Self-Organizing Maps for Fingerprint Image Quality Assessment

Martin Aastrup Olsen[1,2]   Elham Tabassi[2]   Anton Makarov[3]   Christoph Busch[1]

[1]Gjøvik University College, [2]National Institute of Standards and Technology,
[3]Technical University of Denmark

Abstract

Fingerprint quality assessment is a crucial task which needs to be conducted accurately in various phases in the biometric enrolment and recognition processes. Neglecting quality measurement will adversely impact accuracy and efficiency of biometric recognition systems (e.g. verification and identification of individuals). Measuring and reporting quality allows processing enhancements to increase probability of detection and track accuracy while decreasing probability of false alarms. Aside from predictive capabilities with respect to the recognition performance, another important design criteria for a quality assessment algorithm is to meet the low computational complexity requirement of mobile platforms used in national biometric systems, by military and police forces. We propose a computationally efficient means of predicting biometric performance based on a combination of unsupervised and supervised machine learning techniques. We train a self-organizing map (SOM) to cluster blocks of fingerprint images based on their spatial information content. The output of the SOM is a high-level representation of the finger image, which forms the input to a Random Forest trained to learn the relationship between the SOM output and biometric performance. The quantitative evaluation performed demonstrates that our proposed quality assessment algorithm is a reasonable predictor of performance. The open source code of our algorithm will be posted at NIST NFIQ 2.0 website.

1. Introduction

Recognition performance is increased if fingerprints are of sufficiently good quality and when overall database integrity is improved. Therefore, the ability to automatically examine a biometric image and produce a numerical estimate of the utility of the image to a downstream comparison algorithm (*i.e.* matcher) is an operationally desirable function. Particularly, if a low quality value is predictive of recognition failure (primarily a false negative, but possibly a false positive too), then a new sample can be collected while the subject is still present. Without a method to determine whether the quality of a captured sample is sufficient for recognition purposes, an individual can be falsely rejected or actively subvert the system. As such, many large scale biometric deployments mandate measuring and reporting of quality scores. Examples are the Unique Identification Authority of India (UIDAI) [18], the European Union Visa Information System (VIS) [7, 6, 5], or the United States Visitor and Immigrant Status Indicator Technology (US-VISIT). While the ability to inspect low utility images is essential for both the server and client side of these applications, the client side demands low computational cost, both in terms of time and processing power. In fact, requirements for next generation quality assessment includes near frame rate throughput [14].

The major contribution of this paper is a novel method for quality assessment which has a very low computation cost; essentially after training the quality assessment is reduce to a look up table. Our proposed method is based on a combination of unsupervised (self-organizing map [12]) and supervised (Random Forest [2]) machine learning algorithm. A self-organizing map (SOM) is utilized as a receptive field to learn the spatial information content of a fingerprint image creating a high level representation in the form of class labels. Transforming an input from image space to class labels reduces the dimensionality by several orders of magnitude while retaining key characteristics of the finger image. This new representation of the finger image becomes input for the Random Forest which is to give a quality score that is predictive of biometric performance. To the authors' knowledge this is the first application of using a receptive field approach to fingerprint quality assessment.

The paper is organized as follows: Section 2 provides a quick overview of related work in fingerprint quality assessment. Section 3 details our proposed method. Discussion of data, experiments and results are given in Section 4, followed by conclusions and future work in Section 5.

2. Background and Related Work

Several fingerprint quality assessment algorithms have been documented in the literature, a comparative study of

them is given in [1]. The majority of them are based on processing an image in either the spatial or frequency domain to compute particular features such as orientation strength. Subsequently a scalar quality value is computed from a feature or a combination of features. The technical report on fingerprint sample quality [11] by the International Organization for Standardization has a recent overview of recommended fingerprint quality features; *e.g.* Orientation Certainty Level which measures the strength of the energy concentration along the dominant ridge flow orientation within a block by computing the blockwise gradient; Ridge-Valley Structure which computes the blockwise clarity of ridges and valleys by applying linear regression to determine a gray-level threshold and then classify pixels as ridge or valley. A ratio of misclassified pixels is determined by comparing the normalized ridge and valley thickness of that block. Orientation Flow proposed in [3] is a measure of the rate of change in the blockwise ridge flow across the fingerprint. The quality score decreases as the difference between the dominant ridge orientation of the block and its 8 neighboring blocks increases. A block wise Gabor based quality feature was proposed in [16] and a point wise Gabor based quality feature was proposed in [15]. One widely used open source quality assessment algorithm which combines several of the above features is NFIQ [17]. It is based on a neural network trained on an 11-dimensional feature vector derived from finger image characteristics which include minutia count, ridge curvature and local contrast.

Computation of these features happens to be expensive, first a processing unit is required for their computations, and secondly according to our in-house implementations, it takes more than 130 millisecond to compute any of the features. In this paper, we aim to address this shortcoming of the current methods. Once trained, our proposed method is essentially reduced to look-up table operations, which reduces the computation time by about half, and make it suitable for mobile platforms with minimum computational resources.

