

Editorial

# Document Image Processing

Laurence Likforman-Sulem <sup>1,\*</sup> and Ergina Kavallieratou <sup>2</sup>

<sup>1</sup> Institut Mines-Télécom/Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France

<sup>2</sup> Department Information and Communication Systems Engineering, University of the Aegean, Samos 83200, Greece; kavallieratou@aegean.gr

\* Correspondence: laurence.likforman@telecom-paristech.fr or likforman@telecom-paristech.fr

Received: 15 June 2018; Accepted: 15 June 2018; Published: 22 June 2018



**Keywords:** document image processing; preprocessing; document restoration; binarization; slant removal; text-line segmentation; handwriting recognition; indic/arabic/asian scripts; OCR; Video OCR; word spotting; retrieval; document datasets; performance evaluation; document annotation tools

---

The Special Issue “Document Image Processing” in the *Journal of Imaging* aims at presenting approaches which contribute to access the content of document images. These approaches are related to low level tasks such as image preprocessing, skew/slant corrections, binarization and document segmentation, as well as high level tasks such as OCR, handwriting recognition, word spotting or script identification. This special issue brings together 12 papers that discuss such approaches. The first three articles deal with historical document preprocessing. The work by Hanif et al. [1] aims at removing bleed-through using a non-linear model, and at reconstructing the background by an inpainting approach based on non-local patch similarity. The paper by Almeida et al. [2] proposes a new binarization approach that includes a decision-based process for finding the best threshold for each RGB channel. In the paper by Kavallieratou et al. [3], a segmentation-free approach based on the Wigner-Ville distribution is used to detect the slant of a document and correct it.

Once a document image is preprocessed, a next step described in the paper by Ghosh et al. [4] consists in separating text components from non-text ones, using a classifier based on LBP features. Following steps may consist in recognizing text components or searching from word queries. In the paper by Nashwan et al. [5] a holistic-based approach for the recognition of printed Arabic words is proposed, coupled with an efficient dictionary reduction. In the work by Nagendar et al. [6] it is shown that using a query specific fast Dynamic Time Warping distance, improves the Direct Query Classifier (DQC) word spotting system.

Deep neural network-based approaches are now widely used in the domain of document image processing, especially for the recognition of textual elements. The following papers also follow this trend. In the work by Jangid and Srivastava [7], deep convolutional networks trained layer-wise, are applied to the recognition of Devanagari characters. The paper by Kesiman et al. [8] is dedicated to southeast Asian scripts written on palm leaves. Character and word images are recognized by CNNs (Convolutional Neural Networks) and RNNs (Recurrent Neural Networks), respectively. Several binarization and text-line segmentation approaches are also benchmarked on these specific documents. The work by Granell et al. [9] describes an efficient text-line recognition system, based on CNN and stacks of RNNs, that has been developed for the recognition of historical Spanish documents. These documents include out-of-vocabulary ancient words which are handled by a language model based on sub-lexical units.

Annotated datasets are necessary to train systems or to evaluate the various tasks related to document image processing. In several papers published in this special issue, new datasets are released as well as open-source tools that are able to generate synthetic images. A dataset of indic

scripts is released in the paper by Mukhopadhyay et al. [10] and first results are provided with this dataset. The DocCreator software described in the paper by Journet et al. [11] creates additional document samples from input ones, using a degradation model. Such augmented data are used to train deep learning systems or to evaluate system performance. Document images can be extended to videos including text. The paper by Zayenne et al. [12] describes open-source tools for multiple document processing tasks: annotation of Arabic news videos, evaluation of text detection and text recognition. Authors also release the Activ2.0 database of Arabic videos and make it publicly available.

The guest editors would also like to thank all the authors that have submitted papers to this special issue, all the reviewers for their contribution, and the Journal of Imaging Editors.

**Author Contributions:** The two authors have equally contributed to the writing of this editorial.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hanif, M.; Tonazzini, A.; Savino, P.; Salerno, E. Non-Local Sparse Image Inpainting for Document Bleed-Through Removal. *J. Imaging* **2018**, *4*, 68. [[CrossRef](#)]
2. Almeida, M.; Lins, R.D.; Bernardino, R.; Jesus, D.; Lima, B. A New Binarization Algorithm for Historical Documents. *J. Imaging* **2018**, *4*, 27. [[CrossRef](#)]
3. Kavallieratou, E.; Likforman-Sulem, L.; Vasilopoulos, N. Slant Removal Technique for Historical Document Images. *J. Imaging* **2018**, *4*, 80. [[CrossRef](#)]
4. Ghosh, S.; Lahiri, D.; Bhowmik, S.; Kavallieratou, E.; Sarkar, R. Text/Non-Text Separation from Handwritten Document Images Using LBP Based Features: An Empirical Study. *J. Imaging* **2018**, *4*, 57. [[CrossRef](#)]
5. Nashwan, F.M.A.; Rashwan, M.A.A.; Al-Barhamtoshy, H.M.; Abdou, S.M.; Moussa, A.M. A Holistic Technique for an Arabic OCR System. *J. Imaging* **2018**, *4*, 6. [[CrossRef](#)]
6. Nagendar, G.; Ranjan, V.; Harit, G.; Jawahar, C.V. Efficient Query Specific DTW Distance for Document Retrieval with Unlimited Vocabulary. *J. Imaging* **2018**, *4*, 37. [[CrossRef](#)]
7. Jangid, M.; Srivastava, S. Handwritten Devanagari Character Recognition Using Layer-Wise Training of Deep Convolutional Neural Networks and Adaptive Gradient Methods. *J. Imaging* **2018**, *4*, 41. [[CrossRef](#)]
8. Kesiman, M.W.A.; Valy, D.; Burie, J.-C.; Paulus, E.; Suryani, M.; Hadi, S.; Verleysen, M.; Chhun, S.; Ogier, J.-M. Benchmarking of Document Image Analysis Tasks for Palm Leaf Manuscripts from Southeast Asia. *J. Imaging* **2018**, *4*, 43. [[CrossRef](#)]
9. Granell, E.; Chammas, E.; Likforman-Sulem, L.; Martínez-Hinarejos, C.-D.; Mokbel, C.; Cîrstea, B.-I. Transcription of Spanish Historical Handwritten Documents with Deep Neural Networks. *J. Imaging* **2018**, *4*, 15. [[CrossRef](#)]
10. Mukhopadhyay, A.; Singh, P.K.; Sarkar, R.; Nasipuri, M. A Study of Different Classifier Combination Approaches for Handwritten Indic Script Recognition. *J. Imaging* **2018**, *4*, 39. [[CrossRef](#)]
11. Journet, N.; Visani, M.; Mansencal, B.; Van-Cuong, K.; Billy, A. DocCreator: A New Software for Creating Synthetic Ground-Truthed Document Images. *J. Imaging* **2017**, *3*, 62. [[CrossRef](#)]
12. Zayene, O.; Touj, S.M.; Hennebert, J.; Ingold, R.; Ben Amara, N.E. Open Datasets and Tools for Arabic Text Detection and Recognition in News Video Frames. *J. Imaging* **2018**, *4*, 32. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).