Authentication Of Offline Signatures Based On Central Tendency Of Features And Dynamic Time Warping Values Preserved For Genuine Cases

Chitrita Chaudhuri
Computer Sc. & Engg. Dept.
Jadavpur University,
Kolkata-700032, INDIA
cchitrita@hotmail.com

Atal Chaudhuri
Computer Sc. & Engg. Dept.
Jadavpur University,
Kolkata-700032, INDIA
atalc23@gmail.com

Aparajita Khan
Computer Sc. & Engg. Dept.
Jadavpur University,
Kolkata-700032, INDIA
khanaparajita@yahoo.com

This work proposes to authenticate offline signatures using a Case-Based Reasoner (CBR). The case base serves as a repository of sets of genuine signatures for which a central point on the n-dimensional global feature space is preserved along with the Inter-Quartile Range (IQR). These signatures are paired off to perform Dynamic Time Warping (DTW) comparison on their respective contours. Metrics generated from the global features and DTW values for the preserved signatures are utilized to predict authenticity of test signatures. Philosophically, CBR is a good classifier since it does not need any training by forgery models. The overall accuracy of the CBR classifier is maintained at a reasonably high value as a larger False Rejection Rate (FRR) is compensated by a tight False Acceptance Rate (FAR) value when compared with a MLP classifier. Both the classifiers have been tested on a standard offline signature database as well as one collected and prepared during the current research.

Keywords— **Case-Based Reasoner; Dynamic Time Warping; False Acceptance Rate; False Rejection Rate**

1. INTRODUCTION

Offline signatures are still the most prevalent method of biometric authentication. Often, the checking is done manually and generally experts are not called in unless legal controversies occur. Manual detection of fraud presents several complications. There remains ground to suspect personal bias. An organization may suffer loss of face if there is genuine cause for change in writing style due to the ageing process of the signatory and/or for a disease-induced spastic. An expert system, on the other hand, can safely be made scapegoat without jeopardizing goodwill.

The process of signature verification can be posed as a problem to determine whether a particular signature is indeed written by the person claiming to be its author and, if not, whether attempts to forgeries can be established. The idea of applying Case-based Reasoning (CBR) and Dynamic Time Warping (DTW) techniques to solve this problem has been initially stimulated by studying pioneering works in these fields [1] [2]. Yoshimura and Yoshimura (1997)[3] proposed a DTW based signature verification which uses DTW to segment the signature into a fixed number of components and then compute a component wise dissimilarity measure. In [4] Shankar and Rajagopalan propose a modified DTW algorithm that takes into account stability of various components of the signature for enhanced performance in verification.

We employ here a CBR which records a preliminary set of authentic signatures of each person and incorporates values and techniques that help to detect fraud with sufficient accuracy on subsequent presentation of offline signatures . Our choice of classifier was guided by some advantages of the CBR - it does not need any separate training with plausible forged sets. Additionally, since each case preserves the metrics derived from a definite number of pre-recorded signatures of a particular person, there is no question of scalability associated with the growth of the total dataset.

In the next section 2 is discussed the methodologies used to extract specific knowledge from the signature images of each individual and utilize the same to authenticate a new signature supposed to belong to that individual. The steps involved in preprocessing the signature images are detailed in the first subsection. The detailed procedure, for preserving genuine signatures and their metrics in the case base and utilization of these while evaluating a test signature, is discussed in the remaining portion of this section.

To benchmark our system we have tested the signature data on another classifier : the MLP Network. This neural network was designed with the help of the standard data mining software WEKA [12]. Section 3 describes the details of the experiments performed to assess our system. Here we also present the outcome of the experiments in both tabular and graphical representation. This helps us to analyze the efficacy of the system. In the last section 4 is recorded the concluding remarks. Some future scopes of improving the system is also indicated here. The Reference at the end lists the foundation works on which we based our research.

2. METHODOLOGY

2.1 Preprocessing of Signature Images

The following steps were required to prepare the raw signature images and save them in a format from which features needed for classification could be extracted properly :

Image Binarization : From the gray scale scanned image, the global image threshold was calculated using Otsu's

978-1-4799-4272-5/14 $31.00 © 2014 IEEE
DOI 10.1109/EAIT.2014.38

method[5] which performs a clustering based image threshholding.