3. Methodology

This section details how we employed the SOM and Random Forest clustering for calculation of fingerprint image quality. Fig. 1 shows the overview of our proposed method. We give a review of the SOM concept in Section 3.1, and its application to our proposed method in Section 3.2. Section 3.3 describes how we used Random Forest to classify the output of the SOM into five different quality categories.

3.1. Overview of Self-organizing Maps

The Kohonen Map, introduced in [12], is a self-organizing map (SOM) which is trained without supervision. Its objective is to organize the input patterns to a topological structure where the relations between different patterns are preserved. The SOM is a two-layered network consisting of

a number of processing units. The first layer of units can be considered as a group of sensors picking up the signal. It is fully connected to a second, two-dimensional layer: the competitive layer. Fig. 2 shows the topology of the network. The weights associated with the connections are adjusted during training and only one single unit in the competitive layer is active at a time. Due to the training rules explained below, the spatial distance of two units reacting on different input patterns is a measure for the similarity of the two patterns. The training of the network is carried out by presenting data vectors $\mathbf{x}$ to the input layer of the network whose connection weight vectors $\mathbf{m}_i$ of all competitive units i are first initialized by random values. If N is the dimension of the data, we choose N input units of the first layer and define an Euclidean distance d_i between $\mathbf{x}$ and $\mathbf{m}_i$ with

$$d_i = \| \mathbf{x} - \mathbf{m}_i \| = \sqrt{\sum_{j=0}^{N-1}(x_j - m_{ij})^2} \qquad (1)$$

The input signal activates winner unit c whose weight vector has the minimum distance to the given signal. The winning unit is also known as the best matching unit (BMU).

$$d_c = min(d_i) \qquad (2)$$

The updating of the weights m_{ij} associated to the units is only performed for units i within a proximity $d(t)$ defining the time-dependent neighborhood $N_c(t)$ of the winning c. This proximity $d(t)$ is reduced as training of the network is progressing with time t (Eq. (6)). Fig. 2 illustrates the time-dependent neighborhood. The update of weights follows Eq. (3), where $a(t)$ represents a time-dependent learning rate.

$$m_{ij}^{(t+1)} = m_{ij}^{(t)} + \Delta m_{ij}^{(t)} \qquad (3)$$

$$\Delta m_{ij}^{(t)} = \begin{cases} a(t)(x_j - m_{ij}^{(t)}), & \text{if unit } i \in N_c(t) \\ 0, & otherwise \end{cases} \qquad (4)$$

$$a(t) = a_0(1 - \frac{1}{T}), t \in [0, \ldots, T-1] \qquad (5)$$

$$d(t) = d_0(1 - \frac{t}{T}), t \in [0, \ldots, T-1] \qquad (6)$$

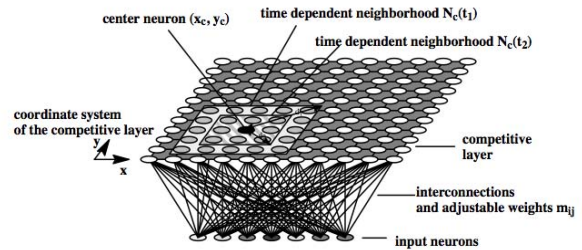Figure 2. Topology of the self-organizing map and the time dependent neighborhood $N_c(t)$ of a unit c.

Binned comparison scores

Training set / validation set

Train self-organizing map

Train random forest

SOM Codebook

Activation histogram

Random forest

Test set

Activation histogram
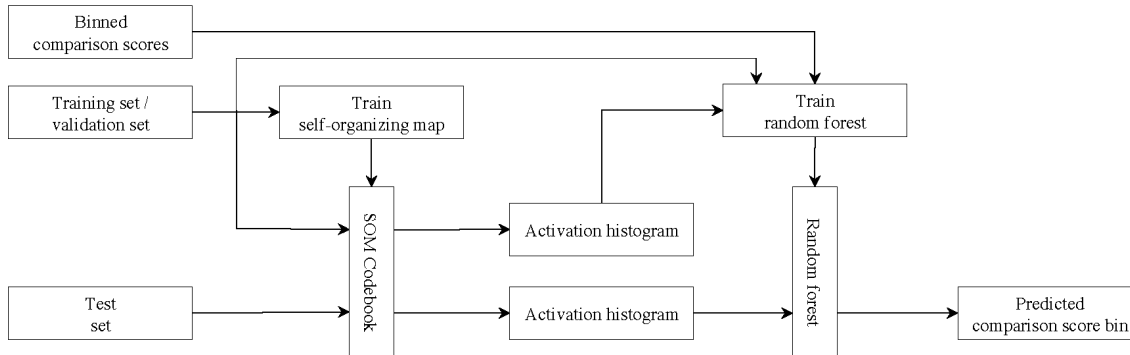
Predicted comparison score bin

Figure 1. Overview of system architecture and components. The training and validation sets and their binned associated comparison scores are utilized to train the self-organizing map and the Random Forest. Training the SOM creates a codebook which is used to create an activation histogram of the training and validation sets. The histogram of activations from the training set is used as the explanatory variables in the Random Forest, while the binned comparison score is the target variable. Once trained, the Random Forest produces a model which is verified using the validation set. The test set, containing the unseen data, is used to determine the accuracy of the pipeline.

