Non-reference image quality assessment for biometric presentation attack detection

Amrit Pal Singh Bhogal, Dominik Söllinger, Pauline Trung, and Andreas Uhl
Department of Computer Sciences, University of Salzburg, Austria
Email: uhl@cosy.sbg.ac.at

***Abstract*—Non-reference image quality measures are used to distinguish real biometric data from data as used in presentation / sensor spoofing attacks. An experimental study shows that based on a set of 6 such measures, classification of real vs. fake iris, fingerprint, and face data is feasible with an accuracy of 90% on average. However, we have found that the best quality measure (combination) and classification setting highly depends on the target dataset. Thus, we are unable to provide any other recommendation than to optimise the choice of quality measure and classification setting for each specific application setting.**

I. Introduction

Biometric authentication techniques have emerged to replace or at least complement the traditional authentication methods (e.g. passwords). Consequently, various attacks have been increasingly observed threatening the reliability of this authentication approach. In particular, artefacts mimicking real biometrics traits or captured and displayed image or video footage of real biometric traits have been used to deceive biometric sensors and systems in so-called "presentation"- or "sensor-spoofing"- attacks. In general, counter-measures to such presentation attacks (or anti-spoofing [1]) in biometrics can be categorised into (1) liveness-based, (2) motion-based, and (3) texture-based methods. Liveness-based methods use signs of vitality to ensure that the image is captured from a living human being. In contrast, motion-based methods utilise unnatural movements on scenes as indication of spoofing, e.g. caused by hand motion when presenting a photo or a display to the sensor. Texture-based methods aim to explore textural artefacts in the images captured by the sensor (e.g. caused by recapturing artefacts). While liveness-based techniques are of course specific for the modality under investigation, texture-based methods often employ general purpose texture descriptors in a machine learning setting to discriminate real biometric data from spoofed variants. For example, [2] compares the attack detection performance of certain local descriptors on collections of spoofed iris, fingerprint, and face data. In order to circumvent the question which texture descriptors to choose, also generative deep learning techniques employing convolutional neural networks have been successfully used to identify spoofed data [3].

An entirely different approach is to consider the quality of the imagery in biometric anti-spoofing which can be interpreted as a specific form of texture-based technique. While this can be done in an approach entirely agnostic of the underlying modality by employing general purpose image quality measures (IQM) [4], a possible alternative is to consider specific properties of the target modality in the quality considerations (see e.g. [5] for quality assessment for face recognition spoofing detection). In this paper we revisit general purpose non-reference IQM (also termed "blind") for their suited-ness in presentation attack detection. Complementing recent results [4], we aim at (i) a different and larger set of non-reference IQM (6 instead of 2) and (ii) do not fuse the results with full-reference IQM but focus on blind IQM as a stand-alone technique (eventually also employing a single metric contrasting to [4] where most results given correspond to fusing a considerable amount of IQM also resulting in significant computational effort).

Section 2 introduces and explains the blind IQM as used in this paper. The databases specifically provided to test presentation attack detection techniques for iris, fingerprint, and face recognition used in the present work are described in Section 3. Section 4 presents corresponding experimental anti-spoofing results while Section 5 provides the conclusions of this paper.

II. Non-reference Image Quality Metrics

Current state-of-the-art non-reference Image Quality Assessment (NR IQM) algorithms are based on models that can learn to predict human judgments from databases of human-rated distorted images. These kinds of IQM models are necessarily limited, since they can only assess quality degradations arising from the distortion types that they have been trained on. However, it is also possible to contemplate sub-categories of general-purpose NR IQM models having tighter conditions. A model is said to be opinion-aware (OA) if it has been trained on a database(s) of human rated distorted images and associated subjective opinion scores.
Algorithms like DIIVINE, BIQI, BLINDS-II and BRISQUE are OA IQM measures. However, IQM like NIQE, and BIQAA are opinion-unaware (OU) and they make only use of measurable deviations from statistical regularities observed in natural images without being trained on human-rated distorted images and indeed without any exposure to distorted images.

