A Survey of Privacy Preserving in Deep Learning

Mak Chun Chung, Desmond

The Hong Kong Polytechnic University

Abstract

Deep Learning applications have become common in people's life, such as Machine Translation, Objects Classification, Recommendation systems, etc. One of the critical factors in achieving an outstanding deep learning application is the dataset. Deep Learning requires using the related dataset to train the model in order to reach the purpose. Facial Recognition is one of the most widely used deep learning applications, and it has been applied everywhere. Its model training needs many images of a human's face. However, the people resist that their photos are collected to train the companies' recognition system or used by the governmental agencies for various purposes. These concerns could be generalized because people are more concerned about whether facial recognition technology could impact their privacy. This report will introduce five privacy-preserving approaches to prevent facial recognition systems trained by using people's photo, which include: cloaking [1]*, Privacy-Preserving Representation-Learning Variational Generative Adversarial Network (PPRL-VGAN)* [2]*, Face-off* [3]*, Gender Suppressing* [4]*, and Privacy-Protective-GAN* [5].

1. Introduction

Deep Learning is widely used in various services and products, and it can provide extreme performance than the traditional approaches or machine learning approaches. Computer vision (CV) is one of the popular areas of applying deep learning to improve performance or handle some tasks that only perform well using deep learning, such as Image Generalization and Object Recognition. Deep learning models are just like machine learning models, which require the related dataset to train a model to reach the purpose. For example, in facial recognition, the model should be trained with many people's face images. Social media are widely applying facial recognition, such as analyzing the user’s activities to recommend the latent information that the user might interested in. Also, the government is another organization of widely using facial recognition. For example, the airports and port of entry (POE) apply facial recognition to catch the criminal people who restrict leaving the country.

The dataset of training a facial recognition model is critical. Today, the approach of collecting the required dataset can be various. Taking an example of social media, which is viral and common these days, people always upload their social life pictures and selfie on the app, which has become a habit. When the users upload their photos to the social media app, it can use these photos to train the recognition model to recognize the user. However, people are disgusted with and resist these collections of their photos for other purposes. Especially, people start concerns about whether deep learning is risky to their privacy.

Hence, some researchers invent various solutions to prevent the recognition models trained by using the people's photos, in case the facial recognition system is able to identify those people. There are three types of solution: (1) Using Generative Adversarial Network (GAN) based model to generate the "special" photo that seems normal but unidentifiable to the facial recognition model, (2) By adding the additional features on the people's face in the photo in order to make the recognition model cannot recognize the people whose face has been covered by the additional features. (3) Removing the identification from the soft biometrics layer.

In this report, we will explain why we adopt these three types and describe them in section 2. Further, we have found five solutions that are developed based on these three types. They will be explained in sections 3 to 7 in detail, and the comparisons of them are described in section 8. Nevertheless, each of them has limitations, and we will describe them in section 9.

2. Problem Statement

People are worried about the deep learning applications that could impact their privacy. Facial Recognition is one of the most widely used deep learning applications, and it also is one of the user worried applications because of its formidable human face identification. At the same time, the rapid growth of the technology and corresponding applications make the people who are over-reliance on those technologies. Thus, people would like to keep sharing their photos on the Internet and enjoy the services of the Internet applications but preserving their privacy. It is challenging work because collecting people's photo for training a recognition system is effortless now. The images for training can be obtained from the companies' services (e.g., social media platform) or web scraping to automatically collect the photos on the Internet. For example, IBM had collected nearly a million photos from Flickr without noticing some of the photographers of those images [6]. Some countries have promulgated corresponding laws to protect the people's privacy on the Internet, such as the Australian Government has a law to forbid personal photos taken without permission [7]. However, these are the penalization mechanisms after those companies or organizations whose actions are exposed. The more particular approach should be making the photos uploaded on the Internet which become un-trainable or unidentifiable to the facial recognition model.

In section 1, we have mentioned that this report will introduce the approaches in three types of solutions if facial recognition exploits users' privacy because they have the outstanding ability to handle the limitations of the traditional solutions. They have novel solutions to achieve particular purposes.

