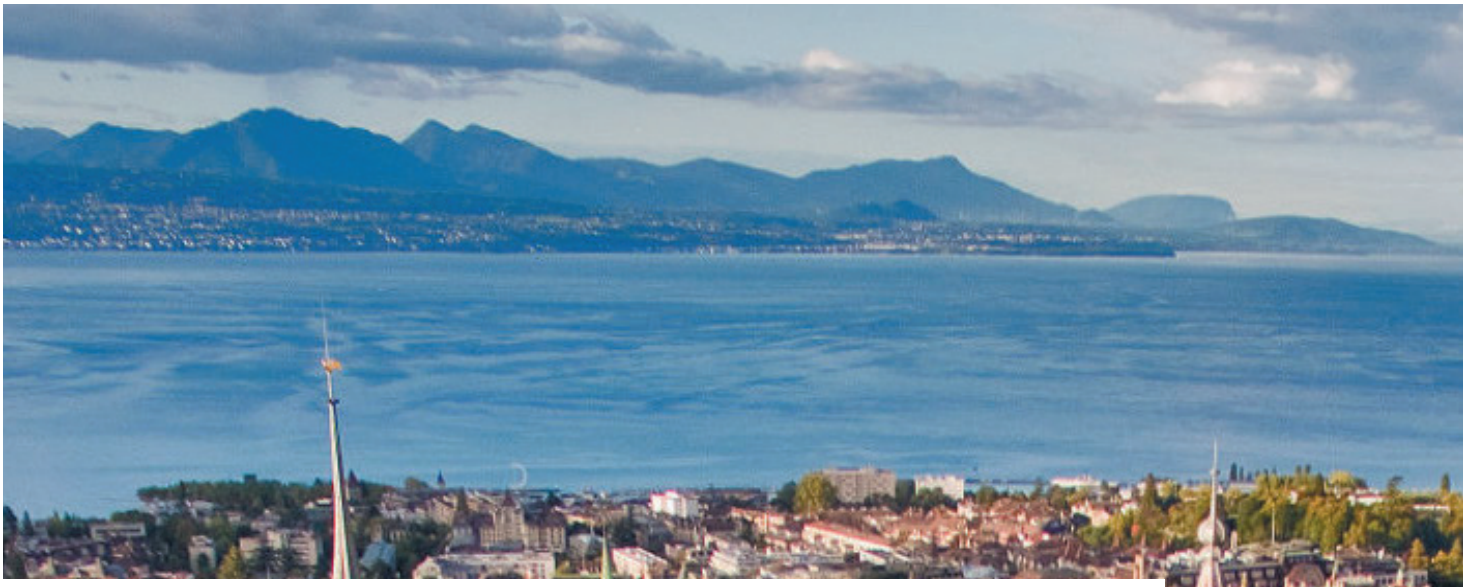


# Program

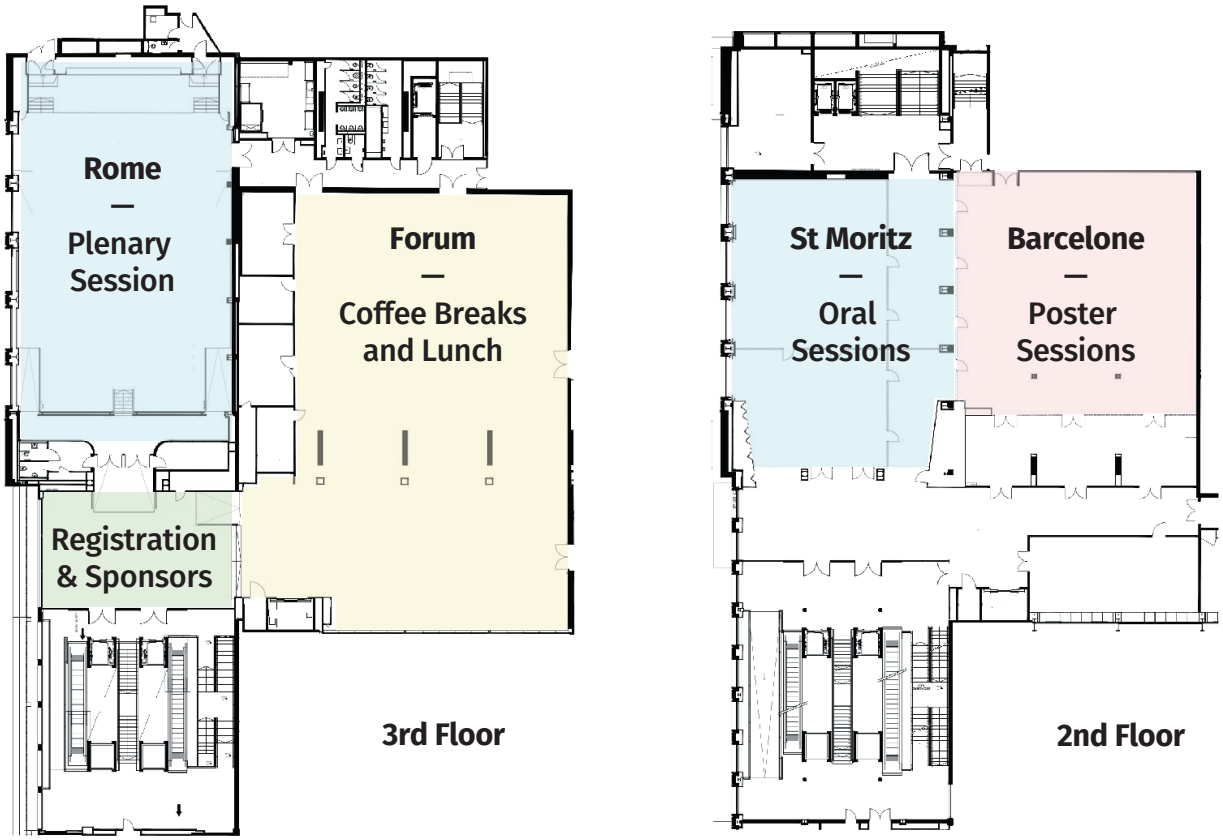
16th International Conference on Document Analysis and Recognition

5 – 10 September 2021

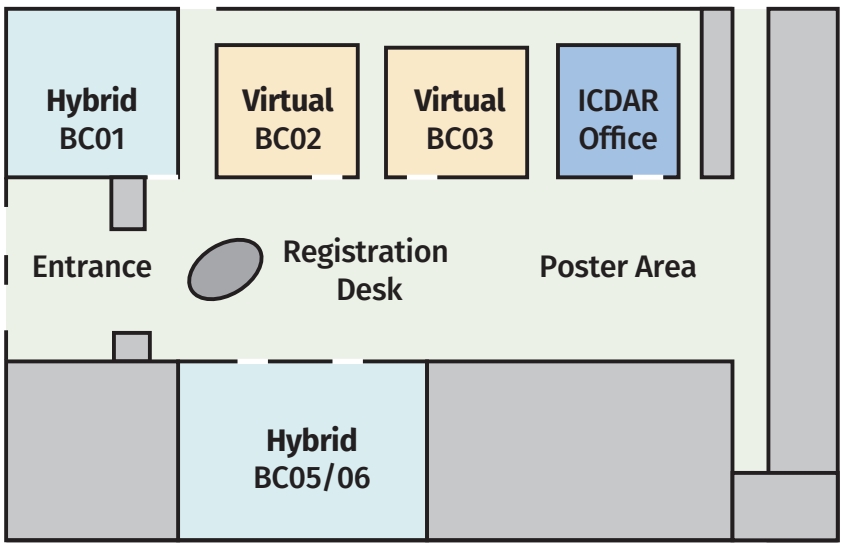
ICDAR

LAUSANNE 2021

# Maps of Beaulieu Convention Center




# Map of BC Building @ EPFL



## Foreword

Our warmest welcome to ICDAR 2021, the 16th IAPR International Conference on Document Analysis and Recognition, which is held in Switzerland for the first time this year. Organizing an international conference of significant size during the COVID pandemic, with the goal of welcoming at least a part 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 foremost our local partners, namely the Beaulieu convention center in Lausanne, the 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.

At the time of the writing, we are still not a hundred percent sure if we will see each other in Lausanne but we remain confident, together with almost half of the attendees who have registered for on-site participation. We will rely on the hybridization support of a motivated team from the Luleå University of Technology during the pre-conference, and the 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 are not able to travel to Switzerland this year. We wish you an exciting virtual conference week, and hope to see you soon in person again at another event of the active DAR community.

With ICDAR 2021, we step 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 was held in St Malo, France in 1991, followed by Tsukuba, Japan ('93), Montreal, Canada ('95), Ulm, Germany ('97), Bangalore, India ('99), Seattle, USA ('01), Edinburgh, UK ('03), Seoul, South Korea ('05), Curitiba, Brazil ('07), Barcelona, Spain ('09), Beijing, China ('11), Washington DC, USA ('13), Nancy, France ('15), Kyoto, Japan ('17), and Syndey, Australia in 2019.

The attentive reader may have remarked that this list of cities includes several venues for olympic games. This year the conference will be hosted in Lausanne, the Olympic City with the headquarters of the International Olympic Committee. Not unlike the athletes who are currently competing in Tokyo, Japan, the researchers will profit from a healthy spirit of competition, aiming at advancing our knowledge on how a machine can understand written communication. Indeed, following the tradition from previous years, 13 scientific competitions are held in conjunction with ICDAR 2021, including for the first time 3 so-called "long-term" competitions, addressing wider challenges that may continue over the next years.

Other highlights of the conference include 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 will talk about "OCR: A Journey through Advances in the Science, Engineering, and Productization of AI/ML", and Beáta Megyesi, professor of computational linguistics at Uppsala University, who will elaborate on "Cracking Ciphers with 'AI-in-the-loop': Transcription and Decryption in a Cross-Disciplinary Field".