Systematic comparisons of the NR IQM as used in this paper have been published [6], [7]. Both, in non-trained [6] as well as in specifically trained manner [7] the correspondence to human vision turns out to be highly dependent on the dataset considered and the type of distortion present in the data. Thus, there has been no "winner" identified among the techniques considered with respect to correspondence to subjective human judgement and objective distortion strength.

A. NIQE - Natural Image Quality Evaluator

A NR OU-DU IQM (no reference, opinion unaware & distortion unaware) is based on constructing a collection of

quality aware features and fitting them to a multivariate Gaussian (MVG) model. The quality aware features are derived from a simple, but highly regular natural scene statistic (NSS) model. NIQE [8] only uses the NSS features from a corpus of natural images while BRISQUE is trained on features obtained from both natural and distorted images and also on human judgments of the quality of these images.

The classical spatial NSS model begins with preprocessing: local mean removal and divisive normalisation. Once the new image pixels calculated by the preprocessing have been computed, the image is partitioned into P×P image patches. Specific NSS features are then computed from the coefficients of each patch. Then the sharpness of each patch is determined and only patches with higher sharpness are selected. A simple model of the NSS features computed from natural image patches can be obtained by fitting them with an MVG density.

NIQE is applied by computing the 36 identical NSS features from patches of the size P×P from the image to be quality analysed, fitting them with the MVG model, then comparing its MVG fit to the natural MVG model. The NIQE Index delivers performance comparable to top performing NR IQA models that require training on large databases of human opinions of distorted images.

B. BLIINDS-II - Blind Image Integrity Notator

BLINDS-II [9] uses natural scene statistics models of discrete cosine transform (DCT) coefficients. The algorithm can be divided into four stages. At the first stage the image is subjected to local 2-D DCT coefficient computation. At this point the image is partitioned into equally sized n×n blocks, then computing a local 2-D DCT on each of the blocks. The DCT coefficient extraction is performed locally in accordance with the HVS (Human Visual System) property of local spatial visual processing (i.e., in accordance with the fact that the HVS processes the visual space locally), thus, this DCT decomposition is accomplished across several spatial scales.

The second stage applies a generalised Gaussian density model to each block of DCT coefficients, as well as for specific partitions within each DCT block. In order to capture directional information from the local image patches, the DCT block is partitioned directionally into three oriented subregions. A generalised Gaussian fit is obtained for each of the oriented DCT coefficient subregions. Another configuration for the DCT block partition reflects three radial frequency subbands in the DCT block. The upper, middle and lower partitions correspond to the low-frequency, mid-frequency, and high-frequency DCT subbands, respectively. A generalised Gaussian fit is obtained for each of the radial DCT coefficient subregions as well.

The third step computes functions of the derived generalised Gaussian model parameters. These are the features used to predict image quality scores. The fourth and final stage is a simple Bayesian model that predicts a quality score for the image. Here the training is required. The prediction model is the only element of BLIINDS-II that carries over from BLIINDS-I. The Bayesian approach maximises the probability that the image has a certain quality score given the model-based features extracted from the image. The posterior probability that the image has a certain quality score from the extracted features is modelled as a multidimensional generalised Gaussian density.

C. BIQAA - Blind Image Quality Assessment through Anisotropy

BIQAA [10] is based on measuring the variance of the expected entropy of a given image upon a set of predefined directions. Entropy can be calculated on a local basis by using a spatial/spatial-frequency distribution as an approximation for a probability density function. The generalised Renyi entropy and the normalised pseudo-Wigner distribution (PWD) have been selected for this purpose. As a consequence, a pixel-by-pixel entropy value can be calculated, and therefore entropy histograms can be generated as well. The variance of the expected entropy is measured as a function of the directionality, and it has been taken as an anisotropy indicator. For this purpose, directional selectivity can be attained by using an oriented 1-D PWD implementation. So, the method is based on measuring the averaged anisotropy of the image by means of a pixel-wise directional entropy. BIQAA aims to show that an anisotropy measure can be used to assess both, the fidelity and quality of images.

