DEEP IRIS COMPRESSION

Ehsaneddin Jalilian, Heinz Hofbauer, and Andreas Uhl

University of Salzburg, Department of Computer Science
Jakob-haringer-straße 2, 5020 Salzburg, Austria
{ejalilian, hhofbaue, uhl}@cs.sbg.ac.at

Abstract. Lossy image compression can reduce the space and bandwidth required for image storage and transmission, which is increasinly in demand by the iris recognition systems developers. Deep learning techniques (*i.e.* CNN, and GAN networks) are quickly becoming a tool of choice for general image compression tasks. But some key quality criteria, such as high perceptual quality and the spatial precision of the images, need to be satisfied when applying such modules for iris images compression tasks. We investigate and evaluate the expediency of a deep learning based compression model for iris data compression. In particular, we relate rate-distortion performance as measured in PSNR, and Multi-scale Structural Similarity Index (MS-SSIM) to the recognition scores as obtained by a concrete recognition system. We further compare the model performance against a state-of-the-art deep learning base image compression technique as well as some lossy compression algorithms currently used for iris compression (namely: the current ISO standard JPEG2000, JPEG, H.265 derivate BPG, and WEBP), to figure out the most suited compression algorithm which can be used for this purpose. The experimental results show superior compression, and promising recognition performance of the model over all other techniques on different iris data.

Keywords: Deep Learning · Iris compression · Iris recognition.

1 Introduction

Efficient storage and rapid transmission of iris biometric records is a driving implementation factor in iris recognition systems (especially on low-powered mobile sensors and for portable devices). The International Organization for Standardization (ISO) specifies that iris biometric data should be recorded and stored in (raw) image form (ISO/IEC FDIS 19794-6), rather than in extracted templates (e.g. iris-codes). Such deployments can directly benefit from future improvements which can be easily incorporated, thus enabling more interoperability and vendor neutrality [5]. Image compression techniques can be generally divided into lossless and lossy compression. Lossless techniques compress an image by removing statistical redundancy while lossy compression algorithms typically use inexact approximations, and partial data discarding, to represent the content, exploiting the fact that the human eye is insensitive to certain visual features.

Rapid development of deep learning theory and neural networks has introduced a new image compression paradigm to the lossy compression technology. Deep learning based methods mostly use convolutional neural networks (CNNs) to design image codecs. Benefiting from the strong learning ability of neural networks, these models can learn image characteristics through back propagation and conduct the compression of image information without too much prior knowledge. Recently, it was shown that deep learning can even synthesize a image using only a semantic segmentation map as input, thanks to the generative adversarial networks (GAN) technology [6]. This advocates the possibility of developing efficient image compression modules employing deep learning networks and the associated image synthesis. The major draw back of applying the GAN networks is their lack of spatial precision, which results in structural distortions in the reconstructed images. This could cripple the functionality of the key iris recognition modules such as the segmentation (which *e.g.* is based on circular features) and/or the feature extraction algorithms. Thus, the ability of the compression module to preserve the unique iris features in the reconstructed images, with high the spatial precision and perceptual quality, is required when it comes to iris image compression.

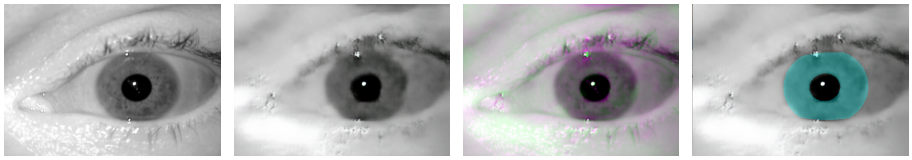
In this work we investigate the expediency of a deep semantic segmentation-based layered image compression (DSSLIC) model [3] for iris compression within a biometric recognition framework. The model leverages the power of GAN networks to encode the key iris features with high precision in the compressed images, while preserving the spatial precision and perceptual quality of the reconstructed iris images. The GAN network takes the segmentation map as the input and tries to learn the missing detail information of the up-sampled version of a compacted input image to minimize the distortion of the synthesized images. The segmentation map of the iris (raw) image is losslessly encoded as the base layer of the bit-stream. At the same time, the input image and the segmentation map are used by a deep network to obtain a low-dimensional compact representation of the input, which is encoded into the bit-stream as the first enhancement layer. The compact image and the segmentation map are then used to obtain a coarse reconstruction of the image. The residual between the input and the coarse reconstruction is encoded as the second enhancement layer in the bit-stream.

We use this model along with four other commonly used compression algorithms (BPG, JPEG2000, JPEG, WEBP) as well as a state-of-the-art deep learning based model to compress iris images in five well-known datasets. The visual quality of the compression in each cases is measured and compared against each other in terms of Peak Signal to Noise Ratio (PSNR), and Multi-scale Structural Similarity Index (MS-SSIM). Then the biometric recognition performance is evaluated, in terms of Equal Error Rate (EER), by using the compressed iris images in a regular iris biometric system. At the end, the compression and the corresponding recognition results are compared and carefully investigated to figure out a well suited compression algorithms to be employed in iris recognition systems.

