ICDAR Athens Greece 2024

International Conference on Document Analysis and Recognition

The 18th International Conference on Document Analysis & Recognition

Programme

30 August - 4 September 2024

ICDAR Athens Greece 2024

International Conference on Document Analysis and Recognition

The 18th International Conference on Document Analysis & Recognition

ORGANISERS

DEMOKRITOS

ATHENA Research & Innovation Information Technologies

LULEÅ UNIVERSITY OF TECHNOLOGY

IAPR

Contents

Welcome Note from Conference Chairs

The Organising Committee of the International Conference on Document Analysis and Recognition with pleasure welcomes you to Athens, Greece for the 18th edition of ICDAR 2024 which will take place from 30 August – 4 September 2024. The format of the conference will build upon best practices of previous ICDAR conferences and will feature keynote talks, main track sessions, presentations, panel discussions, poster sessions, side workshops and social events. The series of ICDAR conferences attracts more than 500 participants from the world over and is the perfect meeting place for the community to come together.

We are honoured to welcome you to the proceedings of ICDAR 2024, the 18th IAPR International Conference on Document Analysis and Recognition, that takes place in Athens, the beautiful and historic capital of Greece. ICDAR 2024 marks the start of the annual basis for the ICDAR series. ICDAR 2024 is the 18th edition of a longstanding conference series that has come of age, sponsored by the International Association of Pattern Recognition (IAPR). It is the premier international event for scientists and practitioners in document analysis and recognition. This field continues to play an important role in document understanding and recognition.

The IAPR TC10 / 11 technical committees endorse the conference. The very first ICDAR 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), Sydney, Australia (2019), Lausanne, Switzerland (2021) and San Jose, USA (2023).

Keeping with its tradition from past years, ICDAR 2024 featured a three-day main conference, including several competitions to challenge the field and a pre-conference slate of workshops, tutorials, and a doctoral consortium. The conference is being held in Athens, Greece on September 2{4, 2024, and the pre-conference tracks on August 30 till September 1, 2024. The highlights of the conference include keynote talks by the recipient of the IAPR/ICDAR Outstanding Achievements Award, and distinguished speakers: Jurgen Schmidhuber, Director of the AI Initiative at KAUST, Swiss AI Lab IDSIA, Univ. Lugano, Switzerland; Maria Kamilaki, Acting Director-General of e-Administration, Library & Publications of the Hellenic Parliament, Greece; Cheng-Lin Liu, State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation of Chinese Academy of Sciences, China.

A total of 263 papers were submitted to the main conference (plus 35 papers to the ICDAR-IJDAR journal track), with 66 papers accepted for oral presentation (plus 17 IJDAR track papers) and 94 for poster presentation. We would like to express our deepest gratitude to our Program Committee Chairs, featuring three distinguished researchers from academia, Elisa Barney Smith, Liangrui Peng, >>

and Marcus Liwicki, who did a phenomenal job in overseeing a comprehensive reviewing process and who worked tirelessly to put together a very thoughtful and interesting technical program for the main conference. We are also very grateful to the members of the Program Committee for their high-quality peer reviews. We extend our gratitude to our competition chairs, George Retsinas and Xiang Bai, for overseeing the competitions. The pre-conference featured 6 excellent workshops, 4 value filled tutorials, and the doctoral consortium. We would like to thank Harold Mouchère and Anna Zhu, the workshop chairs, Vincent Christlein and Alicia Fornès, the tutorial chairs, and KC Santosh and Andreas Fischer, the doctoral consortium chairs, for their efforts in putting together a wonderful pre-conference program. We would like to thank and acknowledge the hard work put in by our Publication Chairs, Giorgos Sfikas and Christophoros Nikou, who worked diligently to compile the camera-ready versions of all the papers and organize the conference proceedings with Springer. Many thanks are also due to our sponsorship, awards, industry, and publicity chairs for their support of the conference.

Finally, we would like to thank our many financial sponsors for their support and the conference attendees and authors, for helping make this conference a success. We sincerely hope that all attendees are having an enjoyable conference, a wonderful stay in Athens, and fruitful academic exchanges with their colleagues.

August 2024

Basilis Gatos

National Centre for Scientific Research "Demokritos"

Vassilis Katsouros

ATHENA Research Center

Foteini Simistira Liwicki

Luleå University of Technology

Topics

- Document image processing
- Physical and logical layout analysis
- Text and symbol recognition
- Handwriting recognition
- Document analysis systems
- Document classification
- Indexing and retrieval of documents
- Document synthesis
- Extracting document semantics
- NLP for document understanding
- Office automation
- Graphics recognition
- Human document interaction
- Document Representation Modeling
- Structured document generation
- Multimedia document analysis
- Mobile text recognition
- Pen-based document analysis
- Scene text detection and recognition
- Recognition of tables and formulas
- Historical document analysis
- Signature verification
- Document summarization and translation
- Document forensics and provenance
- Medical document analysis
- Document analysis for social good
- Document analysis for literature search
- Gold-standard benchmarks and datasets

Committees

General Chairs

Basilis Gatos | National Centre for Scientific Research "Demokritos", Greece

Vassilis Katsouros | Athena Research Center, Greece

Foteini Simistira Liwicki | Luleå University of Technology, Sweden

PC Chairs

Elisa Barney-Smith | Luleå University of Technology, Sweden

Marcus Liwicki | Luleå University of Technology, Sweden

Liangrui Peng | Tsinghua University, Beijing, China

Workshop Chairs

Harold Mouchère | Nantes Université, France

Anna Zhu | Wuhan University of Technology, China

Tutorial Chairs

Vincent Christlein | University of Erlangen-Nuremberg, Germany

Alicia Fornes | Universitat Autònoma de Barcelona, Spain

Competitions Chairs

Xiang Bai | Huazhong University of Science and Technology, China

George Retsinas | National Technical University of Athens, Greece

Publication Chairs

Giorgos Sfikas | University of West Attica, Greece

Christophoros Nikou | University of Ioannina, Greece

Doctoral Consortium Chairs

Andreas Fischer | University of Applied Sciences and Arts Western, Switzerland

KC Santosh | University of South Dakota, USA

Awards Chairs

Michael Blumenstein | University of Technology Sydney, Australia

Ioannis Pratikakis | Democritus University of Thrace, Greece

Local organisation Chairs

Elena Galifianaki | National Centre for Scientific Research "Demokritos", Greece

Pelagia Drosaki | National Centre for Scientific Research "Demokritos", Greece

Anastasios Kesidis | University of West Attica, Greece

Kosmas Kritsis | Athena Research Center, Greece

Friday, August 30, 2024

Parallel Workshops

9:00 AM - 10:45 AM **Parthenon II**
Workshop on Document Analysis Systems (DAS) I

9:00 AM - 10:45 AM **Parthenon I**
Workshop on Comics Analysis, Processing and Understanding (MANPU) I

9:00 AM - 10:45 AM **Metis**
Workshop on Machine Vision and NLP for Document Analysis (VINALDO) I

10:45 AM - 11:15 AM **Coffee Break**

11:15 AM - 1:15 PM **Parthenon II**
Workshop on Document Analysis Systems (DAS) II

11:15 AM - 1:15 PM **Parthenon I**
Workshop on Comics Analysis, Processing and Understanding (MANPU) II

11:15 AM - 1:15 PM **Metis**
Workshop on Machine Vision and NLP for Document Analysis (VINALDO) II

1:15 PM - 2:15 PM **Lunch Break**

2:15 PM - 4:00 PM **Parthenon II**
Workshop on Document Analysis Systems (DAS) III

2:15 PM - 4:00 PM **Parthenon I**
Workshop on Comics Analysis, Processing and Understanding (MANPU) III