D. BRISQUE

BRISQUE [11] does not require any transformation to another coordinate frame like DCT used by BLINDS-II. BRISQUE has very low computational complexity, making it well suited for real time applications. The two main stages of BRISQUE are natural scene statistics in the spatial domain and quality evaluation. At the first stage an image is locally normalised (via local mean subtraction and divisive normalisation). Subsequently, 2 parameters are estimated $(\alpha, \sigma 2)$ from a generalised Gaussian distribution (GGD) fit of the normalised pixel data. These form the first set of features that will be used to capture image distortion. To show that pristine and distorted images are well separated in GGD parameter space, a set of pristine images from the Berkeley image segmentation database was taken. Similar kinds of distortions as present in the LIVE image quality database were introduced in each image at varying degrees of severity to form the distorted image set: JPEG 2000, JPEG, white noise, Gaussian blur, and fast fading channel errors. A model for the statistical relationships between neighboring pixels is also built. While normalised coefficients are definitely more homogeneous for pristine images, the signs of adjacent coefficients also exhibit a regular structure, which gets disturbed in the presence of distortion. To model this structure the empirical distributions of pairwise products of neighboring normalised coefficients along four orientations are used.

At the second stage a mapping is learned from feature space to quality scores using a regression module, yielding a measure of image quality. For that purpose a support vector machine (SVM) regressor (SVR) is used. SVMs are popular as classifiers since they perform well in high-dimensional spaces, avoid over-fitting and have good generalisation capabilities. In contrast to algorithms like NIQE and BLINDS-II BRISQUE requires training.

E. DIIVINE - Distortion Identification-based Image Verity and Integrity Evaluation

DIIVINE [12] is based on a 2-stage framework involving distortion identification followed by distortion-specific quality assessment. Most present-day NR IQM algorithms assume that the distorting medium is known - for example, compression, loss induced due to noisy channel etc. Based on this assumption, distortions specific to the medium are modelled and quality is assessed. By far the most popular distorting medium is compression which implies that blockiness and blurriness should be evaluated. DIIVINE targets three common distortion categories, i.e. JPEG compression, JPEG2000 compression, and blur.

In order to extract statistics from distorted images the steerable pyramid decomposition is utilised. The steerable pyramid is an over-complete wavelet transform that allows for increased orientation selectivity. Since NR IQM algorithms are generally trained and tested on various splits of a single dataset (as described above), it is natural to wonder if the trained set of parameters are database specific. However, the training process of DIIVINE is simply a calibration, and once such training is performed, DIIVINE is capable of assessing the quality of any distorted image, since the performance of the algorithm was evaluated on an alternate database. An SVR is used for the classification into the distortion categories.

DIIVINE was actually not developed under the constraint of real-time analysis of images, given that the performance of DIIVINE is as good as leading full-reference quality assessment (FR QA) algorithms.

F. BIQI -Blind Image Quality Index

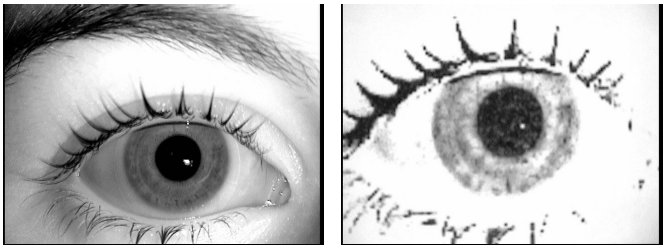
BIQI [13] is also based on a 2-stage framework like DIIVINE. The two steps are image distortion classification based on a measure of how the natural scene statistic (NSS) are modified, followed by quality assessment, using an algorithm specific to the decided distortion. Once trained, an algorithm of the proposed framework does not require further knowledge of the distortion affecting the images to be assessed. The framework is modular in that it can be extended to any number of distortions.

BIQI starts with wavelet transforming an image over three scales and three orientations using the Daubechies 9/7 wavelet basis. The subband coefficients so obtained are parametrised using a GGD. An 18-D vector is formed and it is the representative feature vector for each image.