2 Related Work

Classic Iris Image Compression: Numerous studies are conducted on iris image compression and recognition during the past decades (*e.g.* [5] [9]). Grother [7] explored existing approaches and compared JPEG and JPEG2000 to provide a quantitative support to the revision of the ISO/IEC IS 19794-6, including a cropped format (IREX K3), masked and cropped image format (IREX K7), and an unsegmented polar format (IREX K16). Matschitsch *et al.* [15] investigated the impact of using different lossy compression algorithms on the matching accuracy of iris recognition systems, relating rate-distortion performance to the matching scores. They concluded that JPEG2000, SPIHT and PRVQ are almost equally well suited for iris compression. Korvath *et al.* [8] investigated the impact of dedicated lossless image codecs (lossless JPEG, JPEG-LS, PNG, and GIF), lossless variants of lossy codecs (JPEG2000, JPEG XR, and SPIHT), and some general purpose file compression schemes on the iris images.

Deep Image Compression: In recent years, number of learning based image compression methods have been proposed as well. Toderici *et al.* [24] proposed recurrent neural networks based on convolution and deconvolution long short-term memory (LSTM) to extract binary representations, which are then compressed with entropy coding. In [4], a model that involved a generalized divisive normalization (GDN)-based nonlinear analysis transform, a 2-uniform quantizer, and a nonlinear synthesis transform were proposed. Johnston *et al.* [12] used the structural similarity (SSIM) quality measure and spatially adaptive bit allocation to further improve the performance. Theis *et al.* [23] proposed an auto-encoder where they used smooth approximation instead of quantization to get different rates. Agustsson *et al.* [1], introduced a soft-to-hard vector quantization model along with a unified formulation for both the compression of deep learning models and image compression. The authors in [25] proposed a compression bit allocation algorithm to allow the recurrent neural network (RNN)-based compression network to hierarchically compress the images according to semantic importance maps. Li *et al.* [13] proposed a model based on image content weighting. They used the edge feature map, extracted by the convolution neural network, as the importance map of the original image. In [11], a compact convolutional neural network (ComCNN) and a reconstruction convolutional neural network (RecCNN), were used to encode and decode the original image, respectively. An innovative algorithm solves the non-differentiated calculation in the quantization rounding function to achieve a backward propagation gradient in the standard image algorithm. In [14] the authors combined image compression and classification to reconstruct the images and generate corresponding semantic representations at the same time. Mantzer *et al.* [16] proposed a conditional probability models for deep image compression (CPDIC), focusing on improving the entropy rate of the latent image representation using a context model (a 3D-CNN which learns a conditional probability model of the latent distribution of the auto-encoder). During training the auto-encoder makes use of the context model to estimate the entropy of its representation, and the context model



(a) (b) (c) (d)

Fig. 1: An iris image (a) and its corresponding output using a GAN based model (b), along with their difference image (c), and the overlaid ground-truth (d)

is concurrently updated to learn the dependencies between the symbols in the latent representation.
Some recent works used generative adversarial networks (GAN) in their learning-based image compression schemes. Santurkar *et al.* [21], used a discriminator to help training an decoder. They used perceptual loss based on the feature map of an, ImageNet-pretrained, AlexNet, although only low-resolution image coding results were reported. Ripple *et al.* [20] embedded an auto-encoder in a GAN framework in which the feature extraction adopted a pyramid of inter-scale alignments. They considered the target and its reconstruction jointly as a single example and, instead of producing an output for classification at the last layers of the pipeline, accumulated scalar outputs along branches constructed at different depths. An average of these scalars was used as the final value provided to a sigmoid function. The discriminator also extracted outputs from different layers, similar to the pyramid feature generation. Augustesson *et al.* [2] proposed a segmentation map-based image synthesis model based on the GAN networks operating at extremely low bitrates. The framework combines an encoder, decoder/generator and a multi-scale discriminator, which are trained jointly for a generative learned compression objective. The main draw back of these models (which operate directly on the input image) is that they generate outputs which have some structural distortions and lack the sufficient spatial precision. In fact, as already mentioned, such distortions will change the iris structure and can cause the different system modules to fail. Fig. 1 shows a sample iris image (1a) (from the Notredame iris dataset, as used in this work), and its compressed version (1b), using the last GAN based model, along with their difference image (1c), and the overlaid ground-truth mask (1d). Gray regions in the difference image show where the two images have the same intensities, and magenta and green regions show where the intensities are different. Also the overlaid ground-truth mask shows how the actual iris outer and inner boundaries (as specified by the mask) are distorted in the compressed (reconstructed) image.

3 Deep Compression Model

Fig.2 illustrates the overall scheme of the model used in this work, which is derived from the model already proposed in [3]. As a key distinction to the current

GAN based models, the embedded GAN network in this model does not operate directly on the input the iris image. Instead it takes the segmentation map as the input and tries to learn the missing detail information of the up-sampled version of the compacted input image to minimize the distortion of the synthesized images. This is made practical due to the recent advancement in deep learning techniques which have made it easy to access such segmentations with very high accuracy and in a timely manner (*i.e.* [10]). In fact here we didn't use a segmentation network in our model, and fed the manually segmented labels directly to the model instead. Doing so we helped to improve the models performance by introducing more accurate labeling data to the model. The encoder includes two deep learning networks: CompNet and FiNet (GAN-based network). An input iris image is fed into the ComNet, while the segmentation map is encoded to serve as side information to this network for generating a low-dimensional version of the original image. Both the segmentation map and compact version are losslessly encoded using the FLIF codec [22], which is a state-of-the-art lossless image codec. Given the segmentation map and up-sampled compact image, the FiNet tries to obtain a high-quality reconstruction of the input image. Note that although GAN-based synthesized images from segmentation maps are visually appealing their details can be quite different from the original images. To minimize the distortion of the synthesized images the up-sampled version of the compact image, as an additional input, is added to it. In this way the FiNet is trained to learn the missing detail information of the up-sampled version of a compact image with respect to the input image, which in turn controls the output of the GAN network. After adding the up-sampled version of the compact image and the FiNet's output we get a better estimate of the input. The residual difference between the input and the estimate is then obtained and encoded by a lossy codec (H.265/HEVC intra coding-based BPG). In order to deal with negative values, the residual image is re-scaled to [0, 255] with a min-max normalization before encoding. The min and max values are also sent to decoder for inverse scaling. In this scheme the segmentation map serves as the base layer and the compact image and the residual are the first and second enhancement layers respectively. At the decoder side the segmentation map and the compact representation are decoded to be used by the FiNet to get an estimate of the input image. The output of FiNet is then added to the decoded residual image to get the reconstructed image as output.