Noise Reduction : This step removes single white pixel on black background or single black pixel on white background. The technique employs a 3x3 mask to the image such that if the number of the 8-neighbors of a pixel, that have the same color as the central pixel, is less than two, then the color of the central pixel is reversed.

Minimal Area Cropping : The binarized image was scanned horizontally and vertically to obtain first the leftmost and rightmost, and then the topmost and bottommost black pixels respectively. The image is cropped with respect to a rectangular window, known as the minimum bounding box, passing through these four pixels.

Width Normalization : From the minimum bounding box of the signature we obtain the pure width of the signature. Then all authentic signatures are width normalized using average pure width as the scaling factor, while preserving the aspect ratio of the signatures.

Skeletonization : The skeletonizing technique used here is called *Medial axis transformation* introduced by Blum [13] that produces a unique skeleton for a given object by removing pixels on the boundaries of the object, without allowing the object to break apart.

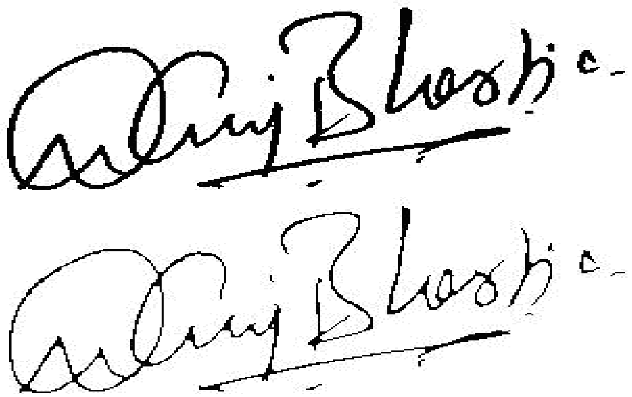

Figure 1. Signature Image before and after Skeletonization

2.2 Feature Extraction

In our work we considered a set of 20 global features[6] [7] [8]. Global features categorize the signature as a whole. These features are usually extracted from the pixels that lie within the region circumscribing the signature image. Some of the advantages associated with global features are that they are easily extractable, less sensitive to noise as small distortions in isolated regions of signature does not cause a major impact on the total image, and they provide information about the structural aspects of the signature and distribution of pixels across the signature image. Moreover, as they are dependent on overall position alignment, these can easily reflect style variations in case of forgeries. Following is a list of global features extracted and utilized by us : *Pure Height, Pure Width, Aspect Ratio, Image Area, Signature Height*, *Vertical Center, Horizontal Center*, *Maximum Vertical Projection*, *Maximum Horizontal Projection, Vertical Projection Peaks*, *Horizontal Projection Peaks*, *Baseline Shift*, *Number of Edge Points*, *Number of Cross Points*, *Number of Closed Loops*, *Top Heaviness, Horizontal Dispersion, Mean Ascender Height, Mean Descender depth,* and *Interior to Exterior pixel ratio*.

Here each signature image S_k can be represented as a feature vector, $F_k = (f_{k1}, f_{k2}, \ldots, f_{k20})$, and the Euclidean distance between two signature images S_i and S_j is given by

$$Dist(S_i,S_j) = \left(\sum_{m=1}^{20} (f_{im} - f_{jm})^2 \right)^{1/2} \qquad (1)$$

2.3 Dynamic Time Warping for signature verification

The DTW algorithm has been widely used in speech processing, bio-informatics and handwriting communities to match one-dimensional sequences. It uses dynamic programming to find an optimal match by allowing stretching and compression of sections of the sequences, the primary objective being non-linearly aligning one or more observation sequences or feature vectors before they are compared. It finds a final alignment between two time series data, $X=x_1, x_2, x_3, \ldots, x_M$ and $Y=y_1, y_2, y_3, \ldots, y_N,$ M and N being the length of the two series respectively, under some constraints. In this work, DTW has been used to find an optimal distance between two signatures' contour lines. (Figure 2 below).

