Signature Segmentation and Recognition from Scanned Documents

Ranju Mandal*, Partha Pratim Roy†, Umapada Pal‡ and Michael Blumenstein*

*School of Information and Communication Technology, Griffith University, Queensland, Australia
ranju.mandal@griffithuni.edu.au, m.blumenstein@griffith.edu.au
†Synchromedia Lab, Automation Engineering Department, Ecole de technologie superieure, Montreal, Canada
2partharoy@gmail.com
‡Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata, India
umapada@isical.ac.in

Abstract—Signature as a query is important for content-based document image retrieval from a scanned document repository. This paper presents a two-stage approach towards automatic signature segmentation and recognition from scanned document images. In the first stage, signature blocks are segmented from the document using word-wise component extraction and classification. Gradient based features are extracted from each component at the word level to perform the classification task. In the 2nd stage, SIFT (Scale-Invariant Feature Transform) descriptors and Spatial Pyramid Matching (SPM)-based approaches are used for signature recognition. Support Vector Machines (SVMs) are employed as the classifier for both levels in this experiment. The experiments are performed on the publicly available "Tobacco-800" and GPDS [1] datasets and the results obtained from the experiments are promising.

Keywords-Document image retrieval; Signature segmentation; Signature recognition; Dense SIFT; Bag-of-Features; Spatial Pyramid Matching

I. INTRODUCTION

Signatures provide rich information about a person as they consist of unique properties of human behaviour, thus they are used for verification/authentication purposes. The signature has been considered for biometric authentication in administrative documents, legal documents, bank cheques, etc. In a document, signature may be examined by forensic document analysis experts for authenticating documents and to restrict fraud. It is a common organisational practice nowadays to store and maintain large 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. "Tobacco-800" [2] dataset) without adequate index information. As a consequence of that, such practice has created a tremendous demand for robust ways to access and manipulate the information that these images contain. Obtaining information resources relevant to the query information from such repositories is the main objective of document retrieval. A sample scanned document from the "Tobacco-800" dataset is shown in Fig.1. So, signatures could be used as key information for searching and retrieval of documents. Thus, the handwritten signature will undoubtedly add advantage for document indexing and searching.

Hence, segmentation and recognition of signatures from documents is very significant because of its various applications.
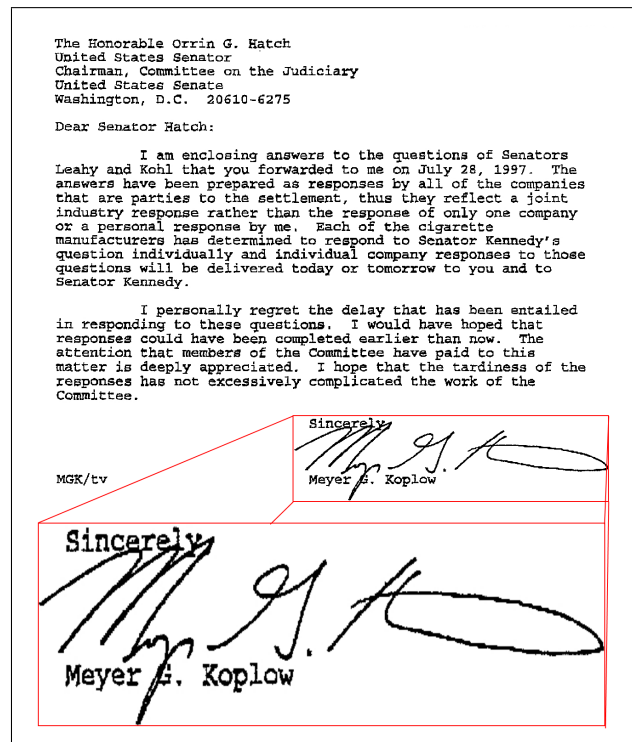

Figure 1. A sample printed signed document from "Tobacco-800" dataset. A zoomed version of the signature component is shown below the document for better visibility. It can be noted that there are touching printed characters with signature component.

