IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 29, 2020, accepted May 9, 2020, date of publication May 15, 2020, date of current version June 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2994960

PE-MIU: A Training-Free Privacy-Enhancing Face Recognition Approach Based on Minimum Information Units

PHILIPP TERHÖRST[1,2], KEVIN RIEHL[1,2], NASER DAMER[1,2], PETER ROT[3], (Student Member, IEEE), BLAZ BORTOLATO[3], FLORIAN KIRCHBUCHNER[1,2], VITOMIR STRUC[3], AND ARJAN KUIJPER[1,2]

[1]Fraunhofer Institute for Computer Graphics Research IGD, 64283 Darmstadt, Germany
[2]Department of Computer Science, Technical University of Darmstadt, 64289 Darmstadt, Germany
[3]Faculty of Electrical Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia

Corresponding author: Philipp Terhörst (philipp.terhoerst@igd.fraunhofer.de)

This work was supported in part by the German Federal Ministry of Education and Research (BMBF), in part by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) through the National Research Center for Applied Cybersecurity (ATHENE), in part by the ARRS Research Project J2−1734, in part by the German Federal Ministry of Education and Research (BMBF) through the Software Campus Project, and in part by the DOD Counterdrug Technology Development Program Office, FERET Database of Facial Images Collected through the FERET Program.

ABSTRACT Research on soft-biometrics showed that privacy-sensitive information can be deduced from biometric data. Utilizing biometric templates only, information about a persons gender, age, ethnicity, sexual orientation, and health state can be deduced. For many applications, these templates are expected to be used for recognition purposes only. Thus, extracting this information raises major privacy issues. Previous work proposed two kinds of learning-based solutions for this problem. The first ones provide strong privacy-enhancements, but limited to pre-defined attributes. The second ones achieve more comprehensive but weaker privacy-improvements. In this work, we propose a Privacy-Enhancing face recognition approach based on Minimum Information Units (PE-MIU). PE-MIU, as we demonstrate in this work, is a privacy-enhancement approach for face recognition templates that achieves strong privacy-improvements and is not limited to pre-defined attributes. We exploit the structural differences between face recognition and facial attribute estimation by creating templates in a mixed representation of minimal information units. These representations contain pattern of privacy-sensitive attributes in a highly randomized form. Therefore, the estimation of these attributes becomes hard for function creep attacks. During verification, these units of a probe template are assigned to the units of a reference template by solving an optimal best-matching problem. This allows our approach to maintain a high recognition ability. The experiments are conducted on three publicly available datasets and with five state-of-the-art approaches. Moreover, we conduct the experiments simulating an attacker that knows and adapts to the systems privacy mechanism. The experiments demonstrate that PE-MIU is able to suppress privacy-sensitive information to a significantly higher degree than previous work in all investigated scenarios. At the same time, our solution is able to achieve a verification performance close to that of the unmodified recognition system. Unlike previous works, our approach offers a strong and comprehensive privacy-enhancement without the need of training.

INDEX TERMS Biometrics, face recognition, privacy, soft-biometrics, soft-biometric privacy.

I. INTRODUCTION

The face is one of the most used biometric modalities [9], [54]. A typical face recognition system contains feature representations (templates) for each enrolled individual. To verify a subject's identity, a template of this subject probe is computed and compared against the template of the claimed identity [33]. Recent works showed that more information than just the person's identity can be deduced from these templates [10], [49]. This includes information about an individual's gender, age, ethnicity, sexual orientation and health status [10], [55]. However, many applications are not

The associate editor coordinating the review of this manuscript and approving it for publication was Michele Nappi.

permitted by the users to have access to this information. Thus, the stored data should be exclusively used for recognition purposes [28]. Consequently, extracting such information without a person's consent is considered a violation of their privacy [21].

In order to prevent this kind of function creep, *soft-biometric privacy* aims at suppressing or hiding privacy-risk information in face representations. This is further challenged by maintaining a high recognition performance at the same time. Previous works proposed privacy-enhancing solutions based on supervised [25], [28], [31] and unsupervised approaches [46], [47]. While unsupervised approaches show a more comprehensive but weaker privacy-enhancement, supervised approaches are limited to the suppression of pre-defined attributes and thus, are vulnerable to unconsidered function creep attacks.

In this work, we propose PE-MIU, a privacy-preserving face recognition approach based on minimum information units. PE-MIU is a novel, training-free, and privacy-preserving face recognition approach that works on the biometric template-level. Exploiting the structural differences between face recognition and the estimation of facial attributes, our approach divides face templates into small blocks of minimal information units and randomly changes their positions in the templates. Since the information of privacy-sensitive attributes is usually distributed across the template, this approach significantly reduces the chance of function creep attacker to successfully estimate privacy-sensitive information from the modified face templates. To compare two modified templates, and thus verify if these belong to the same identity, we introduce an optimal assignment protocol. In this protocol, the minimal information units of both templates are assigned based on their optimal matching. This assignment is used to align and compare the templates.

The experiments in this work were conducted on three publicly available databases in the context of function creep attackers who know and adapt to the used privacy mechanism. To put the results in a broad perspective, we compare our proposed solution against five state-of-the-art approaches that try to suppress the attribute gender on template-level. The experiments show that PE-MIU outperforms all other approaches in terms of suppressing privacy-risk attributes and maintaining recognition performance. It is able to reach significantly higher gender suppression rates than previous works in all investigated cases, and, at the same time, reaches a face recognition performance close to the unmodified face recognition system. The source code for PE-MIU is available at the following link.[1]

II. RELATED WORK

Despite the high performance of face recognition, it is still prone to certain vulnerabilities towards operation scenarios and attacks. Examples of these vulnerabilities are presentation attack detection (spoofing) [6], [37], face morphing attacks [7], [8], and the inherited differential performance (bias between different demographic groups) [12], [50]. As this work focus on the privacy aspect of face recognition, the discussed related work will go deeper into this issue.

In the context of face biometrics, privacy has been studied from from two perspectives. The first kind focuses on preserving facial characteristics such as gender, age, and expression while de-identifying face images [15], [19], [24], [30]. The second kind aims at preventing the estimation of these facial attributes while maintaining its recognition ability [28]. In this work, we will focus on the latter case that is known as soft-biometric privacy [2], [45]. Solutions for this problem are based on image fusion, perturbations, and adversarial learning and are described in the following.

Suo *et al.* [43] proposed an approach that flips the estimated gender by decomposing the face image and replacing the facial components with similar parts of the opposite gender. This aims at suppressing the gender of the face image. Othman and Ross presented a different approach [31] where they proposed a face morphing methodology that iteratively morphs two images and therefore, suppresses gender information at different levels. However, this resulted in morphed images with significant artefacts.

In [40] and [39], adversarial images created by using a fast flipping attribute technique showed that it was able to fool their network in predicting binary facial attributes. An incremental flipping approach was proposed by Mirjalili and Ross [28] with the use of perturbations. In [5], imperceptible noise was used to suppress k attributes at the same time. However, this noise is trained to suppress attributes from only one specific neural network classifier and consequently, does not generalize to other classifiers.

In [25]–[27], Mirjalili *et al.* proposed semi-adversarial networks consisting of a convolutional autoencoder, a gender classifier, and a face matcher. It enhances the soft-biometric privacy on image level. The autoencoder perturbs the input face image such that it minimizes gender classifier performance while trying to preserve the performance of the face matcher. Training this supervised approach requires a large amount of data with the corresponding privacy-sensitive labels. However, this approach is limited at suppressing pre-defined attributes and thus, it is vulnerable to unseen function creep attacks.

All previously mentioned works were based on image-level. However, most biometric representations are stored in templates rather than images [11], [42] and templates offer a less restricted way of encoding information. Consequently, recent privacy-preserving research was done on template-level [2], [29], [45]–[47]. These works investigate the privacy performance in a more critical and challenging context of a function creep attacker that knows the systems privacy-mechanism and adapts to it. Terhörst *et al.* [45] proposed an incremental variable elimination (IVE) approach to eliminate privacy-risk features from the face templates.

[1] https://github.com/pterhoer/PrivacyPreservingFaceRecognition

Morales *et al.* [29] suppress attribute information via a modified triplet loss. In 2020, Bortolato *et al.* [2] proposed PFRNet, an autoencoder approach that learns privacy-enhancing face representations disentangling identity from attribute information. Since these approaches require privacy-sensitive attribute labels during training, their privacy-protection is further limited to the suppression of these pre-defined attributes.

