

Article

Face Beneath the Ink: Synthetic Data and Tattoo Removal with Application to Face Recognition

Mathias Ibsen , Christian Rathgeb, Pawel Drozdowski  and Christoph Busch 

da/sec —Biometrics and Security Research Group, Hochschule Darmstadt, Schöfferstraße 3,
64295 Darmstadt, Germany

* Correspondence: mathias.ibsen@h-da.de

Abstract: Systems that analyse faces have seen significant improvements in recent years and are today used in numerous application scenarios. However, these systems have been found to be negatively affected by facial alterations such as tattoos. To better understand and mitigate the effect of facial tattoos in facial analysis systems, large datasets of images of individuals with and without tattoos are needed. To this end, we propose a generator for automatically adding realistic tattoos to facial images. Moreover, we demonstrate the feasibility of the generation by using a deep learning-based model for removing tattoos from face images. The experimental results show that it is possible to remove facial tattoos from real images without degrading the quality of the image. Additionally, we show that it is possible to improve face recognition accuracy by using the proposed deep learning-based tattoo removal before extracting and comparing facial features.

Keywords: facial tattoos; synthetic data generation; tattoo removal; face recognition



Citation: Ibsen, M.; Rathgeb, C.; Drozdowski, P.; Busch, C. Face Beneath the Ink: Synthetic Data and Tattoo Removal with Application to Face Recognition. *Appl. Sci.* **2022**, *12*, 12969. <https://doi.org/10.3390/app122412969>

Academic Editors: Jana Dittmann and Andrey Makrushin

Received: 11 November 2022

Accepted: 14 December 2022

Published: 16 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Facial analysis systems are deployed in various applications ranging from medical analysis to border control. Such facial analysis systems are known to be negatively affected by facial occlusions [1,2]. A specific kind of facial alteration that partially occludes a face is a facial tattoo. Facial tattoos have become more appealing recently and have been described as a mainstream trend in several major newspapers [3,4]. Ensuring inclusiveness and accessibility for all individuals, independent of physical appearance, is imperative in developing fair facial analysis systems. In this regard, facial tattoos are especially challenging, as they cause permanent alterations where ink is induced into the dermis layer of the skin. For instance, Ibsen et al. investigated in [5] the impact of facial tattoos and paintings on state-of-the-art face recognition systems. The authors showed that tattoos might impair the recognition accuracy, and hence the security of such a facial analysis system. In the paper, the authors considered the scenario where either the reference or the probe image has been altered by tattoos. One way to address this issue is to re-enroll the subject into the reference database. However, this is not always possible, e.g., in some forensic applications and in other applications it requires that new documents are issued, e.g., for automated border control gates. Additionally, it still requires that the identity of the subject is verified before the facial image with tattoos is enrolled.

In coherence with the findings in [5], it is of interest to make facial analysis systems more robust towards facial tattoos. Some research has explored methods for adopting face recognition systems to be more robust towards occlusions, e.g., [6]. Another way to do this is face completion, where missing or occluded parts of a face are reconstructed; such approaches have, for instance, shown to improve face recognition performance for some occlusions [7]. An additional benefit of using face completion over approaches such as occlusion-aware face recognition is the potential to use the reconstructed facial image for other purposes, e.g., visualising how a face might look without the occlusion or preventing

that tattoos are used for recognition purposes, which is something that raises ethical issues as discussed in [8].

However, one major problem with face completion for tattoo removal is the lack of sufficient and high-quality training data, as no extensive database of facial tattoos is currently available.

The main focus of this work is, therefore, two-fold. First, we propose a method for synthetically adding tattoos to facial images, which we use to create a large database of facial images with tattoos. The proposed method uses face detection and landmark localisation to divide the face into regions, whereafter suitable placements of tattoos are found. Subsequently, we approximate depth and construct depth and cut-out maps used to blend tattoos onto a face realistically. It has recently been shown that synthetic data can be beneficial for face analysis tasks and be a good alternative to real data [9,10]. Secondly, we show the usefulness of our synthetic data by training a deep learning-based model for tattoo removal (as illustrated in Figure 1) and evaluate the impact of removing facial tattoos on a state-of-the-art face recognition system using a database comprising real facial images with tattoos.

The approach for synthetically adding tattoos to a facial image in a fully automated way is, to the authors' best knowledge, the first of its kind. The proposed generator can be used to create large databases which can be used in related fields, e.g., tattoo detection or studying the effects of tattoos on human perception. Additionally, we are the first to measure the effect of removing facial tattoos on face recognition systems.

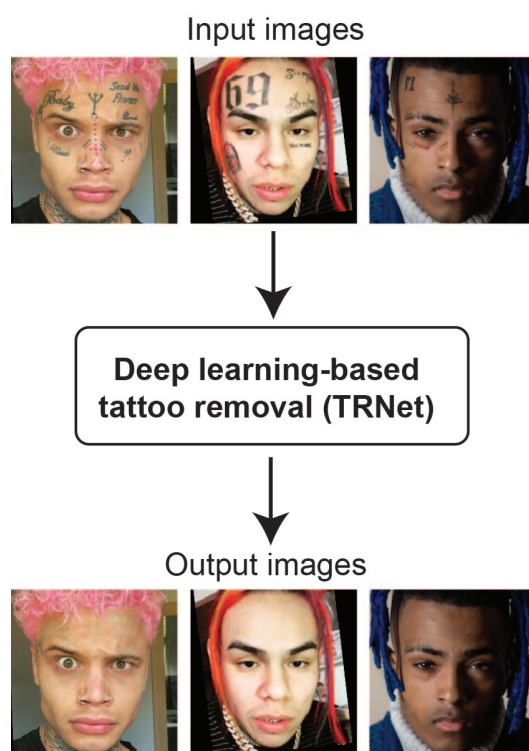


Figure 1. Examples of using deep learning-based tattoo removal.

In summary, this work makes the following contributions:

- A novel algorithm for synthetically adding facial tattoos to face images.
- An algorithm for removing tattoos from facial images trained on only facial images with synthetically added tattoos. We refer to this algorithm as TRNet.
- An experimental analysis of the quality of the tattoo removal.
- Showcasing the application of tattoo removal in a face recognition system by conducting an experimental analysis on the effect of removing facial tattoos on a face recognition system.

The outline of the remaining article is as follows: Section 2 describes prominent related works, Section 3 describes an automated approach for synthetically blending tattoos to facial images, which is used in Section 4 to generate a database of facial images with tattoos. Sections 5 and 6 show the feasibility of the synthetic generation by training a deep learning-based model for tattoo removal and evaluating if it can improve biometric recognition performance, respectively. Finally, Section 7 provides a summary of this work.

2. Related Work

The following subsections summarise some related works w.r.t. synthetic data generation for facial analysis (Section 2.1), facial alterations (Section 2.2), and facial completion (Section 2.3). Readers are referred to the following comprehensive surveys for a more in-depth comparison and overview of different approaches [1,10,11].

2.1. Synthetic Data Generation for Face Analysis

Synthetically generated data have seen many application scenarios in face analysis, most notably for addressing the lack of training data. Synthetic data have especially become relevant with the recent advances in deep learning-based algorithms, which usually require a large amount of training data. Privacy regulations, e.g., the European General Data Protection Regulation [12], make sharing and distributing large-scale face databases impracticable as face images are classified as a special category of personal data when used for biometric identification. As an alternative, researchers have explored the use of synthetic data. The generation of realistic-looking synthetic face data have especially become feasible with the recent advances in Generative Adversarial Networks (GANs), first proposed by Goodfellow et al. in [13]. Prominent work in this field includes StyleGAN, which was first introduced in [14] by Karras et al. and showed, at the time, state-of-the-art performance for synthesising facial images. Since the original work, two improved versions of StyleGAN have been proposed [15,16]. Much current research in this area focuses on GAN-inversion, where existing face images are encoded into the latent space of a generator. After that, the resulting latent code can be shifted in the latent space, whereby the inverted image of the shifted vector results in an alteration of the original image. The technique can, for instance, be used for face age progression [17]. In addition to the face, some research has also been conducted for other biometric modalities, e.g., fingerprint [18–20] and iris [21,22].

