## **Lecture Notes in Computer Science**

## 12824

## Founding Editors

Gerhard Goos

Karlsruhe Institute of Technology, Karlsruhe, Germany

Juris Hartmanis

Cornell University, Ithaca, NY, USA

#### **Editorial Board Members**

Elisa Bertino

Purdue University, West Lafayette, IN, USA

Wen Gao

Peking University, Beijing, China

Bernhard Steffen

TU Dortmund University, Dortmund, Germany

Gerhard Woeginger

RWTH Aachen, Aachen, Germany

Moti Yung

Columbia University, New York, NY, USA

More information about this subseries at http://www.springer.com/series/7412

Josep Lladós · Daniel Lopresti · Seiichi Uchida (Eds.)

# Document Analysis and Recognition – ICDAR 2021

16th International Conference Lausanne, Switzerland, September 5–10, 2021 Proceedings, Part IV



Editors
Josep Lladós

Universitat Autònoma de Barcelona
Barcelona, Spain

Seiichi Uchida (D) Kyushu University Fukuoka-shi, Japan Daniel Lopresti 
Lehigh University
Bethlehem, PA, USA

ISSN 0302-9743 ISSN 1611-3349 (electronic) Lecture Notes in Computer Science ISBN 978-3-030-86336-4 ISBN 978-3-030-86337-1 (eBook) https://doi.org/10.1007/978-3-030-86337-1

LNCS Sublibrary: SL6 - Image Processing, Computer Vision, Pattern Recognition, and Graphics

#### © Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

### **Foreword**

Our warmest welcome to the proceedings of ICDAR 2021, the 16th IAPR International Conference on Document Analysis and Recognition, which was held in Switzerland for the first time. Organizing an international conference of significant size during the COVID-19 pandemic, with the goal of welcoming at least some of the participants physically, is similar to navigating a rowboat across the ocean during a storm. Fortunately, we were able to work together with partners who have shown a tremendous amount of flexibility and patience including, in particular, our local partners, namely the Beaulieu convention center in Lausanne, EPFL, and Lausanne Tourisme, and also the international ICDAR advisory board and IAPR-TC 10/11 leadership teams who have supported us not only with excellent advice but also financially, encouraging us to setup a hybrid format for the conference.

We were not a hundred percent sure if we would see each other in Lausanne but we remained confident, together with almost half of the attendees who registered for on-site participation. We relied on the hybridization support of a motivated team from the Lule University of Technology during the pre-conference, and professional support from Imavox during the main conference, to ensure a smooth connection between the physical and the virtual world. Indeed, our welcome is extended especially to all our colleagues who were not able to travel to Switzerland this year. We hope you had an exciting virtual conference week, and look forward to seeing you in person again at another event of the active DAR community.

With ICDAR 2021, we stepped into the shoes of a longstanding conference series, which is the premier international event for scientists and practitioners involved in document analysis and recognition, a field of growing importance in the current age of digital transitions. The conference is endorsed by IAPR-TC 10/11 and celebrates its 30th anniversary this year with the 16th edition. The very first ICDAR conference was held in St. Malo, France in 1991, followed by Tsukuba, Japan (1993), Montreal, Canada (1995), Ulm, Germany (1997), Bangalore, India (1999), Seattle, USA (2001), Edinburgh, UK (2003), Seoul, South Korea (2005), Curitiba, Brazil (2007), Barcelona, Spain (2009), Beijing, China (2011), Washington DC, USA (2013), Nancy, France (2015), Kyoto, Japan (2017), and Syndey, Australia (2019).

The attentive reader may have remarked that this list of cities includes several venues for the Olympic Games. This year the conference was be hosted in Lausanne, which is the headquarters of the International Olympic Committee. Not unlike the athletes who were recently competing in Tokyo, Japan, the researchers profited from a healthy spirit of competition, aimed at advancing our knowledge on how a machine can understand written communication. Indeed, following the tradition from previous years, 13 scientific competitions were held in conjunction with ICDAR 2021 including, for the first time, three so-called "long-term" competitions addressing wider challenges that may continue over the next few years.

Other highlights of the conference included the keynote talks given by Masaki Nakagawa, recipient of the IAPR/ICDAR Outstanding Achievements Award, and Mickaël Coustaty, recipient of the IAPR/ICDAR Young Investigator Award, as well as our distinguished keynote speakers Prem Natarajan, vice president at Amazon, who gave a talk on "OCR: A Journey through Advances in the Science, Engineering, and Productization of AI/ML", and Beta Megyesi, professor of computational linguistics at Uppsala University, who elaborated on "Cracking Ciphers with 'AI-in-the-loop': Transcription and Decryption in a Cross-Disciplinary Field".