4:00 PM - 4:30 PM Coffee Break

4:30 PM - 5:45 PM **Parthenon II**
Workshop on Document Analysis Systems (DAS) IV

4:30 PM - 5:45 PM **Parthenon I**
Workshop on Comics Analysis, Processing and Understanding (MANPU) IV

Saturday, August 31, 2024

Parallel Workshops

9:00 AM - 10:45 AM	**Parthenon II**	
	Workshop on Document Analysis Systems (DAS) V	
9:00 AM - 10:45 AM	**Parthenon I**	
	Workshop on Computational Paleography (WCP) I	
9:00 AM - 10:45 AM	**Metis**	
	Workshop on Advanced Analysis and Recognition of Parliamentary Corpora (ARPC) I	
10:45 AM - 11:15 AM	**Coffee Break**	
11:15 AM - 1:15 PM	**Parthenon II**	
	Workshop on Document Analysis Systems (DAS) VI	
11:15 AM - 1:15 PM	**Parthenon I**	
	Workshop on Computational Paleography (WCP) II	
11:15 AM - 1:15 PM	**Metis**	
	Workshop on Advanced Analysis and Recognition of Parliamentary Corpora (ARPC) II	
1:15 PM - 2:15 PM	**Lunch Break**	
2:15 PM - 4:00 PM	**Parthenon II**	
	Workshop on Document Analysis Systems (DAS) VII	
2:15 PM - 4:00 PM	**Parthenon I**	
	Workshop on Computational Paleography (WCP) III	
2:15 PM - 4:00 PM	**Metis**	
	Workshop on Automatically Domain-Adapted and Personalized Document Analysis (ADAPDA) I	
4:00 PM - 4:30 PM	**Coffee Break**	
4:30 PM - 5:45 PM	**Parthenon II**	
	Workshop on Document Analysis Systems (DAS) IX	
4:30 PM - 5:45 PM	**Parthenon I**	
	Workshop on Computational Paleography (WCP) IV	
4:30 PM - 5:45 PM	**Metis**	
	Workshop on Automatically Domain-Adapted and Personalized Document Analysis (ADAPDA) II	

Sunday, September 1, 2024

Parallel Tutorials

9:00 AM - 10:45 AM **Metis**

Tutorial I: ICDAR Tutorial on Private, Collaborative Learning in Document Analysis (Part 1)

Organizers: *Dimosthenis Karatzas, Rubèn Tito, Mohamed Ali Souibgui, Khanh Nguyen, Raouf Kerkouche, Kangsoo Jung, Marlon Tobaben, Joonas Jälkö, Vincent Poulain, Aurelie Joseph, Ernest Valveny, Josep Lladós, Antti Honkela, Catuscia Palamidessi, Mario Fritz*

9:00 AM - 10:45 AM **Parthenon II**

Tutorial II: Retrieval Augmented Generation (RAG): Bridging Document Analysis and Recognition with Large Language Models (Part 1)

Organizer: *Falak Shah*

9:00 AM - 10:45 AM **Parthenon I**

Tutorial IV: Hands-On Deep Learning for Document Analysis (Part 1)
Organizer: *Thomas M. Breuel*

10:45 AM - 11:15 AM **Coffee Break**

11:15 AM - 1:15 PM **Metis**

Tutorial I: ICDAR Tutorial on Private, Collaborative Learning in Document Analysis (Part 2)

Organizers: *Dimosthenis Karatzas, Rubèn Tito, Mohamed Ali Souibgui, Khanh Nguyen, Raouf Kerkouche, Kangsoo Jung, Marlon Tobaben, Joonas Jälkö, Vincent Poulain, Aurelie Joseph, Ernest Valveny, Josep Lladós, Antti Honkela, Catuscia Palamidessi, Mario Fritz*

11:15 AM - 1:15 PM **Parthenon II**

Tutorial II: Retrieval Augmented Generation (RAG): Bridging Document Analysis and Recognition with Large Language Models (Part 2)

Organizer: *Falak Shah*

11:15 AM - 1:15 PM **Parthenon I**

Tutorial IV: Hands-On Deep Learning for Document Analysis (Part 2)

Organizer: *Thomas M. Breuel*

1:15 PM - 2:15 PM **Lunch Break**

>>

Sunday, September 1, 2024

2:15 PM - 4:00 PM **Metis**

Tutorial I: ICDAR Tutorial on Private, Collaborative Learning in Document Analysis (Part 3)

Organizers: *Dimosthenis Karatzas, Rubèn Tito, Mohamed Ali Souibgui, Khanh Nguyen, Raouf Kerkouche, Kangsoo Jung, Marlon Tobaben, Joonas Jälkö, Vincent Poulain, Aurelie Joseph, Ernest Valveny, Josep Lladós, Antti Honkela, Catuscia Palamidessi, Mario Fritz*

4:00 PM - 4:30 PM **Coffee Break**

4:30 PM - 5:45 PM **Metis**

Tutorial I: ICDAR Tutorial on Private, Collaborative Learning in Document Analysis (Part 4)

Organizers: *Dimosthenis Karatzas, Rubèn Tito, Mohamed Ali Souibgui, Khanh Nguyen, Raouf Kerkouche, Kangsoo Jung, Marlon Tobaben, Joonas Jälkö, Vincent Poulain, Aurelie Joseph, Ernest Valveny, Josep Lladós, Antti Honkela, Catuscia Palamidessi, Mario Fritz*

8:30 PM - 11:00 PM **Welcome Drinks**

https://icdar2024.net/social-event/

Monday, September 2, 2024

9:00 AM - 9:40 AM Parthenon
Opening Ceremony

9:40 AM - 10:40 AM Parthenon
Keynote Speech: **Thoughts about Machine Learning**
Prof. Jürgen Schmidhuber, Director, AI Initiative, KAUST

10:40 AM - 11:10 AM Coffee Break

Parallel Sessions

11:10 AM - 1:10 PM Parthenon
Oral Session I:
Frontiers in Handwriting Recognition I

11:10 AM - 11:30 AM
O1.1 **Handwritten Document Recognition Using Pre-trained Vision Transformers** *Daniel Parres*

11:30 AM - 11:50 AM
O1.2 **The Learnable Typewriter: A Generative Approach to Text Analysis** *Ioannis Siglidis*

11:50 AM - 12:10 PM
O1.3 **Deep Metric Learning with Cross-Writer Attention for Offline Signature Verification** *Lu-Rong Ling*

12:10 PM - 12:30 PM
O1.4 **Janus-faced handwritten signature attack: a clash between Handwritten Signature Duplicator and a writer independent Metric Meta-Learning Offline Signature Verifier** *Elias Zois*

12:30 PM - 12:50 PM
O1.5 **SAGHOG: Self-Supervised Autoencoder for Generating HOG Features for Writer Retrieval** *Florian Kleber*

12:50 PM - 1:10 PM
O1.6 **Self-Supervised Vision Transformers for Writer Retrieval** *Gernot Fink*

11:10 AM - 1:10 PM Cyclades
Oral Session II:
Layout Analysis and Doc Class

11:10 AM - 11:30 AM
O2.1 **DLAFormer: An End-to-End Transformer for Document Layout Analysis** *Kai Hu*

11:30 AM - 11:50 AM
O2.2 **A LayoutLMv3-Based Model for Enhanced Relation Extraction in Visually-Rich Documents** *Wiam Adnan*

11:50 AM - 12:10 PM
O2.3 **Are Layout Analysis and OCR Still Useful for Document Information Extraction using Foundation Models?** *Anna Scius-Bertrand*

12:10 PM - 12:30 PM
O2.4 **Machine Unlearning for Document Classification** *Mohamed Souibgui*

12:30 PM - 12:50 PM
O2.5 **CICA: Content Injected Contrastive Alignment for Zero-Shot Document Image Classification** *Muhammad Zeshan Afzal*

