International Journal of Digital Crime and Forensics
Volume 14 • Issue 2

Face Anonymity Based on Facial Pose Consistency

Jing Wang, Yunnan Normal University, China
Jianhou Gan, Yunnan Normal University, China
Jun Wang, Yunnan Normal University, China
Juxiang Zhou, Yunnan Normal University, China*
Zeguang Lu, National Academy of Guo Ding Institute of Data Science, China

ABSTRACT

With the development of artificial intelligence, there are more applications related to face images. The recording of face information causes potential cyber security risks and personal privacy disclosure risks to the public. To solve this problem, the authors hope to protect face privacy through face anonymity. This paper designs a conditional autoencoder that uses the data preprocessing method of image inpainting. Based on the realistic generation ability of StyleGAN, their autoencoder model introduces facial pose information as conditional information. The input image only contains pre-processed face-removed images. The method can generate high-resolution images and maintain the posture of the original face. It can be used for identity-independent computer vision tasks. Experiments further prove the effectiveness of the anonymization framework.

KEYWORDS

Cyber Security, Face Anonymity, Generative Adversarial Network, Privacy Protection

INTRODUCTION

Thanks to the vigorous development of network technology and social media, a large number of photos are shared on social platforms, most of which are face pictures. Face information is not only personal core privacy but also personal sensitive information. Face information involves not only personal portraits, but also private information such as health, age, and race. The China Consumer Association has released a personal information collection and privacy policy evaluation report for 100 apps. The report shows that 10 of the 100 apps evaluated are suspected of over-collecting personal biometric information. Therefore, how to protect their biometric information, especially facial information, from being abused by unauthorized software and malicious attackers has become the focus of attention.

The general data protection regulation (GDPR) in Europe regulates data security and affects all personal data processing in Europe. GDPR requires individuals to regularly agree to use their data in any scenario. Fortunately, if the data cannot identify individuals, we can freely use the data without the user's consent. Therefore, we need a robust model to hide the identity information of the original face for face anonymity without changing the original distribution of the face image and retaining the validity of the image, the output should be a data distribution consistent with the given real face.

Face anonymization, also known as face de-identification, refers to generating another face image with a similar appearance without changing the background while hiding the real identity, to protect

DOI: 10.4018/IJDCF.302872 *Corresponding Author

the privacy of the corresponding person. Traditional anonymous methods (Boyle et al., 2000; Gross et al., 2009) are mainly based on fuzzy processing, which can eliminate the given identity to a great extent, but these methods will cause poor visual perception and can no longer be applied to computer vision tasks such as facial expression recognition. Most of the methods based on k-same (Meden et al., 2018; Newton et al., 2005) perform face recognition in a closed set and are not suitable for processing a single image. The method based on antagonistic disturbance (Kingma & Welling, 2013) (Sharif et al., 2016) usually highly depends on the reachability of the target system, requires special training, and has poor robustness. Recent generative-based methods (Hukkelås et al., 2020) (Chen et al., 2020; Guo & Chen, 2019; Hukkelås et al., 2019; Meden et al., 2017; Ren et al., 2018; Sun, Tewari, & Xu, 2018; Zhang, Hu, & Luo, 2018) also have difficulty generating realistic anonymous faces.

Our goal is to preserve the face pose information as much as possible and generate a realistic anonymous face image on the premise of hiding the real identity. We need to strike a balance between privacy protection and the effectiveness of preserving data, which cannot be balanced by previous methods. The model we proposed is a conditional autoencoder model. Our model is based on the generative model StyleGAN (Karras et al., 2020) proposed by Karras et al, which is one of the best generation models at this stage and can generate realistic faces from random noise sampling. Firstly, the anonymous image is preprocessed to obtain the image background and sparse face pose information to ensure that all face privacy-sensitive information is deleted. Then we map this information to the W+ latent space through a feature pyramid network, generate realistic face images through the StyleGAN model, and let the generated images learn the random face identity information generated by random noise. Finally, we can generate realistic anonymous faces and retain the original face pose information.

The main contributions of this paper:

- Conditional pose information is added to the network generation model so that the generated face retains the pose information on the premise of hiding the identity information
- The anonymous face we generate is unknowable. The anonymous face generated each time is random, which only ensures the same image background and pose. Their identity information is not taken from the identity information of existing faces, which will not affect anyone's privacy.
- We weigh the anonymity, pose retention, and generation quality of the generated image, and compare it with some existing methods. Some data show that our method achieves the best effect.

