Training-Free and Segmentation-Free Word Spotting using Feature Matching and Query Expansion

Ekta Vats*, Anders Hast* and Alicia Fornés†
*Department of Information Technology,
Uppsala University, SE-751 05 Uppsala, Sweden
Email: ekta.vats@it.uu.se; anders.hast@it.uu.se
†Computer Vision Center, Computer Science Department,
Universitat Autònoma de Barcelona, Barcelona, Spain
Email: afornes@cvc.uab.es

***Abstract*—Historical handwritten text recognition is an interesting yet challenging problem. In recent times, deep learning based methods have achieved significant performance in handwritten text recognition. However, handwriting recognition using deep learning needs training data, and often, text must be previously segmented into lines (or even words). These limitations constrain the application of HTR techniques in document collections, because training data or segmented words are not always available. Therefore, this paper proposes a training-free and segmentation-free word spotting approach that can be applied in unconstrained scenarios. The proposed word spotting framework is based on document query word expansion and relaxed feature matching algorithm, which can easily be parallelised. Since handwritten words posses distinct shape and characteristics, this work uses a combination of different keypoint detectors and Fourier-based descriptors to obtain a sufficient degree of relaxed matching. The effectiveness of the proposed method is empirically evaluated on well-known benchmark datasets using standard evaluation measures. The use of informative features along with query expansion significantly contributed in efficient performance of the proposed method.**

***Index Terms*—Word spotting; Segmentation-free; Training-free; Query expansion; Feature matching**

I. Introduction

Libraries and cultural organisations contain rich and valuable manuscripts from diverse historical eras, that are to be digitized for preservation and protection from further degradation with time. The strong interest in digitization and transcription of historical manuscripts has led to exploration of automated methods for handwritten text recognition. However, automatic recognition of handwritten manuscripts is a challenging task, since the ancient manuscripts are barely readable and have complex layouts with high variability in handwriting styles. Some issues that hamper manuscript readability include high levels of degradation due to damages on the script surface, faded ink, ink bleed-through, stains, etc. [1].

Word spotting methods have emerged in the recent past, that allows exploring document images without the need of performing a full transcription. In general, word spotting refers to retrieving multiple occurrences of a word on-the-fly from document images. It is of particular interest in the research community as it offers an alternative solution to ad-hoc automatic transcription of historical handwritten text [2]. In recent times, deep learning based methods have achieved significant performance in handwritten word spotting [3], [4], [5], and are considered crucial in advancing handwritten text recognition research. This is due to availability of large scale of annotated data, GPU resources, and efficient architectures (e.g. Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), etc.) [4].

In document analysis literature, most traditional methods for handwritten text recognition are based on Hidden Markov Model (HMM) [6], [7], Dynamic Time Warping (DTW) [8], Support Vector Machine (SVM) [9], RNN [10], Bidirectional long short-term memory (BLSTM) [11], and large multidimensional long short-term memory RNN (MDLSTM) [12]. CNN-RNN hybrid architectures [4] have also emerged in the recent years that are capable of learning invariances specific to handwritten text. However, such methods are training-based, and require a huge amount of annotated training data. Nevertheless, their performance depend on the accuracy of word or line segmentation algorithms. This is because they require an initial segmentation of documents into words or text lines such that the word image representations contain relevant information. These limitations constrain the application of HTR techniques in document collections, because training data or segmented words are not always available. Therefore, this paper proposes a training-free and segmentation-free word spotting approach that can be applied in unconstrained scenarios.

Segmentation-free word spotting approaches are gaining importance these days. For example, Hast et al. [13] proposed a segmentation-free word spotting approach based on keypoints, a Fourier-based descriptor, and the Putative Match Analysis (PUMA). The main limitation of this method is that some parts of the retrieved words may be very similar to parts of the query word (i.e. confusion between similar words). This issue can be taken into account by dividing words into more parts using an efficient algorithm. Also, it has been tested for only single-writer scenario. Howe [14] presented a segmentation-free method based on a flexible ink-ball model where the query model is deformed in order to match the candidate regions. This method achieved significant performance with main drawback in terms of computational speed. A method based on local image descriptor was presented in [15], where