>>

Monday, September 2, 2024

1:10 PM - 2:10 PM Lunch Break

Parallel Sessions

2:10 PM - 3:10 PM Parthenon
Oral Session III: Document Understanding and NLP

2:30 PM - 2:50 PM

03.1 **GDP: Generic Document Pretraining to Improve Document Understanding**

Santanu Chaudhury

2:30 PM - 2:50 PM

03.2 **GeoContrastNet: Contrastive Edge Learning in Graph Attention Networks for Document Understanding**

Sanket Biswas

2:50 PM - 3:10 PM

03.3 **EntityLayout: Entity-level Pre-training Language Model for Semantic Entity Recognition and Relation Extraction**

Cheng-Lin Liu

3:10 PM - 3:40 PM Coffee Break

2:10 PM - 3:10 PM Cyclades
Oral Session IV: Music Recognition

2:10 PM - 2:30 PM

04.1 **Practical End-to-End Optical Music Recognition for Pianoform Music**

Jiří Mayer

2:30 PM - 2:50 PM

04.2 **Source-Free Domain Adaptation for Optical Music Recognition**

Adrián Roselló

2:50 PM - 3:10 PM

04.3 **Sheet Music Transformer: End-To-End Optical Music Recognition Beyond Monophonic Transcription**

Antonio Ríos-Vila

Monday, September 2, 2024

Parallel Sessions

3:40 PM - 5:20 PM **Parthenon**

Oral Session V: Frontiers in Handwriting Recognition II - Journal Track

3:40 PM - 4:05 PM

05.1 Exploring Recursive Neural Networks for Compact Handwritten Text Recognition Models

Jorge Calvo-Zaragoza

4:05 PM - 4:30 PM

05.2 Self-Training for Handwritten Word Recognition and Retrieval

Fabian Wolf

4:30 PM - 4:55 PM

05.3 Neural Models for Semantic Analysis of Handwritten Document Images

Oliver Tueselmann

4:55 PM - 5:20 PM

05.4 Evaluating Learned Feature Aggregators for Writer Retrieval

Vincent Christlein

3:40 PM - 5:20 PM **Cyclades**

Oral Session VI: Journal Track

3:40 PM - 4:05 PM

06.1 Towards Privacy Preserved Document Image Classification - A Comprehensive Benchmark

Saifullah Saifullah

4:05 PM - 4:30 PM

06.2 DocXClassifier: Towards a Robust and Interpretable Deep Neural Network for Document Image Classification

Saifullah Saifullah

4:30 PM - 4:55 PM

06.3 SemiDocSeg: Harnessing Semi-Supervised Learning for Document Layout Analysis

Sanket Biswas

4:55 PM - 5:20 PM

06.4 CHWMaster: Mastering Chinese Handwriting via Sliding-Window Recurrent Neural Networks

Zhouhui Lian

3:10 PM - 3:40 PM Coffee Break

5:20 PM - 6:20 PM Industry Panel

6:20 PM - 7:20 PM TC-10 / TC-11 Meeting

Tuesday, September 3, 2024

Parallel Sessions

9:00 AM - 10:40 AM **Parthenon**

Oral Session VII: Journal Track II

9:00 AM - 9:25 AM

07.1 **End-to-End Semi-Supervised approach with Modulated Object Queries for Table Detection in Documents.**

Tahira Shehzadi

9:25 AM - 9:50 AM

07.2 **Table Image Dewarping with Key Element Segmentation**

Ziyi Zhu

9:50 AM - 10:15 AM

07.3 **ChemScraper: Graphics Extraction, Molecular DiagramParsing, and Annotated Data Generation for PDF Images**

Bryan Amador

10:15 AM - 10:40 AM

07.4 **Am I Readable? Transfer Learning based Document Image Rectification**

Pooja Kumari

9:00 AM - 10:40 AM **Cyclades**

Oral Session VIII: Music and Scene Text Recognition

9:00 AM - 9:25 AM

08.1 **A Unified Representation Framework for the Evaluation of Optical Music Recognition Systems**

Pau Torras

9:25 AM - 9:50 AM

08.2 **Towards Reduced-Complexity Scene Text Recognition (RCSTR) through a Novel Salient Feature Selection**

Rina Buoy

9:50 AM - 10:15 AM

08.3 **PARSTR: Partially Autoregressive Scene Text Recognition**

Rina Buoy

10:15 AM - 10:40 AM

08.4 **A New Unsupervised Approach Text Localization in Shaky and Non-Shaky Scene Video**

Ayan Banerjee

10:40 AM - 11:10 AM **Coffee Break**

11:10 AM - 12:10 PM **Parthenon II**

Keynote Speech: **Towards Explainable Document Recognition**

Prof. Cheng-Lin Liu, Director State Kay Laboratory of Multimodal Institute of Automation of Chinese Academy of Sciences (CASIA)

Tuesday, September 3, 2024

Parallel Sessions

12:10 PM - 1:10 PM **Parthenon**
Oral Session IX: Text & Symbol Recognition

12:10 PM - 12:30 PM

09.1 Class Incremental Learning for Character String Recognition

Qiufeng Wang

12:30 PM - 12:50 PM

09.2 Improving Automatic Text Recognition with Language Models in the PyLaia Open-Source Library

Solène Tarride

12:50 PM - 1:10 PM

09.3 SketchProphet: A Vectorized Autoregressive Modeling for Sketch Generation and Analysis

Sanket Biswas

12:10 PM - 1:10 PM **Cyclades**
Oral Session X: Scene Text Recognition

12:10 PM - 12:30 PM

010.1 Adaptive Scaling and Refined Pyramid Dual Attention Network for Scene Text Segmentation

Heng Zhang

12:30 PM - 12:50 PM

010.2 Knowledge Mining of Scene Text for Referring Expression Comprehension

Chenyang Gao

12:50 PM - 1:10 PM

010.3 PARSTR: Partially Autoregressive Scene Text Recognition

Yuxin Kong

1:10 PM - 2:10 PM **Lunch Break**

2:10 PM - 3:10 PM **Parthenon II**

Keynote Speech **IAPR/ICDAR Young Investigator Award 2024**

Unraveling Scribal Authorship: New Frontiers in Writer Retrieval

Vincent Christlein

3:10 PM - 4:40 PM **Coffee Break**

Poster Session I & Doctoral Consortium

Tuesday, September 3, 2024

Parallel Sessions

4:40 PM - 6:20 PM **Parthenon**

Oral Session XI: Tables & Forms

4:40 PM - 5:00 PM

O11.1 ClusterTabNet: Supervised clustering method for table detection and table structure recognition

Marek Polewczyk

5:00 PM - 5:20 PM

O11.2 SPRINT: Script-agnostic Structure Recognition in Tables

Badri Vishal Kasuba

5:20 PM - 5:40 PM

11.3 Latent Diffusion for Guided Document Table Generation

Sheraz Ahmed

5:40 PM - 6:00 PM

O11.4 Towards End-to-End Semi-Supervised Table Detection with Semantic Aligned Matching Transformer

Tahira Shehzadi

6:00 PM - 6:20 PM

O11.5 Mining and Analyzing Statistical Information from Untranscribed Form Images

José Andrés Moreno

4:40 PM - 6:20 PM **Cyclades**

Oral Session XII: Historical Document Analysis I

4:40 PM - 5:00 PM

O12.1 Adaptive Scaling and Refined Pyramid Dual Attention Network for Scene Text Segmentation

Nimol Thuon

5:00 PM - 5:20 PM

O12.2 Approximate Ground Truth Generation for Semantic Labeling of Historical Documents with Minimal Human Effort