4 Experimental Framework

Datasets: We used five different iris datasets in our experiments: The Notredame dataset (including 835 iris images of 30 different subjects)[1]. The Casia4i dataset (containing 2640 iris images of 249 subjects)[2]. The IITD dataset (containing

[1] https://sites.google.com/a/nd.edu/public-cvrl/data-sets
[2] http://biometrics.idealtest.org

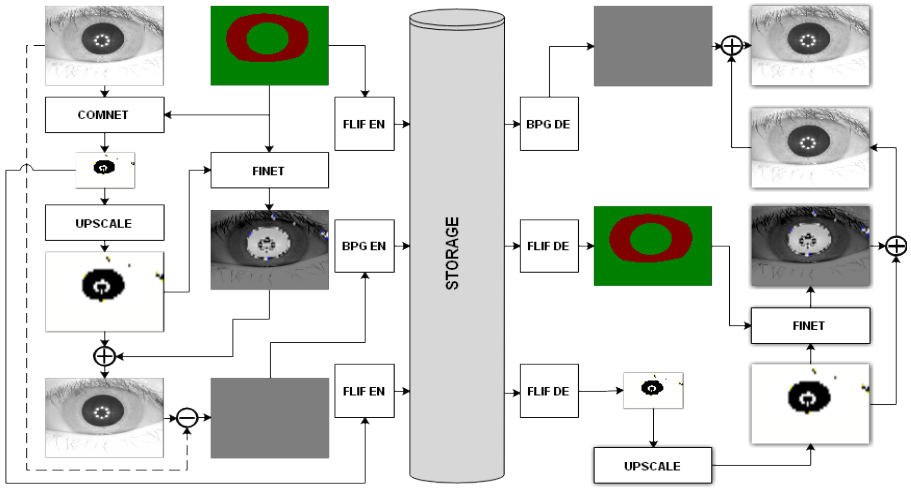


Fig. 2: The deep learning based iris compression model

2240 iris images of 224 subjects)[3]. The Casia5a dataset (including 1880 images of both eyes of 94 users)[4].

Compression Algorithms: To evaluate and compare the model performance against some popular and state-of-the-art lossy compression algorithms we considered: JPEG, the current ISO standard JPEG2000 (J2K), the H.265 derived BPG, the WEBP algorithms, and a deep learning based image compression model termed Conditional Probability Models for Deep Image Compression (CPDIC) [16]. The overall compression model consists of an encoder, a quantizer, and a decoder. The encoder $E : \mathbb{R}^d \rightarrow \mathbb{R}^m$ maps an input image to a latent representation y which is in form of a 3D feature map. The encoder architecture consists of convolution and ReLU layers combined with 15 Residual Blocks, with skip connection between every third layer. The quantizer $Q : \mathbb{R} \rightarrow C$ then discretized the coordinates of the latent representation (y) to the $L = |P|$ centers, which can then be losslessly encoded into the bit-stream. Specifically, given centers $P = \{p_1, \ldots, p_L\} \subset \mathbb{R}$, the quantizer uses nearest neighbor assignments to compute

$$\tilde{y}_i = Q(y_i) := argmin_j ||y_i - p_j||, \tag{1}$$

relying on the following differentiable soft quantization:

$$\tilde{y}_i = \sum_{j=1}^{L} \frac{exp(-\sigma||y_i - p_j||)}{\sum_{l=1}^{L} exp(-\sigma||y_i - p_l||)} p_j. \tag{2}$$

[3] http://www4.comp.polyu.edu.hk/ csajaykr/database.php
[4] http://www.biometrics.idealtest.org

Table 1: Selected compression parameters (par) and their corresponding compression performance in bits per pixel (bpp) for each algorithm.

Dataset	Casia4i				Casia5a				IITD				Notredame			
Method	par	bpp	par	bpp	par	bpp	par	bpp	par	bpp	par	bpp	par	bpp	par	bpp
DSSLIC	23	0.20	16	0.44	23	0.16	14	0.45	27	0.30	19	0.53	23	0.16	14	0.51
BPG	37	0.21	30	0.54	30	0.19	24	0.42	33	0.29	26	0.60	33	0.18	24	0.55
J2K	35	0.23	21	0.55	45	0.18	14	0.55	28	0.27	14	0.55	45	0.18	14	0.55
JPEG	23	0.20	57	0.50	12	0.19	57	0.51	17	0.30	57	0.53	09	0.18	57	0.58
WEBP	1	0.21	82	0.44	45	0.20	82	0.44	1	0.29	45	0.57	25	0.19	82	0.57
CPDIC	11	0.29	22	0.60	11	0.27	22	0.57	11	0.29	22	0.60	11	0.27	22	0.57
bpp	A (0.30)		B (0.60)		A (0.30)		B (0.60)		A (0.30)		B (0.60)		A (0.30)		B (0.60)	