Little work has been conducted regarding synthetic data generation of facial images with tattoos. However, in [23], the authors proposed a method for transforming digital portrait images into realistic-looking tattoos. In [24], the author also shows examples of tattoo images added to facial and body images using an existing GAN for drawing art portraits; however, details about this approach are not scientifically documented.

2.2. Facial Alterations

Facial alterations can occur in either the physical or digital domain and cause permanent or temporary changes in a face. Several studies have explored the impact of physical and digital alterations on face recognition systems. In the physical domain, predominantly, the effects of makeup and plastic surgery on face recognition have been studied [11]. In [25], the authors collected a database of 900 individuals to analyse the effect of plastic surgery and found that the tested algorithms were unable to account for the appearance changes caused by plastic surgery effectively. More recently, Rathgeb et al. showed in [26], using a database of mostly ICAO-quality face images [27] captured before and after various types of facial plastic surgeries, that different tested state-of-the-art face recognition systems maintained almost perfect verification performance at an operationally relevant threshold corresponding to a False Match Rate (FMR) of 0.1%. Numerous works have addressed the impact of temporary alterations on face recognition systems. In [28], Dantcheva et al. found that makeup can hinder reliable face recognition; similar conclusions were drawn by Wang et al. in [29] where they investigated the impact of human faces under disguise and makeup. The previous work shows that makeup might be successfully used for identity

concealment; in [30], the authors additionally showed that makeup could also be used for presentation attacks with the goal of impersonating another identity. In [31], the authors found that especially high-quality makeup-based presentation attacks can hamper the security of face recognition systems. In [32], the authors found that disguised faces severely affect recognition performance, especially for occlusions near the periocular region. The database used by the authors includes different types of disguises, including facial paintings. Coherent with these findings, Ibsen et al. showed in [5] that facial tattoos and paintings can severely affect different modules of a face recognition system, including face detection as well as feature extraction and comparison.

Ferrara et al. were among the first to show that digital alterations can impair the security of face recognition systems. Especially notable is their work in [33], where they showed the possibility of attacking face recognition systems using morphed images. Specifically, they showed that if a high-quality morphed image is infiltrated into a face recognition system (e.g., stored in a passport), it is likely that the biometric system positively authenticates individuals contributing to the morph. Since then, there have been numerous works on face recognition systems under morphing attacks. For a comprehensive survey, the reader is referred to [34]. Facial retouching is another area which has seen some attention in the research community. While some early works showed that face recognition can be significantly affected by retouching, Rathgeb et al. showed more recently that face recognition systems might be robust to slight alterations caused by retouching [35]. Similar improvements have been shown for geometrical distortions, e.g., stretching [36]. A more recent threat that has arrived with the prevalence of deep-learning techniques is so-called DeepFakes [37], which can be used to spread misinformation and, as such, lead to a loss of trust in digital content. Many researchers are working on the detection or generation of deep learning-based alterations. Several arduous challenges and benchmarks have already been established, for instance, the recent Deepfake Detection Challenge [38] where the top model only achieved an accuracy of approximately 65% on previously unseen data. Generation and detection of deep learning-based alterations are continuously evolving and remain a cat-and-mouse game; interested readers are referred to [39] for a comprehensive survey.

2.3. Facial Completion

Most methods for face completion (also called face inpainting) build upon deep learning-based algorithms, which are trained on paired images where each pair contains a non-occluded face and a corresponding occluded face. In [40], the authors proposed an approach for general image completion and showed its applicability for facial completion. In this work, the authors leveraged a fully convolutional neural network trained with global and local context discriminators. Similar work was done in [41] where the authors occluded faces by adding random squares of noise pixels. Subsequently, they trained an autoencoder to reconstruct the occluded part of the face using global and local adversarial losses as well as a semantic parsing loss. Motivated by the prevalence of VR/AR displays which can hinder face-to-face communication, Zhao et al. [42] proposed a new generative architecture with an identity-preserving loss. In [43], Song et al. used landmark detection to estimate the geometry of a face and used it, together with the occluded face image, as input to an encoder-decoder architecture for reconstructing the occluded parts of the face. The proposed approach allows for generating diverse results by altering the estimated facial geometry. More recently, Din et al. [44] employed a GAN-based architecture for the unmasking of masked facial images. The proposed architecture consists of two stages where the first stage detects the masked area of the face and creates a binary segmentation map. The segmentation map is then used in the second stage for facial completion using a GAN-based architecture with two discriminators: one focuses on the global structure and the other on the occluded parts of the face. In [7], it was found that facial completion can improve face recognition performance.

3. Facial Tattoo Generator

To address the lack of existing databases of image pairs of individuals before and after they got facial tattoos, we propose an automated approach for synthetically adding facial tattoos to images. An overview of the proposed generation is depicted in Figure 2. The process of synthetically adding tattoos to a facial image can be split into two main steps, which are described in the following subsections: (1) finding the placement of tattoos in a face and (2) blending the tattoos onto the face.

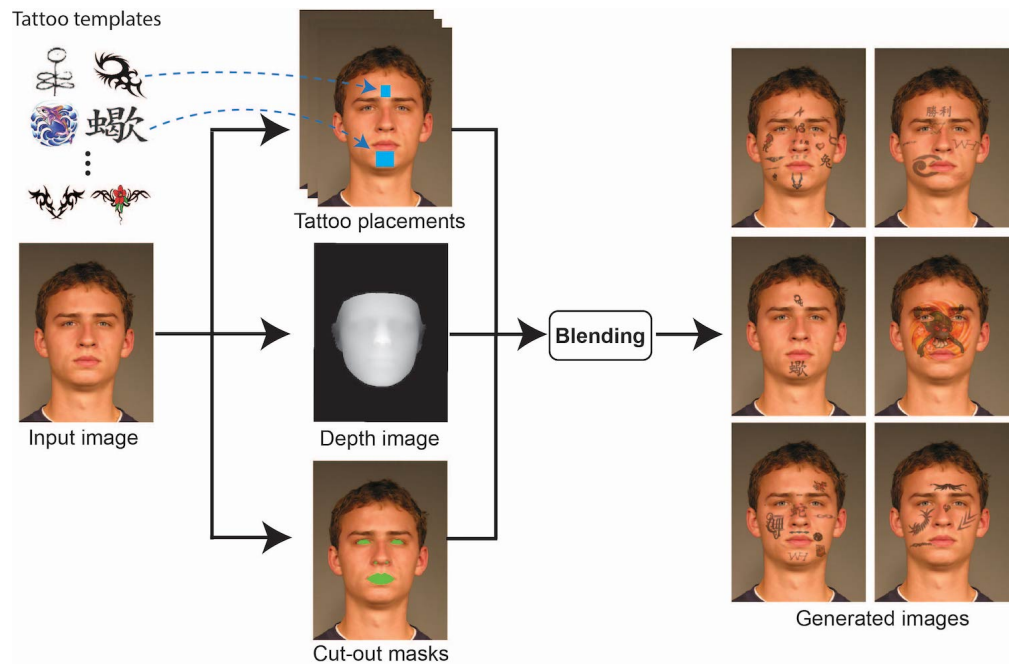


Figure 2. Synthetic facial tattoo generation workflow.

3.1. Placement of Tattoos

To find suitable placements of tattoos on a face, we start by localising the facial region and detecting landmarks of the face. To this end, we use *dlib* [45], which returns a list of 68 landmarks as shown in Figure 3.