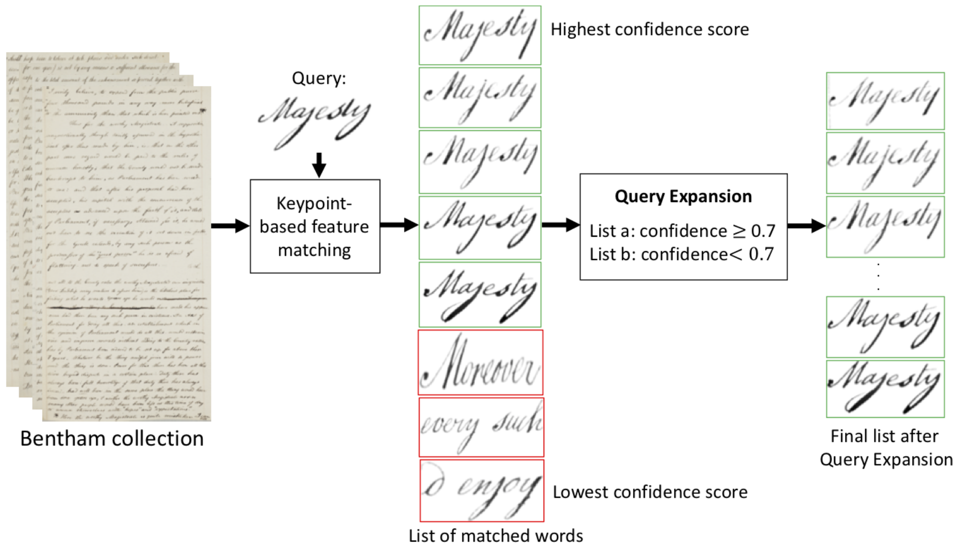

Fig. 1: Overall pipeline of the proposed framework, where a query word $Majesty$ is searched in the Bentham collection. The pre-processing is performed in such a way that the gray-level information crucial for keypoint detection is not affected. Keypoint-based feature matching is performed that generates a list of words with a certain degree of confidence. Using query expansion, the retrieved list of words are spread across two lists based on the confidence scores. As a result, the query expansion algorithm generates a list of best candidate words, with significant improvement in the word spotting performance.

the query is constructed using a glyph book of letter templates. Their word spotting model performs matching with zones of interest i.e. high curvature locations and extrema points in the strokes. This method is computationally slow, and requires computation of the gradient angles for each of the detected zones of interest. The word matching is performed by searching through the tree. Another drawback of this method is its inability to handle complex manuscripts with multiple writers, non-uniform layout, etc.

A patch-based framework for word spotting based on a bag-of-visual-words model was presented in [16], where a document is divided into cells and a SIFT feature vector is computed for each cell. The query word search is performed based on sliding-window algorithm with good computational speed. An efficient and fast segmentation-free word spotting framework based on exemplar SVM and Histogram of Gradient (HOG) descriptor was proposed in [9]. In general, the literature suggests that the algorithms with a better trade-off between accuracy and computational speed are based on gradients such as SIFT, HOG, etc., patches and sliding-window based methods.

This work presents a training-free and segmentation-free approach for word spotting based on document query word expansion and relaxed feature matching algorithm. The present investigation involves feature descriptors from the Fourier domain, use of keypoints based matching (instead of patches), and a document query word expansion algorithm, which is embarrassingly parallel. Since segmentation-free approaches do not rely on an initial segmentation of documents into words, the word representation model lacks relevant information and needs to handle additional noise. Therefore, the algorithm presented herein uses an efficient pre-processing technique based on Gaussian filtering to remove background noise from the document images.

The main motivation behind this work is that handwritten words posses distinct shape and characteristics which makes it difficult to obtain an exact transformation. Therefore, this work uses a combination of different keypoint detectors and a Fourier-based descriptor to obtain a sufficient degree of relaxed matching. By using informative features and efficient feature description, the part-based word matching problem is reduced to a much faster word search problem. An efficient document query word expansion algorithm is presented in this work which helps in improving the word recognition performance significantly.

This paper is organized as follows. Section II presents the proposed word spotting and query expansion approach; section III demonstrates the efficacy of the proposed method on well-known historical datasets using standard evaluation measures; and section IV concludes the paper, with insights on future perspectives.