The decoder D, which has a similar architecture as the encoder, forms the reconstructed image from the quantized latent representation, which is in turn (losslessy) decoded from the bit-stream

Metrics and Measures: To measure the compression performance: Peak Signal-to-Noise Ratio (PSNR), which is a mathematical measure of image quality based on the pixel difference between input images, and Multi-Scale Structural Similarity Index Measure (MS-SSIM) are used. Unlike in Structural Similarity Index Measure (SSIM), where variation in luminance, contrast and structure of "single-scale" input images are compared, MS-SSIM alliteratively down-samples the input images up to M scales. At each scale, the contrast comparison and the structure comparison are calculated. The luminescence comparison is computed only at scale M, and the final MS-SSIM evaluation is obtained by combining the measurements at different scales [26]. As an overall measure of biometric recognition performance the Equal Error Rate (EER) was chosen. It is the operation point on the receiver operating characteristic curse where the false non-match rate and the false match rate are equal.

Recognition Pipeline: We used the contrast adjusted Hough transform (CAHT) [18], and Osiris [17], for iris segmentation, local Gabor filters (LG) for feature extraction, and the Hamming distance with rotation correction for matching. Apart from the Osiris the algorithms from the USIT toolkit [19] were used.

5 Experiments and Analysis

To fit the fixed network dimensions all images are re-scaled to $256 \times 512 \times 1$. Since the networks are trained on RGB format we cloned each image two times to generate 3 channel (RGB) images ($256 \times 512 \times 3$). We applied a cross-fold scheme to train the model. First we partitioned each dataset into two equal parts and then trained and tested the model alternatively on each partition. Doing so, we tested the networks on all samples in each dataset without overlapping training and testing sets. We set the down-scaling factor $\alpha = 8$ to get the compact representation of the inputs. All models were jointly trained for 250 epochs

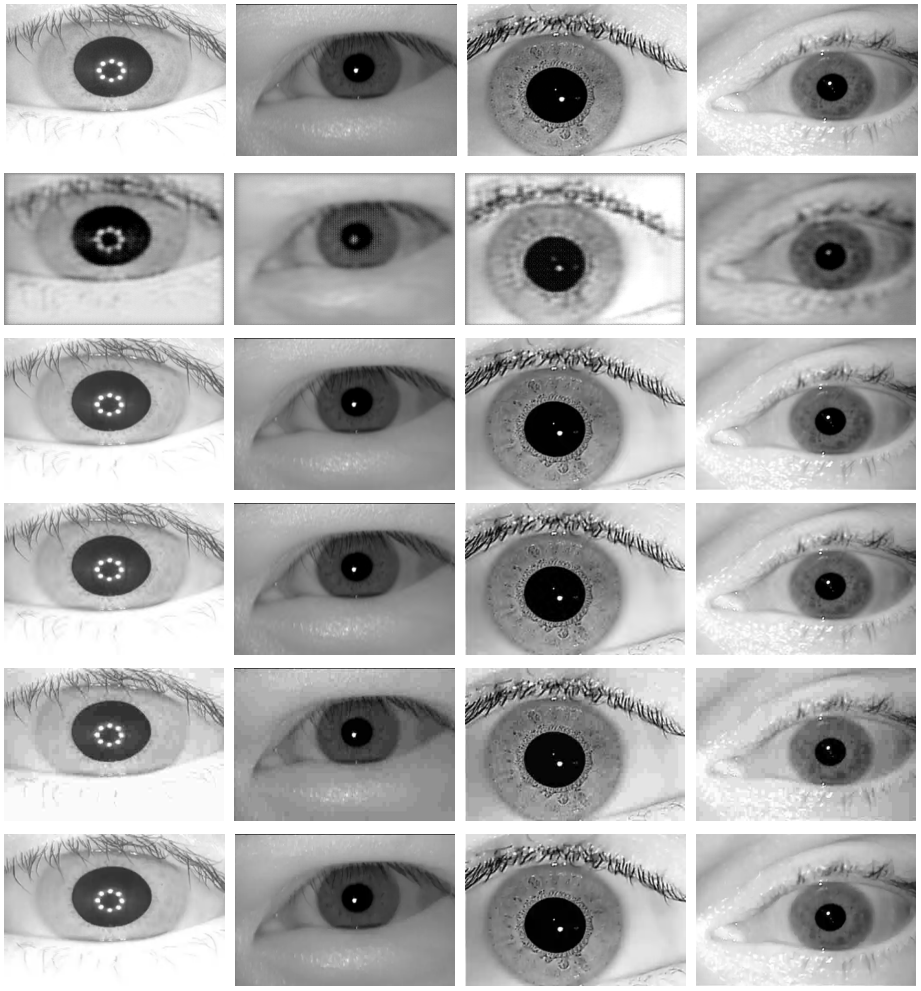


Fig. 3: Samples of highly (A) compressed iris images from the Casia4i, Casia5a, IITD, and Notredame datasets, per column respectively, using DSSLIC, CPDIC, BPG, J2K, JPEG and WEB algorithms per row respectively

with mini-batch stochastic gradient descent (SGD) and a mini-batch sizes of 2 and 8 respectively. To address the fixed bandwidth/storage compression limit requirement we set two bandwidth limits of 0.30 (A) and 0.60 (B), corresponding to the higher and the lower compression levels respectively, for each dataset in terms of bit-per-pixel (bpp). It should be noted that not all algorithms allow to set the exact output file size. Thus, we selected the compression parameter for each algorithm so that the achieved bpp of the resulting compressed images are equal to or less than the fixed bandwidth/storage limit. It is also important to note that the resulting file sizes using the DSSLIC model are the smallest in the