Rolf Ingold

5:20 PM - 5:40 PM

O12.3 Clustering Running Titles to Understand the Printing of Early Modern Books

Nikolai Vogler

5:40 PM - 6:00 PM

O12.4 Active Learning with Relevance Feedback for Handwriting Detection in Historical Print

Jacob Murel

8:30 PM - 11:00 PM **Conference Social Event**

https://icdar2024.net/social-event/

Tuesday, September 3, 2024

Poster Session I

3:10 PM - 4:40 PM

P1.1 **GraphMLLM: A Graph-based Multi-level Layout Language-Independent Model for Document Understanding**
Xiao-Hui Li

P1.2 **One-shot Transformer-based Framework for Visually-Rich Document Understanding**
Chening Yang

P1.3 **Light-Weight Multi-Modality Feature Fusion Network for Visual Rich Document Understanding**
Chening Yang

P1.4 **Embedding Layout in Text for Document Understanding Using Large Language Models**
Mohammad Minouei

P1.5 **Doc-DINO: A Transformer Model for Complex Logical Document Layout Analysis**
Mayire Ibrahim

P1.6 **Font Style Interpolation with Diffusion Models**
Daichi Haraguchi

P1.7 **Learning to Kern — Set-wise Estimation of Optimal Letter Space**
Seiichi Uchida

P1.8 **Geometric-aware control in diffusion model for handwritten Chinese font generation**
Ning Ding

P1.9 **Typographic Text Generation with Off-the-Shelf Diffusion Model**
Daichi Haraguchi

P1.10 **Drawing the Line: Deep Segmentation for Extracting Art from Ancient Etruscan Mirrors**
Robert Sablatnig

P1.11 **Historical Printed Ornaments: Dataset and Tasks**
Zeynep Sonat Baltaci

Tuesday, September 3, 2024

Poster Session I

P1.12 The Socface Project: Large-Scale Collection, Processing, and Analysis of a Century of French Censuses

Christopher Kermorvant

P1.13 Recognition of Components in Taoist Charm Images

Hsiang-An Wang

P1.14 Text Line Segmentation on Ancient Egyptian Papyri: Layout Analysis with Object Detection Networks and Connected Components

Stephan Unter

P1.15 DELINE8K: A Synthetic Pipeline for the Semantic Segmentation of Historical Documents

Taylor Archibald

P1.16 Callico: a versatile open-source document image annotation platform

Christopher Kermorvant

P1.17 A Historical Handwritten Dataset for Ethiopic OCR with Baseline Models and Human-level Performance

Birhanu Hailu Belay

P1.18 HistNERo: Historical Named Entity Recognition for the Romanian Language

Andrei-Marius Avram

P1.19 Fully automatic virtual unwrapping method for documents imaged by X-ray tomography

Vladimir Arlazarov

P1.20 Enhancing Recognition of Historical Musical Pieces with Synthetic and Composed images

Manuel Villarreal

P1.21 On Image Processing and Pattern Recognition for Thermograms of Watermarks in Manuscripts - A First Proof-of-Concept

Matthias Beckmann

Tuesday, September 3, 2024

Poster Session I

P1.22 SegHist: A General Segmentation-based Framework for Chinese Historical Document Text Line Detection

Liangcai Gao

P1.23 Recognition and Link Prediction of Onomatopoeia Texts with Arbitrary Shapes

Hongxi Wei

P1.24 Oracle Bone Inscriptions Image Retrieval Based on Metric Learning

Alimjan Aysa

P1.25 End-to-end information extraction in handwritten documents: understanding Paris marriage records from 1880 to 1940

Thomas Constum

P1.26 LMTextSpotter: Towards Better Scene Text Spotting with Language Modeling in Transformer

Guodong Ding

P1.27 Progressive Evolution from Single-Point to Polygon for Scene Text

Yuliang Liu

P1.28 A New Bottom-up Path Augmentation Attention Network for Script Identification in Scene Images

Kurban Ubul

P1.29 MOoSE: Multi-Orientation Sharing Experts for Open-set Scene Text Recognition

Elisa Barney Smith

P1.30 More and Less: Enhancing Abundance and Refining Redundancy for Text-prior-guided Scene Text Image Super-Resolution

Mayire Ibrahim

P1.31 A Real-Time Scene Uyghur Text Detection Network Based on Feature Complementation

Mayire Ibrahim

P1.32 Controllable text layout generation for synthesizing scene text image

Jiangyang He

Tuesday, September 3, 2024

Poster Session I

P1.33 The First Swahili Language Scene Text Detection and Recognition Dataset
Ling Fu

P1.34 Indic Scene Text on the Roadside
Cheerakkuzhi Veluthemana Jawahar

P1.35 Dataset and Benchmark for Urdu Natural Scenes Text Detection and Recognition and Visual Question Answering
Ling Fu

P1.36 Cross-Domain Image Conversion by CycleDM
Sho Shimotsumagari

P1.37 Self-supervised Pre-training of Text Recognizers
Michal Hradis

P1.38 Counting the Corner Cases: Revisiting Robust Reading Challenge Data Sets, Evaluation Protocols, and Metrics
Dimosthenis Karatzas

P1.39 Coarse-to-Fine Document Image Registration for Dewarping
Qiufeng Wang

P1.40 YOLO Assisted A* Algorithm for Robust Line Segmentation of Degraded Document Images
Ujjwal Bhattacharya

P1.41 DistilDoc: Knowledge Distillation for Visually-Rich Document Applications
Josep Llados

P1.42 Deep Learning Enabled Functional Knowledge Unit Innovation Generation Model
Yahong Hu

P1.43 Global-SEG: Text semantic segmentation based on global semantic pair relations
Mickael Coustaty

P1.44 Multimodal Adaptive Inference with Anytime Early Exiting
Souhail Bakkali

Tuesday, September 3, 2024

Poster Session I

P1.45 Integrating Dependency Type and Directionality into Adapted Graph Attention Networks to Enhance Relation Extraction

Yiran Zhao

P1.46 An Ultra-Lightweight Approach for Machine Readable Zone Detection via Semantic Segmentation and Fast Hough Transform

Vladimir Arlazarov

P1.47 Enhancing CRNN HTR Architectures with Transformer Blocks

Konstantina Nikolaidou

Wednesday, September 4, 2024

Parallel Sessions

9:00 AM - 10:40 AM **Parthenon**

Oral Session XIII: Visual Question Answering and Comics

9:00 AM - 9:20 AM

013.1 **Extractive Question Answering with Contrastive Puzzles and Reweighted Clues.**

Chao Liu

9:20 AM - 9:40 AM

013.2 **CHIC: Corporate Document for Visual Question Answering**

Mickael Coustaty

9:40 AM - 10:00 AM

013.3 **UniVIE: A Unified Label Space Approach to Visual Information Extraction from Form-like Documents**

Kai Hu

10:00 AM - 10:20 AM

013.4 **Federated Document Visual Question Answering: A Pilot Study**

Dimosthenis Karatzas

10:20 AM - 10:40 AM

013.5 **Multimodal Transformer for Comics Text-Cloze**

Ernest Valveny

9:00 AM - 10:40 AM **Cyclades**

Oral Session XIV: Historical Document Analysis II

9:00 AM - 9:20 AM

014.1 **CATMuS Medieval: A multilingual large-scale cross-century dataset in Latin script for handwritten text recognition and beyond**

Malamatenia Vlachou Efstathiou

9:20 AM - 9:40 AM

014.2 **Historical Astronomical Diagrams Decomposition in Geometric Primitives**

Syrine Kalleli

9:40 AM - 10:00 AM

014.3 **Zipf Curves and Basic Text Analytics from Untranscribed Manuscript Images**

Alejandro Toselli

10:15 AM - 10:40 AM

014.4 **Binarizing Documents by Leveraging both Space and Frequency**

Silvia Cascianelli

10:15 AM - 10:40 AM

014.5 **Exploring Knowledge Distillation Towards Document Object Detection with Structured Graph Creation**