In the literature, segmentation and recognition of signatures from scanned documents have been found to be a challenging task. Separation of handwritten annotations/words [3], [4], [5], [6], [7], [8] from scanned documents is focused on in most of the earlier works. Signature detection in scanned documents is discussed in previous works [9]. Zhu et al. [9] proposed a multi-scale structural saliency approach to capture the dynamic curvature using a signature production model for signature detection and segmentation. Signature segmentation techniques from machine printed

documents have been proposed in some works [10], [11], [12]. To segment signatures from bank cheques and other documents Madasu et al. [10] proposed an approach based on sliding window to calculate the entropy and finally fit the window to signature block. A major deficiency of this technique is that a priori information about the location of the signature is assumed. Ahmed et al. [12] proposed a Speeded Up Robust Features (SURF) based approach for signature segmentation from document images.

Signature-based document retrieval methods have been discussed in a few proposed works [13], [14], [15], [16]. Chalechale et al. [13] describes a method for document image decomposition and retrieval based on connected component analysis and geometric properties of the labelled regions. Documents having Arabic/Persian signature are considered for the experiment. Srinivasan and Srihari [14] proposed a method on signature-based retrieval of scanned documents. A model based on Conditional Random Fields (CRF) is used to label extracted segments of scanned documents as machine-printed, signature and noise. Next, a classification technique based on Support Vector Machine (SVM) is used to remove noise and printed text overlapping the signature images. Finally, a global shape-based feature is computed for each signature image. In [16], Roy et al. presented a signature based document retrieval technique from documents with cluttered background. Zernike Moment feature is extracted from each blob and the K-means clustering algorithm is used to create the codebook of blobs. During retrieval, Generalized Hough Transform (GHT) is used to detect the query signature and a voting is casted to find possible location of the query signature in a document.

This paper proposes a two-stage approach for signature segmentation and recognition. Gradient-based features and the SVM classifier are used for signature segmentation. SIFT descriptors with an extended version of the Bag-of-Features (BoF) algorithm is employed for signature recognition task. The rest of the paper is organized as follows. In Section II, we explain the proposed signature segmentation and recognition methodology. The experimental results are presented in Section III. Finally, conclusions are presented in Section IV.

II. PROPOSED METHODOLOGY

As mentioned earlier, the proposed approach comprises of two stages; first stage is dedicated for signature segmentation from a document image and then signatures are recognised. The steps involved in the proposed work are explained as follows.

A. Signature Segmentation

A signature generally consists of some large strokes compare to the strokes of the printed text. So, this distinct feature of signature is very important to get the difference of signature from printed strokes and it is used to discriminate signature from printed text. The proposed method describes a signature segmentation from a machine printed scanned document. A two-step procedure has been proposed for signature segmentation. At first, signature blocks are extracted. The signature blocks are further processed to remove non-signature components, such as touching and/or overlapping printed name and affiliations of signatory. Block-wise word extraction, word level feature extraction and classification techniques are described below.

The scanned documents are in grayscale and Otsu's threshold selection method is used to convert these document images into two-tone (0 and 1). Afterwards, Hough transform-based methods have been used to correct the skew of our documents. Now, the binarized document images are segmented into words based on the inter-character spacing between words. A morphological dilation operation using a 5×5 structuring element is performed and a connected component labelling method is applied to find the bounding boxes of the word patches on the dilated document image.
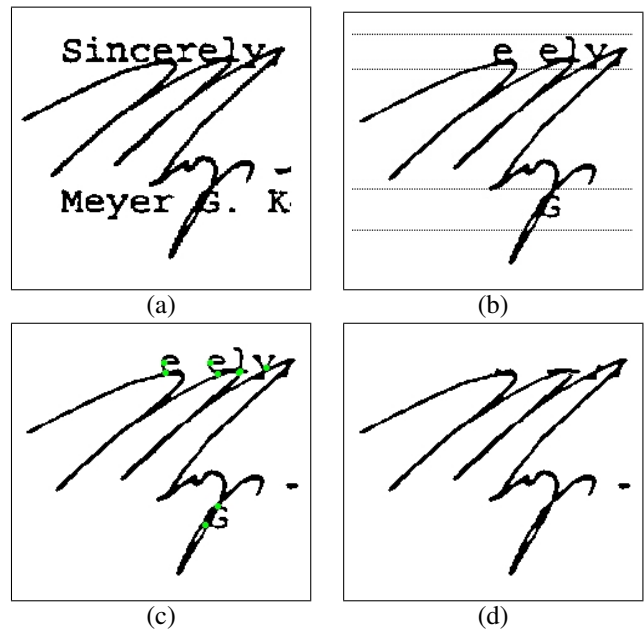