RELATED WORK

Traditional Obfuscation-Based Methods

In early research on traditional face anonymity (Boyle et al., 2000; Gross et al., 2009), face de-recognition technology mainly used fuzzy, pixelated, occlusion, and other methods to blur the privacy-sensitive face areas in the image. Because of their simple operation, these fuzzy technologies have been widely used by Internet news media, social media platforms, and government agencies. However, some studies (McPherson et al., 2016; Oh et al., 2016) show that this kind of blur-based method is not safe and can still recognize the identity of pixelated or blurred images. At the same time, such methods greatly change the data distribution and bring poor visual perception. Moreover, images processed based on traditional anonymous methods often destroy the availability of data.

Adversarial Perturbation-Based Methods

The method based on adversarial perturbation produces an imperceptible but useful worst-case perturbation to the original image to make the model recognize the wrong identity. Sharifet al. (Sharif et al., 2016) designed a special kind of glasses that can cause the wearer to be mistaken. Fawkes (Shan et al., 2020) applied imperceptible pixel-level changes to the picture before the user published the picture to achieve anonymity protection. When used to train the face recognition model, the

functional model generated by these "invisible" images will always lead to the user's normal image being incorrectly recognized. This type of method requires high requirements for the accessibility of the target system, so it usually only works on the recognizer of a specific target.

Image Generative-Based Methods

Image generation models can generate false and realistic images. It is widely used in the fields of image inpainting, image-to-image conversion, and face-swapping. The common generation model includes conditional variational autoencoder (VAE) (Kingma & Welling, 2013), generative adversarial network (GAN) (Goodfellow et al., 2014), and some variants based on them (Chen et al., 2020; Guo & Chen, 2019; Hukkelås et al., 2019; Meden et al., 2017; Ren et al., 2018; Sun, Tewari, & Xu, 2018; Zhang, Hu, & Luo, 2018). The image generative-based method has become popular in the field of face anonymity. It can be divided into two directions: the methods based on face swapping and the methods based on face image inpainting.

Conditional ID-Swapping-Based Methods

The anonymous method based on face id-swapping is to replace the original face area with another person's face, exchange the identity information of two people, and achieve the purpose of anonymizing the original face. Sun et al. (Sun, Tewari, & Xu, 2018) mixed parametric face synthesis technology with a generation countermeasure network for data-driven image synthesis to achieve the effect of anonymity. Meden et al. (2017) proposed a face recognition pipeline to ensure anonymity by synthesizing artificial agent faces using generative neural networks. SimSwap Chen et al., (2020) proposes a weak feature matching loss to ensure that the input target and the generated result are consistent at the attribute semantic level. Although these methods hide the identity information of the target face and protect its privacy, they infringe on the privacy of the target face. Although most of the target faces selected in face exchange are the face images of some public figures, when anonymous scenes involve damage to the character image or even some illegal and criminal scenes, no matter whose face information is exposed, it is reluctant for anyone to see.

Conditional Inpainting-Based Methods

The anonymity method based on image completion completely shields the information of face privacy-sensitive areas and uses GANs to generate semantically consistent results for missing areas. Some methods Guo & Chen, (2019); Zhang, Hu, & Luo, (2018) reduced the damaged area in the natural image by gradually drawing the missing area and producing the desired repair results. Zhang et al. (2019) presented an approach for pluralistic image completion, this method can generate complex and diverse face images. Ren et al. (2018) trained a face anonymizer to delete private information and maximize the performance of spatial behavior detection. DeepPrivacy Hukkelås et al., (2020) generates authentic anonymous faces through existing background and sparse pose annotation. CIAGAN Maximov et al., (2020) uses facial key point coordinates and one hot coded identity vector to generate realistic anonymous faces while retaining other necessary features. The larger the face privacy shielding area is, the more private information is protected, and the more difficult it is to generate realistic faces. This often requires more advanced generation models to understand complex semantic reasoning, which makes the task very difficult.

Although image generative-based methods are becoming increasingly popular in the research of face anonymity, they are affected by various conditional information. In this paper, we design an image generation method based on a conditional autoencoder to perform image inpainting, to achieve of face anonymity, and to try to solve the current face anonymity problem.