A total of 340 publications have been submitted to the main conference, which will be held at the Beaulieu convention center from September 8–10. Based on the reviews, our Program Committee Chairs have accepted 40 papers for oral presentation and 142 papers for poster presentation. In addition, 9 articles accepted for the ICDAR–IJDAR journal track will be presented orally at the conference and 1 workshop will be

integrated in a poster session. Furthermore, 12 workshops, 2 tutorials, and the doctoral consortium will be held during the pre-conference at the EPFL from September 5–7, 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 Llados, Dan Lopresti, and Seiichi Uchida, who have overseen a comprehensive reviewing process and who have 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 have 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 would like to thank for the support we have received from the Sponsorship Chairs, from our valued sponsors, and from our Local Organization Chairs, which together enabled us to put in the extra effort required for a hybrid conference setup.

We are organizing ICDAR 2021 foremost for you, authors of this conference and its satellite events. Thank you for submitting and presenting your outstanding research work. We sincerely hope that you will enjoy the conference and the exchange with your colleagues, be it on-site or online. Let the conference begin!

Andreas Fischer
Rolf Ingold
Marcus Liwicki

## Preface

It gives us great pleasure to welcome you to the proceedings of the Sixteenth 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 in September 2019. A few months later the pandemic situation locked down the world, and the Document Analysis and Recognition (DAR) events under the umbrella of the IAPR had to be held in virtual format (DAS 2020 in Wuhan, China; and ICFHR 2020 in Dortmund, Germany). ICDAR 2021 is held in Lausanne, Switzerland in a hybrid mode. Thus, it opens 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 has 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 orals (12%) and 142 as posters (42%). Among the accepted papers, 112 had a student as first author (62%), and 41 were identified as coming from industry (23%). In addition, a special track was organized consisting of 9 oral papers, accepted for publication in 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 9 accepted papers are 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, totalling more than 1,000 reviews. We recruited 30 senior PC members (SPCs) 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 are represented, with China, India, France, United States, Japan, Germany, and Spain at top of the list. The most popular topics for accepted papers, in order, include: 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 consists of ten oral sessions, two poster sessions, three keynotes, one of them given by the recipient of the ICDAR Outstanding Achievements Awards, 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 ICDAR has 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 Beáta 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 senior PC members, the reviewers, and the ICDAR organizing committee as well as the local arrangements team.

Josep Lladós
Daniel Lopresti
Seiichi Uchida

## Organizing Committee

### General Chairs

Andreas Fischer & Univ. of Applied Sciences and Arts Western 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 Barney Smith & Boise State University, USA
Umapada Pal & 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, Beijing, China
Fouad Slimane & University of Fribourg, Switzerland
Oussama Zayene & University of Applied Sciences and Arts Western 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 & Univ. of Applied Sciences and Arts Western Switzerland
Anna Scius-Bertrand & University of Applied Sciences and Arts Western 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

### Senior Program Committee Members

Apostolos Antonacopoulos & University of Salford, UK
Xiang Bai & Huazhong Univ. of Science and Technology, China
Michael Blumenstein & University of Technology, Sydney, Australia
Jean-Christophe Burie & University of La Rochelle, France
Mickaël Coustaty & University of La Rochelle, France
Bertrand Coüasnon & University of Rennes, France
Andreas Dengel & DFKI, Germany
Gernot Fink & TU of 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 & Telecom-ParisTech, France
Cheng-Lin Liu & Chinese Academy of Sciences, China
Angelo Marcelli & University of Salerno, Italia
Simone Marinai & University of Florence, Italy
Wataru Ohyama & Saitama Institute of Technology, Japan
Luiz Oliveira & Federal University of Parana, Brazil
Liangrui Peng & Tsinghua University, Beijing, 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, SUNY, USA
Faisal Shafait & National Univ. of Sciences and Technology, India
Nicole Vincent & University Paris Descartes, France
Jerod Weinman & Grinnell College, USA
Richard Zanibbi & Rochester Institute of Technology, USA

### Program Committee Members

Sébastien Adam
Irfan Ahmad
Sheraz Ahmed
Younes Akbari
Musab Al-Ghadi
Alireza Alaei
Eric Anquetil
Srikar Appalaraju
Elisa H. Barney Smith
Abdel Belaid
Mohammed Faouzi Benzeghiba
Anurag Bhardwaj
Ujjwal Bhattacharya
Alceu Britto
Jorge Calvo-Zaragoza
Chee Kheng Ch'Ng
Sukalpa Chanda
Bidyut B. Chaudhuri
Jin Chen
Youssouf Chherawala
Hojin Cho
Nam Ik Cho
Vincent Christlein
Christian Clausner
Florence Cloppet
Donatello Conte
Kenny Davila
Claudio De Stefano
Sounak Dey
Moises Diaz
David Doermann
Antoine Doucet
Fadoua Drira
Jun Du
Véronique Eglin
Jihad El-Sana
Jonathan Fabrizio
Nadir Farah
Rafael Ferreira Mello
Miguel Ferrer
Julian Fierrez
Francesco Fontanella
Alicia Fornés
Volkmar Frinken
Yasuhisa Fujii
Akio Fujiyoshi
Liangcai Gao
Utpal Garain
C. Lee Giles
Romain Giot
Lluis Gomez
Petra Gomez-Krämer
Emilio Granell
Mehdi Hamdani
Gaurav Harit
Ehtesham Hassan
Anders Hast
Sheng He
Jean Hennebert
Pierre Héroux
Laurent Heutte
Nina S. T. Hirata
Tin Kam Ho
Kaizhu Huang
Qiang Huo
Donato Impedovo
Reeve Ingle
Brian Kenji Iwana
Motoi Iwata
Antonio Jimeno
Slim Kanoun
Vassilis Katsouros
Ergina Kavallieratou
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 Luqman
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

# Masaki Nakagawa

## Toward automatic recognition and scoring of handwritten descriptive answers

Starting from the brief history of offline and online handwriting recognition, I will talk about my experiences of joint projects with companies, which might be useful for the audience. Then I will present the latest challenge to automate scoring of handwritten answers for descriptive questions. Descriptive questions can test deep understanding and problem-solving ability of examinees much better than selection-type questions asked by most of CBTs and encourage examinees to think rather than select. Full-automatic recognition and scoring of descriptive answers provides immediate feedback to examinees to review their answers when examinees can confirm scoring, while semi-automatic, or computer assisted scoring, provides reliable scoring when examinees cannot confirm scoring. Both decrease time and effort for examiners or teachers to score exams. My dream is to unify online recognition of handwritten answers from tablets and offline recognition from scanners except for several early-stage layers in DNN. The same DNN architecture may learn to recognize Japanese, English, and Math answers. The DNN for handwritten answer recognition will output reliable features to cluster answers for semi-automatic scoring. The DNN for handwriting recognition could be even merged with that for automatic scoring and trained end-to-end. An initial attempt for Japanese language questions for 120,000 examinees shows a promising result.

**Short Bio**

Masaki Nakagawa is emeritus professor at the Tokyo University of Agriculture & Technology. He has been working on handwriting recognition, pen-based user interfaces and educational applications. Since 1980s, he has been collaborating with many companies and has contributed to develop handwriting recognizers for real commercial use. In 2011, he founded a start-up iLabo, which now sells the best handwriting recognizers for touch-based smart phones, tablets and so on in Japan. In 1990, he also introduced User Interfaces for tablet devices and developed several educational applications using various sizes of tablets. His U.S. patents to scroll the window in proportion to the pen speed, called “touch scroll”, were sold from his university for the highest amount among all the Japanese universities in the fiscal year 2010. He is also working on historical document processing to read excavated documents from the Heijo palace (the capital in the 7th century) in Nara, Japan, and to read Chu Nom documents in Vietnam. He received the Minister of Education and Science award of Japan and the Contribution award from the Tokyo Metropolitan Government both in 2016. He is a fellow of IAPR, IEICE (Institute of Electronics, Information and Communication Engineers, Japan) and IPSJ (Information Processing Society of Japan).

**Wednesday 8 September, 09:15–10:00**
**Room “Rome”, second floor**

# Mickaël Coustaty

## Complex Document Analysis and Its Impact

Documents are part of our daily life, in a personal or professional way. Even if they seem easy to handle, to analyze and to manipulate for a human, current trends in document analysis tend to address more and more complex information, mixing textual content (typewritten, handwritten), visual content (logo, signature, pictures) and their semantic. For the textual content, OCR is a common step to process documents and extract their content. Even if its performance is getting better and better, errors remain and may impact further steps. A lot of document image analysis techniques have been proposed over the last 30 years, which are still not satisfying as documents are not only composed of visual elements. The best current approaches tend to join text and images in order to achieve solutions for multimodal analysis of documents. The presentation will propose some results obtained by combining visual and textual elements, while trying to deal with the need of large annotated datasets. I will also share my thoughts on some ideas that could be addressed to open our community to widen our field and extend our future work on complex document analysis.

**Short Bio**

Mickaël Coustaty is a tenured Associate Professor at the L3i laboratory of La Rochelle University since 2015. He is specialized in complex document analysis by using multimodal approaches mixing textual and visual content. He initially worked on historical documents, extended its techniques to administrative documents, and included NLP approaches. He always worked in collaboration with experts or end-users in order to extract information, to index the content and to assess its relevance / trustability. He has been participating in ten externally funded projects, cross-disciplinary research projects and currently serves as the head of a joint private-public lab between L3i and the Yooz Company, co-funded by the French National Research Agency, the Région Nouvelle-Aquitaine and the Yooz Company. Finally, since 2016, he is in charge of a Master degree in digital law where he proposed the first French Master degree mixing computer science classes and law classes in collaboration with the French National Trusted Third-party association in order to connect different fields and to bond academic and industrial fields.

