

A Complete Logo Detection/Recognition System for Document Images

Alireza Alaei, Mathieu Delalandre

Laboratoire d'Informatique (LI EA6300), Université François-Rabelais de Tours, France
{alireza.alaei, mathieu.delalandre}@univ-tours.fr

Abstract—In this paper, a complete logo detection/ recognition system for document images is proposed. In the proposed system, first, a logo detection method is employed to detect a few regions of interest (logo-patches), which likely contain the logo(s), in a document image. The detection method is based on the piece-wise painting algorithm (PPA) and some probability features along with a decision tree. For the logo recognition, a template based recognition approach is proposed to recognize the logo which may present in every detected logo-patch. The proposed logo recognition strategy uses a search space reduction technique to decrease the number of template logo-models needed for the recognition of a logo in a detected logo-patch. The features used for search space reduction are based on the geometric properties of a detected logo-patch. Based on our experimentations on 1290 document images of Tobacco800 dataset, 99.31% of the logos were detected as logo-patches. Among the detected logo-patches 97.90% of logos were fairly recognized. Considering both logo detection and recognition results, 97.22% of the logos in the document images could truly be detected/recognized as the overall performance of the proposed system.

Keywords: *Logo recognition; Logo detection; Under/over segmentation; Template matching.*

I. INTRODUCTION

Logo is a unique graphic, word(s), or combination of both word(s) and graphic symbol(s), which identifies products and services of a company, a public organization, an institution or an individual. It is also considered to be the mean of identification for that company, organization or institution itself all through the world. Although logos can be found in many forms, colors and styles but they are bounded by certain design restrictions as they need to be salient and easily recognized. To have an idea of different logo types, some examples of logos are shown in Fig. 1.

In the literature of document image analysis (DIA), the problem of using logo information as a fundamental characteristic of many document images has been investigated around two main topics of i) logo spotting, and ii) logo recognition. Logo spotting can be viewed as a way to efficiently localize logos by limiting the computational complexity, without concerning on completely recognition of logos [1]. Moreover, the logo spotting methods rely on an indexing step computed off-line, and compare queries online with the index. Logo recognition methods [2, 3] are commonly focused on the segmentation and classification issues, making the systems suitable for real-time processing of document work-flow. Basically, the logo recognition involves two main steps: a) detecting probable logo from a

document image, and b) recognizing the detected logo candidate from a database.

Several approaches for logo detection and recognition/classification have been presented in the literature [2-10]. Logo detection methods are generally interested with initial detection of regions of interest, which contain logos, in complete documents [4], [5]. A typical approach for logo detection in the literature is based on segmentation of document images into a set of connected components, and then describing these components with the help of geometric features such as size, aspect ratio, density, and domain based features (e.g. a priori position of logos in document images) [4, 5]. Techniques such as decision trees or Fisher classifiers are then used to separate logo components from the rest of the components. Compared to other classification approaches these techniques generally require fewer training samples and provide straightforward interpretation. These methods perform usually within a rough localization and performance evaluation metrics tolerate a precision of 25% overlap between the ground-truth and the detection results. In an overall, the time complexity of this type of approaches is of linear complexity, making them suitable for real-time processing of document work-flow.

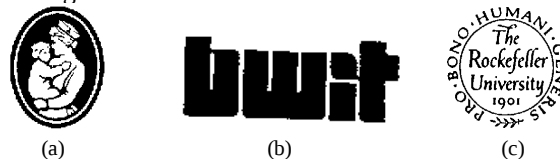


Figure 1. Three different types of logos shown from the Tobacco800 dataset [21]: a) A graphical logo, b) A textual logo, and c) A logo composed of both text and graphic entities.

Logo recognition/classification as a topic closely linked to the logo detection has been discussed in the literature and many recognition systems have been proposed in [6-10] based on shape, local primitives, template, and matching functions. In general, shape extraction techniques are not suitable for the recognition purpose as they are known to be sensitive to noise and shape distortion (e.g. occlusion and strip corruption) [6, 7]. Local methods based on low-level primitives such as connected components described with invariants [7], contour images [8], pixel framing [9], and templates [10] and matching functions as Hausdorff distance [8], L_1 Norm [7], PCA-based matching [10]) or classifiers (MLP) [9] can obtain better recognition performance. Results are reported in terms of robustness to noise, shape distortion (occlusion, strip corruption, and skew). The systems proposed in [6-10] work well where the logos are perfectly segmented from the document images in advance.