METHOD

The method based on image generation has achieved good results on the task of face anonymity. Most studies add some conditional information to make the generated face meet the original data distribution and retain some conditional information. In this paper, we combine the methods of face exchange and image restoration to train a condition generator, which aims to generate realistic anonymous faces while preserving the pose information of the original face. This goal can be expressed as:

$$\varphi\left(X\right) = X'$$

where, $\varphi()$ represents the transformation function of the face de-identification model, which maps the given original face image X to image X'; at the same time, the original face X and the adult face X' have different identity information and the same pose information.

Data Preprocessing

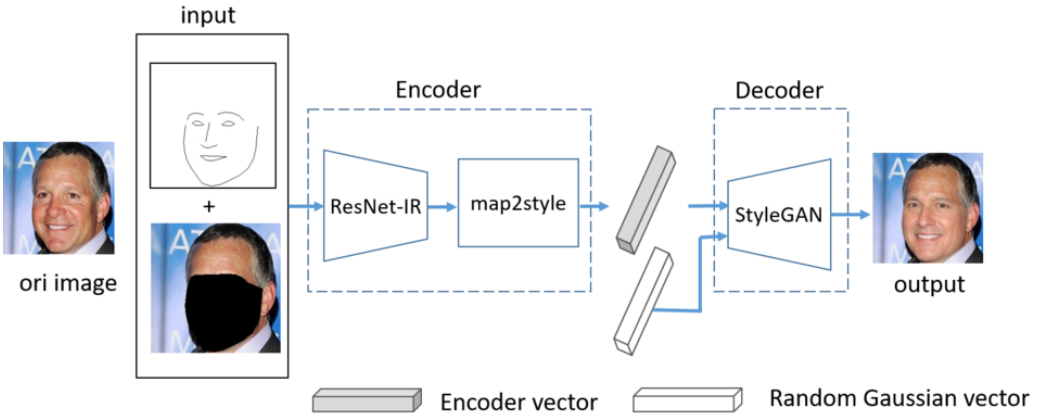
Since the identity information of the image is concentrated in the face area, to make our method focus on the generation of the face area rather than the background, we refer to the similar image preprocessing method in CIAGAN, use a mask to cover all the face areas except the forehead to delete all privacy-sensitive information and use the face pose detector to generate. The face area only retains the forehead information, which allows the network to learn some non-identity information. On the one hand, the forehead of the face contains little identity information, which has little impact on identity and avoids identity disclosure. On the other hand, the forehead information contains the face skin color information. When the network learns the information here, the face skin color of the generated image can be matched with the face skin color of the original image.

To make the generated face image contain the pose information of the original face, our network is designed as a condition generator, which can learn the pose information of the original face. Deepprivacy represents sparse facial pose estimation with pose information, including 7 key point coordinates of the ear, eye, nose, and shoulder. PRVSLi et al., (2019) takes the edge contour as the prior knowledge of the network to generate an image with a consistent attitude. Sun et al. (2018) and CIAGAN (2020) used the Dlib face detector (2014) to detect 68 facial landmark points and obtain a sparse representation of face poses. We also use the Dlib face coordinate detector to generate abstract face pose images. We do not use all 68 key face point coordinates but only the coordinate information of the face contour, eyebrows, eyes, nose, and mouth. This information is represented as a binary image as part of the generator input.

Network Model

Our network model structure refers to the network model in PSP (2021). PSP has achieved good results in image inversion, face normalization, conditional image synthesis, image restoration, and image super-resolution. Our network model adds conditional pose information to the original model, optimizes the loss function according to the specific task, and finally generates an anonymous face while retaining the facial posture information of the original image. The model framework is shown in Figure 1.

Figure 1. *Our input includes the background information and abstract pose representation of the picture, and they are spliced. The generator is a conditional autoencoder model. The image information is embedded into the low-dimensional space through the encoder to generate a potential vector with a size of 18x512, and a false face is generated through the StyleGAN encoder. The generated false face learns the identity information in the false face generated by a random Gaussian vector.*