Table 2: Average MS-SSM scores using high (A) and low (B) compression levels

Dataset	Casia4i		Casia5a		IITD		Notredame	
Compress	B	A	B	A	B	A	B	A
DSSLIC	0.998	0.994	0.995	0.989	0.998	0.994	0.997	0.990
BPG	0.996	0.988	0.994	0.985	0.997	0.992	0.996	0.988
J2K	0.991	0.966	0.992	0.970	0.987	0.945	0.988	0.964
JPEG	0.993	0.950	0.988	0.931	0.994	0.957	0.991	0.949
WEBP	0.993	0.982	0.991	0.965	0.995	0.987	0.992	0.981
CPDIC	0.897	0.889	0.844	0.852	0.881	0.875	0.909	0.902

majority of cases (*i.e.* IITD, Notredame, Casia5a-A, and Casia4i-B) among all the algorithms. Table 1 shows the selected compression parameters (par) and the resulting bpps per algorithm and dataset. Samples of the compressed images in each dataset using the compression methods used are presented in Fig. 3 per column and row respectively.

Table 2 gives the quality results based on the MS-SSIM for each dataset (averaged over all images) for the different compression algorithms. The DSSLIC model shows superior performance over all other codecs for both compression levels considered. This is a quite remarkable result given that the files produces by the DSSLIC are smaller in size than files produces by the competing methods. Visual inspection of the corresponding output iris images as presented in the Fig. 3 (first column) shows that the model is able to preserve spatial precision and the uniqueness of the iris features very well. Across all datasets, and both compression settings, BPG is always the second-best and CPDIC is always the worst. The performance of the other three algorithms are varies depending on the dataset. The performance, in terms of rank, of the WEBP, JPEG and J2k algorithms also can vary for different compression levels (on Notredame and Casia4i). While the order of performance for the higher compression level (A) on these datasets is: WEBP, J2k, JPEG, the order of performance for the lower compression level (B) is: WEBP, JPEG, then J2k. Considering the average performance of each algorithm on the different dataset the order of performance (after DSSLIC and BPG) is: WEBP, J2k, and JPEG. Table 3 shows the corresponding results for each dataset (averaged over all images)in terms of PSNR.

Table 3: Average PSNR scores using high (A) and low (B) compression levels

Dataset	Casia4i		Casia5a		IITD		Notredame	
Compress	B	A	B	A	B	A	B	A
DSSLIC	49.1	44.0	45.2	41.6	45.5	41.3	45.7	40.3
BPG	44.5	39.8	44.0	40.9	44.7	40.5	43.5	40.0
J2K	41.8	35.5	43.1	37.9	40.5	34.4	41.1	35.5
JPEG	39.7	33.0	39.6	32.4	39.6	32.5	39.1	32.5
WEBP	41.0	37.0	41.8	37.5	41.5	37.6	41.0	37.2
CPDIC	16.1	16.0	18.7	18.7	17.8	17.8	16.8	16.4

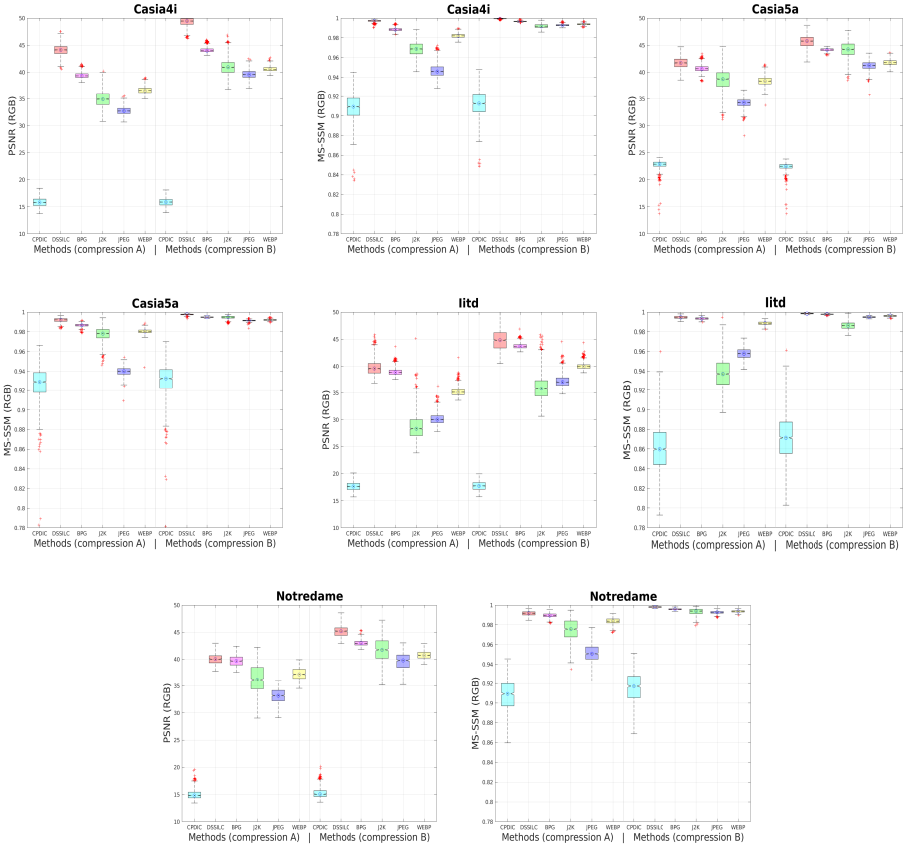


Fig. 4: Compression performance in terms of PSNR and MS-SSIM for the high (left side of the graphs) and low (right side of the graphs) compression levels