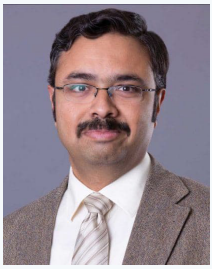
**Wednesday 8 September, 10:00–10:30**
**Room "Rome", second floor**

# Prem Natarajan

## OCR: A Journey through Advances in the Science, Engineering, and Productization of AI/ML

From the very early years of AI, the problem of optical character recognition (OCR) has captured the imagination of researchers; Selfridge and Neisser presented an approach for OCR of hand printed characters in 1960. During last three decades, optical character recognition (OCR) technology for machine printed and handwritten text has evolved in significant ways — from script-specific techniques to script-independent methodologies, and from segmentation-based techniques to hidden Markov models to deep learning. In my talk, I will present my perspective on that evolution and it's interplay with concomitant advances in speech recognition, natural language processing, and computer vision. The presentation will include a discussion of some practical, even if off the beaten path, applications of OCR technology, including work done in partnership with the census bureau in applying a deep learning based OCR framework to census forms. I will also share my views on some of the most interesting open problems in the field of OCR and document processing. The presentation will conclude with a few comments about one of my current areas of research interests — fairness in AI and machine learning.

**Thursday 9 September, 13:00–14:00**
**Room "Rome", second floor**



**Short Bio**

Prem Natarajan is a Vice President at Amazon where he leads research and engineering efforts in dialog systems, natural language understanding, and multimodal/multimedia technologies in the Alexa AI organization. Prior to joining Amazon in June 2018, he was with the University of Southern California (USC) where he was Senior Vice Dean of Engineering in the Viterbi School of Engineering, Executive Director of the Information Sciences Institute (a 300-person R&D organization), and Research Professor of computer science with distinction. Prior to that, as Executive VP and Principal Scientist at Raytheon BBN Technologies, he led the speech, language, and multimedia business unit, which included research and development operations, and commercial products for real-time multimedia monitoring, document analysis, and information extraction. During his tenure at USC and at BBN, Natarajan directed R&D efforts in speech recognition, natural language processing, computer vision, and other applications of machine learning. While at USC, he led nationally influential DARPA and IARPA sponsored research efforts in biometrics/face recognition, OCR, NLP, media forensics, and forecasting. Most recently at Amazon, he helped to launch the Fairness in AI (FAI) program — a collaborative effort between NSF and Amazon for funding fairness focused research efforts in US Universities.

# Beáta Megyesi

## Cracking Ciphers with “AI-in-the-loop”: Transcription and Decryption in a Cross-Disciplinary Field

Accurate transcription of hand-written texts in images is indispensable in many research areas in digital humanities. Manual transcription is error-prone, time-consuming, and expensive to produce. Historical texts with their specific textual qualities require expert knowledge and trained eyes. During the past years, image processing applied to hand-written historical text documents to provide transcription output has been shown great opportunities, but also challenges for users. How can users without knowledge in AI in general and HTR in particular transcribe hand-written documents efficiently with ”AI-in-the- loop”? In my talk, I will focus on encrypted manuscripts from Early Modern times with various symbols systems, hand-writing styles, and languages. The point of departure is the DECRYPT project, aiming at the creation of resources and tools for historical cryptology by bringing the expertise of various disciplines together for collecting images of ciphers and keys, to transcribe them, and to decrypt and contextualize those. I will give an overview of the project, the methods we use to solve various problems from transcription to decryption including historical corpora and natural language processing methods.

**Short Bio**

Beáta Megyesi is a professor of computational linguistics at Uppsala University, the former head of department at the Dept. Linguistics and Philology, Uppsala University, Sweden, and the current president of the North European Association for Language Technology (NEALT). She is specialized in digital philology and natural language processing with a special interest in the automatic analysis of non-standard, noisy language data, from text produced by language learners to historical texts and encrypted documents. She has been participating in ten externally funded, cross-disciplinary research projects and currently serves as the PI of the DECRYPT project, financed by the Swedish Research Council (grant 2018-06074). Bea received her Ph.D. in speech communication from the Royal Institute of Technology (KTH) in Stockholm, Sweden.

**Friday 10 September, 09:00–10:00**
**Room “Rome”, second floor**

# Pre-Conference Program · EPFL

| | Hybrid | | Virtual | |
|---|---|---|---|---|
| | Room BC05/06 | Room BC01 | Room BC02 | Room BC03 |
| | **Sunday · 5 September 2021** | | | |
| 08:30 | Registration | | | |
| 09:00 | **GREC** | | **OST** | |
| 10:30 | Coffee Break | | | |
| 11:00 | **GREC** | | **OST** | |
| 12:30 | Lunch (not included) | | | |
| 14:00 | **GREC** | **IWCDF** | **GLESDO** | |
| 15:30 | Coffee Break | | | |
| 16:00 | **GREC** | **IWCDF** | **GLESDO** | |
| 17:30 | | | | |
| | **Monday · 6 September 2021** | | | |
| 08:30 | Registration | | | |
| 09:00 | **HIP** | | **ASAR** | |
| 10:30 | Coffee Break | | | |
| 11:00 | **HIP** | | **ASAR** | |
| 12:30 | Lunch (not included) | | | |
| 14:00 | **HIP** | **CBDAR** | **DIL** | **DocVQA** |
| 15:30 | Coffee Break | | | |
| 16:00 | **HIP** | **CBDAR** | **DIL** | **DocVQA** |
| 17:30 | | | | |
| 18:00 | | | | |
| | **Tuesday · 7 September 2021** | | | |
| 08:00 | Registration | | | |
| 08:30 | **FDAR** | **Paleo** | **WML** | |
| 10:30 | Coffee Break | | | |
| 11:00 | **FDAR** | **Paleo** | **WML** | |
| 12:30 | Lunch (not included) | | | |
| 14:00 | **Doctoral Consortium** | **Tutorial: NewsEye** | **WML** | **Tutorial: NLP Canvas** |
| 15:30 | Coffee Break | | | |
| 16:00 | **Doctoral Consortium** | **Tutorial: NewsEye** | **WML** | **Tutorial: NLP Canvas** |
| 17:30 | | | | |
| 18:30 | Pre-Conference Apéro | | | |

# Main Conference Program · Beaulieu

| | Room "Rome" · Third floor | Room "St-Moritz" · Second floor | Room "Barcelone" · First floor |
|---|---|---|---|
| | **Wednesday · 8 September 2021** | | |
| 08:00 | Registration & Coffee | | |
| 08:45 | **Opening** | | |
| 09:15 | **Award Keynote: Masaki Nakagawa** | | |
| 10:00 | **Award Keynote: Mickaël Coustaty** | | |
| 10:30 | Coffee Break | | |
| 11:00 | **Oral Session 1: Journal Track 1** | **Oral Session 2: Journal Track 2** | |
| 12:40 | Lunch, Forum, 2nd floor | | |
| 14:00 | **Oral Session 3: Historical Document Analysis 1** | **Oral Session 4: Document Analysis Systems** | |
| 16:00 | Coffee Break | | |
| 16:30 | Group Picture | | |
| 16:45 | **TC10/11 Meeting** | | |
| 17:30 | Reception and Gala Dinner, Olympic Museum | | |
| 18:00 – 23:00 | | | |
| | **Thursday · 9 September 2021** | | |
| 08:00 | Registration & Coffee | | |
| 08:30 | | | **Poster: Session 1** |
| 10:00 | Coffee Break | | |
| 10:30 | | | **Poster: Session 2** |
| 12:00 | Lunch, Forum, 2nd floor | | |
| 13:00 | **Keynote: Prem Natarajan** | | |
| 14:00 | Coffee Break | | |
| 14:30 | **Oral Session 5: Handwriting Recognition** | **Oral Session 6: Scene Text Detection and Recognition** | |
| 16:30 | Coffee Break | | |
| 17:00 | **Industrial Session** | | |
| 17:45 | | | |
| | **Friday · 10 September 2021** | | |
| 08:30 | Registration & Coffee | | |
| 09:00 | **Keynote: Beáta Megyesi** | | |
| 10:00 | Coffee Break | | |
| 10:30 | **Oral Session 7: Historical Document Analysis 2** | **Oral Session 8: Document Image Processing** | |
| 12:00 | **Surprise Session** | | |
| 12:15 | Lunch, Forum, 2nd floor | | |
| 14:00 | **Oral Session 9: NLP for Document Understanding** | **Oral Session 10:Graphics, Diagram, and Math Recognition** | |
| 15:30 | Coffee Break | | |
| 16:00 | **Competition Session** | | |
| 16:45 | **ICDAR Awards Ceremony & Closing** | | |
| 17:15 | Goodby Apéro | | |
| 17:30 – 18:30 | | | |

# Wednesday · 9 September

## Oral Session 1: Journal Track 1

O1.1 11:00 — 11:20

**Learning from similarity and information extraction from structured documents**
Martin Holeček

O1.2 11:20 — 11:40

**Learning-free Pattern Detection for Manuscript Research: An Efficient Approach Toward Making Manuscript Images Searchable**
Hussein Mohammed, Volker Märgner and Giovanni Ciotti

O1.3 11:40 — 12:00

**A two-step framework for text line segmentation in historical Arabic and Latin document images**
Olfa Mechi, Maroua Mehri, Rolf Ingold and Najoua Essoukri Ben Amara

O1.4 12:00 — 12: 20

**Self-Supervised Deep Metric Learning for ancient papyrus fragments retrieval**
Antoine Pirrone, Marie Beurton-Aimar and Nicholas Journet