A total of 340 publications were submitted to the main conference, which was held at the Beaulieu convention center during September 8–10, 2021. Based on the reviews, our Program Committee chairs accepted 40 papers for oral presentation and 142 papers for poster presentation. In addition, nine articles accepted for the ICDAR-IJDAR journal track were presented orally at the conference and a workshop was integrated in a poster session. Furthermore, 12 workshops, 2 tutorials, and the doctoral consortium were held during the pre-conference at EPFL during September 5–7, 2021, focusing on specific aspects of document analysis and recognition, such as graphics recognition, camera-based document analysis, and historical documents.

The conference would not have been possible without hundreds of hours of work done by volunteers in the organizing committee. First of all we would like to express our deepest gratitude to our Program Committee chairs, Joseph Lladós, Dan Lopresti, and Seiichi Uchida, who oversaw a comprehensive reviewing process and designed the intriguing technical program of the main conference. We are also very grateful for all the hours invested by the members of the Program Committee to deliver high-quality peer reviews. Furthermore, we would like to highlight the excellent contribution by our publication chairs, Liangrui Peng, Fouad Slimane, and Oussama Zayene, who negotiated a great online visibility of the conference proceedings with Springer and ensured flawless camera-ready versions of all publications. Many thanks also to our chairs and organizers of the workshops, competitions, tutorials, and the doctoral consortium for setting up such an inspiring environment around the main conference. Finally, we are thankful for the support we have received from the sponsorship chairs, from our valued sponsors, and from our local organization chairs, which enabled us to put in the extra effort required for a hybrid conference setup.

Our main motivation for organizing ICDAR 2021 was to give practitioners in the DAR community a chance to showcase their research, both at this conference and its satellite events. Thank you to all the authors for submitting and presenting your outstanding work. We sincerely hope that you enjoyed the conference and the exchange with your colleagues, be it on-site or online.

September 2021

Andreas Fischer Rolf Ingold Marcus Liwicki

### **Preface**

It gives us great pleasure to welcome you to the proceedings of the 16th International Conference on Document Analysis and Recognition (ICDAR 2021). ICDAR brings together practitioners and theoreticians, industry researchers and academics, representing a range of disciplines with interests in the latest developments in the field of document analysis and recognition. The last ICDAR conference was held in Sydney, Australia, in September 2019. A few months later the COVID-19 pandemic locked down the world, and the Document Analysis and Recognition (DAR) events under the umbrella of IAPR had to be held in virtual format (DAS 2020 in Wuhan, China, and ICFHR 2020 in Dortmund, Germany). ICDAR 2021 was held in Lausanne, Switzerland, in a hybrid mode. Thus, it offered the opportunity to resume normality, and show that the scientific community in DAR has kept active during this long period.

Despite the difficulties of COVID-19, ICDAR 2021 managed to achieve an impressive number of submissions. The conference received 340 paper submissions, of which 182 were accepted for publication (54%) and, of those, 40 were selected as oral presentations (12%) and 142 as posters (42%). Among the accepted papers, 112 had a student as the first author (62%), and 41 were identified as coming from industry (23%). In addition, a special track was organized in connection with a Special Issue of the International Journal on Document Analysis and Recognition (IJDAR). The Special Issue received 32 submissions that underwent the full journal review and revision process. The nine accepted papers were published in IJDAR and the authors were invited to present their work in the special track at ICDAR.

The review model was double blind, i.e. the authors did not know the name of the reviewers and vice versa. A plagiarism filter was applied to each paper as an added measure of scientific integrity. Each paper received at least three reviews, totaling more than 1,000 reviews. We recruited 30 Senior Program Committee (SPC) members and 200 reviewers. The SPC members were selected based on their expertise in the area, considering that they had served in similar roles in past DAR events. We also included some younger researchers who are rising leaders in the field.

In the final program, authors from 47 different countries were represented, with China, India, France, the USA, Japan, Germany, and Spain at the top of the list. The most popular topics for accepted papers, in order, included text and symbol recognition, document image processing, document analysis systems, handwriting recognition, historical document analysis, extracting document semantics, and scene text detection and recognition. With the aim of establishing ties with other communities within the concept of reading systems at large, we broadened the scope, accepting papers on topics like natural language processing, multimedia documents, and sketch understanding.