Given a training and test set of distorted images, a classifier is based on the feature vector to classify the images into five different distortion categories, based on the distortion type JPEG, JPEG2000, WN, Blur, and FF. DIIVINE in contrast only classifies the distortion images into 3 categories. The classifier used is a SVM, which is also utilised in DIIVINE and BRISQUE. BIQI works well for images corrupted by white noise and blur and to some extent for JPEG2000 and FF. However, the performance for JPEG compression is less impressive.

III. Used Spoofing / Presentation Attack Databases

ATVS-FIr DB: The ATVS-Flr database consists of fake and real iris samples of both eyes of 50 subjects and complements the real data of the BioSecure dataset [14]. Four samples of each iris were captured in two acquisition sessions with the LG Iris Access EOU3000. Thus the database holds 800 real image samples (100 irises $\times$ 4 samples $\times$ 2 sessions). The fake samples were also acquired with the LG Iris Access EOU3000 from high quality printed images of the original sample. As the structure is the same as for the real samples, the database comprises 800 fake image samples (100 irises $\times$ 4 samples $\times$ 2 sessions). Fig. 1 displays example images.



(a) Iris; Right eye; Real (b) Iris; Right eye; Fake

Figure 1: ATVS-FIr DB samples

The dataset has been used before in spoofing / presentation attack detection investigations, e.g. [15], [4], [2].

ATVS-FFp DB: The ATVS-FFp database consists of fake and real images taken from a human's index and middle finger of both hands. Those fingerprints can be divided into two categories: *With cooperation (WC)* and *Without cooperation (WOC)*. With cooperation means that acquisition assumes the cooperation of the fingerprint owner, whereas images taken without cooperation are latent fingerprints which had to be lifted from a surface.

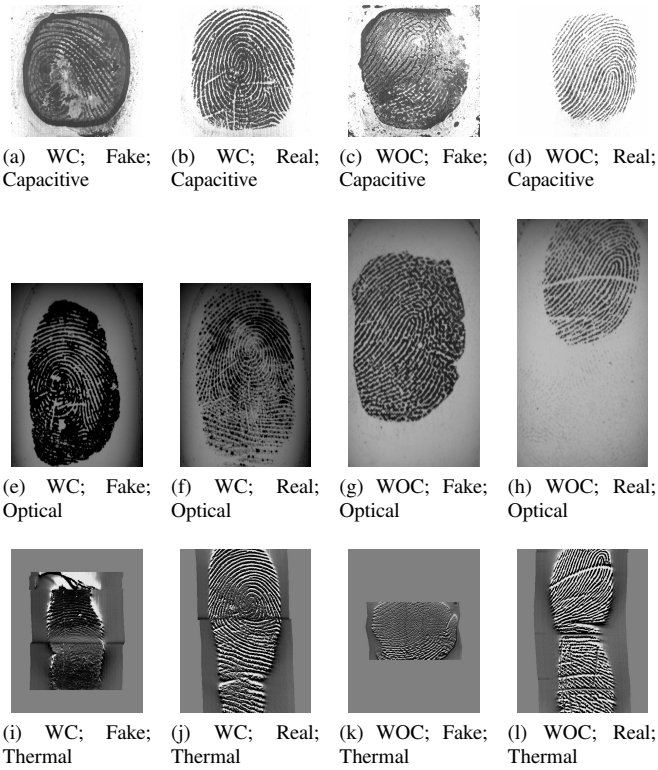
Independent of the category, four samples of each finger were captured in one acquisition session with three different sensors:

- Flat optical sensor Biometrika Fx2000 (512 dpi)
- Sweeping thermal sensor by Yubee with Atmel's Fingerchip (500 dpi)
- Flat capacitive sensor by Precise Biometrics model Precise 100 SC (500 dpi).

As a result the database consists of 816 real/fake images (68 fingers $\times$ 4 samples $\times$ 3 sensors) samples taken with cooperation and 768 real/fake images (64 fingers $\times$ 4 samples $\times$ 3 sensors) samples taken without cooperation. Fig. 2 displays example images from this dataset.

The dataset has been used before in spoofing / presentation attack detection investigations, e.g. in [16], [17], [18].

