AnonFACES: Anonymizing Faces Adjusted to Constraints on Efficacy and Security

Minh-Ha Le, Md Sakib Nizam Khan, Georgia Tsaloli, Niklas Carlsson and Sonja Buchegger
(KTH)

Image data analysis techniques such as facial recognition can threaten individuals' privacy. Whereas privacy risks often can be reduced by adding noise to the data, this approach reduces the utility of the images. For this reason, image de-identification techniques typically replace directly identifying features (e.g., faces, car number plates) present in the data with synthesized features, while still preserving other nonidentifying features. As of today, existing techniques mostly focus on improving the naturalness of the generated synthesized images, without quantifying their impact on privacy. In this paper, we propose a methodology to (a) quantify the privacy-utility trade-off using an information loss metric, and (b) improve the overall trade-off by utilizing deep learning-based feature extraction and a novel equal-size clustering technique that minimizes the information loss. To the best of your knowledge, this is the first work to define and quantify this privacy utility trade-off. We also note that the importance of embedding and clustering algorithms has mostly been neglected. This is an important observation since our results demonstrate that we can achieve significant improvements in the privacy-utility trade-off by better clustering similar images and by exploiting non-linear relations between privacy and utility. By incorporating StyleGAN, a state-of-the-art Generative Neural Network, our model produces more realistic synthesized faces than prior works. Finally, we note that an example benefit of these improvements is that our solution allows car manufacturers to train their autonomous vehicles while complying with privacy laws.