II. Methodology

This section introduces the proposed word spotting framework, and highlights its main contributions. To begin with, the input document images are first pre-processed to remove background noise using a simple two band-pass filtering approach [17]. The pre-processing is performed in such a way that the gray-level information crucial for keypoint detection and feature extraction is not affected. This renders the keypoints detection to be more informative as it is easier to find more key points in a noise-free image. The keypoint detection and feature representation algorithms are described in section II-A, followed by a description of the proposed feature matching approach and query expansion in section II-B. The overall pipeline of the proposed word spotting framework is pictorially described in Figure 1, using an example of query word $Majesty$ from the Bentham collection.

A. Keypoint detection and feature representation

In general, keypoint-based matching is performed using a single keypoint detector. For example, Harris detector find corners; SIFT and SURF finds blob type of features, using difference of Gaussians and determinant of Hessian, respectively. Keypoint detection is performed using a combination of four keypoint detectors in this work, where the keypoints are detected for the document image as well as the query word. This is to capture different features representing a handwritten word. The four keypoint detectors used in this work includes Harris corner detector (capturing corners), square of the Determinant of Hessian (capturing dark and bright blobs), negative of Determinant of Hessian (capturing saddle points), and an edge detector (capturing text stroke edges). The experimental analysis deem a combination of these keypoint detectors to be best suited for the proposed word spotting framework.

Furthermore, handwritten word spotting is a challenging problem for feature descriptors as handwritten words posses distinct shape and characteristics, and suffer from various sources of variability, such as different writing styles, deformations, etc. [18]. Although there exists well-known descriptors such as SIFT, SURF, it is observed that their invariant properties amplify noise in a degraded image [19]. Therefore, it is important to use efficient feature descriptors that are fast to compute, and easy to be integrated in a sliding-window based search. In this work, the feature computed for each key points is based on a Radial Line Fourier descriptor [20], and the feature describing the whole word is based on another fast and efficient Fourier-based feature descriptor [21], both allowing a sufficient degree of relaxed matching of handwritten text.

In both cases, the feature description is fast as it is based on computing the amplitude of a few, yet most significant elements of the Fourier transform, unlike a Fast Fourier Transform (FFT) that requires computing amplitude of all elements. The method computes a low frequency content in the local neighbourhood that represent the most significant part of the signal. Therefore, the feature vector is not affected by high frequency changes such as noise or residuals from neighbouring words. The feature vector is also not sensitive to small variations in shape and form characteristics of handwritten text. The descriptor allows sampling in a small neighbourhood, and therefore constructs a short feature vector, without compromise on the matching performance. Although, Fourier-based descriptors are believed to be less accurate as compared to other popular features, they are indeed fast to compute, and are found to be well-suited for the proposed segmentation-free word spotting framework which requires a sliding-window search. A good trade-off between the accuracy and the computational speed is achieved using the Fourier-based descriptor.

B. Matching and query expansion

The proposed feature matching algorithm is based on keypoint matching using a simple sliding-window nearest neighbour search, performed within the subgroups of the four keypoint detectors. The initial steps of the proposed algorithm involves word partitioning and core text size estimation. In order to avoid confusion between similar words (where a word may share several letters with other different words and hence generate false-positives), a word is partitioned into several parts, relative to the size of the word. The partitions are automatically computed by taking the ratio between the width of the word and the core text size (i.e. no. of partitions = floor (width / coreSize)). This renders the proposed keypoint matching algorithm to follow a part-based matching approach, where parts of a word are matched with the corresponding part of other word. The extent of a word and the bounding box are adjusted such that the word is perfectly encapsulated inside the rectangle region of interest.

Using the nearest neighbor sliding-window search, the bounding box surrounding a word is thus evaluated. The size of the sliding-window is optimally selected and is four times the size of the query word. It is ensured that the word being searched is located in the query window at least once. Consequently, after a word is located, the matched points are removed from the set of points. This is to avoid retrieval of the same word multiple times. Furthermore, the matching keypoints are used to obtain a transformation which allows warping of the tentative word, removing any slant or size differences, without having to use DTW or similar. The confidence scores are computed from the Fourier-based feature of the warped word image and the query, by transforming the tentative word to the form of the query word using their corresponding key points. Furthermore, the feature vectors of the tentative word and the query word are normalized by taking the dot product to compute the confidence scores (normalized in the range 0 and 1). The matching algorithm results in correspondences between the query and the sliding-window and requires further processing to eliminate false-positives.