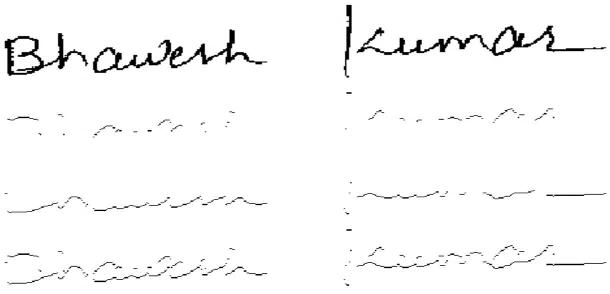

Figure 2. A sample signature with the upper and lower contours separated and together.

Since there are no time sequences associated with an offline signature, we have generated sequences from the vertical offsets of the maxima and minima of the upper and lower contour of a signature against their horizontal positions. Tentatively similar sequences of vertical heights are compared between two signatures by aligning their cumulative distance from one another as they grow, one along the vertical axis and the other along the horizontal axis of an M x N grid. Both sequences start from bottom left corner of the grid, and in each cell the cumulative DTW distance between the corresponding elements of the two sequences is placed. A greedy algorithm determines the best match path through the grid, as near to the diagonal as possible, minimizing the total distance between the sequences using a non linear mapping. The DTW distance $D(i,j)$ is calculated and the grid populated using dynamic programming techniques, starting from location $i=1$ and $j=1$ and assuming the grid position at $(1,1)$ to be already filled up with the Manhattan distance between the corresponding x_1 and y_1 values. The ultimate DTW distance between the two sequences is thus obtained using the following recurrence relation :

$$
\begin{aligned}
D(i, j) &= D(1,j-1) + d(x_1,y_j), \text{ at } i=1 \text{ and } j>1,\\
&= D(i-1,j) + d(x_i,y_1), \text{ at } i>1 \text{ and } j=1,\\
&= \min \{ D(i,j-1), D(i-1,j), D(i-1,j-1)\} + d(x_i ,y_j),\\
&\qquad \text{at } i>2 \text{ and } j>2
\end{aligned}
\tag{2}
$$

where

$d(x_i ,y_j)$ is the Manhattan distance between the i-th element of the x-series and j-th element of the y-series data,

and

$D(i, j)$ is the cumulative DTW distance so far calculated and occupying grid location (i , j).

$D(M,N)$ thus gives us the DTW distance between the two sequences.

As outlined by Sakoe and Chiba [10], the major constraints of the DTW algorithm while calculating the nature of the acceptable paths through the grid, namely monotonic condition, continuity condition, boundary condition, warping window condition and slope constraint condition, were all considered during optimal path extraction.

2.4 Working Procedure

Each case of ten genuine signatures per person is utilized for predicting the authenticity of a test signature. Feature deviations are augmented by the Dynamic Time Warping distance measured for local maxima and minima in both the upper and lower contours of the test signature being compared with each of the authentic signatures stored in the case. Statistical dispersion of both are calculated. A set comprising of the median and IQR value is measured in terms of the global features for the test signature. A second such set is determined by subjecting the test specimen to a central tendency measure with respect to the DTW scores obtained as stated above. These two sets mark the basis of the components on which the authenticity of a test signature depends. A value is obtained by generating the weighted sum of the absolute difference of the median distance of the test signature from the genuine specimens for both these components when compared with their respective IQRs. If this sum value falls within a predetermined threshold limit, the test signature is predicted to be genuine, otherwise it is assumed to be a forgery. The procedure is explained as follows:

