

Bag-of-Visual-Words for Signature-Based Multi-Script Document Retrieval

Ranju Mandal ✉ · Partha Pratim Roy ·
Umapada Pal · Michael Blumenstein

the date of receipt and acceptance should be inserted later

Abstract An end-to-end architecture for multi-script document retrieval using handwritten signatures is proposed in this paper. The user supplies a query signature sample and the system exclusively returns a set of documents that contain the query signature. In the first stage, a component-wise classification technique separates the potential signature components from all other components. A bag-of-visual-words powered by SIFT descriptors in a patch-based framework is proposed to compute the features and a Support Vector Machine (SVM)-based classifier was used to separate signatures from the documents. In the second stage, features from the foreground (i.e. signature strokes) and the background spatial information (i.e. background loops, reservoirs etc.) were combined to characterize the signature object to match with the query signature. Finally, three distance measures were used to match a query signature with the signature present in target documents for retrieval. The ‘Tobacco’ [1] document database and an Indian script database containing 560 documents of Devanagari (Hindi) and Bangla scripts were used for the performance evaluation. The proposed system was also tested on noisy documents and promising results were obtained. A comparative study shows that the proposed method outperforms the state-of-the-art approaches.

Keywords Signature Retrieval · Logo Retrieval · SIFT · Bag-of-Visual-Words · Spatial Pyramid Matching · Content-Based Document Retrieval

Ranju Mandal

School of Information and Communication Technology, Griffith University, Queensland, Australia

E-mail: r.mandal@griffith.edu.au

Partha Pratim Roy

Dept. of Computer Science & Engineering, Indian Institute of Technology, Roorkee, India

E-mail: proy.fcs@iitr.ac.in

Umapada Pal

Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata, India

E-mail: umapada@isical.ac.in

Michael Blumenstein

School of Software, University of Technology Sydney, Australia

E-mail: michael.blumenstein@uts.edu.au

1 Introduction

Handwritten signatures are pure behavioural biometric and have long been used as identifying marks in documents as they provide rich information such as unique properties of an individual’s behaviour. Signature verification and recognition are considered for biometric authentication in administrative documents, legal documents, bank cheques etc. In documents, signatures are often examined by forensic document analysis experts for authenticating documents and to address fraud. Moreover, a document containing a signature may provide richer knowledge about the origin of the document. Thus, the handwritten signature will undoubtedly add an advantage for document indexing and searching. So, signatures could be used as key information for searching and retrieval of relevant documents from large heterogeneous document image databases.

Large institutions and corporations still receive a high volume of communication in paper form because of their legal significance. It is a common organisational practice nowadays to store and maintain large digital databases which is an effort to move towards a paperless office. Large quantities of administrative documents are often scanned and archived as images (e.g. the ‘Tobacco’ [1] dataset) for general-purpose correspondence, however, without adequate indexing information. Consequently, there is a tremendous demand for robust ways to access and manipulate the information that these images contain. For example, a survey had revealed that over 55 billion bank cheques processed annually in North America at a cost of 25 billion dollars [2]. Google and Yahoo have recently announced their intention to make handwritten books accessible through their search engines [3]. In this context, field-based document image retrieval will be a valuable tool for users to browse the contents of these books. Obtaining information resources relevant to the query information from such repositories is the main objective of content-based document retrieval. Hence, detection as well as recognition of signatures from documents is very important because of the various applications that it brings. Thus, the objective of this paper is to present a novel handwritten signature-based document retrieval approach, which can be applied to real-world scenarios. A few samples of scanned documents from the ‘Tobacco’ dataset, as well as a Hindi and a Bangla dataset are shown in Fig. 1. These documents are printed texts with one or more signatures in the document.

Automatic signature detection is the initial stage to a signature-based document image retrieval system. But, detection of signatures from a document page involves challenges due to the free-flow nature of handwriting strokes and writing styles of different individuals. Sometimes, the detection of signatures is challenging due to their overlapping/touching nature with other information (background text and graphical lines) in the document (see Fig. 1 for one such example). Also, signatures often have similar strokes to that of handwritten text, which makes it difficult to detect when signatures overlap with handwritten text in a document. After detection of signatures from document images, the matching process with the query signature is also a challenging task due to various reasons, such as the existence of variability among signatures of the same signatory (interclass variation), and the fact that a signature may contain a different number of components in different documents.

Traditional OCR systems have some limitations in working on handwritten scripts for indexing and searching from document image databases. In this pa-

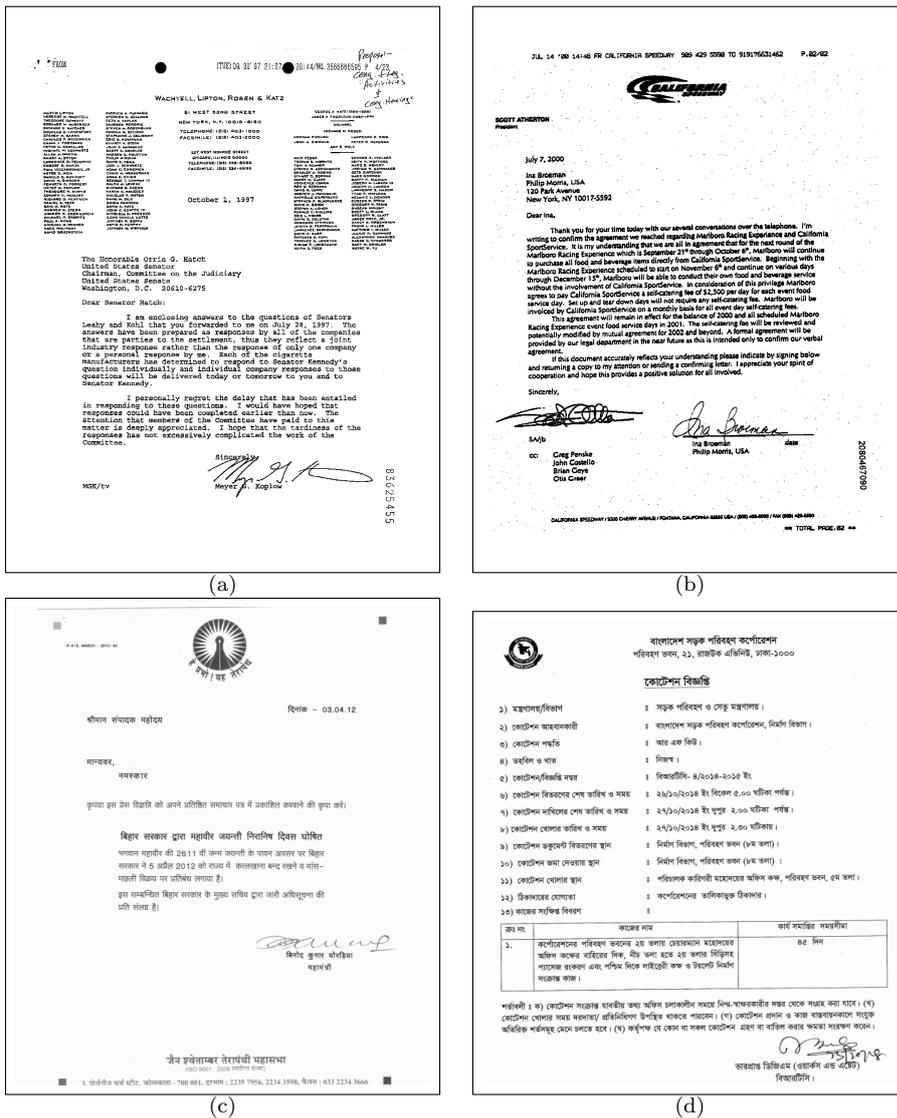


Fig. 1 Signed printed documents of different scripts are shown here. (a,b) Samples of printed signed English documents from the 'Tobacco' dataset. (c) A letter printed in Devanagari script and (d) an official notice printed in Bangla script.

per, a complete end-to-end architecture for automatic document retrieval from a multi-lingual (i.e. English, Devanagari and Bangla scripts) document repository is proposed using handwritten signatures. The system could be used to retrieve documents based on signature information from different databases such as administrative documents, historical archives, postal mail, etc. In the experiments, the properties of signatures/documents are unconstrained in nature, with diverse layout structures and complex backgrounds. Moreover, in multilingual and multi-

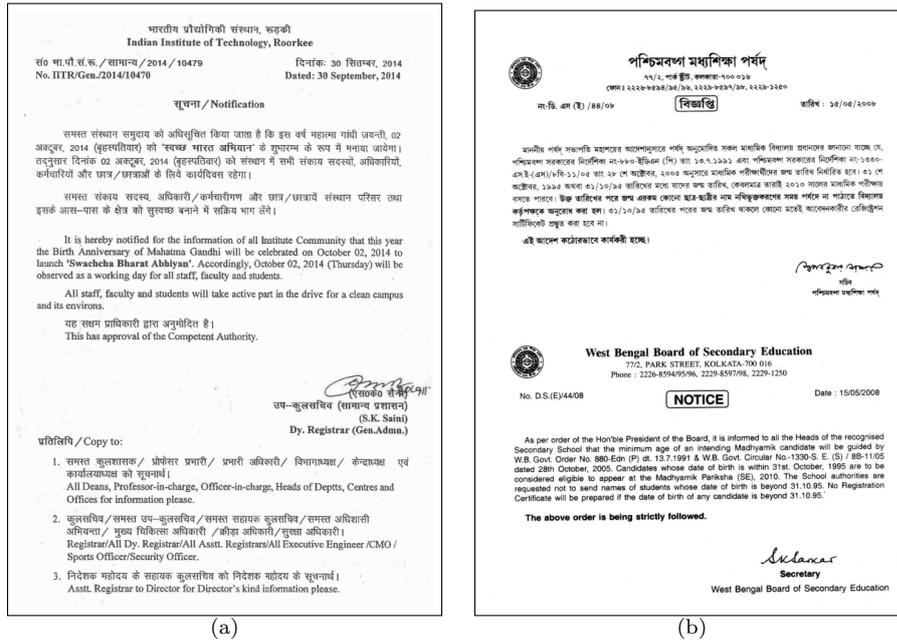


Fig. 2 Samples of Indian multi-script official documents. (a) Document containing English (Roman) and Devanagari scripts (b) Document containing English and Bangla scripts.

script countries such as India, retrieval of multi-script documents using signature information is more challenging due to the presence of signatures as well as the text of different scripts in a single document. An Indian state generally uses three official languages. For example, the West Bengal state of India uses Bangla, Devanagari, and English as official languages. Fig. 2 (b) contains signatures and text of English (Roman) and Bangla scripts. Hence, a single document may contain one or more of these three scripts and such documents as multi-script documents are considered. Fig. 2 shows some examples of signed official multi-script documents containing English scripts along with Devanagari and Bangla scripts.

The following are the contributions of the proposed work:

- A complete end-to-end architecture comprising of signature detection, grouping, and signature matching steps are presented
- Bag-of-Visual words combined with Spatial Pyramid Matching for signature detection is employed to achieve higher performance
- A novel technique based on Harris-Stephens corner points and density-based clustering is applied to group signature components in a robust way
- Finally, the signature's background information is combined with the foreground information in feature extraction, which leads to a significant improvement in signature recognition accuracy
- Proposed method has genericness attribute which has been validated by the encouraging results when applied for Logo detection and matching. The experimental outcomes also prove that the proposed method is also tolerant to noisy documents

It should be pointed out that in the experiments, only printed administrative documents are considered. This is mainly based on the fact that the usage of handwritten documents is effectively outdated in the context of administrative communications. Additionally, the same performance could not be expected at the signature detection level, as signatures and the text would both be handwritten. The ‘Tobacco’ public dataset was considered therefore as this contains administrative documents with machine printed text and signatures.

