Differential Privacy and *k*-Anonymity for Pedestrian Image Data: Impact on Cross-Camera Person Re-Identification and Demographic Predictions

LUCAS MARIS, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan
YUKI MATSUDA, Okayama University, Okayama, Japan and RIKEN Center for Advanced Intelligence Project, Chuo-ku, Japan
KEIICHI YASUMOTO, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan and RIKEN Center for Advanced Intelligence Project, Chuo-ku, Japan

Video cameras are prevalent in large cities but their use outside of public safety remains limited due to legitimate privacy concerns. Nevertheless, the rich information they can capture appears incredibly promising for large-scale smart city applications, as they can function as very powerful and versatile sensors. This ambivalence raises the question of whether such image data can be used in a privacy-responsible manner. Encryption-based solutions assume the end server can be trusted with keeping data safe; data leaks show us this assumption does not necessarily hold true. Traditional image obfuscation methods such as pixelization or blurring on the other hand fail to offer both sufficient privacy and utility. As such, privacy approaches that can provide privacy protection directly on the data itself while retaining practical utility are required. We here extend two such notions, differential privacy and *k*-anonymity, to image data, and extensively evaluate the resulting privacy-utility tradeoff on cross-camera person re-identification and attribute recognition data. Our results show that our proposed approaches can significantly reduce the privacy-sensitivity of image data at source while retaining decent utility for vision-based smart city applications.

CCS Concepts: • **Security and privacy** → **Data anonymization and sanitization**; • **Human-centered computing** → *Ubiquitous and mobile computing design and evaluation methods*; • **Computing methodologies** → Image representations;

Additional Key Words and Phrases: Image Differential Privacy, Image K-anonymity, Person Re-identification, Demographic Predictions

ACM Reference format:
Lucas Maris, Yuki Matsuda, and Keiichi Yasumoto. 2025. Differential Privacy and *k*-Anonymity for Pedestrian Image Data: Impact on Cross-Camera Person Re-Identification and Demographic Predictions. *ACM Trans. Cyber-Phys. Syst.* 9, 4, Article 36 (October 2025), 31 pages.
https://doi.org/10.1145/3743680

This study was supported in part by Sponsor JSPS KAKENHI JP21H03431, JP24K20763, and JP25K03107.
Authors' Contact Information: Lucas Maris, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan; e-mail: lucas.maris.lo3@is.naist.jp; Yuki Matsuda, Okayama University, Okayama, Japan and RIKEN Center for Advanced Intelligence Project, Chuo-ku, Japan; e-mail: yukimat@okayama-u.ac.jp; Keiichi Yasumoto (corresponding author), Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan and RIKEN Center for Advanced Intelligence Project, Chuo-ku, Japan; e-mail: yasumoto@is.naist.jp.

ACM 2378-9638/2025/10-ART36
https://doi.org/10.1145/3743680

1 Introduction

Cities are rapidly evolving, with an ever-growing number of citizens flowing to them and an aspiration to modernize the way they function. Consequently, the interest in smart cities, which exploit data collected by large networks of ubiquitous sensing devices to help improve the city's inner functioning, is peaking, with both large- and small-scale projects being developed around the globe [35]. Most of these projects closely associate the concept of smart cities with the use of a wide array of IoT devices [57], allowing large amounts of data to be collected and enabling, e.g., policymakers to make informed urban planning decisions [52], transport or tourism companies to achieve commercial success [34], or individuals to plan routes with high context awareness [31].

This wide variety of applications largely depends on comprehensive data collection regarding human flows within a city. One type of device that is already prevalent within cities are video cameras. While ordinarily installed as a dissuasive safety device, for use by security services for real-time or retrospective surveillance, their feeds carry incredibly rich information, and thus have much potential for use in typical smart city applications. This however stumbles on particularly valid concerns regarding people's right to privacy. Taken together, these issues form the privacy-utility compromise: How much can visual data be protected while still retaining practical value for smart city applications? Camera-based smart city applications rely on standard computer vision tasks; we here focus our attention on two such tasks, cross-camera person **re-identification (reID)** and demographic predictions, which together are expected to enable route-based smart city applications.

While existing methods such as encryption and traditional image obfuscation methods (e.g., pixelization or blurring) can offer a degree of privacy protection, they may be impractical or insufficient in actual smart city scenarios. Encryption, for instance, assumes the end server can be trusted with the data; the reality is that the data processor may not be trusted by users, or that it may be trusted but susceptible to data leaks [8]. Due to their computational complexity, cryptographic solutions to this issue such as homomorphic encryption remain impractically slow [73], especially for visual data. Pixelization or blurring does not always suffice to guarantee anonymity, as original images have been shown to be recoverable from their obfuscated counterparts [30, 47].

In these circumstances, research on other forms of privacy has flourished over the last decades. Concepts such as differential privacy [21] or k-anonymity [56] aim to model different ways in which a practical equilibrium between privacy and utility can be achieved, but their initial application domain is databases, and there is limited research on their suitability for image data. To bridge this gap, we here broaden the analysis of our previously introduced **image differential privacy (IDP)** mechanism [46], and additionally introduce and analyze a novel non-uniform body part segmented **class-activation mapping (CAM)**-based IDP mechanism. Furthermore, we propose the first extension of k-anonymity to the image domain, and confirm that our proposed IDP mechanisms successfully increase privacy from a k-anonymity perspective.

We make the following contributions:

- We extend the analysis of our recently introduced IDP mechanism with extensive experiments on two large person reID datasets and one facial attribute recognition dataset.
- We introduce a novel IDP mechanism combining an existing body part segmentation model with CAM to better target privacy-sensitive areas within images through non-uniform noising and partial inpainting.
- We newly introduce a method to empirically compute the k-anonymity of an image set and discuss the relationship between its anonymity factor k and IDP's ε.
- We show that our proposed mechanisms outperform traditional image obfuscation methods, providing data with a quantifiable privacy budget ε resulting in a higher anonymity factor k while retaining practical utility, and identify and discuss privacy-utility tradeoff points.

—We conduct an online user survey to collect the perceived acceptability of images protected through our proposed IDP mechanism, and find over half of the respondents are satisfied with the provided privacy protection.

2 Related Work

This section focuses on privacy definitions and methods, specifically those related to images, and introduces cross-camera person reID with a focus on existing privacy-aware studies.

2.1 Privacy

Privacy has grown to a major concern over the last years, which has led to many studies striving to define, protect, or improve privacy. We briefly introduce the concepts most relevant to our work.

2.1.1 ***Differential Privacy (DP).*** Since its introduction by Dwork et al. [20, 21], differential privacy has become an incredibly popular way to model formal data privacy, able to provide quantifiable privacy guarantees for statistical data releases, providing individuals roughly the same privacy that would result from having their data removed. Its robustness against arbitrary, unforeseen attacks has caused the concept to be utilized by Google [22], Apple [6], Microsoft [18], the U.S. Census Bureau [2], and SAP [33]. Initially introduced for use on statistical databases, recent research has focused on extending DP's desirable properties to other forms of data, such as location data [5] or deep learning models [1].

2.1.2 k-Anonymity. As formalized by Sweeney [56], k-anonymity is a property of a dataset that characterizes the re-identifiability of its data records. A given set of data is said to be k-anonymous if each data record is made indistinguishable from at least $k-1$ other data records in terms of a given set of *quasi-identifiers*, essentially setting an upper bound of $1/k$ to each data record's reID probability. While still widely used today, k-anonymized data is known to be susceptible to background knowledge attacks [45, 64], specifically because k-anonymity makes the assumption there exists a clear distinction between quasi-identifiers and non-identifying attribute values, which is not necessarily true.

2.2 IDP

With unstructured data forming the largest part of today's data landscape, multiple studies have investigated the possibility of applying differential privacy to images. While a consensus has yet to emerge on a universal definition for IDP, the common approach in these studies has been to vectorize unstructured data into a structured data form [76], which can then be obfuscated with conventional DP methods. The foremost issue is that image data is inherently high-dimensional; however, to satisfy differential privacy, noise must be scaled to data sensitivity, i.e., just how different two images can be. As different studies tackle this problem in different ways, research on IDP extends in a variety of directions.

2.2.1 Pixel-Level IDP. Pixel-level IDP focuses on perturbing pixel values in a manner that satisfies differential privacy. This implies the data sensitivity varies directly with the size and number of color channels within the image. The first extension of differential privacy to the image domain limits data sensitivity by introducing a pixelization step, prior to differential privacy noise addition, and by working exclusively with grayscale images [23]. A couple studies look into treating grayscale images as data streams, which can then be traversed via a sliding window procedure: Liu et al. [43] use this idea to allocate privacy budget dynamically to different areas in the image but remains obtuse as to how it computes data sensitivity, and [37] uses the same idea to perform neighboring pixel merging but equates local sensitivity (at specific positions in the data stream) to

global sensitivity (throughout the whole data stream), which raises questions about the resulting differential privacy guarantee. Another study introduces a sampling method [49] adapted from [59], which samples and releases a certain number of pixels of the original image and then interpolates the remaining pixels. The IDP mechanisms proposed in this article also fall into the family of pixel-level IDP methods.

2.2.2 Generation-Based IDP. Generation-based IDP relies on the use of a pretrained autoencoder model or generative adversarial network to compress images into a latent vector representation. This reduces the data sensitivity to the sensitivity of the latent vector representation, which can then be obfuscated in a way that satisfies DP, and be fed back to the pretrained model to generate an obfuscated image. Where most existing studies differ is how they calculate the sensitivity of the latent vector, which is real-valued, and therefore not as unambiguously defined as with bound pixel values. Some studies suggest forcing the latent vector into a specific format [10, 75], where the latent vector represents attributes in a binary manner, but only focus on noising a select subset of these attributes, implying only partial differential privacy guarantees. Most studies simply set the sensitivity to some empirical value [42, 61, 72], as observed on a given dataset; this works for protecting that dataset, but is not guaranteed to satisfy differential privacy for new data, which may not fit the distribution of the dataset that was used for empirical sensitivity calculation. This issue can be solved elegantly by introducing a clipping step [40, 53], which forces outlier latent vector values into an acceptable range of values, as observed from a given dataset; this then bounds the latent vector's sensitivity, and provides sound differential privacy, but increases the tendency of generative models to leak information about the training images through their output images.

2.2.3 Metric IDP. To circumvent the sensitivity issue, a line of work has emerged that focuses on applying metric differential privacy [11] rather than pure differential privacy to images. Instead of calibrating noise to fixed data sensitivity, in order to provide the same level of privacy protection to all inputs, metric privacy calibrates noise to the distance between inputs according to a predefined distance function, which provides similar privacy for similar inputs. The first metric IDP study used **singular value decomposition (SVD)** [24] for this purpose, where noise addition occurs at the singular vector level, and the invertibility of SVD is used to reconstruct an image afterwards. This idea is also used in [66], which uses a high-dimensional space instead of SVD. Some generation-based IDP studies [12, 16] also operate under metric differential privacy. As metric differential privacy does not offer the same guarantees as pure differential privacy, their privacy parameters do not line up, making them hard to compare in practice.

2.3 Cross-Camera Person reID

Cross-camera person reID is a computer vision task that is concerned with matching together snapshots from individuals across different points of view [68]. While still a matter of designing informative appearance signatures for individuals [26], the advent of deep convolutional neural networks [36] has shifted this problem to yet another learning problem. It is traditionally formalized as an image retrieval task, where the model is to rank all *gallery* images in order of similarity to a given *query* image, having been trained on a *training* set composed of disjoint identities. This widely accepted definition skims over the object detection, tracking, and segmentation components that are involved in end-to-end reID frameworks to focus more heavily on how to transform images into effective vector representations [9, 26], and on how to score such representations against one another. These two concerns are usually modeled in terms of a backbone, a feature extractor tasked with transforming images into numerical vectors, a machine learning network, which transforms these vectors into identity-specific embeddings, and its appropriately chosen loss function. Provided images can be transformed into meaningful identity-embeddings, pairwise

distance between such embeddings has proven an effective way to build reID models [70]. The backbone conventionally consists of a pretrained deep CNN model, most commonly a **residual neural network (ResNet)** model [28], typically pretrained on ImageNet [51]. The subsequent network is usually a simple feed-forward network trained under Triplet Loss [29], which helps learning semantically meaningful identity features. A large body of research has been concerned with improving onto this base framework, with recent studies aggregating information from multiple person images [62], leveraging spatial-temporal data [58], exploring large-scale unsupervised pretraining [27], exploiting attention mechanisms [14], learning specific body-part representations [54], eliminating clothing bias [67], or focusing on model explainability [13], among others.

2.4 Privacy-Preserving reID

Some research has concerned itself with privacy-preserving reID, with varying focus areas for privacy-preservation: private data collection, private network communications, or private data storage. On the data collection side, various studies explore replacing classic RGB cameras with sensors deemed less privacy-invasive, e.g., continuous-wave radars [25], event cameras [4], or LiDARs [48]. On the network side, privacy-preserving reID focuses on developing frameworks with inherent privacy guarantees, e.g., through encryption [15] or federated learning [63]. On the data storage side, the main goal is to transform data to reduce its inherent sensitive information while retaining utility for reID, e.g., by blurring faces [17], through encoder-based obfuscation [74], through adversarial perturbations [55] or through encryption [69]. Within these data storage approaches, some studies also provide a way for trusted parties to reverse obfuscation and retrieve the original images, through decoder-based recovery [55, 74] or decryption [69]. To the best of our knowledge, the effect of applying a differential privacy mechanism directly onto RGB image data, which then falls in this third category of privacy-preserving reID, has not been explored yet.

3 Privacy Methods and Definitions

This section summarizes our previously proposed ε-IDP mechanism and how it differs from existing pixel-level differential privacy, introduces a novel segmented CAM-based approach to ε-IDP, and details the motivation and the specifics of our proposed image k-anonymity evaluation procedure. A brief discussion on the runtime performance of these different methods is included in Appendix A.

3.1 ε-IDP Mechanism

Our previous study [46] extended Fan's pixel-level IDP mechanism [23].