The privacy leakage issue by facial recognition is not a recent issue. Although there are already some existing solutions, their unsatisfactory performance and limitations cannot convince the users. For example, modifying the faces on the photo (e.g., blurring, obscuring, or morphing) and inserting the particular physical objects (e.g., clothes, eyeglass frames, hat, etc.) on the people who are in the photo [3]. The first one belongs to the methodologies of computer vision. However, they result in significant impacts (e.g., the faces are unidentifiable by human eyes) to the face on the original photos. Users are not willing the faces on their photos are destroyed to prevent face detection and recognition. The latter solution belongs to the evasion attacks [1]. It relies on adding some wear on the people in the photo leading the system to misidentify those people. Nonetheless, it is restricted by two limitations. Some additional wear must be added to the people's faces or bodies in the photo. Also, the attacking model is not robust to the update of the target recognition system. So, it has to keep tracking the update of the target recognition system constantly. Otherwise, the update of the target recognition model can break the attacking model.

Hence, researchers invent novel methods to provide better privacy-preserving ways to users. Adding the additional features on people's faces on the photo [1] [3] and GAN-based methods [2] [5] can provide outstanding performance in preventing the identification of recognition systems. Further, they solve the limitations of the traditional approaches. Apart from disabling the recognition system's identification function or making it to be misidentified the people, we think that the privacy issues can involve the soft biometrics layer. Since the recognition system is not only used to identify who the people are, it can be extended to read the people's information according to its face or body. For example, a trained recognition model can identify people's ages, gender, race, etc. Thus, this report also discusses a method for suppressing a recognition system's soft biometrics identification ability [4].

3. Image Cloaking

Fawkes, a system invented by researchers to prevent the deep learning recognition model from identifying the people on images [1]. It leverages data poison attack to the dataset to affect the deep learning model identification, also called "cloaking." This model shows outstanding performance without the limitations that evasion attacks have, which means that the naked eye cannot notice the difference between the "cloaked" photo and the original photo, but the deep learning recognition system cannot identify the people the photos correctly.

The experiments result provided by the researchers who develop *Fawkes* show excellent performance in different types of experiments [1]. It can provide 100% protection against state-of-the-art facial recognition services from famous companies (e.g., Microsoft and Amazon Web Service). For the user recognition in the web services, it provides over 95% protection. Even giving the "uncloaked" images to the recognition model for training, there still is over 80% success leading the recognition model to misidentify the "cloaked" photos (i.e., the training image is the same person of the inputting testing image, but the training image is "uncloaked" and the testing image is "cloaked").

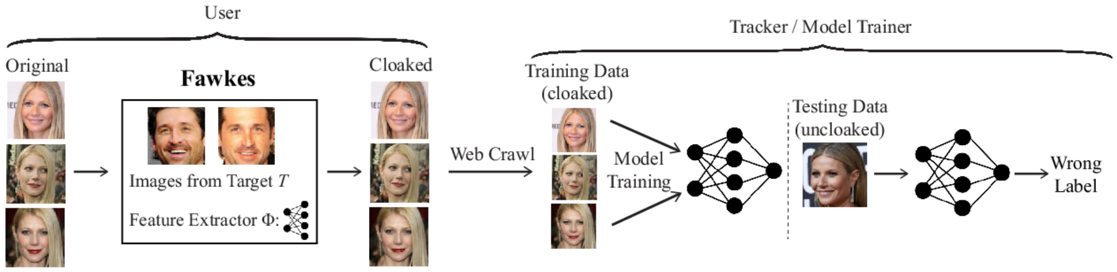


Figure 1 Fawkes system protecting user's privacy by cloaking the training data [1]

The deep learning recognition model is trained to extract features according to those features to achieve the identification. This approach aims to add slight perturbations on the photos in the features layer (i.e., cloaked photos), in which those perturbations are imperceptible so that the trained recognition model misidentifies the people on the normal photos (i.e., uncloaked photos). It results in the recognition model, which thinks itself correctly recognizes the person but in the fact that it is wrong. Since the modification added by *Fawkes* is the wrong feature, the model will classify the label class of that person as another person. Figure 1 shows the above procedure of how Fawkes works.

Thereby, the critical factor of *Fawkes* is computing the cloak perturbations. The following is the "cloaking" procedure: (1) "Cloaking" aims to move the original label of the image from the original position to another label class in the feature space, and the distance of that label class and original label class of the input image is the longest. Thus, *Fawkes* first needs to select the target class, which is the longest from the original label. The target class is selected from the public dataset, which contains multiple groups of images. (2) Then, *Fawkes* randomly pick the images from the target class found on the procedure (1) and using Structural Dis-Similarity Index (DSSIM) to choose the most suitable one as the "cloaks." Also, *Fawkes* will optimize the "cloaks" simultaneously to make the "cloaks" image-specified. The following is the equation of computing the "cloaks" of the people on the image and optimizing the "cloaks" to be image-specified:

$$\min_{\delta} Dist\left(\phi(x_T), \phi\left(x \oplus \delta * (x, x_T)\right)\right) + \lambda \cdot \max\left(|\delta(x, x_T)| - \rho, 0\right)$$

Also, the λ in the above equation represents the parameter of making the resulting image visually similar to the original image. (3) The main computation tasks are finished. The remaining part is that the user has to ensure that no "uncloaked" images are uploaded and prevent user’s friends' tag user on their shared images on the Internet because it can also provide the "uncloaked" images to the recognition model.