The system could also be useful to retrieve documents in a multi-script environment. In addition, the proposed architecture works as a generic method for document retrieval based on signatures as well as logo information. Different experiments for document retrieval based on logo information have also been performed. The main objective of these experiments on the logo is to validate that the system can also be extended for logo-based retrieval as the proposed feature extraction technique is robust. Moreover, to investigate the robustness of the proposed system, some experiments on synthetic noisy signed documents are performed and the results outperform existing methods.

The rest of the paper is organized as follows. Section 2 presents the literature review. In Section 3, the proposed approach is described in three sub-sections. Section 3.1, Section 3.2 and Section 3.3 describe the signature detection, signature component grouping and matching techniques, respectively. The experimental results are presented in Section 4. Finally, conclusions are presented in Section 5.

2 Related work

Significant work has been undertaken in the area of detection, segmentation and recognition of graphical elements [4, 5, 6] from a document image for the purpose of document retrieval. There are also considerable existing methods available [7, 8, 9, 10, 11] for identification/classification of handwritten text at different levels namely word, line, zone, etc. Few recent works are also available on mobile signature verification [12] and signature recognition [13, 14]. Since signatures are also handwritten, some research on handwriting text identification is also discussed here.

Farooq et al. [7] proposed a Gabor filter-based feature extraction approach and an Expectation Maximization (EM)-based probabilistic neural network for handwritten text identification. This work is a simple classification problem of two classes (i.e. handwritten and printed) where word-level features were extracted and classified. Peng et al. [10] used a modified K-Means clustering algorithm for text identification from annotated documents at an initial stage and a Markov Random Field (MRF) was applied for relabelling purposes in the final stage. Although the system is robust for handwriting separation, the same technique cannot be applied on detection of signatures with multiple components. An algorithm for identification and segmentation of handwriting in noisy document images was proposed by Zheng et al. [11] using structural and texture features such as bi-level Co-occurrence, bi-level 2×2 -grams, pseudo Run-Lengths, and Gabor filters. The Fisher classifier was used to distinguish text into two classes as handwritten and printed. The rule-based method which computes spatial proximity in the horizontal direction lacks robustness. There are many existing methods which deal with automatic online/offline signature verification and recognition [15]. However, these

approaches use only isolated signatures, and there is not much work that focuses on document retrieval based on the signature information.

Chalechale et al. [16] proposed an approach for signature-based document retrieval using connected component analysis and a geometric property-based feature. The extracted feature is scale and rotation invariant which is desirable for signature-based document retrieval but the component-based feature extraction assumed signature as a single component. A signature-based document retrieval method was proposed by Zhu et al. [17]. Here, structural salience from the curvature of contour fragments was used for signature detection. The challenge of signature detection remains when the segmentation of contour from the background/touching strokes of signatures is difficult.

A Conditional Random Fields (CRF)-based model was proposed by Srinivasan and Srihari [18] on signature-based retrieval from a scanned document repository. The extracted segments of the scanned documents were labeled as machine-printed, signature and noise. Next, a Support Vector Machine (SVM)-based classification technique was employed to remove noise and printed text overlapping the signature images. Finally, a global shape-based feature was computed for each signature image for the task of retrieval but it is not clear how the system will handle if more than one signature exist. The Generalized Hough Transform (GHT)-based approach was proposed by Roy et al. [19] for signature-based documents retrieval. The spatial correspondence between the blobs of the signature query and the target documents was matched. In the early work by the present authors [20], a Conditional Random Field (CRF)-based technique was used to segment signatures from printed documents.

A signature matching method was proposed by Du et al. [21] based on locality sensitive hashing(LSH). All features of contour points are clustered and then a term-frequency histogram was built for each signature as the high-level feature. The K-Nearest Neighbor (K-NN) search-based technique was used to find the closest sample for a query signature. However, this method does not work on partial signatures, to build the holistic features, local information was used. The time complexity of K-NN search is also high. Briceño et al. [22] proposed an angles based parameterization system of signature edge (2D-shape) for off-line signature recognition. A range of experiments was conducted with three different classifiers, the K-NN, Neural Networks and Hidden Markov Models. This method solves a correspondence problem between point features extracted from signature shapes. A better matching performance is achieved for tolerating lower degrees of rigidity by this type of methods. However, these methods are intractable as computationally expensive when the size of dataset grows [21].

Here some proposed algorithms are mentioned with similar objectives but some other content such as logos, text, etc. that were used instead of signatures for retrieval of documents. A content-based retrieval algorithm based on a hierarchical matching tree was proposed by Dewan et al. [23]. Hough transform-based feature descriptors were extracted from paragraphs and line blocks and based on these descriptors, documents were indexed. The similarity of two images was defined by the Euclidean distance between document feature points in space. Wang [24] proposed an algorithm for logo detection and recognition using a Bayesian model. A multi-level step-by-step approach was used for recognition of logos and the logo matching process involved a logo database. Here, a region adjacency graph (RAG)

was used for representing logos, which models the topological relations between the regions.

Finally, Bayesian belief networks were employed as well in a logo detection and recognition framework. Recently, Alaei and Delalandre [25] proposed a system for detection and recognition from document images. A Piece-wise Painting Algorithm (PPA) and some probability features along with a decision tree were used for logo detection and a template-based recognition approach was proposed to recognize the logo. Significant work has been undertaken [26, 27] to make the handwritten text available for searching and browsing using word spotting. A Recurrent Neural Network-based approach was proposed in [27] to make handwritten documents available for word-based searching and indexing. Neural Networks and CTC Token Passing algorithms were used for the word spotting task. Hidden Markov Model (HMM)-based methods are extensively used for modeling handwritten text, word spotting, etc. In [26], Fischer et al. proposed a learning-based word spotting system that uses HMM sub-word models to spot keywords. The proposed lexicon-free approach can spot arbitrary keywords from the handwritten text. An HMM-based method was employed for word spotting from handwritten documents by Serrano and Perronnin [28]. Local Gradient Histogram (LGH) features were used in this work. Some recently published works [29, 30, 31] are also available in the literature on improving SIFT feature matching for object detection and matching. However, in the proposed approaches, the feature extraction technique (i.e. Bag-of-Features powered by SIFT-descriptors) is completely different to SIFT, and such improved SIFT matching techniques cannot be applied to the problem at hand.

A sample signed machine printed document is shown in Fig. 1. It is to be noted that proper detection of such signatures is a vital step before applying the methods for recognition or a matching scheme.

3 Proposed Methodology

As mentioned earlier in this paper, a technique for signature-based document image retrieval from multi-script documents has been proposed. Three main steps: signature/handwriting components detection, the grouping of signature components and the matching technique between the query signature and the signature of the target document are discussed here in detail. A connected component analysis-based technique is used to extract the components from the document. Very small components were ignored in the classification stage using a stroke width-based component's size threshold. Next, features based on a bag-of-visual-words powered by SIFT descriptors and an SVM-based classifier are used to segment the signature components from the document. Finally, signature components are grouped and matched with the query signature to retrieve the target documents. For signature matching purposes, the signature object is characterized by spatial features from signature strokes (i.e. foreground information) and background loops and reservoirs (i.e. background information). Finally, the foreground and background features are combined and relevant documents are retrieved based on a distance measure between the query signature and the signature in the target documents. A detailed discussion of all the three steps is given below.

3.1 Signature Detection

An efficient patch-based SIFT descriptor with a Spatial Pyramid Matching (SPM)-based pooling scheme was applied for feature extraction in the proposed signature detection task. Here, detection of signatures refers to the classification of components in a document into two classes, i.e. signature components and printed components. The feature extraction module used here has three components. A flow diagram of feature extraction and classification for signature detection is presented in Fig. 3. First, SIFT descriptors were extracted from the components of the signature and the K-means clustering algorithm was used to create the code-book. Next, the SPM-based scheme was applied for the final representation of an image. Finally, the SVM was employed for classification. The general idea of the SIFT-descriptors and the SPM employed in the proposed technique are described below in Section 3.1.1 and Section 3.1.2, respectively. The feature extraction and classification modules are detailed in Section 3.1.3.

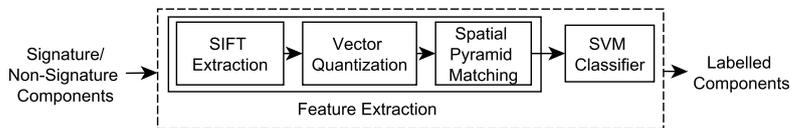


Fig. 3 Flow diagram of the signature detection module.

3.1.1 SIFT descriptor

The SIFT (Scale-Invariant Feature Transform) [32] is a local shape descriptor to characterize local gradient information. Here, a 128-dimensional vector for each keypoint is extracted which stores the gradients of 4×4 locations around a pixel in a histogram bin of 8 directions. The SIFT descriptor is scale and rotation invariant. The gradients are aligned to the main direction, which makes it a rotation invariant descriptor. Different Gaussian scale spaces are considered for the computation of a vector to make it scale invariant. The blue asterisk symbols in Fig. 4 represent the 14×14 SIFT patches of signature and printed components.

3.1.2 Spatial Pyramid Matching (SPM)

The SPM is an extended version of the Bag-of-Features (BoF) model, which is simple and computationally efficient. As the BoF model discards the spatial order of local descriptors, it restricts the descriptive power of the image representation. The limitation of BoF is overcome by the SPM [33] approach, which is successfully applied on image categorization tasks. An image is partitioned into $2^l \times 2^l$ segments where $l = 0, 1, 2, 3, \dots, n$; represents different resolutions. Next, the BoF histograms are computed within each of the 2^l segments, and finally, all the histograms are concatenated to form a vector representation of the image. SPM is equivalent to BoF, when the value of the scale $l = 0$. Here, pyramid matching is performed in

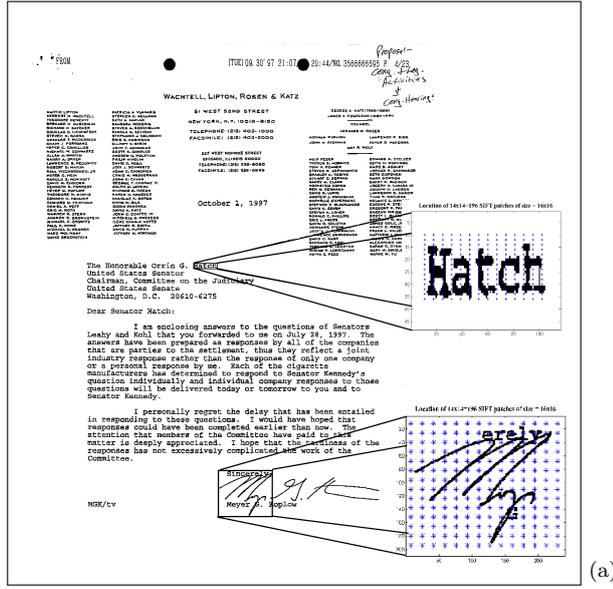


Fig. 4 Blue asterisk symbols represent 196 (of 14x14) SIFT patches of Signature and Printed components.

two-dimensional image space and uses a traditional clustering technique in feature space. The number of matches at level l is given by the histogram intersection function:

$$I(H_X, H_Y) = \sum_{i=1}^D \min(H_X(i), H_Y(i)) \quad (1)$$

where H_X, H_Y represent the histograms obtained from image X, Y respectively and D represents the dictionary size.