Definition 3.1 (ε-IDP). A randomized mechanism $\mathcal{M}$ gives ε-IDP if for any two images i and j of same dimension, and for any possible output $R \subseteq \text{Range}(\mathcal{M})$:

$$\Pr[\mathcal{M}(i) \in R] \leq \exp(\varepsilon) \Pr[\mathcal{M}(j) \in R]. \tag{1}$$

The Laplace mechanism can provide such a guarantee [20], if $\mathcal{M}$ is defined such that the random noise n is calibrated to a chosen privacy budget ε and to the ℓ_1-sensitivity Δf of function f:

$$\mathcal{M}(x) = f(x) + n, \text{ where } n \sim Laplace\left(0, \frac{\Delta f}{\varepsilon}\right). \tag{2}$$

Fan's mechanism defined this function as the pixelization of input images, and made two main assumptions to keep the magnitude of its ℓ_1-sensitivity in check: (1) images are exclusively grayscale, and (2) the sensitive information within a given image is expected to cover at most m of its pixels, i.e., erasing these m pixels suffices to protect said image (this is defined as the m-neighborhood notion). Multiple subsequent studies [24, 40, 76] have argued this definition too strong, introducing too large amounts of both pixelization and random noise and therefore destroying too much of the

data's utility. We here argue the opposite, as both of its underlying assumptions severely restrict the application range and limit the practical usability of the mechanism. We therefore relax both of these assumptions, expecting that: (1) images are RGB, and (2) the sensitive information within a given image can be expected to cover *all* of its pixels.

Relaxing these two assumptions naturally leads to much higher ℓ_1-sensitivity. We counteract this by introducing a second dimensionality reduction step, beyond the pixelization step proposed by Fan. This color quantization step drastically reduces possible color values for pixels, thus decreasing sensitivity by the same factor, but is expected to have a limited effect on data utility, as the RGB color model defines a large number of fine-grained color nuances that can be expected to be redundant from an image-processing perspective. As such, we redefine the function f at the core of the ε-IDP mechanism as the identity function on RGB images of width w and height h, with optional pixelization determined by b and optional color quantization determined by c. The ℓ_1-sensitivity of this function, or maximum difference between two given images, then becomes:

$$\Delta f = \frac{wh}{4^b}\left(\frac{2^8}{2^c} - 1\right)^3. \tag{3}$$

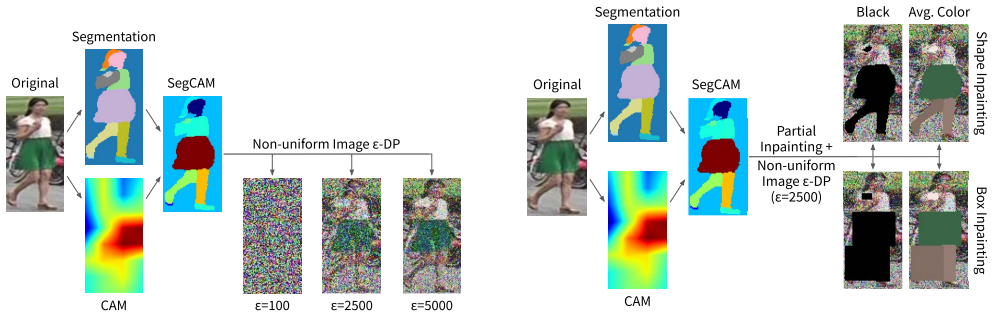
Pixelization is applied such that 4^b pixels are reduced to a single pixel, and color quantization reduces the range of color channels from its original 8 bits to $(8-c)$ bits. Not using either pixelization or color quantization is represented by parameter values $b = 0$ and $c = 0$, and has the ℓ_1-sensitivity function simplified to that of the identity function. Applied to a set of images, this mechanism sets an upper bound of ε to the privacy loss incurred by the individuals represented in this set. Provided the original data is discarded, this privacy loss cannot be increased, i.e., the images cannot be made less private, no matter what auxiliary information might be available. This key property of differential privacy is referred to as post-processing [21].

3.2 Segmented CAM-Based ε-IDP Mechanism (SegCAM-IDP)

We additionally introduce a non-uniform noising mechanism, which uses a human body part segmentation model prior to the noise addition mechanism. Whereas our ε-IDP mechanism adds noise to all pixels of an image equally, the aim here is to distinguish degrees in private information within an image prior to obfuscation, and to use this information to obfuscate accordingly. By concentrating obfuscation efforts on specific areas of an image, to maximize or minimize local impact, we expect to be able to target "desirable" and "undesirable" tasks more efficiently.

Figure 1(a) illustrates our basic workflow for segmentation-based obfuscation: (1) each image is ran through a CE2P human parsing model [50] which segments the image into 20 distinct body parts; (2) each image is ran through a pretrained gender classification CNN with CAM [78] to identify the areas most/least crucial to gender discrimination; (3) the segmentation and CAM information are integrated in order to obtain the average CAM-value per body part area (SegCAM); (4) an IDP mechanism uses this integrated information to do non-uniform noise addition to the image. In essence, this will result in high noise addition for body parts that are deemed unimportant to gender classification, while important body parts will suffer lower noise addition. In this manner, we expect to maximize the suitability of obfuscated images for the gender classification task; maximizing the suitability of obfuscated images for any other task would be possible by using CAM information relevant to that task. However, given that CAM is a relatively computationally expensive process, it is quite difficult to implement on top of inherently computationally expensive tasks such as reID, which is why we here focus on gender-specific CAM information.

This segmented CAM-based obfuscation satisfies IDP. Indeed, provided that the SegCAM information is normalized prior to noise addition, such that the sum of all pixels' weights w_{ij} is equal to 1, we can define the noise added to each pixel x_{ij} as $w_{ij}\varepsilon = \varepsilon_{ij}$. Since $\sum w_{ij} = 1$, and



(a) SegCAM-IDP. (b) SegCAM-IDP with partial inpainting.

Fig. 1. Visualization of the proposed segmented CAM-based IDP mechanism.

through differential privacy's sequential composition property, this leads to $\sum \varepsilon_{ij} = \varepsilon$. In other words, instead of all pixels being noised equally, each pixel value x_{ij} is now perturbed as follows:

$$\mathcal{M}_{\text{SegCAM}}(x_{ij}) = x_{ij} + n_{ij}, \text{ where } n_{ij} \sim Laplace\left(0, \frac{\Delta f_{ij}}{\varepsilon_{ij}}\right). \tag{4}$$

To further exploit the SegCAM information, we also look into partial inpainting for body parts deemed important to gender classification, as illustrated by Figure 1(b). Any (non-background) body part whose average CAM-value is higher than $\frac{1}{wh}$ (the expected CAM-value in situations where all body parts would be equally important) is inpainted. We consider four types of inpainting: black in the shape of the body part, average body part color in the shape of the body part, black box around the body part, and average body part color box around the body part. Non-uniform noise addition is performed on the non-inpainted areas (after normalization of the non-inpainted pixels' SegCAM weights), such that they satisfy IDP. Additional visual examples of images protected with SegCAM-IDP are included in Appendix B.

3.3 Image k-Anonymity

The main obstacle with applying k-anonymity directly onto images resides in its poor generalization ability to high-dimensional data [3], which images inherently are. Indeed, plain k-anonymity essentially states that for any possible combination of values a set of *quasi-identifiers* can take on, there exist at least k samples satisfying this combination. This then guarantees that knowing the values for the quasi-identifiers of a given sample does not suffice to uniquely identify that sample. The ability of samples to "hide in the crowd," as k-anonymity is often described, requires a "crowd," i.e., a set of data whose size is magnitudes larger than the number of possible value combinations for the quasi-identifiers. One can freely define what attributes are to be considered quasi-identifiers, but it is reasonable to assume that the size of this set grows jointly with the overall number of attributes. An RGB image of width w and height h consists of $3wh$ different numerical features, each of which can take 256 possible values; providing a k-anonymity guarantee on a set of these features, as quasi-identifiers, of any meaningful size, is essentially impossible. Even assuming one could get around this issue, how one can choose a meaningful set of quasi-identifiers within raw image features remains a non-trivial question.

We therefore introduce a novel procedure for computing the k-anonymity of a high-dimensional image dataset, which we describe in Algorithm 1 and illustrate in Figure 2. Through this procedure, one can characterize the ability of a machine learning model to learn meaningful attributes from an image set. The primary goal of any image obfuscation is to reduce the ability to extract privacy-sensitive information from image data, but it can be difficult to define a measure of that ability. Our

Algorithm 1: Computing the Image k-Anonymity of a Dataset

Require: Number of identities p, number of attributes n, number of classes per attribute m_n
Require: Attribute set $A = \{a_1, ..., a_n\}$, quasi-identifier set $QI \subseteq A$
Require: Classifier F1-scores $F = \{f_1, ..., f_n\}$, per attribute
Require: Classifier confidence scores $C = \{c_{1,1,1}, ..., c_{p,n,m_n}\}$, per person, attribute and class
Require: Ground-truths $T = \{t_{1,1}, ..., t_{p,n}\}$, per person and attribute, encoded as class indexes

$EQ \leftarrow$ generateCombinations(QI) ▷ *Generate all possible equivalence classes.*
for $i \leftarrow 1$ to length(EQ) **do**
    $EQC_i \leftarrow 0$ ▷ *Initialize a counter for each equivalence class.*

for $ID \leftarrow 1$ to p **do** ▷ *For every identity, initialize an empty set for all valid attr. assignments for this person.*
    $V \leftarrow \{\}$
    for $ATT \leftarrow 1$ to n **do** ▷ *For every attribute and possible attribute value, prediction confidence is evaluated.*
        for $VAL \leftarrow 1$ to m_{ATT} **do**
            if $c_{ID,ATT,VAL} + \frac{1-f_{ATT}}{m_{ATT}-1} > \frac{1}{m_{ATT}}$ **then** ▷ *If a given classifier guess is **confident enough**, the associated attribute value is added to V.*
                $V \leftarrow V \cup \{(ATT, VAL)\}$
            else if $VAL = t_{ID,ATT}$ **then** ▷ *If the classifier fails to identify the ground-truth, the associated attr. value is also added to V.*
                $V \leftarrow V \cup \{(ATT, VAL)\}$
    for $i \leftarrow 1$ to length(EQ) **do** ▷ *For each equivalence class, if it is represented by the current identity, the associated counter is incremented.*
        if $EQ_i \subseteq V$ **then**
            $EQC_i \leftarrow EQC_i + 1$

$k \leftarrow$ minimumExcludingZero(EQC) ▷ *k is the number of appearances of the least-represented equivalence class, excluding entirely unrepresented equivalence classes.*
return k

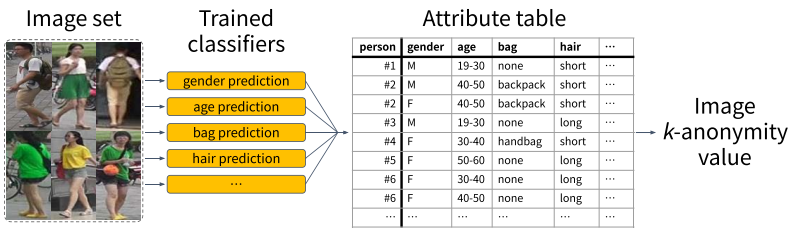


Fig. 2. Visual illustration of the proposed k-anonymity procedure. Notice how persons #2 and #6 appear twice in the table, as the classifiers are unable to decide their gender and age, respectively.

proposed algorithm allows that measure to be k-anonymity, and computing this measure before and after image obfuscation then offers an intuitive view of the ensuing privacy gains.

The proposed evaluation procedure requires a set of images, annotated in terms of an arbitrary number of demographic or visual attributes, and a set of classifiers trained on this image dataset to recognize these attributes. The key idea behind the proposed algorithm is that it simulates how k-anonymity would be computed on a database table that contains attribute values for every person; such a table can be constructed through these trained classifiers. The algorithm examines the performance, confidence, and correctness of the classifiers' predictions, and interprets poor, unconfident, or wrong predictions as instances where the machine learning model is unable to identify a meaningful attribute value. If the model cannot decide on a definite value for an attribute, that means multiple of its values remain admissible; the procedure reflects this by including multiple rows in the database table, for each of the admissible possibilities. For spatial complexity reasons, the

intermediate database table is not constructed; the algorithm instead directly reasons on equivalence classes (i.e., possible combinations for quasi-identifier values), which yields the same results.

Since the procedure from Algorithm 1 relies on the predictions of a trained machine learning model, the threshold defining what constitutes a *confident enough* guess is adjusted relative to the actual prediction quality `[F1-score]` of said model, as well as to chance-level $\frac{1}{m}$:

$$\texttt{[confidence]} + \frac{1 - \texttt{[F1-score]}}{m-1} > \frac{1}{m}. \tag{5}$$

If the classifier's prediction quality is flawless, i.e., `[F1-score]` $= 1$, this simplifies to the following, simply meaning a guess is to be considered confident enough whenever the confidence exceeds chance-level:

$$\texttt{[confidence]} > \frac{1}{m}. \tag{6}$$

On the other hand, for worst-case classifier prediction ability, i.e., `[F1-score]` $= \frac{1}{m}$, this simplifies to the following, meaning every guess is to be considered confident enough, as the classifier is predicting entirely randomly:

$$\texttt{[confidence]} + \frac{1}{m} > \frac{1}{m}. \tag{7}$$

A well-chosen attribute classifier model is expected to fall somewhere in between these two extreme cases. In this manner, the threshold thus gets adjusted relative to the actual prediction ability of trained classifiers, which is then expected to yield reasonable estimates for how well attribute values can be acquired from a set of images.

4 Experimental Method

This section introduces the datasets used in experiments, the explored privacy parameters, the evaluated reID, and attribute prediction models and the baseline obfuscation methods.

4.1 Datasets

Two different reID datasets and one attribute recognition dataset are evaluated in this study. Market1501 [77] is the most widely used public reID dataset, containing 32,668 cropped pedestrian images of 1,501 different individuals, collected across 6 cameras in front of a supermarket in Tsinghua University. Every individual appears on at least 2 cameras, and we use the standard evaluation protocol where 751 identities are used for training (12,936 images), while the remaining 750 are used for testing (13,115 images for the gallery set, when excluding so-called "junk" images, and 3,368 images for the query set). Market1501 is annotated in terms of 27 attributes [41], including 2-class gender (male, female) and 4-class age (young, teenager, adult, old); the remaining 25 binary attributes are related to clothing type and color. All images in Market1501 are 128×64 in size.