The image generation model of the network adopts a conditional autoencoder structure, including two parts: encoder and decoder. The decoder structure used in this paper is the best StyleGAN network structure at this stage. In StyleGAN, the author shows that the latent vector input of different styles corresponds to different levels of detail, which can be roughly divided into three levels: rough, medium, and high quality. At the same time, some studies based on StyleGAN show embedding the input image into the W+ latent space, and then affine transformation can generate images with higher quality and more realistic details. The potential W+ vector can be expressed as the style representation of different network levels, with a size of 18x512. Therefore, the purpose of the encoder is to map the input image to the W+ latent space. According to this goal, we use the feature pyramid based on ResNet to extract different levels of features of the image and then map the extracted features to the W+ latent space through a simple intermediate network map2style to generate a style potential representation with a size of 18x512. The intermediate network map2style is a small full convolution network that uses a set of 2x2 convolutions. Then, it activates Leaky ReLU to gradually reduce the space size and map the feature map to 1x1x512. To anonymize the face generated by our model, we randomly sampled an 18x512 potential vector before the StyleGAN decoder to generate a false face that does not exist in the real world and let the network learn the identity information of the random false face so that the final generated face can achieve the purpose of anonymity.

Loss Function

We use the weighted sum of various losses, including L2 loss of posture consistency, identity loss, and perceptual similarity loss, which are optimized by the gradient descent method.

To make the generated image have the same facial pose as the original image, we use the pixel level $\mathcal{L}2$ loss, $L()$ represents the Dlib face coordinate detection model, and the processing method is the same as the data preprocessing method. Our goal is not to reconstruct and generate an image similar to the original image, so we only calculate the $\mathcal{L}2$ loss of face coordinates of the original image X and the generated image X'.

$$\mathcal{L}_2 = L\left(X\right) - L\left(X'\right)_2^2$$

Loss of identity is defined as:

$$\mathcal{L}_{id} = 1 - \cos\left(R\left(X\right), R\left(G\left(Z\right)\right)\right)$$

where $R()$ is the pretrained model of face recognition ArcFace (2019), Z represents the random Gaussian vector sampled from the hidden space, the size is 18x512, $G\left(Z\right)$ represents the random identity face image generated from the random Gaussian noise sampling input StyleGAN pretrained model, and $\cos()$ represents the cosine similarity of the two identity vectors.

To learn the perceptual similarity between the original image X and the generated image X', we use LPIPS loss (2018). Compared with the more standard perceptual loss (2016), this method has been proven to better maintain image quality:

$$\mathcal{L}_{LPIPS} = F\left(X\right) - F\left(X'\right)_2$$

Therefore, the total loss of the network consists of multiple losses with different weights, which is defined as:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{lk} + \lambda_2 \mathcal{L}_{id} + \lambda_3 \mathcal{L}_{LPIPS}$$

where $\lambda 1$, $\lambda 2$, and $\lambda 3$ are the hyper-parameters that define the loss weight. Since different losses have different effects on the convergence of the network, we judge our training results according to some visualization tools and make adjustments in time.

EXPERIMENT AND DISCUSSION

In this section, we compare our method with several common anonymous methods based on tradition and learning. Our method achieves the most advanced qualitative and quantitative results in the face image dataset.

Dataset

The CelebA-HQ dataset randomly samples 30000 pieces from the CelebA dataset and generates 1024x1024 high-quality face images from celebrity face images with an original image size of 178x218 through the generation strategy of ProGAN (2017). The dataset contains different demographic information such as age, gender, and race. According to the task goal, we preprocess the CelebA-HQ dataset to obtain the image background of deleted irregular face regions and diluted face pose representations. After the Dlib face detector, 29613 effective faces were detected, and the data availability rate was 98.91%. We randomly selected 3000 images as the testing dataset and the remaining 26613 images as the training dataset.

Implementation Details

In the training, the input image size is 256x256. A total of 1024x1024 images are generated by the StyleGAN decoder and downsampled to the original input size. The face region of the generated image is fused with the original background to generate the final anonymous face. In the training process, we use the Ranger optimizer to update the parameters. The Ranger optimizer combines the advantages of radam and lookahead to achieve a better optimization effect. In terms of parameter

setting, our learning rate is 0.0001, the batch size is 4, and the weight of the loss function is $\lambda 1 = 1$, $\lambda 2 = 0.1$, and $\lambda 3 = 0.1$. In the training process, the encoder uses the pretrained ResNet structure for face feature extraction, which can speed up the convergence. Although the encoder StyleGAN can generate realistic face images, StyleGAN adopts a progressive training strategy, which takes too long and is cumbersome training. Therefore, we used StyleGAN's pretrained model, fixed the parameters, and trained only the encoder part. Our experiment was trained on a single NVIDIA RTX 3090.