A more comprehensive privacy-protection is provided by unsupervised methodologies, because these approaches have a more generalized goal of encoding information that does not apply attention mechanisms to single characteristics. In [46], the problem was tackled with two similarity-sensitive noise transformations, cosine-sensitive noise (CSN) and euclidean-sensitive noise (ESN) transformation. These transformations apply specific noise-injections to the face templates that alter the identity information in a controlled manner. In [47], Terhörst *et al.* proposed a negative face recognition (NFR) approach. While ordinary (positive) face templates contain information of the persons identity, negative face templates provide random complementary information about this individual. Storing only negative templates in the database, prevents function creep attackers from successfully predicting privacy-sensitive attributes. While these unsupervised approaches provide a more comprehensive privacy-protection, it is harder to reach high suppression rates while maintaining a high recognition rate as well.

The privacy-issue in biometrics can also be seen from the perspective of cancelable biometrics. Similar to soft-biometric privacy, cancelable biometrics approaches apply one-way functions to transform biometric data [1], [23], [32] and store the transformed data [4]. However, the solutions from both areas target different goals. In cancelable biometrics, the privacy-preservation comes from the computational difficulty to recover the original biometric from the transformed one [32]. The transformed representations aim to achieve irreversibility, revocability, and unlinkability [36]. In contrast to this, soft-biometric privacy does not aim at revocability and non-linkability. It aims at suppressing soft-biometric information in biometric representations, while maintaining a high recognition ability [26], [46].

In this work, we proposed a training-free and template-based privacy-preserving face recognition approach that exploits the structural differences between face recognition and the estimation of facial characteristics by function creep attacker. While the later only requires the input of one face template, for face recognition two templates are necessary. The additional information (in form of a second template) is used to to make the estimation of privacy-sensitive attributes a difficult task. In our PE-MIU approach, this is achieved by representing the template of an identity in a randomized fashion of template blocks. Due to this kind of representations, function creep attackers can only use minimum information units for their attacks, while for face recognition we can exploit the second template to align, and thus compare, both representations.

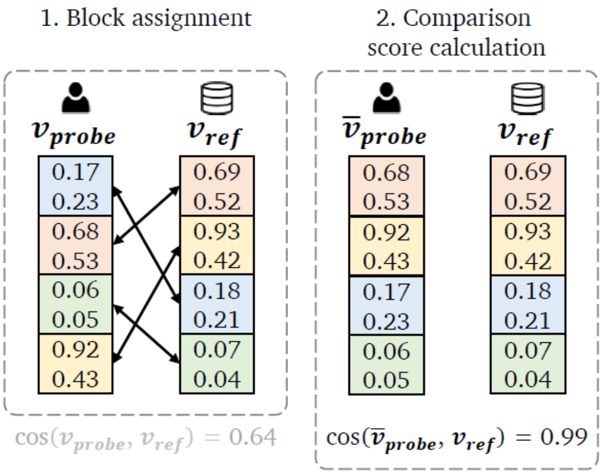


FIGURE 1. Illustration of the comparison of two MIU templates. In the first step, an optimal assignment of the MIU blocks per template are computed. In the second step, this assignment is used to align both templates such that they can be compared with a standard similarity function.

III. METHODOLOGY

Enhancing soft-biometric privacy aims at preventing function creep attackers from successfully predicting privacy-risk characteristics. This task is further challenged by simultaneously maintaining a high recognition ability. With our PE-MIU approach, we exploit the structural differences between a face recognition scenario and the scenario of a function creep attacker. While the function creep attacker aims at the predicting privacy-sensitive information from *one template*, in face recognition *two templates* are compared to decided if they belong to the same identity or not. In this work, we propose a training-free approach for privacy-preserving face recognition, PE-MIU. Our PE-MIU approach divides face templates into small blocks and randomly changes their positions. These blocks are noted as minimum information units (MIU). Consequently, it is hard to reliably predict these characteristics. For the purpose of recognition, two templates are given and their relation to each other can be used to find corresponding MIUs. In the first step, the optimal assignment between the MIU's per template are calculated, to verify if two MIU-based templates belong to each other. In the second step, this assignment is used to align the templates and further compute their comparison score. This idea is illustrated in Figure 1 and detailed in the rest of this section.

A. ENROLMENT PHASE

In the enrolment phase, given a face image I, the corresponding MIU template v is computed and stored in the database. The computation of the MIU-based template is described in Algorithm 1. Given a face image I, the corresponding face embedding $x \in \mathcal{R}^L$ is extracted (*createEmbedding*) from I, where L is the size of the embedding. This face embedding x is divided (*divideMIU*) into L/s MIU blocks of size s. Then, the positions of these units are exchanged randomly (*shuffle*) resulting in a face template v where every entry

Algorithm 1 ComputeMIUTemplate(I, s)

```
Input: Face image $I$, bin size $s = 16$
Output: Face template $v$ to be stored in the database
$x \leftarrow createEmbedding(I)$
$v_{org} \leftarrow divideMIU(x, s)$
$v \leftarrow shuffle(v_{org})$
return $v$
```

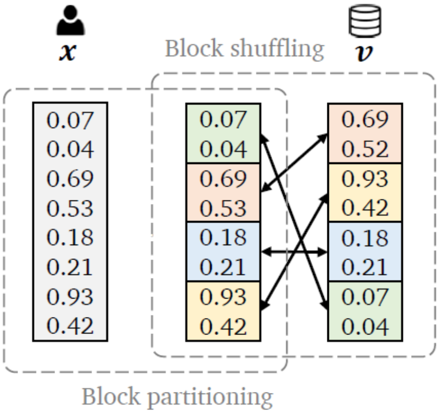


FIGURE 2. Illustration of the key parts during enrolment: first, the face embedding *x* is divided into equally-sized blocks, the MIU. Second, the position of these MIU blocks are randomly shuffled.

consists of a feature block. This MIU template is then stored in the database. The process of dividing the embedding into MIU blocks and shuffling the block positions is illustrated in Figure 2.

B. VERIFICATION PHASE

In the verification phase, an MIU reference template v_{ref} stored in the database is compared with an MIU probe template v_{probe} from a captured individual. The verification is done in two steps: first, the MIU-blocks of v_{probe} and v_{ref} have to be assigned such that there is an optimal pair-wise matching between the blocks of both templates. Second, the probe template v_{probe} is aligned to v_{ref} such that the matched MIUs are at the same entries of the templates. The aligned probe template $\hat{v}_{probe}$ is then compared to v_{ref} using a similarity metric. In our case, as will be explained later, we use the cosine similarity metric as it was recommended to be used with the original templates in our experiments.

1) MIU-BLOCK ASSIGNMENT

In order to compare a probe embedding x with a block-wise reference template v_{ref}, an MIU representation of x have to be computed and the optimal matching of the MIU-blocks of each template has to be found. Similarly to the block partitioning during the enrolment (see Figure 2), the probe face embedding x is divided into MIU of size s, resulting in v_{probe}. Then, the best matching between the two MIU templates is computed. In graph theory, this problem is known as weighted bipartite matching problem [38] and is equivalent to the following optimization

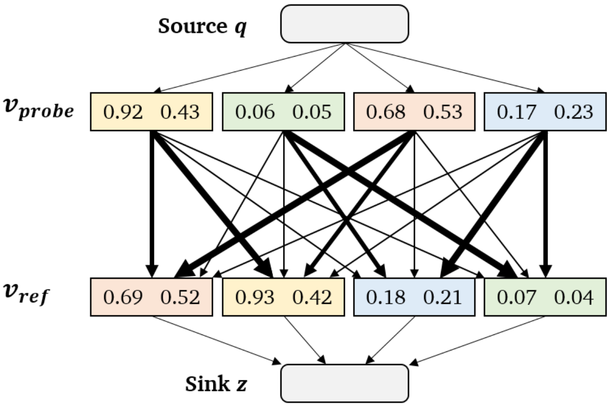


FIGURE 3. The MIU-block assignment can be solved as a maximum-flow problem where the maximum flow from source *q* to sink *z* is defined by the optimal matchings. The weights between the MIU-blocks of v_{probe} and v_{ref} are given by their euclidean distances. Broader edges represent higher weight values.

$$\min_{\chi} \sum_{i,j} C_{i,j}\, \chi_{i,j}. \tag{1}$$

Applied to our problem, the cost matrix $C_{i,j}$ describes the euclidean distance between then MIU i and j and $\chi_{i,j}$ is the resulting binary assignment matrix with $\chi_{i,j} = 1$ if and only if the i^{th} probe MIU is assigned to the j^{th} reference MIU. This problem can be solved via the Hungarian [22] or the Ford-Fulkerson [14] algorithm.