- Every authentic signature is assumed to be a point in a 20 dimensional feature-space.
- Let ξ be the set of genuine specimens in a case, with say $n=|\xi|$. Euclidean distance between each pair of signatures $S_i, S_j \in \xi$ in the feature space is computed according to Eqn. (1) and these ${}^nC_2$ distance values are utilized to obtain the median distance value M_{ftr} and the inter quartile range IQR_{ftr} .
- For each of these signature pairs $S_i, S_j \in \xi$, the upper and lower profiles are extracted. Next the dynamic time warping (DTW) distance between the upper profile local maxima sequences of signature S_i and S_j is calculated and stored in $Dupr_{ij}$ according to Eqn. (2) . Similarly, DTW distance between the lower profile local minima sequences of signature S_i and S_j are calculated and stored in $Dlwr_{ij}$. These two values $Dupr_{ij}$ and $Dlwr_{ij}$ are summed up for each pair of signatures S_i and S_j where $i, j = 1,2,\dots,n$ and $i \neq j$ and is accumulated to generate the median M_{dtw} and the inter quartile distance range IQR_{dtw} from the ${}^nC_2$ total DTW distance values.
- These four parameters M_{ftr}, IQR_{ftr}, M_{dtw} and IQR_{dtw} along with the signatures and their integer scores are all preserved in the case for further processing.
- When a new test signature T arrives, global features are extracted from it after preprocessing the image. Euclidean distance between feature set of T and each genuine signature $S_i \in \xi$, $\forall i = 1,2, \dots, n$ in the case is computed and the median distance MT_{ftr} for the test signature is obtained.
- DTW distances $Dupr_i$ and $Dlwr_i$ for the upper profile local maxima sequences and lower profile local minima sequences of T from the corresponding sequences for each genuine signature in the case are obtained. The total DTW distance, DTW_{TSi} between T and S_i , $\forall i = 1,2, \dots, n$, is calculated as sum of $Dupr_i$ and $Dlwr_i$ and the median of all such distances is computed in MT_{dtw}.
- The test signature T is classified as a genuine signature if it satisfies the following inequality:-

$$\alpha * DTW\_Comp + \beta * FTR\_Comp \leq 1 \tag{3}$$

where,
$DTW\_Comp = (abs(MT_{dtw} - M_{dtw}) / \gamma * IQR_{dtw})$,
$FTR\_Comp = (abs(MT_{ftr} - M_{ftr}) / \delta * IQR_{ftr})$,
α is the DTW similarity weight with value ranging between 0 and 1,
β is the feature similarity weight, such that $\beta = 1 - \alpha$,
γ is the allowed percentage of IQR_{dtw} with value ranging between 0.1 and 1,
δ is the allowed percentage of IQR_{ftr} with value ranging between 0.1 and 1.5.

The above four pre-defined authentic bounds α, β, γ, and δ help detect forgery. In our experiments, we have varied these values at a rate of 0.1 gradation to produce all possible combinations and detected the bound values for which the total error was lowest for both datasets. In the next section we have recorded the results for the lowest total error positions and analyzed the performance of the CBR classifier.

3. Experimental Results and Performance Analysis

3.1 Description of the Datasets

Two sets of data have been utilized in our experiments. The first is a standard database *MCYT Bimodal Biometric Database (MCYT- SignatureOff-75)* of off-line signatures with

15 genuine and 15 skilled forgeries for 75 persons [9] scanned at a resolution of 300 dpi. The second one is our own collection of off-line signatures prepared by us over a period of one and half year containing twenty authentic signatures for 75 persons. To check the system's performance while detecting forgery, we added three different categories of forged signatures for each of these persons, namely *Random Forgery, Unskilled Forgery, Skilled Forgery*[11]**.**

Our forged set consists of 10 random forgeries, 5 unskilled and 5 skilled forgeries for each of the 75 persons. We shall henceforth designate these two data sets as Dataset1 and Dataset2.

All the signatures in Dataset2 were collected on paper using black ball point pens with 0.5 micron tip points. By strategy, the collection procedure was phased out over a duration of a year and a half. Out of the twenty genuine signatures, ten were set aside for building up the initial case base. The remaining ten authentic signatures and the set of forged signatures were used to assess the classifier accuracy. All the signatures were scanned at a resolution of 200 dpi to obtain gray scale images.

3.2 Tables and Graphs depicting result

In Figures 3 and 4, we plot the FAR, FRR and the Total Error (FAR+FRR)**,** obtained for the lowest total error condition(/s) selected from all combinations of values of ***α, β,*** and ***γ***, and for a series of values of ***δ***, for both Dataset1 and Dataset2 respectively. The particular values of ***α*** and ***β*** were found to be ***0.7*** and ***0.3*** for both datasets. The ***γ*** value was found to be ***0.6*** for Dataset1 and ***1.0*** for Dataset2. In each case, the error values are plotted along the vertical axis increasing from 0 upwards, against the ***δ*** values plotted along the horizontal axis varying between ***0.1*** and ***1.5*** from left to right.

We find the Equal Error Rate (EER) position with respect to the point where the FAR and FRR values are exactly the same, i.e. at the intersection of the FRR and FAR curves. We also find out the Lowest Total Error position on the Total Error curve.