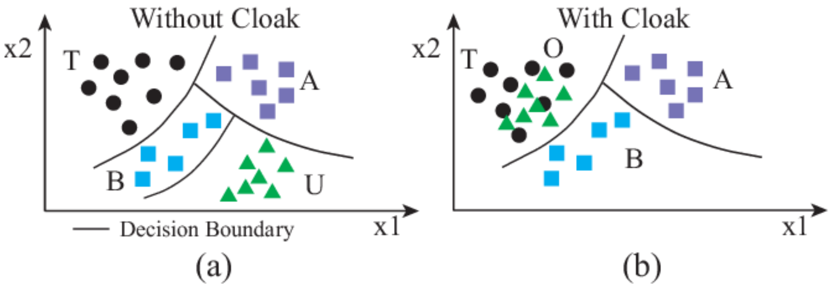


Figure 2 Example of Fawkes system can find the target class on the dataset, even if the dataset has not to target T [1].

Fawkes is a solid system for disabling the identification purpose of the facial recognition system. It converts the images that misled the recognition system into thinking it is another user. Besides, *Fawkes*' latent robustness and transferability are also the crucial factors resulting in its outstanding performance [1]. The authors of *Fawkes* have proposed an issue on *Fawkes's* approach. As mentioned, the *Fawkes* relies on finding the longest distance of class from the class of input images as the target class. How if the target class is not in the features space of the dataset? The authors of *Fawkes* think *Fawkes* can still find another class as the input image's target class. Figure 2 visualizes this assumption. Figure 2a is the typical case, class T is the target class of input class U. Figure 2b is the abnormal case, in which the original target T does not exist in this dataset. In Figure 2b, however, the input class U can still find the target class. Since the original target T's area in the feature space will have another class, input U can still find it through the distance computation. For the transferability, the authors of *Fawkes* have indicated that the *Fawkes* generated images are works, even if the recognition model applied transfer learning to train its feature extractor or train the whole model from scratch.

4. Face-off

Before introducing the Face-off, the two essential factors of this approach have to be briefly introduced: (1) Adversarial Learning and (2) Metric Learning. Adversarial Learning Adversarial Learning is a type of machine learning that aims to modify the input data imperceptibly, resulting in the model misclassifying or misidentifying that input. For example, a normal image can be classified by the classifier. After an adversarial model process the image, the classifier cannot classify it or classify it with a wrong label. Metric Learning is grouping classifying the input class according to the computation distance of each input in the space of metric embedding. Face-off is implemented by extending the adversarial learning model from a typical classification model to metric learning [3].

Face-off provides an image modification to adding imperceptible features to the image so that the facial recognition system cannot identify the people on the image. In fact, it is similar to the Image Cloaking introduced in section 1. Face-off also is not requiring using the evasion attack to modify the images, so the input image still seems normal by human's eyes but misidentified to the recognition system.

As mentioned, the Face-off consists of adversarial learning and metric learning, and they are the main challenges in the Face-off implementation. The main challenges consist of the following [3]: (1) Face-off needs the significant loss function to pull the face on the input image from a cluster that should belong to the user of the face on the input image in the embedding space the recognition systems misidentify the output image. (2) Face-off should show its robustness in transferability, which means that the output image generated by the adversarial model for one targeting model should also be applied to another model if both of the targeting models have a similar purpose.

For challenge (1), Face-off's researching team proposed two new loss functions to generate the images that the recognition system cannot identify correctly. The loss functions are described in the following: (1) It aims to launch the untargeted attack on the metric space, which means finding a minimum perturbation that can make the input's embedding to a closer centroid than the original's embedding. (2) It aims to launch the targeted attack on the metric space, which means finding a minimum perturbation that can make the input's embedding to a centroid of the target label class. In conclusion of these loss function, they both aims to move the original input's far away from its original position in the metric space, the two functions correspond to the cases of un-specified target label and specified target label.