However, the methods proposed in [6-10] for the recognition of logos cannot be applied in the context of detection as the issues related with logo detection e.g. the under/over segmentation and rough localization of logo images have not been considered. Therefore, the challenge is to design a recognition method which can deal with the detection failures (rough localizations and false alarms). To the best of our knowledge, only in the work described in [2, 3] both detection and recognition have been considered in a single framework to deal with the problem of logo detection/recognition. In [2], a top-down matching strategy, which looks similar to logo spotting, is proposed. Features are extracted for the connected components obtained from the document and then compared with a trained index. Any positive match will trigger a geometry reconstruction step to detect and recognize the logo. Since, this method [2] is based on a compact features vectors of global primitives along with a primitives-driven geometry reconstruction, it seems to be time-efficient as well. The work described in [3] is more related to a system level approach. In this system [3], the detection and recognition stages interact with each other employing a feedback strategy which provides information of the recognition stage to the detection stage. Regions of interest are detected using a text/graphics separation method. Connected components composing a region are described using local invariants and compared with a trained database. Relations between the components are also modeled through a Region Adjacency Graph (RAG). While comparing a region of interest with the models, any incorrect structures in the RAG will trigger a modification of the region of interest by sending a feedback to the detection stage [3].

Nevertheless, the systems presented in [2, 3] are not very accurate as the results reported on public datasets highlight a moderate accuracy. In addition, though some considerations about the system level are discussed in [3], no discussions are carried out about the overall performance of detection/recognition strategy. In this paper, we present a contribution on these aspects proposing a complete detection and recognition system with an integrated strategy. In our proposed system, the detection stage aims to obtain probable logo-patches (regions of interest) with a high recall value whereas the precision value may be small resulting in many false alarms and rough localization of logos. In other words, the main goal at the detection level is to obtain a large pruning of the search space. The recognition stage relies on a strong matching technique to support the position-invariance property, under/over segmentation problems and also false alarms detection. Additional information coming from the detection level, including a priori logo-models, regions of interest with their confidence rates, a priori knowledge about the geometrical places of the logo-models are also utilized in the recognition process to speed up the recognition process. This strategy makes the system more accurate and time-efficient, without bounding the accuracy at the detection level.

The organization of rest of the paper is as follows: Section II describes our proposed system. Section III discusses the experimental results and comparative analysis. Finally, some conclusions are drawn in Section IV.

II. PROPOSED SYSTEM

In Fig. 2, an overview of our proposed logo detection/recognition system is depicted. The detection process is based on the method proposed in [20]. In the proposed logo detection scheme [20], a number of regions of interest (RIs) are obtained for each document image as probable logo areas or logo-patches. Then employing a template matching based recognition method the recognition of logos is achieved. Details of the steps involved in the proposed system are described in the subsequent subsections.

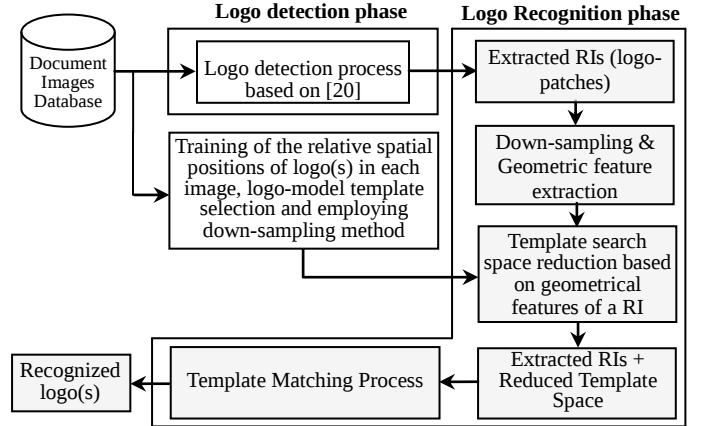


Figure 2. Overview of the proposed logo detection and recognition scheme.

A. Logo detection

In this research work, the method proposed by the authors in [20] is used for logo detection purpose. In this method, at first, piece-wise painting algorithm (PPA) used for text-line segmentation is employed to represent a document image as a set of patches. Content of the input document image represented by the patches is then pruned utilizing a decision tree and a small number of features such as frequency probability (FP), Gaussian probability (GP), height, width, and average density computed for the patches to obtain a set of patches with more probability of containing logos. A morphological dilation operation is applied to find the final logo-patches (regions of interest). Three detected logo-patches based on the proposed logo detection method [20] considering Tobacco800 dataset [21] are shown in Fig.3. These regions of interest are then considered for final template matching process in the recognition phase of the proposed system in this research work. More details about the logo detection method can be found in [20].