A query expansion algorithm is thus proposed, that refers to a sliding-window search performed on a list of words, and aims at improving the word matching results. Figure 1 pictorially represents the query expansion algorithm, using an example of query word $Majesty$ from the Bentham collection. The retrieved list of words are spread across two lists based on the confidence scores obtained from the matching algorithm. *List a* consists of words with a confidence score higher than or equal to 0.7; and *List b* consists of words with a confidence score lower than 0.7. Keypoint matching is performed on the list of words in *List a* to validate the matching result. A local query expansion algorithm is used where the list of words found on a page locally are taken into account, instead of performing query expansion on the entire set of pages. For example, in case of documents from the Esposalles dataset, the query word $reberé$ was found 8 times on the first page. Therefore, local query expansion is performed on the first page using the 8 retrieved instances of $reberé$. This procedure is repeated for each page, and can be fully implemented in parallel. The result from each query expansion is gathered generating a new list of found words with a higher degree of confidence. If too few words are found after performing local query expansion on *List a*, matching is performed on words in *List b*. As a result, the query expansion algorithm generates the best candidate words obtained from the simple sliding-window search and keypoint matching, rendering the word spotting system more accurate.

III. Experiments and results

This section describes the datasets used in the experiments, and empirically evaluates the proposed method.

A. Datasets

The dataset used for the experimental analysis consisted of documents written in old Catalan, German and English, with high levels of degradation and complexities.

- *Barcelona Historical Handwritten Marriages Dataset (BH2M)*: The BH2M database [22] consists of 550,000 marriage records stored in 244 books in the archives of Barcelona cathedral, with marriages held between the 15th and 19th century. A subset of BH2M database is used for experimental evaluations, where 50 pages are selected from the 17th century, handwritten by a single writer in old Catalan. In general, the ground truth consists of bounding boxes for each word, along with its corresponding transcription.
- *Bentham Dataset*: The Bentham collection, prepared as part of the $tranScriptorium$ project, consists of handwritten manuscripts by Jeremy Bentham over a period of 60 years, and his secretarial staff. The Bentham dataset was a part of the ICDAR 2015 Handwritten Keyword Spotting Competition [23]. For experiments analysis, all document pages from the dataset are employed, which have been written by multiple writers in different styles, font sizes, and contains crossed-out words.
- *Konzilsprotokolle Dataset*: The Alvermann Konzilsprotokolle dataset, prepared as part of the READ project, consist of a series of handwritten documents. The collection contains around 18000 pages in good preservation state from the University Archives in Greifswald [24]. It includes fair copies of handwritten minutes from formal meetings by the central administration, held between 1794-1797. The documents were digitized and provided by the University Library in Greifswald, and the transcripts were provided by the University Archives (Dirk Alvermann). This dataset was a part of the ICFHR 2016 Handwritten Keyword Spotting Competition [24].

B. Results

The performance of the proposed method is empirically evaluated against popular methods in the literature, using the same set of images and query words, with focus on segmentation-free word spotting scenario. The evaluation measure used is the mean Average Precision (mAP) metric, where a higher value of mAP is desirable. In general, a region is classified as a positive region if it overlaps more than 50% of the area in the ground truth corpora.

TABLE I: Experimental results for the BH2M dataset.

Method	mAP
Almazán et al. [25]	51.30
Zagoris et al. [19]	53.00
Hast et al. [20]	78.30
Proposed method	**85.72**

TABLE II: Experimental analysis of word spotting results for the BH2M dataset, w.r.t. random query words.

Query word	Proposed	Radial Line Fourier [20]	Fourier [13]	Graphs [26]	HOG [27]
reberé	**99.08**	95.25	94.10	75.67	74.78
pages	**98.78**	86.37	84.07	70.20	76.55
habitant	**99.35**	92.76	96.74	55.83	74.10
viuda	**96.82**	79.36	77.24	49.67	69.47
viudo	59.92	63.42	**64.35**	40.92	33.13
fill	49.72	46.50	**60.80**	56.66	42.09
mAP	**83.94**	77.27	79.55	58.15	61.68

In the first set of experiments, the proposed method is evaluated against the related methods: [25], [19], [20] on marriage records from the BH2M dataset. The method proposed in [25] is based on an exemplar-SVM framework for word spotting; [19] uses Document-oriented Local Features for word spotting; and [20] is based on a Radial Line Fourier descriptor. Table I illustrates the experimental results, where it is observed that the proposed method achieves a significantly higher mAP against other methods. Further investigation on the results obtained by [25] and [19] suggests that these methods are not well-equipped in handling cases where a word shares several characters with other words. The proposed method performed better in comparison with [20] due to efficient query expansion strategy. Nevertheless, it terms of computational speed, the proposed method is found to be the slowest, but most accurate.