For challenge (2), Face-off's team also proposed using the surrogate face recognition models to achieve transferability. Face-off will find a small multiplicative factor to achieve the amplification. Through the amplification, the result of the adversarial model can be transferred to other target models.

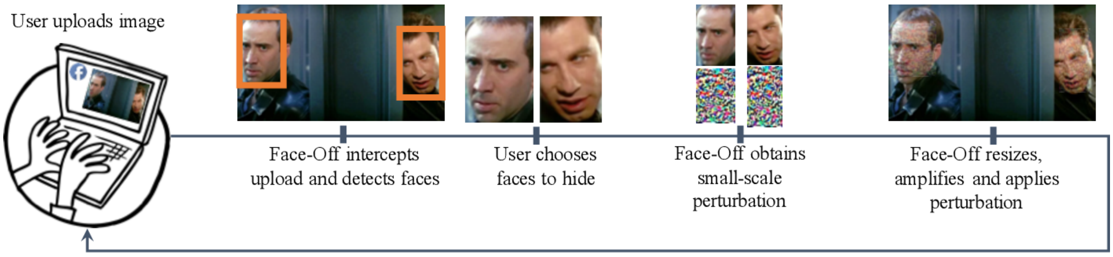

User uploads image
Face-Off intercepts upload and detects faces
User chooses faces to hide
Face-Off obtains small-scale perturbation
Face-Off resizes, amplifies and applies perturbation

Figure 3 Overview of Face-off's operation

Figure 3 shows how Face-off's operation and provides privacy-preserving by achieving image perturbation. Firstly, Face-off will detect the face on the input image because the following perturbation procedures are focusing the faces on the input image. Also, the model of Face-off requires the fixed size of the input, so the extracted faces need to be resized. Then, Face-off will generate a suitable perturbation mask for each resized input face at the pixel level. Lastly, if it is required, the generated masks might require amplification because of model transferring. Otherwise, the generated masks can be applied to the input images and return as the final result.

5. PPRL-VGAN

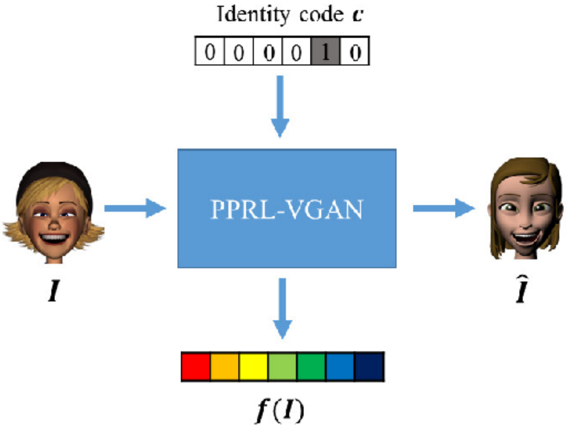


Figure 4 Overview of PPRL-VGAN [2]

The last two approaches are about adding a mask or "cloaking" on the person's face on the image leading the recognition system to misidentify the person. Some researchers have applied another type of approach to generate a new image that the person on the image cannot be identified by the facial recognition systems, extending the Generative Adversarial Network (GAN). GAN is a type of machine learning model. It creates two models: the generator and discriminator simultaneously, and then both of them are used to train each other during the training task (i.e., the generator needs to generate something that the discriminator needs to distinguish whether the input sample is actual or fake. Also, the discriminator returns a loss function to optimize the generator). It aims to achieve specific purposes, which can be either generate something by the trained generator or classify something by the trained discriminator. Thus, researchers proposed an approach by extending typical GAN and combining Variational Auto-Encoder (VAE) to develop a new model called Variational Generative Adversarial Networks (VGANs) that can generate an image from the input image leading the recognition system to misclassify the person on the image, but the same facial expression as the person on the original image [2] (see Figure 4).

VAE consists of the encoder network and decoder network. The encoder network is used to mapping the input data to the latent representation. The decoder network is used to mapping the latent representation to the original format of data. The latent representation from the output of the encoder will be the input of the generator, and the output of the generator will be the input of the decoder. VAE aims to minimize the cost function for the generator to learn the latent representation of input data and achieve realistic data generation. Also, it ensures sufficient diversity in the generated data.

In using VGANs against facial recognition, the generator of GAN is replaced by the encoder-decoder of VAE (see Figure 5). Thus, the encoder converts the input image to latent representation. Then, the decoder generates the face from the latent representation via computation. The computation consists of the target identity code but having the same facial expression as the face on the input image. Therefore, the generated face is different from the original face on the input image to the target recognition system. The discriminator in VGANs can return three classification results: (1) Classify whether the input face image is actual or fake. (2) Identify the person in the input image. (3) Identify the facial expression of the person in the input image.