Finally, the representation of the image for classification is the total number of matches from all the histograms, which is given by the definition of a pyramid match kernel:

$$K_{\Delta}(\Psi(X), \Psi(Y)) = \sum_{i=0}^L \frac{1}{2^i} N_i \quad (2)$$

where N_i is the number of newly matched pairs at level i and the value is determined by subtracting the number of matches at the previous level from the current level and $\Psi(X), \Psi(Y)$ represent the histogram pyramid obtained from X, Y respectively.

$$N_i = I(H_i(X), H_i(Y)) - I(H_{i-1}(X), H_{i-1}(Y)) \quad (3)$$

3.1.3 Feature Extraction and Classification

This section briefly describes the feature extraction and classification method at the component level for signature detection. First, the image was divided into

14×14 patches (see Fig 5a) to obtain a dense regular grid, instead of interest points, which was based on the comparative evaluation of Fei-Fei and Perona [34]. A total of 196 patches were extracted from the image. The higher dimensional SIFT descriptors [32] of the 16×16 pixel patch were computed over each patch. A set of 196 vectors of dimension 128 were finally obtained at the end. Next, the K-means clustering technique was applied on the extracted SIFT descriptors from the training set for the generation of the codebook. The typical vocabulary size for the experiments was 256. The number of patches ($14 \times 14 = 196$) and the size of the vocabulary (256) was selected experimentally as any significant increase in performance beyond these numbers was not achieved. The size of the vector obtained after codebook matching was 256 which is equal to the vocabulary size. The codebook matching process always returns 256 vector regardless input number of patches.

Finally, an SPM scheme was employed to generate the actual feature vector using the vector of dimension 256 obtained from the previous step, which was then fed to the SVM classifier [35]. In the experiment, the image vector was divided into $2^l \times 2^l$ segments in three different scales $l = 0, 1$ and 2. 21 ($16+4+1$) BoF histograms were computed (SPM configuration was adopted from Lazebnik et al. [33]) from these three levels and all the histograms were concatenated to get the final vector representation of size 5376 (21×256) from an image.

For example, 196×128 dimensional vector was obtained from Fig 5a as a result of computing 128 dimensional SIFT descriptor each from 196 patches. The dimension of our dictionary is 256×128 which was computed from the SIFT descriptors of patches from the training dataset. In the next step, our dictionary matching process always returns a vector of 256 dimensions when matching between the dictionary and a set of SIFT descriptors. We, therefore, obtained 256-dimensional features from one matching process and we continue this process for 21 times at three scales (i.e. I_0, I_1 and I_2) as illustrated in Fig 5b. The equation below represents the pyramid match kernel for three scales:

$$K_{\Delta} = I_2 + \frac{1}{2}(I_1 - I_2) + \frac{1}{4}(I_0 - I_1) \quad (4)$$

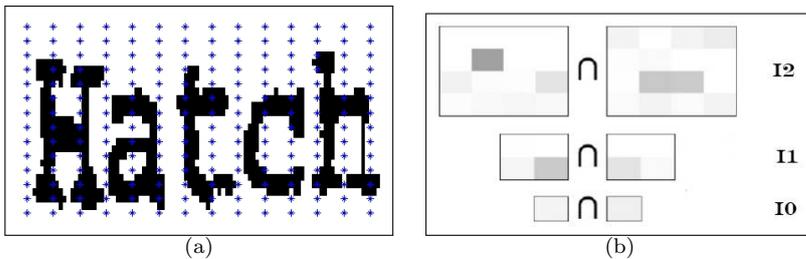


Fig. 5 (a) Location of $14 \times 14 = 196$ SIFT patches of size 16×16 pixels (b) Three scales of pyramid matching. I_0, I_1 and I_2 represent a global, 4 locals and 16 locals matching respectively.

Classifier: SVM is a popular classification technique which can successfully be applied to a wide range of applications [35]. So, in the experiments, the SVM classifier was used. SVMs are defined for two-class problems and they look for

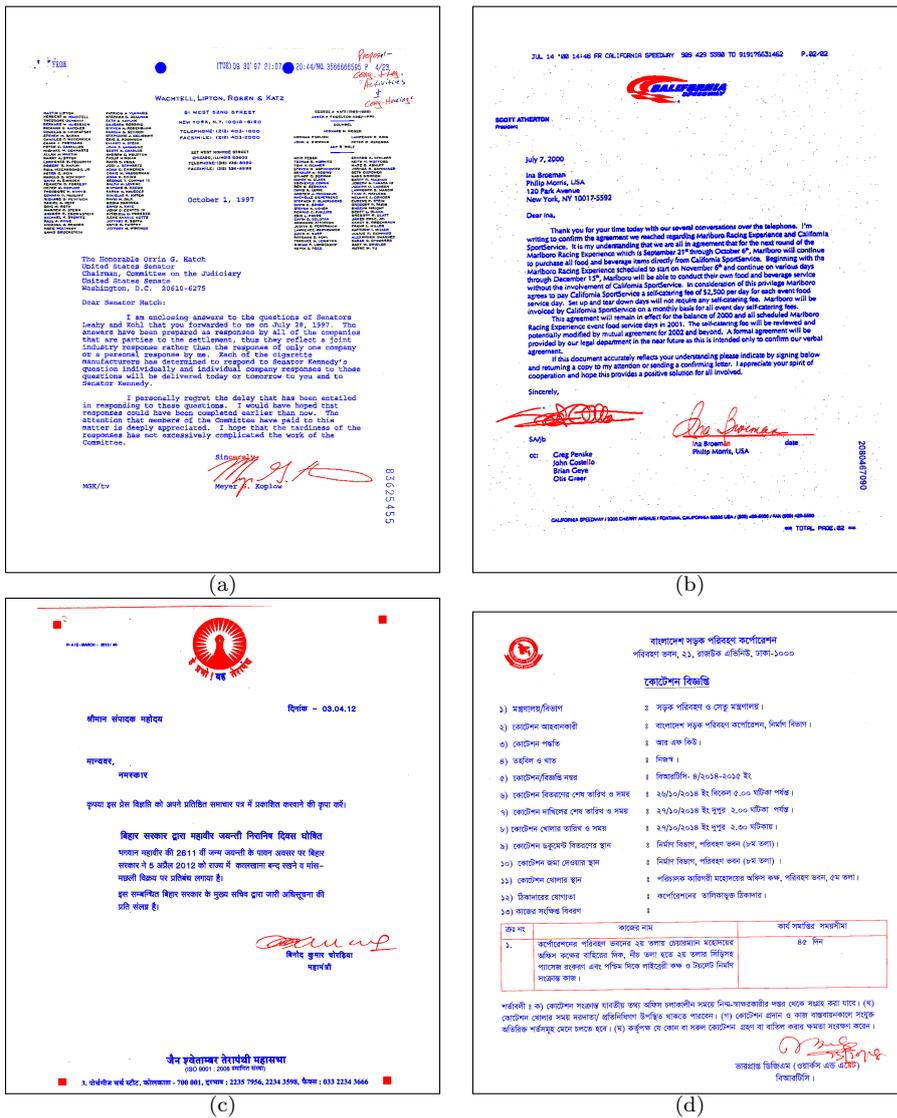


Fig. 6 Classification result of printed and signature/handwritten components on the documents shown in Fig. 1. (a,b) English ('Tobacco'), (c) Hindi and (d) Bangla . Printed text and signature/handwritten components are marked by blue and red respectively. PDF version of this paper is recommended as color codes are used in this figure for better visibility.

the optimal hyperplane which maximizes the distance, the margin, between the nearest examples of both classes, named support vectors (SVs). Given a training database of M data: $x_m | m = 1, \dots, M$, the linear SVM classifier is then defined as: $f(x) = \sum_j \alpha_j x_j + b$ where x_j are the set of support vectors and the parameters α_j and b have been determined by solving a quadratic problem. The linear SVM can be extended to various non-linear variants, and details can be found in [35]. In the experiments, the Gaussian kernel SVM outperformed other non-linear SVM

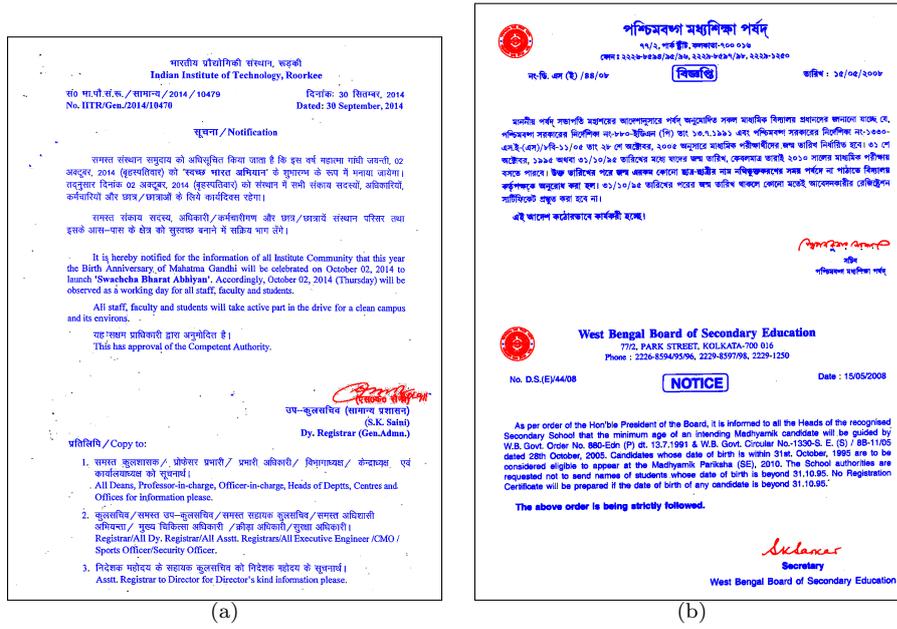


Fig. 7 Signature detection results from Indian bi-script official documents. (a) Results on bi-script document containing English and Devanagari scripts (b) Results on bi-script document containing English and Bangla scripts. Printed text and signature components are marked by blue and red colour, respectively. PDF version of this paper is recommended as color codes are used in this figure for better visibility.

kernels; hence the reported recognition results are based on the Gaussian kernel only. The hyperparameters of the SVM were set as follows; kernel type = RBF, $\gamma = 1$ and $C = 1$. The best results have been achieved by setting the above values of these parameters which were applied using a validation process. The Gaussian kernel is of the form:

$$k(x, y) = e^{-\gamma \|x - y\|^2} \quad (5)$$

The qualitative signature detection results on single script documents are shown in Fig. 6. Fig. 7 shows the signature detection results on two sample multi-script documents.