O1.5 12:20 — 12:40

**Data Augmentation using Geometric, Frequency, and Beta Modeling approaches for Improving Multi-lingual Online Handwriting Recognition**
Yahia Hamdi, Houcine Boubaker and Adel Alimi

## Oral Session 2: Journal Track 2

O2.1 11:00 — 11:20

**EAPML: Ensemble Self-Attention-based Positive Mutual Learning Network for Document Image Classification**
Souhail Bakkali, Zuheng Ming, Mickael Coustaty and Marcal Rusinol

O2.2 11:20 — 11:40

**Beyond Document Object Detection: Instance-Level Segmentation of Complex Layouts**
Sanket Biswas, Pau Riba, Josep Llados and Umapada Pal

O2.3 11:40 — 12:00

**Asking Questions on Handwritten Document Collections**
Minesh Mathew, Lluis Gomez, Dimosthenis Karatzas and C V Jawahar

O2.4 12:00 — 12: 20

**Revealing a History: Palimpsest Text Separation with Generative Networks**
Anna Starynska, David Messinger and Yu Kong

## Oral Session 3: Historical Document Analysis 1

O3.1 14:00 — 14:20

**BoundaryNet: An Attentive Deep Network with Fast Marching Distance Maps for Semi-automatic Layout Annotation**
Abhishek Trivedi and Ravi Kiran Sarvadevabhatla

O3.2 14:20 — 14-40

**Pho(SC)Net: An Approach Towards Zero-shot Word Image Recognition in Historical Documents**
Anuj Rai, Narayanan C. Krishnan and Sukalpa Chanda

O3.3 14:40 — 15:00

**Versailles-FP dataset: Wall Detection in Ancient Floor Plans**
Wassim Swaileh, Dimitrios Kotzinos, Suman Ghosh, Michel Jordan, Ngoc-Son Vu and Yaguan Qian

O3.4 15:00 — 15:20

**Graph Convolutional Neural Networks for Learning Attribute Representations for Word Spotting**
Fabian Wolf, Andreas Fischer and Gernot A. Fink

O3.5 15:20 — 15:40

**Context Aware Generation of Cuneiform Signs**
Kai Brandenbusch, Eugen Rusakov and Gernot A. Fink

O3.6 15:40 — 16:00

**Adaptive Scaling for Archival Table Structure Recognition**
Xiao-Hui Li, Fei Yin, Xu-Yao Zhang and Cheng-Lin Liu

## Oral Session 4: Document Analysis Systems

O4.1 14:00 — 14:20

**LGPMA: Complicated Table Structure Recognition with Local and Global Pyramid Mask Alignment**
Liang Qiao, Zaisheng Li, Zhanzhan Cheng, Peng Zhang, Shiliang Pu, Yi Niu, Wenqi Ren, Wenming Tan and Fei Wu

O4.2 14:20 — 14-40

**VSR: A Unified Framework for Document Layout Analysis combining Vision, Semantics and Relations**
Peng Zhang, Can Li, Liang Qiao, Zhanzhan Cheng, Shiliang Pu, Yi Niu and Fei Wu

O4.3 14:40 — 15:00

**LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis**
Zejiang Shen, Ruochen Zhang, Melissa Dell, Benjamin Charles Germain Lee, Jacob Carlson and Weining Li

O4.4 15:00 — 15:20

**Understanding and Mitigating the Impact of Model Compression for Document Image Classification**
Shoaib Ahmed Siddiqui, Andreas Dengel and Sheraz Ahmed

O4.5 15:20 — 15:40

**Hierarchical and Multimodal Classification of Images from Soil Remediation Reports**
Korlan Rysbayeva, Romain Giot and Nicholas Journet

O4.6 15:40 — 16:00

**Competition and Collaboration in Document Analysis and Recognition**
Daniel Lopresti and George Nagy

# Thursday · 9 September

## Oral Session 5: Handwriting Recognition

O5.1 14:30 — 14:50

**2D Self-Attention Convolutional Recurrent Network for Offline Handwritten Text Recognition**
Nam Tuan Ly, Hung Tuan Nguyen and Masaki Nakagawa

O5.2 14:50 — 15:10

**Handwritten Text Recognition with Convolutional Prototype Network and Most Aligned Frame Based CTC Training**
Likun Gao, Heng Zhang and Cheng-Lin Liu

O5.3 15:10 — 15:30

**Online Spatio-Temporal 3D Convolutional Neural Network for Early Recognition of Handwritten Gestures**
William Mocaër, Eric Anquetil and Richard Kulpa

O5.4 15:30 — 15:50

**Mix-Up Augmentation for Oracle Character Recognition with Imbalanced Data Distribution**
Jing Li, Qiu-Feng Wang, Rui Zhang and Kaizhu Huang

O5.5 15:50 — 16:10

**Radical Composition Network for Chinese Character Generation**
Mobai Xue, Jun Du, Jianshu Zhang, Zi-Rui Wang, Bin Wang and Bo Ren

O5.6 16:10 — 16:30

**SmartPatch: Improving Handwritten Word Imitation with Patch Discriminators**
Alexander Mattick, Martin Mayr, Mathias Seuret, Andreas Maier and Vincent Christlein

## Oral Session 6: Scene Text Detection & Recognition

O6.1 14:30 — 14:50

**Reciprocal Feature Learning via Explicit and Implicit Tasks in Scene Text Recognition**
Hui Jiang, Yunlu Xu, Zhanzhan Cheng, Shiliang Pu, Yi Niu, Wenqi Ren, Fei Wu and Wenming Tan

O6.2 14:50 — 15:10

**Text Detection by Jointly Learning Character and Word Regions**
Deyang Wu, Xingfei Hu, Zhaozhi Xie, Haiyan Li, Usman Ali and Hongtao Lu

O6.3 15:10 — 15:30

**Vision Transformer for Fast and Efficient Scene Text Recognition**
Rowel Atienza

O6.4 15:30 — 15:50

**Look, Read and Ask: Learning to Ask Questions by Reading Text in Images**
Soumya Jahagirdar, Shankar Gangisetty and Anand Mishra

O6.5 15:50 — 16:10

**CATNet: Scene Text Recognition Guided by Concatenating Augmented Text Features**
Ziyin Zhang, Lemeng Pan, Lin Du, Qingrui Li and Ning Lu

O6.6 16:10 — 16:30

**Explore Hierarchical Relations Reasoning and Global Information Aggregation**
Lei Li, Chun Yuan and Kai Fan

# Friday · 10 September

## Oral Session 7: Historical Document Analysis 2

O7.1 10:30 — 10:50

**One-Model Ensemble-Learning for Text Recognition of Historical Printings**
Christoph Wick and Christian Reul

O7.2 10:50 — 11:10

**On the use of attention in deep learning based denoising method for ancient Cham inscription images**
Tien-Nam Nguyen, Jean-Christophe Burie, Thi-Lan Le and Anne-Valerie Schweyer

O7.3 11:10 — 11:30

**Visual FUDGE: Form Understanding via Dynamic Graph Editing**
Brian Davis, Bryan Morse, Brian Price, Chris Tensmeyer and Curtis Wiginton

O7.4 11:30 — 11:50

**Annotation-Free Character Detection in Historical Vietnamese Stele Images**
Anna Scius-Bertrand, Michael Jungo, Beat Wolf, Andreas Fischer and Marc Bui

## Oral Session 8: Document Image Processing

O8.1 10:30 — 10:50

**DocReader: Bounding-Box Free Training of a Document Information Extraction Model**
Shachar Klaiman and Marius Lehne

O8.2 10:50 — 11:10

**Document Dewarping with Control Points**
Guo-Wang Xie, Fei Yin, Xu-Yao Zhang and Cheng-Lin Liu

O8.3 11:10 — 11:30

**Unknown-box Approximation to Improve Optical Character Recognition Performance**
Ayantha Randika, Nilanjan Ray, Xiao Xiao and Allegra Latimer

O8.4 11:30 — 11:50

**Document Domain Randomization for Deep Learning Document Layout Extraction**
Meng Ling, Jian Chen, Torsten Möller, Petra Isenberg, Tobias Isenberg, Michael Sedlmair, Robert S. Laramee, Han-Wei Shen, Jian Wu and C. Lee Giles

# Friday · 9 September

## Oral Session 9: NLP for Document Understanding

O9.1 14:00 — 14:20

**Distilling the Documents for Relation Extraction by Topic Segmentation**
Minghui Wang, Ping Xue, Ying Li and Zhonghai Wu

O9.2 14:20 — 14:40

**LAMBERT: Layout-Aware Language Modeling for Information Extraction**
Łukasz Garncarek, Rafał Powalski, Tomasz Stanisławek, Bartosz Topolski, Piotr Halama, Michał Turski and Filip Graliński

O9.3 14:40 — 15:00

**ViBERTgrid: A Jointly Trained Multi-Modal 2D Document Representation for Key Information Extraction from Documents**
Weihong Lin, Qifang Gao, Lei Sun, Zhuoyao Zhong, Kai Hu, Qin Ren and Qiang Huo

O9.4 15:00 — 15:20

**Kleister: Key Information Extraction Datasets Involving Long Documents with Complex Layouts**
Tomasz Stanisławek, Filip Graliński, Anna Wróblewska, Dawid Lipiński, Agnieszka Kaliska, Paulina Rosalska, Bartosz Topolski and Przemysław Biecek

## Oral Session 10: Graphics, Diagram, and Math Recognition

O10.1 14:00 — 14:20

**Towards an efficient framework for Data Extraction from Chart Images**
Weihong Ma, Hesuo Zhang, Shuang Yan, Guangshun Yao, Yichao Huang, Hui Li, Yaqiang Wu and Lianwen Jin