Ayan Banerjee

10:40 AM - 11:10 AM **Coffee Break**

>>

Wednesday, September 4, 2024

11:10 AM - 12:10 PM — **Parthenon II**

Keynote Speech

Sharing the past, preparing the future. The digital transformation of the Hellenic Parliament Library

Maria Kamilaki, Acting Director General D.G. for e-Administration, Library & Publications Hellenic Parliament

Parallel Sessions

12:10 PM - 1:10 PM — **Parthenon**

Oral Session XV: Chinese Text Recognition

12:10 PM - 12:30 PM

O15.1 Visual Prompt Learning for Chinese Handwriting Recognition

Ning Ding

12:30 PM - 12:50 PM

O15.2 Puzzle Pieces Picker: Deciphering Ancient Chinese Characters with Radical Reconstruction

Pengjie Wang

12:50 PM - 1:10 PM

O15.3 UniVIE: Context-Aware Confidence Estimation for Rejection in Handwritten Chinese Text Recognition

Yangyang Liu

12:10 PM - 1:10 PM — **Cyclades**

Oral Session XVI: Fonts & Scripts

12:10 PM - 12:30 PM

O16.1 Script Identification in the Wild with FFT-Multi-Grained Mix Attention Transformer

Malamatenia Vlachou Efstathiou

12:30 PM - 12:50 PM

O16.2 Font Impression Estimation in the Wild

Daichi Haraguchi

12:50 PM - 1:10 PM

O16.3 Impression-CLIP: Contrastive Shape-Impression Embedding for Fonts

Yugo Kubota

1:10 PM - 2:10 PM — **Lunch Break**

Wednesday, September 4, 2024

Parallel Sessions

2:10 PM - 2:50 PM **Parthenon**
Oral Session XVII: Mathematical Expression Recognition

2:10 PM - 2:30 PM

017.1 ICAL: Implicit Character-Aided Learning for Enhanced Handwritten Mathematical Expression Recognition

Jianhua Zhu

2:30 PM - 2:50 PM

017.2 Stroke-Level Graph Labeling with Edge-weighted Graph Attention Network for Handwritten Mathematical Expression Recognition

Harold Mouchère

2:10 PM - 2:50 PM **Cyclades**
Oral Session XVIII: Transformers

2:10 PM - 2:30 PM

018.1 Dynamic Relation Transformer for Contextual Text Block Detection

Kai Hu

2:30 PM - 2:50 PM

018.2 End-to-End Table Transformer

Taehoon Kim

2:50 PM - 4:20 PM **Coffee Break**

Poster Session II & Doctoral Consortium

4:20 PM - 5:30 PM

Competitions I **Parthenon**

Competitions II **Cyclades**

5:30 PM - 6:20 PM

Closing Session & Award Ceremony **Parthenon**

Wednesday, September 4, 2024

Poster Session II

2:50 PM - 4:20 PM

P2.15 **Deepfake In-air Signature Verification via Two-channel Model**

Hongxi Wei P2.1 ViT-ED: Transformer network for image similarity measurement

Marie Beurton

P2.2 **Synergistic Diverse Perspective for Topic Evolution Analysis on Weibo**

Jianing Zhang

P2.3 **KVP10k: A Comprehensive Dataset for Key-Value Pair Extraction in Business Documents**

Oshri Naparstek

P2.4 **Weakly Supervised Training for Hologram Verification in Identity Documents**

Joseph Chazalon

P2.5 **Multi-task Learning for License Plate Recognition in Unconstrained Scenarios**

Zhenlun Mo

P2.6 **Recurrent Few-Shot model for Document Verification**

Oriol Ramos

P2.7 **SlideCraft: Synthetic Slides Generation for Robust Slide Analysis**

Travis Seng

P2.8 **A Multiclass Imbalanced Dataset Classification of Symbols from Piping and Instrumentation Diagrams**

Eyad Elyan

P2.9 **Document Specular Highlight Removal with Coarse-to-fine Strategy**

Xin Yang

P2.10 **Radical Similarity Based Model Optimization and Post-correction for Chinese Character Recognition**

Jun Du

P2.11 **PCA-Based Adversarial Attacks on Signature Verification Systems**

Faisal Shafait

Wednesday, September 4, 2024

Poster Session II

P2.12 Analysis of the Calibration of Handwriting Text Recognition Models

Eric Ayllon

P2.13 Test Time Augmentation as a Defense Against Adversarial Attacks on Online Handwriting

Brian Kenji Iwana

P2.14 Robust Handwritten Signature Representation with Continual Learning of Synthetic Data over Predefined Real Feature Space

Talles Brito Viana

P2.16 Content-based Similarity for Automatic Scoring of Handwritten Descriptive Answers

Tuan Nam Ly

P2.17 LABT: A Sequence-to-Sequence Model for Mongolian Handwritten Text Recognition with Local Aggregation BiLSTM and Transformer

Hongxi Wei

P2.18 From Handwriting Analysis to Alzheimer's Disease Prediction: An Experimental Comparison of Classifier Combination Methods

Tiziana D'Alessandro

P2.19 Text Enhancement for Historical Handwritten Documents

Jihad El-Sana

P2.20 Bridging the Gap in Resource for Offline English Handwritten Text Recognition

Cheerakkuzhi Veluthemana Jawahar

P2.21 StylusAI: Stylistic Adaptation for Robust German Handwritten Text Generation

Nauman Riaz

P2.22 Reading Order Independent Metrics for Information Extraction in Handwritten Documents

Solène Tarride

P2.23 BRESSAY: A Brazilian Portuguese Dataset for Offline Handwritten Text Recognition

Byron Leite Dantas Bezerra

Wednesday, September 4, 2024

Poster Session II

P2.24 A Domain Adaptive Hybrid Approach for Document Layout Analysis in Document images

Tahira Shehzadi

P2.25 A region-based approach for layout analysis of music score images in scarce data scenarios

Francisco J. Castellanos

P2.26 LAPDoc: Layout-Aware Prompting for Documents

Darko Obradovic

P2.27 CREPE: Coordinate-Aware End-to-End Document Parser

Youngmin Baek

P2.28 DocXplain: A Novel Model-Agnostic Explainability Method for for Document Image Classification

Saifullah Saifullah

P2.29 Improving Efficiency and Performance through CTC-based Transformers for Mathematical Expression Recognition

Dan Anitei

P2.30 The KuiSCIMA Dataset for Optical Music Recognition of Ancient Chinese Suzipu Notation

Tristan Repolusk

P2.31 WikiDT: Visual-based Table Recognition and Question Answering Dataset

Hui Shi

P2.32 Synthesizing Realistic Data for Table Recognition

Qiyu Hou

P2.33 Multi-Cell Decoder and Mutual Learning for Table Structure and Character Recognition

Takaya Kawakatsu

Wednesday, September 4, 2024

Poster Session II

P2.34 DTSM: Toward Dense Table Structure Recognition with Text Query Encoder and Adjacent Feature Aggregator

Xinhong Chen

P2.35 AltChart: Multi-Pretext Tasks for Better Chart Summaries

Omar Moured

P2.36 DocTabQA: Answering Questions from Long Documents Using Tables

Liangcai Gao

P2.37 ChartReformer: Natural Language-Driven Chart Image Editing

David Doermann

P2.38 DT-VQA: A Visual Question Answering Dataset for Exploring the Capabilities of LMMs on Dense Text

Yuliang Liu

P2.39 ConClue: Conditional Clue Extraction for Multiple Choice Question Answering

Wangli Yang

P2.40 Multi-Page Document Visual Question Answering using Self-Attention Scoring Mechanism