Evaluating Indicator

We evaluated the indicators of face images generated by all models in face detection and face de-identification. We use Dlib and MTCNN(2016) as public face detectors. For the detection performance of the model, we use the percentage of detected faces to evaluate. To verify the performance of face anonymity, we use the cos distance between the original face image and the generated face image, and we use ArcFace for feature extraction. In addition, we use structural similarity (SSIM) and frè Chet perception distance (FID) (2017) to measure the quality of picture generation. SSIM measures the similarity between pictures from brightness, contrast, and structure. The larger SSIM, the more similar pictures are. FID is a measure that compares the statistics of the generated sample with the real sample. The lower the FID is, the better, and the more similar the corresponding real sample is to the generated sample.

Comparison Results

We compare traditional anonymous methods with learning-based anonymous methods. Based on traditional anonymous methods, including pixelation and Gaussian blur, (1) pixelization: cluster the pixels of the face area close to each other in the two-dimensional space and color space, and then cluster each pixel Replace with its average. (2) Gaussian blur: pass the image through a Gaussian filter to make the pixels around the image average. We select filter core sizes of 15 and 25 for comparison. Learning-based anonymous methods include DeepPrivacy (2018) and Fawkes (2020).

In this method, the anonymous region only includes the irregular face region detected by the Dlib face detector. For the generated images, we hope to preserve the data utility as much as possible, so that they can still be detected by the face detector, retain a high detection rate, and can be applied to other tasks. Therefore, in the experimental part, we first use Dlib (2005) and MTCNN (2016) for evaluation. In Table 1 and Figure 2, we show the detection results and visualization results of our method and the traditional anonymous method on the CelebA-HQ dataset. The results show that the larger the pixel block and Gaussian filter core, the more blurred the observable information and the worse the detectable performance. Compared with the traditional face anonymity method, all anonymous images generated by our method can be detected and show the best visual effect in visual perception.

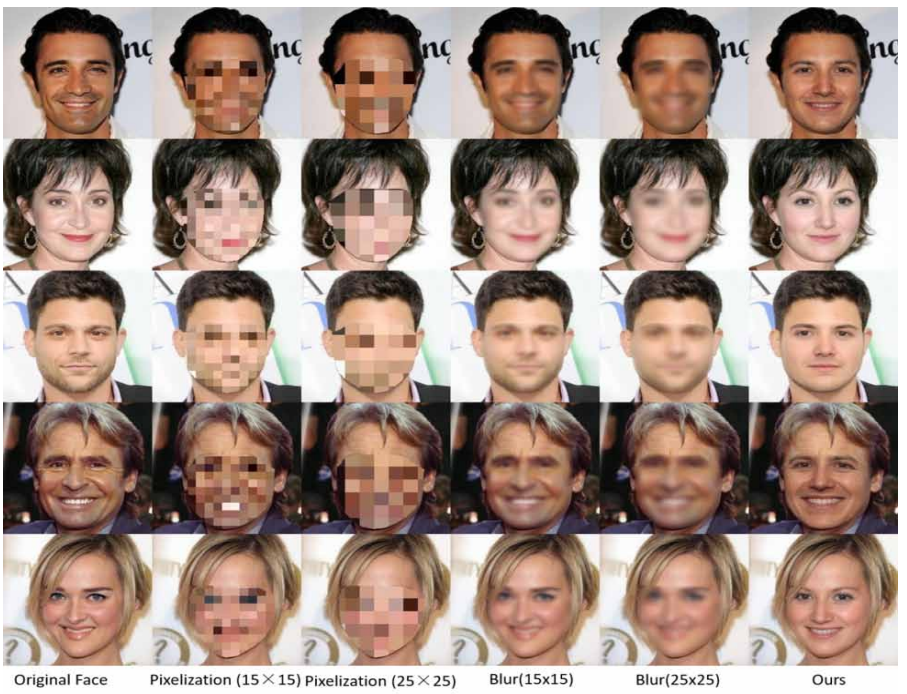
Table 1. Test results of common test models. The values in the table represent the percentage of accuracy of detected faces. The larger the value, the more detectable faces and the better the model retains the detection ability.