The **richly annotated pedestrian (RAP)** dataset [38] is a more recent public attribute recognition and reID dataset, which includes 84,928 cropped pedestrian images, 26,638 of which are identity-annotated for a total of 2,589 different individuals, collected over 25 different cameras inside a shopping mall. While the RAP dataset does include a standard train/test split, the exact way the test set is separated into a gallery and query set is left unspecified. For our evaluation protocol, we thus proceeded as follows. For each identity/camera pair in the test set, an image was randomly sampled. Within this auxiliary subset, for each identity, a number of images were sampled, up to half the amount of cameras the identity appears in, but no more than five images. The resulting subset is our query set; all remaining test images make up our gallery set. Overall, 1,294 identities are used for training (13,148 images), the remaining 1,293 are used for testing (10,667 images for the gallery set and 2,763 images for the query set). RAP is annotated in terms of 111 attributes,

(a) Regular reID model.
The query image is erroneously matched with the nearest image's identity.

(b) Centroid-based reID model.
The query image is correctly matched with the nearest mean centroid's identity.

Fig. 3. Difference between reID models. Red (green) image borders represent wrong (correct) matches.

including 3-class gender (male, female, undetermined) and 5 binary age attributes (16 or less, 17–30, 31–45, 46–60, 61 or more). We discard the latter two age attributes as they contain no samples in the reID subset of the RAP dataset. The remaining 105 attributes binary attributes are related to clothing type and color, body shape, activity, and occupation. Images in RAP have varying sizes, ranging from 114×28 to 634×389; we here resize all images to 128×64, in line with Market1501.

As both Market1501 and RAP suffer from class imbalance for demographic attributes, which are arguably the most important attributes from a privacy perspective, we additionally consider FairFace [32], a face image dataset that provides balanced 2-class gender (male, female), 9-class age (0–2, 3–9, 10–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70, or more), and 7-class race (White, Black, Indian, East Asian, Southeast Asian, Middle Eastern, Latino) annotations for all of its 97,698 images. All images in FairFace are 224×224 in size.

4.2 Person reID Model

We use a simple yet effective reID model in the form of the widely used Bag of Tricks model [44] with a simple ResNet50 [28] CNN backbone. As a means to further enhance the robustness of this model against privacy perturbations, we additionally consider the use of the Centroid Triplet Loss function, as introduced by Wieczorek et al. [62], and their class centroid representations for both the training and testing stages, which achieves state-of-the-art performances on classic reID datasets. Figure 3 illustrates this differences between regular and centroid-based reID models.

This centroid-based model slightly shifts the person reID task from ranking images (i.e., specific identity-instances) to classifying into actual identities, by averaging image embeddings for every person in the dataset. Not only does this arguably make more sense for practical reID applications, but we here hypothesize that these aggregated representations can help the model to magnify the identity-specific latent features that remain underneath the privacy perturbations added to images. As such, we train and test our centroid-based models directly onto noised training, gallery, and query samples, meaning the model never accesses the original, unprotected images. We report reID performances in terms of **mean average precision (mAP)**, which measures the mean of the average precision score for each image within the query set, and Rank-1, defined as the percentage of query set images for which the predicted highest confidence gallery match is a true match. Experiments are repeated three times and we report average metric values.

4.3 Attribute Recognition Model

The attribute predictions reported in this work and used to empirically evaluate k-anonymity as per our proposed method are obtained using simple fully connected layers on top of a pretrained

Table 1. Parameter Settings Discussed in This Study

	b	**c**	Δ**f** (Market1501/RAP)	Δ**f** (FairFace)
(A) High color quantization	0	6	221,184	1,354,752
(B) Mixed pixelization and color quantization	1	5	702,464	4,302,592
(C) High pixelization	2	4	1,728,000	10,584,000
(D) No pixelization nor color quantization	0	0	135,834,624,000	831,987,072,000

ResNet50 [28], a common backbone for image attribute recognition. One layer is trained per attribute, and performances are evaluated in terms of the F1-score as obtained on the test set, which aggregates both the gallery and the query sets for Market1501 and RAP. Experiments are repeated three times and average metric values are reported. For the reID datasets Market1501 and RAP, which contain multiple image representations of the same identity, we report the F1-scores on a per-identity basis, where the predictions obtained on single images are aggregated by identity via a majority vote. This allows for a fairer comparison with the centroid-based reID model, which also benefits from using multiple images per identity.

4.4 Privacy Parameters

To explore the effect of dimensionality reduction parameters b and c of our ε-IDP mechanism on privacy budgets ε and reID performances, we report results for four parameter combinations throughout this study, which are summarized in Table 1. These parameters were chosen empirically, for their performance and illustrative purposes. Privacy budgets ε were made to vary between $\{1, 2.5, 5\} \times 10^x$, where $x \in \{0, ..., 12\}$. Reported sensitivity values Δf are computed both for images with $w = 64$ and $h = 128$, as is the case in Market1501 and RAP, and for images with $w = 224$ and $h = 224$, as in the case for FairFace. It clearly appears the use of dimensionality reduction vastly reduces Δf values, with color quantization parameter c having the largest impact.

4.5 Baseline Methods

This study explores the feasibility of vision tasks on image data protected by differential privacy. Considering the limited studies in this direction, we first compare our proposed method to a variety of traditional image obfuscation methods. Blurring is a simple privacy-preserving method that applies Gaussian kernels to modify each pixel relative to neighboring pixels; we here consider kernel size $k = 25$. Deleting face information by superimposing black boxes on faces can also provide some form of privacy [17]. We here use a CE2P human parsing model [50] to segment images into labeled areas, and consider those labeled as *face*, *hair*, *hat*, and *sunglasses* as sensitive; this area is enlarged into a rectangular shape, and all pixels within are set to zero. Pixelization remains a common way of achieving privacy; we here consider pixelization factor $b = 2$. We also include results with simple color quantization, when reducing color information by a factor $c = 6$.

Besides these traditional image obfuscation methods, we also compare our results to two pixel-level IDP mechanisms, which were reproduced from their respective papers. One of these mechanisms is DP-Pix [23], the mechanism we expanded upon, for which the pixelization factor is set to $b = 2$. Results are also compared to DP-Samp [49], which clusters pixel intensities into k clusters, then samples some of the most common pixel intensities and releases a number of them. The number of clusters is set to $k = 48$. For both of these methods, the m-neighborhood, i.e., the amount of sensitive pixels m, is set to wh (all pixels in the image), in line with our own mechanism, and privacy budgets ε are selected empirically depending on the dataset, for ease of comparison. Note that DP-Pix and DP-Samp work exclusively with grayscale images. An illustration of the selected baselines is provided in Figure 4.

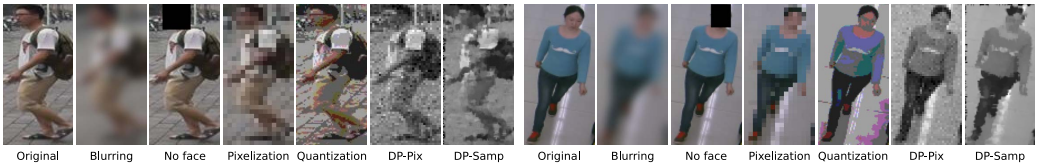


Fig. 4. Baseline methods on example images from Market1501 (left) and RAP (right).

Directly comparing to other IDP methods is difficult. To the best of our knowledge, there exist few other pixel-level IDP mechanisms, and even fewer publicly available implementations of any IDP mechanisms at all. Beyond pixel-level IDP, generative-based IDP methods appear particularly unsuitable for the target task of reID, as they cannot offer consistent obfuscation: Two images of the same person are likely to be noised into two images of completely different people, which completely erases data utility in terms of reID. IDP mechanisms operating under metric differential privacy are also difficult to compare to, as they operate under a different differential privacy definition, which leads to a different privacy granularity in the sense of [21] and results in different scales of privacy budgets ε, rendering comparisons ill-defined.

5 Results

This section details the effect of the proposed ε-IDP and SegCAM-IDP mechanisms on reID performances, on attribute predictions, and on image k-anonymity values, then compares the proposed mechanism to existing baselines. It also includes a subjective analysis of perceived privacy, with example output images and the results of an online user survey.

5.1 Cross-Camera Person reID on Differentially Private Images

The effect of our proposed ε-IDP mechanism on reID performances is reported in Figure 5 for both Market1501 and RAP. Pixelization and color quantization parameters were made to vary by column; both the performance of regular reID models (light lines) and centroid-based reID models (dark lines) are included. It immediately appears from all graphs that centroid-based reID models are much more robust to differentially private noise, with all dark lines outperforming light lines. For both types of models, a decrease in privacy budget ε (which improves privacy) degrades performance metrics, but regular reID models suffer these decreases from higher privacy budgets ε (i.e., at a worse privacy level) than centroid-based reID models. This indicates the latter are indeed a better choice when attempting to do reID on images protected through differential privacy.

Having established this, we now further discuss the observed differences between dimensionality reduction parameter settings A-D. These are juxtaposed in Figure 6 for both datasets. Market1501 performs noticeably better than RAP; this has been noted by the authors of the latter dataset [38], and is likely due to the higher number of camera point-of-views and occlusions in RAP, as well as larger variations in the number of images per pedestrian. Whereas Market1501 performances are hardly impacted by varying dimensionality reduction parameters b and c in scenarios where little noise is added (high privacy budget ε), RAP performances are more largely impacted. Indeed, the right-hand side of the graph for RAP shows the high color quantization setting (blue line, A) scoring noticeably lower than the setting with no pixelization nor quantization (red line, D).

In both these graphs and for all dimensionality reduction parameter settings, we mark with $\star$ the privacy budgets ε around which reID performances go from almost unaffected to nearly chance-level; this behavior is due to color channels having a relatively limited range of values, which means values get forced into their minimum or maximum value once the privacy budget ε drops under a certain threshold. We define these as tradeoff points, at which image data is made

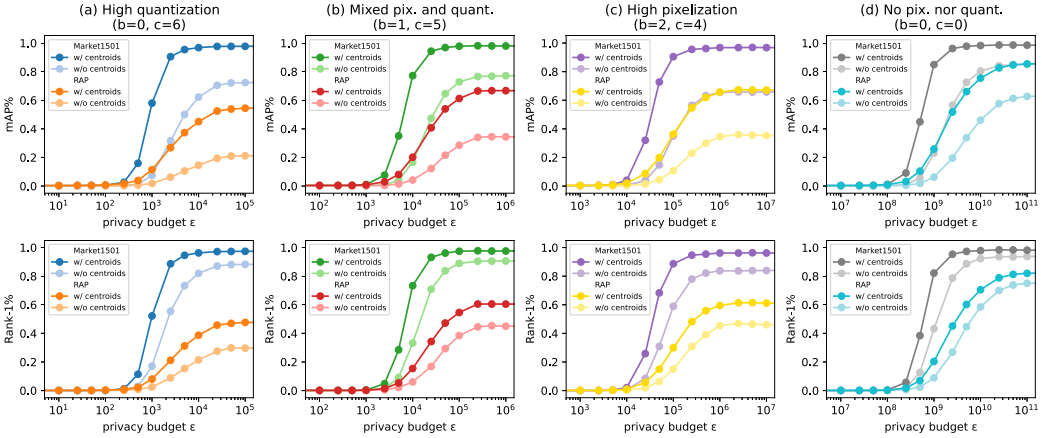


Fig. 5. ReID performances (mAP on top, Rank-1 below) of centroid-based (dark line) and regular (light line) models on both Market1501 and RAP after ε-IDP. Note the different scales on the x-axes.

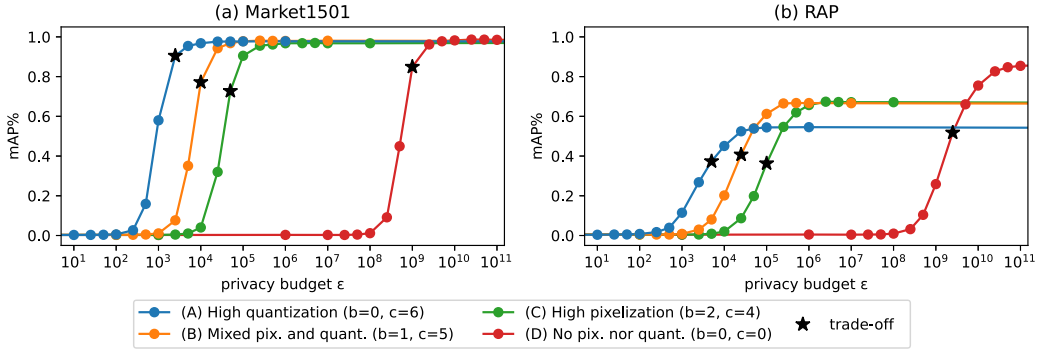


Fig. 6. ReID performances of centroid-based models on Market1501 and RAP after ε-IDP. Only mAP is reported for brevity, as Rank-1 behaves almost the same.

maximally differentially private while remaining reasonably useful (mAP% above half of what it would be with a high privacy budget ε, i.e., under low privacy constraints). Both datasets display the same order in which dimensionality reduction parameter settings encounter these tradeoff points. Without any dimensionality reduction (red line, D), any ε value under 10^9 essentially erases all information within pictures, whereas using either or both pixelization and color quantization can allow for much lower privacy budgets ε where performances remain reasonable. Using high quantization (blue line, A) can achieve a privacy budget as low as $\varepsilon = 2{,}500$ even with a mAP $\geq 90\%$ on Market1501. This shift in ε values for different parameter settings is related to their difference in sensitivity values Δf, which we reported in Table 1. Reducing the amount of information in images, or sensitivity, through dimensionality reduction prior to differentially private obfuscation reduces the redundancy inherent to the way raw images store information, which in turn helps privacy.

