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