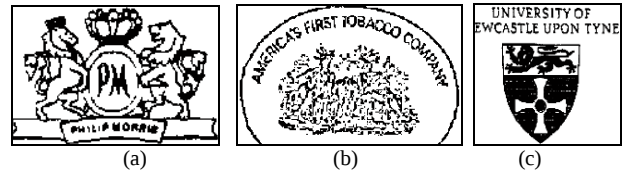


Figure 3. Results of the proposed logo detection phase: a) A perfect detected logo-patch, b) An instance of over-segmentation, and c) An under-segmentation example (based on the ground truth information).

B. Logo recognition

Based on the aforementioned discussion in the section I, we propose here a recognition method based on area-based

template matching [11], as this method provides all the robustness and position-invariant requirements to deal with the logo recognition problem in the context of detection problems (over/under segmentation and rough localization). To accelerate the recognition process, a down-sampling technique and a template search space reduction are also employed before the template matching procedure.

1) *Feature extraction*: For each patch obtained from the logo detection phase presented in [20], a number of simple geometrical features such as the center of bounding box obtained from the patch, center of gravity, height, width, left most point, right most point, top most point, bottom most point of the patch are extracted. This information mainly characterizes the size and spatial position of each patch in a document image. Furthermore, to speed up the template matching process and consequently achieve better performance in terms of computation time, a down-sampling method [22] is employed on the extracted patches considering only alternate rows (1,3,5,...) and columns (1,3,5,...) of the extracted patches.

2) *Template search space reduction*: Spatial positions of logos in document images have been used for logo detection [4, 20]. Concerning the Tobacco800 dataset, spatial positions of logos have been defined as means (centers) of 3 different clusters obtained during learning phase [4, 20]. In this research work, in addition to the centers of those clusters, standard deviations of features are also considered to take care of the distribution of features in all clusters. Based on the geometrical features extracted for each patch, and the spatial positions (cluster center and standard deviation) of the templates obtained during the training phase, the patch is classified to either one or more of those clusters. The templates included in the cluster (clusters) of which the patch is classified to are considered for template matching. Since, the number of templates in each cluster is comparably less than the total number of templates in all the clusters, the search space of templates for the matching process is considerably reduced. In theory, this process can reduce the time complexity for template matching into 1/3 of the time needs for template matching using whole set of templates, if the templates are equally distributed in all the clusters.

3) *Template Matching*: Template matching is concerned with different open problems including size, scale, position invariance, and selection of the distance-based functions for image comparison [12]. In the context of logo recognition, scale and rotation invariance are not concerned as the logos appear at fix orientation and the resolution in document work-flow is also fixed. Problems to be solved include the choice of a robust distance-based function, and to deal with the position invariance and the complexity issues related to the image and template sizes. However, in case of skew and scalability problems some preprocessing methods are applicable.

Distance-based comparison of binary images is mainly concerned with three main methods: i) the Chamfer Matching (CM), ii) the Hausdorff Distance (HD), and iii) the direct comparison [12]. The CM increases the false alarms,

which mainly happen because of background with high level of clutter noise [13], making this method not suitable for the recognition of the logos extracted based on a detection approach from document images. In the literature, the HD has been reported to perform well for logo recognition [8] concerning several distortion issues. However, HD is computationally expensive having a $O(m \times n^2)$ complexity to match a template of size n to an image of size m . An approximation of the HD have been proposed working with a $O(n \times \log n)$ complexity [14] with the cost of dipping the accuracy. Pruning methods for reducing the search space in the image can also be applied to speed up the process [15]. Nevertheless their performance are mainly emphasized when facing a small in size problem e.g. to find a 32×32 pattern in a 256×256 image [15]. In this context, application of the HD to the logo recognition problem may not be time efficient.