Figure 2. (a) Extracted signature block (b) Signature block after computation of printed zones and removal of isolated printed characters (c) Signature after selection of interest points (i.e. junction points belongs to printed character zones, which are marked by green dots) for touching character segmentation (d) Signature after touching character segmentation.

Next, we compute word-block level features of the segmented word patches and feed the feature vector into an SVM classifier [17]. A robust gradient based feature extraction technique and SVM as classifier have been used to classify those segmented words as signature or printed words. The grey-scale local-orientation histogram of the component is used for gradient-based feature [18] extraction. The classification result based on Gaussian kernel has been reported. The Gaussian kernel is of the form:

$$k(x,y) = exp\frac{-\|x - y\|^2}{2\sigma^2}$$

The printed text characters which are included in the signature block and are touched/overlapped with the signature are removed using touching character analysis in hypothetical zones of printed text characters. Sample resultant images of the signature are presented in (Fig. 2), which show signature blocks before and after touching/overlapping character removal. The detailed method of touching character removal has been described in [11].

B. Signature Recognition

An efficient patch-based SIFT descriptors with Spatial Pyramid Matching (SPM)-based pooling scheme is applied for the proposed signature recognition task. The feature extraction module has three components. A flow diagram of signature recognition system is presented in Fig. 3. First, SIFT descriptors are extracted from the signature and quantised using the K-means clustering algorithm. Next, the SPM-based scheme is applied for the representation of an image. Finally, the SVM is employed for classification. The modules are described in the following sub-sections.
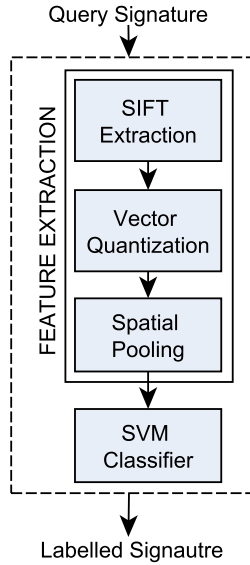

Figure 3.   Flow diagram of signature recognition module.

1) SIFT descriptor: The SIFT (Scale-Invariant Feature Transform) [19] is a local shape descriptor to characterize local gradient information. Here, 128-dimensional vector for each SIFT 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 circles in Fig. 4(a) represent the 16×16 SIFT patches and Fig. 4(b)