The final program consisted of ten oral sessions, two poster sessions, three keynotes, one of them given by the recipient of the ICDAR Outstanding Achievements Award, and two panel sessions. We offer our deepest thanks to all who contributed their time

and effort to make ICDAR 2021 a first-rate event for the community. This year's ICDAR had a large number of interesting satellite events as well: workshops, tutorials, competitions, and the doctoral consortium. We would also like to express our sincere thanks to the keynote speakers, Prem Natarajan and Beta Megyesi.

Finally, we would like to thank all the people who spent time and effort to make this impressive program: the authors of the papers, the SPC members, the reviewers, and the ICDAR organizing committee as well as the local arrangements team.

September 2021

Josep Lladós Daniel Lopresti Seiichi Uchida

## **Organization**

## **Organizing Committee**

**General Chairs** 

Andreas Fischer University of Applied Sciences and Arts Western

Switzerland, Switzerland

Rolf Ingold University of Fribourg, Switzerland

Marcus Liwicki Luleå University of Technology, Sweden

**Program Committee Chairs** 

Josep Lladós Computer Vision Center, Spain

Daniel Lopresti Lehigh University, USA Seiichi Uchida Kyushu University, Japan

**Workshop Chairs** 

Elisa H. Barney Smith
Umapada Pal
Boise State University, USA
Indian Statistical Institute, India

**Competition Chairs** 

Harold Mouchère University of Nantes, France

Foteini Simistira Luleå University of Technology, Sweden

**Tutorial Chairs** 

Véronique Eglin Institut National des Sciences Appliquées, France

Alicia Fornés Computer Vision Center, Spain

**Doctoral Consortium Chairs** 

Jean-Christophe Burie La Rochelle University, France

Nibal Nayef MyScript, France

**Publication Chairs** 

Liangrui Peng Tsinghua University, China

Fouad Slimane University of Fribourg, Switzerland

Oussama Zayene University of Applied Sciences and Arts Western

Switzerland, Switzerland

**Sponsorship Chairs** 

David Doermann University at Buffalo, USA

Koichi Kise Osaka Prefecture University, Japan Jean-Marc Ogier University of La Rochelle, France

### **Local Organization Chairs**

Jean Hennebert University of Applied Sciences and Arts Western

Switzerland, Switzerland

Anna Scius-Bertrand University of Applied Sciences and Arts Western

Switzerland, Switzerland

Sabine Süsstrunk École Polytechnique Fédérale de Lausanne,

Switzerland

Industrial Liaison

Aurélie Lemaitre University of Rennes, France

Social Media Manager

Linda Studer University of Fribourg, Switzerland

## **Program Committee**

#### **Senior Program Committee Members**

Apostolos Antonacopoulos University of Salford, UK

Xiang Bai Huazhong University of Science and Technology,

China

Michael Blumenstein University of Technology Sydney, Australia

Jean-Christophe Burie University of La Rochelle, France University of La Rochelle, France University of Rennes, France

Andreas Dengel DFKI, Germany

Gernot Fink TU Dortmund University, Germany

Basilis Gatos Demokritos, Greece Nicholas Howe Smith College, USA

Masakazu Iwamura Osaka Prefecture University, Japan

C. V. Javahar IIIT Hyderabad, India

Lianwen Jin South China University of Technology, China

Dimosthenis Karatzas Computer Vision Center, Spain Laurence Likforman-Sulem Télécom ParisTech, France

Cheng-Lin Liu Chinese Academy of Sciences, China

Angelo Marcelli University of Salerno, Italy Simone Marinai University of Florence, Italy

Wataru Ohyama Saitama Institute of Technology, Japan Luiz Oliveira Federal University of Parana, Brazil