IDIAP Replay-Attack DB [19]: The Replay-Attack database for face spoofing consists of 1300 video clips of photo and video attack attempts to 50 clients under different lighting conditions. All videos were generated by either having a real



(a) WC; Fake; Capacitive (b) WC; Real; Capacitive (c) WOC; Fake; Capacitive (d) WOC; Real; Capacitive

(e) WC; Fake; Optical (f) WC; Real; Optical (g) WOC; Fake; Optical (h) WOC; Real; Optical

(i) WC; Fake; Thermal (j) WC; Real; Thermal (k) WOC; Fake; Thermal (l) WOC; Real; Thermal

Figure 2: ATVS-FFp DB samples

client trying to access a laptop through its webcam or by displaying a photo/video to the webcam. Real as well as fake videos were taken under two different lighting conditions:

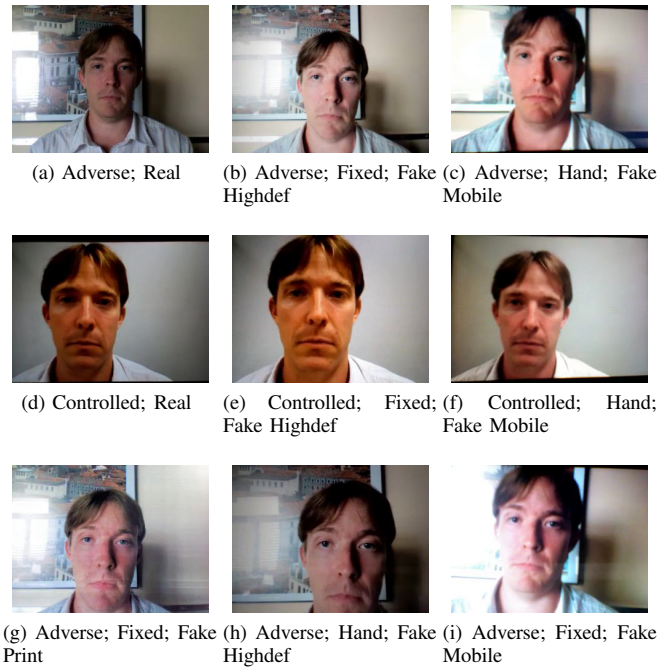
- **controlled**: The office light was turned on, blinds are down, background is homogeneous.
- **adverse**: Blinds up, more complex background, office lights are out.

To produce the attack, high-resolution videos were taken with a Canon PowerShot SX150 IS camera. The way to perform the attacks can be divided into two subsets: the first subset is composed of videos generated using a tripod to present the client biometry ("fixed"). For the second set, the attacker holds the device used for the attack with his/her own hands.

In total, 20 attack videos were registered for each client, 10 for each of the attacking modes just described:

- 4 × mobile attacks using an iPhone 3GS screen (with resolution 480x320 pixels)
- 4 × high-resolution screen attacks using an iPad (first generation, with a screen resolution of 1024x768 pixels)
- 2 × hard-copy print attacks (produced on a Triumph-Adler DCC 2520 color laser printer) occupying the whole available printing surface on A4 paper

As the algorithms used in our experiment are not compatible with videos, we extracted every Xth frame from each video and used them as test data in our experiment. Fig. 3 displays example images used in experimentation.



(a) Adverse; Real (b) Adverse; Fixed; Fake Highdef (c) Adverse; Hand; Fake Mobile

(d) Controlled; Real (e) Controlled; Fixed; Fake Highdef (f) Controlled; Hand; Fake Mobile

(g) Adverse; Fixed; Fake Print (h) Adverse; Hand; Fake Highdef (i) Adverse; Fixed; Fake Mobile

Figure 3: IDIAP Replay-Attack DB samples

The dataset has been used before in spoofing / presentation attack detection investigations, e.g. in [19], [4], [5], [3], [2].