The SOM offers two properties during training that are strongly related to our challenge namely a) an approximation of the presented data by mean vectors $\mathbf{m_i}$ that are associated as weights to the units thus essentially realizing a cluster analysis of the data and b) a topological ordering of the competitive units in such a way that neighboring units in the layer represent similar clusters in multidimensional space and thus a dimensionality reduction. This can be interpreted as a nonlinear, topology preserving, associative mapping process. Generally, the Kohonen Map refers to the well-known K-means clustering [4]. As the units in the competitive layer are ordered topologically, *i.e.* neighboring units react to data vectors similar in the input space, the mapping can be interpreted as a reduction of any N-dimensional space into two dimensions preserving the topology of the data as much as possible. The resolution of this discrete 2D-space is given by the number of competitive units, *i.e.* clusters, which should be configured such that the map serves the requirements with regard to grading (granularity) of the quality score. The limited use case of the SOM for clustering, topological organization and for subspace mapping has been described. Due to the fixed number of units in the SOM and its topology preserving property, some units may have very similar mean vectors $\mathbf{m_i}$. With a trained SOM we then assign a meaning or class to each unit, such that its activation can be related to biometric recognition performance. Our SOM input is blocks of the fingerprint image, and therefore we do not know the contribution of each individual block towards the resulting comparison score. Furthermore, comparison scores are pairwise while quality scores only relate to a single impression. Thus we cannot know if a low comparison score is due to both impressions being of low quality or only one of them. We uncover the relation between each unit in the SOM and biometric performance through another layer of machine learning. In particular we utilize Random Forests to determine

the relation between a certain activation pattern and a given comparison score (see Section 3.3). During the organization process the goal is to find a limited set of centroids representing the data in a C-means sense. As a result the centroids represent the variety of feature vectors in an optimal way leading to a minimum of the error function, which can be easily calculated by the Euclidean distance between all feature vectors and their nearest codebook vector in the map. This search does not need to cover the entire space of codebook vectors since efficient approximations can exclude the vast majority of codebook vectors from consideration [19]. The unit which is closest to the input vector is known as the best matching unit (BMU).

3.2. Application of Self-organizing Maps

First, to provide a fixed length input vector for the SOM the finger images are subdivided without padding or overlap into blocks of size $n \times n$ pixels after intensity normalization. One-dimensional vectors of length n^2 are constructed from the pixel intensities in each block by concatenating the rows. For the training a SOM of size $dim \times dim$ is initialized with randomly generated weight vectors. During the training phase each training input vector (*i.e.* the pixel intensities of a block) is presented to the network and modification of network weights is performed. After T iterations (assuming T is large enough) the SOM clusters the training input vectors into $dim \times dim$ units representing different spatial characteristics of fingerprints. The trained SOM produces a codebook with $dim \times dim$ entries each of length n^2 where each codebook entry represents a cluster of input training vectors. Due to the update function (Eq. (3)) the codebook vector (or entry) associated with each unit will stabilize such that nearby units will have similar codebook vectors. Figure 3 shows the codebook vectors of a SOM of size 24 by 24 ($dim = 24$) trained using input vectors of length 576 ($n = 24$) for $T = 1 \times 10^6$ iterations.
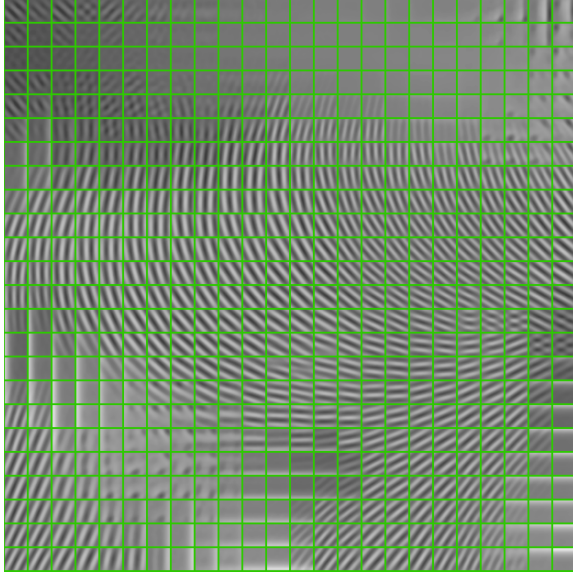
Figure 3. Visualization of the codebook for a SOM of size 24 by 24 ($dim = 24$) trained with 1.7×10^7 input vectors of length 576 ($n = 24$) for $T = 1 \times 10^6$ iterations shown in a grid structure. Unit 0 is located in upper left corner and unit 575 is located in the lower right corner.
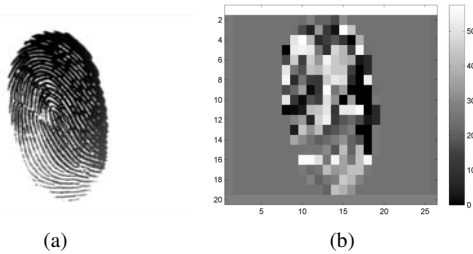