The best matching task can be formulated and solved as a minimum cost maximum-network-flow problem [3]. Therefore, an acyclic graph is constructed as shown in Figure 3. The edge weights from source q to the nodes (MIU blocks) of v_{probe} are set to 1. The same applies for the weights of the edges from the v_{ref} nodes to the sink z. The weights for the edges connecting the blocks between v_{probe} and v_{ref} are determined by its euclidean distances resulting in the cost matrix C. The optimal assignment of these MIUs is then defined by the maximum flow from source q to sink z.

2) ALIGNED COMPARISON SCORE

After the optimal block assignment is found, the order of the blocks of the probe template are chosen such that the matched blocks are at the same positions. This results in an aligned probe template $\overline{v}_{probe}$. After the MIU-blocks of the probe and the reference templates $\overline{v}_{probe}$ and v_{ref} are aligned, the comparison score $cs(\overline{v}_{probe}, v_{ref})$ of these templates is computed. In this work, we use cosine similarity for this comparison score calculation. However, this metric should be chosen according to the utilized face embeddings.

3) SUMMARY OF THE VERIFICATION PHASE

To summarize the verification phase using MIU-templates, Algorithm 2 describes how a comparison score between a probe face image I of an identity and a reference template v_{ref}

(stored in the database) of the claimed identity is computed. First, a MIU-based template v_{probe} is computed from image i using Algorithm 1 (*computeMIUTemplate*). Then, the probe template v_{probe} is aligned to v_{ref} (*eqnarray*) as described in Section III-B1. Finally, the aligned probe template $\overline{v}_{probe}$ is compared with v_{ref} using a pre-defined comparison score metric $cs(\overline{v}_{probe}, v_{ref})$ such as cosine similarity. This comparison score is then used to determine if the person belongs to the claimed identity.

Algorithm 2 Compare(I_{probe}, v_{ref})

```
Input: Face image $I_{probe}$, claimed-identity template $v_{ref}$
Output: Comparison score $score$
1: $v_{probe} \leftarrow computeMIUTemplate(I, s = 16)$
2: $\overline{v}_{probe} \leftarrow eqnarray(v_{probe}, v_{ref})$
3: $score \leftarrow cs(\overline{v}_{probe}, v_{ref})$
4: return $score$
```

C. PROPERTIES OF BLOCK-WISE REPRESENTATIONS

The soft-biometric privacy-protection of the proposed method lies in the randomized nature of the MIU representation. Due to the fact that the previous order of the MIU-blocks is unknown and can only be reconstructed with an unmodified face embedding of the same identity, function creep attackers can only use the set of the minimal information units for their attacks.

Soft-biometric privacy usually describes a trade-off between suppressing privacy-sensitive attributes and maintaining the recognition ability of its templates. In this work, this trade-off is determined by the size of the MIU blocks s. Higher MIU sizes result in a weaker privacy-protection, due the fact that higher block sizes contain more attribute information. However, higher MIU sizes also leads to less misassigned MIU-blocks and thus, it leads to a lower recognition errors as well. For this work, we choose an MIU size of $s = 16$ to balance these two points. The effect of changing this parameter is investigated in Section V-C.

The key part of verifying a persons identity with the proposed method is the MIU-block assignment. As indicated in Figure 1, the comparison of two not-aligned MIU templates results in a weak recognition performance. The block assignment, needed for the computation of the aligned MIU templates, is done via the Hungarian algorithm [22], since it provides stable and optimal assignments. This method scales with $\mathcal{O}(n^3)$, where $n = L/s$ is the number of MIU-blocks per template. Consequently, higher privacy-protection (smaller s) comes at the cost of higher computation times. However, this can be mapped to a complexity of $\mathcal{O}(n^2 \log n)$ by using the approach presented from Ramshaw and Tarjan [38].

IV. EXPERIMENTAL SETUP

A. DATABASE

We conduct experiments on the publicly available ColorFeret [35] and Adience [13] and Labeled Faces in the Wild (LFW) [17] databases to evaluate and compare our solution to related works. ColorFeret [35] consists of 14,126 images from 1,199 different individuals with different poses under controlled conditions. The Adience dataset [13] consists of 26,580 images from over 2,284 different subjects under uncontrolled imaging conditions. Labeled Faces in the Wild (LFW) [17] provides 13,233 face images from 5749 identities. The databases cover a wide range of variations in illumination, focus, blurriness, pose, and occlusions. Moreover, the databases include information about the identities and their gender. This allows to deeply investigate the privacy-preservation techniques of the attribute gender, as well as their recognition performances.

B. EVALUATION METRICS

Preserving soft-biometric privacy is challenged by a trade-off between the desired degradation of the attribute estimation performance by function creep attackers and the desired preservation of the recognition ability. In the experiments, we report the verification performances in terms of false non-match rate (FNMR) at fixed false match rates (FMR). We further report the equal error rate (EER), which equals the FMR at the threshold where FMR $= 1 -$ FNMR. Both verification performance measures are defined in the ISO standard [18]. In order to evaluate the attribute suppression performance, we report the results in terms of balanced attribute classification accuracy, since this allows an unbiased performance measure on testing data with unbalanced attribute information. This balanced accuracy is equivalent to the standard accuracy with class-balanced sample weights. A value of 50% is the best possible case for a privacy-preserving methodology and the worst outcome for a function creep attacker. In order to evaluate if the privacy enhancing method is beneficial, we use the privacy gain identity loss (PIC) coefficient defined in [46] and reported in recent works [2]. The PIC is defined as

$$PIC = \frac{AE' - AE}{AE} - \frac{RE' - RE}{RE}. \tag{2}$$

The value is defined by attribute prediction errors AE' and AE and the verification errors RE' and RE with and without the privacy-preserving methodology. Positive values indicate that the privacy gain is higher than the loss in the identity preservation performance. Since it measures how beneficial it is to apply the privacy transformation, a higher PIC coefficients indicates a better privacy-enhancing technique.

C. FACE RECOGNITION MODEL

In this work, our block-assignment-based approach builds on arbitrary face embeddings of certain dimensions. In the experiments, we utilize the widely used FaceNet model[2] [41] pretrained on MS-Celeb-1M [16]. In order to extract an embedding of a face image, the image is aligned, scaled, and cropped as described in [20]. The preprocessed face image is then passed into the face recognition model to obtain

[2] https://github.com/davidsandberg/facenet

a 128-dimensional face embeddings. The comparison of two such embeddings is performed using cosine-similarity.

D. FUNCTION CREEP ATTACKS

In this work, we consider two kinds of function creep attacks, the standard attack (S-ATK) and the advanced attack (A-ATK). We decided to introduce A-ATK due to the limited effectiveness of S-ATK on our proposed approach. Both attacks evaluate the attribute suppression performance and simulate the critical scenario of a function creep attacker that knows the systems privacy mechanism and adapts to it.

For the S-ATK, the adaptation is done by training (function creep) classifiers on the privacy-enhanced templates to predict the privacy-sensitive attributes. Before the training of these classifiers, the transformed templates are further normalized and scaled to unit-length. The utilized classifiers include random forest (RF), support vector machines (SVM), k-nearest neighbours (kNN), and logistic regression (LR). The hyperparameters of these classifiers are fine-tuned with Bayesian optimization.

During the experiments, we realized that these naive function creep attacks (S-ATKs) show only a very limited effect on our proposed approach, meaning that the classification performance with optimized function creep classifiers show a close to random behaviour. Therefore, we additionally considered more challenging attack classifier approaches for our proposed solution that is directly customised to achieve the highest classification accuracies. The most successful kind of attacks were the ones that learn to predict the gender for each MIU-block separately. During prediction, each of these blocks of a face template is classified separately and the predicted scores per classed are fused with a mean-fusion-rule [44], [48]. In this work, we refer to this attack as A-ATK.

For the evaluation, we consider function creep attacks to the privacy-sensitive attribute gender as done in previous works [2], [25]–[27], [29], [46], [47]. The reason for this choice is that gender information can be estimated from face templates with very high accuracies [48], [49]. Moreover, it requires only a binary decision, which makes it an easy target for function creep attackers and a challenge for privacy-preserving methodologies.

E. BASELINE APPROACHES

To evaluate our proposed training-free and template-based solution in a broad setting, we compare it against 5 recent template-based privacy-preserving face recognition approaches. These include the two supervised solutions PFRNet [2] and IVE [45] and three unsupervised solutions NFR [47], CSN [46], and ESN [46]. PFRNet [2] aims at learning a feature representation that disentangle identity from gender. The original network was optimized for an embedding size of 512. In our evaluation setting an embedding size of 128 is used. Consequently, the network was adapted such that the encoder consists of two layer with size 128 and 100+28 dimensions and the decoder consists of two layer with 128 dimensions as this adaptation showed the best privacy-preserving performance while maintaining high verification rates. IVE [45] incrementally eliminate the most privacy-risk features from a face template to suppress the attribute information. CSN and ESN [46] are based on geometric-inspired noise-injections that alter the inherent identity information in a controlled manner. In contrast to mentioned approaches, NFR [47] stores only complementary information about an individual in a face template and during deployment, it compares the probe template with a reference template in the complementary domain by calculating its dissimilarity.