IV. Experiments

A. Experimental Setup

For each image in the databases quality scores were calculated with the IQM described in section II. We used the MATLAB implementations from the developers of BIQI, BLIINDS-2, NIQE, DIVINE, BRISQUE[1] and BIQAA[2]. In all cases, we used the default settings. We normalised the result data with the result that 0 represents a good quality and 100 the bad one which is already the default result in all cases except BIQAA. Originally the data of BIQAA is between 0 and 1. However, the values are so small that we had to define our own limits for the normalisation. A thorough analysis shows that our values are all between 0.00005 and 0.05 therefore we used these figures as our limits. Moreover we had to change the "orientation" of the BIQAA quality scores to be comformable to our definition. Summarising, the following formula (1) was built:

$$x' = 100 - \frac{x - 0.05}{0.00005 - 0.05} \cdot 100 \qquad (1)$$

In the first experimental stage we consider the distribution of the quality scores only. Our aim was to eventually find a

[1]All available from http://live.ece.utexas.edu/research/quality/

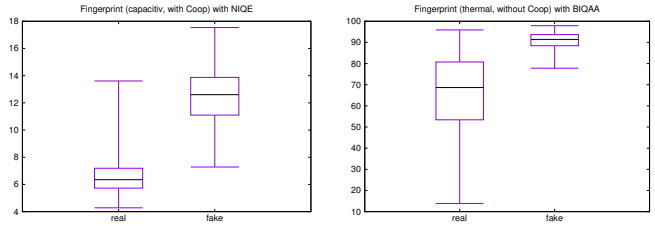
[2]Available at https://www.mathworks.com/matlabcentral/fileexchange/30800-blind-image-quality-assessment-through-anisotropy

threshold between the values of the real data and the fake ones for the various IQM.

Afterwards, in the second stage, we used the quality scores for a leave-one-out cross validation to get an exact assertion about the classification possibility with NR IQM. To classify our data we used k-nearest neighbours (kNN) classification. Our used k were 1, 3, 5, 7 and 9 for this experiment. First, we only used one quality score for the classification. In the next step, we combined several quality scores of the different measures into one vector and used this for the kNN-classification. This method allowed us to test all possible combinations of IQM in a simple way. The distance for the kNN-classification was in the first case the difference between the two values and in the second case the distance between the two vectors. At the end, we got the classification accuracy for discriminating real from fake images for all IQM combinations.

B. *Experimental Results*

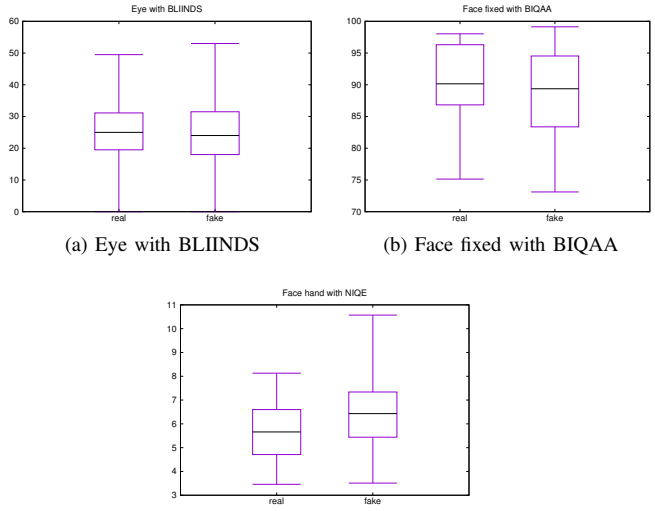
In Fig. 4, we display the distribution of IQM values for real and fake data. For some cases, we notice a decent separation of the values almost allowing to specify a separation threshold. However, this is not possible for most configurations. In many cases (see e.g. Fig. 5) we could not recognise any differences between the distributions because they exhibited the same spread for real and the fake data. That was the reason for employing kNN-Classification.