(a)   (b)

Figure 4. Illustration of relationship between finger image and SOM codebook shown in Fig. 3. 4(a) is the input images which is divided into blocks of size 24×24. 4(b) shows the best-matching unit for each image block where black is unit 0 and white is unit 575. The image is FVC 2004's image id 102 [13].

It is clear that the SOM has identified clusters representing various ridge orientations, ridge strengths and background. Variation in image intensity in the codebook is absent or limited due to the image normalization performed as explained in Section 4.1. The trained SOM assigns an index of the corresponding best-matching unit to each block of an input image. Fig. 4 shows the unit activation patterns for a finger image. Note that the white background appearing in the input images shown in Fig. 4(a) activates the same units in the SOM (shown in Fig. 4(b)) indicating that the background has a cluster of its own. To summarize, a trained SOM transforms a fingerprint image from its spatial domain into a vector of the SOM unit activations. The length of this vector varies for different images because the number of blocks depend on the image size. To get a fixed-length

representation we compute the histogram of unit activations normalized by the block count per image. Each image block in a finger image has exactly one BMU and thus we can form a histogram of unit activations. This is motivated by the expectation that the activation patterns for high quality and low quality images are different. Therefore, an image of size $U \times V$ is represented by a vector of length $dim \times dim$ with each component being the frequency of occurrence of the codebook entries.

To determine how the SOM behaves when presented with low quality (*i.e.* difficult to match) and high quality (*i.e.* easy to match) fingerprint images we select 64 images that result in false non-matches at the threshold corresponding to false match rate of 1×10^{-4} (this set is denoted Low) and 64 images involved in comparisons giving the highest possible comparison score (this set is denoted $High$). Note that the threshold was computed over 120 000 impostor comparisons. Fig. 5 shows 10 randomly chosen examples of activation histograms from each of the Low and $High$ sets. The SOM is of size 32×32 ($dim = 32$) and the network was trained for $T = 1 \times 10^6$ iterations using image blocks of 24×24 pixels ($n = 24$). Visual inspection reveals that the histograms of images in the low and $high$ sets look different. Fig. 6 shows the sum of SOM unit activations of the 64 images in the $High$ and Low sets respectively. The difference of these two aggregated histogram ($Low - High$) is depicted in Fig. 7. Bars that extend below zero represent units with stronger activation in the $High$ set than the Low. Similarly bars that extend above zero represents units with stronger activation in the Low set. Since the difference between the two sets is non-zero we conclude that the histograms of the two sets are substantially different. In other words, we have shown that the activation histogram of images involved in a false reject are significantly different from that of those which are highly likely to be correctly recognized. Therefore we used this activation histogram as our feature vector for a supervised learning to predict biometric performance.

3.3. Classification using Random Forest

We chose bins of genuine comparison scores as the response variable and tried several classifiers: K-Nearest Neighbors, Support Vector Machine, Recursive Partitioning tree, and Random Forest. Due to the 8-page limit, we only present results of Random Forests which performed better than the others. Random Forests [2] basically grows many classification trees. To classify a new object from an input vector, it puts the input vector down each of the trees in the forest. Each tree gives a classification, in other words the tree votes for that class. The forest chooses the classification having the most votes. Random Forest has the advantages of being fast, capable of handling large datasets, and that it does not over fit. We investigate network depen-
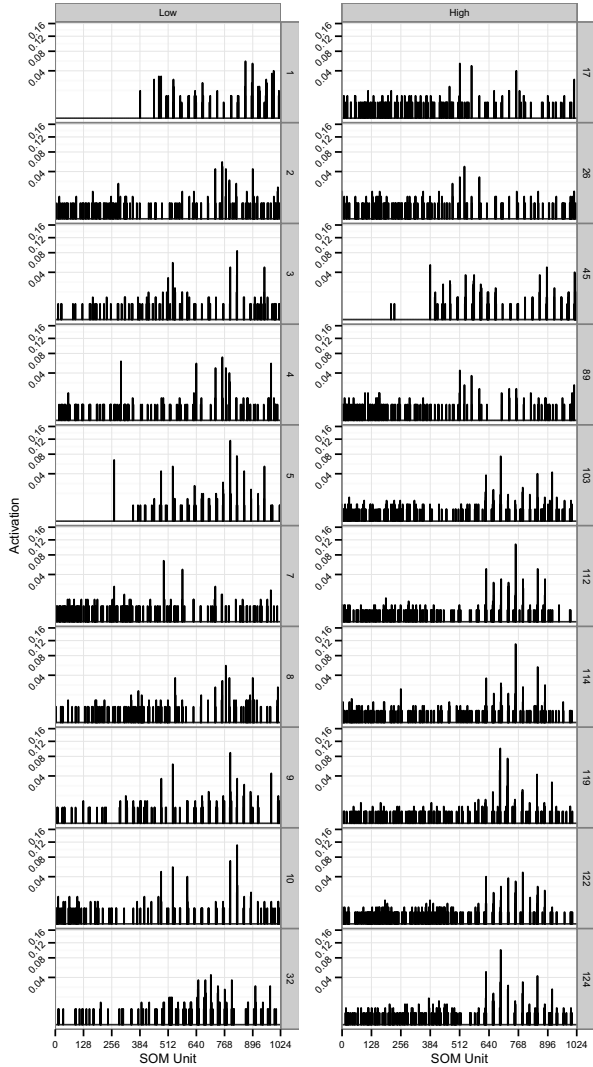
Figure 5. On the left, histogram of activations for 10 images from the set of images which occurred in comparisons resulting in false non-matches at FMR = 1×10^{-4} (set *High*). On the right, histogram of activations for 10 images from the set of images which occurred in comparisons with the highest possible comparison score (set *Low*). The numbers on the right side of each panel (shaded in gray) indicate the image anonymized identification key.