shows graphical representation of SIFT descriptors of 4 bins and 8 orientations.
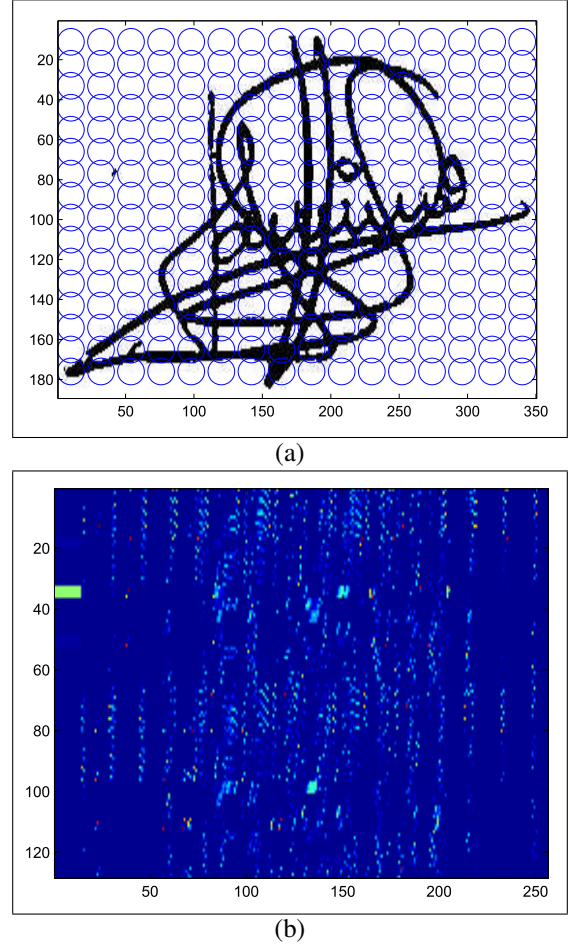

Figure 4.   (a) Blue circles represent 256 (of 16x16) SIFT patches (b) SIFT descriptors with 4 bins and 8 orientations

2) Spatial Pyramid Matching (SPM): The SPM is an extended version of Bag-of-Features (BoF) model, which is simple and computationally efficient. As BoF model discards the spatial order of local descriptors, it restricts the descriptive power of the image representation. The limitation of BoF is vanquished by SPM [20] approach, which is successfully applied on image categorization tasks. An image is partitioned into $2^l \times 2^l$ segments where $l = 0, 1, 2, ..., n$; 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 reduces to BoF, when the value of the scale $l = 0$. Here, the pyramid matching is performed in two-dimensional image space and use a traditional clustering technique in feature space. The number of matches at level l is given by the histogram intersection

function:

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

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:

$$K_\Delta(X, Y) = \sum_{i=1}^{L} \frac{1}{2^i}(I_{i-1} - I_i) \qquad (2)$$

3) Feature Extraction and Classification: This section briefly describes the feature extraction method from signature for signature recognition. First, the signature image is divided into 16×16 patches. The higher dimensional SIFT descriptors of 16×16 pixel patches are computed over a patch. Next, K-means clustering technique is applied on the patches from the training set for the generation of codebook. The typical vocabulary size for our experiments is 1024. Finally, SPM scheme is employed to generate the feature vector, which is then fed to the SVM classifier. In our experiment, the image is divided into $2^l \times 2^l$ segments in three different scales $l = 0, 1, 2$. 21 (16+4+1) BoF histograms are computed from these three levels, and all the histograms are concatenated to get the final vector representation of an image. The equation below represent the pyramid match kernel for three scales:

$$K_\Delta = I_2 + \frac{1}{2}(I_1 - I_2) + \frac{1}{4}(I_0 - I_1) \qquad (3)$$

The SVM using the one-versus-all is employed for multi-class signature classification. The signature recognition experiment is repeated 6 times with different randomly selected training and test images.

III. EXPERIMENT RESULTS

A. Experimental Dataset

3080 signatures from GPDS dataset and 7684 English printed words collected mainly from books, newspapers, magazines, journals, etc. are used to train the SVM classifier for signature detection task. For the testing of signature segmentation method, all the signed documents from the dataset of 'Tobacco-800' industrial archives [2] is used. The documents are written in English and the signatures on these documents also contain handwritten English characters. 300 classes of genuine signatures from GPDS dataset and 50 classes of Devnagari genuine signatures are used in the experiment. 24 genuine signatures are available in each class.

B. Signature Segmentation

The features computed from the patches obtained from morphological dilation of 'Tobacco-800' dataset, are fed to classifier and an overall accuracy of 95.58% is achieved for signature block detection. The errors are mainly due to segmentation problem at block level. Some broken parts of

signatures are identified as non-signature and some patches which contain printed words of two consecutive rows are misclassified as signature block. To get a comparative idea, the performance of our proposed method and the performance of an earlier similar work on the same dataset are given in Table I.