Lei Kang

P2.41 Privacy-Aware Document Visual Question Answering

Aurélie Joseph

P2.42 Information Extraction from Visually Rich Documents using Directed Weighted Graph Neural Network

Hamza Gbada

P2.43 IndicBART alongside Visual Element: Multimodal Summarization in Diverse Indian Languages

Deepak Prakash

P2.44 Improving Retrieval-Based Dialogue Systems: Fine-grained Post-training Prompt Adaptation and Pairwise Optimization

Alimjan Aysa

Wednesday, September 4, 2024

Poster Session II

P2.45 Improving Automatic Text Recognition With Explicit Language Modeling

Solène Tarride

P2.46 Dynamic Reasoning with Language Model and Knowledge Graph for Question Answering

Yujie Liu

Tuesday & Wednesday, September 3 & 4, 2024

Doctoral Consortium

3:40 PM - 4:05 PM

DC.1 Multi-Modal Structural Reasoning for Historical Document Information Extraction

Adria Molina

DC.2 Self-Supervised Learning for Handwritten Text Recognition

Carlos Peñarrubia Morcillo

DC.3 Historical Handwritten Isolated Glyph and Text Recognition for Palm Leaf Manuscripts

Nimol Thuon

DC.4 Mathematical Equation Recognition for Patent Publication and Accessibility

François Wieckowiak

DC.5 Studying the effect of integrating numerical methods on efficient learning of graph-based representations

Carlos Boned

DC.6 Multi-Modal Models for Explainable Document Understanding

Artemis Llabres

DC.7 Sketching Imagination: Explorations of Generative Models for Visual Creativity

Ayan Banerjee

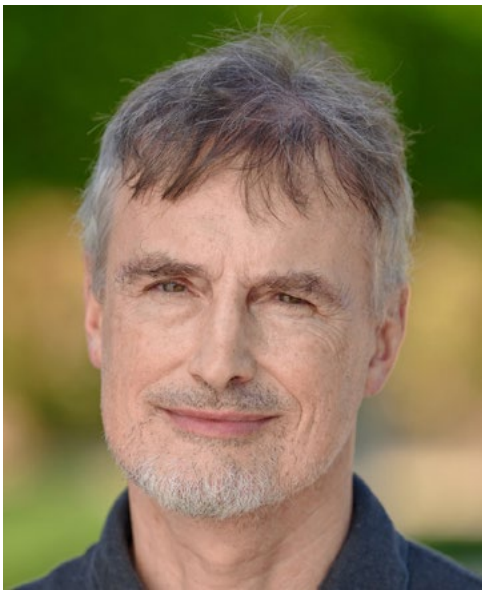
DC.8 Domain Generalization for Handwritten Text Recognition

Carlos Garrido-Munoz

DC.9 Graph Neural Networks for Handwriting Recognition

Tim Hallyburton

Keynote Speakers



Prof Jürgen Schmidhuber

Director, AI Initiative, KAUST

Monday, September 2, 2024

9:40 AM - 10:40 AM, Parthenon I

Thoughts about Machine Learning

Speaker Bio

The New York Times headlined: "When A.I. Matures, It May Call Jürgen Schmidhuber 'Dad'." Since age 15, his main goal has been to build a self-improving A.I. smarter than himself, then retire. His lab's deep learning artificial neural networks based on ideas published in the "Annus Mirabilis" 1990-1991 have revolutionised machine learning and A.I. By 2017, they were on over 3 billion smartphones, and used billions of times per day, for Facebook's automatic translation, Google's speech recognition, Google Translate, Apple's Siri & QuickType, Amazon's Alexa, etc. He pioneered generative adversarial networks (1990, now widely used), artificial curiosity, Transformers with linearized self-attention (1991 – Transformers are the basis of the famous ChatGPT), and meta-learning machines that learn to learn (since 1987). Today, the most cited neural networks all build on work done in his labs. Elon Musk tweeted: "Schmidhuber invented everything."
He is recipient of numerous awards, Director of the AI Initiative at KAUST in KSA, Scientific Director of the Swiss AI Lab IDSIA, Adj. Prof. of A.I. at Univ. Lugano, and Co-Founder & Chief Scientist of the company NNAISENSE. He is a frequent keynote speaker at major events, and advising various governments on A.I. strategies.

Keynote Speakers

Dr Maria Kamilaki

Acting Director General D.G.
for e-Administration, Library
& Publications Hellenic Parliament

Wednesday, September 4, 2024

11:10 AM - 12:10 PM, Parthenon II

Sharing the past, preparing the future. The digital transformation of the Hellenic Parliament Library

Speaker Bio

Dr Maria Kamilaki is Acting Director-General of e-Administration, Library & Publications of the Hellenic Parliament. She holds a PhD in Sociolinguistics (University of Athens), a MSc in Applied Linguistics & English Language Teaching (University of Edinburgh) and a MSc in Cultural Management (Panteion University of Social & Political Sciences). She teaches at the Hellenic Open University postgraduate course Current trends in Linguistics for Teachers, and works as a Training Program Developer at the National Centre for Public Administration and Local Government. She is co-author of the book Pepper in the mouth! Aspects of taboo words in Standard Modern Greek and author of 'Words that smile, words that hurt': Verbal bullying in the school environment. A teacher's Guide. She has also published a long series of papers in the field of Sociolinguistics and Language Teaching. Her research interests currently lie in parliamentary discourse analysis, carrying out a postdoctoral research, entitled From language attitudes to language policies: Aspects of the Greek Language Question in parliamentary discourse (University of the Aegean). She has a long-standing experience in designing educational programs and outreach activities at the Hellenic Parliament, such as Glossopolis: A multimodal exhibition on Modern Greek linguistic variety.

>>

Keynote Speakers

Prof Cheng-Lin Liu

Director State Kay Laboratory of Multimodal Institute of Automation of Chinese Academy of Sciences (CASIA)

Tuesday, September 3, 2024

11:10 AM - 12:10 PM, Parthenon II

Towards Explainable Document Recognition

Speaker Bio

Cheng-Lin Liu is a Professor at the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation of Chinese Academy of Sciences. He is a vice president of the Institute of Automation, a vice dean of the School of Artificial Intelligence, University of Chinese Academy of Sciences. He received the PhD degree in pattern recognition and intelligent control from the Chinese Academy of Sciences, Beijing, China, in 1995. He was a postdoctoral fellow in Korea and Japan from March 1996 to March 1999. From 1999 to 2004, he was a researcher at the Central Research Laboratory, Hitachi, Ltd., Tokyo, Japan. His research interests include pattern recognition, machine learning and document image analysis. He has published over 400 technical papers in journals and conferences. He is an Associate Editor-in-Chief of Pattern Recognition Journal and Acta Automatica Sinica, an Associate Editor of International Journal on Document Analysis and Recognition, Cognitive Computation, IEEE/CAA Journal of Automatica Sinica, Machine Intelligence Research, CAAI Trans. Intelligence Technology, CAAI Artificial Intelligence Research and Chinese Journal of Image and Graphics. He is a Fellow of the CAA, CAAI, the IAPR and the IEEE.

Keynote Speakers

Vincent Christlein

Pattern Recognition Lab,
Friedrich-Alexander-Universität
Erlangen-Nürnberg (FAU)

Unraveling Scribal Authorship: New Frontiers in Writer Retrieval

Speaker Bio

Vincent Christlein heads the Computer Vision group at the Pattern Recognition Lab, Friedrich-Alexander University of Erlangen-Nürnberg (FAU), Germany. He received his diploma and Dr.-Ing. degrees from FAU in 2012 and 2018, respectively. His primary research focus is on document analysis, including writer identification and handwriting imitation, as well as environmental projects such as glacier front segmentation and bird detection. In the field of document analysis, his work has earned recognition through several international competition wins and multiple awards.