dence on block size, network size and number of iterations by varying n and dim.

Table 1 summarizes the accuracy of the Random Forest for the different parameters. It shows smaller block size gives better performance. Increasing the size of the network does not seem to have a significant effect, which might be due to lack of sufficient training data. Theoretically, the size of the map is only limited by the memory capacity of the computing unit and availability of sufficient number of training imagery. A larger SOM network is more capable of representing all possible spatial information of its training data, but it requires a larger number of images for training.
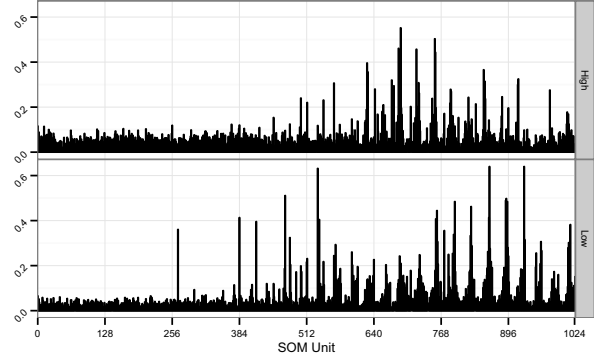

Figure 6. Aggregated histograms of the *Low* and *High* sets. The upper figure shows the aggregate histogram of 64 images that gave the highest possible comparison score (easy to match and therefore high quality). The lower figure shows the aggregate histogram of 64 images that produces false non-match at threshold corresponding to false match rate of 1×10^{-4}.
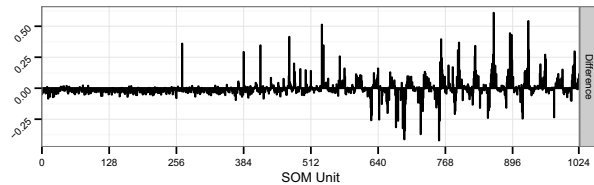

Figure 7. Difference between aggregated histograms of the *Low* and *High* sets. Negative values indicate that the unit is more active in images that gave the highest possible comparison score (easy to match and therefore high quality), while positive values indicate that the unit is more active in images that gave a false reject (therefore hard to match or low quality).

4. Experiments and Results

4.1. Data set

Data we used comprised of 7976 right index finger images of 3988 subjects (2 images per subject) selected from a large operational repository. We took care to construct a balanced training set (1000 pairs) with a representation of real world difficult to match samples to the extend possible. Comparison scores are generated using one of the leading commercial comparison algorithms. We want the training data to be a balanced representation of all possible (or to the extend possible) fingerprint impairments, as well as pristine images. We found that the current available public datasets were not suitable for our work. This is mainly due to small size of these datasets (a few hundred users at the most) and also absence of a reasonably wide range of quality problems in these collections, which is expected and understandable given the careful and at times controlled design of these collections. To ensure inclusion of a wide range of quality in our dataset, we formed a balanced representation from all possible comparison score values generated by the biometric comparator used in our study. Given that different levels of fingerprint quality result in different

values of comparison scores, a set of images representing all possible comparison scores, would represent the obtainable diversity in the quality. Specifically, we chose images from a large operational repository using comparison scores of a leading commercial fingerprint comparison provider. First, we quantized the raw comparison scores by dividing their range into 100 bins. Then for each bin of comparison scores, we randomly selected 10 pairs of images with the corresponding quantized comparison score. This gave 2000 fingerprint images (10×100 comparisons, 2 images per comparison) from which we used 1800 for training and 200 for validation. Our test set consists of images from 2988 comparisons (5976 fingerprint images) selected at random from the initial corpus excluding images occurring in any of the 1000 comparisons selected for the training or validation sets. We further grouped the quantized comparison scores (which are in [1-100]) into 5 levels such that quantized scores in [1, 20] belongs to bin 1, [21, 40] belongs to bin 2 and so on. These 5 levels quantify the biometric performance for each pair of images and are used as the target (response) values for a Random Forest whose explanatory variables are the SOM activation histogram. Images are normalized to give equal intensity according to the pyramid based enhancement method from [8]. An example of the enhancement process is shown in Fig. 8.