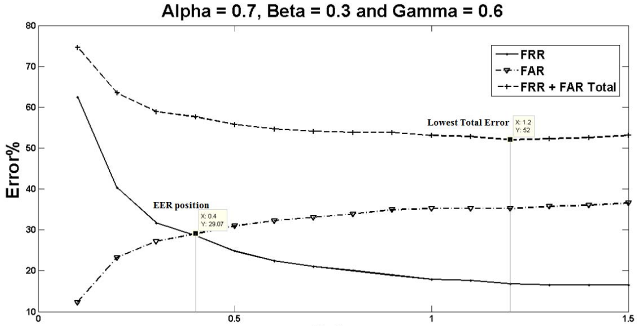

Figure 3. FAR, FRR, Total Error and EER values for Dataset1

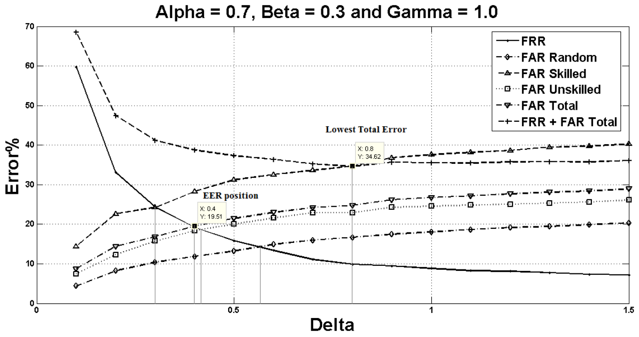

Figure 4. FAR, FRR, Total Error and EER values for Dataset2

The values marked on the above graphs have been tabulated for the different types of forgeries in Table I below for both datasets. The ***α, β, γ*** and ***δ*** values, found for the lowest total error values obtained by varying DTW and feature value dispersions for different percentage participation of their respective scores using Eqn. (3), are also indicated in the Figures 3 and 4 above and Table I below.

TABLE I. EER AND LOWEST ERROR VALUES

Data Set	***α=0.7, β=0.3,* and *γ=0.6(Dataset1) ,1.0(Dataset2)***			
	Forgery	***EER%***	**δ_{EER}**	**Lowest Error δ**
1	**Alltype**	**29.07**	**0.4**	**1.2**
2	**Alltype**	**19.51**	**0.4**	**0.8**
2	**Skilled**	**24.32**	**0.3**	
2	**Unskilled**	**18.61**	**0.41**	
2	**Random**	**14.29**	**0.56**	

We have further compared the performance of two classifiers, MLP and CBR, in terms of accuracy. In Table II below are given the percentage of signatures correctly identified as genuine or fraud for both sets of data and for the two classifiers MLP and CBR. The CBR system was assessed in each case with the ***α, β, γ*** and ***δ*** value settings at the Lowest Total Error position on the Total Error curve as obtained from the Figures 3 and 4 and Table I above. The corresponding bargraphs for Dataset1 is shown in Figure 5 next.

TABLE II. COMPARISON OF ACCURACY % : MLP AND CBR

Dataset	**Classifier Model**	**Overall**	**Authentic**	**Alltype Forgery**
1	**CBR**	**74**	**83.2**	**64.8**
1	**MLP**	**71.6**	**88.26**	**54.93**
2	**CBR**	**81.55**	**90.13**	**75.24**
2	**MLP**	**77.18**	**87.067**	**63.45**

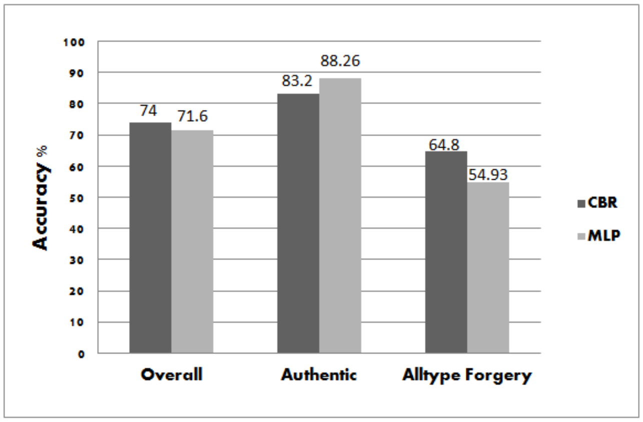

Figure 5. **Accuracy BarGraph for Dataset1**

In Figure 6 below are shown the accuracy percentage of the different categories of forged and genuine as well as the overall accuracy of data in the second set. The values of the different types of forgeries : the skilled, unskilled and random, are given in Table III below, the CBR values being shown at the ***α, β, γ*** and ***δ*** at lowest total error as indicated in Table I.