In order to make the experiments as comparable as possible, we calibrated the hyperparameters of these baselines in such a way that they reach a similar verification EER performance if possible. For all experiment scenarios, the same subject-exclusive 5-fold cross-validation setup is utilized. This includes training the function creep classifiers, as well as training the baseline approach for privacy-enhancing face recognition. The setup is shown in Table 1. For the three utilized databases, it provides details about each fold properties. It should be noted that for LFW, the gender distribution is unbalanced. For this reason, we choose the balanced attribute classification accuracy as described in Section IV-B.

F. INVESTIGATIONS

The investigations of this work are divided in four parts:

A. We analyse the face verification performance of our privacy-enhancing solution in comparison to previous works.
B. We investigate the attribute prediction performance in a qualitative and quantitative manner.
   1) The qualitative investigation provides a qualitatively aided analysis of the gender separability of the original and the MIU-based templates. This is done by providing a visual understanding of the proposed approach.
   2) The quantitative investigation analyses the attribute prediction performance of the original template, on our solution, and on state-of-the-art. This is done in the critical scenario of a function creep attacker that adapts to the systems privacy mechanism using the attack scenarios S-ATK and A-ATK.
C. We analyse the parameter space of our solution to provide a deeper understanding of the influence of the MIU-block size on several aspects of our methodology.
D. Lastly, we focus on the strongest attack for each method and summarize the methods recognition ability, as well as the privacy-protection in a joint manner. This includes reporting the privacy-gain identity-loss coefficients (PIC) to measure and compare the usefulness of the studied approaches in the context of the most successful function creep attacks.

V. RESULTS

A. FACE VERIFICATION PERFORMANCE

In Figure 4, the face verification performance is shown on three databases. The performance of the original

TABLE 1. Properties of the used cross-validation setup for the three utilized datasets, ColorFeret, Adience, and LFW. The number of samples, identities, and the percentage of female individuals are reported per training and testing fold.

		Testing fold					Training fold				
Dataset		0	1	2	3	4	0	1	2	3	4
ColorFeret	# samples	2160	2162	2159	2159	2159	8639	8637	8640	8640	8640
	# identities	189	191	190	190	190	761	759	760	760	760
	Ratio of Female	63.6%	68.2%	58.6%	60.5%	67.9%	63.8%	62.6%	65.0%	64.6%	62.7%
Adience	# samples	3868	3868	3868	3868	3867	15471	15471	15471	15471	15472
	# identities	456	456	457	457	456	1826	1826	1825	1825	1826
	Ratio of Female	55.5%	47.2%	46.2%	45.9%	47.0%	46.6%	48.7%	48.9%	49.0%	48.7%
LFW	# samples	2629	2629	2628	2628	2628	10513	10513	10514	10514	10514
	# identities	1137	1145	1146	1146	1147	4584	4576	4575	4575	4574
	Ratio of Female	20.7%	23.1%	26.0%	23.0%	19.8%	23.0%	22.4%	21.7%	22.4%	23.2%

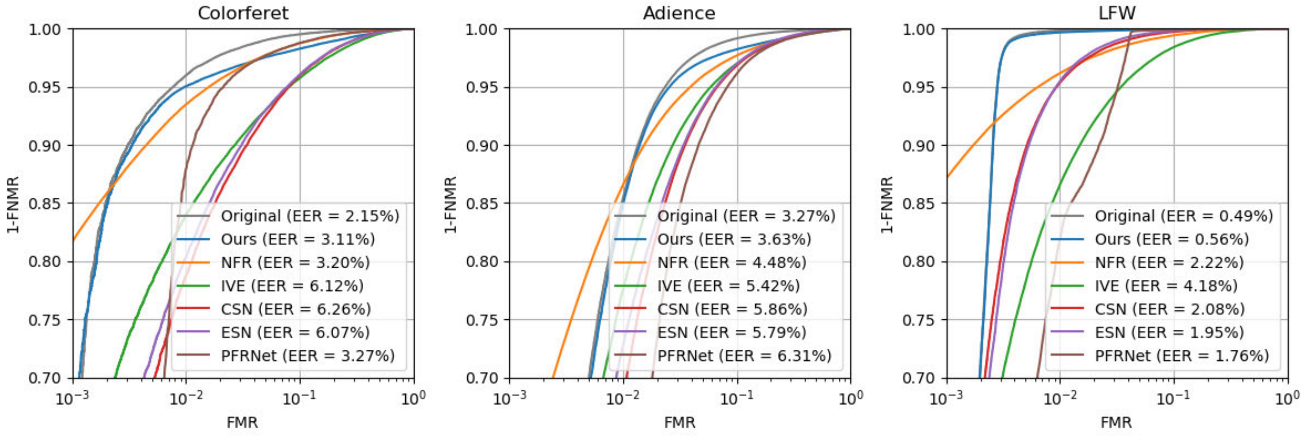


FIGURE 4. Face recognition performance of our MIU-based solution in comparison with five state-of-the-art approaches on three databases. In addition, the face verification performance of the unmodified (original) face templates are shown.

FaceNet embeddings is shown along the performance of six privacy-enhancing approaches including our proposed approach. It can be seen that all approaches show a degraded face verification performance compared to the original embeddings. This is shown in every privacy-enhancing work [27], [29], [45], [46], since soft-biometric privacy defines a trade-off between maintaining identity information and suppressing privacy-sensitive attributes. For lower FMR, NFR [47] is an exception of this trade-off. Since in the NFR approach the comparison score is computed by the dissimilarity between the positive probe and the negative reference template, it is more robust to embeddings with more intra-class variations. In total, our proposed approach shows the most similar verification performance to the original templates. This can be noticed by both, the ROC curves and the EER values. The recognition performance is mostly maintained, due to nearly error-free MIU assignments.