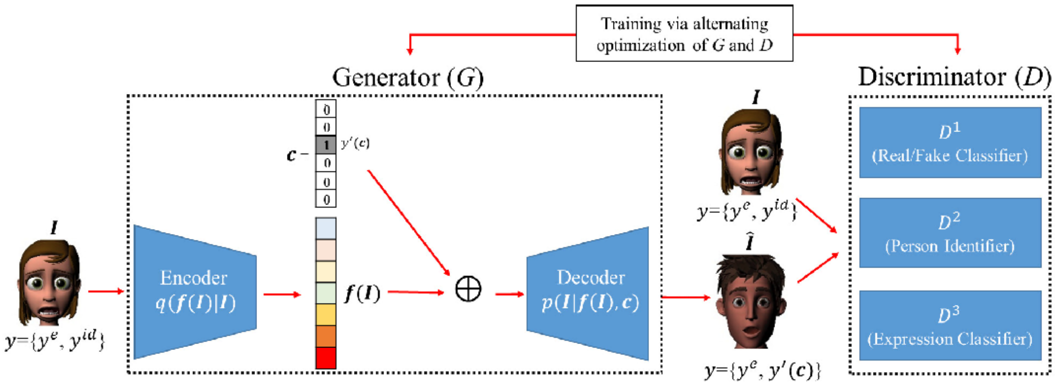


Figure 5 Operation of VGANs [2]

In fact, the authors of this model aim the resulting image to preserve the facial expression but removing the identity. However, its designed discriminator is powerful, which is containing three loss functions. Therefore, the model can be applied to both facial expression recognition and facial identification cases. Figure 5 shows the entire operation of the VGANs. The generator of VGANs first extracts the latent representation of the input face image's facial expression and then uses the target identity code to generate a new face image. The discriminator output three loss function about the generated image from the generator, and it returns the loss functions to the generator achieving the optimization. In final, the well-trained generator can be used to fool the facial expression recognition system that the person on the input image is a real person and have the same facial expression as the original person’s, but the person's identity on the input image depends on the identity code.

6. Privacy-Protective-GAN

The previous section has introduced a VGANs-based approach to fool the recognition system that the person on the input image is a real person but different from the person on the original image. Here is another approach to achieve the face de-identification to the input image, which is also GAN-based. It is called Privacy-Protective-GAN (PP-GAN) [5]. However, it is different from the previous VGAN model. PP-GAN has capable of retaining most of the structure of the face on the input image, but it can de-identify the face on the image. The research team has indicated three objectives of PP-GAN: (1) The generated face image should look real, (2) The generated face is de-identified, it cannot be identified as the same person on the original image by the recognition system, and (3) the generated image should not be too different than the original image on the pixel-level.

PP-GAN is different from the typical GAN model. There are two additional mechanisms to help it to achieve the purpose of face de-identification. First, since the purpose of PP-GAN is to de-identify the face, it will enlarge the range away from the input data in the identity-related embedding feature space. Second, constraining the variation of input image structure, in case the output's structure is too different from the original image.

For the PP-GAN structure, it is also different from the traditional GAN because it needs to achieve those additional mechanisms and face de-identification. Except for the generator and discriminator, PP-GAN also consists of a verificator and regulator. The verificator is used to compute the contrastive loss in order to add the prior identity information in the embedding space to help remove the biometric information. The regulator computes the Structural Similarity Index (SSIM) via the luminance, contrast, and structural similarities of the generator's output. The SSIM is used to improve the generator's ability to maintain similar images.