3.2 Grouping of Signature Components

After the separation of signature components from a document, multiple components might be present in the document. A signature can consist of one or more components and a document can contain more than one signature. Moreover, some misclassified non-signature components can also be present in the document. Therefore, all the components belonging to a signature were grouped, which is required to match with the query signature. To group signature components, first, corner points were computed from the document image and then a density-based clustering algorithm (DBSCAN [36]) was applied for discovering clusters of

points, which represent signature components. The algorithm computes the number of clusters starting from the estimated density distribution.

3.2.1 Corner points computation

First, corner points were computed from the components of a document using Harris-Stephens combined corner/thin edge detector algorithm [37] which is invariant to rotation, shift or even an affine change of intensity. The variance of light was computed using the local autocorrelation energy function:

$$E(x, y) = \sum_{u, v} W_{u, v} (I(x + u, y + v) - I(x, y))^2 \quad (6)$$

where (u, v) denote a neighborhood of (x, y) . A smooth Gaussian circular window with

$$W_{u, v} = \exp\left(-\frac{u^2 + v^2}{2\sigma^2}\right) \quad (7)$$

is the window function, and normally its value is 1, whereas $I(x + u, y + v)$ is the shifted intensity. Fig. 8 shows two sample signatures where corners points have been plotted using blue markers. Next, the co-ordinates of corner points are fed for processing by density-based spatial clustering.

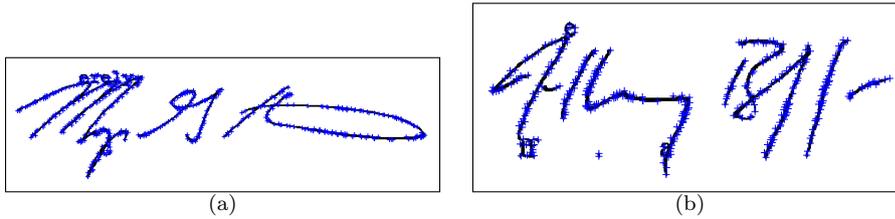


Fig. 8 (a,b) Signature images after computation of corner points. Blue markers represent corner points.

3.2.2 Density-based clustering

DBSCAN is a clustering algorithm proposed by Ester et al. [36], which finds the number of clusters starting from the estimated density distribution of corresponding nodes. It shows its efficiency on a large spatial database of synthetic data as well as real data by discovering the clusters of arbitrary shape. In comparison to other clustering algorithms, it requires minimal domain knowledge. The algorithm prerequisite is only one parameter i.e. distance threshold which is used to determine the maximum distance among points in a cluster and the algorithm also supports the user in determining the appropriate value for it.

In the component grouping work, an iterative method was used to set the threshold value, and for the iteration some clusters were computed from the corner points obtained from the segmented documents. First, a maximum threshold was computed for a density-based clustering algorithm which was based on the size of the query signature. 10 percent of the maximum threshold was used as an

initial threshold for the clustering in the first iteration. Next, the bounding boxes of all the clusters were computed and then the features from each of the clusters' bounding boxes were computed. The details of the cluster level feature extraction and matching techniques are described in Section 3.3. In the next step, those features were matched with the query signature's feature and stored the minimum matching distance obtained from this iteration. The distance threshold was increased by 10 percent for the next iteration. If the minimum matching distance from any iteration was larger than the previous one then the iteration was stopped and the minimum distance from the previous iteration was considered as the final minimum distance.

The step-by-step algorithm is presented in Algorithm 1. Although, the component grouping algorithm has scope for ten iterations, it was noticed from the experiments that signature components were properly grouped within the first three iterations. Fig. 9 shows some sample results from the signature component grouping experiment. In Fig. 9(a1) components are grouped into 6 clusters and the components of the actual signature are grouped into two clusters after the second iteration. Fig. 9(a2) shows the result after the second iteration where the actual signature components are grouped properly into one cluster.

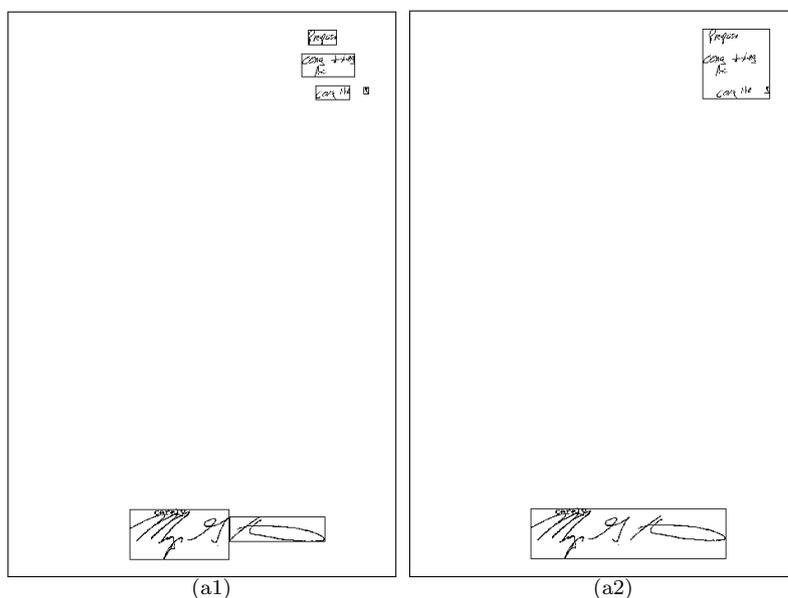


Fig. 9 Example of clustering results. Component clusters are shown after (a1) first (a2) second iterations on the document shown in Fig. 1(a).

3.3 Matching with the Query Signature

In this section, the signature shape encoding technique and matching procedure for the retrieval of documents are described. The encoding of signature images is

Algorithm 1 Grouping of signature components and matching with the query signature

Require: A query signature with the document to be matched

Ensure: Return a matching score with the query signature

```

\*Computation of Maximum Threshold (MaxTh) for DBSCAN clustering. Height and
Width refer to the query signature's height and width*\
Step 1:  $MaxTh \leftarrow \max(Height, Width)$ 
Step 2:  $InitTh \leftarrow MaxTh \times 0.1$ 
Step 3:  $Dist_{Match} \leftarrow -1$ 
Step 4:  $MinDist_{PreviousStep} \leftarrow -1$ 
for  $k \leftarrow InitTh$  to  $MaxTh$  step  $InitTh$  do
  Step 5:  $C \leftarrow DBSCAN(CornerPoints, k, MinPoints)$ 
  \*C refers to clustered corner points, CornerPoints refer to Harris-Stephens corner points
  computed from the documents and MinPoints refers to the minimum points threshold.
  The cluster bounding box refers to a rectangle computed using the boundary points of
  the cluster*\
  for Each Cluster in C do
    Step 6: Extract feature from cluster bounding box image
    Step 7:  $Dist \leftarrow FuncMatchDist(QuerySignature, TargetSignature)$ 
    if  $Dist_{Match} < 0$  then
      Step 8:  $Dist_{Match} \leftarrow Dist$ 
    else
      if  $Dist_{Match} \geq Dist$  then
        Step 9:  $Dist_{Match} \leftarrow Dist$ 
      end if
    end if
  end for
  if  $MinDist_{PreviousStep} < 0$  then
    Step 10:  $MinDist_{PreviousStep} \leftarrow Dist_{Match}$ 
  else
    if  $MinDist_{PreviousStep} > Dist_{Match}$  then
      Step 11:  $MinDist_{PreviousStep} \leftarrow Dist_{Match}$ 
    else
      Step 12: Return  $Dist_{Match}$ 
    end if
  end if
end for

```

almost the same as the proposed feature extraction technique described for signature components detection in Section 3.1. However, here the signature background information along with the foreground information was incorporated for encoding the signature.

3.3.1 Foreground-based feature

The shape coding technique of signatures also involves three steps as discussed in Section 3. First, to code the shape of the signature, the signature image is divided into densely sampled local patches and a descriptor has been computed from each of the patches. Here, signature images are divided into 900 (30×30) patches and one SIFT descriptor is computed from each patch. The number of patches determined in this stage is based on experimentation. Next, 900 SIFT-descriptors are used in the next process of computation of features based on codebook learning and a 3 level Spatial Pyramid Matching-based technique. Fig. 11(a1), Fig. 11(b1) and Fig.

11(c1) show 900 descriptor patches from three samples of foreground signatures namely English, Hindi and Bangla respectively.

3.3.2 Background-based feature

The cavity regions and loops in a signature are referred to as background information in this work. The cavity regions are obtained using the Water Reservoir concept [38]. The water reservoir in all four directions (top, bottom, left, right) and loops present in an image are used. Fig. 10 shows reservoirs from all four directions extracted from a signature. Here, the background signature image is also divided into 900 (30×30) patches and one SIFT descriptor is computed from each patch. Next, 900 SIFT-descriptors are used in the next step for computation of features using codebook learning and a Spatial Pyramid Matching-based technique. Fig. 11(a2), Fig. 11(b2) and Fig. 11(c2) show three sample signatures from English, Hindi, and Bangla, respectively, where the images are divided into 30×30 grid patches and the patch centers are marked. Finally, the foreground and background features are concatenated to get the final features.



Fig. 10 Loops and water reservoirs in three signature images are shown and red is used to mark reservoirs. The original signature, loops and the water reservoir from top, left, right and bottom sides are shown respectively in (a1- a6) for English, (b1- b6) Hindi and (c1- c6) Bangla signatures.

3.3.3 Distance between signature images

Three matching distances such as Euclidean distance, rank correlation, and DTW-based methods were considered for computation between the query signature and signatures from the document images. Given the two feature vectors $X_m | m = 1, 2, \dots, n$

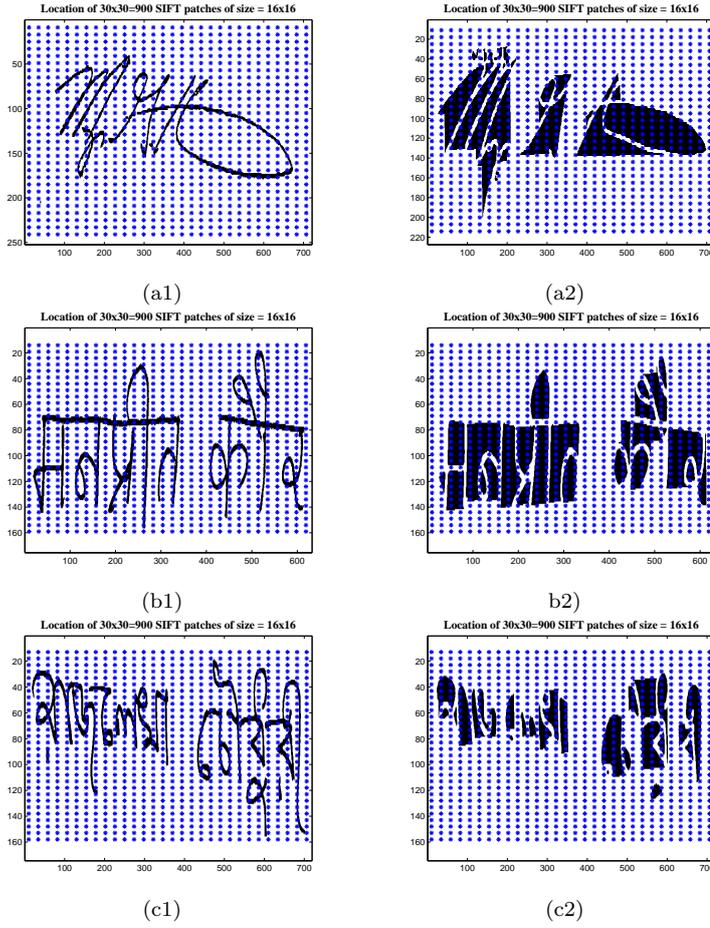


Fig. 11 (a1,b1,c1) Samples of English, Hindi and Bangla foreground signatures after grid-based 900 (30×30) SIFT patches are marked. (a2,b2,c2) Samples of background signatures after grid-based 900 (30×30) SIFT patches are marked on background information.