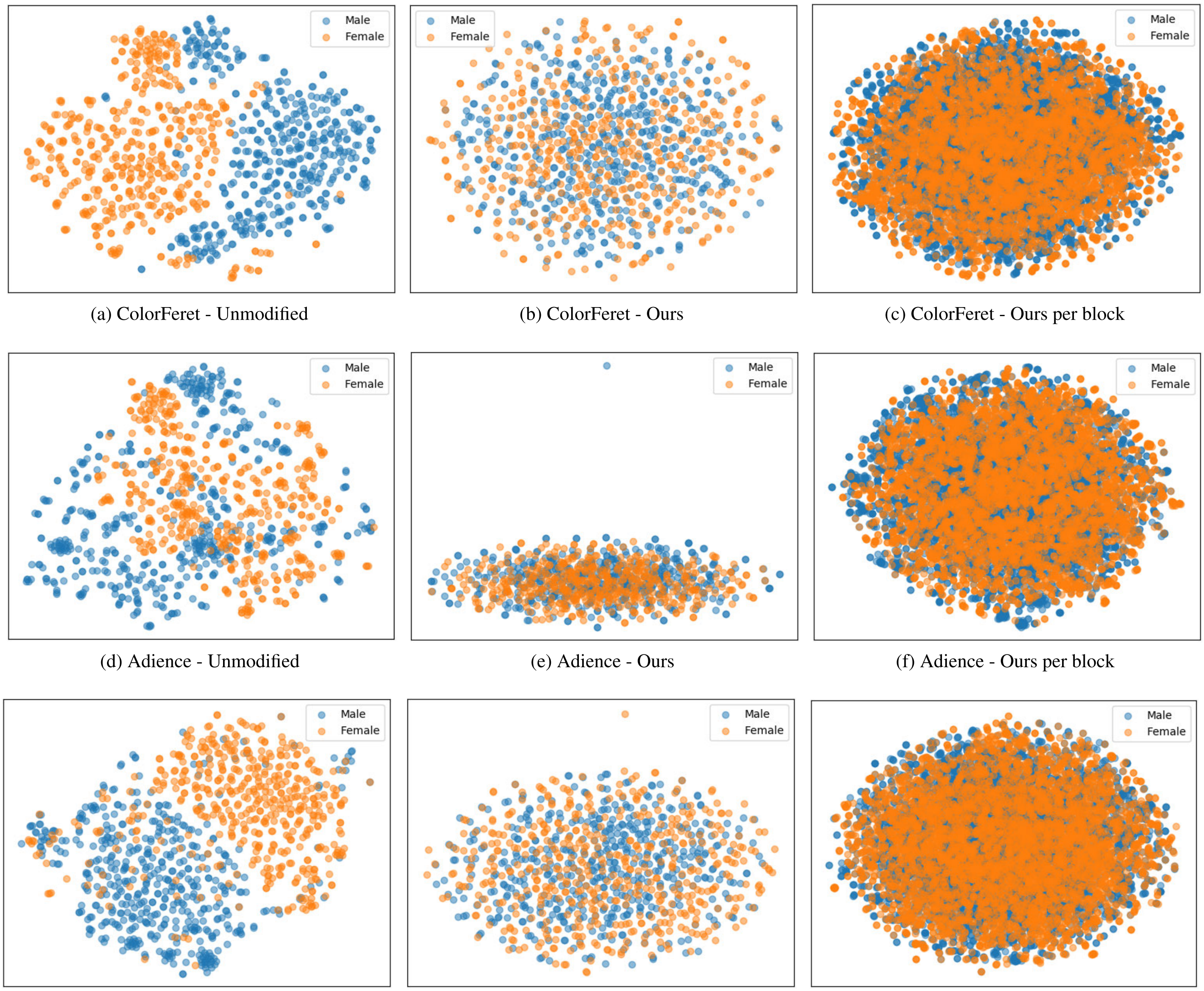
B. ATTRIBUTE SUPPRESSION PERFORMANCE

1) QUALITATIVE ANALYSIS

In order to provide a visual understanding of the proposed approach towards the suppression of gender characteristics, Figure 5 represents a 2D-visualization of 1000 randomly chosen identity-embeddings. The visualization was done by utilizing t-distributed stochastic neighbour embeddings (t-SNE) [51]. Female samples a characterized by orange dots, while male samples are represented by male points. The visualizations are provided on three databases for the unmodified (original) templates (a, d, g), for the MIU-based templates of our PE-MIU approach (b, e, h), and for individual MIU-blocks individually (c, f, i). A clear separation is observed in the visualizations of the unmodified images (a, d, g) indicating that the attribute gender can be correctly predicted to a high degree. In contrast, the plots visualizing our approach (b, e, h) show highly randomized patterns indicating that it is hard to reliably estimate the correct attribute. In order to show that the same applies for individual MIUs separately, Figure 5 (e, f, i) show the same visualization per MIU. Similarly as for our full approach (e, f, i), no pattern between the different gender classes is easily observable.

2) QUANTITATIVE ANALYSIS

To deeply understand the privacy-enhancement of our solution along with previous works, Table 2 shows the balanced accuracies for four optimized function creep classifiers on the three databases. The gender prediction performance is shown



(a) ColorFeret - Unmodified

(b) ColorFeret - Ours

(c) ColorFeret - Ours per block

(d) Adience - Unmodified

(e) Adience - Ours

(f) Adience - Ours per block

(g) LFW - Unmodified

(h) LFW - Ours

(i) LFW - Ours per block

FIGURE 5. **Visualization of different face representations of the three databases, ColorFeret (5a–5c), Adience (5d–5f), and LFW (5g–5i). For the visualizations, 500 female and 500 male identities were chosen randomly and their templates are reduced to two dimensions using t-distributed stochastic neighbour embedding (t-SNE) [51]. The first row (a, d, g) shows the t-SNE plots for the original facenet embeddings. The second row (b, e, h) shows the same plots for templates modified by our PE-MIU approach. The last row represents the t-SNE plots created from each block of the templates. The lower separability introduced by PE-MIU in comparison to the unmodified templates is demonstrated by the increasingly overlapping samples.**

TABLE 2. **Balanced gender decision accuracies on the three databases ColorFeret, Adience, and LFW. The gender prediction performance is determined by four function creep classifiers on the unmodified (original) templates, on our MIU-based templates, and the templates created by previous works using S-ATK. The results showing the most randomized accuracies per classifier are highlighted. Ours* represents the highly challenging attack methodology (A-ATK) that was specifically designed to maximize successful predictions on modified MIU-based templates as explained in Section IV-D.**

	ColorFeret				Adience				LFW			
Method	KNN	LR	RF	SVM	KNN	LR	RF	SVM	KNN	LR	RF	SVM
Original	97.55%	96.90%	91.62%	97.62%	84.37%	87.32%	84.52%	89.81%	85.99%	86.02%	66.61%	89.50%
IVE [45]	89.57%	79.35%	75.28%	95.34%	77.53%	63.72%	76.37%	84.48%	65.76%	62.12%	64.78%	74.76%
ESN [46]	96.01%	91.37%	83.17%	95.60%	84.11%	78.21%	80.89%	88.23%	77.16%	71.20%	56.02%	84.02%
CSN [46]	92.06%	89.33%	82.88%	86.40%	80.03%	75.28%	74.09%	62.15%	69.41%	66.46%	55.56%	82.56%
NFR [47]	91.95%	89.13%	72.68%	92.24%	79.15%	72.30%	61.18%	80.51%	87.86%	76.27%	77.55%	85.07%
PRFNet [2]	81.17%	84.93%	65.95%	82.00%	**54.53%**	58.25%	52.70%	64.63%	82.16%	78.97%	54.43%	83.38%
Ours	**67.64%**	**63.32%**	**51.87%**	**68.71%**	45.45%	**44.71%**	**50.63%**	**49.94%**	**55.38%**	**52.69%**	**50.23%**	**63.54%**
Ours*	69.73%	66.01%	54.97%	70.51%	66.75%	56.12%	66.36%	70.90%	59.98%	53.84%	53.41%	67.93%

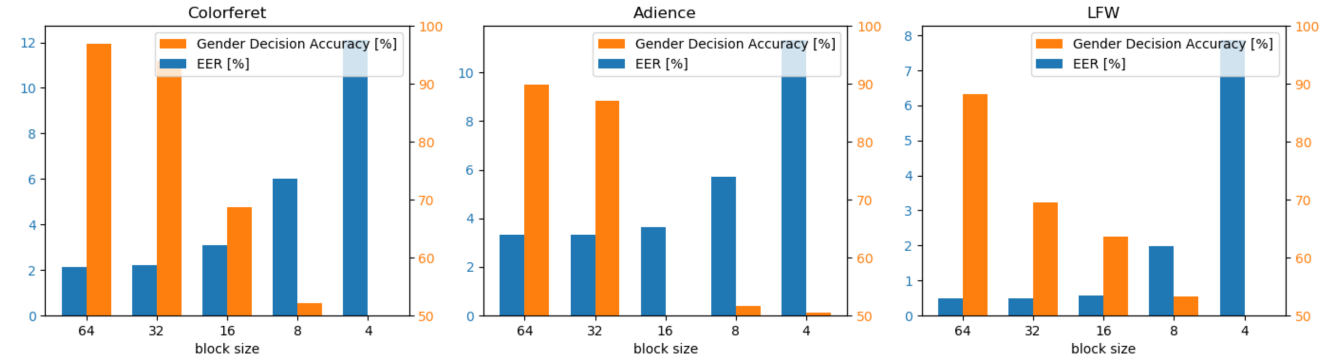


FIGURE 6. Investigation of the block size: the gender decision performance and the recognition EER over the different block sizes are shown for the three databases. The gender decision accuracy comes from the most successful function creep estimator in Table 2, the SVM.

for the unmodified (original) templates, for the our PE-MIU approach (Ours), and for five state-of-the-art solutions. Moreover, the results for a highly challenging attack methodology (A-ATK), directly designed to maximize successful attacks on our approach, is shown as Ours*.

The gender decision accuracies of the original FaceNet embeddings show high values, demonstrating the need for privacy-enhancing technologies. The state-of-the-art privacy-preserving face recognition approaches lead to degraded estimation performances. However, the gender decision accuracies, and thus the resulting attribute suppression, varies a lot depending on the utilized database and function creep classifier. Generally, the highest privacy-improvement is observed for our proposed approach. The function creep classifiers achieve correct classification performances close to a random decision behaviour of 50% in most cases. One exception is the scenario where the KNN estimator was used on the Adience database. Here, PFRNet reaches a slightly more randomized behaviour (54.53%) than our proposed approach (45.45%). However, this comes at the cost of a lower verification performance, e.g in terms of EER where the original templates achieve an EER of 3.27%, our PE-MIU approach reaches an EER of 3.63%, and PFRNet reaches an EER of 6.31%.

In the last row of Table 2, the suppression performance of our PE-MIU approach (Ours*) in the context of an highly advanced and adapted attack methodology (A-ATK) is shown. It demonstrates a strong gender suppression performance can be achieved even in this more critical and challenging attack scenario.

