly where 0 defines original visual appearance, and -1 and 1 maximum allowed change in opposite directions. The value selected

from the interval [-1,1]. For example, eye size modification parameter value from the interval [-1,1], is linearly mapped into the value $modified\_eye\_size$ from the eye size interval $[minsye\_size, max\_eye\_size]$ which is calculated as follows

$$[min_eye_size, max_eye_size] = [\frac{distance\_between\_eyes}{K1} + \frac{head\_width}{K2}] \quad (1)$$

where $distance\_between\_eyes$ and $head\_width$ are determined for each face that is a subject of de-identification. The parameters k_1 and k_2 are obtained based on analyses of near frontal face database. The corresponding affine parameter of a scale is obtained as $modifiedeye\_size/originaleve\_size$, where $original\_eye\_size$ is defined as a distance between leftmost and rightmost landmarks od an eye. Note that affine parameters of translation and rotation, respectively, directly correspond to the values obtained by linear mapping of face modification parameters.

A heuristic formula for defining range of allowed modification is determined for each face component modification parameter. These formulas are determined based on face component characteristics obtained from distributions of their values in a database of near profile faces. The face component modifications with corresponding parameters are listed in Tab. 1. The symbols in Tab. 1 have the following meaning: $s_i^{x/y}$ - scale, $t_i^{x/y}$ - translation and α_i - rotation, and index i corresponds to the face component, while super index x or y denotes axis directions. For example, for an eye the typical values are $s_i^{x/y}$ $[0.85, 1.15]$, $t_i^{x/y}$ $[-15, 15]$ pixels, and $\alpha \in [-0.10, 0.15]$ radians.

Tab. 1: Face component modifications

	Face components			
Modifications	**Eyes**	**Eyebrows**	**Nose**	**Lips**
Size horizontal	s_1^x	–	s_3^x	s_4^x
Size vertical	s_1^y	–	s_3^y	s_4^y
Position horizontal	t_1^x	t_2^x	–	–
Position vertical	t_1^y	t_2^y	t_3^y	t_4^y
Rotation	α_1	α_2	–	–

3.2 Face de-identification - affine transformations of face components

For each face component all corresponding modification parameters (see Tab. 1) are used to determine affine transformation. The affine transformation is then used on all Delaunay triangles belonging to corresponding face component. Note that the parameters of affine modification used for left and right eye, as well as for both eyebrows must preserve symmetrical appearance of a face. The following steps are performed in the process of transforming face components:

STEP 1: For each face component select a corresponding set of face feature points $FFP_i = \{(x_j, y_j)\}_i$, where $i \in (1, ..., 6)$ is index of a face component, and $j \in (1, ..., 68)$ is an index of face feature points, $\{\quad\}_i$ denotes a set of face feature points that belong to a face component i.

Determine a center of face feature points for a face component i, denoted as (x_i^c, y_i^c). The (x_i^c, y_i^c) is used to define the displacement of the face component center from an image origin.

STEP 2: Use all face modification parameters that correspond to face component *i* to determine parameters of the affine transformation: scale s_i^x and s_i^y, rotation α_i, and translations t_i^x and t_i^y.

STEP 3: Transform original vertices $\{FFP_i = (x_j, y_j)\}_i$, that overly face component i, by using affine transformation into new vertices of de-identified face component $FFP_i^* = \{(x_j^*, y_j^*)\}_i$.

$$\begin{bmatrix} x_j^* \\ y_j^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_i^x \\ 0 & 1 & t_i^y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} cos(\alpha_i) & sin(\alpha_i) & 0 \\ -sin(\alpha_i) & cos(\alpha_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_i^x & 0 & 0 \\ 0 & s_i^y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_j - x_i^c \\ y_j - y_i^c \\ 1 \end{bmatrix}$$

STEP 4: Backward mapping is performed as follows. For all face component pixels' coordinates $FCP_i^* = \{(x_l^*, y_l^*)\}_i$ that are inside the triangles defined with vertices $FFP_i^* = \{(x_j^*, y_j^*)\}_i$, use the inverse transformation for mapping pixels' coordinates $FCP_i^* = \{(x_l^*, y_l^*)\}_i$ into corresponding pixels' coordinates $FCP_i = \{(x_k, y_k)\}_i$ defined over the original face component image.

$$\begin{bmatrix} x_k^* \\ y_k^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_i^c \\ 0 & 1 & t_i^c \\ 0 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 & t_i^x \\ 0 & 1 & t_i^y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} cos(\alpha_i) & sin(\alpha_i) & 0 \\ -sin(\alpha_i) & cos(\alpha_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_i^x & 0 & 0 \\ 0 & s_i^y & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} x_l^* \\ y_l^* \\ 1 \end{bmatrix}$$

STEP 5: Use spline interpolation to estimate RGB pixel values on transformed face component image on coordinates $FCP_i^* = \{(x_l^*, y_l^*)\}_i$ from RGB pixel values in the original face component image. Spline interpolation is necessary for pixel value mapping because pixel coordinates $FCP_i^* = \{(x_k, y_k)\}_i$ are real numbers while $FCP_i^* = \{(x_l^*, y_l^*)\}_i$ are all integer numbers.