O10.2 14:20 — 14:40

**Geometric Object 3D Reconstruction from Single Line Drawing Image Based on a Network for Classification and Sketch Extraction**
Zhuoying Wang, Qingkai Fang and Yongtao Wang

O10.3 14:40 — 15:00

**DiagramNet: Hand-drawn Diagram Recognition using Visual Arrow-relation Detection**
Bernhard Schäfer and Heiner Stuckenschmidt

O10.4 15:00 — 15:20

**Formula Citation Graph Based Mathematical Information Retrieval**
Ke Yuan, Liangcai Gao, Zhuoren Jiang and Zhi Tang

# Posters

Session 1


P1-1

**Towards Document Panoptic Segmentation with Pinpoint Accuracy: Method and Evaluation**
Rongyu Cao, Hongwei Li, Ganbin Zhou and Ping Luo

P1-2

**A Math Formula Extraction and Evaluation Framework for PDF Documents**
Ayush Kumar Shah, Abhisek Dey and Richard Zanibbi

P1-3

**Toward Automatic Interpretation of 3D Plots**
Laura E. Brandt and William T. Freeman

P1-4

**Can Text Summarization Enhance the Headline Stance Detection Task? Benefits and Drawbacks**
Marta Vicente, Robiert Sepúlveda-Torrres, Cristina Barros, Estela Saquete and Elena Lloret

P1-5

**The Biased Coin Flip Process for Nonparametric Topic Modeling**
Justin Wood, Wei Wang and Corey Arnold

P1-6

**CoMSum and SIBERT: A Dataset and Neural Model for Query-Based Multi-Document Summarization**
Sayali Kulkarni, Sheide Chammas, Wan Zhu, Fei Sha and Eugene Ie

P1-7

**RTNet: An End-to-End Method for Handwritten Text Image Translation**
Tonghua Su, Shuchen Liu and Shengjie Zhou

P1-8

**NTable: A Dataset for Camera-based Table Detection**
Ziyi Zhu, Liangcai Gao, Yibo Li, Yilun Huang, Lin Du, Ning Lu and Xianfeng Wang

P1-9

**Label Selection Algorithm Based on Boolean Interpolative Decomposition with Sequential Backward Selection for Multi-label Classification**
Tianqi Ji, Jun Li and Jianhua Xu

P1-10

**GSSF: A Generative Sequence Similarity Function based on a Seq2Seq model for clustering online handwritten mathematical answers**
Huy Quang Ung, Cuong Tuan Nguyen, Hung Tuan Nguyen and Masaki Nakagawa

P1-11

**C2VNet: A Deep Learning Framework Towards Comic Strip to Audio-Visual Scene Synthesis**
Vaibhavi Gupta, Vinay Detani, Vivek Khokar and Chiranjoy Chattopadhyay

P1-12

**LSTMVAEF: Vivid Layout via LSTM-based Variational Autoencoder Framework**
Jie He, Xingjiao Wu, Wenxin Hu and Jing Yang

P1-13

**HCRNN: A Novel Architecture for Fast Online Handwritten Stroke Classification**
Andrii Grygoriev, Illya Degtyarenko, Ivan Deriuga, Serhii Polotskyi, Volodymyr Melnyk, Dmytro Zakharchuk and Olga Radyvonenko

P1-14

**RFDoc: memory efficient local descriptors for ID documents localization and classification**
Daniil Matalov, Elena Limonova, Natalya Skoryukina and Vladimir V. Arlazarov

P1-15

**Dynamic Receptive Field Adaptation for Attention-Based Text Recognition**
Haibo Qin, Chun Yang, Xiaobin Zhu and Xucheng Yin

P1-16

**Context-Free TextSpotter for Real-Time and Mobile End-to-End Text Detection and Recognition**
Ryota Yoshihashi, Tomohiro Tanaka, Kenji Doi, Takumi Fujino and Naoaki Yamashita

P1-17

**MIDV-LAIT: a challenging dataset for recognition of IDs with Perso-Arabic, Thai, and Indian scripts**
Yulia Chernyshova, Ekaterina Emelianova, Alexander Sheshkus and Vladimir V. Arlazarov

P1-18

**Determining optimal frame processing strategies for real-time document recognition systems**
Konstantin Bulatov and Vladimir V. Arlazarov

P1-19

**Embedded Attributes for Cuneiform Sign Spotting**
Eugen Rusakov, Turna Somel, Gerfrid G.W. Müller and Gernot A. Fink

P1-20 324

**Date Estimation in the Wild of Scanned Historical Photos: An Image Retrieval Approach**
Adrià Molina, Pau Riba, Lluis Gomez, Oriol Ramos-Terrades and Josep Lladós

P1-21 259

**Two-Step Fine-Tuned Convolutional Neural Networks for Multi-Label Classification of Children's Drawings**
Muhammad Osama Zeeshan, Imran Siddiqi and Momina Moetesum

P1-22 306

**DCINN: Deformable Convolution and Inception Based Neural Network for Tattoo Text Detection through Skin Region**
Tamal Chowdhury, Palaiahnakote Shivakumara, Umapada Pal, Tong Lu, Ramachandra Raghavendra and Sukalpa Chanda

P1-23 56

**Sparse Document Analysis using Beta-Liouville Naive Bayes with Vocabulary Knowledge**
Fatma Najar and Nizar Bouguila

P1-24 158

**Automatic Signature-based Writer Identification in Mixed-script Scenarios**
Sk Md Obaidullah, Mridul Ghosh, Himadri Mukherjee, Kaushik Roy and Umapada Pal

P1-25 327

**Learning to Rank Words: Optimizing Ranking Metrics for Word Spotting**
Pau Riba, Adrià Molina, Lluis Gomez, Oriol Ramos-Terrades and Josep Lladós

P1-26 146

**A-VLAD: An End-to-End Attention-based Neural Network for Writer Identification in Historical Documents**
Trung Tan Ngo, Hung Tuan Nguyen and Masaki Nakagawa

P1-27 172

**Manga-MMTL: multimodal multitask transfer learning for manga character analysis**
Nhu-Van Nguyen, Christophe Rigaud, Arnaud Revel and Jean-Christophe Burie

P1-28 131

**Probabilistic Indexing and Search for Hyphenated Words**
Enrique Vidal and Alejandro H. Toselli

P1-29 175

**SandSlide: Automatic Slideshow Normalization**
Sieben Bocklandt, Gust Verbruggen and Thomas Winters

P1-30 366

**Digital Editions as Distant Supervision for Layout Analysis of Printed Books**
Alejandro H. Toselli, Si Wu and David A. Smith

P1-31 129

**Palmira: A Deep Deformable Network for Instance Segmentation of Dense and Uneven Layouts in Handwritten Manuscripts**
S P Sharan, Sowmya Aitha, Amandeep Kumar, Abhishek Trivedi, Aaron Augustine and Ravi Kiran Sarvadevabhatla

P1-32 149

**Page Layout Analysis System for Unconstrained Historic Documents**
Oldřich Kodym and Michal Hradiš

P1-33 235

**Improved Graph Methods for Table Layout Understanding**
Jose Ramón Prieto and Enrique Vidal

P1-34 251

**Unsupervised learning of text line segmentation by differentiating coarse patterns**
Berat Kurar Barakat, Ahmad Droby, Raid Saabni and Jihad El-Sana

P1-35 108

**Rethinking Table Structure Recognition Using Sequence Labeling Methods**
Yibo Li, Yilun Huang, Ziyi Zhu, Lemeng Pan, Yongshuai Huang, Lin Du, Zhi Tang and Liangcai Gao

P1-36 119

**TabLeX: A Benchmark Dataset for Structure and Content Information Extraction from Scientific Tables**
Harsh Desai, Pratik Kayal and Mayank Singh

P1-37 50

**Handwritten Mathematical Expression Recognition with Bidirectionally Trained Transformer**
Wenqi Zhao, Liangcai Gao, Zuoyu Yan, Shuai Peng, Lin Du and Ziyin Zhang

P1-38 338

**TabAug: Data Driven Augmentation for Enhanced Table Structure Recognition**
Umar Khan, Sohaib Zahid, Muhammad Asad Ali, Adnan Ul-Hasan and Faisal Shafait

P1-39 257

**An Encoder-Decoder Approach to Handwritten Mathematical Expression Recognition with Multi-Head Attention and Stacked Decoder**
Haisong Ding, Kai Chen and Qiang Huo

P1-40 271

**Global Context for improving recognition of Online Handwritten Mathematical Expressions**
Cuong Tuan Nguyen, Thanh-Nghia Truong, Hung Tuan Nguyen and Masaki Nakagawa

P1-41 96

**Image-based Relation Classification Approach for Table Structure Recognition**
Koji Ichikawa

P1-42 115

**Image to LaTeX with Graph Neural Network for Mathematical Formula Recognition**
Shuai Peng, Liangcai Gao, Ke Yuan and Zhi Tang

P1-43 316

**A Novel Method for Automated Suggestion of Similar Software Incidents using 2-Stage Filtering : Findings on Primary Data**
Badal Agrawal, Mohit Mishra and Varun Parashar

P1-44 37

**Research on pseudo-label technology for multi-label news classification**
Lianxi Wang, Xiaotian Lin and Nankai Lin

P1-45 254

**Information Extraction from Invoices**
Ahmed Hamdi, Elodie Carel, Aurélie Joseph, Mickael Coustaty and Antoine Doucet

P1-46 256

**Are You Really Complaining? A Multi-task Framework for Complaint Identification, Emotion and Sentiment Classification**
Apoorva Singh and Sriparna Saha

P1-47 318

**Going Full-TILT Boogie on Document Understanding with Text-Image-Layout Transformer**
Rafał Powalski, Łukasz Borchmann, Dawid Jurkiewicz, Tomasz Dwojak, Michał Pietruszka and Gabriela Pałka