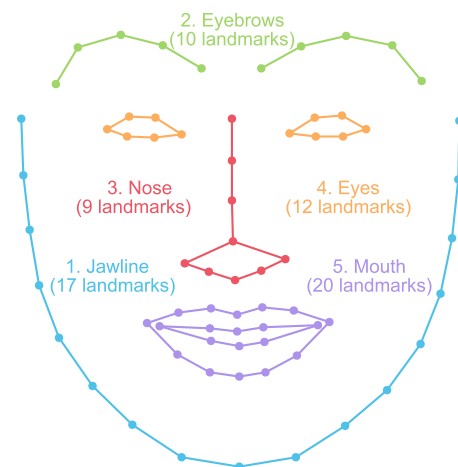


Figure 3. Facial landmarks detected by dlib.

The landmarks are used to divide the face into small regions of triangles by performing a fixed Delaunay triangulation. The regions are then extended to the forehead by using the length of the nose as an estimate. Each region now constitutes a possible placement

of a tattoo; however, such a division is inadequate for the placement of larger tattoos. Therefore, the face is divided into six larger regions. The division of the face into large and small regions gives high controllability in the data generation. As indicated, some regions are excluded, i.e., regions around the nostrils, mouth and nose. These regions are estimated based on the detected landmarks. The division of a face into regions is illustrated in Figure 4. The regions make it possible to avoid placing tattoos in heavily bearded areas or on top of glasses if such information is available about the facial images during the generation phase. In our work, we do not use beard or glasses detectors; however, for some of the images, information about beard or glasses is available, which we use to avoid placing tattoos in the affected regions.

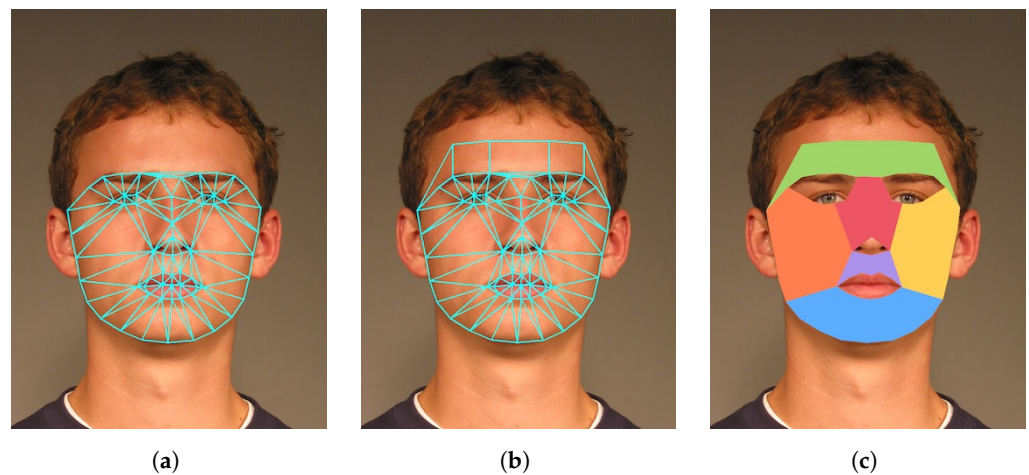


Figure 4. (a) Division of a facial image into regions from landmarks, (b) extended to the forehead, and (c) division into six pre-defined regions.

A tattoo can now be placed in one of the six pre-defined regions, or the regions can be further combined to place the tattoos in larger areas of the face. A combined region is simply a new region consisting of several smaller regions. The exact placement of a tattoo within a region depends on a pre-selected generation strategy. The generation strategy determines (1) possible regions where a tattoo can be placed, (2) the selection of tattoos, and (3) the size and placement of a tattoo within a region. An example is illustrated in Figure 5 where one of the cheeks is selected as a possible region, whereafter the largest unoccupied subset within that region is found. Thereafter, the tattoo is placed by estimating its largest possible placement within the selected subset without altering the original aspect ratio of the tattoo. In this work, we use a database comprising more than 600 distinct tattoo templates, mainly consisting of real tattoo designs collected from acquired tattoo books. Selecting which tattoos to place depends on the generation strategies, which are further described in Section 3.3.

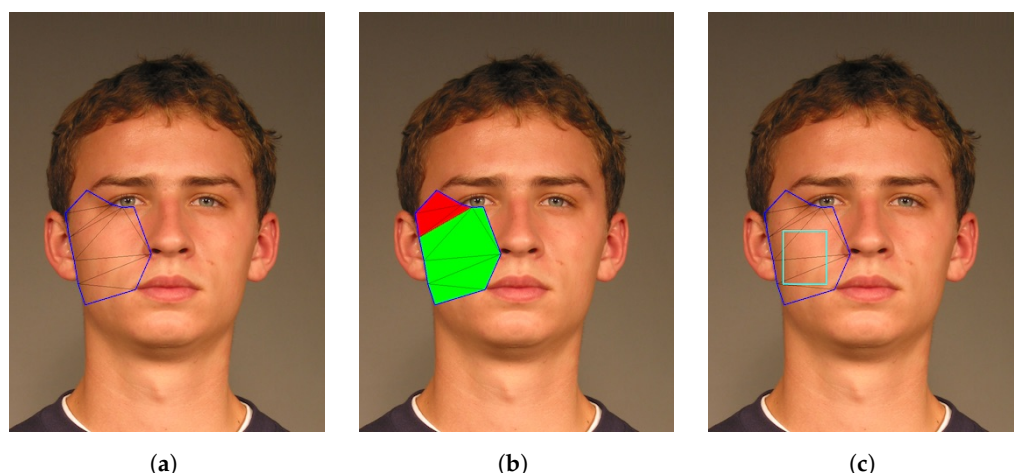


Figure 5. Illustration shows an example of how a tattoo placement in a region can be found. The red area in (b) illustrates that there might be some areas within a selected region where a tattoo cannot be placed, e.g., if the area is reserved for another tattoo. (a) Selected region. (b) Find a subset of the region not occupied (the green area). (c) Find a placement for the tattoo.

3.2. Blending

To blend the tattoos onto faces, various image manipulations are performed.

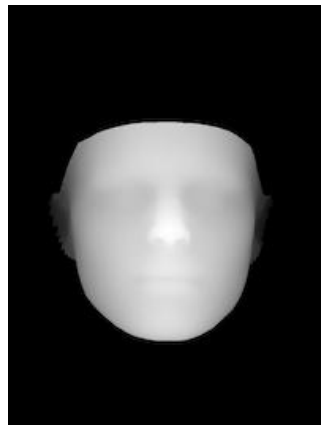
Given a facial image and placement of tattoos (see Section 3.1); each tattoo is placed and overlaid on the facial image by multiplying the tattoo layer with the facial image. Afterwards, the tattoo is displaced to match the contours of the face using displacement mapping. Areas of the tattoo which have been displaced outside the face or inside the mouth, nostrils and eyes are cut out. This is achieved by using cut-out maps (see Figure 2), which are calculated from the landmarks detected by dlib in the placement phase. Lastly, the tattoo is made more realistic by colour adjustment, Gaussian blurring, and lowering the opacity of the tattoo.

As previously stated, displacement mapping is used for mapping tattoos to the contours of a face. It is a technique which utilises depth information of texture maps to alter the positions of pixels according to the depth information in the provided map. Contrary to other approaches, such as bump mapping, it alters the source image by displacing pixels. In displacement mapping, a map M containing values in the range 0–255 is used to displace pixels in a source image I . As seen in the equation, a specific pixel, $I(x, y)$, is displaced in one direction if the corresponding pixel in the displacement map, $M(x, y)$, is less than the theoretical average pixel value of the map (127.5); otherwise, it is displaced in the opposite direction. For the displacement technique used in this work, a pixel in the source image is displaced both vertically and horizontally.