(a)   (b)

Figure 8. Example of normalization process using image from FVC 2004 [13]. (a) is the input image and (b) is the enhanced version.

4.2. Performance measurements

This section documents methods and results for the quantitative evaluation of our proposed quality assessment approach. Evaluations are done by quantifying the association between quality scores and the observed recognition performance. We used three widely accepted and documented methods to assess the goodness of our proposed quality algorithm in predicting performance. We examined the relationship of quality scores with a) genuine comparison scores, b) recognition error rate (FNMR) for each level of quality, and c) improvement of FNMR as poor quality samples are removed. Per the definition of biometric quality [10], quality values q should be a monotonic increasing function of performance, where high-quality samples give high similarity scores when involved in genuine comparisons. Figure 9 shows the relationship of our proposed quality scores with biometric comparison scores, we gen-

	$dim = 24$	$dim = 32$	$dim = 40$
$n = 24$	26.38%	26.76%	26.23%
$n = 32$	25.85%	24.32%	26.34%
$n = 64$	23.18%	22.73%	24.18%

Table 1. Accuracy of prediction of comparison score bin for various parameter choices for the SOM dimensions (dim) and block sizes (n). The percentage reflects the accuracy of the predicted bin matching the actual bin as determined from the comparison score.

erated heatmaps (*i.e.* surface plots) of the three variables q_{verify}, q_{enrol} and *score* of the images in our test sets. The x-axis represents quality of the enrolment samples. Quality of verification samples are on the y-axis. The color in each cell represents the magnitude of genuine comparison (or similarity) scores. If quality is an indicator of performance, higher similarity scores (shown in light gray) will be achieved for high-quality images of the same finger. The heatmap in the lower right corner shows the magnitude of samples in each of the combinations of *score* for q_{verify}, q_{enrol}. Of the block sizes investigated, block size 24 gives the best performance; images with higher quality score (4 or 5) result in the highest similarity scores (visualized by light gray values in the upper right corner) and images with low quality scores (1 or 2) result in the lowest similarity score (visualized by black color) in the lower left corner. Block size 64 does not perform well.

A quality assessment algorithm is useful if it can at least give an ordered indication of an eventual performance. This is the case for our proposed method as shown in Figure 10. Figure 10 shows the FNMR for each $L = 5$ levels of our proposed quality method. The highest FNMR is observed for the lowest quality samples ($q = 1$), and lowest FNMR is observed for the highest quality samples ($q = 5$). Biometric comparison involves at least two samples and the challenge is then to relate performance (which involves two samples) to quality values q_{verify} and q_{enrol}. We simplify the analysis by combining the two qualities according to $q_i = M(q_{verify}, q_{enrol})$. It is usually the case that operationally a quality assessment algorithm can be used to assess whether an enrolment sample is of high quality. The enrolment sample will be compared later with a sample that typically is of less controlled quality. We used $M(x, y) = \min(x, y)$ which captures the concept that the worse of two samples drives the similarity score. Figure 10 shows that our proposed method provides a rank order indication of FNMR: the highest error rate is observed for the lowest quality sample ($q = 1$) and the lowest error rate for the highest quality ($q = 5$). As such, our proposed method is successful in predicting performance.

Another metric for comparative evaluation of quality assessment algorithms is the error versus reject curves [9]. The goal is to demonstrate how efficiently rejection of low-quality samples results in improved performance. This
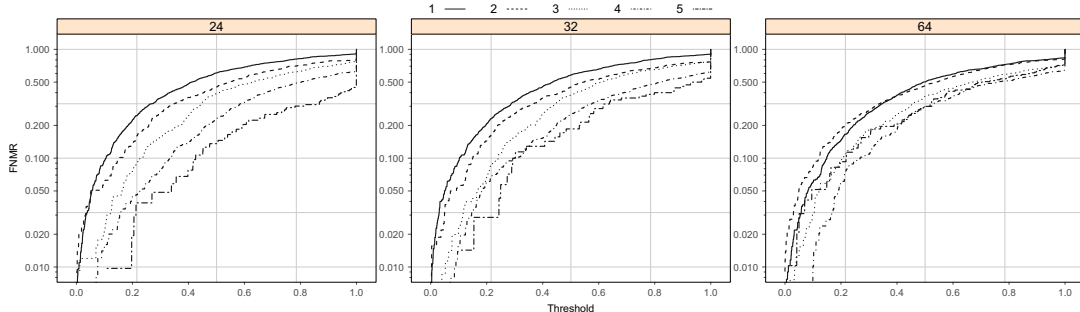
Figure 10. FNMR vs score threshold for each level of quality for different block sizes. The SOM trained on block sizes 24 and 32 performs better than block size 64, as is indicated by a better separation of FNMR for the five levels of quality.
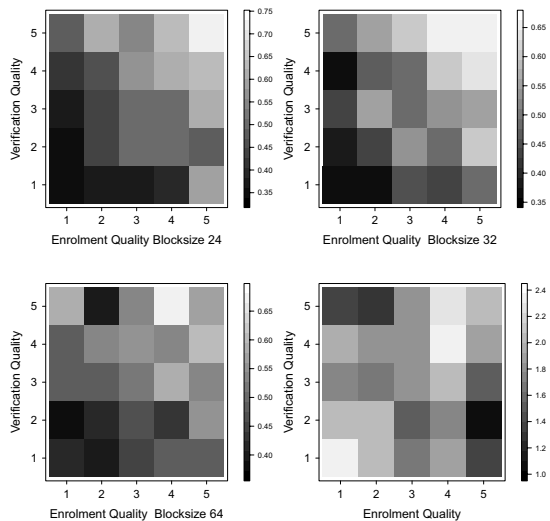