5.2 Attribute Recognition on Differentially Private Images

The impact of the proposed ε-IDP mechanism on demographic predictions is included under Figure 7. Both Market1501 and RAP offer very similar graphs for gender predictions, with good

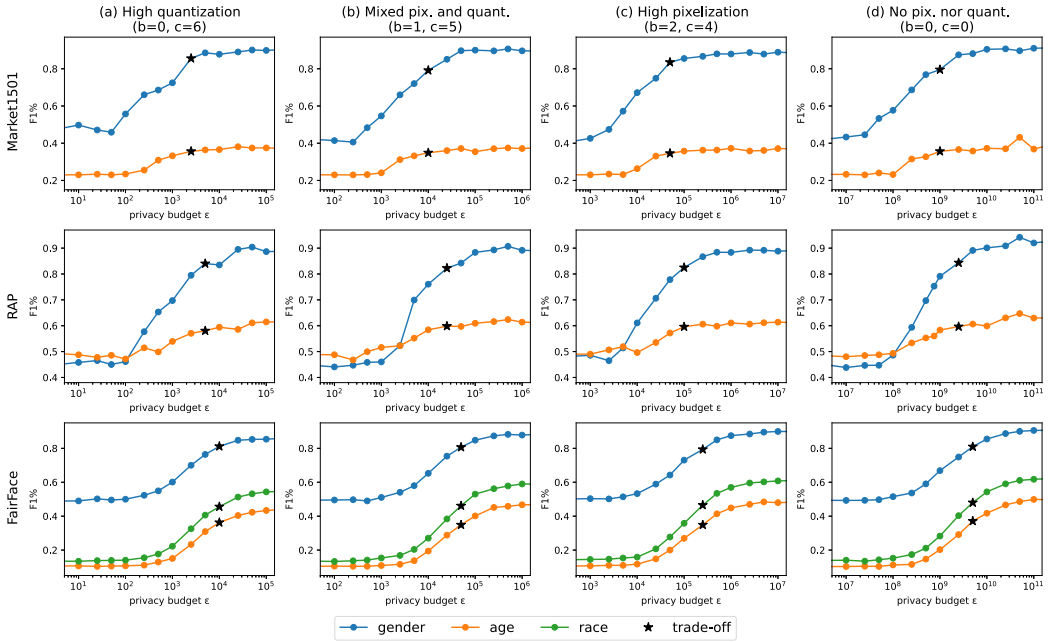


Fig. 7. Demographic prediction ability on Market1501, RAP, and FairFace, after ε-IDP. Note the different scales on the x-axes, per column. For RAP, age F1-scores are averaged across the dataset's binary age attributes.

prediction ability at the ⋆-marked tradeoff points (which are kept consistent with the previous section), but age predictions on these datasets are already relatively poor before even applying any privacy mechanism. As with the reID performances on these datasets, the curves look very similar for all parameter settings (A-D), differing only by the privacy budgets ε on the x-axes. FairFace offers the most complete demographic annotations, and therefore performs noticeably better for all tasks. As before, we identify tradeoff points for this dataset and mark them with ⋆; they are chosen as the lowest privacy budgets ε before prediction abilities start to drop significantly. Compared with the reID task, the slopes around these tradeoff points are gentler for all datasets. We attribute this behavior to demographic predictions being a simpler task than reID.

5.3 Image k-Anonymity Analysis

We here show and discuss the behavior of our proposed image k-anonymity evaluation metric. Using the attribute classifiers trained in the previous section, we apply the procedure we proposed in Algorithm 1. One major hurdle in this process is defining which of the available attributes are to be considered *quasi-identifiers* in the sense of k-anonymity, i.e., which attributes are to be considered most sensitive. One could argue that quasi-identifiers should include the *gender* or *age* attributes, being demographic information, over *sleeve length* or *carrying backpack*; we include such results under Figure 8, which reports how k varies with the privacy budget ε when considering either *gender*, *age*, or *race* quasi-identifiers, subject to availability within datasets.

From Figure 8, it clearly appears that a lower privacy budget ε leads to a higher anonymity metric k for the considered quasi-identifiers, for all three datasets. These graphs also include ⋆-marked tradeoff points consistent with those identified in previous sections, at which minor improvements in k-values can be observed. The choice of quasi-identifier however has a big impact on the precise value for k. Indeed, considering Market1501's binary *gender* attribute a quasi-identifier naturally

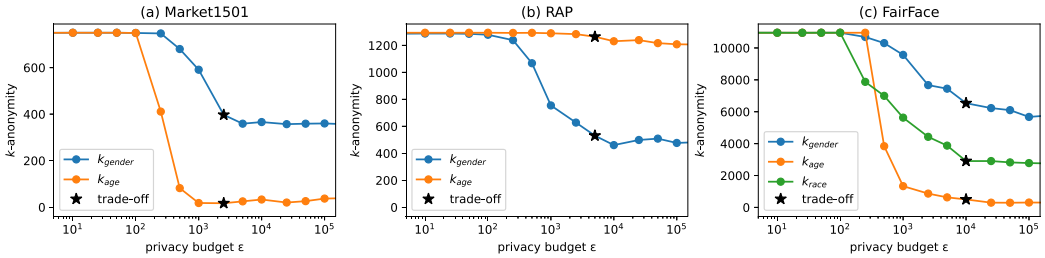


Fig. 8. k-anonymity as computed on Market1501, RAP, and FairFace, after ε-IDP in parameter setting (A), high color quantization. A single quasi-identifier is chosen, and denoted by $k_{\text{quasi-identifier}}$. Note that RAP only includes binary *age* attributes; their values are averaged to avoid cluttering the graph.

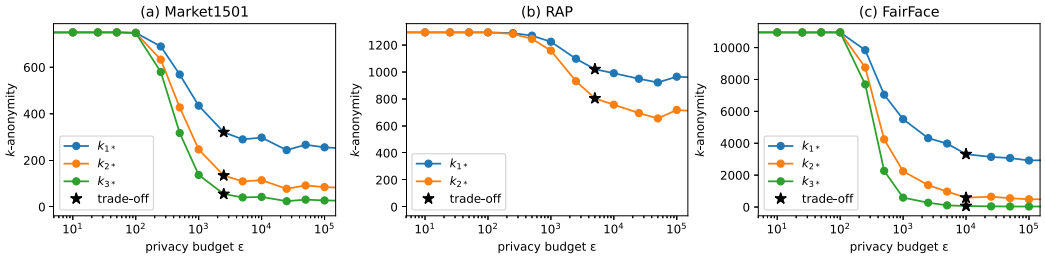


Fig. 9. Mean k-anonymity as computed on Market1501, RAP, and FairFace, after ε-IDP in parameter setting (A), high color quantization. All possible single, pairs, and triplets of quasi-identifiers were considered and averaged to obtain these lines, denoted by k_{1*}, k_{3*}, and k_{3*}, respectively. Note that RAP includes too many attributes to be able to compute a k_{3*} value for quasi-identifier triplets in reasonable time.

leads to a higher k value than when considering its 4-class *age* attribute a quasi-identifier, as there are more individuals in the dataset that share the same *gender* than individuals sharing the same *age*. This observation can also be made for FairFace, where the k value is higher when considering 7-class *race* a quasi-identifier rather than 9-class *age*.

We believe simply this way of choosing quasi-identifiers to be somewhat arbitrary, dependent on some perceived sensitivity ranking of attributes and on available attribute annotations. We therefore consider a more agnostic approach under Figure 9, which reports the mean k-anonymity when every available attribute is considered a quasi-identifier separately; we refer to this metric as k_{1*}. It also includes the mean k-anonymity when every possible pair and every possible triplet of attributes are considered quasi-identifiers; these metrics are referred to as k_{2*} and k_{3*}, respectively. In this manner, we can obtain a dataset and annotation-agnostic view of image k-anonymity, without the need to manually pick quasi-identifiers, under the assumption that every attribute that can be extracted from an image is to some extent privacy-sensitive.

From Figure 9, it appears a low ε, i.e., a high selected privacy level, leads to a high k, i.e., a high observed anonymity level; this is true for any quasi-identifier combination and across all three datasets. The identified ⋆-marked tradeoff points consistently show improved anonymity values k. This indicates that our proposed ε-IDP mechanism is able to provide increased k-anonymity levels without being designed to explicitly optimize for this metric, which suggests it offers broad privacy protection. Another observation to be made from this graph is that larger sets of quasi-identifiers, i.e., larger n in k_{n*}, lead to lower anonymity values. Given that more quasi-identifiers also mean more equivalence classes (i.e., quasi-identifier combinations), this simply reflects how the pool of individuals sharing the same quasi-identifiers shrinks as possible combinations increase.

Table 2. Comparison of Our Proposed Method (under Parameter Settings A to D) on Market1501, RAP, and FairFace with the Original Data (or SOTA When Available) and Existing Obfuscation Baselines

			reID		Gender	Age	Race	Anonymity			
	Method	**Privacy $\varepsilon\downarrow$**	**mAP↑**	**Rank-1↑**	**F1↑**	**F1↑**	**F1↑**	$k_{1*}\uparrow$	$k_{2*}\uparrow$	$k_{3*}\uparrow$	**SSIM↓**
Market1501	Original (*SOTA* [62])	-	*98.3%*	*98.0%*	92.7%	41.9%	-	197	50	13	1.000
	Blurring (k=25)	-	71.5%	87.3%	82.8%	36.4%	-	266	92	31	0.469
	Face blackout	-	82.1%	**92.4%**	85.1%	**37.7%**	-	197	50	13	0.914
	Pixelization (b=2)	-	67.6%	85.2%	83.2%	35.2%	-	266	92	31	0.661
	Quantization (c=6)	-	71.3%	87.4%	84.4%	36.2%	-	264	90	30	0.785
	DP-Pix (b=2) [23]	50,000	41.1%	68.2%	81.7%	34.0%	-	**381**	183	82	0.618
	DP-Samp (k=48) [49]	50,000	35.5%	63.0%	77.4%	34.2%	-	379	185	80	0.707
	Proposed ε-IDP (A)	**2,500**	**90.5%**	88.6%	**85.6%**	35.6%	-	321	135	55	0.220
	Proposed ε-IDP (B)	10,000	77.2%	73.3%	79.2%	34.9%	-	**381**	**190**	**92**	0.232
	Proposed ε-IDP (C)	50,000	72.7%	68.3%	83.5%	34.6%	-	361	170	78	0.330
	Proposed ε-IDP (D)	10^9	84.9%	82.1%	79.5%	35.6%	-	379	188	91	**0.144**
RAP	Original (*SOTA* [38])	-	*47.8%*	*70.7%*	*96.7%*	*74.2%*[a]	-	806	501	-	1.000
	Blurring (k=25)	-	25.9%	35.4%	82.9%	56.6%[a]	-	833	534	-	0.570
	Face blackout	-	34.4%	44.8%	81.7%	59.0%[a]	-	806	501	-	0.921
	Pixelization (b=2)	-	29.6%	39.5%	83.2%	57.6%[a]	-	842	546	-	0.734
	Quantization (c=6)	-	13.1%	19.6%	82.5%	56.5%[a]	-	934	672	-	0.821
	DP-Pix (b=2) [23]	100,000	12.8%	20.9%	53.2%	57.3%[a]	-	1,011	789	-	0.730
	DP-Samp (k=48) [49]	100,000	11.8%	19.1%	51.4%	57.1%[a]	-	1,000	772	-	0.753
	Proposed ε-IDP (A)	**5,000**	37.3%	31.2%	**84.2%**	58.0%[a]	-	**1,021**	**804**	-	0.320
	Proposed ε-IDP (B)	25,000	40.7%	34.3%	81.0%	**59.8%**[a]	-	1,008	785	-	0.395
	Proposed ε-IDP (C)	100,000	36.3%	29.9%	79.2%	59.6%[a]	-	982	745	-	0.450
	Proposed ε-IDP (D)	$2.5 * 10^9$	**51.8%**	**45.1%**	82.3%	59.7%[a]	-	971	728	-	**0.281**
FairFace	Original (*SOTA* [32])	-	-	-	*94.4%*	*60.7%*	*75.4%*[b]	2,746	348	15	1.000
	Blurring (k=25)	-	-	-	90.1%	47.8%	61.8%	2,756	487	34	0.836
	Pixelization (b=2)	-	-	-	**90.8%**	**49.3%**	**61.9%**	2,779	471	18	0.869
	Quantization (c=6)	-	-	-	87.0%	45.0%	55.5%	2,928	442	25	0.742
	DP-Pix (b=2) [23]	250,000	-	-	86.3%	45.0%	55.1%	3,158	521	45	0.367
	DP-Samp (k=48) [49]	250,000	-	-	89.1%	48.0%	58.1%	2,884	382	19	0.880
	Proposed ε-IDP (A)	**10,000**	-	-	81.2%	36.3%	45.6%	3,314	587	64	**0.032**
	Proposed ε-IDP (B)	50,000	-	-	80.6%	34.9%	46.1%	**3,551**	711	**96**	0.041
	Proposed ε-IDP (C)	250,000	-	-	79.4%	34.9%	46.5%	3,507	**796**	94	0.076
	Proposed ε-IDP (D)	$5 * 10^9$	-	-	81.0%	37.1%	48.0%	3,175	499	45	0.038

Best non-SOTA/original values are highlighted in bold, per dataset. Italics denote SOTA values.
[a]Age attributes being binary in RAP, these values correspond to the average F1-scores for all age-related attributes.
[b]The authors [32] simplify the task from 7-class to 4-class to obtain this result.

5.4 Baseline Comparison

We compare the performance of our reID and attribute prediction models with the proposed IDP approach to a number of baseline methods, as introduced in Section 4.5, and to SOTA performances on the evaluated datasets, and report all of these in Table 2. For our proposed method, we include the performance and privacy budget ε at the tradeoff points identified with $\star$ in Figure 6 for parameter settings A-D. Since traditional image obfuscation methods do not rely on a random noising mechanism, no privacy budget ε can be estimated for these. We also include mean k-anonymity values, as reported in the previous section, as well as **structural similarity index measure (SSIM)** [60]. SSIM is a measure for assessing the perceived similarity between images, which we here compute between the original and transformed dataset. The lower this SSIM value, the more different the datasets; a low SSIM metric is expected to correlate with higher privacy.