The superior performance of DSSLIC over the other algorithms is visible in these results too. Likewise, BPG ranks the second-best and CPDIC ranks the worst in terms of peak signal-to-noise ratio. Despite slight differences in the ranking orders (compared to those of the MS-SSM experiments) the other algorithms are also ranked the same when considering the average performance: WEBP, J2K, then JPEG. Figure 4 presents further details about the experimental results in the form of box-plots for each dataset. Unsurprisingly, the higher (left in the graphs) compression rate results in a decreases in performance over all datasets and algorithms (including DSSLIC).

Next, we applied the biometric recognition system to all obtained compressed iris images and evaluated the biometric comparison accuracy, in terms of EER, for the two levels of compression. In addition to the two segmentation algorithms used, Osiris and CAHT, we used the perfect segmentation produced by manual segmentation (manually annotated segmentation drop masks). The manual seg-

Table 4: EERs for the different datasets using the CAHT algorithm

Dataset	Casia4i		Casia5a		IITD		Notredame	
Compress	B	A	B	A	B	A	B	A
DSSLIC	1.2	1.0	21.1	21.2	1.4	1.8	29.9	29.9
BPG	1.0	1.2	21.6	21.3	1.6	2.4	29.6	30.3
J2K	1.1	1.3	20.6	22.3	2.0	2.6	30.0	30.1
JPEG	1.2	2.8	20.6	26.1	1.9	2.5	29.9	32.4
WEBP	1.2	1.7	21.5	23.0	2.0	2.6	30.3	31.5
CPDIC	3.4	4.0	28.8	29.4	2.3	2.8	32.5	34.1

mentation was used to disentangle the compression effect on iris texture from the segmentation performance on compressed data, and their possible failures. Tables 4 and 5 show the results for the CAHT segmentation and manual segmentation respectively. When using the CAHT segmentation recognition does not work at all for the Casia5a and Notredame datasets. For the IITD and Casia4i data the DSSLIC compression frequently shows the best performance, especially for the high compression level. When using manual segmentation results (Table 5), recognition still does not work for Notredame data, while for the remaining datasets, DSSLIC results are never surpassed by any other compression scheme. Given the fact that the DSSLIC also produces the smallest actual files these results imply that DSSLIC compression is able to preserve iris texture very well. Certainly better than the other algorithms under test, as the segmentation effects can ruled out due to manual segmentation. Table 6 shows the results when the OSIRIS algorithm is used for segmentation. ecognition on the Notredame data does not work either, but otherwise the ranking of the algorithms is fairly different. DSSLIC is the best performer only for Casia5a, while it is actually the worst performing algorithm on the IITD dataset. When comparing CAHT and OSIRIS segmentation results it's clear that the segmentation methods, and the logic behind them, react quite differently to the artifacts in the compressed images, and thus deliver very different results considering identical compressed iris images. Overall, the clearly higher rate-distortion compression performance of the DSSLIC algorithm is **not** directly translated into best recognition accuracy, except where a manual segmentation is used.

Table 5: EERs for the different datasets using manual masks

Dataset	Casia4i		Casia5a		IITD		Notredame	
Compress	B	A	B	A	B	A	A	B
DSSLIC	0.4	0.4	2.5	2.9	0.4	0.5	23.8	23.9
BPG	0.4	0.6	2.9	3.9	0.4	0.5	23.8	23.9
J2K	0.4	0.6	2.7	5.1	0.4	0.5	23.8	24.0
JPEG	0.5	1.7	3.0	14.0	0.4	0.5	23.8	25.7
WEBP	0.5	0.7	3.4	5.4	0.4	0.5	24.0	24.6
CPDIC	1.6	2.0	15.1	18.2	0.5	0.6	26.6	29.3