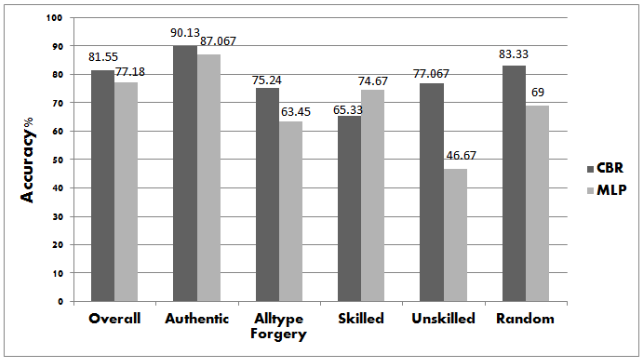

Figure 6. **Accuracy BarGraph for Dataset2**

TABLE III. **FORGERY DETECTION ACCURACY FOR DATASET2**

Model	Alltype Forgery	Skilled Forgery	Unskilled Forgery	Random Forgery
CBR	**75.24**	**65.33**	**77.067**	**83.33**
MLP	**63.45**	**74.67**	**46.67**	**69**

As is evident from the above figures and tables, the overall accuracy percentage as well as the total fogery detection rate is perceptably higher in case of CBR for both the datasets under observation. The recognition rate for the more common types of forgeries, i.e. the unskilled and the random varieties, is also high for our Dataset2. The only variety of forgery available with Dataset1 is skilled. We found an interesting negative correlation between the detection of the authentic signatures and the skilled forgeries, which supports our view that so far CBR imitates the human form of knowledge inference – a signature, perceptibly different from the preserved ones, has got a high chance of rejection even if it is authentic, while a skillfully forged one may be accepted if it is a close enough copy.

The percentage of accuracy in detecting fraud, which left ground for dissatisfaction since higher rates have been claimed by others employing DTW techniques [14], were slightly enhanced by introducing additional grid feature vectors in the form of pixel densities and angle values [15] for the CBR classifier. The modified results have been displayed in the following tables.

TABLE IV. **RESULT ENHANCED BY PIXEL DENSITY AND ANGLE VALUES**

Dataset	Classifier Model	Overall	Authentic	Alltype Forgery
1	**CBR**	**74.13**	**83.2**	**67.2**
1	**MLP**	**68.26**	**88**	**48.4**
2	**CBR**	**85.02**	**90.13**	**79.42**
2	**MLP**	**82**	**93.067**	**66.36**

TABLE V. **FORGERY RESULTS ENHANCED FOR DATASET2**

Model	Alltype Forgery	Skilled Forgery	Unskilled Forgery	Random Forgery
CBR	**79.42**	**65.6**	**81.06**	**91.6**
MLP	**66.36**	**79.43**	**46**	**73.66**

The CBR results for this enhanced model are being shown at ***α=0.4, β=0.6*** for both Datasets, and for ***γ=0.6***, and ***δ=0.7*** for Dataset1 and ***γ=0.8***, and ***δ=1.3*** for Dataset2 at lowest total error position.

The MLP values shown in Tables IV and V have been obtained by incorporating the same set of pixel density and angle value features, as mentioned before, on both the datasets. FRR here decreased but at the cost of slightly increased FAR in case of unskilled variety of forgery for Dataset2. Cross-validation was omitted as, even for a 3-fold validation, the number of training samples for skilled and unskilled were too small to avoid overfitting. Moreover, we wanted both the classifiers to start with the same handicap by presenting them with identical model-building and testing set from the original signatures of each person. But we found that, although MLP enjoys the benefit of extra training in the form of sample forged signature sets, addition of grid attributes seem to have detrimental effect on its performance for the standard database.