From Table 2, it appears our proposed method (A-D), just like other obfuscation methods, degrades reID and demographic prediction performances, which is to be expected when trying to increase privacy. Nonetheless, the proposed method generally achieves better reID performances

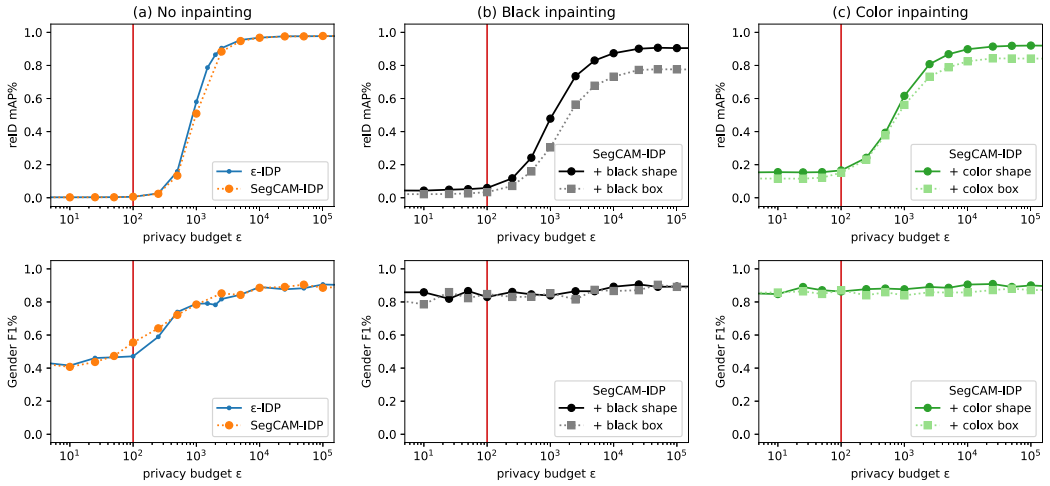


Fig. 10. Comparison of (a) ε-IDP and SegCAM-IDP, (b) black shape-inpainting and box-inpainting for SegCAM-IDP, and (c) color shape-inpainting and box-inpainting for SegCAM-IDP, in terms of reID mAP performances (first row) and gender F1-scores (second row), for Market1501 in parameter setting A.

than traditional obfuscation methods and noticeably better performance than existing pixel-level IDP mechanisms on both Market1501 and RAP, with only face blackout performing similarly well. When it comes to demographic predictions on these two datasets, our proposed method is generally on-par or slightly better than other obfuscation methods. For FairFace, however, we see our proposed method leads to worse demographic predictions than other obfuscation methods, which might be due to the different nature of this dataset (face-only images, whereas the other two datasets contain full-body images) or due its larger images, which might suffer less from the other obfuscation methods at the chosen parameters. When looking at mean k-anonymity, computed from all possible single, pairs, and triplets of quasi-identifiers in the same manner as in the previous section, our proposed method shows a successful increase of k_{1*}, k_{2*}, and k_{3*} over the original data, with significantly better k values than other obfuscation methods. It also clearly appears that our proposed method leads to the lowest SSIM values, implying our method introduces larger visual data distortions while retaining acceptable data usability.

5.5 Effect of SegCAM-IDP

This section summarizes the results of our obfuscation approach combining body part segmentation with CAM. For the sake of brevity, we focus on parameter setting A (high color quantization) for this portion of our analysis, as it allows for the lowest privacy budgets ε, as shown in previous sections; other parameter settings are expected to behave in a similar fashion. Figure 10(a) compares SegCAM-IDP with ε-IDP in terms of reID and gender classification performances after obfuscation. From these graphs, we can observe a minimal decrease in reID performance and a minimal increase in gender classification performance when noise distribution happens non-uniformly, for any given privacy budget ε. Given that the CAM information is obtained via a gender-classifying CNN, this behavior of SegCAM-IDP is in line with our prior expectations, although the difference with regular ε-IDP is not very significant.

This motivates the use of our alternative methods, which use inpainting to further exploit the body-part-segmented CAM information. When inpainting gender-important body parts black or with their average color, as in Figure 10(b) and (c), respectively, it appears reID performances are

Table 3. Comparison of the Proposed ε-IDP and SegCAM-IDP Mechanisms on Market1501 When $\varepsilon = 100$

		reID		Gender	Age	Anonymity			
Method	**Privacy $\varepsilon\downarrow$**	**mAP↑**	**Rank-1↑**	**F1↑**	**F1↑**	$k_{1*}\uparrow$	$k_{2*}\uparrow$	$k_{3*}\uparrow$	**SSIM↓**
Original (*SOTA* [62])	-	*98.3%*	*98.0%*	92.7%	41.9%	197	50	13	1.000
ε-IDP (A)	100	0.6%	0.1%	55.7%	23.4%	748	748	747	0.020
SegCAM-IDP (A)	100	0.6%	0.1%	55.5%	23.3%	745	739	734	0.021
+ inpainting black box	100	3.4%	1.7%	84.8%	27.2%	679	613	552	0.017
+ inpainting black shape	100	5.9%	3.6%	83.0%	32.2%	660	579	506	0.025
+ inpainting color box	100	15.2%	11.1%	87.0%	32.5%	440	253	142	0.102
+ inpainting color shape	100	16.5%	11.9%	86.3%	34.1%	453	269	156	0.103

Italics denote SOTA values.

generally lower for a given privacy budget ε. The same does not apply to gender classification performances, which incur small drops as the privacy budget ε decreases but remain much more stable than when using a regular ε-IDP mechanism, even when the privacy budget ε is set as low as 100, as highlighted in red on the graphs. These results suggest that this inpainting process, which generalizes the visual information in pedestrian images, might be an effective way to reduce the singularity of visual characteristics, making images unsuitable for person reID, while preserving general visual characteristics necessary for attribute recognition tasks.

Our proposed SegCAM-IDP mechanism exhibits its most interesting behavior at low privacy budgets; Table 3 briefly compares this mechanism with ε-IDP when $\varepsilon = 100$. While the original ε-IDP mechanism essentially destroys all information in images at such a low privacy budget, with performances across tasks reaching chance-level, SegCAM-IDP with inpainting manages to retain acceptable gender classification and higher age classification performance. The type of inpainting used has a limited effect on attribute classification performances, but does impact reID performance, which is lowest with black box-inpainting, and then increases when using black shape-inpainting, color box-inpainting, and color shape-inpainting, in that order. Regardless of the type of inpainting, SegCAM-IDP achieves higher k-anonymity levels and lower SSIM values than the original data. These values are in line (or better) with those observed at identified tradeoff points for the ε-IDP mechanism, which we reported in Table 2.

Our results suggest that the proposed SegCAM-IDP mechanism with partial inpainting can provide meaningful obfuscation and is a valid alternative to the regular ε-IDP mechanism for applications where it is desirable to collect demographic information from pedestrians. We here focused specifically on the gender classification task, as it is less computationally expensive to obtain CAM information for the gender classification task than for the reID task. This limitation in our work leaves room for further work that instead explores the effect of using CAM information from the reID task.

5.6 Subjective Evaluation of Images after ε-IDP

We include samples from ε-IDP-protected datasets under different parameter settings and different privacy budgets ε in Figure 11. The previously identified tradeoff points are highlighted with a red border. Both the dimensionality reduction and the actual differentially private mechanism affect the visual aspect of images. Images under higher pixelization degrade faster as the privacy budget ε decreases, whereas higher color quantization streamlines color regions and appears to increase robustness to noise. Tradeoff point images are undoubtedly close to their original, especially when compared side-by-side, but do appear to successfully mask fine-grained information about pedestrians, which are reduced to their general color and shape information. Visually, it appears the

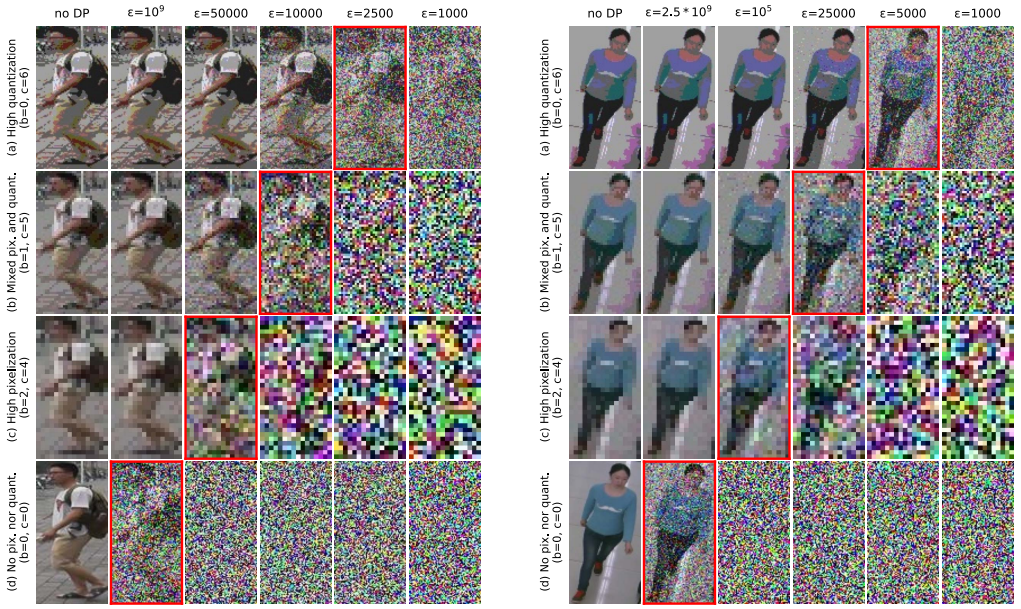


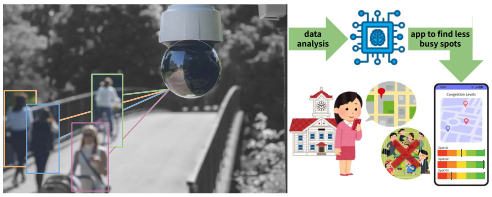
Fig. 11. Visual effect of ε-IDP on example images from Market1501 (left) and RAP (right). The bottom left image is the original. Images at the tradeoff points selected in Table 2 are lined with a red frame.

selected tradeoff points do indeed offer a reasonable tradeoff between performance and leakage of private information. Due to the higher privacy budgets ε, the tradeoff points for RAP (on the right) appear to offer slightly less visual obfuscation than those for Market1501 (on the left). Provided one is willing to further sacrifice data utility, a higher privacy protection level can be achieved with ε-IDP by selecting a lower privacy budget ε. When comparing with traditional obfuscation methods from Figure 4, our proposed method shows a stronger distortion of pedestrians' visual features.

We additionally conducted a user survey through Yahoo! Crowdsourcing[1] to evaluate the perceived acceptability of images protected through our proposed ε-IDP mechanism in two different scenarios. Participants were presented with one of the text and image introductions shown in Figure 12. We considered two different smart city scenarios: (1) the public city administration wishes to collect data for infrastructure improvements, and (2) a private transport company wishes to collect data to improve their services. After confirming they understood the purpose of the survey, participants were asked for their gender (male, female, other) and age group (18 or less, 19–30, 31–45, 46–60, 61, or more). The survey was conducted in Japanese.

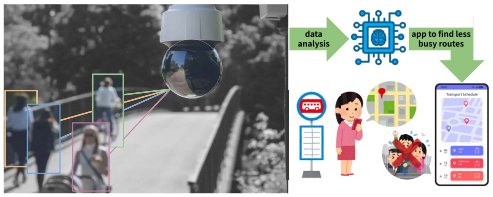
Each participant was asked to judge the acceptability of five ε-IDP-noised images in a binary manner; example images are included under Figure 13. A total of 80 such images were prepared, where 4 original images were noised under 4 different parameter settings and 5 different privacy budgets ε each; these privacy budgets are centered around the previously identified tradeoff points. Each image was shown to 20 different participants, 10 per scenario; this implies each combination of dimensionality reduction parameters and privacy budgets was evaluated by 40 individuals per scenario (10 respondents $\times$ 4 distinct images). A total of 320 people participated in the online survey, 160 per scenario. Participants were compensated 10 JPY for their participation. The results were

[1] https://crowdsourcing.yahoo.co.jp/.



(a) Public city administration

"We are a team of researchers interested in evaluating subjective privacy perceptions. We are developing a system to improve the privacy of data captured by cameras for smart city projects. Assume the images you are about to see are collected by the public city administration to inform and improve urban planning decisions. By sharing your data, you can expect lower congestion throughout the city and better availability of pedestrian crossings and public benches. The city administration also intends to use this data to develop a website where you can check the number of people at specific locations in real-time."



(b) Private transport company

"We are a team of researchers interested in evaluating subjective privacy perceptions. We are developing a system to improve the privacy of data captured by cameras for smart city projects. Assume the images you are about to see are collected by a private transport company to inform and improve its services. By sharing your data, you can expect lower congestion when using busses or trains and better availability of seating options for busy stops. The transport company also intends to use this data to improve its route planning app, so you can use it to find routes that are less busy in real-time, allowing you to travel more comfortably."

Fig. 12. Survey introduction scenarios.

Fig. 13. Example images shown to survey participants. Participants were asked the following question: "*Assume you are the person in the photo. The camera protects the photo on the left and transforms it into the image on the right. Would you feel comfortable if this noised image was collected by a public camera?*".

collected from 226 male respondents and 94 female respondents, with 13 of them in the 19–30 age range, 107 in the 31–45 age range, 159 in the 46–60 age range, and 41 in the 61 or more age range.

Figure 14 shows the number of participants that selected "*Yes, I am comfortable with this level of privacy protection*," per parameter and privacy setting. Results confirm participants are generally more satisfied with higher privacy protection (lower privacy budgets ε), yet the previously identified tradeoff points (the middle bar in each graph) appear generally sufficient to satisfy user's concerns, with over half of the respondents saying they are satisfied with the provided privacy protection. This is on par with the number of positive responses at the lowest surveyed privacy budgets ε, and noticeably better than the number of positive responses at the highest surveyed privacy budgets ε, which suggests our proposed ε-IDP method and the identified tradeoff points are generally able to provide satisfactory *perceived* privacy protection. Responses from people presented with the public city administration and private transport company scenarios appear to vary rather similarly. A two-sample t-test between the two groups, with null hypothesis that they have equal means, yields a p-value of 0.826, which does not suffice to reject the null hypothesis (for any reasonable α). This suggests there might be no statistically significant difference between the way participants view their privacy in the public city administration and private transport company scenarios.

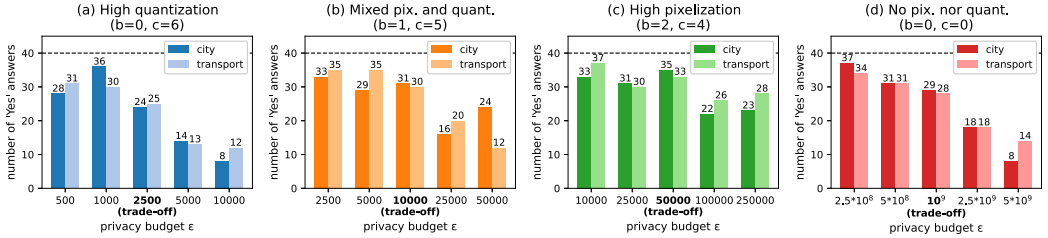


Fig. 14. Number of respondents selecting they were satisfied with the level of privacy protection, per parameter setting and privacy setting, for each of the two respondent groups.

6 Discussion

This section explores the use of empirical data sensitivity, of an alternate noising mechanism, and of an alternate anonymity measure.