Abstract talk

How have recent advancements in machine learning transformed the field of writer identification and retrieval from handwritten text images? This presentation delves into the evolution from conventional techniques to state-of-the-art deep learning approaches. The first part of the talk dissects the process of writer retrieval, highlighting essential components and discussing significant contributions to the field. It will examine how modern deep learning technologies have significantly improved the accuracy and efficiency of identifying writers.

The second half of the talk shifts focus to the rapidly advancing area of handwriting generation and imitation. Initially facilitated by generative adversarial networks for handwriting synthesis, diffusion-based methods have recently taken the lead as more robust alternatives, capable of producing more diverse and realistic handwritten text. What implications do these emerging technologies hold for the future of document analysis? The discussion will highlight the potential impacts, emphasizing how these developments could reshape the landscape of writer identification and retrieval.

Tuesday, September 3, 2024, 2:10 PM - 3:10 PM, Parthenon II

SPONSORS

snowflake

Goodnotes™

Baidu 百度

wacom

Adobe

Diamond Sponsor

Snowflake makes enterprise AI easy, efficient, and trusted. Thousands of companies around the globe, including hundreds of the world's largest, use Snowflake's AI Data Cloud to share data, build applications, and power their business with AI. The era of enterprise AI is here. There is only one AI Data Cloud. Snowflake's founders started from scratch and built a data and AI platform that would harness the immense power of the cloud. But their vision didn't stop there. They engineered Snowflake to power the AI Data Cloud, where thousands of organizations unite siloed data, discover and securely share data, power data applications, and execute diverse AI & ML workloads.

Canva Ad:

Join us and redefine what's possible in the realm of data and AI. We're committed to substantial investments in AI & ML - people, technology, and infrastructure. With numerous zero-to-one opportunities in the nascent field of AI & ML, you'll dive into early product building and tackle the toughest challenges head-on. Be a part of a team focused on getting things done efficiently in a small, flat organization where visibility is high, collaboration is key, and politics are minimal. Embrace our strong data gravity and well-established brand to become an integral part of the AI & ML revolution.

Our team in Warsaw works on:

DocAI (TILT model):

https://arxiv.org/abs/2102.09550 eration

Snowflake Copilot:

https://www.snowflake.com/blog/copilot-ai-powered-sql-assistant/

Arctic model:

https://www.snowflake.com/blog/arctic-open-efficient-foundation-language-models-snowflake/

Explore our scientific publications below to learn more about what we do and our exciting vision for the future. Our team is composed of highly skilled specialists from diverse backgrounds, including engineering, machine learning, mathematics, statistics, data science, computational linguistics, philosophy, and more. Over half of our team members hold a Ph.D., making us a hub of expertise and knowledge. It's an amazing group of people who are passionate about shaping the future.

List of recent publications:

STable: Table Generation Framework for Encoder-Decoder Models

Document Understanding Dataset and Evaluation (DUDE)

CCpdf: Building a High Quality Corpus for Visually Rich Documents from Web Crawl Data

WE'RE HIRING

SNOWFLAKE AI & ML ENGINEERING

WHERE RESEARCH MEETS FULL-SCALE PRODUCTION

Join us and redefine what's possible in the realm of data and AI. We're committed to substantial investments in AI & ML - people, technology, and infrastructure. With numerous zero-to-one opportunities in the nascent field of AI & ML, you'll dive into early product building and tackle the toughest challenges head-on. Be a part of a team focused on getting things done efficiently in a small, flat organization where visibility is high, collaboration is key, and politics are minimal. Embrace our strong data gravity and well-established brand to become an integral part of the AI & ML revolution.

Our team in Warsaw works on:

- DocAI (TILT Model)
- Snowflake Copilot
- Arctic Model

Explore our scientific publications below to learn more about what we do and our exciting vision for the future. Our team is composed of highly skilled specialists from diverse backgrounds, including engineering, machine learning, mathematics, statistics, data science, computational linguistics, philosophy, and more. Over half of our team members hold a Ph.D., making us a hub of expertise and knowledge. It's an amazing group of people who are passionate about shaping the future.

List of recent publications:

- STable: Table Generation Framework for Encoder-Decoder Models
- Document Understanding Dataset and Evaluation (DUDE)
- CCpdf: Building a High-Quality Corpus for Visually Rich Documents

EXPLORE ROLES

Goodnotes is a beloved digital notetaking app that was named Apple's 2022 iPad App of the Year and is used by millions of educators and learners around the world. We offer our users the possibility to create, learn and collaborate through our multi-functional app and then to share their content in a way that is completely paperless. Our mission is to revolutionize the way ideas are shaped and amplified.

Goodnotes™

Apple iPad App of the Year 2022

Build the future of digital paper with us

Goodnotes is a top-rated, award winning productivity app, used by millions of students, working professionals and paperless enthusiasts all around the world.

Research, train, and deploy your models for millions of users worldwide.

We're looking for machine learning engineers to work on:

- **Multi-lingual handwriting recognition**
- **O C R**
- **Natural language processing**
- **Ask Goodnotes LLM**

See our openings and apply at **goodnotes.com/careers** today!

What machine learning at Goodnotes looks like

As an AI Product Designer, I know that being part of the ML team means you're not just coding; you're shaping the future of digital note-taking. I collaborate closely with engineers at every stage of the product lifecycle. Together, we transform innovative ideas into intuitive features that enhance user experience. This teamwork allows us to create solutions that truly resonate with our users, making digital note-taking smarter and more efficient.

Izzy, AI Product Designer, Ask Goodnotes Team

The AANT team at Goodnotes brings constant innovation to its note-taking application. In Goodnotes, we have the opportunity to work on cutting-edge technology and have the freedom to design solutions end-to-end. We run many interesting projects involving organising and beautifying note contents, handwriting recognition, spellcheck and word-complete for handwritten words, automated math solvers and many more. The unique experience of adopting state-of-the-art works for the on-device deployment makes an engineer think out of the box.

Ranajit, Senior ML Engineer, AI Assisted Note Taking Team

We're excited to create tools like auto-grading and answer clustering to ease teachers' tasks and develop AI textbooks with clear insights and outlines for students. Every day, we tackle challenges involving diverse input types, writing styles, languages, and complex layouts. You'll have the chance to use innovative techniques like LLMs, layout recognition, and image processing while balancing creativity with high quality and fast deployment. You'll work closely with teachers and be involved in the entire AI development process, seeing firsthand how our solutions make a meaningful impact around the world.

Trung, ML Engineer, Education Team

Are you up for the challenge?

See our openings and apply at goodnotes.com/careers today!

Gold Sponsor

Founded in 2000, Baidu's mission is to make the complicated world simpler through technology. Baidu is a leading AI company with a strong Internet foundation. We are one of the few companies worldwide operating on a full-stack layout in AI. Driven by the company's deep learning platform PaddlePaddle and pre-trained model ERNIE with real- world applications across industries, Baidu stands at the forefront of the industry, achieving end-to-end optimization and significantly improving operational efficiency.

Over the past decade, we've invested nearly 170 billion RMB in research and development. This sustained investment has propelled us to the forefront of the AI landscape, with pioneering self-developed technologies and leading positions in AI patent applications.