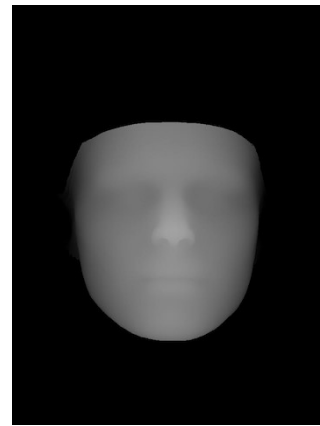
More specifically, let c be a coefficient, let $(x, y) \in I$, and let $(x, y) \in M$. The distance for displacing a pixel, $I(x, y)$, in the vertical and horizontal direction is then:

$$D(x, y) = c \cdot \frac{M(x, y) - 127.5}{127.5} \quad (1)$$

To generate depth maps, PRNet is used [46]. PRNet is capable of performing 3D face reconstruction from 2D face images, and as such, it can also approximate depth maps from 2D facial images. PRNet proposes to use so-called UV position maps to represent 3D facial structure. The position map stores 3D positions as a 2D image in UV space. An encoder-decoder network is trained to regress the UV position map from a 2D facial image. An example of a depth map generated using PRNet is shown in Figure 6a.



(a) Depth image generated by PRNet



(b) Transformed depth image

Figure 6. Example of (a) a depth map generated from a facial image using PRNet and (b) after it has been transformed. Note, that the original depth map (a) corresponds to the input image in Figure 2.

As seen in Figure 6a, the pixel values in the face region are rather bright, and there is little contrast. The small contrast between the pixel values and the high offset from the theoretical average pixel value implies that the depth map will not work very well, as tattoos will be displaced too much in certain regions and too little in others. Therefore, to make the displacement more realistic, the depth map generated by PRNet is transformed by increasing the contrast and lowering the brightness of the map. Figure 6b shows an example of a transformed depth map, and as can be seen, the pixel values are much closer to the theoretical average value than the unaltered map, while the contrast around the nose, eyes and mouth are still high. Figure 7 shows an example where two facial tattoos are displaced to match the contours of a face and figure 8 shows examples where tattoos placed in undesired areas of a face has been cut out.



(a) not displaced



(b) displaced

Figure 7. Facial images with tattoos (a) before and (b) after applying the displacement technique. For (b), the tattoo is blended around the anticipated 3D shape of the nose. Best viewed in electronic format (zoomed in).



Figure 8. Examples of facial images where parts of one or more tattoos have been cut out.

Black ink tends to change in colour slightly over time due to the pigment used in black ink. Therefore, for colour adjustment, all pixels of a tattoo which are similar to pure black are selected and changed to simulate different colours of grey, green, and blue, which causes black tattoos to appear differently for different facial images. The colour adjustments of black pixels are determined per tattoo, and as such slight variations can occur between different tattoos in the same facial image. Examples are given in Figure 9.



Figure 9. Examples of black tattoos blended to facial images.

3.3. Generation Strategies

By varying how tattoos are selected and placed (Section 3.1), many different types of facial images with tattoos can be generated. For the database used in this work, we employed two different strategies. In the first strategy, the desired coverage percent of tattoos on a face is randomly chosen from a specified range. Subsequently, tattoos are arbitrarily selected and placed on facial regions until the resulting coverage approximates the desired coverage. The coverage of a tattoo on a face is calculated based on the total amount of pixels in all the facial regions (see Figure 4c) and the number of non-transparent pixels in the placed tattoos. In the second strategy, a specific region is always selected. Using the first strategy, it is possible to create databases where tattoos are placed arbitrarily until a selected coverage percent has been reached (see Figure 10a–c). Using the latter approach allows for more controlled placement of tattoos, e.g., placing tattoos in the entire face region (Figure 10d) or in a specific region (Figure 10e,f).

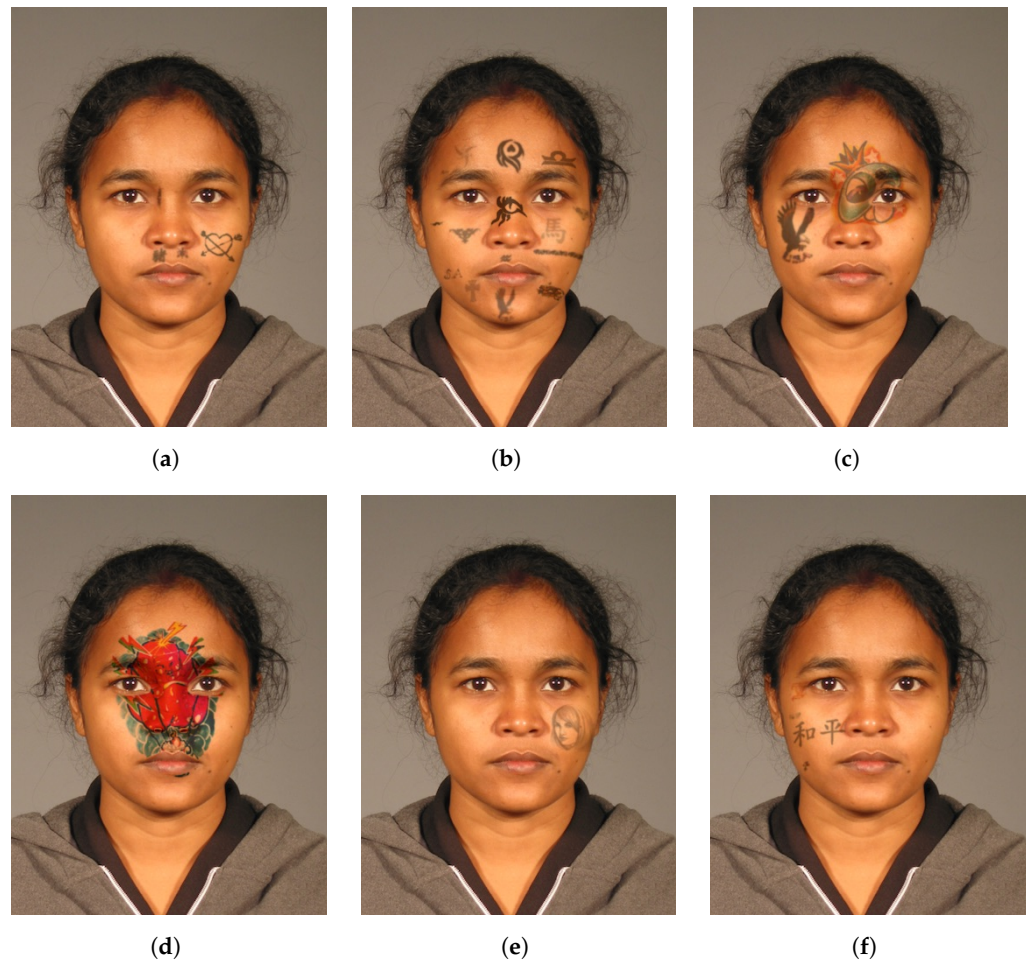


Figure 10. Examples for different types of tattooed faces that can be generated: (a) 5%, (b) 15%, (c) 25% coverage, (d) entire face, (e) single tattoo, and (f) specific region.

4. Synthetic Tattoo Database

This section describes the generation of a large database of facial images with tattoos. The database is used in Section 5 to train deep learning-based models for removing tattoos from facial images. To generate the synthetic tattoo database, subsets of original images from the FERET [47], FRGCv2 [48], and CelebA [49] datasets were used. An overview of the generated database is given in Table 1. For the FERET and FRGCv2 datasets, different generation strategies were used, including facial images where tattoos have been placed randomly, with specific coverage ranging from 5% to 25% as well as placement of single tattoos. We generated two versions for the single tattoos: one version where the tattoo is placed in the entire facial region and another where portrait tattoos are blended to a random region in the face. For the CelebA database, which is more uncontrolled, facial tattoos were placed randomly. Data augmentation was performed to simulate varying image qualities by randomly applying differing degrees of JPEG compression or Gaussian blur to all the images. Tattoo images and corresponding original (bona fide) images were paired such that similar augmentation was applied to corresponding images.