6.1 On Sensitivity

For a given data format (i.e., image size and valid color ranges), data sensitivity Δf in the sense of differential privacy is defined as the maximum difference that can be observed between two images. Throughout this study, we have worked under the worst-case assumption regarding these sensitivity values, i.e., we have assumed the maximum difference between two images is the difference between an image that is entirely white (every pixel's RGB values set to 0) and an image that is entirely black (every pixel's RGB values set to 255). One could argue this very conservative assumption is likely to overestimate the actual sensitivity of input images to the privacy mechanism, as a number of valid RGB images are rather unlikely to appear in realistic image data.

To explore this aspect, Table 4 compares the theoretical upper bounds on sensitivity we have used with the empirical sensitivities of the studied datasets. The latter are computed as the maximum pairwise difference between all images in the dataset. From this table, it appears there is a very significant difference between theoretical and empirically computed sensitivities Δf. This distinction is most marked in parameter setting D (no dimensionality reduction) and becomes less intense as color quantization is increased, with the lowest difference between sensitivities under parameter setting A. This suggests the universe of realistic images is much smaller than the universe of valid RGB images, and that reducing the redundancy of the latter (through higher color quantization) may help reduce this gap. It also appears the difference between sensitivities, which is reported in the last column, remains relatively similar across datasets within a given parameter setting, suggesting there may be a dataset-independent way of considering this issue.

As a consequence of the sensitivity differences, we may have used higher noise levels than necessary to provide differential privacy on the data, as the noise addition process described in Equation (2) is directly proportional to data sensitivity and inversely proportional to ε. As such, the ε values throughout this article may actually be overestimated by the factor reported in the last column of Table 4. In these circumstances, one may prefer to work with empirical sensitivity rather than theoretical sensitivity, but this approach has its own drawbacks. First, empirical sensitivity is difficult to compute, as it relies on comparing every pair of images, which is computationally expensive already for the datasets we used. Second, relying on empirical sensitivity rather than theoretical sensitivity implies a privacy risk: If future data does not fit within the data range of past data (on which the empirical sensitivity was computed), the differential privacy mechanism will fail to properly calibrate noise to sensitivity. This weakens the associated privacy guarantee, which may not be accurately reflected anymore by the privacy budget ε. Provided these issues

Table 4. Differences between Theoretical and Empirical Sensitivity, per Parameter Setting and Dataset

Parameter setting	Dataset	Theoretical Δf	Empirical Δf	Difference factor
(A) High color quantization	Market	221,184	57,193	$\times$ 3.9
	RAP	221,184	62,445	$\times$ 3.5
	FairFace	1,354,752	447,858	$\times$ 3.0
(B) Mixed pix. and color quant.	Market	702,464	30,443	$\times$ 23.1
	RAP	702,464	33,544	$\times$ 20.9
	FairFace	4,302,592	251,343	$\times$ 17.1
(C) High pixelization	Market	1,728,000	15,110	$\times$ 114.4
	RAP	1,728,000	16,975	$\times$ 101.8
	FairFace	10,584,000	131,570	$\times$ 80.4
(D) No pix. nor color quant.	Market	135,834,624,000	3,919,222	$\times$ 34, 658.6
	RAP	135,834,624,000	4,396,641	$\times$ 30,895.1
	FairFace	831,987,072,000	34,440,168	$\times$ 24,157.5

can be addressed, by designing a more efficient but sound way to estimate an upper bound on the empirical sensitivity, we believe this opens up a promising direction for future work.

6.2 On Other Noising Mechanisms

Our proposed ε-IDP mechanism relies on Laplacian noising, as detailed in Section 3.1. The most common alternative to the Laplacian mechanism is the Gaussian mechanism [19], which cannot provide pure ε-differential privacy but instead provides a relaxed guarantee, referred to as approximate differential privacy or (ε, δ)-differential privacy. This additional δ parameter represents the failure probability of the mechanism: With probability $1 - \delta$, the mechanism satisfies pure ε-differential privacy; with probability δ, it provides no privacy guarantee at all. The Gaussian mechanism can be extended to the image domain analogously to our previous mechanism.

Definition 6.1 ((ε, δ)-IDP). A randomized mechanism $\mathcal{M}$ gives (ε, δ)-IDP if for any two images i and j of same dimension, and for any possible output $R \subseteq \text{Range}(\mathcal{M})$:

$$\Pr[\mathcal{M}(i) \in R] \leq \exp(\varepsilon) \Pr[\mathcal{M}(j) \in R] + \delta. \tag{8}$$

The Gaussian mechanism can provide such a guarantee [19] provided $\mathcal{M}$ calibrates random noise n to some variance value σ^2, which is related to the ℓ_2-sensitivity $\Delta_2 f$ of function f:

$$\mathcal{M}(x) = f(x) + n, \text{ where } n \sim Gaussian(0, \sigma^2). \tag{9}$$

However, the classical Gaussian mechanism is unable to define this value σ^2 for cases where $\varepsilon > 1$, which are inevitable when working with high-dimensional images. The analytic Gaussian mechanism [7] lifts this restriction by calibrating noise through an optimization problem involving the Gaussian distribution's cumulative distribution function Φ. Analogously to our ℓ_1-sensitivity definition, we define the ℓ_2-sensitivity of RGB images of width w and height h, with optional pixelization determined by b and optional color quantization determined by c, as:

$$\Delta_2 f = \sqrt{\frac{wh}{4^b}\left(\left(\frac{2^8}{2^c} - 1\right)^3\right)^2}. \tag{10}$$

With this relaxed, approximate IDP definition, we can compare the effect of the Laplacian and the Gaussian mechanism on our target tasks; for the sake of brevity, we only include results for

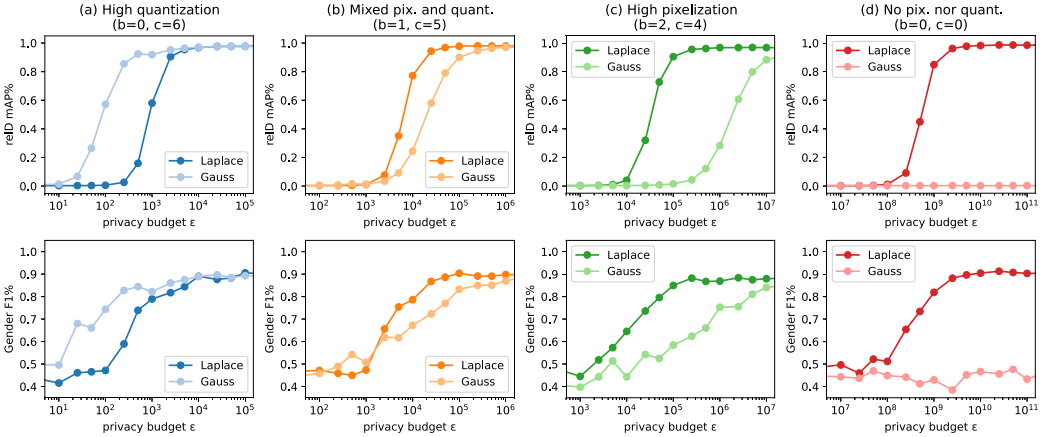
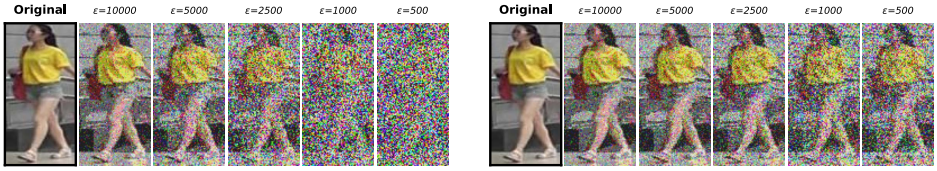


Fig. 15. Comparison of Laplacian ε-IDP (dark line) and Gaussian (ε, δ)-IDP (light line) on Market1501, for the reID (first row) and gender prediction tasks (second row). Note the different scales on the x-axes.



(a) Laplacian ε-IDP. (b) Gaussian (ε, δ)-IDP.

Fig. 16. Comparison of noising mechanisms on an example image from Market1501 in parameter setting A.

Market1501 under Figure 15. Privacy budgets ε are in line with those evaluated previously, and the Gaussian mechanism's failure probability δ is set to 0.001.

From these graphs, it appears using Gaussian noise is generally less desirable for performances, as settings B through D result in lower reID mAP and gender prediction F1-scores. The Gaussian mechanism outperforms the Laplacian mechanism in setting A, where it is able to offer a significant ε decrease for the same task performance. However, the Gaussian mechanism only guarantees approximate differential privacy, i.e., does not provide any privacy guarantee with a 0.001 probability, whereas the Laplacian mechanism cannot fail to provide a privacy guarantee; this makes it difficult to directly compare ε values. Visualization of the noised images under both mechanisms in setting A, in Figure 16, confirms this intuition, as images noised under the Gaussian mechanism appear less protected than images noised under the Laplacian mechanism, for the same ε.

In this study, we chose to work with the Laplacian mechanism for its stricter privacy definition; we acknowledge the use of other noising mechanisms may be beneficial for performances on IDP-protected data. Given the limited work on IDP mechanisms, and the discrepancies in how differential privacy mechanisms define their privacy guarantees, we believe more in-depth mechanism comparisons to be a good direction for future work in this domain.

6.3 On Other Anonymity Measures

Being one of the most prominently studied privacy metrics, we here used k-anonymity to characterize the anonymity of IDP-protected images. Several metrics have aimed to extend k-anonymity to better characterize anonymity, one of them being l-diversity [45]. It measures the variability of a

Algorithm 2: Computing the Image l-Diversity of a Dataset (Modified from Algorithm 1)

Require: (...)
Require: Sensitive attribute $SA \subseteq A$
(...)
for $i \leftarrow 1$ to length(EQ) **do**
  $EQS_i \leftarrow \{\}$ ▷ *Initialize an empty set for each equivalence class.*
(...)
  for $i \leftarrow 1$ to length(EQ) **do** ▷ *For each equivalence class,*
    if $EQ_i \subseteq V$ **then** *if it is represented by the current identity,*
      $EQS_i \leftarrow EQS_i \cup V[SA, *]$ *their sensitive attribute values are appended.*
▷ *l is the size of the smallest sensitive attribute*
$l \leftarrow$ minimumExcludingZero(size(EQS)) *value set across equivalence classes, excluding*
return l *entirely unrepresented equivalence classes.*

given sensitive attribute SA: The higher this variance l across equivalence classes, the lower the vulnerability to background knowledge attacks for that sensitive attribute. We can derive a (single sensitive attribute) image l-diversity algorithm from our k-anonymity algorithm; we include the necessary modifications to our previous algorithm under Algorithm 2. It can be extended to multi sensitive attribute settings; however, as the number of equivalence classes grows exponentially with the number of quasi-identifiers, which grow with the number of sensitive attributes in multi-sensitive attribute settings [45], the number of equivalence classes quickly becomes intractable and unpractical to compute. As such, we limit our analysis to single sensitive attribute l-diversity. Analogously to the quasi-identifier agnostic approach for k-anonymity, we propose a sensitive attribute agnostic approach for l-diversity, and report average l-diversity for any sensitive attribute and any single, pair, triplet of quasi-identifiers; we refer to these metrics as l_1*, l_2*, and l_3*, respectively.

The l-diversity values obtained on Market1501, as reported in Figure 17(a), follow a similar trend as k-anonymity values. Lower privacy budgets ε increase the l-diversity values, suggesting the ε-IDP mechanism successfully increases the diversity of observable sensitive attribute values. Considering a larger set of quasi-identifiers, i.e., computing l_2* or l_3* over l_1*, leads to lower diversity values l. This is a consequence of the increase in equivalence classes, all of which need to feature diverse sensitive attribute values for l to remain high, which gets increasingly difficult with more equivalence classes. The l-diversity values obtained on RAP and FairFace, observable in Figure 17(b) and (c), respectively, are less informative. RAP shows little variation in l-diversity values, regardless of the privacy budget ε. We attribute this behavior to the fact RAP has over 100 attributes, almost all of which binary; by averaging l-diversity values across so many sensitive attributes, we may fail to capture their finer variational trends. FairFace on the other hand shows maximal l-diversity for all privacy budgets ε. This is due to the dataset being balanced by design, i.e., already offering maximal diversity in all its attributes, regardless of privacy budgets ε.

Further extensions to k-anonymity and l-diversity have been proposed, e.g., t-closeness [39], which measures the distribution of sensitive attributes across equivalence classes. Given the low relevance of l-diversity on image data with binary attribute labels observed from our results, the direct applicability of t-closeness or further extensions on such image data remains uncertain, and may warrant further research on the relationship between these metrics and differential privacy's ε.

7 Conclusion

Visual data is ubiquitous in the current data landscape and often used in smart city research and applications, due to the common availability and high potential of video cameras. Despite this,

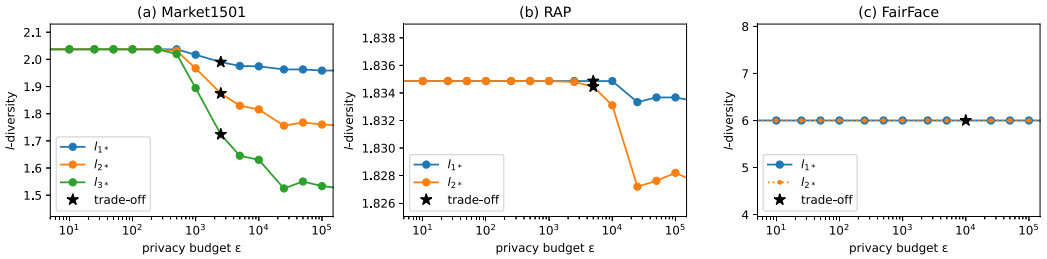


Fig. 17. Mean l-diversity as computed on Market1501, RAP, and FairFace, after ε-IDP in parameter setting (A), high color quantization. All possible single, pairs, and triplets of quasi-identifiers and all possibles single sensitive attributes were considered and averaged to obtain these lines, denoted by l_{1*}, l_{3*}, and l_{3*}, respectively. Note that RAP includes too many attributes to be able to compute a l_{3*} value in reasonable time, and that FairFace contains too few attributes to compute anything higher than a l_{2*} value.

methods that can provide privacy guarantees to affected citizens while retaining practical data utility remain limited, which hurts the social acceptance of IoT-based smart city systems. We previously introduced a strict IDP mechanism that allows system operators to quantify and control the privacy leakage of such systems, and we here extend its analysis to two additional datasets for both cross-camera person reID and attribute prediction tasks. Through extensive experiments, we confirm our proposed mechanism achieves better obfuscation than existing methods while retaining acceptable utility for the target tasks, at practical tradeoff points we have identified for the Market1501, RAP, and FairFace datasets.