P1-48 248

**Data Centric Domain Adaptation for Historical Text with OCR Errors**
Luisa März, Stefan Schweter, Nina Poerner, Benjamin Roth and Hinrich Schütze

P1-49 236

**Temporal Ordering of Events via Deep Neural Networks**
Nafaa Haffar, Rami Ayadi, Emna Hkiri and Mounir Zrigui

P1-50 272

**Document Collection Visual Question Answering**
Rubèn Tito, Dimosthenis Karatzas and Ernest Valveny

P1-51 83

**Dialogue Act Recognition using Visual Information**
Jiří Martínek, Pavel Král and Ladislav Lenc

P1-52 95

**Are End-to-End Systems Really Necessary for NER on Handwritten Document Images?**
Oliver Tüselmann, Fabian Wolf and Gernot A. Fink

P1-53 160

**Training Bi-Encoders for Word Sense Disambiguation**
Harsh Kohli

P1-54 234

**DeepCPCFG: Deep Learning and Context Free Grammars for End-to-End Information Extraction**
Freddy C. Chua and Nigel P. Duffy

P1-55 157

**Consideration of the word's neighborhood in GATs for information extraction in semi-structured documents**
Djedjiga Belhadj, Yolande Belaïd and Abdel Belaïd

P1-56 369

**MiikeMineStamps: A Long-Tailed Dataset of Japanese Stamps via Active Learning**
Paola A., Buitrago, Evgeny Toropov, Rajanie Prabha, Julian Uran and Raja Adal

P1-57 334

**Deep Learning for Document Layout Generation: A First Reproducible Quantitative Evaluation and a Baseline Model**
Romain Carletto, Hubert Cardot and Nicolas Ragot

P1-58 342

**MRD: A Memory Relation Decoder for Online Handwritten Mathematical Expression Recognition**
Jiaming Wang, Qing Wang, Jun Du, Jianshu Zhang, Bin Wang and Bo Ren

P1-59 22

**Full Page Handwriting Recognition via Image to Sequence Extraction**
Sumeet S. Singh and Sergey Karayev

P1-60 76

**SPAN: a Simple Predict & Align Network for Handwritten Paragraph Recognition**
Denis Coquenet, Clément Chatelain and Thierry Paquet

P1-61 163

**IHR-NomDB: The Old Degraded Vietnamese Handwritten Script Archive Database**
Manh Tu VU, Van Linh LE, and Marie BEURTON-AIMAR

P1-62 48

**Sequence Learning Model for Syllables Recognition Arranged in Two Dimensions**
Valerii Dziubliuk, Mykhailo Zlotnyk and Oleksandr Viatchaninov

P1-63 59

**Transformer for Handwritten Text Recognition using Bidirectional Post-Decoding**
Christoph Wick, Jochen Zöllner and Tobias Grüning

P1-64 60

**Zero-Shot Chinese Text Recognition via Matching Class Embedding**
Yuhao Huang, Lianwen Jin and Dezhi Peng

P1-65 67

**Text-conditioned Character Segmentation for CTC-based Text Recognition**
Ryohei Tanaka, Kunio Osada and Akio Furuhata

P1-66 141

**Towards Fast, Accurate and Compact Online Handwritten Chinese Text Recognition**
Dezhi Peng, Canyu Xie, Hongliang Li, Lianwen Jin, Zecheng Xie, Kai Ding, Yichao Huang and Yaqiang Wu

P1-67 246

**HCADecoder: A Hybrid CTC-Attention Decoder for Chinese Text Recognition**
Siqi Cai, Wenyuan Xue, Qingyong Li and Peng Zhao

P1-68 350

**Meta-learning of Pooling Layers for Character Recognition**
Takato Otsuzuki, Heon Song, Seiichi Uchida and Hideaki Hayashi

P1-69 29

**Text-line-up: Don't Worry about the Caret**
Chandranath Adak, Bidyut B. Chaudhuri, Chin-Teng Lin and Michael Blumenstein

P1-70 219

**Multimodal Attention-based Learning for Imbalanced Corporate Documents Classification**
Ibrahim Souleiman Mahamoud, Joris Voerman, Mickaël Coustaty, Aurélie Joseph, Vincent Poulain d'Andecy and Jean-Marc Ogier

P1-71 273

**Light-weight Document Image Cleanup using Perceptual Loss**
Soumyadeep Dey and Pratik Jawanpuria

P2-1 180

**A New Semi-Automatic Annotation Model via Semantic Boundary Estimation for Scene Text Detection**
Zhenzhou Zhuang, Zonghao Liu, Kin-Man Lam, Shuangping Huang and Gang Dai

P2-2 63

**Searching from the Prediction of Visual and Language Model for Handwritten Chinese Text Recognition**
Brian Liu, Weicong Sun, Wenjing Kang and Xianchao Xu

P2-3 127

**Towards an IMU-based Pen Online Handwriting Recognizer**
Mohamad Wehbi, Tim Hamann, Jens Barth, P eter Kaempf, Dario Zanca and Bjoern Eskofier

P2-4 132

**2D vs 3D online writer identification: a comparative study**
Antonio Parziale, Cristina Carmona-Duarte, Miguel Angel Ferrer and Angelo Marcelli

P2-5 139

**A Handwritten Signature Segmentation Approach for Multi-resolution and Complex Documents Acquired by Multiple Sources**
Celso A. M. Lopes Junior, Murilo C. Stodolni, Byron L. D. Bezerra and Donato Impedovo

P2-6 46

**Attention based Multiple Siamese Network for Offline Signature Verification**
Yu-Jie Xiong and Song-Yang Cheng

P2-7 179

**Attention to Warp: Deep Metric Learning for Multivariate Time Series**
Shinnosuke Matsuo, Xiaomeng Wu, Gantugs Atarsaikhan, Akisato Kimura, Kunio Kashino, Brian Kenji Iwana and Seiichi Uchida

P2-8 347

**Customizable Camera Verification for Media Forensic**
Huaigu Cao and Wael AbdAlmageed

P2-9 24

**Density Parameters of Handwriting in Schizophrenia and Affective Disorders Assessed Using the Raygraf Computer Software**
Barbara Gawda

P2-10 138

**Language-Independent Bimodal System for Early Parkinson's Disease Detection**
Catherine Taleb, Laurence Likforman-Sulem and C hafic Mokbel

P2-11 340

**TRACE: A Differentiable Approach to Line-level Stroke Recovery for Offline Handwritten Text**
Taylor Archibald, Mason Poggemann, Aaron Chan and Tony Martinez

P2-12 357

**Segmentation and graph matching for online analysis of student arithmetic operations**
Arnaud Lods, Éric Anquetil and Sébastien Macé

P2-13 289

**Applying End-to-end Trainable Approach on Stroke Extraction in Handwritten Math Expressions Images**
Elmokhtar Mohamed Moussa, Thibault Lelore and Harold Mouchère

P2-14 55

**A Novel Sigma-Lognormal Parameter Extractor for Online Signatures**
Jianhuan Huang and Zili Zhang

P2-15 8

**Near-perfect Relation Extraction from Family Books**
George Nagy

P2-16 154

**Estimating Human Legibility in Historic Manuscript Images — A Baseline**
Simon Brenner, Lukas Schügerl and Robert Sablatnig

P2-17 112

**A Modular and Automated Annotation Platform for Handwritings: Evaluation on Under-resourced Languages**
Chahan Vidal-Gorène, Boris Dupin, Aliénor Decours-Perez and Thomas Riccioli

P2-18 122

**Reducing the Human Effort in Text Line Segmentation for Historical Documents**
Emilio Granell, Lorenzo Quirós, Verónica Romero and Joan Andreu Sánchez

P2-19 186

**DSCNN: Dimension Separable Convolutional Neural Networks for character recognition based on inertial sensor signal**
Fan Peng, Zhendong Zhuang and Yang Xue

P2-20 370

**DocSynth: A Layout Guided Approach for Controllable Document Image Synthesis**
Sanket Biswas, Pau Riba, Josep Lladós and Umapada Pal

P2-21 140

**Font Style that Fits an Image — Font Generation Based on Image Context**
Taiga Miyazono, Brian Kenji Iwana, Daichi Haraguchi and Seiichi Uchida

P2-22 45

**Bayesian Hyperparameter optimization of Deep Neural Network algorithms based on Ant Colony optimization**
Sinda Jlassi, Imen Jdey and Hela Ltifi

P2-23 49

**End-to-End Approach for Recognition of Historical Digit Strings**
Mengqiao Zhao, Andre Gustavo Hochuli and Abbas Cheddad

P2-24 177

**Generating Synthetic Handwritten Historical Documents With OCR Constrained GANs**
Lars Vögtlin, Manuel Drazyk, Vinaychandran Pondenkandath, Michele Alberti and Rolf Ingold

P2-25 91

**Synthesizing Training Data for Handwritten Music Recognition**
Jiří Mayer and Pavel Pecina

P2-26 274

**Towards Book Cover Design via Layout Graphs**
Wensheng Zhang, Yan Zheng, Taiga Miyazono, Seiichi Uchida and Brian Kenji Iwana

P2-27 199

**Complete Optical Music Recognition via Agnostic Transcription and Machine Translation**
Antonio Ríos-Vila, David Rizo and Jorge Calvo-Zaragoza

P2-28 61

**Improving Machine Understanding of Human Intent in Charts**
Sihang Wu, Canyu Xie, Yuhao Huang, Guozhi Tang, Qianying Liao, Jiapeng Wang, Bangdong Chen, Hongliang Li, Xinfeng Chang, Hui Li, Kai Ding, Yichao Huang and Lianwen Jin