Furthermore, a deeper experimental analysis of the word spotting results generated by several popular methods such as [20], [13], [26], [27] is conducted, where the performance is analysed on a variety of random query words. In general, [20] is based on a Radial Line Fourier descriptor; [13] uses a Fourier-based descriptor; [26] proposed a structural descriptor based on graph representation; and [27] employed a statistical descriptor based on Histogram of Gradients (HOG). Table II investigates the performance of different methods for six random and distinct query words. It is observed from Table II that the proposed method achieved a higher mAP overall, and performed well for words such as $reberé$, $pages$, $habitant$ and $viuda$. When search is performed on a longer query word e.g. $habitant$, the possibility of finding a query word as part of other word is low. Therefore, the method achieved a mAP of 99.35 for $habitant$. However, it is very common for a query word to share several characters with other words.

TABLE III: Experimental results for the Bentham dataset.

Method	mAP
PRG, TU Dortmund	29.30
CVC, Spain	11.60
Zagoris et al. [19]	32.60
Hast et al. [20]	78.60
Proposed method	**81.26**

TABLE IV: Experimental analysis of word spotting results for the Bentham dataset, w.r.t. 20 different query words from the query set for the validation partition.

Query word	Proposed	Hast et al. [20]
Aforesaid	92.02	**96.10**
Appoint	**78.52**	69.95
Bentham	**78.93**	68.21
Intended	**87.40**	87.31
Jeremy	**95.10**	85.32
Majesty	**98.79**	89.82
Number	**68.24**	58.54
Overseer	86.26	**86.50**
Remainder	47.12	**72.47**
Samuel	82.93	**85.44**
Decease	87.21	**90.61**
Personal	56.56	**65.28**
Survivor	95.70	**96.45**
Offender	**78.61**	67.40
Place	43.18	**48.66**
Amount	**81.71**	61.37
Supernumerary	**100**	98.67
Prisoner	78.27	**78.29**
Prisoners	**88.68**	86.52
Parliament	**100**	80.00
mAP	**81.26**	78.60

For example, a challenging case observed corresponds to the query words $viudo$ and $viuda$ with a sequence of overlapping characters. The proposed method is equipped in handling such cases to a significant extent, owing to the advantages from the part-based search model. More often than not, $viuda$ is retrieved while searching for $viudo$ (or vice versa). The proposed method handles this effectively by performing part-based keypoint matching, and then validating the results using query expansion. This helps in reducing the false-positives to a great extent, as can be seen in Table II.

Second set of experiments are performed on the Bentham dataset, where the performance of the proposed method is compared against the winner algorithms from the ICDAR 2015 Handwritten Keyword Spotting competition (i.e. teams from PRG, TU Dortmund and CVC, Spain); as well as recent methods [19] and [20]. Table III highlights the experimental results, where it is observed that the proposed method performed significantly in comparison with other methods. A mAP of 81.26 is achieved by the proposed method, and the performance is further analysed by evaluating the results generated from different query words.

Table IV illustrates the word spotting results obtained for 20 distinct query words. The query set for the validation partition included 95 images of 20 different words, which were used in the experiments. For comparison purpose, the word spotting performance of a recent work [20] is studied. It is observed from Table IV that the proposed method performs comparable

TABLE V: Experimental results for the Konzilsprotokolle dataset.

Method	mAP
PRG, TU Dortmund	52.20
TAU, Israel	**61.78**
CVC, Spain	0.0
Proposed method	50.91

TABLE VI: Experimental analysis of word spotting results for the Konzilsprotokolle dataset, w.r.t. random query words.

Query word	Proposed
wolle	47.31
Anfrage	68.33
Magn	41.21
Schreiben	53.12
Auch	76.45
Direct	75.68
Conclusum	65.39
Gadenbusch	49.85
Sicherheit	52.78
Academie	41.91
mAP	**57.20**

to [20], with a high mAP overall. This is a challenging dataset as it contains handwritten documents from multiple authors, with crossed-out words, written in different styles, font sizes, etc. For words such as $Supernumerary$ and $Parliament$, our method achieved perfectly accurate result because of very few instances of these words in the document set, and all the instances were correctly retrieved by the word spotting model.