Figure 9. Genuine comparison scores vs (q_{verif}, q_{enrol}) for the 3 different block sizes. The proper behavior is to have low-quality scores for low genuine scores, as is the case for block sizes 24 and 32. The figure in bottom right shows the count of images involved in each cell.
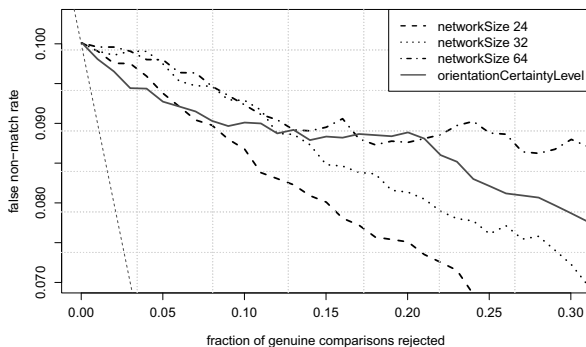

Figure 11. FNMR vs reject curves for our proposed method. The threshold is set to give an initial FNMR = 0.1. The gray dotted line shows the ideal case where the rejection of the ten percent lowest quality results in zero FNMR. Blocksize = 24 perform the best as it is evident by the sharpest decrease in FNMR as poor quality samples are rejected. An almost flat curve suggests that the quality algorithm is not effective in predicting the performance. The solid gray line shows the rejection curve for Orientation Certainty Level metric. Our proposed method outperforms Orientation Certainty Level for rejection rates higher than 5%.

models the operational case in which quality is maintained by reacquisition after a low-quality sample is detected. Consider that a pair of samples (from the same subject) with qualities q_{verif} and q_{enrol} are compared to produce a genuine score, and this is repeated for N such pairs. If the quality values are perfectly correlated with the genuine comparison scores, setting threshold τ to give an overall FNMR of x and then rejecting x percent with the lowest qualities should result in FNMR of zero after recomputing FNMR. Figure 11 shows the error vs. reject curves of our proposed method as applied to different block, namely 24, 32, and 64. We set the value of τ to give a false non-match rate of ten percent. Pairwise quality is computed using the geometric mean of the quality scores of the two images being compared. Similar results were obtained using $\min(x, y)$. Behavior of a perfect quality assessment algorithm is displayed by the gray dotted line where the rejection of the lowest 10% quality would result in an FNMR of zero. The most operationally

relevant part of the error vs. reject curves is usually on the left side where a small fraction, x, of low-quality rejections would be tolerable from the perspective of forcing a second enrolment attempt. We set $x = 0.3$, because rejection of more than 30% is not operationally feasible. SOM block size of 24 performs the best as it is evident by the sharpest decrease in FNMR when poor quality samples are rejected. Rejection of 10% lowest quality decreases FNMR to 0.085 from its initial value of 0.1. To give a point of reference for benchmarking our method and existing methods, Fig. 11 includes an existing known quality assessment method - the Orientation Certainty Level [11]. Our proposed method outperforms Orientation Certainty Level for rejection rates higher than 5%.

Computation efficiency: We noted the processing time required to compute quality values on a PC equipped with a 2.3 GHz Intel Core i7 and 16 GB of memory. With a network size of $dim = 24$ the quickest processing time was

recorded at a block size of $n = 24$ with 503 seconds to generate quality values for 7973 images (65 ms/image). With the same network size but block size $n = 64$ we recorded a total processing time of 654 seconds (82 ms/image).

We conclude this section by noting that all three evaluation methods employed here demonstrated that our proposed quality algorithm is capable of predicting performance of a commercial biometric comparison subsystem.