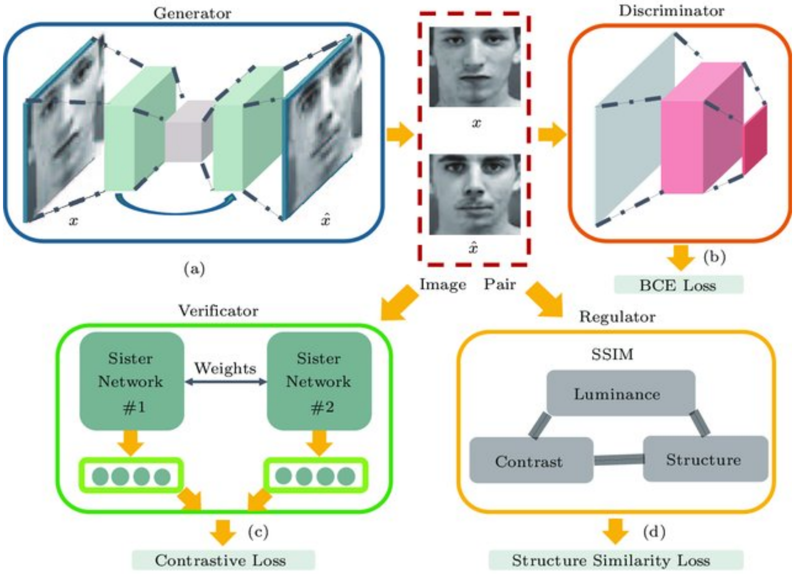
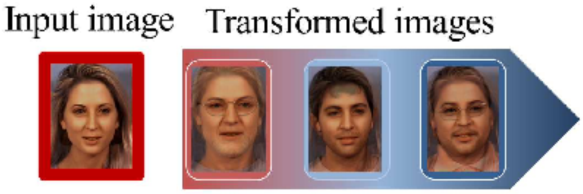


Figure 6 Overview of PP-GAN

The generator and the discriminator are extending the conditional GAN (cGAN). It can fit a conditional data distribution. Firstly, the generator needs to generate a de-identified image from the original input image to challenge the discriminator. For the generator optimization, the verificator computes the verification loss between the original image and the de-identified image in the identity embedding space. Also, the regulator computes the SSIM to evaluate the similarity between the original image and the de-identified image according to the luminance, contrast, and structure. Figure 6 visualizes the operation of the PP-GAN model.

7. Gender Suppressing

Besides preventing facial recognition, some researchers investigate privacy-leakage concerns from personal photos to the soft biometrics' layers, and they proposed the corresponding defending approach. The soft biometric includes age, gender, race, etc. For example, some companies collect the user's photo that aims to know which group of people is the biggest clients in their services by analyzing them so that they can adjust the business strategy or show the advertisements according to the analysis's result. People are resistant to other people using their photos for another purpose, but facial recognition is sometimes crucial. For example, facial recognition can provide an automated photo tagging so that it might help to extend the user's social circle in social media. How to modify a photo that can only be identified by facial recognition but cannot be used to read soft biometrics information? Researchers invented an approach by adding a mask on the face of input images to suppress the gender classification but retain facial identification [4].



Identity	Alice	Alice	Alice	Alice
Gender	Female (confident)	Female (less confident)	Male (less confident)	Male (confident)

Figure 7 An illustration of how the gender suppressing work [4]

The gender suppressing model is shown in Figure 7. The input image is transferred by the gender suppressing model and makes the target classification model misclassify the correct gender of the person in the input image, but the face on the image can still be identified by the facial recognition system correctly. Thus, the suppressing model requires computing the degree of the suppression, and it has to be accurate, as the purpose is only to suppress the gender instead of the other features on the face. The researchers proposed using the automated gender classifier to evaluate the best degree that can suppress the gender of the face on the input image. Since the gender classifier can output a confidence value of the input image representing the possibility of either the person on the image being male or female, in a gender suppressing case, the confidence value or gender prediction result is low so that the target model cannot correctly classify the gender of the person on the input image.

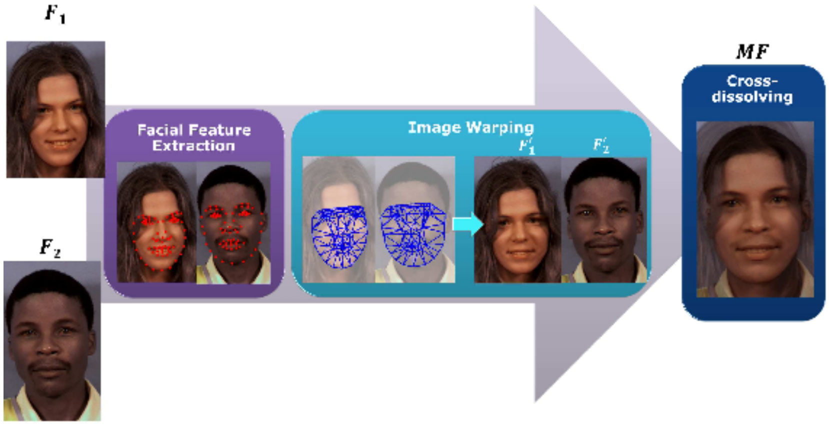


Figure 8 Operation of Gender Suppressing model [4]