Liangrui Peng Tsinghua University, China Ashok Popat Google Research, USA

Partha Pratim Roy Indian Institute of Technology Roorkee, India

Marçal Rusiñol Computer Vision Center, Spain

Robert Sablatnig Vienna University of Technology, Austria

Marc-Peter Schambach Siemens, Germany

Srirangaraj Setlur University at Buffalo, USA

Faisal Shafait National University of Sciences and Technology, India

Nicole Vincent Paris Descartes University, France

Jerod Weinman Grinnell College, USA

Richard Zanibbi Rochester Institute of Technology, USA

#### **Program Committee Members**

Sébastien Adam Nadir Farah

Irfan Ahmad Rafael Ferreira Mello

Sheraz Ahmed Miguel Ferrer Younes Akbari Julian Fierrez

Musab Al-Ghadi Francesco Fontanella Alireza Alaei Alicia Fornés

Eric Anquetil Volkmar Frinken
Srikar Appalaraju Yasuhisa Fujii
Elisa H. Barney Smith Akio Fujiyoshi
Abdel Belaid Liangcai Gao
Mohammed Faouzi Benzeghiba Utpal Garain
Anurag Bhardwaj C. Lee Giles

Ujjwal Bhattacharya Romain Giot Alceu Britto Lluis Gomez

Jorge Calvo-Zaragoza Petra Gomez-Krämer

Chee Kheng Ch'Ng
Sukalpa Chanda
Bidyut B. Chaudhuri
Jin Chen

Emilio Granell
Mehdi Hamdani
Gaurav Harit
Ehtesham Hassan

Youssouf Chherawala Anders Hast Hojin Cho Sheng He Nam Ik Cho Jean Hennebert Vincent Christlein Pierre Héroux Christian Clausner Laurent Heutte Nina S. T. Hirata Florence Cloppet Donatello Conte Tin Kam Ho Kenny Davila Kaizhu Huang Claudio De Stefano Oiang Huo

Sounak Dev Donato Impedovo Moises Diaz Reeve Ingle David Doermann Brian Kenji Iwana Antoine Doucet Motoi Iwata Fadoua Drira Antonio Jimeno Jun Du Slim Kanoun Véronique Eglin Vassilis Katsouros Jihad El-Sana Ergina Kavallieratou

Jonathan Fabrizio Klara Kedem

Christopher Kermorvant Khurram Khurshid Soo-Hyung Kim Koichi Kise Florian Kleber Pramod Kompalli

Alessandro Lameiras Koerich

Bart Lamiroy Anh Le Duc Frank Lebourgeois Gurpreet Lehal

Byron Leite Dantas Bezerra

Aurélie Lemaitre Haifeng Li Zhouhui Lian Minghui Liao Rafael Lins Wenyin Liu Lu Liu

Georgios Louloudis

Yue Lu Xiaoqing Lu

Muhammad Muzzamil Lugman

Sriganesh Madhvanath Muhammad Imran Malik R. Manmatha

Volker Märgner
Daniel Martín-Albo

Carlos David Martinez Hinarejos

Minesh Mathew Maroua Mehri Carlos Mello Tomo Miyazaki Momina Moetesum

Harold Mouchère

Masaki Nakagawa Nibal Nayef

Atul Negi Clemens Neudecker

Cuong Tuan Nguyen Hung Tuan Nguyen Journet Nicholas Jean-Marc Ogier Shinichiro Omachi Umapada Pal

Shivakumara Palaiahnakote

Thierry Paquet Swapan Kr. Parui Antonio Parziale Antonio Pertusa Giuseppe Pirlo Réjean Plamondon Stefan Pletschacher Utkarsh Porwal

Vincent Poulain D'Andecy

Ioannis Pratikakis
Joan Puigcerver
Siyang Qin
Irina Rabaev
Jean-Yves Ramel
Oriol Ramos Terrades
Romain Raveaux
Frédéric Rayar
Ana Rebelo
Pau Riba
Kaspar Riesen
Christophe Rigaud

Syed Tahseen Raza Rizvi

Leonard Rothacker Javad Sadri

Rajkumar Saini Joan Andreu Sanchez

K. C. Santosh Rosa Senatore Amina Serir Mathias Seuret

Badarinath Shantharam

Imran Siddiqi Nicolas Sidère

Foteini Simistira Liwicki

Steven Simske Volker Sorge

Nikolaos Stamatopoulos

Bela Stantic H. Siegfried Stiehl Daniel Stoekl Ben Ezra

Tonghua Su Tong Sun Yipeng Sun Jun Sun

Suresh Sundaram Salvatore Tabbone

Kazem Taghva Ryohei Tanaka

Christopher Tensmeyer Kengo Terasawa Ruben Tolosana Alejandro Toselli

Cao De Tran Szilard Vajda Ernest Valveny Marie Vans Eduardo Vellasques

Ruben Vera-Rodriguez Christian Viard-Gaudin Mauricio Villegas

Qiu-Feng Wang

Da-Han Wang Curtis Wigington Liang Wu

Mingkun Yang Xu-Cheng Yin

Fei Yin

Guangwei Zhang Heng Zhang Xu-Yao Zhang Yuchen Zheng Guoqiang Zhong

Yu Zhou Anna Zhu Majid Ziaratban

# **Contents – Part IV**

Scene	<b>Text</b>	<b>Detection</b>	and	Recognition
-------	-------------	------------------	-----	-------------