and $Y_m | m = 1, 2, \dots, n$, similarity distance between X and Y using the Euclidean distance is calculated using Equation 8. Equation 9 shows the formula for the linear correlation coefficient, which measures the strength and direction of a linear relationship between the vectors of a query signature and signatures from the documents.

$$Distance_{Euclidean}(X, Y) = \sqrt{\sum (X_i - Y_i)^2} \quad (8)$$

$$Corr(X, Y) = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{n(\sum X^2) - (\sum X)^2} \sqrt{n(\sum Y^2) - (\sum Y)^2}} \quad (9)$$

Here DTW is used on two sequences of feature vectors. The DTW distance between two vectors X and Y are calculated using a matrix D . Where

$$D(i, j) = \min \begin{pmatrix} D(i, j-1) \\ D(i-1, j) \\ D(i-1, j-1) \end{pmatrix} + d(x_i, y_j) \quad (10)$$

$$d(x_i, y_j) = \sum (X_i - Y_j)^2 \quad (11)$$

Finally, this matching cost was normalized by the length of the warping path. Here, it was observed that slant and skew angle of a signature class are usually constant but the larger variation normally lies in character spacing. DTW performed better in the experiments because of the flexibility to compensate such variations.

4 Results and discussion

This section evaluates the performance of different levels of the proposed approach by considering various measures. The different datasets used in the different level of experiments are described. Qualitative and quantitative results are detailed which shows the about the efficiency of the proposed approach.

4.1 Dataset

No standard dataset consisting of signatures and printed components of English, Devanagari and Bangla scripts exists to train the SVM classifier at the signature detection stage. Hence, a dataset has been created using components of English, Devanagari and Bangla scripts. Printed components were extracted from different types of documents such as newspaper, books, magazines etc. English signatures used in the experiment were extracted from the ‘Tobacco’ dataset. The Hindi and Bangla signatures used for training the SVM classifier were taken from the dataset created by Pal et al. [39]. The signatures were collected from 300 and 200 writers of Hindi and Bangla, respectively.

Table.1 shows the details of the training and test data used in the proposed experiments. It should be noted that the training and test datasets were different in the experiments. 7390 and 5854 components of printed and signature/handwriting respectively have been used from the English script to train the SVM classifier for the signature detection experiment on English documents. Likewise, 7670 and 5618 components of printed and signature/handwriting from Devanagari script and 5575 and 6950 components of printed and signature/handwriting from Bangla script were used. These components were also used to train the classifier for bi-script document classification (i.e. documents shown in Fig. 2). The document retrieval system was tested on three sets of document data for the three scripts considered in this experiment. The ‘Tobacco’ dataset was used for testing the system on English scripts. A database of 560 official notices and letters written in Devanagari, Bangla, and bi-lingual scripts has also been created. 300 documents of Devnagari and 260 documents of Bangla script are present in the collected dataset. The dataset of logos from the Laboratory for Language and Media Processing, University of Maryland [40] along with 400 downloaded logos has been used for document retrieval experiments based on logo information. A few samples of logos are presented in Fig. 12.

Table 1 The dataset used for training and testing the SVM classifier for signature detection.

Training Data			
Types of Data	English	Hindi	Bangla
Printed components	7390	7670	5575
Signature/Handwritten components	5854	5618	6950
Logos	106+400	-	-
Test Data			
Full page documents	'Tobacco' [1]	300	260

**Fig. 12** Few samples from logo dataset.

4.2 Performance Evaluation

The signature detection experiments on the ‘Tobacco’ dataset demonstrate the excellent performance of the proposed approach. Accuracy rates of 99.68%, 99.94%, and 99.97% were obtained in signature detection experiments from English, Devanagari and Bangla scripts, respectively. The accuracy rate of 99.21% was obtained from the experiments of the multi-script (English, Devanagari and Bangla) combined dataset. The ratio between True Positive Rate (TPR) and False Positive Rate (FPR) (i.e. Receiver Operating Characteristic (ROC) curve) obtained from the signature detection experiment is presented in Fig. 13. Fig. 13(a) shows the ROC curves obtained from the experiment on the ‘Tobacco’, Hindi and Bangla datasets. Fig. 13(b) shows the performance of signature/handwriting detection on the combined dataset of English, Hindi, and Bangla. Table 2 shows the confusion matrix of a classification among printed text, handwritten text, and signature. This experiment helps to understand that 2% handwritten texts are wrongly classified as the signature if handwritten text and signature are considered as separate classes.

Table 2 Confusion Matrix: Printed Text, Handwritten Text and Signature classification

	Printed Text	Handwritten Text	Signature
Printed Text	0.99	-	-
Handwritten Text	-	0.98	-
Signature	-	0.02	1.00

All the documents from three document datasets and the signature images were used for the experimentation of the proposed system for signature retrieval. Four separate experiments were carried out on English, Devanagari, Bangla and the combined dataset of all the three scripts. Three different features based on the foreground, background and combined information of foreground and background

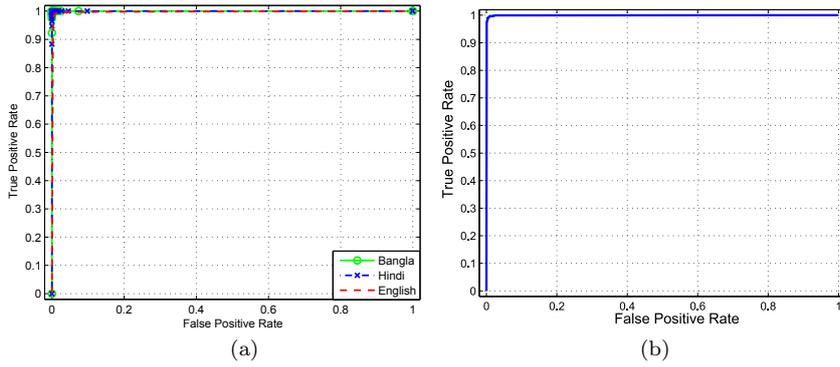


Fig. 13 (a) ROC curves obtained from signature/handwritten detection experiment on English (Roman), Devanagari (Hindi) and Bangla single script documents. Here, ROC curves on English, Devanagari and Bangla are almost overlapped because of the similar accuracy. (b) ROC curves obtained from signature/handwritten detection on multi-script documents of the combined dataset.

have been used in this work. Moreover, the signature retrieval performances based on three different distances have been measured for each case. Fig. 14, Fig. 15 and Fig. 16 show the precision-recall curves on English, Hindi, and Bangla documents respectively using Correlation, Euclidean, and DTW-based distance measures. Fig. 17 shows the precision-recall curve on multi-script documents using all the three distance measures employed for the scripts individually. It was noticed from the experiment that features containing combined information of foreground and background outperformed the performance of features that either contained only foreground or background information.

As an example, English script Fig. 14 shows 91.84% precision and a recall of 82.57% have been obtained from the foreground information when a linear correlation threshold was set to 0.63. An overall precision of 92.07% and recall of 85.32% were achieved on the same dataset using background information when the threshold for linear correlation is fixed to 0.59. Finally, 92.23% precision and 87.15% recall were obtained from the combined information of foreground and background when the linear correlation threshold was fixed to 0.60. It should be noted that there is a basic difference in the pattern of English signatures with non-English Indian script signatures. In signatures of Indian scripts in our dataset, we found many character components, whereas in English signatures we found fewer characters are used to represent the whole signature. Thus, during DTW, profile information of Hindi and Bangla signatures is richer than English signatures which leads to better performance.

4.3 Comparison with other systems

The previously proposed approaches on signature detection and recognition were tested on different publicly available datasets such as ‘Tobacco’ and a few experiments were conducted on the dataset on Hindi and Bangla scripts. Table 3 shows the performance of the previously proposed approaches on signature detection from documents. In [17], the result was reported in two stages: signature detection

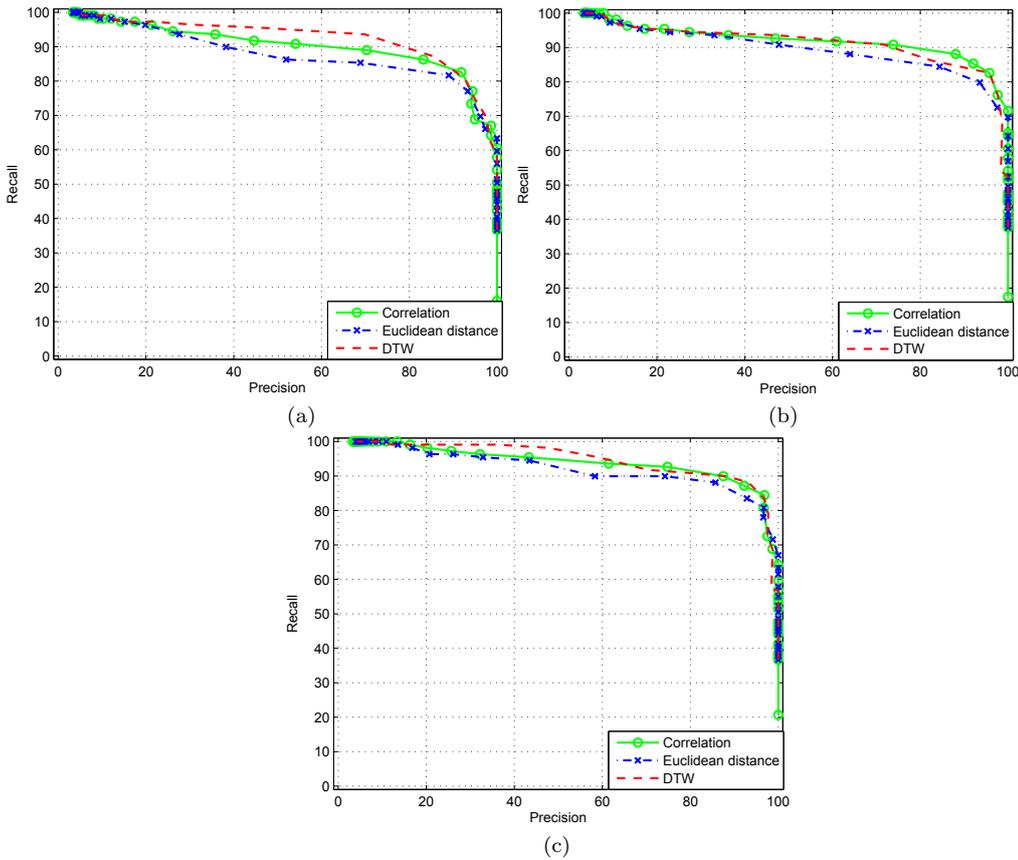


Fig. 14 Precision-Recall curves of signature retrieval on English (Roman) script using (a) Foreground information (b) Background information and (c) Combined information of foreground and background. Three measures such as Correlation, Euclidean and DTW distance have been applied for all cases.