C. PARAMETER ANALYSIS

The block size s of PE-MIU is the key to determine the privacy trade-off between reaching high attribute suppression rates and maintaining a high recognition performance. Therefore, this parameter is investigated in this Section on the three databases ColorFeret, Adience, and LFW. Figure 6 analyses the influence of the block size on the two aspects of the mentioned trade-off. High block sizes lead to lower recognition EER, since the number of possible wrongly-assigned MIU-blocks is lower. At the same time, high block sizes contain more patterns that allow function creep classifiers to successfully predict privacy-risk attributes. Figure 6 shows that a block size $s = 16$ represents a good balance between both aspects of the soft-biometric privacy trade-off.

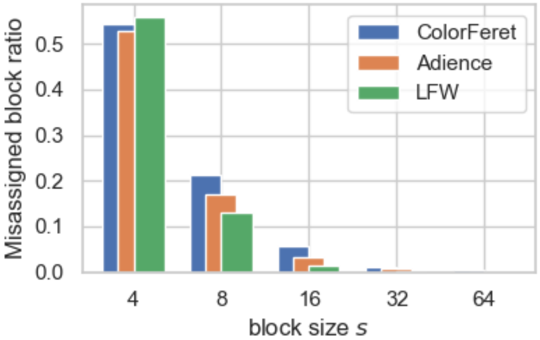


FIGURE 7. Analysis of misassigned MIU-blocks. The average ratio of misassigned blocks per genuine pair comparison is shown for different block sizes on the used databases. Higher block sizes reduce the possibility of misassignments.

The block size s also determines the computational complexity of our proposed MIU-based privacy-preserving face recognition approach. Table 3 shows the average computation time needed for the different MIU-steps. All computational efficiency analyses are based on using a personal computer with an Intel(R) Core(TM) i7-7700 processor. During enrolment, the MIU-template must be generated. This can be implemented efficiently[3] and thus, can be performed in the order of a few microseconds per template. During the verification phase, the MIU-blocks are assigned[4] and then, the aligned templates are compared.[5] The biggest part of the computation time is needed for the MIU-block assignment. The average comparison time of previous works [2], [46] is around 0.10*ms* using the same CPU. Consequently, the strong

[3] In this work, we implemented this part with Numpy [52].

[4] For the block assignment, the Hungarian algorithm implementation from SciPy [53] was used.

[5] The comparison calculation with cosine similarity was computed with Scikit-learn [34].

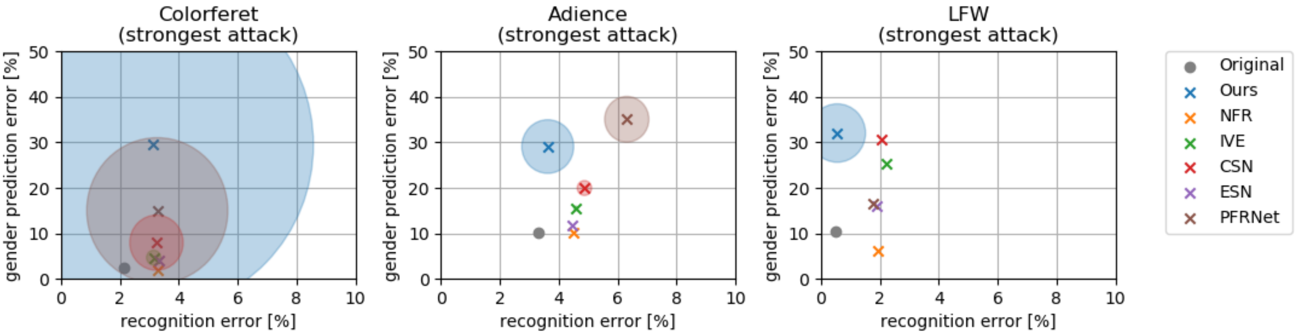


FIGURE 8. Joint analysis of the privacy-risk attribute suppression performance (in terms of balanced accuracies) and the verification performance (in terms of EER). The performance of the unmodified (original) templates (grey dot) is shown in comparison with our MIU-based approach (ours) and five approaches from previous work. The prediction errors refer to the individually most successful attack classifier. To visually encode the usefulness in this challenging attack scenario, PIC values are calculated and represented as shaded areas around the method markers. Negative PIC values are neglected in the plot.

TABLE 3. Average computational time (in *ms*) of the different MIU steps for different block sizes *s*. The values represent the computational time on for an Intel(R) Core(TM) i7-7700 CPU with 3.60 GHz on 128-dimensional templates. The template generation refers to the enrolment phase, while the other steps refer to the verification phase.

Timings $[ms]$	Block size s				
	4	8	16	32	64
MIU-template generation	0.0039	0.0022	0.0014	0.0010	0.0008
MIU-block assignment	7.10	2.23	0.67	0.33	0.06
Comparison score calculation	0.10	0.10	0.10	0.10	0.10
Complete MIU-verification	7.20	2.33	0.77	0.43	0.16

privacy-enhancement and recognition performance of our proposed approach comes at the cost of higher comparison times.

In order to analyse the susceptibility to errors, Figure 7 shows the average ratio of misassigned blocks per genuine pair comparison. These statistics are shown for different block sizes s and for the three utilized datasets. For small block sizes (e.g. $s = 4$), the MIU-blocks contain few information for a reliable assignment. In this case, around 50% of the blocks are incorrectly assigned, explaining the relatively low recognition performance for $s = 4$ in Figure 6. On the other hand, for large block sizes (e.g. $s = 32, 64$), the ratio of misassigned blocks is close to zero (0.5% on ColorFeret, 0% on Adience and LFW, when $s = 64$) and thus, the templates are perfectly aligned in nearly all cases. The perfect alignment leads to low recognition errors. However, it also leads to higher gender decision accuracies of the function creep estimators, as it is demonstrated in Figure 6. In this work, beside analysing different block sizes, we decided to we decided to use a block size of $s = 16$, since it provides a suitable trade-of between maintaining the recognition ability and achieving a high privacy-enhancement, as supported by the information presented in Figure 7. For $s = 16$, the ratio of misassigned blocks varies between $1-5\%$ on the different databases (5% on ColorFeret containing profile face images). This shows that in most cases, two genuine MIU-templates are close to perfectly aligned. At this block size, genuine MIU-blocks that are very similar can be wrongly assigned. However, since these are very similar to each other, the aligned MIU-templates are similar as well and thus, these misassigned blocks have only a minor impact on genuine comparisons.

D. SUMMARY AND USABILITY

Soft-biometric privacy is challenged by maintaining a high recognition performance and degrading the prediction performance of privacy-sensitive attributes. In Figure 8, both aspects can be observed simultaneously under the critical scenario of the most successful individual function creep attack. This is shown for each of the three databases. The x-axes represents the recognition error in terms of EER while the y-axes shows the balanced gender prediction error. Consequently, a highly successful privacy-enhancing solution can be found in the top left corner. Moreover, the PIC coefficient is calculated and represented by the radius of the shaded area around a marker. Since PIC measures the advantageous of applying the privacy-preserving methodology (see Section IV-B, a bigger shaded area represents a high usefulness of applying a solution. As demonstrated our proposed solution achieves the lowest recognition error on all scenarios. Moreover, it also leads to the highest gender prediction errors in most cases and to the highest PIC coefficients in all cases. Consequently, the PIC values (represented as the shaded areas) indicate that our proposed approach is significantly more effective than previous work.

VI. CONCLUSION

Face recognition systems extract and store face templates during enrolment. This enables the recognition of individuals during deployment. However, these templates contain various privacy-sensitive information that can be automatically obtained by function creep estimators. For many applications, these templates are expected to be used for recognition purposes only, raising major privacy issues. Previous work proposed two kinds of solutions. The first kind offers a stronger privacy-enhancement that is however limited to pre-defined attributes. The second kind does not have this limitation but does not achieve the same level of privacy-enhancement. In this work, we propose PE-MIU, a training-free and privacy-preserving face recognition approach based on minimum information units (MIUs). Our solution exploits the structural differences between the different setups of face recognition and facial attribute estimation. This is achieved by dividing a face template into several MIU-blocks and randomly changing their position in the template. This kind of randomized representations changes the pattern of its attributes for each template. Consequently, it is hard for function creep attackers to predict these privacy-sensitive attributes. The experiments were conducted on three publicly-available databases comparing our solution to five state-of-the-art approaches. In the experiments, we simulated function creep attackers that know about the systems privacy mechanism and adapt their attacks based on it. The results show that our novel face recognition approach is able to consistently reach low attribute prediction rates in all investigates scenarios, outperforming all state-of-the-art approaches in most cases. Simultaneously, our solution maintains its recognition ability to a significantly higher degree then previous work. Consequently, unlike previous work, the proposed methodology is characterized by its ability of maintaining a high recognition performance while reaching high attribute suppression rates which are not limited to the suppression of predefined attribute. However, this highly effective solution comes at the cost of a higher comparison time which grows from 0.10 *ms* to 0.77 *ms* per comparison. Future work might focus on solving this issue.