Models	Detection	
	Dlib	MTCNN
Original	100	100
Pixelization(15x15)	21.71	64.42
Pixelization(25x25)	0.13	8.17
Blur(15x15)	97.61	99.93
Blur(25x25)	92.22	99.83
Ours	100	100

Table 2. Quantitative evaluation using the CelebA-HQ dataset under different image generation quality indicators

Models	ID_DIS(Arcface)	SSIM	FID
Pixelization(15x15)	0.7385	0.8483	210.82
Burl(25x25)	0.3126	0.9272	68.27
Fawkes	0.4212	0.9821	16.42
DeepPrivacy	0.8741	0.8071	34.77
Ours	0.7007	0.9157	13.93

Figure 2. Compared with traditional methods, from left to right are the original face, pixelization (15x15,25x25), Gaussian blur (15x15,25x25), and the image generated by our method.



In Table 2, we show the quantitative results of our method and all models under different indicators. In the FID index, we have achieved the best results, which proves that our method has achieved the best results in the generated image quality. In the SSIM index, although both Gaussian blur and Fawkes exceed our method, Gaussian blur has poor performance in image perception and face detection. Fawkes adopts the method based on adversarial perturbation, which adds imperceptible disturbance to pixels and naturally has characteristics similar to the original image. For the de-recognition performance, we use the cosine distance between faces to measure, and id_dis is between 0 and 2. The id_dis of our generated image is close to the pixelization (15x15), which has the anonymous effect of similar pixels and has reached the effect of being visually inconsistent with the original face identity. We compared our method with the images generated by Fawkes and DeepPrivacy. From Figure 3, we found that Fawkes can generate a face that looks very much like the original face while anonymously retaining the posture of the original face, but there will be some strange spots occasionally. The image generated by DeepPrivacy can well maintain the facial pose, which is visually different from the original image but does not retain the pose of the original image. The generated face is prone to distortion and occasionally produces abnormal skin color, which is unnatural. Our method is slightly inferior to DeepPrivacy in terms of privacy protection performance measured by identity distance, but the data utility is significantly improved, and it is better than DeepPrivacy in visual effect and quantitative evaluation index (SSIM, FID). In general, in all models, the face image generated by our method achieves the effect of anonymity in vision and id_dis, retains the pose and skin color of the original face, and the generated image quality is the highest.

Figure 3. Qualitative comparison of our method with Fawkes and DeepPrivacy. The first column is the original face. The second column is the anonymous face generated by Fawkes, the third column is the anonymous face generated by DeepPrivacy, and the last column is our result

CONCLUSION

In this paper, we have designed a conditional autoencoder, which uses an image inpainting method to generate anonymous faces through background information and sparse pose information, while retaining the facial posture of the original face image to generate realistic effects. Our method consists of two stages. Firstly, the irregular region of the face is detected by the face detector, and the image background and sparse face pose expression that deletes the privacy-sensitive region of the face are generated. Then, a conditional generation model is used to generate anonymous faces while preserving the original face pose. Experiments show the validity of our method in terms of privacy protection and image retention utility. Compared with the traditional and latest methods, satisfactory results are obtained.

FUNDING AGENCY

Open Access Funding for this article has been covered by the authors of this manuscript.

ACKNOWLEDGMENT

This work is supported by Yunnan Expert Workstation of Xiaochun Cao.

REFERENCES

Abdal, R., Qin, Y., & Wonka, P. (2019). Image2stylegan: How to embed images into the stylegan latent space? *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4432-4441.

Abdal, R., Qin, Y., & Wonka, P. (2020). Image2stylegan++: How to edit the embedded images? *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8296-8305.

Boyle, M., Edwards, C., & Greenberg, S. (2000). The effects of filtered video on awareness and privacy. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work* (pp. 1–10). ACM. doi:10.1145/358916.358935

Chen, R., Chen, X., & Ni, B. (2020). SimSwap: An Efficient Framework For High Fidelity Face Swapping. *Proceedings of the 28th ACM International Conference on Multimedia*, 2003-2011.

Dalal, , & Triggs, . (2005). Histograms of oriented gradients for human detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR 2005)*, 886–893.

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition. Proc. CVPR, 4690–4699.

Goodfellow, I., Pouget-Abadie, J., & Mirza, M. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27.

Gross, R., Sweeney, L., Cohn, J., De la Torre, F., & Baker, S. (2009). Face de-identification. In *Protecting privacy in video surveillance* (pp. 129–146). Springer. doi:10.1007/978-1-84882-301-3_8

Guo, Z., & Chen, Z. (2019). Progressive image inpainting with full-resolutionresidual network. In *Proceedings of the 27th ACM International Conference on Multimedia*. ACM. doi:10.1145/3343031.3351022

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, 6626–6637.