(a) Fingerprint (capacitiv, with coop) with NIQE (b) Fingerprint (thermal, without coop) with BIQAA

Figure 4: Quality score distribution (positive examples)

In the case of kNN-classification with only one IQM, we already obtain surprisingly good results. In table I we can see that we got over 99% classification accuracy for a fingerprint database (thermal, with coop). In this case we already could see the differences of the distributions of the real and the fake values of the quality scores (see Fig. 4b). For this reason, a high accuracy with kNN-classification was already expected. Nevertheless, we are above 80% overall classification accuracy for all but a single database. Except for BIQAA, all measures are present in the table, thus, we are not able to identify a single IQM specifically well suited for the target task. In contrast, it seems that the different distortions present in the spoofed data are quite specific in terms of the nature and characteristic of the distortions, which is the only explanation of different IQM performing best on different datasets. In fact, our results confirm the general results on IQM quality prediction performance [6], [7] in that it is highly dataset and distortion dependent which IQM provides the best results.



(a) Eye with BLIINDS (b) Face fixed with BIQAA

(c) Face hand with NIQE

Figure 5: Quality score distribution (negative examples)

Table I: Best results for kNN-classification with only one IQM

Database	Algorithm	k	Accuracy
Eye	BRISQUE	7	76.44%
Fingerprint (optical, with coop)	BIQI	9	68.20%
Fingerprint (capacitive, with coop)	NIQE	9	93.38%
Fingerprint (thermal, with coop)	BIQI	5	99.08%
Fingerprint (optical, without coop)	DIVINE	7	82.03%
Fingerprint (capacitive, without coop)	NIQE	9	84.38%
Fingerprint (thermal, without coop)	BIQI	5	94.92%
Face (hand)	BLIINDS-2	5	84.76%
Face (fixed)	BLIINDS-2	9	80.00%

A further increase in classification accuracy was obtained by the combination of several IQM. Table II shows the best combinations for the considered databases from an exhaustive search. On average, we could improve our results by 7% compared to the single measure results and so most of the results are over 90%.

From the latter table we notice that there is a trend of getting best results when combining a larger number of IQM, confirming earlier results in this direction [4]. In order to look into this effect more thoroughly (and to clarify the role of the k-parameter in kNN-classification) we have systematically plotted the results of the exhaustive classification scenarios.

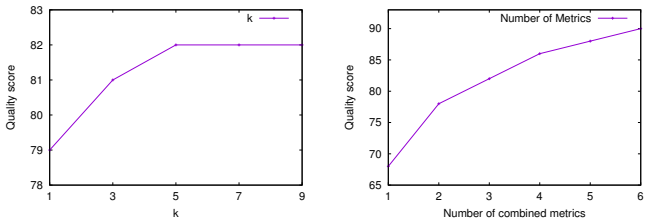
We average all classification results by keeping the number of combined metrics fixed (Figure 6b) and and by keeping the parameter k fixed (Figure 6a). Combining more metrics and choosing k large leads to better results on average, whereas the top results as shown in table II are achieved when using 3 – 6 metrics depending on the considered dataset. In this table, k is also found to be 1 for two datasets.

V. Conclusion

We have found a high dependency on the actual dataset under investigation when trying to answer the question about the optimal choice of an image quality measure. All but

Table II: Best IQM combinations

Database	Combination	k	Accuracy
Eye	BIQI, BLIINDS, NIQE, DIVINE, BRISQUE, BIQAA	9	85.81%
Fingerprint (optical, with coop)	BIQI, BLIINDS, DIVINE, BIQAA	7	81.25%
Fingerprint (capacitive, with coop)	BIQI, BLIINDS, NIQE, DIVINE, BRISQUE, BIQAA	3	96.69%
Fingerprint (thermal, with coop)	BIQI, BLIINDS, (NIQE), DIVINE, BRISQUE	1	99.63%
Fingerprint (optical, without coop)	BIQI, BLIINDS, (NIQE), DIVINE, BRISQUE	7	87.69%
Fingerprint (capacitive, without coop)	BLIINDS, NIQE, BIQAA	5	92.19%
Fingerprint (thermal, without coop)	BIQI, BLIINDS, NIQE, BRISQUE, BIQAA	1	98.44%
Face (hand)	BLIINDS, (NIQE), DIVINE, BRISQUE, BIQAA	7	92.86%
Face (fixed)	BIQI, BLIINDS, NIQE, DIVINE, BRISQUE, BIQAA	9	92.38%