By design, increasing privacy, or limiting information leakage, comes at the expense of utility, or information itself. While our proposed methods can outperform traditional obfuscation baselines at certain tradeoff points, privacy budgets ε and thus private information leakage remain non-negligible. Nevertheless, provided smart city system operators can identify the privacy-utility tradeoff best suited to their data and goals, we believe our IDP mechanisms valuable for storing visual data with a reduced privacy footprint. Our findings confirm that reID images protected through differential privacy can remain valuable as a whole but are of limited use in fragments, a key property which we expect especially useful in distributed IoT settings, where data leaks can be expected to concern subsets of data.

This article additionally introduced a novel IDP mechanism, which utilizes body-part information together with CAM models to locate and rank areas within images in terms of their importance to gender identification, and applies noise and/or inpainting accordingly. Our experiments show that this process is able to preserve demographic information while stripping images of a majority of their person reID potential. We have also introduced a new way to empirically evaluate the k-anonymity of a set of images, and report its behavior on the considered datasets. Our experiments confirm that data protected by our differential privacy mechanisms results in higher anonymity levels k than unprotected or traditionally obfuscated data. We expect the introduced IDP and image k-anonymity methods to be useful for the public release of image datasets and the use of video cameras in smart city systems, assisting data owners and system operators in choosing privacy budgets fit to their privacy ambitions and desired applications.

References

[1] Martin Abadi, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. 2016. Deep learning with differential privacy. In *ACM SIGSAC Conference on Computer and Communications Security (CCS '16)*. ACM, New York, NY, 308–318. DOI: https://doi.org/10.1145/2976749.2978318

[2] John M. Abowd. 2018. The U.S. Census Bureau adopts differential privacy. In *24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18)*. ACM, New York, NY, 2867. DOI: https://doi.org/10.1145/3219819.3226070

[3] Charu C. Aggarwal. 2005. On K-anonymity and the curse of dimensionality. In *31st International Conference on Very Large Data Bases (VLDB '05). VLDB Endowment*, 901–909. Retrieved from https://dl.acm.org/doi/10.5555/1083592.1083696

[4] Shafiq Ahmad, Gianluca Scarpellini, Pietro Morerio, and Alessio Del Bue. 2022. Event-driven Re-Id: A new benchmark and method towards privacy-preserving person re-identification. In *IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW '22)*, 459–468. DOI: https://doi.org/10.1109/WACVW54805.2022.00052

[5] Miguel E. Andrés, Nicolás E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2013. Geo-indistinguishability: Differential privacy for location-based systems. In *ACM SIGSAC Conference on Computer and Communications Security (CCS '13)*. ACM, New York, NY, 901–914. DOI: https://doi.org/10.1145/2508859.2516735

[6] Apple. 2017. Differential Privacy Team. Learning with Privacy at Scale. Retrieved from https://machinelearning.apple.com/research/learning-with-privacy-at-scale

[7] Borja Balle and Yu-Xiang Wang. 2018. Improving the Gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. In *35th International Conference on Machine Learning (ICML '18), Vol. 80*. PMLR, 394–403. Retrieved from http://proceedings.mlr.press/v80/balle18a.html

[8] Ero Balsa, Helen Nissenbaum, and Sunoo Park. 2022. Cryptography, trust and privacy: It's complicated. In *ACM Symposium on Computer Science and Law (CSLAW '22)*. ACM, New York, NY, 167–179. DOI: https://doi.org/10.1145/3511265.3550443

[9] Loris Bazzani, Marco Cristani, Alessandro Perina, Michela Farenzena, and Vittorio Murino. 2010. Multiple-shot person re-identification by HPE signature. In *20th International Conference on Pattern Recognition (ICPR '10)*, 1413–1416. DOI: https://doi.org/10.1109/ICPR.2010.349

[10] Jingyi Cao, Bo Liu, Yunqian Wen, Yunhui Zhu, Rong Xie, Li Song, Lin Li, and Yaoyao Yin. 2022. Hiding among your neighbors: Face image privacy protection with differential private k-anonymity. In *2022 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '22)*, 1–6. DOI: https://doi.org/10.1109/BMSB55706.2022.9828699

[11] Konstantinos Chatzikokolakis, Miguel E. Andrés, Nicolás Emilio Bordenabe, and Catuscia Palamidessi. 2013. Broadening the scope of differential privacy using metrics. In *Privacy Enhancing Technologies*. Emiliano De Cristofaro and Matthew Wright (Eds.), Springer, Berlin, 82–102. DOI: https://doi.org/10.1007/978-3-642-39077-7_5

[12] Jia-Wei Chen, Li-Ju Chen, Chia-Mu Yu, and Chun-Shien Lu. 2021. Perceptual indistinguishability-net (PI-Net): Facial image obfuscation with manipulable semantics. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR '21)*. IEEE Computer Society, Los Alamitos, CA, 6474–6483. DOI: https://doi.org/10.1109/CVPR46437.2021.00641

[13] Xiaodong Chen, Xinchen Liu, Wu Liu, Xiao-Ping Zhang, Yongdong Zhang, and Tao Mei. 2021. Explainable person re-identification with attribute-guided metric distillation. In *IEEE/CVF International Conference on Computer Vision (ICCV '21)*, 11813–11822. DOI: https://doi.org/10.1109/ICCV48922.2021.01160

[14] Yifan Chen, Han Wang, Xiaolu Sun, Bin Fan, Chu Tang, and Hui Zeng. 2022. Deep attention aware feature learning for person re-identification. *Pattern Recognition* 126 (2022), 108567. DOI: https://doi.org/10.1016/j.patcog.2022.108567

[15] Hang Cheng, Huaxiong Wang, Ximeng Liu, Yan Fang, Meiqing Wang, and Xiaojun Zhang. 2021. Person re-identification over encrypted outsourced surveillance videos. *IEEE Transactions on Dependable and Secure Computing* 18, 3 (2021), 1456–1473. DOI: https://doi.org/10.1109/TDSC.2019.2923653

[16] William Croft, Jörg-Rüdiger Sack, and Wei Shi. 2021. Obfuscation of images via differential privacy: From facial images to general images. *Peer-to-Peer Networking and Applications* 14, 3 (2021), 1705–1733. DOI: https://doi.org/10.1007/s12083-021-01091-9

[17] Julia Dietlmeier, Joseph Antony, Kevin McGuinness, and Noel E. O'Connor. 2021. How important are faces for person re-identification? In *25th International Conference on Pattern Recognition (ICPR '20)*, 6912–6919. DOI: https://doi.org/10.1109/ICPR48806.2021.9412340

[18] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. 2017. Collecting telemetry data privately. In *Advances in Neural Information Processing Systems (NIPS '17)*, Vol. 30. Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2017/file/253614bbac999b38b5b60cae531c4969-Paper.pdf

[19] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. 2006. Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology (EUROCRYPT '06)*. Springer, Berlin, 486–503. DOI: https://doi.org/10.1007/11761679_29

[20] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*. Shai Halevi and Tal Rabin (Eds.), Springer, Berlin, 265–284. DOI: https://doi.org/10.1007/11681878_14

[21] Cynthia Dwork and Aaron Roth. 2014. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science* 9, 3–4 (2014), 211–407. DOI: https://doi.org/10.1561/0400000042

[22] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. 2014. RAPPOR: Randomized aggregatable privacy-preserving ordinal response. In *ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*. ACM, New York, NY, 1054–1067. DOI: https://doi.org/10.1145/2660267.2660348

[23] Liyue Fan. 2018. Image pixelization with differential privacy. In *Data and Applications Security and Privacy XXXII*, Vol. 10980. Springer International Publishing, Cham, 148–162. DOI: https://doi.org/10.1007/978-3-319-95729-6_10

[24] Liyue Fan. 2019. Practical image obfuscation with provable privacy. In *2019 IEEE International Conference on Multimedia and Expo (ICME '19)*, 784–789. DOI: https://doi.org/10.1109/ICME.2019.00140

[25] Lijie Fan, Tianhong Li, Rongyao Fang, Rumen Hristov, Yuan Yuan, and Dina Katabi. 2020. Learning longterm representations for person re-identification using radio signals. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR '20)*, 10696–10706. DOI: https://doi.org/10.1109/CVPR42600.2020.01071

[26] Michela Farenzena, Loris Bazzani, Alessandro Perina, Vittorio Murino, and Marco Cristani. 2010. Person re-identification by symmetry-driven accumulation of local features. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, 2360–2367. DOI: https://doi.org/10.1109/CVPR.2010.5539926

[27] Dengpan Fu, Dongdong Chen, Jianmin Bao, Hao Yang, Lu Yuan, Lei Zhang, Houqiang Li, and Dong Chen. 2021. Unsupervised pre-training for person re-identification. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR '21)*, 14745–14754. DOI: https://doi.org/10.1109/CVPR46437.2021.01451

[28] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '16)*, 770–778. DOI: https://doi.org/10.1109/CVPR.2016.90

[29] Alexander Hermans, Lucas Beyer, and Bastian Leibe. 2017. In defense of the triplet loss for person re-identification. arXiv:1703.07737. DOI: https://doi.org/10.48550/arXiv.1703.07737

[30] Steven Hill, Zhimin Zhou, Lawrence Saul, and Hovav Shacham. 2016. On the (in)effectiveness of mosaicing and blurring as tools for document redaction. *Proceedings on Privacy Enhancing Technologies* 2016, 4 (2016), 403–417. DOI: https://doi.org/10.1515/popets-2016-0047

[31] Shogo Isoda, Masato Hidaka, Yuki Matsuda, Hirohiko Suwa, and Keiichi Yasumoto. 2020. Timeliness-aware on-site planning method for tour navigation. *Smart Cities* 3, 4 (2020), 1383–1404. DOI: https://doi.org/10.3390/smartcities3040066

[32] Kimmo Karkkainen and Jungseock Joo. 2021. FairFace: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation. In *IEEE Winter Conference on Applications of Computer Vision (WACV '21)*, 1547–1557. DOI: https://doi.org/10.1109/WACV48630.2021.00159

[33] Stephan Kessler, Jens Hoff, and Johann-Christoph Freytag. 2019. SAP HANA goes private: From privacy research to privacy aware enterprise analytics. *Proceedings of the VLDB Endowment* 12, 12 (2019), 1998–2009. DOI: https://doi.org/10.14778/3352063.3352119

[34] Jamil Khan, Robert Hrelja, and Fredrik Pettersson-Löfstedt. 2021. Increasing public transport patronage—An analysis of planning principles and public transport governance in Swedish regions with the highest growth in ridership. *Case Studies on Transport Policy* 9, 1 (2021), 260–270. DOI: https://doi.org/10.1016/j.cstp.2020.12.008

[35] JungHoon Kim. 2022. Smart city trends: A focus on 5 countries and 15 companies. *Cities* 123 (2022), 103551. DOI: https://doi.org/10.1016/j.cities.2021.103551

[36] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems (NIPS '12)*. Curran Associates, Inc., 1097–1105. Retrieved from https://dl.acm.org/doi/10.5555/2999134.2999257

[37] Ming Lang, Wang Yongli, Huang Yuanyuan, and Chen Junyu. 2024. Differential privacy face image publishing method based on neighboring pixels merging. In *International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 1–4. DOI: https://doi.org/10.1109/ICCWAMTIP64812.2024.10873720

[38] Dangwei Li, Zhang Zhang, Xiaotang Chen, and Kaiqi Huang. 2019. A richly annotated pedestrian dataset for person retrieval in real surveillance scenarios. *IEEE Transactions on Image Processing* 28, 4 (2019), 1575–1590. DOI: https://doi.org/10.1109/TIP.2018.2878349

[39] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. 2007. t-closeness: Privacy beyond k-anonymity and l-diversity. In *IEEE International Conference on Data Engineering (ICDE '07)*, 106–115. DOI: https://doi.org/10.1109/ICDE.2007.367856

[40] Tao Li and Chris Clifton. 2021. Differentially private imaging via latent space manipulation. In *IEEE Symposium on Security and Privacy 2021 (IEEE S &P '21)*. DOI: https://doi.org/10.48550/arXiv.2103.05472

[41] Yutian Lin, Liang Zheng, Zhedong Zheng, Yu Wu, Zhilan Hu, Chenggang Yan, and Yi Yang. 2019. Improving person re-identification by attribute and identity learning. *Pattern Recognition* 95 (2019), 151–161. DOI: https://doi.org/10.1016/j.patcog.2019.06.006

[42] Bo Liu, Ming Ding, Hanyu Xue, Tianqing Zhu, Dayong Ye, Li Song, and Wanlei Zhou. 2021. DP-Image: Differential privacy for image data in feature space. arXiv:2103.07073. DOI: https://doi.org/10.48550/arXiv.2103.07073

[43] Chao Liu, Jing Yang, Weinan Zhao, Yining Zhang, Jingyou Li, and Chunmiao Mu. 2021. Face image publication based on differential privacy. *Wireless Communications and Mobile Computing* 2021, 1 (2021), 1–20. DOI: https://doi.org/10.1155/2021/6680701

[44] Hao Luo, Youzhi Gu, Xingyu Liao, Shenqi Lai, and Wei Jiang. 2019. Bag of tricks and a strong baseline for deep person re-identification. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW '19)*, 1487–1495. DOI: https://doi.org/10.1109/CVPRW.2019.00190

[45] Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. 2007. L-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data* 1, 1 (March 2007), 3–es. DOI: https://doi.org/10.1145/1217299.1217302

[46] Lucas Maris, Yuki Matsuda, and Keiichi Yasumoto. 2024. Protecting cross-camera person re-identification data with image differential privacy. In *IEEE International Conference on Smart Computing (SMARTCOMP '24)*. IEEE Computer Society, Los Alamitos, CA, 386–391. DOI: https://doi.org/10.1109/SMARTCOMP61445.2024.00086

[47] Richard McPherson, Reza Shokri, and Vitaly Shmatikov. 2016. Defeating image obfuscation with deep learning. arXiv:1609.00408. DOI: https://doi.org/10.48550/arXiv.1609.00408

[48] Masakazu Ohno, Riki Ukyo, Tatsuya Amano, Hamada Rizk, and Hirozumi Yamaguchi. 2023. Privacy-preserving pedestrian tracking using distributed 3D LiDARs. In *IEEE International Conference on Pervasive Computing and Communications (PerCom '23)*, 43–52. DOI: https://doi.org/10.1109/PERCOM56429.2023.10099061

[49] Dominick Reilly and Liyue Fan. 2021. A comparative evaluation of differentially private image obfuscation. In *3rd IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA '21)*, 80–89. DOI: https://doi.org/10.1109/TPSISA52974.2021.00009

[50] Tao Ruan, Ting Liu, Zilong Huang, Yunchao Wei, Shikui Wei, and Yao Zhao. 2019. Devil in the details: Towards accurate single and multiple human parsing. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 1 (2019), 4814–4821. DOI: https://doi.org/10.1609/aaai.v33i01.33014814

[51] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. 2015. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision* 115, 3 (2015), 211–252. DOI: https://doi.org/10.1007/s11263-015-0816-y

[52] Gerhard Schrotter and Christian Hürzeler. 2020. The digital twin of the city of Zurich for urban planning. *PFG—Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 88, 1 (2020), 99–112. DOI: https://doi.org/10.1007/s41064-020-00092-2

[53] Hisaichi Shibata, Shouhei Hanaoka, Yang Cao, Masatoshi Yoshikawa, Tomomi Takenaga, Yukihiro Nomura, Naoto Hayashi, and Osamu Abe. 2023. Local differential privacy image generation using flow-based deep generative models. *Applied Sciences* 13, 18 (2023), 10132. DOI: https://doi.org/10.3390/app131810132

[54] Vladimir Somers, Christophe De Vleeschouwer, and Alexandre Alahi. 2023. Body part-based representation learning for occluded person re-identification. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV '23)*, 1613–1623.