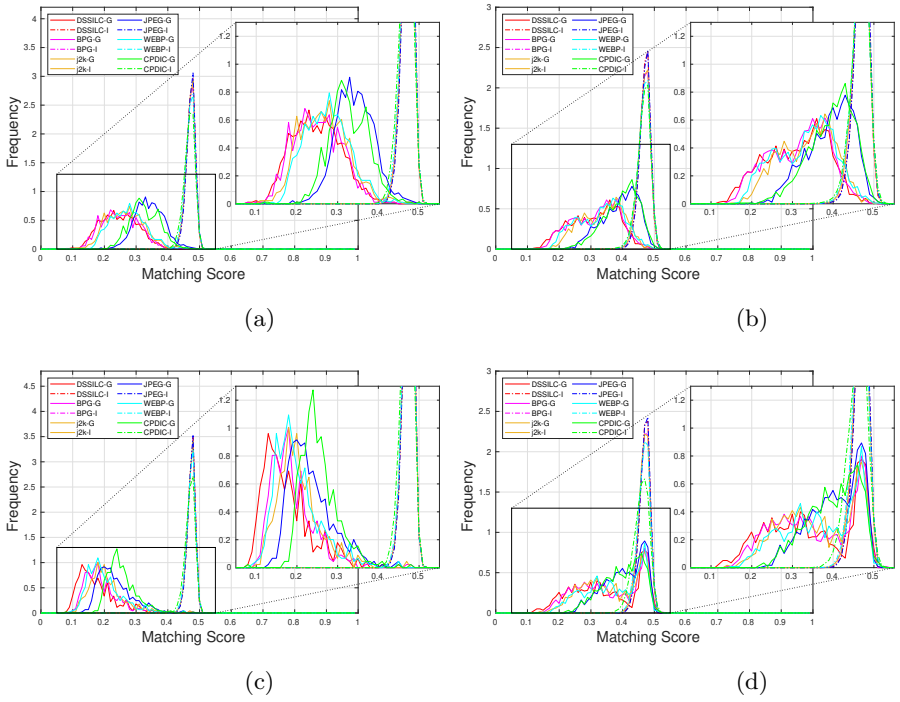


Fig. 5: Genuine and impostor distributions of the different compression methods for Casi4i (a), Casi5a (b), IITD (c), and Notredame (d) when applying the manual drop masks

It is also interesting that the recognition performance of CPDIC on Casi5a is much lower compared to other compression algorithms on the same dataset. When inspecting the iris images generated by this algorithm, along with their corresponding iris features extracted, we noticed some artifacts which were distributed uniformly over all the images in a block-like pattern. These artifacts are more severe and intense in areas of high texture, specifically the iris texture areas. For an example of this behavior see Fig. 6 where an image from the Notredame datasets is compared to an image from the Casia5a dataset. This effect seems to be due to the different performance of the encoder network on the input iris images when generating the latent representations combined with the subsequent quantization technique. The persistence of these artifacts over all images of the Casia5a dataset clearly undermined the recognition performance.

In order to clarify how the quality of the images generated by different compression algorithms affect the actual recognition performance we analysed the distribution of the genuine and impostor scores obtained using each algorithm and dataset. Figure 5 shows the genuine and impostor distributions for the different compression methods for each dataset when using the manual drop masks

Table 6: EERs for different the datasets using the Osiris algorithm

Dataset	Casia4i		Casia5a		IITD		Notredame	
Compress	B	A	B	A	B	A	A	B
DSSLIC	1.1	1.0	2.0	2.2	0.7	0.8	25.2	25.5
BPG	0.9	1.0	2.0	2.5	0.3	0.3	26.9	26.4
J2K	0.8	0.9	2.0	3.1	0.4	0.7	25.7	25.1
JPEG	0.8	1.8	2.4	9.7	0.5	0.6	24.7	24.7
WEBP	0.8	0.9	2.9	4.0	0.4	0.4	25.1	25.0
CPDIC	2.2	2.6	15.9	19.2	0.6	0.6	26.1	27.8

(to exclude the influence of segmentation errors). Each pair of curves (genuine and impostor) are indicated by color while linetype distinguishes between impostor (dash-dotted) and genuine (solid). The impostor curves remains virtually unchanged, while the genuine curves fluctuate almost in all cases. This leads us to the argument that the compression process affects the genuine scores, by introducing artifacts into the iris images which alter the distinct patterns that are present in the genuine samples, making the compressed images more dissimilar. This effect appears as higher fluctuation in genuine scores, and thus EERs.

6 Conclusion

We investigated the performance of a deep learning model (DSSLIC) for iris image compression in terms of rate-distortion and recognition accuracy. The model showed superior compression performance over all other algorithms using different datasets and compression rates. Unlike the other algorithms, the DSSLIC

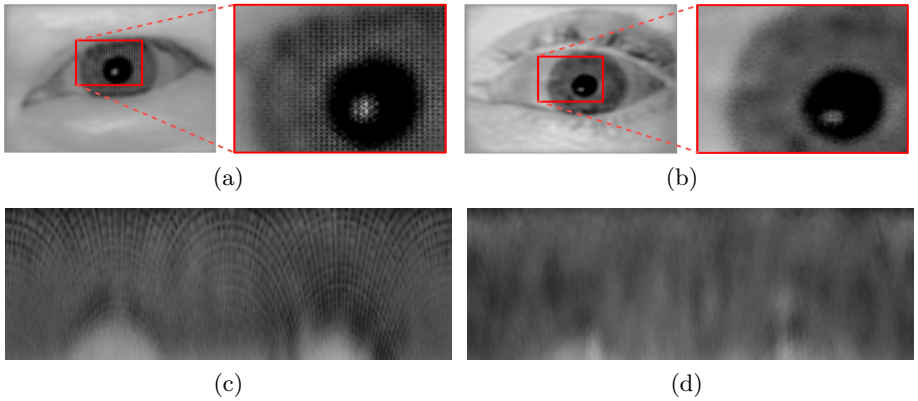


Fig. 6: A sample iris image in the Casia5a dataset (a) vs a sample image in the Notredame datasets (b), both generated by CPDIC algorithm, and their corresponding normalized outputs (c and d) respectively

model was able to cope with iris images with complex feature characteristic, and possessed stable performance on all different types of iris data. The results of the recognition experiments showed that the higher compression performance of the DSSLIC algorithm is directly translated into better recognition rates in majority of the cases. Yet, further experiments with an alternative segmentation algorithm (Osiris) revealed that the segmentation techniques, and the logic used in them, could react quiet differently to the compress images. In the case of Osiris the segmentation performance was degraded considerably. The results obtained using the manual drop masks supported this argument. The experiments also showed that an increase in compression results in reduction of recognition performance in aa majority of cases. Analysis of the genuine and the impostor scores indicated that compression process introduces artifacts into the iris images which alter the distinct patterns that are present in the genuine samples, making the compressed images more dissimilar. Overall, the results showed that the presented deep learning based model is capable of efficient iris image compression for the use in an iris biometric recognition system.

