Anonymized person re-identification in surveillance cameras

Arthur van Rooijen, Henri Bouma[*], Raimon Pruim, Jan Baan, Wouter Uijens, Jelle van Mil

TNO, Oude Waalsdorperweg 63, 2597 AK Den Haag, The Netherlands

ABSTRACT

Person re-identification (Re-ID) is a valuable technique because it can assist in finding suspects after a terrorist attack. However, the machine learning algorithms for person Re-ID are usually trained on large datasets with images of many different people in a public space. This could pose privacy concerns for the people involved. One way to alleviate this concern is to anonymize the people in the dataset. Anonymization is important to minimize the storage and processing of personal information, such as facial information in a surveillance video. However, anonymization typically leads to loss of information and could lead to severe deterioration of the Re-ID quality. In this paper, we show that it is possible to store only anonymized person detections while still achieving a high quality person Re-ID. This leads to the conclusion that for the development of re-identification algorithms in situations where privacy is of great importance it is not necessary to store facial information in person re-identification datasets.

Keywords: Re-identification, anonymization, surveillance, CCTV, privacy enhancing technologies.

1. INTRODUCTION

Person re-identification (Re-ID) is the technique that associates images of the same person from different cameras. It is a valuable technique because it can assist in finding suspects after an incident, for example a terrorist attack. However Re-ID can also be useful in preventing an incident like a terrorist attack from occurring. Several weak indicator of a potential attack might be measured over time and can be combined to a stronger signal using Re-ID. For example by coupling the results of an explosive detector and a firearm detector, neither of which is perfectly reliable on its own.

However the machine learning algorithms for person Re-ID are usually trained on large datasets with images of many different people in a public space [Rooijen, 2018]. This could pose privacy concerns for the people involved. One way to alleviate this concern is to anonymize the people in the dataset. Anonymization is important to minimize the storage and processing of personal information, such as facial information in a surveillance video. Just storing encoded information (e.g., machine readable Re-ID descriptors) rather than raw images could be sufficient for some applications. However if the images are needed to facilitate retraining of the Re-ID to improve the system over time or to verify of the results, storing the encoded information is not enough.

There is a tension between Re-ID and anonymization. Good anonymization requires the removal of identity-related information whereas Re-ID algorithms in fact require discriminative information between persons. In this paper, facial information is removed and similarity between persons is determined based on non-facial appearance characteristics. This enables the development of tools to track people from one camera to neighboring cameras, or to trace a suspect after an incident, under the condition that their (non-facial) appearance is similar over time. However, it prohibits the identification of people and it is less suitable to couple to unique personal information, or establish links from one day to another day in case the appearance changes (e.g., different clothing). Therefore, this approach is not easily modified to allow for the creation of white lists (e.g., to permit access for employees) or black lists (e.g., to detect criminals).

Others already showed that re-identification can reach a high Rank-1 accuracy of 95.4% [Luo, 2019] on the Market-1501 dataset [Zheng, 2015]. However, they did not show the effect of anonymization. In this paper, we show that it is possible to store only anonymized person detections while still achieving high quality person re-identification.

[*] E-mail: henri.bouma@tno.nl, phone; +31 888 66 4054

A. van Rooijen, H. Bouma, R. Pruim, J. Baan, W. Uijens, J. van Mil, "Anonymized person re-identification in surveillance cameras," Proc. SPIE, vol. 11542, (2020). https://doi.org/10.1117/12.2571943

The outline of this paper is as follows. Section 2 describes our method for detection, tracking and anonymized re-identification. Section 3 presents the experiments and results. Finally, Section 4 summarizes the conclusion.

2. METHOD

2.1 System architecture

The system proposed in this paper aims to track people from one camera to neighboring cameras in a local environment. The complete system architecture is shown in Figure 1. The architecture consists of person detection, anonymization, signature computation, tracking, and matching. The camera generates a video stream, which consists of a sequence of frames. Person detection is the step to localize people in the frames [Marck, 2014]. Detection is performed using the SSD detector [Liu, 2016] with a ResNet50 backbone [He, 2016], and a pre-trained model from the TensorFlow detection Model Zoo[†]. Each detected person is extracted from the frame and transferred to the anonymization module. The extracted person is also called a snippet. Anonymization removes the facial information from the snippet (see Sec. 2.2). Signature computation extracts a descriptor from the anonymized snippet. Tracking combines multiple snippets in subsequent frames over time and thus generates a track. A track consists of multiple observations of the same person. Finally, the signatures in a track from one camera are compared with the signatures in other tracks from another camera. The matcher uses the similarity between signatures to decide whether a person in one camera is the same as a person in another camera. In this paper, we will focus on anonymization and re-identification. The method is evaluated on a dataset that contains snippets where the persons are already detected and extracted from the complete video.