(a) Dependent on k (b) Dependent on number of metrics

Figure 6: Average quality score

BIQAA are listed at least once as being the best option for a specific dataset. While BIQAA is not seen among the best performing IQM (and thus seems to be a candidate for the worst-performing IQM), it is found in several IQM fusion settings and thus obviously complements the other IQM (which are all somehow based on NSS) in some way. Therefore, we are not able to identify a clear "winner" or "looser" among the IQM based on the results analysed. The same is true when it comes to classifier settings, which are also rather dependent on the dataset. Overall, there is a trend that more IQM being combined lead to better classification accuracy.

Since the optimal choice of IQM is so dependent on the dataset, it is probably also the nature of attack type that plays a certain role (e.g. if the attack is based on replayed data or if actual artefacts are being used). Thus, the generalisation of the results to unseen attack types might be not straightforward.

ACKNOWLEDGMENT

This work has been partially supported by the Austrian Science Fund, project no. 27776, and by the ICT COST Action IC1206 "De-identification for privacy protection in multimedia content".

REFERENCES

[1] S. Marcel, M. Nixon, and S. L. (Eds.), *Handbook of Biometric Anti-Spoofing*. Springer, 2014.

[2] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "An investigation of local descriptors for biometric spoofing detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 849–861, 2015.

[3] D. Menotti, G. Chiachia, A. Pinto, W. Schwartz, H. Pedrini, A. X. Falcao, and A. Rocha, "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 864–879, 2015.

[4] J. Galbally, S. Marcel, and J. Fierrez, "Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 710–724, 2014.

[5] D. Wen, H. Han, and A. Jain, "Face spoof detection with image distortion analysis," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 746–761, 2015.

[6] A. Nouri, C. Charrier, A. Saadane, and C. Fernandez-Maloigne, "Statistical comparison of no-reference images quality assessment algorithms," in *Proceedings of the Colour and Visual Computing Symposium (CVCS'13)*, 2013.

[7] C. Charrier, A. Saadane, and C. Fernandez-Maloigne, "Comparison of no-reference image quality assessment machine learning-based algorithms on compressed images," in *Image Quality and System Performance XII*, ser. Prooceedings of SPIE, vol. 9396, 2015.

[8] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making image quality assessment robust," in *Proceesings of the 46th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 2012.

[9] M. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339–3352, 2012.

[10] S. Gabarda and G. Cristobal, "Blind image quality assessment through anisotropy," *J. Opt. Soc. Am. A*, vol. 24, December 2007.

[11] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, 2012.

[12] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: from natural scene statistics to perceptual quality," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3350–3364, 2011.

[13] ——, "A two-step framework for constructing blind image quality indices," *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 513–516, May 2010.

[14] J. Fierrez, J. Ortega-Garcia, D. Torre-Toledano, and J. Gonzalez-Rodriguez, "BioSec baseline corpus: A multimodal biometric database," *Pattern Recognition*, vol. 40, no. 4, pp. 1389–1392, April 2007.

[15] J. Galbally, J. Ortiz-Lopez, J. Fierrez, and J. Ortega-Garcia, "Iris liveness detection based on quality related features," in *Proceedings of the IAPR/IEEE International Conference on Biometrics (ICB'12)*, March 2012, pp. 271–276.

[16] J. Galbally, F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia, "A high performance fingerprint liveness detection method based on quality related features," *Future Generation Computer Systems*, vol. 28, pp. 311–321, January 2012.

[17] I. Bhardwaj, N. Londhe, and S. Kopparau, "A spoof resistant multibiometric system based on the physiological and behavioral characteristics of fingerprint," *Pattern Recognition*, vol. 62, pp. 214–224, 2017.

[18] M. Lu, Z. Chen, and W. Sheng, "Fingerprint liveness detection based on pore analysis," in *Biometric Recognition – Proceedings of the Chinese Conference on Biometric Recognition (CCBR'15*, ser. Springer LNCS, vol. 9428, 2015, pp. 233–240.

[19] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG'16)*, September 2012.