[55] Zhigang Su, Dawei Zhou, Nannan Wang, Decheng Liu, Zhen Wang, and Xinbo Gao. 2023. Hiding visual information via obfuscating adversarial perturbations. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 4333–4343. DOI: https://doi.org/10.1109/ICCV51070.2023.00402

[56] Latanya Sweeney. 2002. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 5 (2002), 557–570. DOI: https://doi.org/10.1142/S0218488502001648

[57] Abbas Shah Syed, Daniel Sierra-Sosa, Anup Kumar, and Adel Elmaghraby. 2021. IoT in smart cities: A survey of technologies, practices and challenges. *Smart Cities* 4, 2 (2021), 429–475. DOI: https://doi.org/10.3390/smartcities4020024

[58] Guangcong Wang, Jianhuang Lai, Peigen Huang, and Xiaohua Xie. 2019. Spatial-temporal person re-identification. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 1 (2019), 8933–8940. DOI: https://doi.org/10.1609/aaai.v33i01.33018933

[59] Han Wang, Shangyu Xie, and Yuan Hong. 2020. VideoDP: A flexible platform for video analytics with differential privacy. *Proceedings on Privacy Enhancing Technologies* 2020 (October 2020), 277–296. DOI: https://doi.org/10.2478/popets-2020-0073

[60] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. 2004. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4 (2004), 600–612. DOI: https://doi.org/10.1109/TIP.2003.819861

[61] Yunqian Wen, Bo Liu, Ming Ding, Rong Xie, and Li Song. 2022. IdentityDP: Differential private identification protection for face images. *Neurocomputing* 501, C (Aug. 2022), 197–211. DOI: https://doi.org/10.1016/j.neucom.2022.06.039

[62] Mikołaj Wieczorek, Barbara Rychalska, and Jacek Dąbrowski. 2021. On the unreasonable effectiveness of centroids in image retrieval. In *Neural Information Processing (ICONIP '21)*, 212–223. DOI: https://doi.org/10.1007/978-3-030-92273-3_18

[63] Guile Wu and Shaogang Gong. 2021. Decentralised learning from independent multi-domain labels for person re-identification. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 4 (2021), 2898–2906. DOI: https://doi.org/10.1609/aaai.v35i4.16396

[64] Xiaokui Xiao and Yufei Tao. 2007. M-invariance: Towards privacy preserving re-publication of dynamic datasets. In *ACM SIGMOD International Conference on Management of Data (SIGMOD '07)*. ACM, New York, NY, 689–700. DOI: https://doi.org/10.1145/1247480.1247556

[65] Valikhujaev Yakhyokhuja. 2024. Face-Parsing. Retrieved from https://github.com/yakhyo/face-parsing

[66] Haonan Yan, Xiaoguang Li, Wenjing Zhang, Qian Chen, Bin Wang, Hui Li, and Xiaodong Lin. 2024. CODER: Protecting privacy in image retrieval with differential privacy. *IEEE Transactions on Dependable and Secure Computing* 21, 6 (2024), 5420–5430. DOI: https://doi.org/10.1109/TDSC.2024.3376532

[67] Zhengwei Yang, Meng Lin, Xian Zhong, Yu Wu, and Zheng Wang. 2023. Good is bad: Causality inspired cloth-debiasing for cloth-changing person re-identification. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR '23)*, 1472–1481. DOI: https://doi.org/10.1109/CVPR52729.2023.00148

[68] Mang Ye, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao, and Steven C. H. Hoi. 2022. Deep learning for person re-identification: A survey and outlook. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 6 (2022), 2872–2893. DOI: https://doi.org/10.1109/TPAMI.2021.3054775

[69] Mang Ye, Wei Shen, Junwu Zhang, Yao Yang, and Bo Du. 2024. SecureReID: Privacy-preserving anonymization for person re-identification. *IEEE Transactions on Information Forensics and Security* 19 (2024), 2840–2853. DOI: https://doi.org/10.1109/TIFS.2024.3356233

[70] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z. Li. 2014. Deep metric learning for person re-identification. In *22nd International Conference on Pattern Recognition (ICPR '14)*, 34–39. DOI: https://doi.org/10.1109/ICPR.2014.16

[71] Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang. 2018. BiSeNet: Bilateral segmentation network for real-time semantic segmentation. In *Computer Vision (ECCV '18)*. Springer-Verlag, Berlin, 334–349. DOI: https://doi.org/10.1007/978-3-030-01261-8_20

[72] Jinao Yu, Hanyu Xue, Bo Liu, Yu Wang, Shibing Zhu, and Ming Ding. 2020. GAN-based differential private image privacy protection framework for the internet of multimedia things. *Sensors* 21, 1, Article 58 (2020). DOI: https://doi.org/10.3390/s21010058

[73] Junxue Zhang, Xiaodian Cheng, Liu Yang, Jinbin Hu, Ximeng Liu, and Kai Chen. 2024. SoK: Fully homomorphic encryption accelerators. *ACM Computing Surveys* 56, 12, Article 316 (Oct. 2024), 32 pages. DOI: https://doi.org/10.1145/3676955

[74] Junwu Zhang, Mang Ye, and Yao Yang. 2022. Learnable privacy-preserving anonymization for pedestrian images. In *30th ACM International Conference on Multimedia (MM '22)*. ACM, New York, NY, 7300–7308. DOI: https://doi.org/10.1145/3503161.3548766

[75] Xiaoting Zhang, Tao Wang, and Junhao Ji. 2024. SemDP: Semantic-level differential privacy protection for face datasets. arXiv:2412.15590. DOI: https://doi.org/10.48550/arXiv.2412.15590

[76] Ying Zhao and Jinjun Chen. 2022. A survey on differential privacy for unstructured data content. *ACM Computing Surveys* 54, 10s (2022), 1–28. DOI: https://doi.org/10.1145/3490237

[77] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. 2015. Scalable person re-identification: A benchmark. In *2015 IEEE International Conference on Computer Vision (ICCV '15)*, 1116–1124. DOI: https://doi.org/10.1109/ICCV.2015.133

[78] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. 2016. Learning deep features for discriminative localization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '16)*. IEEE Computer Society, Los Alamitos, CA, 2921–2929. DOI: https://doi.org/10.1109/CVPR.2016.319

Appendices

A Performance

We here include average runtimes of the different components of our system, reported under Table A1, using images from Market1501. These measurements were made on a machine with an Intel Core i9-9900K CPU @ 3.60 GHz and an NVIDIA GeForce RTX 2080 GPU. From these measurements, it appears the computational load is primarily on training the reID model, and, to a lesser extent, on training the attribute classification models. We do not propose these models and believe they can be swapped out for other, similar models, that can be selected for computational efficiency. These models can be trained punctually, at spaced-out points in time, on a central server that only has access to the protected images, so as to lessen the computational cost of training. The privacy components of our system, in italics, which are our main contribution, do not require a training phase and have rather low computational cost. Figure A1 confirms that the runtime performances of most of the different components do not vary much depending on the chosen

Table A1. Average Wall Clock Times on Market1501, Separated by Train/Testing

Component		Training (12,936 images)	Testing (3,368 images)	Testing (1 image)
ε-IDP mechanism	(CPU)	-	1.83 s (±8.09 ms)	0.5 ms
Centroid-based reID	(GPU)	1 h 39 min 6 s (±7.18 s)	1 min 19 s (±1.92 s)	23.5 ms
Attribute recognition	(GPU)	9 min 3 s (±906 ms)	2.55 s (±11.4 ms)	0.8 ms
k_{*1}*-anonymity computation*	(CPU)	-	52 ms (±811 µs)	15.4 µs
k_{*2}*-anonymity computation*	(CPU)	-	655 ms (±7.41 ms)	194.5 µs
k_{*3}*-anonymity computation*	(CPU)	-	5.47 s (±48.3 ms)	1.6 ms
Total		1 h 48 min 9 s (±8.1 s)	1 min 29.6 s (±2.0 s)	26.6 ms

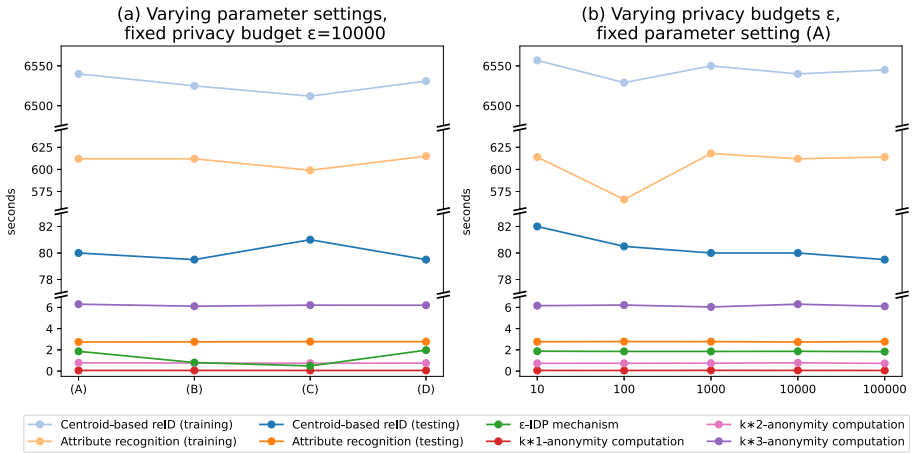


Fig. A1. Average wall clock times of the different components of our system when varying parameters.

Table A2. Average Wall Clock Times of SegCAM-IDP, per Component, on Market1501

Component		3,368 images	1 image
CE2P human parsing	(GPU)	1 min 42 s (±493 ms)	30.3 ms
CAM generation	(CPU)	6 min 59 s (±7.1 s)	124.4 ms
SegCAM-IDP mechanism	(CPU)	3 min 19 s (±312 ms)	59.1 ms
Total		12 min (±7.9 s)	213.8 ms

parameter setting or privacy budget ε. Model-based components (reID and attribute recognition, in blue and orange) show a little variation between runs, which is likely due to GPU availability, but the other components basically do not vary. The only exception to this is the ε-IDP noising mechanism itself (in green), which runs faster when working under higher levels of pixelization (parameter settings B and C).

We also include average runtimes of our proposed SegCAM-IDP mechanism under Table A2, using images from Market1501. The SegCAM-IDP mechanism relies on a pretrained CE2P segmentation model and a pretrained CAM-ready classification model. These models are somewhat computationally expensive to train; however, they only really need to be trained once, and once trained, their practical use is not very resource-intensive.

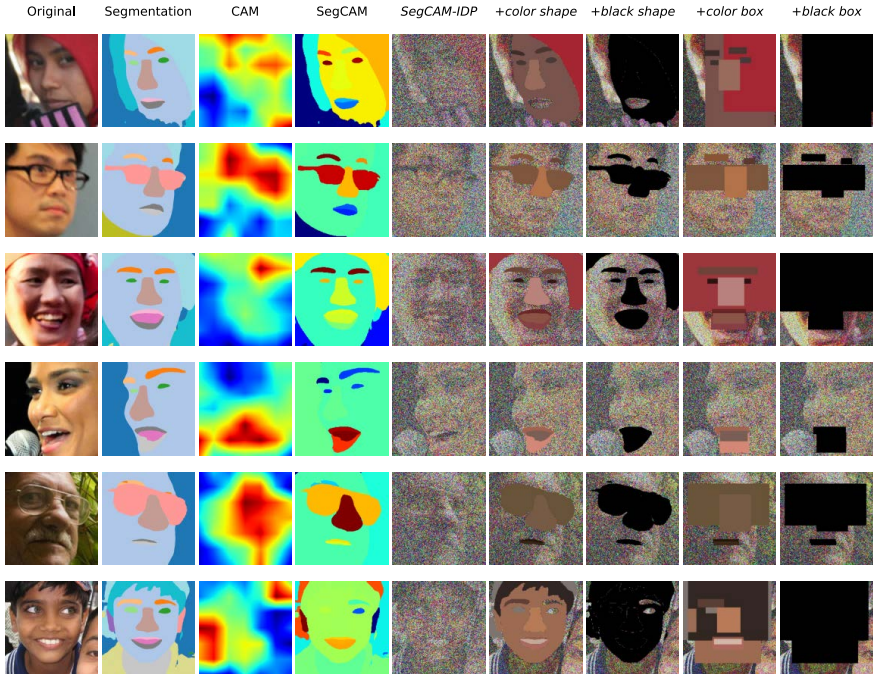


Fig. B1. Additional visualization of SegCAM-IDP (ε = 5,000) on FairFace.

B Visuals

We do not consider SegCAM-IDP for FairFace in the main article, and instead focused our analysis on Market1501, a full-body dataset that is suitable for more tasks. We include under Figure B1 additional visualization when applying the SegCAM-IDP approach on face data from FairFace. We use a face segmentation model based on BiSeNet [65, 71] instead of the previous CE2P human parsing model, which is better suited to facial data. Due to the different nature of images (square-shaped rather than rectangular), the boxing methods (in the last two columns) tend to affect larger parts of the images, which may have a more significant impact on data utility for demographic predictions.

Received 14 September 2024; revised 3 May 2025; accepted 14 May 2025