P2-29 105

**DeMatch: Towards Understanding the Panel of Chart Documents**
Hesuo Zhang, Weihong Ma, Lianwen Jin, Yichao Huang, Kai Ding and Yaqiang Wu

P2-30 86

**Sequential Next-Symbol Prediction for Optical Music Recognition**
Enrique Mas-Candela, Maria Alfaro-Contreras and Jorge Calvo-Zaragoza

P2-31 241

**Which Parts Determine the Impression of the Font?**
Masaya Ueda, Akisato Kimura and Seiichi Uchida

P2-32 287

**Impressions2Font: Generating Fonts by Specifying Impressions**
Seiya Matsuda, Akisato Kimura and Seiichi Uchida

P2-33 361

**HRRegionNet: Chinese Character Segmentation in Historical Documents with Regional Awareness**
Chia-Wei Tang, Chao-Lin Liu and Po-Sen Chiu

P2-34 192

**Fast Text v. Non-text Classification of Images**
Jiri Kralicek and Jiri Matas

P2-35 43

**Mask Scene Text Recognizer**
Haodong Shi, Liangrui Peng, Ruijie Yan, Gang Yao, Shuman Han and Shengjin Wang

P2-36 53

**Rotated Box Is Back: An Accurate Box Proposal Network for Scene Text Detection**
Jusung Lee, Jaemyung Lee, Cheoljong Yang, Younghyun Lee and Joonsoo Lee

P2-37 88

**Heterogeneous Network Based Semi-supervised Learning For Scene Text Recognition**
Qianyi Jiang, Qi Song, Nan Li, Rui Zhang and Xiaolin Wei

P2-38 90

**Scene Text Detection with Scribble Line**
Wenqing Zhang, Yang Qiu, Minghui Liao, Rui Zhang, Xiaolin Wei and Xiang Bai

P2-39 104

**EEM: An End-to-end Evaluation Metric for Scene Text Detection and Recognition**
Jiedong Hao, Yafei Wen, Jie Deng, Jun Gan, Shuai Ren, Hui Tan, and Xiaoxin Chen

P2-40 188

**SynthTIGER: Synthetic Text Image GEneratoR Towards Better Text Recognition Models**
Moonbin Yim, Yoonsik Kim, Han-Cheol Cho and Sungrae Park

P2-41 218

**Fast Recognition for Multidirectional and Multi-Type License Plates with 2D Spatial Attention**
Qi Liu, Song-Lu Chen, Zhen-Jia Li, Chun Yang, Feng Chen and Xu-Cheng Yin

P2-42 244

**A Multi-level Progressive Rectification Mechanism for Irregular Scene Text Recognition**
Qianying Liao, Qingxiang Lin, Lianwen Jin, Canjie Luo, Jiaxin Zhang, Dezhi Peng and Tianwei Wang

P2-43 346

**Representation and Correlation Enhanced Encoder-Decoder Framework for Scene Text Recognition**
Mengmeng Cui, Wei Wang, Jinjin Zhang and Liang Wang

P2-44 34

**FEDS - Filtered Edit Distance Surrogate**
Yash Patel and Jiří Matas

P2-45 360

**Bidirectional Regression for Arbitrary-Shaped Text Detection**
Tao Sheng and Zhouhui Lian

P2-46 253

**VML-HP: Hebrew paleography dataset**
Ahmad Droby, Berat Kurar Barakat, Daria Vasyutinsky Shapira, Irina Rabaev and Jihad El-Sana

P2-47 30

**Open Set Authorship Attribution toward Demystifying Victorian Periodicals**
Sarkhan Badirli, Mary Borgo Ton, Abdulmecit Gungor and Murat Dundar

P2-48 28

**A More Effective Sentence-Wise Text Segmentation Approach using BERT**
Amit Maraj, Miguel Vargas Martin and Masoud Makrehchi

P2-49 73

**Data Augmentation for Writer Identification Using a Cognitive Inspired Model**
Fabio Pignelli, Yandre M. G. Costa, Luiz S. Oliveira and Diego Bertolini

P2-50 208

**Key-guided Identity Document Classification Method by Graph Attention Network**
Xiaojie Xia, Wei Liu, Ying Zhang, Liuan Wang and Jun Sun

P2-51 296

**Document Image Quality Assessment via Explicit Blur and Text Size Estimation**
Dmitry Rodin, Vasily Loginov, Ivan Zagaynov and Nikita Orlov

P2-52 297

**Analyzing the potential of Zero-Shot Recognition for Document Image Classification**
Shoaib Ahmed Siddiqui, Andreas Dengel and Sheraz Ahmed

P2-53 124

**Gender Detection Based on Spatial Pyramid Matching**
Fahimeh Alaei and Alireza Alaei

P2-54 70

**EDNets: Deep Feature Learning for Document Image Classification based on Multi-view Encoder-Decoder Neural Networks**
Akrem Sellami and Salvatore Tabbone

P2-55 283

**Fast End-to-end Deep Learning Identity Document Detection, Classification and Cropping**
Guillaume Chiron, Florian Arrestier and Ahmad Montaser Awal

P2-56 232

**Image Collation: Matching illustrations in manuscripts**
Ryad Kaoua, Xi Shen, Alexandra Durr, Stavros Lazaris, David Picard and Mathieu Aubry

P2-57 319

**Revisiting the Coco Panoptic Metric to Enable Visual and Qualitative Analysis of Historical Map Instance Segmentation**
Joseph Chazalon and Edwin Carlinet

P2-58 247

**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

P2-59 78

**GNHK: A Dataset for English Handwriting in the Wild**
Alex W. C. Lee, Jonathan Chung and Marco Lee

P2-60 202

**Personalizing Handwriting Recognition Systems with Limited User-Specific Samples**
Christian Gold, Dario van den Boom and Torsten Zesch

P2-61 307

**An Efficient Local Word Augment Approach for Mongolian Handwritten Script Recognition**
Haoran Zhang, Wei Chen, Xiangdong Su, Hui Guo and Huali Xu

P2-62 331

**IIIT-INDIC-HW-WORDS: A Dataset for Indic Handwritten Text Recognition**
Santhoshini Gongidi and C V Jawahar

P2-63 221

**AT-ST: Self-Training Adaptation Strategy for OCR in Domains with Limited Transcriptions**
Martin Kišš, Karel Beneš and Michal Hradiš

P2-64 320

**TS-Net: OCR Trained to Switch Between Text Transcription Styles**
Jan Kohút and Michal Hradiš

P2-65 71

**Handwriting Recognition with Novelty**
Derek S. Prijatelj, Samuel Grieggs, Futoshi Yumoto, Eric Robertson and Walter J. Scheirer

P2-66 173

**Vectorization of Historical Maps Using Deep Edge Filtering and Closed Shape Extraction**
Yizi Chen, Edwin Carlinet, Joseph Chazalon, Clément Mallet, Bertrand Duménieu and Julien Perret

P2-67 148

**Data Augmentation Based on CycleGAN for Improving Woodblock-printing Mongolian Words Recognition**
Hongxi Wei, Kexin Liu, Jing Zhang and Daoerji Fan

P2-68 198

**SauvolaNet: Learning Adaptive Sauvola Network for Degraded Document Binarization**
Deng Li, Yue Wu and Yicong Zhou

P2-69 126

**Recognizing Handwritten Chinese Texts with Insertion and Swapping Using A Structural Attention Network**
Shi Yan, Jin-Wen Wu, Fei Yin and Cheng-Lin Liu

P2-70 270

**Strikethrough Removal From Handwritten Words Using CycleGANs**
Raphaela Heil, Ekta Vats and Anders Hast

P2-71 305

**Iterative Weighted Transductive Learning for Handwriting Recognition**
George Retsinas, Giorgos Sfikas and Christophoros Nikou

# Competitions

C1

**ICDAR 2021 Competition on Scientific Literature Parsing**
Antonio Jimeno Yepes, Peter Zhong, Douglas Burdick

C2

**ICDAR 2021 Competition on Historical Document Classification**
Mathias Seuret, Anguelos Nicolaou, Dalia Rodríguez-Salas, Nikolaus Weichselbaumer, Dominique Stutzmann, Martin Mayr, Andreas Maier, and Vincent Christlein

C3

**ICDAR 2021 Competition on Document Visual Question Answering**
Rubèn Tito, Minesh Mathew, C.V. Jawahar, Ernest Valveny, and Dimosthenis Karatzas

C4

**ICDAR 2021 Competition on Scene Video Text Spotting**
Zhanzhan Cheng, Jing Lu, Baorui Zou, Shuigeng Zhou, and Fei Wu

C5

**ICDAR 2021 Competition on Integrated Circuit Text Spotting and Aesthetic Assessment**
Chun Chet Ng, Akmalul Khairi Bin Nazaruddin, Yeong Khang Lee, Xinyu Wang, Yuliang Liu, Chee Seng Chan, Lianwen Jin, Yipeng Sun, and Lixin Fan

C6

**ICDAR 2021 Competition on Components Segmentation Task of Document Photos**
Celso A. M. Lopes Junior, Ricardo B. Neves Junior, Byron L. D. Bezerra, Alejandro H. Toselli, and Donato Impedovo

C7

**ICDAR 2021 Competition on Historical Map Segmentation**
Joseph Chazalon, Edwin Carlinet, Yizi Chen, Julien Perret, Bertrand Duménieu, Clément Mallet, Thierry Géraud,Vincent Nguyen, Nam Nguyen, Josef Baloun, Ladislav Lenc, and Pavel Král

C8

**ICDAR 2021 Competition on Time-Quality Document Image Binarization**
Rafael Dueire Lins, Rodrigo Barros Bernardino, Elisa Barney Smith, and Ergina Kavallieratou