The direct image comparison achieved by comparing the pixels of same coordinates in the search image and in the template. This approach has commonly been used in the literature due to its simplicity, robustness to noise and occlusion and its favorable computational complexity [15]. With gray-level images, the direct image comparison is obtained by using the L_p norm function. This function can be derived in the forms of Sum of Absolute Differences (SAD) and Sum of Squared Distances (SSD). The Cross Correlation (CC) and the Normalized Cross-Correlation (NCC) between a template and an image is obtained by minimizing this distance on the whole image. Optimization is achieved using the Fast Fourier Transform (FFT). This solves the position-invariant template matching within a complexity of $O(m \times \log m) + O(m)$ on a whole image of size m , of which $O(m \times \log m)$ belongs to the FFT computation and $O(m)$ corresponds the frequency domain multiplication and the maximum value search [16].

The major problem with the use of FFT may be the online computation cost of $O(m \times \log m)$ in practical applications. To reduce the computation time, the FFT of the template models can be computed either off-line during a training stage or by the use of a Graphics Processing Unit (GPU) [17]. The computation cost is directly linked to the image size m . In our case, m is set to the maximum size of detected regions of interest during the training.

Another alternative to reduce the complexity is employment of the convolution theorem to exploit the image features. Template matching can support particular implementations when using binary images. The direct image comparison can be achieved with exclusive NOR operator between an image and a template. This binary form requires significantly less resource than an L_p norm implementation with gray level images. NCC in the binary form has been investigated in [18] as a normalized AND product. It is however sensitive to the universal 1 image (i.e. full black region) and the template matching algorithm must predict these cases. Meanwhile, for a universal 0 image it becomes not necessary to have a matching calculation.

When using binary images, correlation computation is also made easier as images contain many zeros (i.e. the foreground pixels) and the task can be further simplified by using matrix factorization technics [19]. Experiments report

in [19] highlight that the correlation operation using this approach can be obtained with 2 to 3 times less operations than the conventional FFT computation. The correlation is also performed in absence of trigonometric, multiplications, summing up operations and floating-point coding required while computing the FFT. Another advantage is to relax the constraint of matrix dimensions, as correlation can be performed with matrixes of arbitrary dimensions shifting the comparison from an $m \times m$ to an $m \times n$ problem.

Concerning the aforementioned discussion, in this research work a NCC template matching using FFT [16] is utilized as the basis of template matching procedure. Apart from down-sampling [22] process employed to speed up the template matching process, the geometrical information is also used to reduce the number of reference temple logo-models to be matched with the detected logo-patch to decrease the computation time of the proposed logo recognition strategy.

III. EXPERIMENTAL RESULTS AND COMPARISON

A. Datasets and metrics of evaluation

In this research work, two different datasets were used to evaluate the performance of the proposed system. The first one is the Tobacco800 dataset [21]. The second one is the Itesoftware dataset which contains 8200 real-life document images of which 5748 images contain logos. In this research work, two different subsets of document images from the Itesoftware dataset were considered. These two subsets are simply the document images having logos belong to the classes from 1 to 100. In Table I, descriptions of these 2 subsets of data samples and the Tobacco800 are drawn.

For the logo detection, precision (Pre.) and recall (Rec.) as the metrics of performance evaluation are computed. Accuracy (Acc.) is computed for the recognition results. Overall performance is also figured as the multiplication of recall and accuracy.

TABLE I. DESCRIPTIONS OF DATASETS.

Property Dataset	Number of document mages	Number of logos	Number of logo classes
Tobacco800 [21]	1290	432	35
Subset 1	261	263	23
Subset 2	685	705	82

B. Results and discussion

Concerning the Tobacco800 dataset [21], 100 document images which contain logos were considered for the training of detection phase. The training documents selected on the basis to cover all the classes of logos placed in different positions. The rest were considered to test the proposed system. Between one and 15 templates were also considered as template logo-models for template matching of the proposed recognition strategy. The results obtained for the detection, recognition and overall performance based on the proposed system in this research work are shown in Table II.

For evaluating the proposed logo detection/recognition system based on the first subset of data samples, 23 document images (one document per class) were considered

for the training. The number of templates used for the template matching in the recognition step was 29. The experimental results obtained from the proposed system on the first subset of Itesoftware dataset are drawn in Table II.

In the experimentation on the second subset of Itesoftware dataset, 94 document images were considered for the training. A number of 109 template logo-models were used for the template matching and the results are demonstrated in Table II. From the Table II, it can be noted that the proposed system provided promising results in terms of detection and recognition on different datasets. The overall performance of the proposed system is also reasonable when considering real-world documents for the experimentations.

TABLE II. THE RESULTS OBTAINED FROM THE PROPOSED TEMPLATE MATCHING APPROACH USING TOBACCO800 AND 2 DIFFERENT SUBSETS OF ITESOFTWARE DATASET.