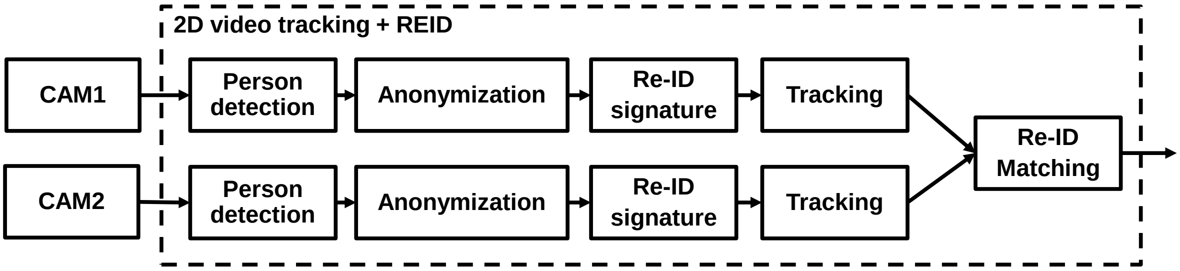
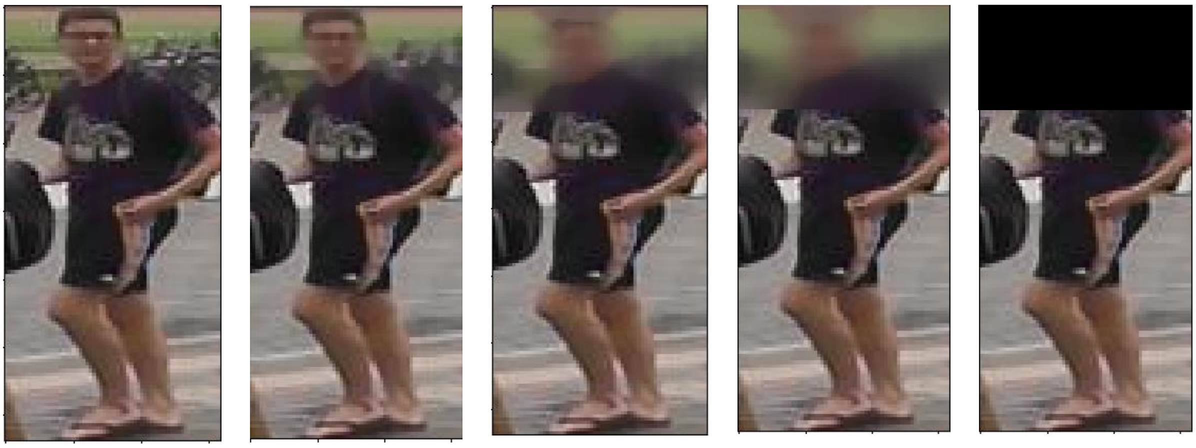


FIGURE 1: SYSTEM ARCHITECTURE.

2.2 Anonymization

We used the following approach for anonymization. Anonymization was performed by applying a Median filter from the Python Imaging Library (PIL)to images of 256x128 pixels. The median filter is a non-linear filtering, which is required to avoid inverse operations. The top 25% of the image was filtered. We investigated the effect of filtering with three different kernel sizes (3, 9 and 15 pixels) and complete removal of information.

[†] https://github.com/tensorflow/models/blob/v1.13.0/research/object_detection/g3doc/detection_model_zoo.md



0 px 3 px 9 px 15 px Black out

FIGURE 2. EXAMPLE IMAGE (256X128 PIXELS) FROM THE MARKET-1501 DATASET FILTERED WITH KERNEL SIZES: 0 (NO BLUR), 3, 9, 15 PIXELS AND TOP FULLY BLACKED OUT.

2.3 Re-identification

For re-identification, we used the “strong baseline” implementation of Lue et al. [Luo, 2019] because it achieves a high performance and allows us to show the principle of balancing accuracy with anonymization. Re-identification algorithms typically consist of two steps: signature computation and matching. Signature computation is computationally the most expensive step, which generates a descriptor of every snippet using a deep neural network (ResNet50 [He, 2016]). Matching computes a distance between pairs of signatures which is used to determine the identity dissimilarity. We use the Euclidean distance as a dissimilarity metric, but other options are also available such as the cosine distance.

3. EXPERIMENTS AND RESULTS

The aim of our experiments is to evaluate the effects of anonymization on the re-identification accuracy. Evaluation of person detection and tracking is not in scope of this paper. In the experiments, we use the Market-1501 dataset [Zheng, 2015], which contains more than 20.000 snippets of 1501 labeled persons. This dataset is commonly used for re-identification. The default dataset division is used, which is displayed in Table 1. For a more thorough explanation of the dataset configuration we refer to the paper of the original authors [Zheng, 2015].

TABLE 1. MARKET-1501 DATASET CONFIGURATION.