Table 1. Overview of the generated database (before augmentation).

Database	Subjects	Images	
		Bona fide	Tattooed
FERET	529	621	6743
FRGCv2	533	1436	16,209
CelebA	6872	6872	6872

Examples of images in the generated database are depicted in Figure 11.



Figure 11. Examples of generated facial images with tattoos.

5. Tattoo Removal

To show the feasibility of the synthetic data and their potential use in real-world applications, we investigate a concrete case where the synthetic data are used for removing real tattoos from facial images. To this end, two models are trained for the task of tattoo removal using the synthetic data described in Section 4.

Section 5.1 briefly describes the different models used for removing tattoos. Section 5.2 describes different metrics for evaluating the quality of the tattoo removal which is then evaluated in Section 5.3.

5.1. Models

Two different deep learning-based methods were trained for removing tattoos from facial images:

- pix2pix is a supervised conditional GAN for image-to-image translation [50]. For the generator, a U-Net architecture is used, whereas the discriminator is based on a PatchGAN classifier which divides the image into $N \times N$ patches and discriminates between bona fide (i.e., real images) and fake images.
- Tattoo Removal Net (TRNet) is a U-net architecture [24] [51] which utilizes spectral normalization and self-attention. The network was training using only the synthetic data described in Section 4. An illustration of the used U-Net architecture is shown in

Figure 12. The encoder of the network is based on ResNet34, and the decoder consists of four main blocks and utilizes PixelShuffling [52]. The loss function is a combination of feature loss (perceptual loss) from [53], gram matrix style loss [54], and pixel (L1) loss. For the gram matrix loss and the feature loss, blocks from a pre-trained VGG-16 model are used [55] [51].

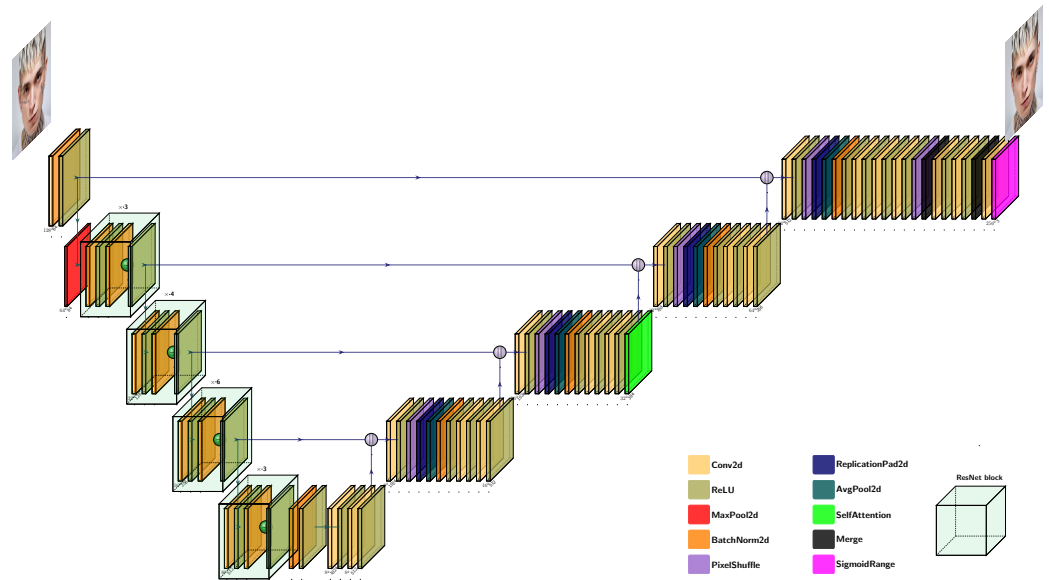


Figure 12. Architecture of the tattoo removal network (TRNet).

5.2. Quality Metrics

To evaluate the quality of the different tattoo removal models, we use three different metrics commonly used in the literature:

- Peak signal-to-noise ratio (PSNR) is a measurement of error between an input and an output image and is calculated as follows:

$$PSNR(X, Y) = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE(X, Y)}} \right) \tag{2}$$

where MAX_I is the theoretical maximum pixel value (i.e., 255 for 8 bit channels) and $MSE(X, Y)$ is the mean squared error between the ground truth image X and the inpainted image Y . The PSNR is measured in decibel, and a higher value indicates better quality of the reconstructed image.

- Mean Structural Similarity Index (MSSIM) as given in [56], is defined as follows:

$$MSSIM(X, Y) = \frac{1}{M} \sum_{i=1}^M SSIM(x_i, y_i) \tag{3}$$

where X and Y are the ground truth image and inpainted image, respectively, M is the number of local windows in an image and x_i and y_i are the image content of the i 'th local window. The SSIM over local window patches (x, y) is defined as:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1) + (\sigma_x^2 + \sigma_y^2 + C_2)} \tag{4}$$

where μ_x and μ_y are the mean values of the local window patches x and y , respectively; $\sigma_x^2 + \sigma_y^2$ are their local variances and σ_{xy} is the local covariance of x and y . C_1 and C_2 are constants set based on the same parameter settings as Wang et al. [56], i.e.,

$C_1 \approx 6.55, C_2 \approx 58.98$. MSSIM returns a value in the range of 0 to 1, where 1 means that X and Y are identical.

- Visual Information Fidelity (VIF) is a full reference image quality assessment measurement proposed by Sheikh and Bovik in [57]. VIF is derived from a statistical model for natural scenes as well as models for image distortion and the human visual system. $VIF(X, Y)$ returns a value in the range of 0 to 1, where 1 indicates that the ground truth and inpainted images are identical. We use the pixel domain version as implemented in [58].

We estimate the different quality metrics both on portrait images, i.e., where the entire face is visible and on the inner part of the face (corresponding to the area covered by the 68 dlib landmark points; see Figure 3) where we focus on only the area from the eyebrows to the chin; these regions are shown in Figure 13.

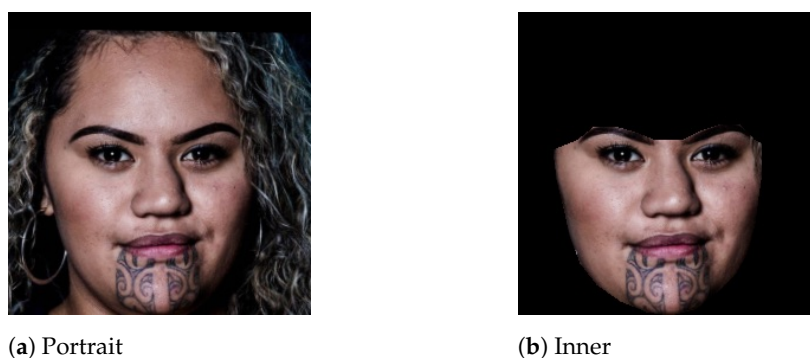


Figure 13. Examples of (a) a full portrait image where the entire face is visible and (b) a crop of the inner face region.

5.3. Removal Quality Results

We use a total of 41 facial images with tattoos from [5] where the tattoos have been manually removed using PhotoShop; we refer to these as our ground truth images. Examples of using the different deep learning-based methods for removing tattoos are given in Figure 14. As seen, the best model (TRNet) is able to remove most tattoos with only a few artefacts, whereas the other models perform less well and, for some images, alter the face or fail to remove all tattoos accurately.

Different quality scores are reported in Table 2, which shows that the TRNet model performs best in most scenarios, especially when only looking at the inner part of the face.

Table 2. Quality measurements of the reconstructed images compared to ground truth images where tattoos have been manually removed. "Tattooed" denotes the baseline case where the tattooed images are compared to the ground truth images.