Hukkelås, H., Lindseth, F., & Mester, R. (2020). Image Inpainting with Learnable Feature Imputation. In *DAGM German Conference on Pattern Recognition*. Springer.

Hukkelås, H., Mester, R., & Lindseth, F. (2019). DeepPrivacy: A generative adversarial network for face anonymization. In *International Symposium on Visual Computing*. Springer. doi:10.1007/978-3-030-33720-9_44

Johnson, J., Alahi, A., & Li, F.-F. (2016). Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision* (pp. 694–711). Springer. doi:10.1007/978-3-319-46475-6_43

.Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196.

.Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4401–4410. doi:10.1109/CVPR.2019.00453

Karras, T., Laine, S., & Aittala, M. (2020). Analyzing and improving the image quality of stylegan. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8110-8119.

Kazemi, V., & Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. *IEEE Conference on Computer Vision and Pattern Recognition*, 1867–1874. doi:10.1109/CVPR.2014.241

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.

Li, J., He, F., & Zhang, L. (2019). Progressive reconstruction of visual structure for image inpainting. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 5962-5971. doi:10.1109/ICCV.2019.00606

Maximov, M., Elezi, I., & Leal-Taixé, L. (2020). Ciagan: Conditional identity anonymization generative adversarial networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,* 5447–5456. doi:10.1109/CVPR42600.2020.00549

McPherson, R., Shokri, R., & Shmatikov, V. (2016). Defeating image obfuscation with deep learning. arXiv preprint arXiv:1609.00408.

Meden, B., Emeršič, Ž., Štruc, V., & Peer, P. (2018). k-Same-Net: k-Anonymity with generative deep neural networks for face deidentification. *Entropy (Basel, Switzerland)*, *20*(1), 60. doi:10.3390/e20010060 PMID:33265147

Meden, B., Mallı, R. C., Fabijan, S., Ekenel, H. K., Štruc, V., & Peer, P. (2017). Face deidentification with generative deep neural networks. *IET Signal Processing*, *11*(9), 1046–1054. doi:10.1049/iet-spr.2017.0049

Newton, E. M., Sweeney, L., & Malin, B. (2005). Preserving privacy by de-identifying face images. *IEEE Transactions on Knowledge and Data Engineering*, 232–243. .10.1109/TKDE.2005.32

Oh, S. J., Benenson, R., & Fritz, M. (2016). Faceless person recognition: Privacy implications in social media. In *European Conference on Computer Vision*. Springer.

Ren, Z., Jae Lee, Y., & Ryoo, M. S. (2018). Learning to anonymize faces for privacy preserving action detection. *Proceedings of the European Conference on Computer Vision (ECCV),* 620–636. doi:10.1007/978-3-030-01246-5_38

Richardson, E., Alaluf, Y., & Patashnik, O. (2021). Encoding in style: a stylegan encoder for image-to-image translation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2287-2296. doi:10.1109/CVPR46437.2021.00232

Shan, S., Wenger, E., & Zhang, J. (2020). Fawkes: Protecting privacy against unauthorized deep learning models. *29th USENIX Security Symposium (USENIX Security 20),* 1589-1604.

Sharif, M., Bhagavatula, S., & Bauer, L. (2016). Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition. *Proceedings of the 2016 acm sigsac conference on computer and communications security*, 1528-1540.

Sun, Q., Ma, L., & Oh, S. J. (2018). Natural and effective obfuscation by head inpainting. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5050-5059.

Sun, Q., Tewari, A., & Xu, W. (2018). A hybrid model for identity obfuscation by face replacement. *Proceedings of the European Conference on Computer Vision (ECCV)*, 553-569. doi:10.1007/978-3-030-01246-5_34

Zhang, H., Hu, Z., & Luo, C. (2018). Semantic image inpainting with progressive generative networks. *Proceedings of the 26th ACM International Conference on Multimedia*, 1939-1947.

Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, *23*(10), 1499–1503. doi:10.1109/LSP.2016.2603342

Zhang, R., Isola, P., Alexei A Efros, E. S., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. Proceedings of the IEEE conference on computer vision and pattern recognition, 586–595. doi:10.1109/CVPR.2018.00068

Zheng, C., Cham, T. J., & Cai, J. (2019). Pluralistic image completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1438-1447). IEEE.