Dataset	Accuracy	Detection		Rec. Acc. %	Overall performance %
		Recall %	Precision %		
Tobacco800 [21]		99.31	20.58	97.90	97.22
Subset 1 of Itesoftware dataset		97.21	23.85	98.82	96.06
Subset 2 of Itesoftware dataset		96.31	23.19	94.10	90.63

C. Comparative analysis

To compare the performance of the recognition strategy used in the proposed system, a method (shape context features along with Hausdorff Distance measure [8]) utilized in the literature for logo recognition is employed on the results obtained from our detection phase using the Tobacco800 [21] and Itesoftware datasets. The results are tabulated in Table III and IV. From Table III and IV, it is evident that the results obtained from the recognition strategy employed in the proposed system in this research work are better than the one used shape context features and the Hausdorff Distance measure [8] for the recognition of logo in the proposed detection context.

TABLE III. THE RESULTS OBTAINED BASED ON SHAPE CONTEXT FEATURES AND THE PROPOSED TEMPLATE MATCHING APPROACH USING THE TOBACCO800 DATASET.

Feature	Accuracy	Detection Acc. %	Recognition Acc. %	Overall performance of the system %
Shape Context features and Hausdorff Distance		99.31	62.01	61.57
The proposed Template Matching approach		99.31	97.90	97.22

TABLE IV. THE RESULTS OBTAINED BASED ON SHAPE CONTEXT FEATURES AND THE PROPOSED TEMPLATE MATCHING APPROACH USING THE SECOND SUBSET OF ITESOFTWARE DATASET.

Feature	Accuracy	Detection Acc. %	Recognition Acc. %	Overall performance of the system %
Shape Context features and Hausdorff Distance		96.31	51.98	50.06
The proposed Template Matching approach		96.31	94.10	90.63

Furthermore, we compared the performance of our proposed system with the one presented in [3] and the results are provided in Table V. Comparing the results presented in Table V, we noted that the performance of our proposed system is better than the system presented in [3]. This is

because in our proposed system, most of the possible logo-patches are detected/localized and logo-patches are rarely missed during the detection phase. Moreover, the recognition strategy used in our proposed system can fairly recognize the logos in the roughly detected logo-patches.

TABLE V. COMPARISON OF THE RESULTS OBTAINED FROM THE PROPOSED SYSTEM AND THE SYSTEM PRESENTED IN [3] ON THE TOBACCO800 DATASET.

Method \ Accuracy	Detection Acc. %	Recognition Acc. %	Overall performance of the system %
Wang [3]	94.70	92.90	87.98
Proposed approach	99.31	97.90	97.22

D. Erroneous results

To have an idea about the types of errors occurred during our experimentation on the Tobacco800 dataset, a document image which the proposed system could not detect the logo in image is shown in Fig. 4-(a). In Fig. 4-(b), an image that the proposed system failed to recognize the logo included in the detected patch is shown. The main problem for the failure of the system was very bad quality of the images; since, both images were affected by too much of noise and degradation.

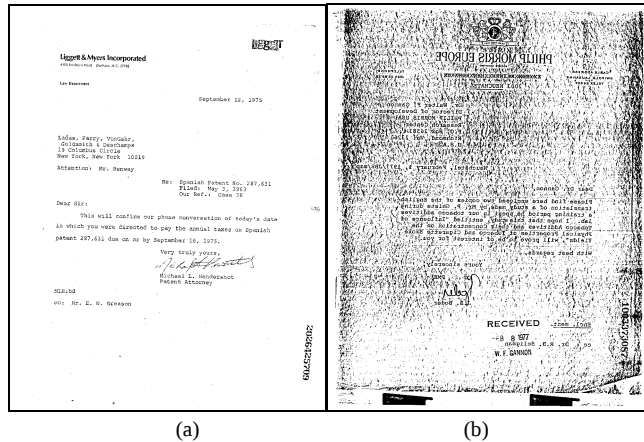


Figure 4. Two erroneous results: a) Missing the detection of logo in the document image, and b) Failure in the recognition phase for recognizing the logo.

IV. CONCLUSIONS

A complete system for logo detection/recognition in document images is proposed in this research work. The problems of under and over segmentation, which frequently occurs during detection step, are discussed and a template based recognition strategy suitable for such kind of problems is proposed. A search space reduction for decreasing the number of template logo-models in the recognition strategy using geometrical properties of logo-patches and template logo-models is proposed to speed up the template matching process during recognition stage. The results obtained for the logo detection and recognition using a real-world document dataset clearly show that the overall performance of the proposed system is promising and it is suitable for practical use.