Table I
THE REPORTED SIGNATURE SEGMENTATION PERFORMANCE IN 'TOBACCO-800' DATASET.

Approach	Dataset	Accuracy (%)
Multi-scale structural saliency [15]	Tobacco-800	92.80
Conditional Random Field [14]	101 documents	91.20
Proposed Method	Tobacco-800	95.58

C. Signature Recognition

The signature recognition experiment on GPDS dataset demonstrates the excellent performance of our proposed approach. Table II shows the results when the experiment is repeated for 6 times for both the datasets using Linear SVM as a classifier. First and second rows show the results on 300 and 100 classes of GPDS signature dataset. The third row shows the accuracy from the experiment on 50 classes of Devnagari signature dataset. Overall 99.95%, 99.98%, 99.60% accuracy have been achieved from 300 classes, 100 classes of GPDS and 50 classes of Devnagari dataset, respectively. The ratio between True Positive Rate (TPR) and False Positive Rate (FPR) (i.e. ROC curve) is presented in Fig. 5. It shows the performance of the signature recognition experiment of English and Devnagari scripts, which is based on the combination of SIFT descriptor, SPM with the SVM classifier.

Table II
6-FOLD CROSS VALIDATION RESULTS FROM SIGNATURE RECOGNITION EXPERIMENT. F1-F6 REPRESENT THE ACCURACY IN PERCENT OF 6-FOLDS

Signature Data	F1	F2	F3	F4	F5	F6
GPDS (300)	99.97	99.97	99.91	99.96	99.97	99.96
GPDS (100)	99.98	99.99	99.98	99.98	99.98	99.95
Devnagari (50)	99.53	99.53	99.56	99.53	99.71	99.75

The signature recognition experiments have also been performed under two other configurations. The results obtained from the experiments are shown in Table III. A 54.33% accuracy is obtained from a HMM-based classification technique. Geometrical features [21] are computed and are fed to HMM classifier. The gradient-based feature and the SVM-based classification technique, which are used for signature segmentation, have also been employed for the signature recognition task and a 69.80% accuracy is obtained. The proposed method outperformed the results obtained on the
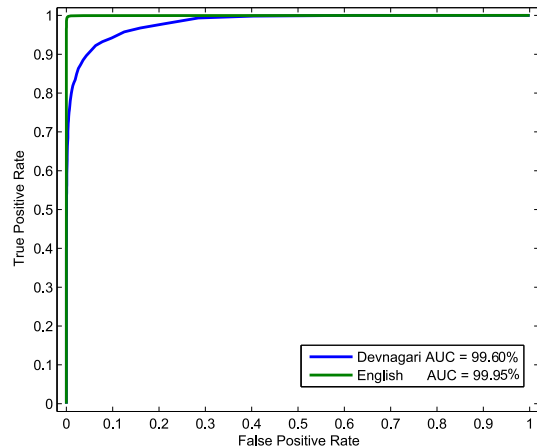
Figure 5. The ROC curves represent the performance of the signature recognition experiment. The green and the blue curves show the performance on the English and Devnagari datasets, respectively.

experiments using 300 classes of genuine signatures from the GPDS dataset.

Table III
PERFORMANCE OF SIGNATURE RECOGNITION UNDER DIFFERENT EXPERIMENTAL SETUPS

Approach	Accuracy (%)
Geometrical features [21] and HMM classifier	54.33
Gradient-based feature and SVM classifier	69.80
Proposed Method	99.95

The previously proposed approaches on signature segmentation and recognition have been tested on different publicly available datasets such as "Tobacco-800" and a few experiments have been conducted on the author's own collected dataset. A recall of 78.4% and 84.2% precision is reported by Srinivasan and Srihari [14] for the signature based document retrieval task. 96.13% accuracy is reported by [13] on Arabic/Persian documents. In [15], 93.20% MAP and 89.5% MRP have been reported for document retrieval based on signatures.