Scenario	Portrait			Inner		
	MSSIM	PSNR	VIF	MSSIM	PSNR	VIF
Tattooed	0.947 (± 0.053)	31.31 (± 5.04)	0.884 (± 0.093)	0.974 (± 0.027)	35.37 (± 6.63)	0.879 (± 0.097)
pix2pix	0.943 (± 0.043)	33.24 (± 4.82)	0.732 (± 0.081)	0.978 (± 0.021)	37.66 (± 5.39)	0.779 (± 0.087)
TRNet	0.967 (± 0.034)	36.22 (± 6.00)	0.883 (± 0.079)	0.987 (± 0.015)	42.34 (± 6.74)	0.891 (± 0.083)

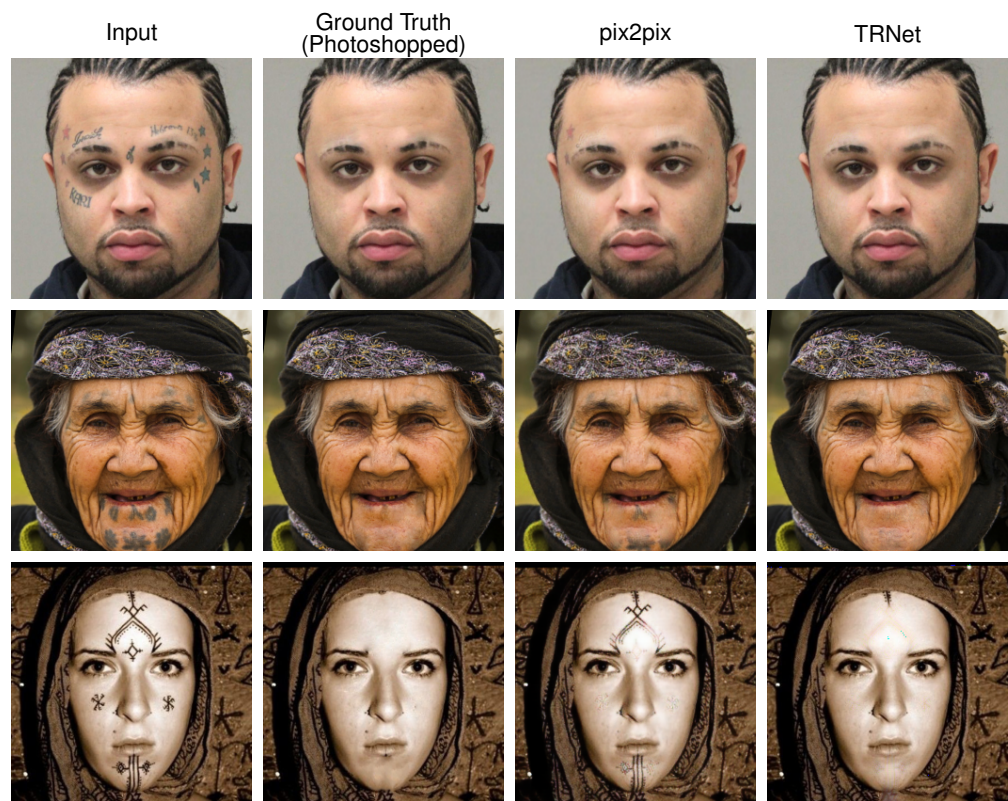


Figure 14. Examples of using deep learning-based algorithms for facial tattoo removal. Best viewed in electronic format (zoomed in).

The results indicate that the synthetic data can be used to train an algorithm for removing real tattoos from facial images, which performs well in many scenarios. However, as shown in Figure 15, the depicted images illustrate the limitation of the presented approach of not being able to remove tattoos entirely.



Figure 15. Facial images with extreme coverage of tattoos, which remain challenging for our tattoo removal approach. Before (left) and after (right) tattoo removal.

6. Application to Face Recognition

This section, describes how tattoo removal can be integrated and used in a face recognition system. A face recognition system consists of several preprocessing modules, such as face alignment and quality estimation. These modules help minimise factors that

are unimportant for face recognition and ensure that only images of sufficient quality are used during authentication. As part of the preprocessing, we propose to use the deep learning-based removal algorithms described in Section 5. While facial tattoos can be seen as distinctive and helpful in identifying individuals, tattoo removal is useful for face recognition in cases where only one of the face images in a comparison contains tattoos [5]. In our experiments, we trained the classifiers to remove facial tattoos from aligned images and as such will assume that our input images have already been aligned since our focus is on improving feature extraction and comparison. Note that the proposed tattoo removal method could also be retrained on unaligned images and placed before the detection module to improve detection accuracy.

6.1. Experimental Setup

In the following, we describe the database, the employed face recognition system, and metrics used to evaluate the biometric performance:

- **Database:** for the evaluation, we use the publicly available database *HDA Facial Tattoo and Painting Database* <https://dasec.h-da.de/research/biometrics/hda-facial-tattoo-and-painting-database> (accessed on 15 December 2022), which consists of 250 image pairs of individuals with and without real facial tattoos. The database was originally collected by Ibsen et al. in [5]. The images have all been aligned using the RetinaFace facial detector [59]. Examples of original image pairs (before tattoo removal) are given in Figure 16. These pairs of images are used for evaluating the performance of a face recognition system. For evaluating the effect of tattoo removal, the models described in Section 5.1 are employed on the facial images containing tattoos, whereafter the resulting images are used during the evaluation.

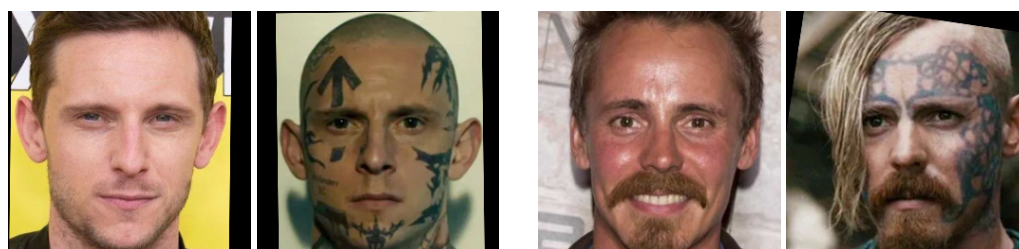


Figure 16. Examples of image pairs in the HDA facial tattoo and painting database.

- **Face recognition system:** to evaluate the applicability of tattoo removal for face recognition, we use the established ArcFace pre-trained model (LResNet100E-IR,ArcFace@ms1m-refine-v2) with the RetinaFace facial detector.
- **Recognition performance metrics:** the effect of removing facial tattoos is evaluated empirically [60]. Specifically, we measure the FNMR at operationally relevant thresholds corresponding to a FMR of 0.1% and 1%:
 - False Match Rate (FMR): the proportion of the completed biometric non-mated comparison trials that result in a false match.
 - False Non-Match Rate (FNMR): the proportion of the completed biometric mated comparison trials that result in a false non-match.

Additionally, we report the Equal Error Rate (EER), i.e., the point where FNMR and FMR are equal. To show the distribution of comparison scores, boxplots are used. The comparison scores are computed between pairs of feature vectors using the Euclidean distance.

A subset of semi-controlled images from the FRGCv2 dataset are used to obtain non-mated comparison scores during the experiments. These scores are used during the experimental evaluating for calculating the operating points corresponding to a FMR of 0.1% and 1% and together with mated scores, obtained on the tattoo database, for computing the EER before and after tattoo removal.

6.2. Experimental Results

The effect of removing tattoos on the computed comparison scores is visualised in Figure 17. As can be seen, the comparison scores are not significantly affected for the pix2pix model, which only showed moderate capabilities of removing tattoos from facial images. However, for TRNet, which has been trained on the synthetic database, it is shown that the dissimilarity scores, on average, gets lower, which indicates that the recognition performance might improve.