4. CONCLUSIONS AND FUTURE SCOPE

The experimental results with our database indicate that MLP can better authenticate original signatures, although our CBR system can be tuned to perform appreciably well - albeit at the cost of raising the FAR for skilled forgeries. At an optimal tuning, CBR outperforms MLP by rejecting the false more accurately. The overall performance of the CBR was found to be higher for both the datasets used by us. It was further enhanced later by introducing some more feature vectors at grid level.

In case of MLP, lowering of accuracy in detecting overall fraud is aggravated due to the intrinsic variance of form suffered by training samples for forged signatures. For CBR, end-comparison by DTW, which works directly on image data, leads to easier detection of unskilled and random types of forgery. DTW, being a comparison technique between two inputs, is not applicable at any stage of MLP.

Another prominent advantage of CBR lies in the fact that it needs no prior training by forged samples as required by the MLP system. The underlying principle of CBR is by far nearer to human learning process in this particular environment and suits the practical implementation of the scheme perfectly.

In CBR, an overall raise in rejection of genuine signatures may well be a cause of concern. To rectify this situation, our future endeavor would be oriented towards utilizing a more refined feature space from the input data. Higher complexities incurred therein may require improved indexing and data mining techniques, opening a vista of future research work in CBR.

References

[1] Sankar K. Pal, Simon C.K. Shiu, “Foundations of Soft Case-Based Reasoning”, John Wiley & Sons, Inc., 2004.

[2] S. Chen and S. N. Srihari, "Use of Exterior Contours and Word Shape in Off-line Signature Verification," *Proc. International Conference on Document Analysis and Recognition*, Seoul, Korea, pp. 1280-1284, August 2005.

[3] Yoshimura, M., Yoshimura, I.,“An application of the sequential dynamic programming matching method to off-line signature verification”,Lecture Notes in Computer Science. In: Proc. of First Brazilian Symposium on Advances in Document Image Analysis 1339. pp. 299–310, 1997.

[4] Shankar A.P and A. N. Rajagopalan, “Off-line signature verification using DTW”, Pattern Recognition Letters, Vol.28, pp. 1407-1414, 2007.

[5] Nobuyuki Otsu, "A threshold selection method from gray-level histograms". IEEE Transactions on Systems, Man and Cybernetics. 9 (1): 62–66,1979.

[6] H. Baltzakis and N. Papamarkos, "A new signature verification technique based on a two-stage neural network classifier", Engineering Applications of Artificial Intelligence 12, 95-103, 2001.

[7] Alan McCabe, Jarrod Trevathan and Wayne Read, "Neural network-based handwritten signature verification", Journal of Computers, Vol. 3, No. 8, August 2008.

[8] Kai Huang and Hong Yan, "Off-line signature verification based on geometric feature extraction and neural network classification", Pattern Recognition, Vol. 30, No. 1, pp. 9-17, 1997.

[9] http://atvs.ii.uam.es/mcyt75so.html (ATVS - Biometric Recognition Group >> Databases >> MCYT - SignatureOff - 75).

[10] H. Sakoe, S. Chiba, " Dynamic Programming Optimization for Spoken Word Recognition", IEEE Transactions on Acoustics, Speech and Signal Processing,Vol.ASSP-26,No.1, pp 43-49, 1978.

[11] Sanjay N. Gunjal, Manoj Lipton, " Robust Offline Signature Verification Based on Polygon Matching Technique", International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 1, Issue 1, November 2011.

[12] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten , The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1,2009.

[13] H.Blum, “ A Transformation for Extracting New Descriptors of Shape”, in *Symposium Models for Perception of Speech and Visual Form,* W. Whaten-Dunn,Ed., MIT Press, Cambridge, MA, 1967

[14] P. S. Deng, H.-Y. M. Liao, C. W. Ho, and H.-R. Tyan, “Wavelet-based offline handwritten signature verification,” *Comput. Vis. Image Underst.*, vol. 76, no. 3, pp. 173–190, Dec. 1999.

[15] V.K.Madasu,B.C. Lovell, “An Automatic Offline Signature Verification and Forgery Detection System”, Pattern Recognition Technologies and Applications: Recent Advances, p.p. 63-89, 2008.