IV. CONCLUSION

Signature segmentation and recognition is a task of interest for content-based document retrieval based on signature information. In this paper, we propose an approach for efficient segmentation and recognition of signatures from document images. The signature region is detected in machine printed documents using the classification of components at the word level. The gradient-based feature and the SVM classifier are employed for signature detection. The signature recognition task is performed using SIFT descriptors with an SPM scheme. The empirical results of the experiments are encouraging and compare well with other state-of-the-art approaches in the literature.

REFERENCES

[1] C. M. T. J. B. A. J. F. Vargas, M. A. Ferrer, "Off-line handwritten signature gpds-960 corpus," In *Proc. International Conference on Document Analysis and Recognition*, 2007, 764–768.

[2] http://legacy.library.ucsf.edu/, "The legacy tobacco document library (ltdl)," University of California, San Francisco, 2007.

[3] J. Guo and M. Ma, "Separating handwritten material from machine printed text using hidden markov models," In *Proc. International Conference on Document Analysis and Recognition(ICDAR)*, 2001, 439–443.

[4] Y. Zheng, H. Li, and D. Doermann, "Machine printed text and handwriting identification in noisy document images," *IEEE Transaction on Pattern Analysis and Machine Intelligence(PAMI)*, vol. 26, no. 3, 337–353, 2004.

[5] J. Kumar, R. Prasad, H. Cao, W. Abd-Almageed, D. Doermann, and P. Natarajan, "Shape codebook based handwritten and machine printed text zone extraction," *Proc. SPIE 7874, 787406 (2011); doi:10.1117/12.876725.*

[6] 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 09)*, 2009, 431–435.

[7] X. Peng, S. Setlur, V. Govindaraju, and R. Sitaram, "Overlapped text segmentation using markov random field and aggregation," In *Proc. International Workshop on Document Analysis System (DAS)*, 2010, 129–134.

[8] X. Peng, V. Govindaraju, S. Setlur, and R. Sitaram, "Text separation from mixed documents using a tree-structured classifier," In *Proc.International Conference On Pattern Recogniton(ICPR)*, 2010, 241–244.

[9] G. Zhu, Y. Zheng, D. Doermann, and S. Jaeger, "Multi-scale structural saliency for signature detection," In *Proc. CVPR*, 2007, 1–8.

[10] V. K. Madasu, M. H. M. Yusof, M. Hanmandlu, and K. Kubik, "Automatic extraction of signatures from bank cheques and other documents," In *Proc. Digital Image Computing: Techniques and Applications(DICTA)*, 2003, 591–600.

[11] 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)*, vol. 26, no. 7, 1 253 003(1–25), 2012.

[12] S. Ahmed, M. I. Malik, M. Liwicki, and A. Dengel, "Signature segmentation from document images," In *Proc. International Conference on Frontiers in Handwriting Recognition(ICFHR)*, 2012, 425–429.

[13] A. Chalechale, G. Naghdy, and A. Mertins, "Signautre-based document retrieval," In *International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2003, 597–600.

[14] H. Srinivasan and S. N. Srihari, "Signature-based retrieval of scanned documents using conditional random fields," *Computational Methods for Counterterrorism*, 17–32, 2009.

[15] 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)*, vol. 31, no. 11, 2015–2031, 2009.

[16] 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)*, 2012, 225–230.

[17] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer-Verlag, 1995.

[18] M. Shi, Y. Fujisawa, T. Wakabayashi, and F. Kimura, "Handwritten numeral recognition using gradient and curvature of gray scale image," *IEEE Transaction on Pattern Recognition*, vol. 35, no. 10, 2051–2059, 2002.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision (IJCV)*, vol. 60, no. 2, 91–110, 2004.

[20] 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)*, vol. 2, 2006, 2169–2178.

[21] A. Fischer, A. Keller, V. Frinken, and H. Bunke, "Hmm-based word spotting in handwritten documents using sub-word models," In *Proc. International Conference on Pattern Recognition*, 2010, 3416–3419.