C9

**ICDAR 2021 Competition on On-Line Signature Verification**
Ruben Tolosana, Ruben Vera-Rodriguez, Carlos Gonzalez-Garcia, Julian Fierrez, Santiago Rengifo, Aythami Morales, Javier Ortega-Garcia, Juan Carlos Ruiz-Garcia, Sergio Romero-Tapiador, Jiajia Jiang, Songxuan Lai, Lianwen Jin, Yecheng Zhu, Javier Galbally, Moises Diaz, Miguel Angel Ferrer, Marta Gomez-Barrero, Ilya Hodashinsky, Konstantin Sarin, Artem Slezkin, Marina Bardamova, Mikhail Svetlakov, Mohammad Saleem, Cintia Lia Szücs, Bence Kovari, Falk Pulsmeyer, Mohamad Wehbi, Dario Zanca, Sumaiya Ahmad, Sarthak Mishra, and Suraiya Jabin

C10

**ICDAR 2021 Competition on Script Identification in the Wild**
Abhijit Das, Miguel A. Ferrer, Aythami Morales, Moises Diaz, Umapada Pal, Donato Impedovo, Hongliang Li, Wentao Yang, Kensho Ota, Tadahito Yao, Le Quang Hung, Nguyen Quoc Cuong, Seungjae Kim, and Abdeljalil Gattal

C11

**ICDAR 2021 Competition on Scientific Table Image Recognition to LaTeX**
Pratik Kayal, Mrinal Anand, Harsh Desai, and Mayank Singh

C12

**ICDAR 2021 Competition on Multimodal Emotion Recognition on Comics Scenes**
Nhu-Van Nguyen, Xuan-Son Vu, Christophe Rigaud, Lili Jiang, and Jean-Christophe Burie

C13

**ICDAR 2021 Competition on Mathematical Formula Detection**
Dan Anitei, Joan Andreu Sánchez, José Manuel Fuentes, Roberto Paredes, and José Miguel Benedí

**Notes**

**Notes**

# Routes

**Routes to BC building @ EPFL**

**From Lausanne Flon**:
— Metro M1 (for Renens Gare), stop at Ecublens VD, EPFL (13 min)
— Walk to BC Building (7 min)

**From Lausanne Gare**:
— Metro M2 (for Lausanne Sallaz, Epalinges, Croisette) stop at Lausanne Flon (3 min)
— Metro M1 (for Renens Gare), stop at Ecublens VD, EPFL (13 min)
— Walk to BC Building (7 min)

or alternatively:
— Train SBB/CFF to Renens Gare (7 min)
— Metro M1 (for Lausanne Flon), stop at Ecublens VD, EPFL (6 min)
— Walk to BC Building (7 min)

**Routes to Beaulieu Convention Center**

**From Place Riponne**:
— Bus #2 (for Lausanne Désert), leaving from Rue Neuve, stop at Beaulieu (5 min)
— Walk to convention center (2 min)

**From Lausanne-Gare** (train station):
— Bus #21 (for Lausanne Blécherette), stop at Beaulieu (after Beaulieu-Jomini, 12 min)
— Walk to convention center (2 min)

or alternatively:
— Bus #3 (for Lausanne Bellevaux), stop at Beaulieu-Jomini (11 min)
— Walk to convention center (5 min)

Walk from the city (Place Riponne): 15–20 minutes

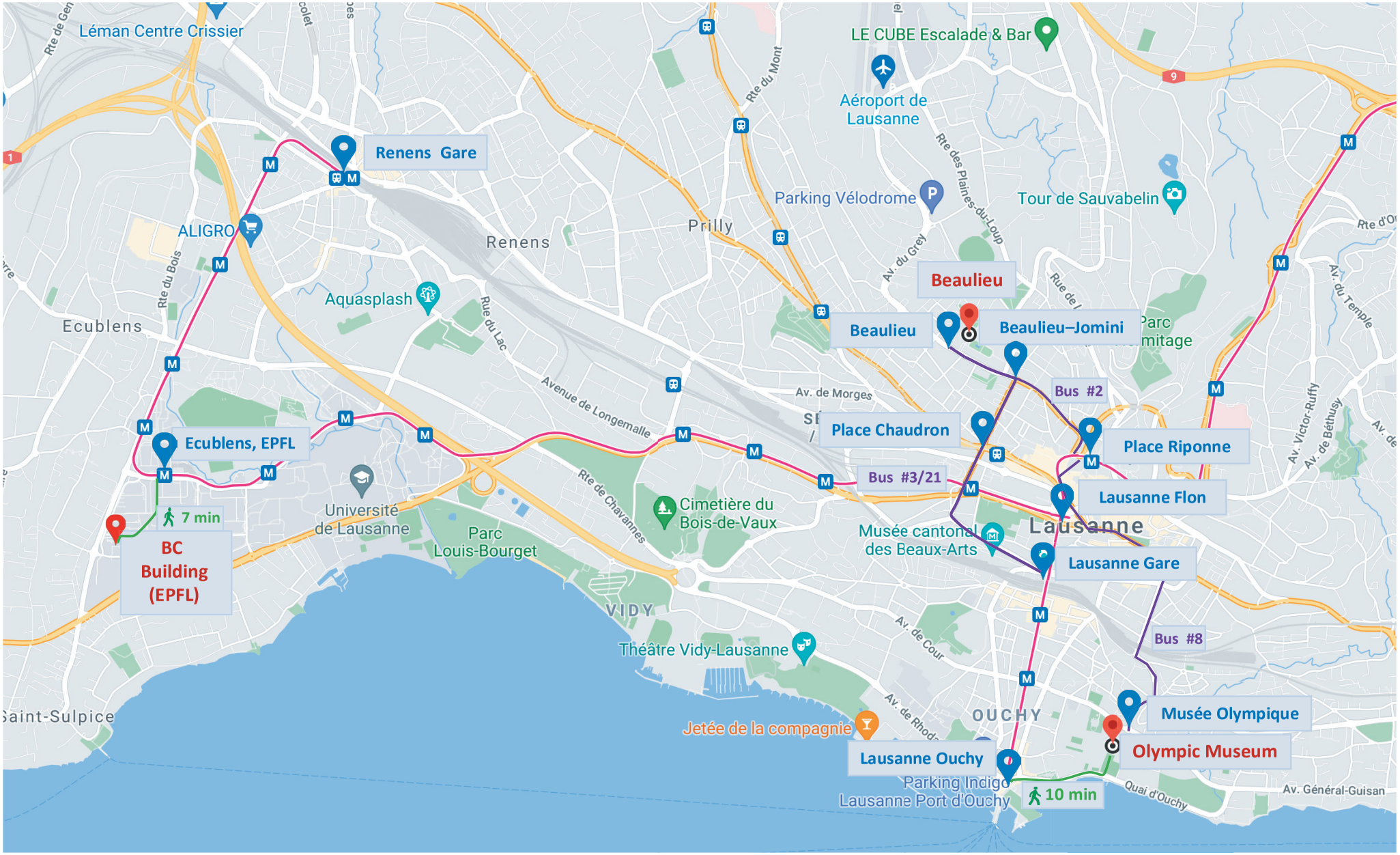
**Routes To Olympic Museum**

Recommended **from Beaulieu** via Place Riponne (reachable from Beaulieu by bus #2):
— Bus #2 (for Lausanne, Maladière-Lac), stop at Rue Neuve (5 min)
— Metro M2, at Place Riponne, to terminus Ouchy Olympique (8 min)
— Walk to Olympic Museum (10 min)

or via Lausanne Gare:
— Bus #21 to terminus Lausanne Gare (11 min)
— Metro M2 to terminus Ouchy Olympique (6 min)
— Walk to Olympic Museum (10 min)

Alternatively, for people with reduced mobility, **from Place Riponne**:
— Bus #8 (for Paudex, Verrière), stop at Musée Olympique (12 minutes)
— Walk to Olympic Museum (2 min)

Colophon — ICDAR2021 logo and booklet design by Vlad Atanasiu · Logo typeface CPMono by Tino Meinert · Booklet typeface Fira Sans by Carrois Corporate & Edenspiekermann AG / The Mozilla Corporation


Léman Centre Crissier
LE CUBE Escalade & Bar
Aéroport de Lausanne
Renens Gare
Parking Vélodrome
Tour de Sauvabelin
ALIGRO
Renens
Prilly
Aquasplash
Beaulieu
Beaulieu
Beaulieu–Jomini
Ecublens
Bus #2
Av. de Morges
Place Chaudron
Place Riponne
Ecublens, EPFL
Avenue de Longemalle
Bus #3/21
Lausanne Flon
7 min
Université de Lausanne
Cimetière du Bois-de-Vaux
Parc Louis-Bourget
Musée cantonal des Beaux-Arts
Lausanne
BC Building (EPFL)
Lausanne Gare
VIDY
Théâtre Vidy-Lausanne
Bus #8
Musée Olympique
Jetée de la compagnie
OUCHY
Olympic Museum
Lausanne Ouchy
Parking Indigo Lausanne Port d'Ouchy
10 min
Quai d'Ouchy
Av. Général-Guisan

## Platinum Sponsors

wacom®

GoodNotes

## Gold Sponsor

Google

## Bronze Sponsors

iTESOFT

yooz
Cloud P2P Automation. Easy. Powerful. Smart.

HITACHI
Inspire the Next

## Academic Sponsors

IAPR

TC10
GRAPHICS RECOGNITION

TC-11
Reading Systems

UNI FR
UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

AIVA

HEIA-FR
HTA-FR

iCoSys
Institute of Complex Systems

Hes·so
Haute Ecole Spécialisée de Suisse occidentale
Fachhochschule Westschweiz
University of Applied Sciences and Arts Western Switzerland

LULEÅ TEKNISKA UNIVERSITET

## Proceedings Publisher

Springer