Table 3 shows the biometric performance scores calculated on the tattooed images and the inpainted facial images for the different used models. The scores indicate that realistic removal of tattoos (TRNet) might improve face recognition performance since we can observe that, compared to the baseline (tattooed), the EER is halved, and the FNMR at an FMR of 1% is reduced to 0%. The results indicate that a tattoo removal module can be integrated into the processing chain of a face recognition system and help make it more robust towards facial tattoos.

Table 3. Biometric performance results for ArcFace.

Type	EER%	FNMR%	
		FMR = 0.1%	FMR = 1%
Tattooed	0.80	1.20	0.80
pix2pix	0.80	1.60	0.80
TRNet	0.40	1.20	0.00

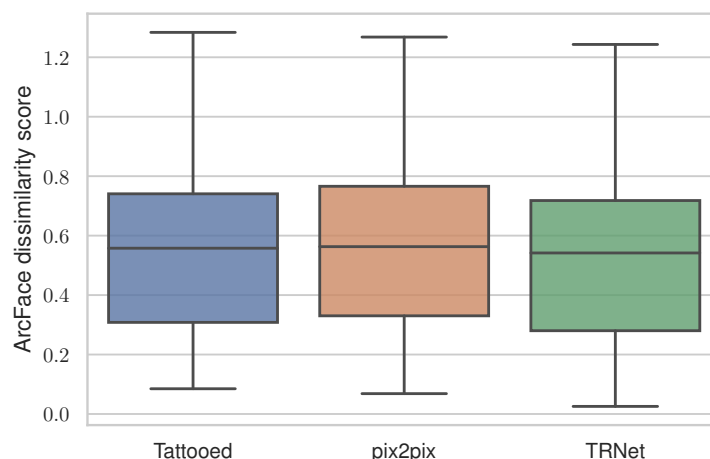


Figure 17. Boxplots showing the effect of tattoo removal on biometric comparison scores.

7. Conclusions

In this paper, we proposed an automatic approach for blending tattoos onto facial images and showed that it is possible to use synthetic data to train a deep learning-based facial tattoo removal algorithm, thereby enhancing the performance of a state-of-the-art face recognition system. To create a facial image with tattoos, the face is first divided into face regions using landmark detection whereafter tattoo placements can be found. Subsequently, deep reconstruction maps and cut-out maps can be estimated from the input image. Thereafter, the information is combined to realistically blend tattoos onto the facial image. Using this approach, we created a large database of facial images with tattoos and used it to train a deep learning-based algorithm for removing tattoos. Experimental results show a high quality of the tattoo removal. To further show the feasibility of the reconstruction, we evaluated the effect of removing facial tattoos on a state-of-the-art face recognition system and found that it can improve automated face recognition performance. Hence, the findings of this paper demonstrate the usefulness of synthetic data for the facial analysis tasks of tattoo removal. Further experiments and optimisations are out of scope

of this work but could be investigated in the future. Additionally, it could be relevant to explore if synthetic data can be generated for other types of facial manipulations and leveraged for facial analysis tasks on real data.

Author Contributions: Conceptualization, Mathias Ibsen, Christian Rathgeb, Pawel Drozdowski and Christoph Busch; Methodology, Mathias Ibsen, Christian Rathgeb, Pawel Drozdowski and Christoph Busch; Software, Mathias Ibsen; Formal analysis, Mathias Ibsen; Investigation, Mathias Ibsen, Christian Rathgeb, Pawel Drozdowski and Christoph Busch; Data curation, Mathias Ibsen; Writing—original draft, Mathias Ibsen; Writing—review and editing, Christian Rathgeb, Pawel Drozdowski and Christoph Busch; Visualization, Mathias Ibsen and Christian Rathgeb; Supervision, Christian Rathgeb, Pawel Drozdowski and Christoph Busch; Funding acquisition, Christian Rathgeb and Christoph Busch. All authors have read and agreed to the published version of the manuscript.

Funding: This research work has been funded by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 860813—TReSPAsS-ETN.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Zeng, D.; Veldhuis, R.; Spreeuwiers, L. A survey of face recognition techniques under occlusion. *IET Biom.* **2021**, *10*, 581–606.
2. Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Taleb-Ahmed, A. Past, Present, and Future of Face Recognition: A Review. *Electronics* **2020**, *9*, 1188.
3. Kurutz, S. Face Tattoos Go Mainstream. 2018. Available online: <https://www.nytimes.com/2018/08/04/style/face-tattoos.html> (accessed on 3 December 2022).
4. Abrams, M. Why Are Face Tattoos the Latest Celebrity Trend. 2020. Available online: <https://www.standard.co.uk/insider/style/face-tattoos-celebrity-trend-justin-bieber-presley-gerber-a4360511.html> (accessed on 3 December 2022).
5. Ibsen, M.; Rathgeb, C.; Fink, T.; Drozdowski, P.; Busch, C. Impact of facial tattoos and paintings on face recognition systems. *IET Biom.* **2021**, *10*, 706–719.
6. Zhao, S.; Liu, W.; Liu, S.; Ge, J.; Liang, X. A hybrid-supervision learning algorithm for real-time un-completed face recognition. *Comput. Electr. Eng.* **2022**, *101*, 108090.
7. Mathai, J.; Masi, I.; AbdAlmageed, W. Does Generative Face Completion Help Face Recognition? In Proceedings of the International Conference on Biometrics (ICB), Crete, Greece, 4–7 June 2019; pp. 1–8.
8. Bacchini, F.; Lorusso, L. A tattoo is not a face. Ethical aspects of tattoo-based biometrics. *J. Information, Commun. Ethics Soc.* **2017**, *16*, 110–122.
9. Wood, E.; Baltrusaitis, T.; Hewitt, C.; Dziadzio, S.; Cashman, T.J.; Shotton, J. Fake It Till You Make It: Face Analysis in the Wild Using Synthetic Data Alone. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10–17 October 2021; pp. 3681–3691.
10. Joshi, I.; Grimmer, M.; Rathgeb, C.; Busch, C.; Bremond, F.; Dantcheva, A. Synthetic Data in Human Analysis: A Survey. *arXiv* **2022**, arXiv:2208.09191.
11. Rathgeb, C.; Dantcheva, A.; Busch, C. Impact and Detection of Facial Beautification in Face Recognition: An Overview. *IEEE Access* **2019**, *7*, 152667–152678.
12. European Council. Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). 2016. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679> (accessed on 13 December 2022).
13. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.C.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the Annual Conference on Neural Information Processing Systems 2014, Montreal, QC, Canada, 8–13 December 2014; Volume 27.
14. Karras, T.; Laine, S.; Aila, T. A Style-Based Generator Architecture for Generative Adversarial Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; pp. 4396–4405.

15. Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; Aila, T. Analyzing and Improving the Image Quality of StyleGAN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 8107–8116.
16. Karras, T.; Aittala, M.; Laine, S.; Härkönen, E.; Hellsten, J.; Lehtinen, J.; Aila, T. Alias-Free Generative Adversarial Networks. In Proceedings of the NeurIPS, virtual, 6–14 December 2021.
17. Grimmer, M.; Raghavendra, R.; Christoph, C. Deep Face Age Progression: A Survey. *IEEE Access* **2021**, *9*, 83376–83393.
18. Cappelli, R.; Maio, D.; Maltoni, D. SFinGe: An Approach to Synthetic Fingerprint Generation. In Proceedings of the International Workshop on Biometric Technologies, Calgary, Canada, June 2004.
19. Priesnitz, J.; Rathgeb, C.; Buchmann, N.; Busch, C. SynCoLFinGer: Synthetic Contactless Fingerprint Generator. *arXiv* **2021**, arXiv:2110.09144.
20. Wyzkowski, A.B.V.; Segundo, M.P.; de Paula Lemes, R. Level Three Synthetic Fingerprint Generation. In Proceedings of the 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 9250–9257.
21. Drozdowski, P.; Rathgeb, C.; Busch, C. SIC-Gen: A Synthetic Iris-Code Generator. In Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG), Darmstadt, Germany, 20–22 September 2017; pp. 61–69.
22. Dole, J. Synthetic Iris Generation, Manipulation, & ID Preservation. 2021. Available online: <https://eab.org/cgi-bin/dl.pl?/upload/documents/2256/06-Dole-SyntheticIrisPresentation-210913.pdf> (accessed on 3 December 2022),
23. Xu, X.; Matkowski, W.M.; Kong, A.W.K. A portrait photo-to-tattoo transform based on digital tattooing. *Multimed. Tools Appl.* **2020**, *79*, 24367–24392.
24. Madhavan, V. SkinDeep. 2021. Available online: <https://github.com/vijishmadhavan/SkinDeep> (accessed on 3 December 2022).
25. Singh, R.; Vatsa, M.; Bhatt, H.S.; Bharadwaj, S.; Noore, A.; Nooreyzedan, S.S. Plastic Surgery: A New Dimension to Face Recognition. *IEEE Trans. Inf. Forensics Secur.* **2010**, *5*, 441–448.
26. Rathgeb, C.; Dogan, D.; Stockhardt, F.; Marsico, M.D.; Busch, C. Plastic Surgery: An Obstacle for Deep Face Recognition? In Proceedings of the 15th IEEE Computer Society Workshop on Biometrics (CVPRW), Seattle, WA, USA, 14–19 June 2020; pp. 3510–3517.
27. International Civil Aviation Organization. Machine Readable Passports—Part 9—Deployment of Biometric Identification and Electronic Storage of Data in eMRTDs, 2021. Available online: https://www.icao.int/publications/documents/9303_p9_cons_en.pdf (accessed on 15 December 2022)
28. Dantcheva, A.; Chen, C.; Ross, A. Can facial cosmetics affect the matching accuracy of face recognition systems? In Proceedings of the IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), Arlington, VA, USA, 23–27 September 2012; pp. 391–398.
29. Wang, T.Y.; Kumar, A. Recognizing human faces under disguise and makeup. In Proceedings of the IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), Sendai, Japan, 29 February–2 March 2016; pp. 1–7.
30. Chen, C.; Dantcheva, A.; Swearingen, T.; Ross, A. Spoofing faces using makeup: An investigative study. In Proceedings of the IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), New Delhi, India, 22–24 February 2017; pp. 1–8.
31. Rathgeb, C.; Drozdowski, P.; Fischer, D.; Busch, C. Vulnerability Assessment and Detection of Makeup Presentation Attacks. In Proceedings of the International Workshop on Biometrics and Forensics (IWBF), Porto, Portugal, 29–30 April 2020; pp. 1–6.
32. Singh, M.; Singh, R.; Vatsa, M.; Ratha, N.K.; Chellappa, R. Recognizing Disguised Faces in the Wild. *Trans. Biom. Behav. Identity Sci. (TBIOM)* **2019**, *1*, 97–108.
33. Ferrara, M.; Franco, A.; Maltoni, D. The magic passport. In Proceedings of the IEEE International Joint Conference on Biometrics, Clearwater, FL, USA, 29 September–2 October 2014; pp. 1–7.
34. Scherhag, U.; Rathgeb, C.; Merkle, J.; Breithaupt, R.; Busch, C. Face Recognition Systems Under Morphing Attacks: A Survey. *IEEE Access* **2019**, *7*, 23012–23026.
35. Rathgeb, C.; Botaljov, A.; Stockhardt, F.; Isadskiy, S.; Debiassi, L.; Uhl, A.; Busch, C. PRNU-based Detection of Facial Retouching. *IET Biom.* **2020**, *9*, 154–164.
36. Hedberg, M.F. Effects of sample stretching in face recognition. In Proceedings of the 19th International Conference of the Biometrics Special Interest Group, online, 16–18 September 2020; pp. 1–4.
37. Verdoliva, L. Media Forensics and DeepFakes: An Overview. *IEEE J. Sel. Top. Signal Process.* **2020**, *14*, 910–932.
38. Ferrer, C.C.; Pflaum, B.; Pan, J.; Dolhansky, B.; Bitton, J.; Lu, J. Deepfake Detection Challenge Results: An Open Initiative to Advance AI. 2020. Available online: <https://ai.facebook.com/blog/deepfake-detection-challenge-results-an-open-initiative-to-advance-ai/> (accessed on 3 December 2022),
39. Tolosana, R.; Vera-Rodriguez, R.; Fierrez, J.; Morales, A.; Ortega-Garcia, J. Deepfakes and beyond: A Survey of face manipulation and fake detection. *Inf. Fusion* **2020**, *64*, 131–148.
40. Iizuka, S.; Simo-Serra, E.; Ishikawa, H. Globally and Locally Consistent Image Completion. *ACM Trans. Graph.* **2017**, *36*, 107.
41. Li, Y.; Liu, S.; Yang, J.; Yang, M.H. Generative Face Completion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 5892–5900.
42. Zhao, Y.; Chen, W.; Xing, J.; Li, X.; Bessinger, Z.; Liu, F.; Zuo, W.; Yang, R. Identity Preserving Face Completion for Large Ocular Region Occlusion. In Proceedings of the 29th British Machine Vision Conference (BMVC), Newcastle, UK, 3–6 September 2018.

43. Song, L.; Cao, J.; Song, L.; Hu, Y.; He, R. Geometry-Aware Face Completion and Editing. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, USA, 27 January–1 February 2019; pp. 2506–2513.
44. Din, N.U.; Javed, K.; Bae, S.; Yi, J. A Novel GAN-Based Network for Unmasking of Masked Face. *IEEE Access* **2020**, *8*, 44276–44287.
45. King, D. Dlib-ml: A Machine Learning Toolkit. *J. Mach. Learn. Res.* **2009**, *10*, 1755–1758.
46. Feng, Y.; Wu, F.; Shao, X.; Wang, Y.; Zhou, X. Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network. In Proceedings of the ECCV, Munich, Germany, 8–14 September 2018.
47. Phillips, P.J.; Wechsler, H.; Huang, J.; Rauss, P.J. The FERET database and evaluation procedure for face-recognition algorithms. *Image Vis. Comput.* **1998**, *16*, 295–306.
48. Phillips, P.J.; Flynn, P.J.; Scruggs, W.T.; Bowyer, K.W.; Chang, J.; Hoffman, K.; Marques, J.; Min, J.; Worek, W.J. Overview of the Face Recognition Grand Challenge. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 947–954.
49. Liu, Z.; Luo, P.; Wang, X.; Tang, X. Deep Learning Face Attributes in the Wild. In Proceedings of the IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, 7–13 December 2015.
50. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 5967–5976.
51. Howard, J.; et al. fastai. 2018. Available online: <https://github.com/fastai/fastai> (accessed on 15 December 2022).
52. Shi, W.; Caballero, J.; Huszar, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1874–1883.
53. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 11–14 October 2016.
54. Gatys, L.A.; Ecker, A.S.; Bethge, M. A Neural Algorithm of Artistic Style. *arXiv* **2015**, arXiv:1508.06576.
55. Liu, S.; Deng, W. Very deep convolutional neural network based image classification using small training sample size. In Proceedings of the 3rd IAPR Asian Conf. on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 3–6 November 2015; pp. 730–734.
56. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612.
57. Sheikh, H.R.; Bovik, A.C. Image information and visual quality. *IEEE Trans. Image Process.* **2006**, *15*, 430–444.
58. Khalel, A. Sewar. 2021. Available online: <https://github.com/andrewekhalel/sewar> (accessed on 3 December 2022),
59. Deng, J.; Guo, J.; Ververas, E.; Kotsia, I.; Zafeiriou, S. RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, 13–19 June 2020.
60. ISO/IEC JTC1 SC37 Biometrics; ISO/IEC 19795-1:2021. Information Technology—Biometric Performance Testing and Reporting—Part 1: Principles and Framework. International Organization for Standardization: Geneva, Switzerland, 2021.