Applying the above-described Steps 1-6 for all face components i =1, 2, ..., 6, a modified face is obtained, counted as a de-identified face.

The proposed face de-identification process is reversible. By applying a sequence of inverse transformations and using backward mapping and spline interpolation, the original appearance of the face can be restored.

4 Experiment setup and results

We performed two types of experiments. In the first type of the experiments we used crowdsourcing approach and in the second automatic face recognition based on ResNet [7].

4.1 Experiment setup and results

We compiled a set of 30 face images of famous people (7 females and 23 males with ages ranging from 30 to 75) from politics, sports, business and entertainment. The used testing procedure is like the one used for diagnosing of the prosopagnosia a neurological disorder characterized by the inability to recognize human faces. Images of de-identified faces of famous persons are presented to the test subjects with a request to try to recognize them. Obtained answers are recorded for later matching with ground truth answers. The main aim of the performed testing is to evaluate the impact of geometrical and texture modifications on human ability to recognize faces. The evaluation is performed by means of crowdsourcing performed by 150 test subjects (20 females and 130 males). The test subjects were informed that faces in the tests are de-identified faces of famous people. The background (ie context) and biometrical ques like hair and ears, that a user can use for face identification are removed in all tests. For example, some of them are shown in (Fig. 1. fifth row). We have performed five experiments.

First experiment The geometry of a face is left unchanged and texture of a face is replaced with the average texture obtained from 30 faces. All de-identified faces thus have the same texture and the original geometry (Fig. 1. first row). The following results are obtained in the crowdsourcing experiment: From the total of 4500 de-identified faces (30 faces x 150 test subjects) for only 60 de-identified faces original identity was revealed (ie. 1.33 % fail rate of de-identification).

Second experiment The geometry of a face is changed as described in Section 3. A texture of a face is obtained by blending 50 % of an original texture and 50% of an average face texture (Fig. 1. second row). The following results are obtained: From the total of 4500 de-identified faces for 260 de-identified faces original identity was revealed (i.e. 5.78 % fail rate of de-identification),

Third experiment The geometry of a face is changed as described in the Section 3. An original face texture was left unchanged (Fig. 1. third row). The following results are obtained: From the total of 4500 de-identified faces, original identities of 1410 de identified faces were revealed (i.e. 31.33 % fail rate of de-identification).

Fourth experiment The geometry of a face is changed to an average geometry obtained from 30 faces (i.e. average positions of 68 face feature points), while texture is left unchanged (Fig. 1. fourth row). The following results are obtained: From the total of 4500 de-identified faces, original identities of 1980 de-identified faces were revealed (i.e. 44.00 % fail rate of de-identification)

Fifth experiment Original images of famous person are used (Fig. 1. fifth row). The following results are obtained: From the total of 4500 de-identified faces, total of 3020 true identities were known (ie 67.11 %). In papers [8] authors have concluded that shape and texture are about equally important" for humans to recognize faces. The results of above first four experiments have shown that texture is even more important than face shape (face trigonometry).