and signature matching. A 92.8% accuracy was reported on the ‘Tobacco’ dataset for signature detection using a multi-scale structural saliency-based [17] approach. After signature detection, signature matching was performed with a dissimilarity measure. With a combination of dissimilarity measures, the best matching accuracy MAP (Mean Average Precision) obtained was 90.5%. Though, there was no report of the full signature retrieval result, theoretically, the combination of detection and matching results would provide approximately 84% ($92.8\% \times 90.5\%$) MAP as 92.8% accuracy was obtained for detection and 90.5% for matching. A recall of 78.4% and a precision of 84.2% were reported by Srinivasan and Srihari [18] for the signature-based document retrieval task. A 96.13% accuracy (298 signatures were correctly identified out of 310 documents) was reported by [16] on signature detection from Arabic/Persian documents. In the previously proposed approach [41], 95.58% accuracy was achieved on signature components detection. The gradient-based features and the SVM classifier were applied on the patch-wise classification of signatures and printed text from signed documents. In addition,

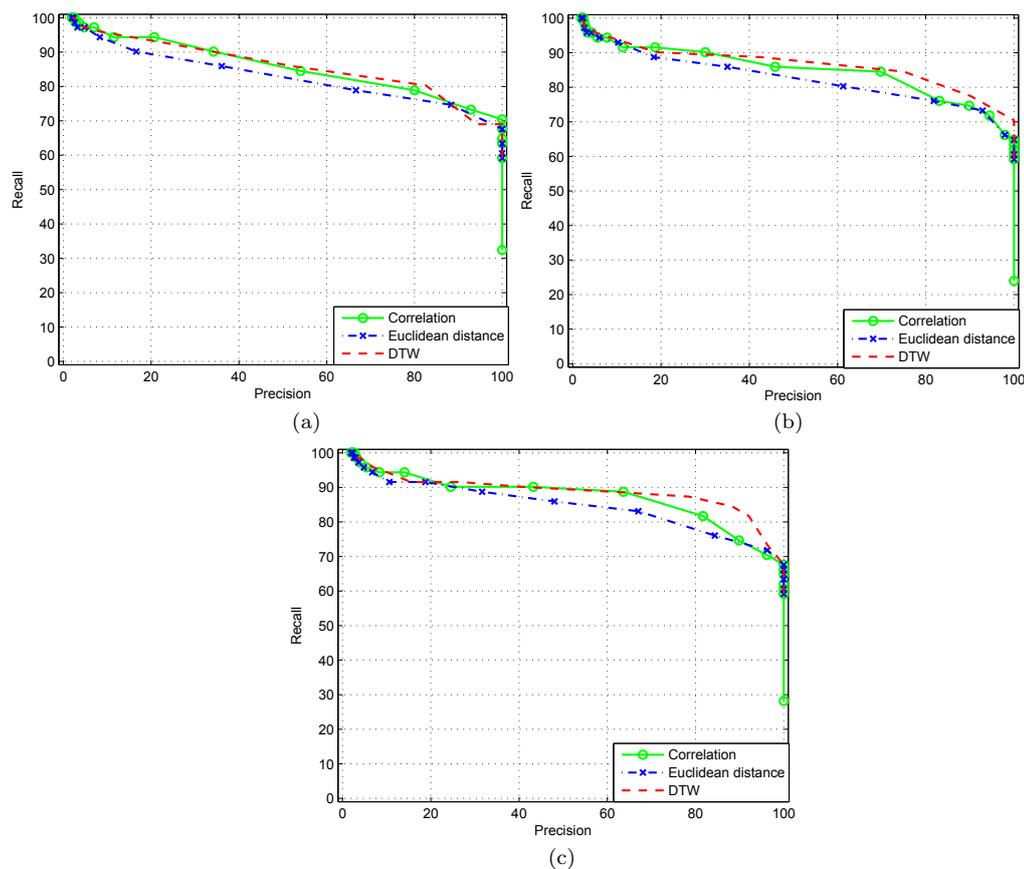


Fig. 15 Precision-Recall curves of signature retrieval on Devanagari script using (a) Foreground information (b) Background information and (c) Combined information of foreground and background. Three measures such as Correlation, Euclidean and DTW distance have been applied for all cases.

signature-based document retrieval using SURF/SIFT features with RANSAC-based matching was implemented but unfortunately performed poorly (precision was below 20%) when the parameter values of Transform Type and Max Distance were set to affine and 10 respectively. The reason for the poor performance is due to the large variation among handwritten strokes, which exists among samples from the same signature class. These variations were not captured properly using a traditional SIFT/SURF-based method.

Though the proposed system achieves better accuracy than previously proposed approaches, the primary advantage is that the feature extraction technique is simpler and more robust than previous methods and works in a multi-script environment. The proposed system does not need pre-processing or noise correction of signature portions for matching in an earlier stage. The empirical results of the experiments are encouraging and compare well with other state-of-the-art approaches in the literature.

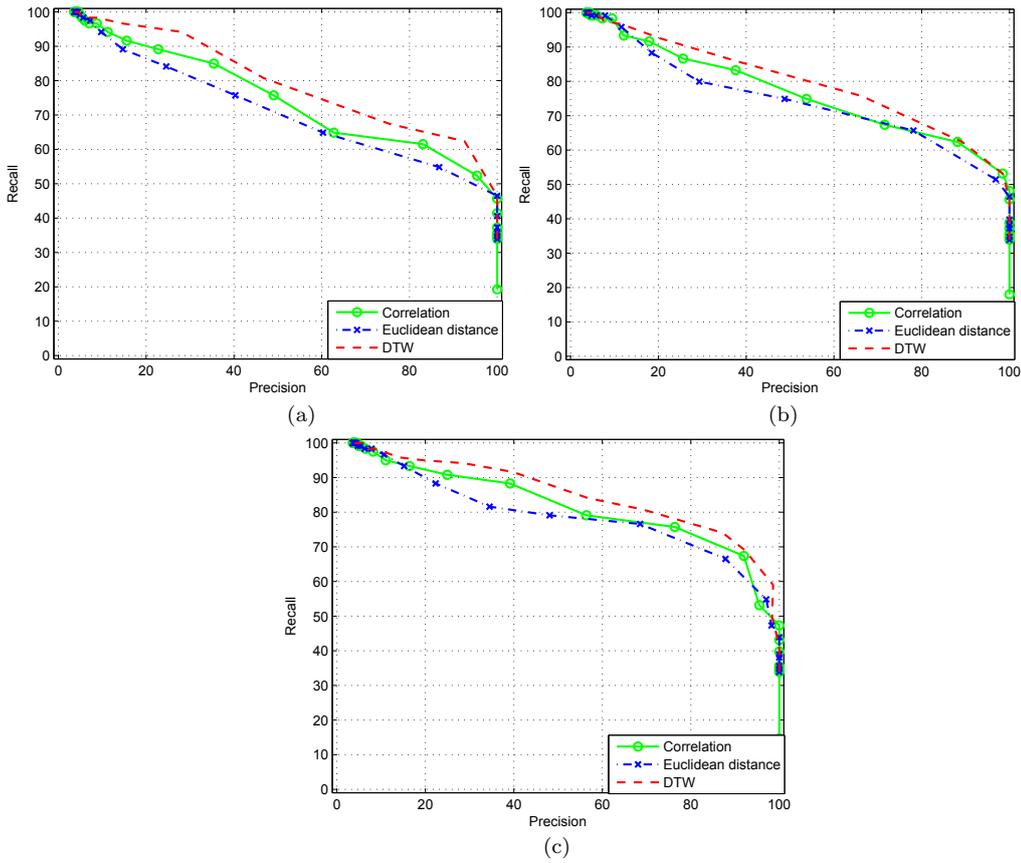


Fig. 16 Precision-Recall curves of signature retrieval on Bangla script using (a) Foreground information (b) Background information and (c) Combined information of foreground and background. Three measures such as Correlation, Euclidean and DTW distance have been applied for all cases.

Table 3 Comparison of signature detection performance on ‘Tobacco’ document repository.

Approach	Dataset	Accuracy (%)
Multi-scale structural saliency [17]	Tobacco-800	92.80
Conditional Random Field [18]	101 documents	91.20
Gradient-based feature with SVM [41]	Tobacco-800	95.58
Proposed Method	Tobacco-800	99.68

4.4 Error Analysis

Here, some errors are described that resulted from the experiments. In the signature detection stage, some printed components such as logos, seals, and figures were incorrectly classified as handwritten/signature components. It is to be noted that small components such as small dots were ignored in this classification stage

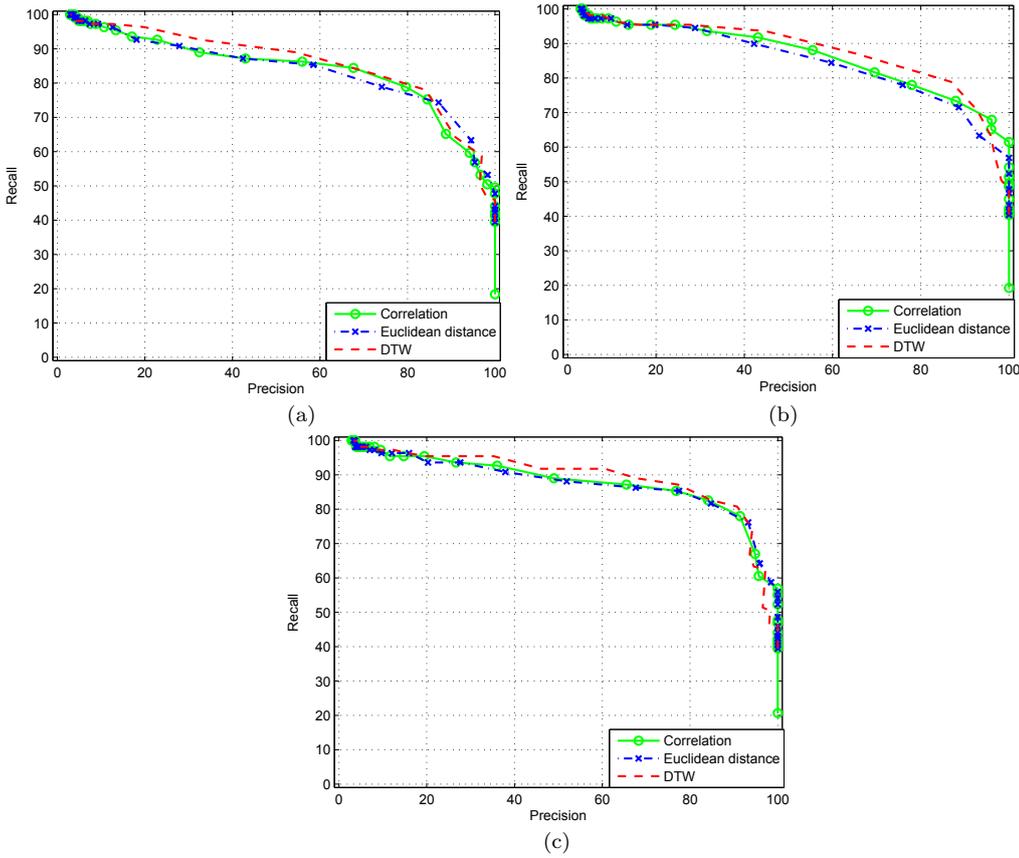


Fig. 17 Precision-Recall curves of signature retrieval on a multi-script dataset (English (Roman), Devanagari and Bangla) using (a) Foreground information (b) Background information and (c) Combined information of foreground and background. Three measures such as Correlation, Euclidean and DTW distance have been applied to all cases.

and the average stroke-width of the components-based threshold values were used. Since in the experiments, only two classes (text and non-text) were considered, the graphical components shown in Fig. 18 were identified as non-text. Here, signatures and graphical components were all considered as a non-text class.