HRRegionNet: Chinese Character Segmentation in Historical Documents with Regional Awareness	3
Chia-Wei Tang, Chao-Lin Liu, and Po-Sen Chiu	
Fast Text vs. Non-text Classification of Images	18
Mask Scene Text Recognizer	33
Rotated Box Is Back: An Accurate Box Proposal Network for Scene	40
Text Detection	49
Heterogeneous Network Based Semi-supervised Learning for Scene	
Text Recognition	64
Scene Text Detection with Scribble Line	<b>7</b> 9
EEM: An End-to-end Evaluation Metric for Scene Text Detection	
and Recognition	95
SynthTIGER: Synthetic Text Image GEneratoR Towards Better Text	
Recognition Models	109
Fast Recognition for Multidirectional and Multi-type License Plates	
with 2D Spatial Attention	125

A Multi-level Progressive Rectification Mechanism for Irregular Scene Text Recognition	140
Qianying Liao, Qingxiang Lin, Lianwen Jin, Canjie Luo, Jiaxin Zhang, Dezhi Peng, and Tianwei Wang	
Representation and Correlation Enhanced Encoder-Decoder Framework	
for Scene Text Recognition	156
FEDS - Filtered Edit Distance Surrogate	171
Bidirectional Regression for Arbitrary-Shaped Text Detection	187
Document Classification	
VML-HP: Hebrew Paleography Dataset	205
Open Set Authorship Attribution Toward Demystifying Victorian Periodicals	221
Sarkhan Badirli, Mary Borgo Ton, Abdulmecit Gungor, and Murat Dundar	
A More Effective Sentence-Wise Text Segmentation Approach Using	22/
BERT	236
Data Augmentation for Writer Identification Using a Cognitive Inspired	251
Model	251
Key-Guided Identity Document Classification Method by Graph Attention	265
Network	267
Document Image Quality Assessment via Explicit Blur and Text Size	281
Estimation	261
Analyzing the Potential of Zero-Shot Recognition for Document Image	200
Classification	293

Contents – Part IV	xvii
Gender Detection Based on Spatial Pyramid Matching	305
EDNets: Deep Feature Learning for Document Image Classification Based on Multi-view Encoder-Decoder Neural Networks	318
Fast End-to-End Deep Learning Identity Document Detection, Classification and Cropping	333
Gold-Standard Benchmarks and Data Sets	
Image Collation: Matching Illustrations in Manuscripts	351
Revisiting the Coco Panoptic Metric to Enable Visual and Qualitative Analysis of Historical Map Instance Segmentation	367
A Large Multi-target Dataset of Common Bengali Handwritten Graphemes.  Samiul Alam, Tahsin Reasat, Asif Shahriyar Sushmit, Sadi Mohammad Siddique, Fuad Rahman, Mahady Hasan, and Ahmed Imtiaz Humayun	383
GNHK: A Dataset for English Handwriting in the Wild	399
Personalizing Handwriting Recognition Systems with Limited	
User-Specific Samples	413
An Efficient Local Word Augment Approach for Mongolian Handwritten Script Recognition	429
IIIT-INDIC-HW-WORDS: A Dataset for Indic Handwritten Text Recognition Santhoshini Gongidi and C. V. Jawahar	444
Historical Document Analysis	
AT-ST: Self-training Adaptation Strategy for OCR in Domains with Limited Transcriptions	463

xviii

Pratik Kayal, Mrinal Anand, Harsh Desai, and Mayank Singh

## Contents - Part IV

 $\mathbf{X}\mathbf{X}$ 

ICDAR 2021 Competition on Multimodal Emotion Recognition	
on Comics Scenes	767
Nhu-Van Nguyen, Xuan-Son Vu, Christophe Rigaud, Lili Jiang,	
and Jean-Christophe Burie	
ICDAR 2021 Competition on Mathematical Formula Detection	783
Dan Anitei, Joan Andreu Sánchez, José Manuel Fuentes,	
Roberto Paredes, and José Miguel Benedí	
Author Index	797