REFERENCES

[1] H. B. Alwan and K. R. Ku-Mahamud, "Cancellable face biometrics template using alexnet," in *Applied Computing to Support Industry: Innovation and Technology*, M. I. Khalaf, D. Al-Jumeily, and A. Lisitsa, eds. Cham, Switzerland: Springer, 2020, pp. 336–348.

[2] B. Bortolato, M. Ivanovska, P. Rot, J. Krizaj, P. Terhörst, N. Damer, P. Peer, and V. Struc, "Learning privacy-enhancing face representations through feature disentanglement," in *Proc. 15th IEEE Int. Conf. Autom. Face Gesture Recognit., FG*, Buenos Aires, Argentina, May 2020, pp. 1–8.

[3] R. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2012.

[4] A. Castiglione, K.-K.-R. Choo, M. Nappi, and F. Narducci, "Biometrics in the cloud: Challenges and research opportunities," *IEEE Cloud Comput.*, vol. 4, no. 4, pp. 12–17, Jul. 2017.

[5] S. Chhabra, R. Singh, M. Vatsa, and G. Gupta, "Anonymizing k facial attributes via adversarial perturbations," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 656–662.

[6] N. Damer and K. Dimitrov, "Practical view on face presentation attack detection," in *Proc. Brit. Mach. Vis. Conf. BMVC*, R. C. Wilson, E. R. Hancock, W. A. P. Smith, eds. New York, NY, USA: BMVA Press, Sep. 2016, pp. 1–11.

[7] N. Damer, A. M. Saladie, A. Braun, and A. Kuijper, "MorGAN: Recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–10.

[8] N. Damer, A. M. Saladie, S. Zienert, Y. Wainakh, P. Terhorst, F. Kirchbuchner, and A. Kuijper, "To detect or not to detect: The right faces to morph," in *Proc. Int. Conf. Biometrics (ICB)*, Jun. 2019, pp. 1–8.

[9] N. Damer, Y. Wainakh, V. Boller, S. V. D. Berken, P. TerhOrst, A. Braun, and A. Kuijper, "CrazyFaces: Unassisted circumvention of watchlist face identification," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–9.

[10] A. Dantcheva, P. Elia, and A. Ross, "What else does your biometric data reveal? A survey on soft biometrics," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 3, pp. 441–467, Mar. 2016.

[11] S. Dey, S. Barman, R. K. Bhukya, R. K. Das, B. C. Haris, S. R. M. Prasanna, and R. Sinha, "Speech biometric based attendance system," in *Proc. 20th Nat. Conf. Commun. (NCC)*, Feb. 2014, pp. 1–6.

[12] P. Drozdowski, C. Rathgeb, A. Dantcheva, N. Damer, and C. Busch, "Demographic bias in biometrics: A survey on an emerging challenge," 2020, *arXiv:2003.02488*. [Online]. Available: https://arxiv.org/abs/2003.02488

[13] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 12, pp. 2170–2179, Dec. 2014.

[14] D. R. Ford and D. R. Fulkerson, *Flows in Networks*. Princeton, NJ, USA: Princeton Univ. Press, 2010.

[15] R. Gross, L. Sweeney, F. de la Torre, and S. Baker, "Model-based face de-identification," in *Proc. Conf. Comput. Vis. Pattern Recognit. Workshop (CVPRW)*, Jun. 2006, p. 161.

[16] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "Ms-celeb-1m: A dataset and benchmark for large-scale face recognition," in *Computer Vision—ECCV* (Lecture Notes in Computer Science), vol. 9907, B. Leibe, J. Matas, N. Sebe, and M. Welling, eds. Amsterdam, The Netherlands: Springer, Oct. 2016, pp. 87–102.

[17] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Univ. Massachusetts, Amherst, MA, USA, Tech. Rep. 07-49, Oct. 2007.

[18] *Information Technology—Biometric Performance Testing and Reporting—Part 1: Principles and Framework*, Standard ISO/IEC 19795-1:2006, International Organization for Standardization, Geneva, Switzerland, 2016.

[19] A. Jourabloo, X. Yin, and X. Liu, "Attribute preserved face de-identification," in *Proc. Int. Conf. Biometrics (ICB)*, May 2015, pp. 278–285.

[20] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1867–1874.

[21] E. J. Kindt, *Biometric Data, Data Protection Right to Privacy*. Dordrecht, The Netherlands: Springer Netherlands, 2013.

[22] H. W. Kuhn, "The hungarian method for the assignment problem," *Nav. Res. Logistics*, vol. 52, no. 1, pp. 7–21, Feb. 2005.

[23] K. N. Manisha, "Cancelable biometrics: A comprehensive survey," *Artif. Intell. Rev.*, vol. 53, pp. 3403–3446, 2020, doi: 10.1007/s10462-019-09767-8.

[24] B. Meden, P. Peer, and V. Struc, "Selective face deidentification with end-to-end perceptual loss learning," in *Proc. IEEE Int. Work Conf. Bioinspired Intell. (IWOBI)*, Jul. 2018, pp. 1–7.

[25] V. Mirjalili, S. Raschka, A. Namboodiri, and A. Ross, "Semi-adversarial networks: Convolutional autoencoders for imparting privacy to face images," in *Proc. Int. Conf. Biometrics (ICB)*, Feb. 2018, pp. 82–89.

[26] V. Mirjalili, S. Raschka, and A. Ross, "Gender privacy: An ensemble of semi adversarial networks for confounding arbitrary gender classifiers," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–10.

[27] V. Mirjalili, S. Raschka, and A. Ross, "FlowSAN: Privacy-enhancing semi-adversarial networks to confound arbitrary face-based gender classifiers," *IEEE Access*, vol. 7, pp. 99735–99745, 2019.

[28] V. Mirjalili and A. Ross, "Soft biometric privacy: Retaining biometric utility of face images while perturbing gender," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 564–573.

[29] A. Morales, J. Fiérrez, and R. Vera-Rodríguez, "Sensitivenets: Learning agnostic representations with application to face recognition," 2019, *arXiv:1902.00334*. [Online]. Available: https://arxiv.org/abs/1902.00334

[30] E. M. Newton, L. Sweeney, and B. Malin, "Preserving privacy by de-identifying face images," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 2, pp. 232–243, Feb. 2005.

[31] H. Patil, A. Kothari, and K. Bhurchandi, "Expression invariant face recognition using semidecimated DWT, patch-LDSMT, feature and score level fusion," *Appl. Intell.*, vol. 44, no. 4, pp. 913–930, Jun. 2016, doi: 10.1007/s10489-015-0735-1.

[32] V. M. Patel, N. K. Ratha, and R. Chellappa, "Cancelable biometrics: A review," *IEEE Signal Process. Mag.*, vol. 32, no. 5, pp. 54–65, Sep. 2015.

[33] H. Patil, A. Kothari, and K. Bhurchandi, "Expression invariant face recognition using semidecimated dwt, patch-ldsmt, feature and score level fusion," *Appl. Intell.*, vol. 44, no. 4, pp. 913–930, 2016.

[34] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

[35] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.

[36] P. Punithavathi and G. Subbiah, "Can cancellable biometrics preserve privacy?" *Biometric Technol. Today*, vol. 2017, no. 7, pp. 8–11, Jul. 2017.

[37] R. Raghavendra and C. Busch, "Presentation attack detection methods for face recognition systems: A comprehensive survey," *ACM Comput. Surv.*, vol. 50, no. 1, pp. 8:1–8:37, Mar. 2017.

[38] L. Ramshaw and R. E. Tarjan, "A weight-scaling algorithm for min-cost imperfect matchings in bipartite graphs," in *Proc. IEEE 53rd Annu. Symp. Found. Comput. Sci.*, Oct. 2012, pp. 581–590.

[39] A. Rozsa, M. Gunther, E. M. Rudd, and T. E. Boult, "Are facial attributes adversarially robust?" in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 3121–3127.

[40] A. Rozsa, M. Günther, E. M. Rudd, and T. E. Boult, "Facial attributes: Accuracy and adversarial robustness," 2018, *arXiv:1801.02480*. [Online]. Available: https://arxiv.org/abs/1801.02480

[41] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823.

[42] M. Stokkenes, R. Ramachandra, and C. Busch, "Biometric authentication protocols on smartphones: An overview," in *Proc. 9th Int. Conf. Secur. Inf. Netw. SIN*, 2016, pp. 136–140.