Table 4 and Table 5 show Type I and Type II errors respectively obtained from the signature retrieval step in the experiments. The first column of Table 4 shows two sample query signatures and the second row of Table 4 shows the retrieved signatures present in the target documents. Although query signatures and retrieved signatures belong to different classes, the correlation between query signatures and retrieved signatures are high. Likewise, Euclidean and DTW distances are low among these samples.

Similarly, Table 5 shows similarity measures of two different sets of signatures. Samples of query signatures are written in a slanted style whereas signatures present in the target documents are written in a standard style. As a result, the correlation measure is low between the query signature and the retrieved signa-

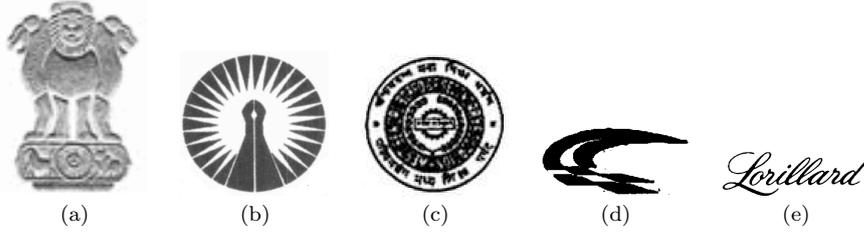


Fig. 18 Samples of logos and printed components recognised as signatures or handwritten components.

Table 4 Three different distances among different signature samples show Type I error (false positive) cases in the signature retrieval experiments.

Signature Samples	Correlation		Euclidean Distance		DTW	
	<i>R.M. Neel</i>	<i>R.M. Neel</i>	<i>R.M. Neel</i>	<i>R.M. Neel</i>	<i>R.M. Neel</i>	<i>R.M. Neel</i>
<i>Thomas Willert</i>	0.606	0.604	1.141	1.140	38.665	38.705
<i>Wolke</i>	0.608	0.612	1.132	1.123	38.818	38.745

Table 5 Three different distances among different signature samples show Type II error (false negative) cases in signature retrieval experiment.

Signature Samples	Correlation		Euclidean Distance		DTW	
	बाबुराज गंध	इम जिते कुरु	बाबुराज गंध	इम जिते कुरु	बाबुराज गंध	इम जिते कुरु
बाबुराज गंध	0.508	0.545	1.765	1.692	44.386	42.387
इम जिते कुरु	0.528	0.557	1.725	1.666	43.673	41.669

ture in the target document belonging to the same class. Likewise, the Euclidean distance and the DTW distance are high. So, a Type II error occurs in this case.

4.5 Experiments on noisy documents

To evaluate the robustness of the proposed system on noisy documents, a synthetic noisy document dataset was created. Gaussian noises of two different variances (i.e. 0.005 and 0.01) were applied on the ‘Tobacco’ database for this work. Fig. 19(a) and Fig. 19(b) show the same document with Gaussian noise of 0.005 and 0.01 variances respectively. The qualitative performance of the signature detection results on these two sample noisy documents is shown in Fig. 19(c) and Fig. 19(d) respectively.

Fig. 20(a) shows the ROC curve obtained based on the experiments for signature detection from noisy document images. The area under the curve was 99.91%. The accuracy dropped by 0.74% (98.94% accuracy was obtained in contrast to

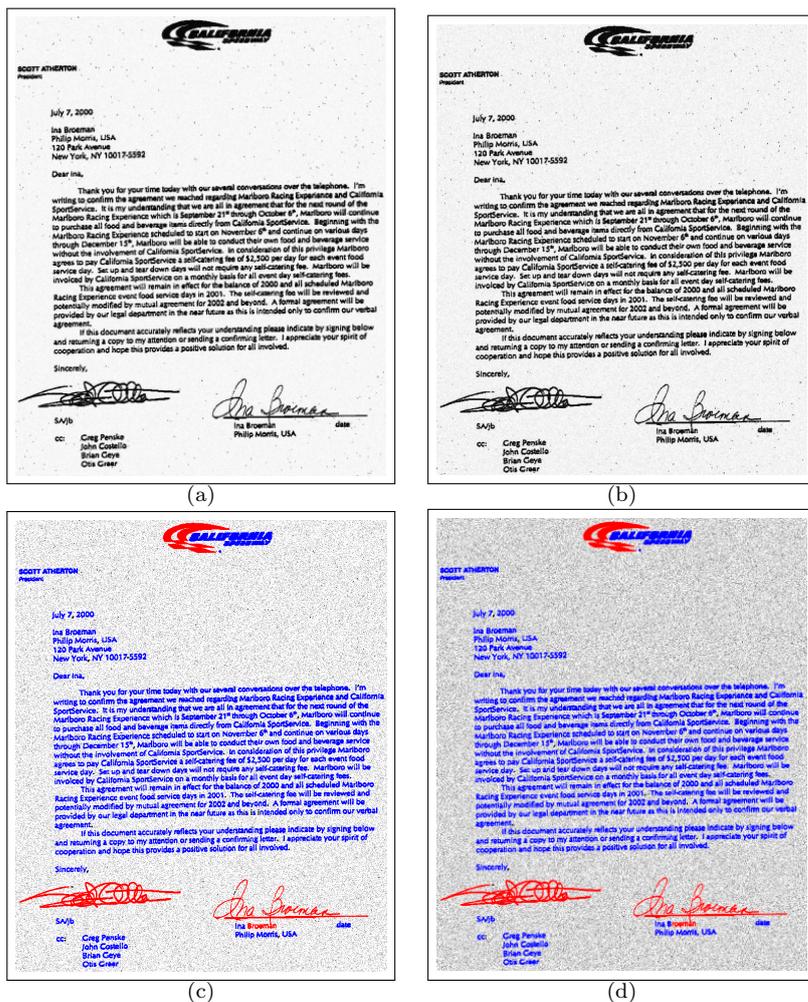


Fig. 19 Samples of English official documents after addition of Gaussian noise (a) with variance 0.005 (b) with variance 0.01 (c,d) signature detection results on the binary version of (a) and (b), respectively. Printed text and signature components are marked by blue and red, respectively. PDF version of this paper is recommended as color codes are used in this figure for better visibility.

99.68% in the experiment on normal, less noisy documents) in the experiments on synthetic noisy documents. Fig. 20(a) and Fig. 20(b) show precision-recall curves obtained from signature-based document retrieval experiments on noisy documents with different Gaussian noise. In this experiment, two different variances such as 0.005 and 0.01 were used to create the synthetic Gaussian noisy document images. The performance of the system during the retrieval stage decreased by approximately 8% in comparison to the original documents. The Gaussian noise affects the 16×16 pixel grids and the computation of the SIFT-descriptors is also affected. Thus, this is the reason that the performance has dropped.

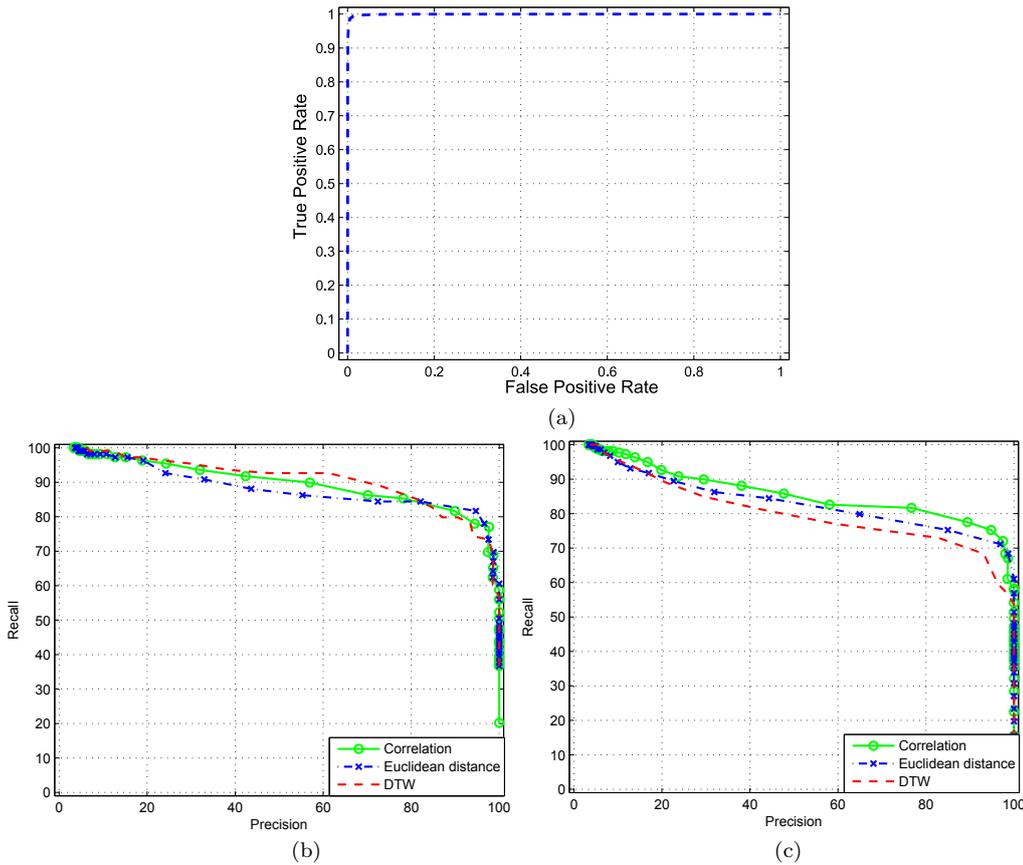


Fig. 20 (a)ROC curve obtained from the experiment of signature detection from Gaussian noisy 'tobacco' documents. Precision-Recall curves of signature retrieval on Gaussian noisy dataset (b) with variance 0.005 (c) with variance 0.01. Three measures such as Correlation, Euclidean, and DTW distance were taken for all cases.

4.6 Document retrieval based on logo information

As stated earlier in Section 1, an experiment on logo-based retrieval was performed and the outcomes of the experiments are presented using the ROC curves. Fig. 21(a) shows three ROC curves obtained from the experiments of logo detection from documents. The area under the ROC curves quantifies the overall performance obtained from the experiments. In the logo detection experiment, three different cases were considered. The first experiment was a two-class problem where classes contain logos and printed text and no classification errors were obtained. The second experiment also contained two classes. The printed and handwritten components were kept in one class and the other class contained logos. A 99.61% accuracy was obtained from this experiment for logo detection. Finally, in the third experiment logos, printed text, and signature/handwritten texts were considered as three different classes and an accuracy of 98.46% was achieved. It was observed

that 5.5% and 1.38% of logos were confused as signature/handwritten text and printed text, respectively.