Acknowledgment

This project was partly funded from the FFG KIRAS project AUTFingerATM under grant No. 864785 and the FWF project "Advanced Methods and Applications for Fingervein Recognition" under grant No. P 32201-NBL.

References

1. Agustsson, E., Mentzer, F., Tschannen, M., Cavigelli, L., Timofte, R., Benini, L., Van-Gool, L.: Soft-to-hard vector quantization for end-to-end learned compression of images and neural networks. CoRR **abs/1704.00648** (2017)
2. Agustsson, E., Tschannen, M., Mentzer, F., Timofte, R., Gool, L.V.: Generative adversarial networks for extreme learned image compression. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 221–231 (2019)
3. Akbari, M., Liang, J., Han, J.: Dsslic: deep semantic segmentation-based layered image compression. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 2042–2046. IEEE (2019)
4. Ballé, J., Laparra, V., P-Simoncelli, E.: End-to-end optimized image compression. In: Int'l Conf on Learning Representations (ICLR). Toulon, France (April 2017)
5. Daugman, J., Downing, C.: Effect of severe image compression on iris recognition performance. IEEE Transactions on information Forensics and Security **3**(1), 52–61 (2008)
6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in neural information processing systems. pp. 2672–2680 (2014)
7. Grother, P.: Quantitative standardization of iris image formats. biometrics and electronic signatures (BIOSIG) (2009)
8. Horvath, K., Uhl, H.S.A., Weinhandel, G.: Lossless compression of polar iris image data. In: Iberian Conference on Pattern Recognition and Image Analysis. pp. 329–337. Springer (2011)

9. Ives, R., Bishop, D., Du, Y., Belcher, C.: Iris recognition: The consequences of image compression. EURASIP Journal on Advances in Signal Processing (1) (2010)
10. Jalilian, E., Uhl, A.: Iris segmentation using fully convolutional encoder–decoder networks. In: Deep Learning for Biometrics, pp. 133–155. Springer (2017)
11. Jiang, F., Tao, W., Liu, S., Ren, J., Guo, X., Zhao, D.: An end-to-end compression framework based on convolutional neural networks. IEEE Transactions on Circuits and Systems for Video Technology **28**(10), 3007–3018 (2017)
12. Johnston, N., Vincent, D., Minnen, D., Covell, M., Singh, S., Chinen, T., Hwang-Jin, S., Shor, J., Toderici, G.: Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4385–4393 (2018)
13. Li, M., Zuo, W., Gu, S., Zhao, D., Zhang, D.: Learning convolutional networks for content-weighted image compression. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3214–3223 (2018)
14. Luo, S., Yang, Y., Yin, Y., Shen, C., Zhao, Y., Song, M.: Deepsic: Deep semantic image compression. In: International Conference on Neural Information Processing. pp. 96–106. Springer (2018)
15. Matschitsch, S., Tschinder, M., Uhl, A.: Comparison of compression algorithms' impact on iris recognition accuracy. In: International Conference on Biometrics. pp. 232–241 (2007)
16. Mentzer, F., Agustsson, E., Tschannen, M., Timofte, R., Gool, L.V.: Conditional probability models for deep image compression. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4394–4402 (2018)
17. Othman, N., Dorizzi, B., Garcia-Salicetti, S.: Osiris: An open source iris recognition software. Pattern Recognition Letters **82**, 124 – 131 (2016)
18. Rathgeb, C., Uhl, A., Wild, P.: Iris Recognition: From Segmentation to Template Security, Advances in Information Security, vol. 59. Springer Verlag (2013)
19. Rathgeb, C., Uhl, A., Wild, P., Hofbauer, H.: Design decisions for an iris recognition sdk. In: Bowyer, K., Burge, M.J. (eds.) Handbook of Iris Recognition. Advances in Computer Vision and Pattern Recognition, Springer, second edition edn. (2016). https://doi.org/10.1007/978-1-4471-6784-6
20. Rippel, O., Bourdev, L.: Real-time adaptive image compression. In: Precup, D., Teh, Y.W. (eds.) Proceedings of the 34th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 70, pp. 2922–2930. PMLR, International Convention Centre, Sydney, Australia (06–11 Aug 2017)
21. Santurkar, S., Budden, D., Shavit, N.: Generative compression. In: 2018 Picture Coding Symposium (PCS). pp. 258–262 (2018)
22. Sneyers, J., Wuille, P.: Flif: Free lossless image format based on maniac compression. In: IEEE International Conference on Image Processing. pp. 66–70 (Sep 2016)
23. Theis, L., Shi, W., Cunningham, A., Huszár, F.: Lossy image compression with compressive autoencoders. In: International Conference on Learning Representations (2017)
24. Toderici, G., M.O-Malley, S., Jin-Hwang, S., Vincent, D., Minnen, D., Baluja, S., Covell, M., Sukthankar, R.: Variable rate image compression with recurrent neural networks. international conference on learning representations (2016)
25. Wang, C., Han, Y., Wang, W.: An end-to-end deep learning image compression framework based on semantic analysis. Applied Sciences **9**(17), 3580 (2019)
26. Wang, Z., P-Simoncelli, E., C-Bovik, A.: Multiscale structural similarity for image quality assessment. In: The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003. vol. 2, pp. 1398–1402. Ieee (2003)