In Figure 8, the input image is F_1 and the output image is MF. By mixing the F_1 with another image F_2 to suppressing the gender of the face on F_1. The authors of this model proposed the above implementation by using face morphing, and it consists of three stages (see Figure 8): (1) Facial Feature Extraction, (2) Image Warping, and (3) Cross-dissolving. Firstly, the model will extract the facial features from both F_1 and F_2 images. This task aims to get the control points and they are used to describe the prominent facial features. The controls points are stored in vector format and the following is an example:

$$X_j = \left[x_{1j}, x_{2j}, x_{3j}, \dots, x_{nj}, y_{1j}, y_{2j}, y_{3j}, \dots, y_{nj}\right]^T,$$

where j belongs to the image class (e.g., F_1 and F_2) and n is the number of control points. Moreover, the researchers of this model indicated that the correct control points could effectively minimize the ghosting problem.

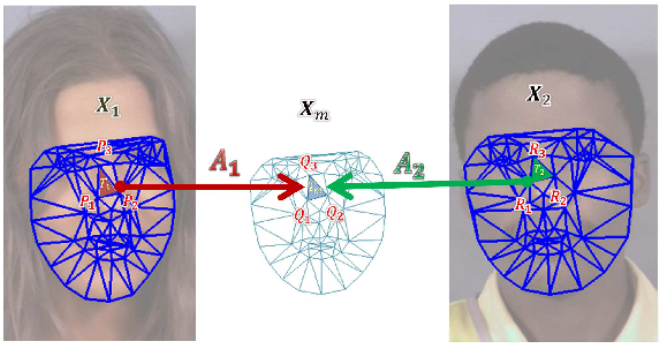


Figure 9 Generating the corresponding triangle according to the intermediate image [4]

Then, the model will perform image warping to both images so that the shape of both faces on the two images can correspond to each other. This stage relies on computing the intermediate image based on both images and using it to warp both images. The model will first determine the intermediate control points set and then using it to dissect the region of the face on the image (see Figure 9). After that, the model utilizes the affine transformation to warp both images according to the intermediate image of both images. Figure 8 shows the result of image warping.

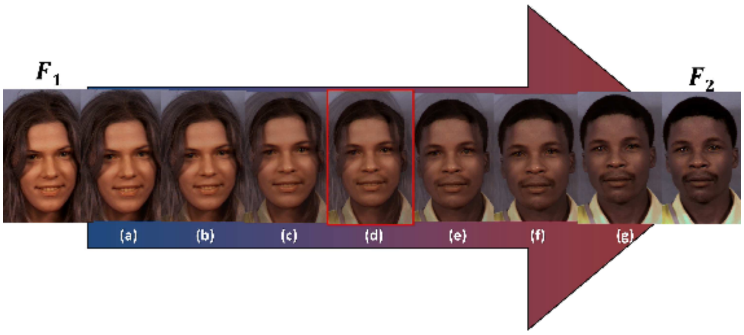


Figure 10 Warped Image cross-dissolving [4]

Finally, both warped images will be cross-dissolved, and it is the output image (see Figure 10).

8. Solution Comparison

The five solutions against privacy leakage issues in facial recognition can perform outstandingly in different aspects. *Fawkes* just relies on adding the imperceptible modification to the image at pixel-level to result in the recognition system misidentifying the person on the image [1]. Face-off can generate a mask to the face on the input image leading the recognition system to misidentify the person [3]. PP-GAN applies the cGAN and extends it by adding the verificator and regulator to generate the image that the image's structure is similar to the original input image, but it is de-identified [5]. VGAN extends the typical GAN by adding the variational encoder and decoder to generate the image keeping the same facial expression but changing the identity [2]. The gender suppressing model is a solution to defend the privacy leakage in the soft biometrics layer. The image processed by the model can keeping the identity but cannot be classified or predicted the gender [4].

In these five approaches, we think that *Fawkes* is the most effective model to protect people's privacy with just a tiny modification to the images. Its implementation is impressive, as well as the results. Moreover, *Fawkes* provides high transferability performance, which can provide over 95% protection rate. Although Face-off and PP-GAN similar results, their design makes them weaker than *Fawkes* in different aspects. The face-off has indicated itself cannot provide well transferability because it is a black-box attack. It means that it does not know the architecture of the target recognition model, but the Face-off model requires an understanding of the model. PP-GAN is an approach relying on the GAN model, and its architecture is complicated. Although it can achieve face de-identification, its modification is visible.