ACKNOWLEDGMENTS

Authors would like to thank The Anh Pham of Université de Tours, V.P. D'Andecy and S. Kébairi of ITESOFT for their help, support and suggestions.

REFERENCES

- [1] R. Jain and D. Doermann, "Logo retrieval in document images," In Proceedings of DAS, pp. 135–139, 2012.
- [2] Z. Li, M. Schulte-Austum, and M. Neschen, "Fast logo detection and recognition in document images," In Proceedings of ICPR, pp. 2716–2719, 2010.
- [3] H. Wang, "Document logo detection and recognition using bayesian model," In Proceedings of ICPR, pp. 1961–1964, 2010.
- [4] G. Zhu and D. Doermann, "Automatic document logo detection," In Proceedings of ICDAR, pp. 864–868, 2007.
- [5] H. Wang and Y. Chen, "Logo detection in document images based on boundary extension of feature rectangles," In Proceedings of ICDAR, pp. 1335–1339, 2009.
- [6] S. Lowther, V. Chandran, and S. Sridharan, "Recognition of logo images using invariants defined from higher-order spectra," In Proceedings of Asian Conference on Computer Vision, 2001.
- [7] J. Neumann, H. Samet and A. Soffer, "Integration of Local and Global Shape Analysis for Logo Classification," Pattern Recognition Letters, 23(12), pp. 1449–1457, 2002.
- [8] J. Chen, M. K. Leung and Y. Gao, "Noisy logo recognition using line segment Hausdorff distance," Pattern Recognition, 36, pp. 943–955, 2003.
- [9] M. Gori, M. Maggini, S. Marini, J. Q. Sheng and G. Soda, "Edge-back propagation for noisy logo recognition," Pattern Recognition, 36, pp. 103–110, 2003.
- [10] H. Sali and M. Dzulkifli, "Logo Matching Technique Based on Principle Component Analysis," International Journal of Computer Vision and Applications, 1(1), pp. 20–26, 2011.
- [11] T. Mahalakshmi, R. Muthaiah, and P. Swaminathan, "An overview of template matching technique in image processing," Research Journal of Applied Sciences, Engineering and Technology, 4(24), pp. 5469–5473, 2012.
- [12] V. D. Gesu and V. Starovoitov, "Distance-based functions for image comparison," Pattern Recognition Letters, 20(2), pp. 207–214, 1999.
- [13] T. Ma, X. Yang, and L. Latecki, "Boosting chamfer matching by learning chamfer distance normalization," In Proceedings of ECCV, LNCS Vol. 6315, pp. 450–463, 2010.
- [14] F. Correa-Tome and R. Sanchez-Yanez, "Fast similarity metric for real-time template-matching applications," Journal of Real-Time Image Processing (JRTP), 2013.
- [15] Y. Hel-or and H. Hel-or, "Real-time pattern matching using projection kernels," Pattern Analysis and Machine Intelligence, 27(9), pp. 1430–1445, 2005.
- [16] J. Lewis, "Fast template matching," in Vision Interface, pp. 120–123, 1995.
- [17] O. Fialka and M. Cadik, "Fft and convolution performance in image filtering on gpu," In Proceedings of International Conference on Information Visualisation, pp. 609–614, 2006.
- [18] G. Zhang, M. Lei, and X. Liu, "Novel template matching method with sub-pixel accuracy based on correlation and fourier-mellin transform," Optical Engineering, 48(5), p. 057001, 2009.
- [19] R. Bogush, S. Maltsev, S. Ablameyko, S. Uchida, and S. Kamata, "An efficient correlation computation method for binary images based on matrix factorisation," In Proceedings of ICDAR, pp. 312–316, 2001.
- [20] A. Alaei, M. Delalandre, N. Girard, "Logo Detection Using Painting Based Representation and Probability Features," In Proceedings of ICDAR, pp. 1267–1271, 2013.
- [21] G. Agam, S. Argamon, O. Frieder, D. Grossman, D. Lewis, "The complex document image processing (CDIP) test collection project," Illinois Institute of Technology, 2006, <http://ir.iit.edu/projects/CDIP.html>
- [22] R.C. Gonzalez, R.E. Woods, "Digital image processing," Prentice Hall India, Second Edition, 2009.