5. Conclusions and Future Work

We presented a novel approach for fingerprint quality assessment which is fast and of low computational complexity. Our approach is predictive of biometric performance and avoids the explicit quality feature extraction that often has high computational cost. This is achieved by training a self organizing map to represent the variety of spatial information content of fingerprints in our training data. The training of the SOM network may require a significant computational effort, however the training is performed only once, in an offline setting, using suitable computing resources. Once the code book is generated, the operational mode is extremely efficient as the implicit quality assessment is reduced to a loading-task of the fingerprint-image blocks to the input layer of the trained SOM, and a subsequent lookup of the quality level in the associated Random Forest. Thus our suggested approach has the advantage of having a very low computational cost which makes it particularly suitable for applications constrained in processing resources such as low-end mobile fingerprint devices used by the Unique Identification Authority of India (UIDAI). We investigated the effect of finger image block size (input to the SOM), and the size of the SOM network. Our experimental results suggest that smaller image block size (24 pixels by 24 pixels in our case) performed better than larger sizes (32 or 64 pixels). A larger SOM network while more capable of representing all possible spatial information of its training data, requires a larger number of images for training. We tried network sizes of 24 (576 units), 32 (1024 units), and 40 (1600 units). Network size of 32 outperformed the others, noting that our training data might not have been large enough for network size of 40. Quantitative results presented in Section 4.2 demonstrate the predictive power of our proposed finger quality assessment method.

Future work will address improvements in the SOM training, presentation of SOM output to the learning algorithm (Random Forest), and its fine tuning. A potentially valuable source of information is the distance between an image block and its associated BMU. This information could be used to create a weighted SOM activation histogram. Combining units with low distance between weight vectors may provide more robust classification. Additionally, we plan to repeat the study using comparison scores of one or more comparison subsystem (or algorithms) and investigate its generalizability to comparison algorithms not used at the training time. Finally, we are eager to try this approach for the quality assessment of latent impressions.

References

[1] F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, H. Fronthaler, K. Kollreider, and J. Bigun. A comparative study of fingerprint image-quality estimation methods. *Information Forensics and Security, IEEE Trans. on*, 2(4):734 –743, dec. 2007.

[2] L. Breiman. Random forests. *Mach. Learn.*, 45(1):5–32, Oct. 2001.

[3] T. Chen, X. Jiang, and W. Yau. Fingerprint image quality analysis. In *Image Processing, 2004. ICIP '04. 2004 International Conf. on*, volume 2, pages 1253 – 1256 Vol.2, oct. 2004.

[4] R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, New York, NY, USA, 1973.

[5] European Commission. Decision of 30 november 2009 determining the first regions for the start of operations of the visa information system (VIS) (notified under document c(2009) 8542), 2009.

[6] European Council. 2004/512/EC: Council decision of 8 june 2004 establishing the visa information system (VIS), 2004.

[7] European Council. Proposal for a council decision establishing the visa information system (VIS), 2004.

[8] H. Fronthaler, K. Kollreider, and J. Bigun. Local features for enhancement and minutiae extraction in fingerprints. *Image Processing, IEEE Trans. on*, 17(3):354 –363, march 2008.

[9] P. Grother and E. Tabassi. Performance of biometric quality measures. *Pattern Analysis and Machine Intelligence, IEEE Trans. on*, 29(4):531 –543, april 2007.

[10] ISO/IEC. *ISO/IEC IS 29794-1:2009 Information technology — Biometric sample quality — Part 1: Framework*. International Organization for Standardization, Geneva, Switzerland, Jan. 2009.

[11] ISO/IEC. *ISO/IEC TR 29794-4:2010 Information technology — Biometric sample quality — Part 4: Finger image data*. International Organization for Standardization, Geneva, Switzerland, 2010.

[12] T. Kohonen. The self-organizing map. *IEEE_J_PROC*, 78(9):1464–1480, 1990.

[13] D. Maio, D. Maltoni, R. Cappelli, J. Wayman, and A. Jain. Fvc2004: Third fingerprint verification competition. In *Biometric Authentication*, volume 3072 of *Lecture Notes in Computer Science*, pages 1–7. Springer Berlin Heidelberg, 2004.

[14] National Institute of Standards and Technology. Proceeding of the 2nd international biometric performance conference, 2012.

[15] M. A. Olsen, H. Xu, and C. Busch. Gabor filters as candidate quality measure for NFIQ 2.0. In *Proceedings of the 5th IAPR International Conf. on Biometrics (ICB), March 29-April 1, New Dehli, India.*, 2012.

[16] L. Shen, A. Kot, and W. Koo. Quality measures of fingerprint images. In *IN: PROC. AVBPA, SPRINGER LNCS-2091*, pages 266–271, 2001.

[17] E. Tabassi, C. L. Wilson, and C. I. Watson. NIST-IR 7151 - fingerprint image quality. Technical report, NIST, August 2004.

[18] Unique Identification Authority of India. Role of biometric technology in aadhaar enrollment. uidai.gov.in, 2012.

[19] D. Willett, C. Busch, and F. Seibert. Fast image analysis using kohonen maps. In *Proceedings of the IEEE Workshop NNSP*, pages 461–470, 1994.