9. Discussion

Although those five solutions are solid and robust, each of them has limitations in different situations.

Fawkes is a powerful model to preventing the privacy leakage problem from the facial recognition system. It applies the cloaking to the people on the target images so that the facial recognition system cannot identify the cloaked people on the image. However, there is a limitation in *Fawkes*. If a recognition system only targets a person and puts as many resources as possible to train the recognition model, *Fawkes* might not provide enough protection. Since the recognition model can be trained to be very complicated, such as the target person's movement, identifying the target person is exceptionally high.

Face-off is a solution that is similar to the *Fawkes*. It relies on generating a mask and placing it on the people's faces on the image to lead the recognition to misidentify the people. Since it is a black-box attack, the protection of the model is according to the structure of the target model. Thus, the transferability of Face-off is low if it does not know the architecture of the target model. The authors of Face-off have suggested that choosing a correct surrogate model can enhance the transferability of the model.

Face-off and *Fawkes* are the same types of approach, so they have common limitations. Since they both need to process the images before uploading to the Internet or social media to prevent the recognition system from identifying the people on the image, the processing might impact the real-time image uploading. Besides, there is a situation that is hard to escape. In social media applications, they always allow the user to tag their friends on the uploaded photos. It gives an opportunity for the recognition model to study the people's identity features. To prevent this issue, the user should turn off the corresponding settings, such as forbidding users to tag me on their images.

The PP-GAN is a solution extending the GAN model and using a verificator and regulator to generate an image de-identified but keeping most of the structure. However, this solution has the limitation that the face image has to be mostly frontal. If there is a deviation in the head pose of the image, it can affect the generating results. Another GAN-based solution is the VGAN model, which aims to generate the image that retains the facial expression of the face on the original image but changing the identity. The recognition system is hard to identify the correct person to the output image from VGAN because the identity is changeable. However, this solution might not be suitable for social media. In social media, people would like to upload their photos and keep the structure of them. If the identities of the people on the image are changed, the users will not use it.

The gender suppressing model is a solution that only suppresses the gender features of the people on the images but retaining the identity of the people. This is impressive for preventing the recognition system from reading the soft biometrics information of the people on the images. However, if the users want the recognition system that also misidentifies the people on the images, this solution does not work.

10. Conclusion

This discussed the privacy leakage problems brought from facial recognition and introduced the three types of approaches that provide corresponding protection in different aspects. The first type is generating a mask for or cloaking the people on the images to make the facial recognition system misidentify them. The second type uses the GAN model to de-identify or changes the people's identity on the target image. The last type provides the privacy-preserving approach in the soft biometrics layer, forbidding the recognition system to classify or predict the soft biometrics information about the people on the target image. In the future, we will keep looking for novel approaches that can provide outstanding protection against the privacy leakage problems from facial recognition systems.

Reference

[1] Shawn Shan and Emily Wenger and J. Zhang and Huiying Li and Haitao Zheng and B. Zhao, "Fawkes: Protecting Privacy against Unauthorized Deep Learning Models," in *USENIX Security Symposium*, 2020.

[2] J. Chen and J. Konrad and P. Ishwar, "VGAN-Based Image Representation Learning for Privacy-Preserving Facial Expression Recognition," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),* pp. 1651-165109, 2018.

[3] Chuhan Gao and V. Chandrasekaran and Kassem Fawaz and S. Jha, "Face-Off: Adversarial Face Obfuscation," *Proceedings on Privacy Enhancing Technologies,* vol. 2021, pp. 369 - 390, 2021.

[4] Asem A. Othman and A. Ross, "Privacy of Facial Soft Biometrics: Suppressing Gender But Retaining Identity," in *ECCV Workshops*, 2014.

[5] Y. Wu and Fan Yang and Yong Xu and Haibin Ling, "Privacy-Protective-GAN for Privacy Preserving Face De-Identification," *Journal of Computer Science and Technology,* vol. 34, pp. 47-60, 2019.

[6] O. Solon, "https://www.nbcnews.com/tech/internet/facial-recognition-s-dirty-little-secret-millions-online-photos-scraped-n981921," NBC News, 12 March 2019. [Online]. Available: https://www.nbcnews.com/tech/internet/facial-recognition-s-dirty-little-secret-millions-online-photos-scraped-n981921.

[7] O. o. t. A. I. Commissioner, "Photos and videos," [Online]. Available: https://www.oaic.gov.au/privacy/your-privacy-rights/social-media-and-online-privacy/photos/.