In the third set of experiments, the proposed method is evaluated on the Konzilsprotokolle dataset, used in the ICFHR 2016 Handwritten Keyword Spotting Competition. The competition has four tracks, of which segmentation-free query-by-example track is in line with this work. Contrary to our proposal, the approaches from the teams from CVC, Spain; PRG, TU Dortmund; and TAU, Israel are training-based and take great advantage of the available training material. As observed in Table V, the method proposed by TAU, Israel performed the best in comparison with other methods. Nevertheless, our method is very competitive, and completely training-free, unlike other methods compared herewith. For insights on the word spotting results for sample query words, the reader is referred to Table VI.

C. Discussion

These days most of the research on word spotting is focused on designing training-based methods (deep learning being the most popular choice), and segmentation-free methods are also not explored much in the recent past. However, in real-world it is not realistic to have access to large amounts of pre-transcribed texts to train a word spotting system. This also limits the applicability and usability of training-based methods in scenarios where text can be written in any script, style and language, and a computer may necessarily not be pre-trained to identify such texts. Therefore, research on developing sophisticated training-free method is essential for advancing handwritten text recognition. The method proposed in this work is an effort towards developing efficient training-free and segmentation-free word spotting approach, which achieves decent performance in terms of accuracy. However, the proposed method is computationally slow due to exhaustive keypoint matching, and requires further improvement in future work.

IV. Conclusion and perspectives

This paper presented a segmentation-free and training-free handwritten word spotting approach, based on document query word expansion and relaxed feature matching. Sufficient degree of relaxed matching of handwritten words was obtained by using a combination of different keypoint detectors and a Fourier-based descriptor. The experimental results on a variety of historical document images from well-known datasets demonstrate the effectiveness of the proposed method. It is empirically demonstrated that the proposed query expansion algorithm significantly improved the word recognition performance. Under the experimental settings, the proposed method significantly performed in comparison with the state-of-the-art training-free methods, including the winners of the popular keyword spotting competitions.

As future work, the ideas presented herein will be tested for complex handwritten text with multi-writer document collections and heavily degraded archival databases. This is to ensure robustness of the proposed word spotting framework for a variety of handwriting styles. Furthermore, we plan to advance the current experimental framework by automatically fine-tuning the key parameters using sophisticated methods such as Bayesian optimization. Lastly, the computational speed will be improved in future work.

Acknowledgment

This work has been partially supported by the Swedish strategic research programme eSSENCE, the Riksbankens Jubileumsfond (Dnr NHS14-2068:1), the Spanish project RTI2018-095645-B-C21 and the Ramon y Cajal Fellowship RYC-2014-16831. The computations were performed on resources provided by SNIC through UPPMAX under project SNIC 2018/3-232.

References

[1] A. P. Giotis, G. Sfikas, B. Gatos, and C. Nikou, "A survey of document image word spotting techniques," *Pattern Recognition*, vol. 68, pp. 310–332, 2017.

[2] L. Rothacker and G. A. Fink, "Segmentation-free query-by-string word spotting with bag-of-features hmms," in *International Conference on Document Analysis and Recognition*. IEEE, 2015, pp. 661–665.

[3] P. Krishnan, K. Dutta, and C. Jawahar, "Word spotting and recognition using deep embedding," in *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*. IEEE, 2018, pp. 1–6.

[4] K. Dutta, P. Krishnan, M. Mathew, and C. Jawahar, "Improving cnn-rnn hybrid networks for handwriting recognition," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*. IEEE, 2018, pp. 80–85.

[5] S. Sudholt and G. A. Fink, "Phocnet: A deep convolutional neural network for word spotting in handwritten documents," in *ICFHR*. IEEE Computer Society, 2016, pp. 277–282.

[6] J. A. Rodríguez-Serrano and F. Perronnin, "Handwritten word-spotting using hidden markov models and universal vocabularies," *Pattern Recognition*, vol. 42, no. 9, pp. 2106–2116, 2009.

[7] J. A. Rodríguez-Serrano, F. Perronnin *et al.*, "A model-based sequence similarity with application to handwritten word spotting," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2108–2120, 2012.