[43] J. Suo, L. Lin, S. Shan, X. Chen, and W. Gao, "High-resolution face fusion for gender conversion," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 2, pp. 226–237, Mar. 2011.

[44] P. Terhorst, N. Damer, A. Braun, and A. Kuijper, "Deep and multi-algorithmic gender classification of single fingerprint minutiae," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Jul. 2018, pp. 2113–2120.

[45] P. Terhorst, N. Damer, F. Kirchbuchner, and A. Kuijper, "Suppressing gender and age in face templates using incremental variable elimination," in *Proc. Int. Conf. Biometrics (ICB)*, Jun. 2019, pp. 4–7.

[46] P. Terhörst, N. Damer, F. Kirchbuchner, and A. Kuijper, "Unsupervised privacy-enhancement of face representations using similarity-sensitive noise transformations," *Appl. Intell.*, vol. 49, pp. 3043–3060, Feb. 2019.

[47] P. Terhörst, M. Huber, N. Damer, F. Kirchbuchner, and A. Kuijper, "Unsupervised enhancement of soft-biometric privacy with negative face recognition," 2020, *arXiv:2002.09181*. [Online]. Available: https://arxiv.org/abs/2002.09181

[48] P. Terhörst, M. Huber, J. N. Kolf, N. Damer, F. Kirchbuchner, and A. Kuijper, "Multi-algorithmic fusion for reliable age and gender estimation from face images," in *Proc. 22st Int. Conf. Inf. Fusion, FUSION*, Ottawa, ON, Canada, Jul. 2019, pp. 1–8.

[49] P. Terhörst, M. Huber, J. N. Kolf, I. Zelch, N. Damer, F. Kirchbuchner, and A. Kuijper, "Reliable age and gender estimation from face images: Stating the confidence of model predictions," in *Proc. 10th IEEE Int. Conf. Biometrics Theory, Appl. Syst., BTAS*, Tampa, FL, USA, Sep. 2019, pp. 1–8.

[50] P. Terhörst, J. N. Kolf, N. Damer, F. Kirchbuchner, and A. Kuijper, "Post-comparison mitigation of demographic bias in face recognition using fair score normalization," 2020, *arXiv:2002.03592*. [Online]. Available: https://arxiv.org/abs/2002.03592

[51] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

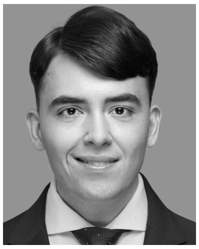
[52] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy array: A structure for efficient numerical computation," *Comput. Sci. Eng.*, vol. 13, no. 2, pp. 22–30, Mar. 2011.

[53] P. Virtanen *et al.*, "SciPy 1.0: Fundamental algorithms for scientific computing in python," *Nature Methods*, vol. 17, no. 3, pp. 261–272, 2020.

[54] C.-P. Wang, W. Wei, J.-S. Zhang, and H.-B. Song, "Robust face recognition via discriminative and common hybrid dictionary learning," *Appl. Intell.*, vol. 48, no. 1, pp. 156–165, 2018.

[55] Y. Wang and M. Kosinski, "Deep neural networks are more accurate than humans at detecting sexual orientation from facial images," *J. Personality Social Psychol.*, vol. 114, no. 2, p. 246, 2018.

PHILIPP TERHÖRST received the degree in physics from the Technical University of Darmstadt, in 2017. Since 2017, he has been working with the Smart Living and Biometric Technologies Department, Fraunhofer Institute for Computer Graphics Research (IGD) as a Research Assistant. In the context of his doctorate, his field of work includes research into biometric solutions based on machine learning algorithms, specialising in dealing with reliability, privacy, and bias. Furthermore, he is involved in the Software Campus Program, a Management Program of the Federal Ministry of Education and Research (BMBF). He has served as a Reviewer for various conferences and journals (e.g., the IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Access, BTAS, ICB, *Fusion*).



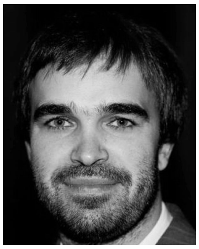
KEVIN RIEHL received the B.Sc. degree in business engineering from the Department of Electrical Engineering, Technical University of Darmstadt, in 2017. He is currently pursuing the master's degree. He worked as a Research Assistant for the Chair of International Economics. He performed two internships in Frankfurt, Germany and Shanghai, Beijing, China. He is currently working with the Fraunhofer IGD, Darmstadt.

NASER DAMER received the master of science degree in electrical engineering from the Technical University of Kaiserslautern, in 2010, and the Ph.D. degree in computer science from the Technical University of Darmstadt, in 2018. He has been a Researcher with Fraunhofer IGD, since 2011, where he is performing applied research, scientific consulting, and system evaluation. He is currently a Senior Researcher with the Competence Center, Smart Living and Biometric Technologies, Fraunhofer IGD. His main research interests include the fields of biometrics, machine learning, and information fusion. He published more than 60 scientific articles in these fields. Dr. Damer is a Principal Investigator with the National Research Center for Applied Cybersecurity ATHENE, Darmstadt, Germany. He serves as a reviewer for a number of journals and conferences and as an Associate Editor for the *Visual Computer* journal. He represents the German Institute for Standardization (DIN) in ISO/IEC SC37 Biometrics Standardization Committee.

PETER ROT (Student Member, IEEE) received the bachelor's and master's degrees from the Faculty of Computer and Information Science, University of Ljubljana, Slovenia, in 2015 and 2018, respectively, where he is currently pursuing the Ph.D. degree in computer science. He is currently a Researcher with the Laboratory for Machine Intelligence, Faculty of Electrical Engineering, and the Laboratory for Computer Vision, Faculty of Computer and Information Science, University of Ljubljana. During his undergraduate studies, he completed several internships, including internships at the Jozef Stefan Institute, Ljubljana, Slovenia, and Philips, Belfast, U.K. His research interests include privacy aspects of face biometrics, sclera recognition, and deep learning. He is a Reviewer for top-tier conferences and journals, e.g., IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, FG, and CVWW.

BLAZ BORTOLATO is currently with the University of Ljubljana.

FLORIAN KIRCHBUCHNER received the M.Sc. degree in computer science from the Technical University of Darmstadt, in 2014. He is currently pursuing the Ph.D. degree with the Technical University of Darmstadt on the topic Electric Field Sensing for Smart Support Systems: Applications and Implications. He was trained as an Information and Telecommunication Systems Technician and served as the IT Expert for the German Army, from 2001 to 2009. He has been working with Fraunhofer IGD, since 2014, most recently as the Head of the Department for Smart Living and Biometric Technologies. He is currently the Principal Investigator with the National Research Center for Applied Cybersecurity ATHENE. He participated at Software Campus, a Management Program of the Federal Ministry of Education and Research (BMBF).



VITOMIR STRUC received the Ph.D. degree from the Faculty of Electrical Engineering, Ljubljana, in 2010. He is currently an Associate Professor with the University of Ljubljana, Slovenia. His research interests include problems related to biometrics, computer vision, image processing, pattern recognition, and machine learning. He has coauthored more than 100 research articles for leading international peer-reviewed journals and conferences in these and related areas. He served in different capacities on the organizing committees of several top-tier vision conferences, including the IEEE Face and Gesture, ICB, WACV and others. He is currently a Program Co-Chair for the 2020 International Joint Conference on Biometrics (IJCB). He is an Associate Editor for the IEEE Transactions on Information Forensics and Security, *Signal Processing*, and *IET Biometrics*. He served as an Area Chair for WACV 2018, 2019, 2020, ICPR 2018, Eusipco 2019 and FG 2020. Dr. Struc is a member of IAPR, EURASIP, Slovenia's National Contact Point for the EAB and the Current President of the Slovenian Pattern Recognition Society (Slovenian branch of IAPR).

ARJAN KUIJPER received the M.Sc. degree in applied mathematics from Twente University, The Netherlands, the Ph.D. degree from Utrecht University, The Netherlands, and the Habitation degree from TU Graz, Austria. He was an Assistant Research Professor with the IT University of Copenhagen, Denmark, and a Senior Researcher with RICAM, Linz, Austria. He has authored over 300 peer-reviewed publications. His research interests include all aspects of mathematics-based methods for computer vision, graphics, imaging, pattern recognition, interaction, and visualization. Dr. Kuijper is a member of the Management of Fraunhofer IGD, where he is responsible for scientific dissemination. He holds the Chair in mathematical and applied visual computing with TU Darmstadt. He is an Associate Editor of CVIU, PR, and TVCJ, the Secretary of the International Association for Pattern Recognition (IAPR), and serves both as a reviewer for many journals and conferences, and as a program committee member and the organizer of conferences.