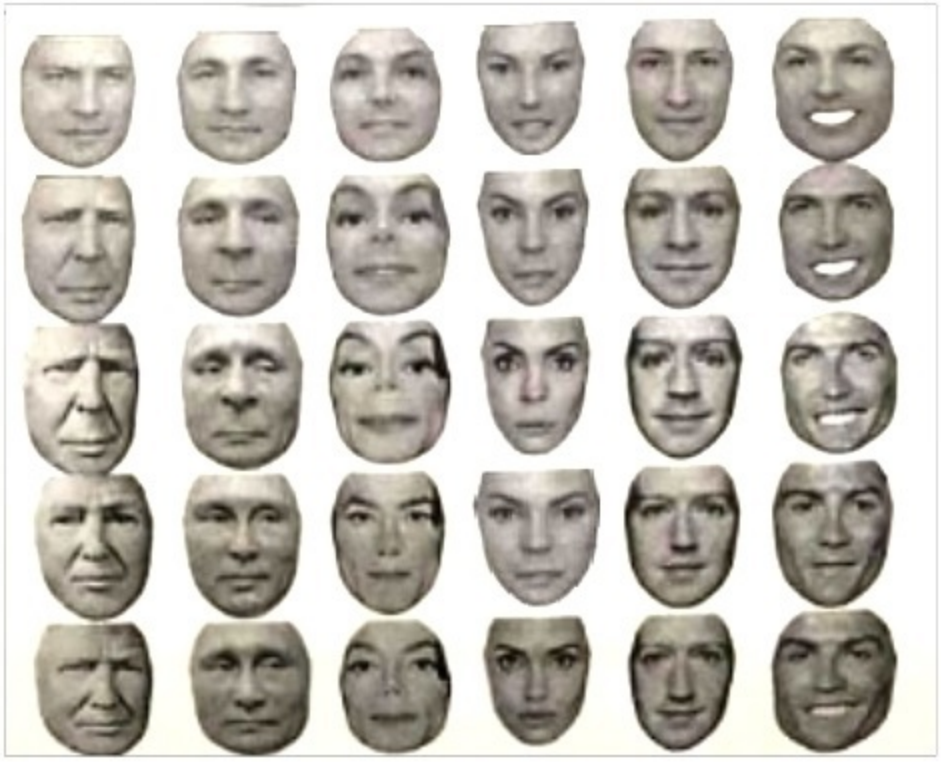

Fig. 1: Examples of six faces used in experiment 1-5 are depicted consecutively in row 1-5. Each row corresponds to one experiment and each column to one face

4.2 Automatic face recognition based on ResNet

ResNet network [7] with 29 convolution layers implemented in Dlib [9] is used. The network was trained on dataset of about 3 million faces [9]. Identical test samples from 1-5 experiments described above are used to evaluate a level of privacy protection when machine recognition approach is used. We are interested in comparing results obtained by humans and machine. Tab. 2. depict results. From the results we can conclude that the machine is much better than humans in the task of recognition of de identified faces (i.e., with altered visual appearances).

Tab. 2: Failed rate of de-identification (recognition rate) for humans and machine

Experiment	Human/Croudsourcing	Machine/ Resnet
1	0.0133	0.1333
2	0.0578	0.2333
3	0.3133	0.8667
4	0.4400	0.0667
5	0.6711	1

4.3 Automatic face recognition based on ResNet on KinFaceW-I Dataset

ResNet network with 29 convolution layers implemented in Dlib is used. Identical test samples from 1-5 experiments are used to evaluate a level of privacy protection on KinFaceW-I dataset [10] when machine recognition approach is used. KinFaceW-I has 500 kinship image face pairs. We are interested in comparing results obtained by humans and machine. Tab. 3. depict results. From the results we can conclude that the machine is much better than humans in the task of recognition of de identified faces (i.e., with altered visual appearances).

Tab. 3: Failed rate of de-identification (recognition rate) for humans and machine

Experiment	Human/Croudsourcing	Machine/ Resnet
1	0.0183	0.1233
2	0.0568	0.2233
3	0.3143	0.8767
4	0.4440	0.0267
5	0.6731	1

First experiment The geometry of a face is left unchanged and texture of a face is replaced with the average texture obtained from 500 faces. All de-identified faces thus have the same texture and the original geometry (Fig. 2. first row).

Second experiment The geometry of a face is changed as described in Section 3. A texture of a face is obtained by blending 50 % of an original texture and 50% of an average face texture (Fig. 2. second row).

Third experiment The geometry of a face is changed as described in the Section 3. An original face texture was left unchanged (Fig. 2. third row).

Fourth experiment The geometry of a face is changed to an average geometry obtained from 500 faces (i.e. average positions of 68 face feature points), while texture is left unchanged (Fig. 2. fourth row).

Fifth row It depicts the original images in the dataset without any modification.

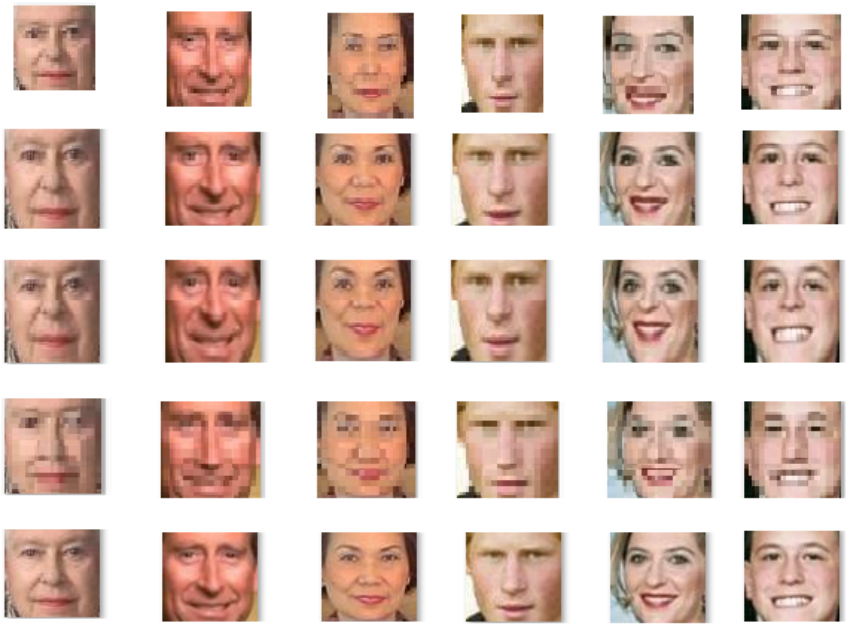

Fig. 2: Examples of six faces used in experiment 1-5 are depicted consecutively in row 1-5. Each row corresponds to one experiment and each column to one face

