International Journal on Document Analysis and Recognition (IJDAR) (2025) 28:327–328
https://doi.org/10.1007/s10032-025-00551-9

EDITORIAL

Special issue on advanced topics in document analysis (2025 ICDAR-IJDAR journal track)

Daniel Lopresti[1] · Dimosthenis Karatzas[2] · Xu-Cheng Yin[3]

Published online: 9 September 2025
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2025

The field of document analysis research has a long, rich history extending back for decades. At times, the progress has been slow but steady. Recently, however, we have seen huge leaps forward, due in part to tremendous advances in artificial intelligence and machine learning, including fundamentally new approaches to computer vision, natural language understanding, and the other basic techniques that underlie document analysis. With this context in mind, we are very pleased to bring you the fifth installment in the successful series that highlights papers from the Journal Track at the 2025 International Conference on Document Analysis and Recognition (ICDAR 2025) held in Wuhan, China.

After a broadly publicized open call for papers, we received a total of 46 submissions. The three guest editors of this volume screened all of them for their scope and their contributions to the field. We passed on those that met our expectations to a set of dedicated reviewers. Each submission was reviewed to the high standards of IJDAR and underwent a round of revisions as requested by the reviewers. In the end, we accepted 10 papers for the 2025 journal track. We briefly summarize the contributions of each of them here.

Several of the papers draw from problems in historical document analysis, which presents a myriad of interesting challenges. Josef Baloun, Martin Prantl, Ladislav Lenc, Jiří Martínek, and Pavel Král examine the problem of applying self-supervision techniques in the context of historical handwritten document analysis, with a particular focus on text line segmentation. They explore a range of pre-training setups for self-supervised features, and illustrate their approach using a wide variety of public datasets for training and three standard datasets for testing.

Nicholas R. Howe, Feiran Chang, Isabella Falbo, Tajhini Brown, and Aaron Hershkowitz examine the problem of recognizing text on Greek squeezes, which are paper casts taken from ancient stone inscriptions. The collection they are working with consists of approximately 30,000 squeezes from museums and archeological sites that are now housed at the Institute of Advanced Study in Princeton. Squeezes are inherently three-dimensional objects which presents its own interesting challenges. Beyond developing and testing a text recognition pipeline, they have also made their annotated dataset available to other researchers.

Loitongbam Gyanendro Singh and Stuart E. Middleton address the task of optical character recognition for large-scale historical document collections presented in tabular format. They employ a fine-tuning approach that utilizes contextual information from neighboring cells, reducing cascading failures and thereby enhancing the accuracy of table reconstruction tasks. The ByT5 language model is used for post-processing error correction. The dataset of interest consists of images of ship log books drawn from a number of sources dating from the 19th and 20th Centuries. A mix of handwritten and typed text can be found in the tabular data.

A new dataset of historical census records is the subject of the paper by José Andrés, Casey Wall, Solène Tarride, Mickaël Coustaty, Alejandro H. Toselli, and Enrique Vidal. Their PARES Dataset consists of 535 tables that present a unique set of challenges from French archives dating from the 16th through the 19th Centuries. Alongside this dataset, they examine three different baseline methods for information extraction, the first two employing traditional sequential approaches while the third uses an end-to-end model that directly extracts information from table images without prior row detection.

In their paper, Nimol Thuon, Jun Du, Panhapin Theang, and Ratana Thuon study the problem of palm leaf manuscript

✉ Daniel Lopresti
lopresti@cse.lehigh.edu

[1] Lehigh University, Pennsylvania, USA

[2] Universitat Autonoma de Barcelona, Barcelona, Spain

[3] University of Science and Technology, Beijing, China

analysis. They employ a low-intervention, dual-loop iterative framework designed to efficiently expand datasets and enhance glyph classification. Their findings include highlighting the value of dynamic stride shift patch tokenization, multi-scale attention mechanisms, and dual backbones including both vision and Swin transformers.

For the analysis of historical handwritten documents, a two-branch neural network architecture that integrates layout structural information to help deal with complex layout distributions is the topic of the paper by Fei Wu, Mathias Seuret, Martin Mayr, Florian Kordon, Jochen Zöllner, Sebastian Wind, Andreas Maier, and Vincent Christlein. Their network learns global layout structural information and local handwritten character details simultaneously. Experimental results are reported for a popular benchmark dataset and also for a newer dataset of historical manuscripts.

Turning next to the topic of handwriting recognition, Gang Yao, Liangrui Peng, Zhiyu Li, Tianqi Zhao, Kemeng Zhao, Ning Ding, and Yao Tao propose a multi-level style-controlled method for generating Chinese handwritten text lines in their contribution to this special issue. The goal is to move beyond simple character and word generation to provide more complete synthetic data which is more representative of the challenges found in handwritten text recognition. A special effort is made to manage radical-level style. They compare their approach to other state-of-the-art techniques.

In the area of document layout analysis, adapting the Monte Carlo Dropout technique from classification to layout analysis is the topic of the paper by Samuel B. Oliva-Bulpitt, Juan P. Martinez-Esteso, Alejandro Galan-Cuenca, Francisco J. Castellanos, and Antonio Javier Gallego. Their target application is identifying staff regions in images of historical music scores. They demonstrate the effectiveness of their approach on three datasets of scanned sheet music ranging from the 16th Century to the 20th Century, across a variety of musical notations.

Finally, two of our papers can best be described as addressing systems-level issues. The first, SlimDoc, is an approach introduced by Marcel Lamott, Muhammad Armaghan Shakir, Adrian Ulges, Yves-Noel Weweler, and Faisal Shafait for distilling lightweight student models for document understanding applications. The goal is to overcome some of the challenges associated with deploying document understanding models where resources, particularly labeled training data, are scarce. Their unsupervised approach to lightweight model distillation yields student model performance that approaches the teacher's, which they demonstrate on a range of standard datasets.

In our last paper for the special issue, Arthur Hemmer, Mickaël Coustaty, Nicola Bartolo, and Jean-Marc Ogier explore using syntactic-, task-, and domain-level validation to confirm that information extracted from transactional documents satisfies various arithmetic constraints. They note that the goal of neurosymbolics is to merge the controllability and interpretability of rule-based systems with the scalability and adaptability of probabilistic neural models. Beyond the methodology, their work also involved re-annotating two existing datasets which provide a further resource for others wishing to explore this area.

The research being done in document analysis today is impressive and we believe that the papers contained within this special issue present a wonderful overview of some of the most interesting topic areas now under study. We wish to thank all of the authors for choosing to share their work with us, and the reviewers for allowing us to take advantage of their expertise and devoting much time to the review process. In addition, we thank the IJDAR team members who made this special issue possible, including Katherine Moretti who is our primary contact, along with SimsonJohnpaul and Priya Verma. We also thank IJDAR's two other Editors-in-Chief, Koichi Kise and Josep Lladós, for their continued support of the journal and its ongoing success.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.