[8] T. M. Rath and R. Manmatha, "Word spotting for historical documents," *International Journal on Document Analysis and Recognition*, vol. 9, no. 2, pp. 139–152, 2007.

[9] J. Almazán, A. Gordo, A. Fornés, and E. Valveny, "Word spotting and recognition with embedded attributes," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 12, pp. 2552–2566, 2014.

[10] 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*, vol. 34, no. 2, pp. 211–224, 2012.

[11] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, "A novel connectionist system for unconstrained handwriting recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 5, pp. 855–868, 2009.

[12] P. Voigtlaender, P. Doetsch, and H. Ney, "Handwriting recognition with large multidimensional long short-term memory recurrent neural networks," in *Frontiers in Handwriting Recognition (ICFHR), 2016 15th International Conference on*. IEEE, 2016, pp. 228–233.

[13] A. Hast and A. Fornés, "A segmentation-free handwritten word spotting approach by relaxed feature matching," in *Document Analysis Systems (DAS), 2016 12th IAPR Workshop on*. IEEE, 2016, pp. 150–155.

[14] N. R. Howe, "Part-structured inkball models for one-shot handwritten word spotting," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 582–586.

[15] Y. Leydier, A. Ouji, F. LeBourgeois, and H. Emptoz, "Towards an omnilingual word retrieval system for ancient manuscripts," *Pattern Recognition*, vol. 42, no. 9, pp. 2089–2105, 2009.

[16] M. Rusinol, D. Aldavert, R. Toledo, and J. Lladós, "Efficient segmentation-free keyword spotting in historical document collections," *Pattern Recognition*, vol. 48, no. 2, pp. 545–555, 2015.

[17] E. Vats, A. Hast, and P. Singh, "Automatic document image binarization using bayesian optimization," in *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing*. ACM, 2017, pp. 89–94.

[18] J. Almazán, A. Gordo, A. Fornés, and E. Valveny, "Segmentation-free word spotting with exemplar svms," *Pattern Recognition*, vol. 47, no. 12, pp. 3967–3978, 2014.

[19] K. Zagoris, I. Pratikakis, and B. Gatos, "Unsupervised word spotting in historical handwritten document images using document-oriented local features," *IEEE Transactions on Image Processing*, 2017.

[20] A. Hast and E. Vats, "Radial line fourier descriptor for historical handwritten text representation," *Journal of WSCG*, vol. 26, no. 1, pp. 31–40, 2018.

[21] A. Hast, V. Sablina, I.-M. Sintorn, and G. Kylberg, "A fast fourier based feature descriptor and a cascade nearest neighbour search with an efficient matching pipeline for mosaicing of microscopy images," *Pattern Recognition Image Analysis*, vol. 28, no. 2, pp. 261–272, 2018.

[22] D. Fernández-Mota, J. Almazán, N. Cirera, A. Fornés, and J. Lladós, "Bh2m: The barcelona historical, handwritten marriages database," in *International Conference on Pattern Recognition (ICPR)*. IEEE, 2014, pp. 256–261.

[23] J. Puigcerver, A. H. Toselli, and E. Vidal, "Icdar2015 competition on keyword spotting for handwritten documents," in *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*. IEEE, 2015, pp. 1176–1180.

[24] I. Pratikakis, K. Zagoris, B. Gatos, J. Puigcerver, A. H. Toselli, and E. Vidal, "Icfhr2016 handwritten keyword spotting competition (h-kws 2016)," in *Frontiers in Handwriting Recognition (ICFHR), 2016 15th International Conference on*. IEEE, 2016, pp. 613–618.

[25] J. Almazán, A. Gordo, A. Fornés, and E. Valveny, "Efficient exemplar word spotting." in *BMVC*, vol. 1, no. 2, 2012, p. 3.

[26] P. Wang, V. Eglin, C. Garcia, C. Largeron, J. Lladós, and A. Fornés, "A coarse-to-fine word spotting approach for historical handwritten documents based on graph embedding and graph edit distance," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*. IEEE, 2014, pp. 3074–3079.

[27] J. Almazán, D. Fernández, A. Fornés, J. Lladós, and E. Valveny, "A coarse-to-fine approach for handwritten word spotting in large scale historical documents collection," in *International Conference on Frontiers in Handwriting Recognition (ICFHR)*. IEEE, 2012, pp. 455–460.