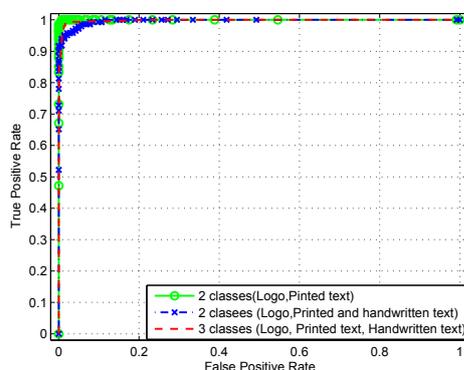


Fig. 21 ROC curves obtained from the experiments of detection of logos from documents.

The background information of logos is not always present. So, only the foreground information was used in the experiments of document retrieval based on logo information. The precision was always 100% for all recall values. Different recall values based on different thresholds are presented in Table 6.

Table 6 Threshold vs. Recall from logo-based document retrieval using three similarity measures (Correlation, Euclidean Distance, and DTW).

Similarity Measure: Correlation				
Threshold	0.30	0.25	0.20	0.15
Recall(%)	88.92	95.30	98.13	99.41
Similarity Measure: Euclidean Distance				
Threshold	1.71	1.61	1.51	1.41
Recall(%)	99.76	98.14	91.95	74.55
Similarity Measure: DTW				
Threshold	71.71	61.61	55.55	53.54
Recall(%)	99.89	99.55	90.07	76.89

Table 7 shows the comparative study of logo detection and recognition performance on the ‘Tobacco’ document dataset. The proposed approach outperformed the recently proposed approaches on logo detection and recognition. 99.50% (99.61% × 99.89%) overall accuracy was achieved on logo detection from the ‘Tobacco’ dataset. Here, the best accuracy obtained from the experiments was considered for the comparison with the recently proposed approaches. The experiments were performed in a system of Core i5 2.5GHz CPU with 8GB of RAM. Matlab environment was used for the implementation. The proposed algorithm takes approximately 1.5 to 2.0 seconds to detect the signature in a document and 0.000062 seconds for matching using the correlation technique. However, the performance can be improved with C++ environment and fine tuning of the algorithm.

Table 7 Comparison of logo detection and recognition performance on the ‘Tobacco’ document repository.

Approach	Detection Accuracy (%)	Recognition Accuracy (%)	Overall Performance(%)
Alaei and Delalandre [25]	99.31	97.90	97.22
Wang [24]	94.70	92.90	87.98
Proposed Method	99.61	99.89	99.50

5 Conclusion

A novel end-to-end architecture for handwritten signature detection and matching for signature-based document retrieval is proposed in this paper. A component-wise bag-of-visual-words-based feature extraction powered by SIFT descriptors and an SVM-based classification technique achieved a high accuracy on signature detection. The proposed Spatial Pyramid Matching-based feature extraction technique is proved to be robust and has high discriminative features as it concatenates global and local features. Experiments on three languages (i.e. English, Hindi, and Bangla) were conducted to show that the system works in the multi-script environment. In addition to signatures, an experiment of document retrieval based on logos was performed. The proposed approach produced encouraging results due to its robustness to signature and logo variability even though it retains its simplicity. The experimental outcomes from the logo-based retrieval show the genericness of the architecture. Finally, the combined feature derived from foreground and background information leads to significant improvement in the signature matching stage.

The following contributions were achieved by the proposed work:

- A complete end-to-end system comprising of three steps which outperformed the state-of-the-art approaches
- Spatial Pyramid Matching-based method for signature detection achieved higher performance
- The genericness property has been validated by the experimental results when applied for Logo detection and matching.
- Finally, the signature’s background and foreground information together for feature extraction leads to a significant improvement in signature recognition accuracy

Conflict of Interest: The authors declare that they have no conflict of interest.

References

1. <http://legacy.library.ucsf.edu/>. The Legacy Tobacco Document Library (LTDL). University of California, San Francisco, 2007.
2. C. Y. Suen, Q. Xu, and L. Lam. Automatic recognition of handwritten data on cheques - Fact or fiction? *Pattern Recognition Letters*, 20:1287–1295, 1999.
3. S. Levy. “google’s two revolutions”. Newsweek, <http://www.newsweek.com/googles-two-revolutions-123507>, 2004.

4. P. P. Roy, E. Vazquez, J. Lladós, R. Baldrich, and U. Pal. A system to segment text and symbols from color maps. In *Proc. International Workshop on Graphics Recognition (GREC)*, pages 245–256, 2008.
5. G. Zhu and D. Doermann. Logo matching for document image retrieval. In *Proc. International Conference on Document Analysis and Recognition (ICDAR)*, pages 606–610, 2009.
6. G. Zhu, S. Jaeger, and D. Doermann. A robust stamp detection framework on degraded documents. In *Proc. of SPIE Conference on Document Recognition and Retrieval*, pages 1–9, 2006.
7. F. Farooq, K. Sridharan, and V. Govindaraju. Identifying handwritten text in mixed documents. In *Proc. International Conference on Pattern Recognition (ICPR)*, pages 1–4, 2006.
8. J.K. Guo and M.Y. Ma. Separating handwritten material from machine printed text using Hidden Markov Models. In *Proc. International Conference on Document Analysis and Recognition (ICDAR)*, pages 439–443, 2001.
9. J. Kumar, R. Prasad, H. Cao, W. Abd-Almageed, D. Doermann, and P. Natarajan. Shape codebook based handwritten and machine printed text zone extraction. In *Proc. SPIE*, volume 7874, page doi:10.1117/12.876725, 2011.
10. X. Peng, S. Setlur, V. Govindaraju, R. Sitaram, and K. Bhuvanagiri. Markov Random Field-based text identification from annotated machine printed documents. In *Proc. International Conference on Document Analysis and Recognition (ICDAR)*, pages 431–435, 2009.
11. Y. Zheng, H. Li, and D. Doermann. The segmentation and identification of handwriting in noisy document images. In *Proc. Document Analysis Systems (DAS)*, pages 95–105, 2002.
12. M. Martinez-Diaz, J. Fierrez, R.P. Krish, and J. Galbally. Mobile signature verification: feature robustness and performance comparison. *IET Biometrics*, 3, 2014.
13. J. Galbally, M. Diaz-Cabrera, M. A. Ferrer, M. Gomez-Barrero, A. Morales, and J. Fierrez. On-line signature recognition through the combination of real dynamic data and synthetically generated static data. *Pattern Recognition*, 48 (9):2921–2934, 2015.
14. D. Morocho, A. Morales, J. Fierrez, and R. Vera-Rodriguez. Towards human-assisted signature recognition: improving biometric systems through attribute-based recognition. In *Proc. International Conference on Identity, Security and Behavior Analysis (ISBA)*, 2016.
15. M. Blumenstein, Miguel A. Ferrer, and J.F. Vargas. The 4NSIGCOMP2010 off-line signature verification competition: Scenario 2. In *Proc. International Conference on Frontiers in Handwriting Recognition (ICFHR)*, volume 4, pages 721–726, 2010.
16. A. Chalechale, G. Naghdy, and A. Mertins. Signautre-based document retrieval. In *Proc. International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 597–600, 2003.
17. G. Zhu, Y. Zheng, D. Doermann, and S. Jaeger. Signature detection and matching for document image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 31(11):2015–2031, 2009.
18. H. Srinivasan and S. N. Srihari. Signature-based retrieval of scanned documents using Conditional Random Fields. *Computational Methods for Countert-*

- errorism*, pages 17–32, 2009.
19. P.P. Roy, S. Bhowmick, U. Pal, and J. Y. Ramel. Signature based document retrieval using GHT of background information. In *Proc. International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pages 225–230, 2012.
 20. R. Mandal, P.P. Roy, and U. Pal. Signature segmentation from machine printed documents using Conditional Random Field. In *Proc. International Conference on Document Analysis and Recognition (ICDAR)*, pages 1170–1174, 2011.
 21. X. Du, W. AbdAlmageed, and D. Doermann. Large-scale signature matching using multi-stage hashing. In *Proc. ICDAR*, pages 976–980, 2013.
 22. J.C. Brice no, C.M. Travieso, M.A. Ferrer, J.B. Alonso, and F. Vargas. Angular contour parameterization for signature identification. In *LNCS EUROCAST*, volume 5717, 2009.
 23. H. Dewan, W. Xichang, and L. Jiang. A content-based retrieval algorithm for document image database. In *Proc. International Conference On Multimedia Technology (ICMT)*, pages 1–5, 2010.
 24. H. Wang. Document logo detection and recognition using Bayesian model. In *Proc. International Conference On Pattern Recogniton (ICPR)*, pages 1961–1964, 2010.
 25. A. Alaei and M. Delalandre. A complete logo detection/recognition system for document images. In *Proc. International Workshop on Document Analysis Systems (DAS)*, pages 324–328, 2014.
 26. A. Fischer, A. Keller, V. Frinken, and H. Bunke. Hmm-based word spotting in handwritten documents using subword models. In *Proc. International Conference on Pattern Recognition (ICPR)*, pages 3416–3419, 2010.
 27. V. Frinken, A. Fischer, R. Manmatha, and H. Bunke. A novel word spotting method based on recurrent neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 3(3):211–224, 2012.
 28. J.A. Rodríguez-Serrano and F. Perronnin. Handwritten word-spotting using hidden markov models and universal vocabularies. *Pattern Recognition*, 42(9): 2106–2116, 2009.
 29. F. Alhwarin, C. Wang, D. R. Durrant, and A. Gräser. Improved sift-features matching for object recognition. In *Proc. Vision of Computer Science*, pages 179–190, 2008.
 30. Y. Hua, J. Lin, and C. Lin. An improved sift feature matching algorithm. In *Proc. World Congress on Intelligent Control and Automation (WCICA)*, pages 6109–6113, 2010.
 31. W. Kai, C. Bo, and T. Long. An improved sift feature matching algorithm based on maximizing minimum distance cluster. In *Proc. International Conference on Computer Science and Information Technology (ICCSIT)*, pages 255–259, 2011.
 32. D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision (IJCV)*, 60(2):91–110, 2004.
 33. S. Lazebnik, C. Schmid, and J. Ponce. Beyond Bags of Features: Spatial Pyramid Matching for recognizing natural scene categories. In *Proc. Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 2169–2178, 2006.
 34. L. Fei-Fei and P. Perona. A bayesian hierarchical model for learning natural scene categories. In *Proc. Computer Vision and Pattern Recognition (CVPR)*, pages 524–531, 2005.

35. V. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, 1995.
36. M. Ester, H. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 226–231, 1996.
37. C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. Alvey Vision Conference (AVC)*, pages 147–151, 1988.
38. U. Pal, A. Belaid, and Ch. Choisy. Touching numeral segmentation using water reservoir concept. *Pattern Recognition Letters*, 24(1-3):261–272, 2003.
39. S. Pal, A. Alaei, U. Pal, and M. Blumenstein. Multi-script off-line signature identification. In *Proc. International Conference Hybrid Intelligent Systems (HIS)*, pages 236–240, 2012.
40. <http://lamp.cfar.umd.edu/>. Logo dataset. University of Maryland, Laboratory for Language and Media Processing (LAMP), 2014.
41. R. Mandal, P. P. Roy, and U. Pal. Signature segmentation from machine printed documents using contextual information. *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)*, 26(7), 2012.