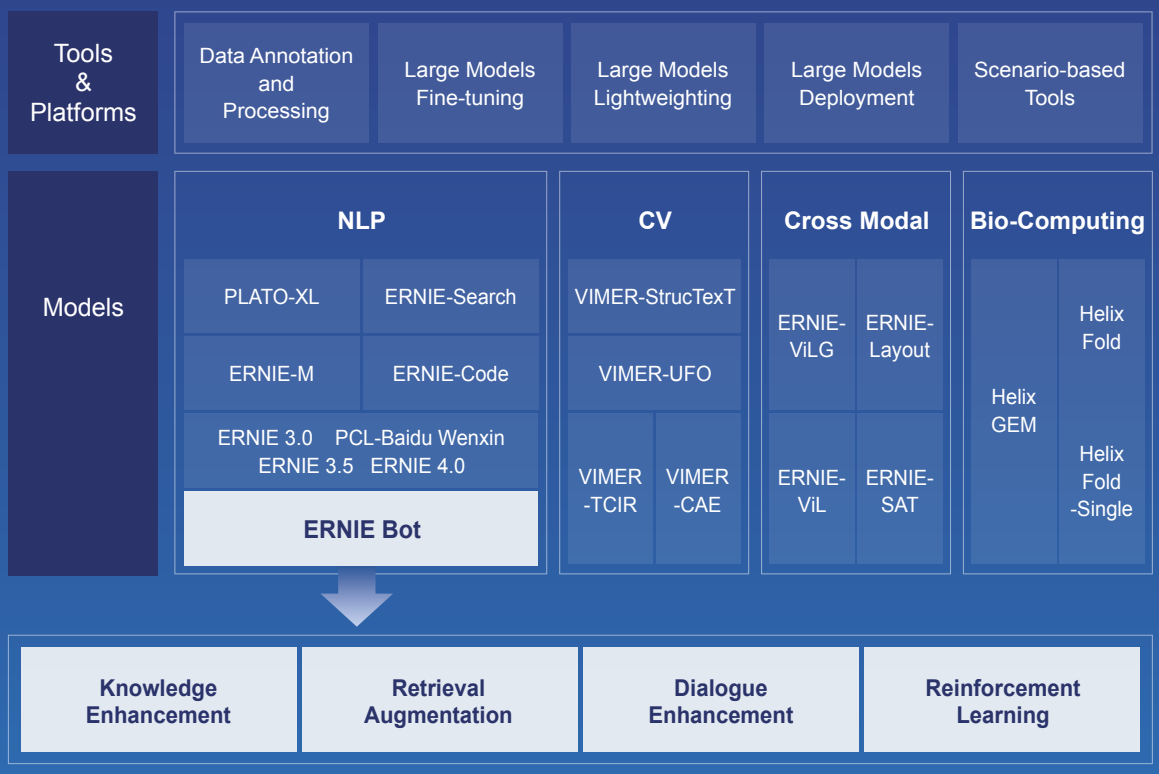
Baidu has long been at the forefront of deep learning research. Its PaddlePaddle, open-sourced in 2016, is the first open-source deep learning framework in China. Its ERNIE 1.0, released in 2019, marks a significant step forward in LLM development. Based on these technological advancements, its AI chatbot "ERNIE Bot" was developed and publicly launched in 2023. Baidu's consistent commitment to AI has led to a thriving ecosystem, with ERNIE Bot reaching 200 million users and the ERNIE API handling 200 million daily queries.

文心大模型

BAIDU WENXIN

A series of knowledge-enhanced large models, from general-purpose to industry-specific, independently developed by Baidu, speed up AI innovations and empower the industry to upgrade.

ERNIE Knowledge - Enhanced Large Models



Gold Sponsor

Handwriting Analysis with Wacom Ink Technologies

Wacom Ink Technologies transform digital ink into a powerful medium for content creation, conversion, and derivation, facilitating knowledge structuring and enhancing interactive experiences. The underlying Universal Ink Model is a language-neutral, hardware- and platform-independent data model for representing and manipulating digital ink data captured with an electronic pen, stylus, or touch input. In addition, Semantic Ink technology stack leverages artificial intelligence (AI) to understand and interpret ink strokes, converting them into machine-understandable data, making it a valuable technology for various industries needing efficient content organization, exploration and searchability.
Wacom offers a wide range of Software Development Kits (SDKs) and Cloud Services built on Wacom Ink Technologies, providing a versatile toolkit for developers to enhance signature tools and handwriting software. These software tools empower developers to integrate advanced features for digital pen-based data capturing and customize applications to meet specific needs and workflows.

About Wacom

Wacom is a global leader in digital pen and touch input technology, providing innovative solutions for businesses and a broad range of professional users. Our cutting-edge pen tablets, displays, and styluses offer precision and ergonomic design, seamlessly integrating handwriting, signatures and artistic techniques into the digital workspace. Partner with Wacom to unlock transformative business, academic and creative solutions, leveraging our industry-leading technology to enhance productivity and redefine possibilities in document analysis, technology and education. For more information go to ***developer.wacom.com***

wacom

wacom.com

Digitalize and analyze your handwriting

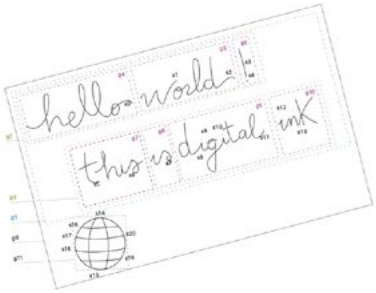
Digital paper and digital pens break through the wall of paper digitization and promote the digitalization of education.

- Precision and Accuracy by capturing detailed handwriting input
- Natural Writing Experience helps to collect authentic handwriting samples
- Advanced features enabling comprehensive data on handwriting dynamics
- Compatibility and Integration allows for seamless use in diverse research projects and classroom settings

Wacom One Family

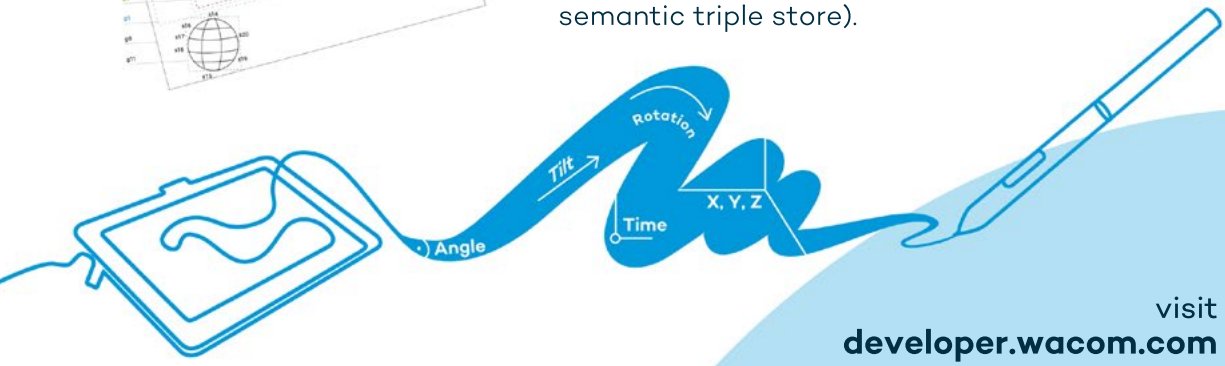
Wacom Movink

IoT Paper*



Wacom Ink Technologies transform digital ink into a powerful medium for content creation and interaction.

With Wacom's Universal Ink Model you have access to digital ink data (input data, ink data, metadata, ink trees and views, and semantic triple store).



visit
developer.wacom.com
to learn more

*Research Prototype Product

Silver Sponsor

With a team of world-class research scientists, engineers, artists, and designers, Adobe Research combines cutting-edge academic discovery with industry impact. Our researchers shape early-stage ideas into innovative technologies. Many of Adobe Research's breakthroughs are incorporated into Adobe's products, building the company's reputation as a leader in fostering new forms of creativity and in advancing document and content intelligence. Our team collaborates with interns and faculty from universities across the globe. Visit us ***online***, on ***X*** and on ***Facebook.***

For future reference, the language is here: ***https://research.adobe.com/about-adobe-research/***

ICDAR Athens Greece 2024

International Conference on Document Analysis and Recognition

icdar2024.net