4.4 Comparision with state-of-art

Identical test samples from 1-5 experiments are used to evaluate a level of privacy protection with [11] [12] [13] [14] [15]. When machine recognition approach is used for various datasets. We are comparing results obtained by machine. Tab. 4. depict results.

Tab. 4: Failed rate of de-identification (recognition rate) for machine

Comparision	**Machine/ Resnet**
[11]	0.8204
[12]	0.758
[13]	0.654
[14]	0.79
[15]	0.1173
Proposed model	**0.0667**

From the results we can conclude that our method outperforms than the existing recognition of de identified faces (i.e., with altered visual appearances).

5 Conclusion

In this paper we proposed a hybrid reversible face de-identification pipeline that combines the good qualities of naive and complex face-de identification methods. The texture of a face can be adaptively modified based on an original texture as in naive approaches, Consequently, we have removed need for using tedious texture model of faces. Geometry of a face can be completely modified as in complex methods. These geometrical modifications performed by affine transformation are pseudo reversible what makes them compliant with security requirements. Only fourteen parameters for geometrical modification are used, what makes them suitable for steganographic encoding into the de identified image. For texture modification fixed texture template is used. We have investigated impact of various geometrical and texture modifications of face components like eyes, eyebrows, nose and lips on ability of humans and machines to recognize faces, The crowdsourcing and machine face recognition experiments have shown that modification of a face texture has stronger impact on a level of privacy protection then face geometry (shape) modifications and that the machine is superior to humans in the task of recognition of de-identified faces. Our future work will beverage scale evaluation.

References

[1] Jain, Anil K., and Stan Z. Li. Handbook of face recognition. Vol. 1. New York: springer, 2011.

[2] Ribaric, Slobodan, Aladdin Ariyaeeinia, and Nikola Pavesic. ”De-identification for privacy protection in multimedia content: A survey.” Signal Processing: Image Communication 47 (2016): 131-151.

[3] Chen, Renwang, Xuanhong Chen, Bingbing Ni, and Yanhao Ge. "Simswap: An efficient framework for high fidelity face swapping." In Proceedings of the 28th ACM International Conference on Multimedia, pp. 2003-2011. 2020.

[4] Liao, Shengcai, Anil K. Jain, and Stan Z. Li. "A fast and accurate unconstrained face detector." IEEE transactions on pattern analysis and machine intelligence 38, no. 2 (2015): 211-223.

[5] Kazemi, Vahid, and Josephine Sullivan. "One millisecond face alignment with an ensemble of regression trees." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1867-1874. 2014.

[6] Paul Chew, L. "Constrained delaunay triangulations." Algorithmica 4, no. 1 (1989): 97-108.

[7] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.

[8] Sinha, Pawan, Benjamin Balas, Yuri Ostrovsky, and Richard Russell. "Face recognition by humans: Nineteen results all computer vision researchers should know about." Proceedings of the IEEE 94, no. 11 (2006): 1948-1962.

[9] King, Davis E. "Dlib-ml: A machine learning toolkit." The Journal of Machine Learning Research 10 (2009): 1755-1758.

[10] Jiwen Lu, Xiuzhuang Zhou, Yap-Peng Tan, Yuanyuan Shang, Jie Zhou. Neighborhood Repulsed Metric Learning for Kinship Verification. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 36, no. 2, pp. 331-345, 2014.

[11] Li, Tao, and Lei Lin. "Anonymousnet: Natural face de-identification with measurable privacy." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2019.

[12] Wu, Y., Yang, F., Xu, Y. et al. Privacy-Protective-GAN for Privacy Preserving Face De-Identification. J. Comput. Sci. Technol. 34, 47-60 (2019). https://doi.org/10.1007/s11390-019-1898-8

[13] Zhu, Bingquan, et al. "Deepfakes for medical video de-identification: Privacy protection and diagnostic information preservation." Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. 2020.

[14] Li, Yuezun, and Siwei Lyu. "De-identification without losing faces." Proceedings of the ACM Workshop on Information Hiding and Multimedia Security. 2019.

[15] Du, Liang, et al. "GARP-face: Balancing privacy protection and utility preservation in face de-identification." IEEE international joint conference on biometrics. IEEE, 2014.