Subset	# ids	# images	# cameras
Train	750	12936	6
Query (test)	751	3368	6
Gallery (test)	751	15913	6

We use the “strong baseline” method for re-identification [Luo, 2019]. First, we train and test on the dataset without anonymization. Second, we train and test on the dataset with various degrees of anonymization, using kernel sizes for median filtering of respectively 3, 9 and 15 pixels. Last, we train and test on the dataset with full anonymization by completely painting the top part of the image black, thereby removing all available information in the top region. The results are shown in Table 2. The table shows that with the most extensive anonymization (i.e., painting the top of the image black), the mAP is only 6.9% lower and the Rank-1 result is only 3.3% lower than without anonymization. If we use blurring instead of blacking out we see that the accuracy improves. The results show that the kernel size acts as a trade-off parameter between anonymization and accuracy. This leads to the conclusion that re-identification can be done with high accuracy on anonymized images. However, there exists a trade-off between the obtained accuracy and the severity of the anonymization.

TABLE 2. RE-IDENTIFICATION PERFORMANCE WITH AND WITHOUT ANONYMIZATION.

Method	Kernel Size (pixels)	Kernel Size (percentage of image width)	mAP	Rank 1	Rank 5	Rank 10
Full image	-	-	85.1	93.6	98.0	98.8
Top blur	3 px	2.3%	84.6	93.8	98.2	99.0
Top blur	9 px	7.0%	82.5	92.7	97.7	98.3
Top blur	15 px	12.7%	80.9	91.6	97.4	98.5
Black out	-	-	78.2	90.3	96.7	98.0

4. CONCLUSIONS

In this paper, we showed that anonymization of person detections minimally affects the quality of person re-identification. This leads to the conclusion that for the development of re-identification algorithms in situations where privacy is of great importance it is not necessary to store facial information in person re-identification datasets.

5. DISCUSSION

The proposed method anonymizes the face, so the Re-ID algorithm can only use the information related to non-facial characteristics of the persons’ appearance. The results show that anonymization of the face hardly decreases the accuracy. This implies that “strong baseline” [Luo, 2019] – and probably many other current state-of-the-art Re-ID algorithms – are (also) under-utilizing facial features. The focus on non-facial features leads to the following limitations.

The first limitation is that the person-of-interest most probably will be lost when he/she changes clothes. This is a small problem in case of neighboring sensors, but it might become a bigger problem when tracking in a larger area or over longer time.

The second limitation is that the current approach may focus on the dominant color of the shirt and the dominant color of the pants. This might be a problem with very common clothing choices (e.g., at locations where people wear a uniform, or at a business convention where half the participants wear the same navy-blue suit). This may be resolved by a coarse-to-fine triplet reranking approach [Katsaros, 2020], where minor unique differences from the same viewing angle are emphasized to improve the matching.

Current state-of-the-art Re-ID algorithms appears to under-utilize the facial features. The limited added value of facial information in our experiments may be caused by the low-resolution images in the used dataset. Anonymization removes facial information, while Re-ID algorithms could potentially benefit from higher-resolution images and face-recognition technology. Future work could include the use of higher-resolution images and the investigation of alternative strategies, such as homomorphic encryption [Erkin, 2012] or federated learning [Li, 2020], to enhance privacy and facilitate multi-camera tracking and re-identification.

ACKNOWLEDGEMENTS

The INSTEAD (“INtegrated System for Threats EArly Detection”) project has received funding in the DEXTER (Detection of EXplosives and firearms to counter TERrorism) program from NATO Science for Peace and Security (SPS) under grant agreement number G5605.

REFERENCES

[1] K. He, X. Zhang, S. Ren, J. Sun, “Deep residual learning for image recognition,” IEEE CVPR, 770-778 (2016).

[2] Z. Erkin, T. Veugen, T. Toft, R. Lagendijk, “Generating private recommendations efficiently using homomorphic encryption and data packing,” IEEE Trans. information forensics and security 7(3), 1053-1066 (2012).

[3] E. Katsaros, H. Bouma, A. van Rooijen, E. Dusseldorp, “A triplet-learnt coarse-to-fine reranking for vehicle re-identification,” ICPRAM, (2020).

[4] T. Li, A. Sahu, A. Talwalkar, A., V. Smith, V., “Federated learning: Challenges, methods, and future directions,” IEEE Signal Processing Magazine 37(3), 50-60 (2020).

[5] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, A. Berg, “SSD: Single shot multibox detector,” ECCV, 21-37 (2016).

[6] H. Luo, Y. Gu, X. Liao, S. Lai, W. Jiang, “Bag of tricks and a strong baseline for deep person re-identification,” IEEE CVPR Workshop, (2019).

[7] J. Marck, H. Bouma, J. Baan, J. de Oliveira-Filho, M. van den Brink, “Finding suspects in multiple cameras for improved railway protection,” Proc. SPIE 9253, (2014).

[8] A. van Rooijen, H. Bouma, F. Verbeek, “Fast and accurate person re-identification with Xception Conv-Net and C2F,” CIARP, (2018).

[9] L. Zheng, L. Shen, L. Tian, et al., “Scalable Person Re-identification: A Benchmark,” IEEE ICCV, (2015).