

Workshop on Scaling-up Document Image Understanding

24th August, 2023 at ICDAR 2023, San Jose, California, USA

Organizers

Anand Mishra, IIT Jodhpur

Andreas Fischer, HES-SO Switzerland

Dimosthenis Karatzas, CVC, Universitat Autònoma de Barcelona

Seiichi Uchida, Kyushu University

C V Jawahar, IIIT Hyderabad

The workshop on [Scaling Up Image Document Understanding](#) was held at the Adobe World Headquarters in San Jose, California on August 24th, 2023.

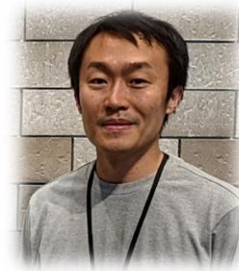
The workshop was aimed at opening the discussion on possible ways to widen our community using data preparation efforts and defining the large-scale (grand) challenges that drive progress in the field. This is meant to be one of a series of such events to be organized on our scientific forums in the near future. The [Deep Document Workshop](#) at ICFHR 2022 was our first successful event in this series. The aspiration was to set the seed for an initiative to create our own community's document-oriented "ImageNet", over which multiple long-term challenges can be defined.

While most of the participants joined the workshop physically at San Jose, a couple of participants, who could not travel due to several reasons, attended online. A total of 60 participants, including researchers from academia and industry, from various countries attended the event.

The activities of the workshop included two very informative keynote sessions, an innovative pitch session and an exciting panel discussion. The invited talks were 45 minutes long.

Keynotes:

Keynote 1: "What are grand challenges in DAR? Do we already know?"



Yasuhisa Fujii (Google Research, Tokyo)

Keynote 2: "Beyond OCR: Non-Textual opportunities and challenges at the Hathitrust Research Center"



J. Stephen Downie (University of Illinois, Urbana-Champaign)

Pitch session:



The session was moderated by Anand Mishra of IIT Jodhpur. The participants were asked to give a 2 min presentation on datasets, challenges and competitions in the document research community. The topics were shared with the participants prior to the workshop. The pitch participants were Anna Scius-Bertrand (HES-SO), Christopher Kermorvant (Founder-TEKLIA), Lukasz Borchmann (Snowflake), Shangbang Long (Google), Ravi Kiran S (IIIT Hyderabad), Shubhi Asthana (IBM), Vincent Christlein (University Erlangen-Nuremberg) and Stepan Simsa (Rossum.ai).



Panel Discussion:

The interactive panel discussion led by Andreas Fischer consisted of Mickael Coustaty (La Rochelle Universite), Joseph Chazalon (LRDE, EPITA), Anand Mishra (IIT Jodhpur), Stephen Downie (University of Illinois), Yasuhisa Fujii (Google Research) as panel members.



The main topic of discussion was on the existing or future challenges and datasets of DAR that will help us to scale up DAR research.

The three representatives of TC10/11 in the panel were trying to understand how they can help in dataset preparations. The panel discussed on why document analysis and recognitions exist, how it evolved over the years from simple text recognition to more complex recognitions, more content, as well as topics like HCI, Interactive documents, better tools for the end users, more visibility for datasets, Joining multiple isolated datasets into Single one, interaction with digital library communities etc.

The young people who attended the ScalDoc workshop were so motivated that a Slack channel was created soon after to gather academic and industrial researchers willing to collaborate on the creation of a “dataset of datasets” and is mentioned in the August TC11 newsletter.

We are deeply thankful to the keynote speakers, panel members, pitch speakers, organizers, all the workshop participants and all the members of the ICDAR program committee whose collaboration has been fundamental for the success of this workshop.