

Minesh Mathew

<https://mineshmathew.netlify.app/> | [publications](#) |
minesh.mathew@gmail.com

Summary

Presently, I work as a senior machine learning scientist at Wadhvani AI where I lead ML teams solving problems in public health space using computer vision, NLP and document image analysis. My PhD research dealt with the problem of Question Answering on document images.

Work experience

[Senior] Machine learning scientist - Wadhvani AI April 2023 - present

I lead ML teams for multiple projects in public health, agriculture and education. My job is to coordinate the ML work, mentor the associate ML scientists and architect ML solutions. Additionally, I work with the partnerships team to look into new use cases, study the feasibility from ML perspective and help with writing proposals for grants.

Graduate research assistant - CVIT, IIIT Hyderabad Jan 2016 - April 2023

Coordinating document understanding projects (printed, handwritten and scene text OCR and text-based VQA works) and mentoring Master's and undergraduate students and interns.

Trainee Engineer - CVIT, IIIT Hyderabad Apr 2013 - Jan 2016

Worked as part of a government of India project to build OCRs for Indian languages. My responsibilities included defining annotation requirements, coordinating annotation, development of an RNN+CTC based recognition model and leading a project where the OCRs were used for making DAISY format audio books

Skill development Executive - ASAP Kerala Jan 2011 - Apr 2011

ASAP is an initiative by the Government of Kerala to increase the employability of Higher Secondary and undergrad students in government schools/colleges. My job was to help the students improve their basic communication and IT skills, and to equip them for the next levels in the training

Business Analyst - S&P Global Market Intelligence (previously -S&P Capital IQ) Jun 2009 - Jun 2011

I worked as part of the Financial Estimates team. We were responsible for building and maintaining tools for collection of data from financial estimates. My work included collecting requirements from the clients and translating it into software design documents and most importantly to work as a liaison between the tech team and the clients/users.

Career break

Indian civil service exam attempt Jun 2011 - Jan 2013

I spent nearly 18 months preparing and appearing for Indian civil service examination. I appeared for the prelims and mains exams in the year 2012.

Education

2023 PhD Computer Science (IIIT Hyderabad, advisor: Prof. CV Jawahar)
2016 MS - Computer Science (IIIT Hyderabad, advisor: Prof. CV Jawahar)
2009 Btech - Computer Science and Engineering (NIT Warangal)

Achievements and fellowships

2025 CVPR 2025 outstanding reviewer
2023 CVPR 2023 outstanding reviewer
2021 ICCV 2021 outstanding reviewer
2021 Selected for WACV doctoral consortium
2021 Selected for ICCV doctoral consortium
2017 Best Paper; Int'l Workshop on Arabic Script Analysis and Recognition (ASAR)
2016 Awarded TCS PhD fellowship
2015 Runner up - Microsoft Azure ML Hackathon held at IIIT Hyderabad

Professional/academic activities

- Reviewer for conferences — ECCV ‘22, SIGGRAPH ‘22, CVPR ‘22, ‘23 and ‘24, WACV ‘21 and ‘22, ICDAR ‘21 and ICCV ‘21 and 23, and ACCV ‘24
- Reviewer for journals — TMLR, Pattern Recognition, IEEE TNNLS, TPAMI, IJCV, IJDAR, Visual Computer, Concurrency n’ comput., and Engineering science and tech.
- 2024 Challenge chair, National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics (NCVPRIPG) 2024.
- 2024 Organizer, RRC challenge on occluded scene text.
- 2023 Organizer, Competition on Text-based Video Question Answering on News Videos, ICDAR 2023
- 2023 Organizer, Competition on RoadText Video Text Detection, Tracking and Recognition, ICDAR 2023
- 2021 Organizer, Document Visual Question Answering Workshop and challenge ICDAR 2021
- 2020 Organizer and competition chair, Text and documents workshop and DocVQA challenge, CVPR 2020
- 2019 Organizer, Scene text VQA challenge , ICDAR 2019
- 2016-18 Organizer, Computer vision summer schools at IIIT Hyderabad

Talks and presentations

- December 2024 - Invited talk - 'Using AI responsibly for social good', at Advanced Computing and Communications (ADCOM) conference, IIIT Bangalore
- August 2024 - Invited talk at SOSCON (Samsung Open Source Conference) India 2024.
- July 2024 - Tutorial on 'LLMs for information extraction from visually rich PDFs and document images', NCVPRIPG 2024, Thiruvananthapuram.
- May 2024 - AI readiness workshop at Piramal workshop, Guwahati.
- Nov 2023 - AI-readiness workshop at IIPH Gandhinagar
- 2022 - *DocVQA*, Amex AI-CVIT workshop
- 2017 - *Sequence learning using RNNs*, Computer vision summer school, IIIT Hyderabad
- 2017 - Text recognition for Indian languages. CIE, IIIT Hyderabad startups meet
- 2016 - *Indian languages audio books using OCR and TTS*, 'Technology for the visually impaired' seminar at University of Hyderabad
- 2013 - *Audio books playback app demonstration* at various schools and State Council of Education Research and Training(SCERT) Kerala.

Projects

Past projects at Wadhvani AI

1. **Computer vision for automating digitization of treatment adherence.** This work involved use of table detection and table understanding to digitize data in a handwritten table indicating a patient’s medication adherence. Table detection, table understanding and symbol classifier models are used.
2. **Development of essential language technologies for a low resource Himalayan language.** We worked on development of OCR of manuscripts (woodblock and the ancient ones), speech-to-text and machine translation. The larger objective was to use these tools for the preservation of cultural artefacts of the community.
3. **Epidemiological and agriculture related events extraction from online media.** Built a system that scans the internet for news talking about events of interest to public health and agriculture and extract structured data (place, type of incident, severity etc.)from the articles. We use BERT for text classification and NER, LLMs (Llama, Gemma and OpenAI models) for structured data extraction and text embeddings for de-duplication.

Current projects at Wadhvani AI

1. **Oral reading fluency assessment and remediation (ORF)** To assess reading fluency of children using ASR and report relevant metrics to the school teachers.

Skills

ML/CV libraries Pytorch, OpenCV, Numpy

References

Available on request.

Publications

Total citations: 3519 ; h-index=20; i10-index=19 source: Google Scholar as on May 29, 2026

Journals

1. **Minesh Mathew**, Lluís Gomez, Dimosthenis Karatzas, CV Jawahar, *Asking questions on Handwritten Document Collections* - Int'l journal on Document Analysis and Recognition (IJ DAR); ICDAR-IJ DAR special issue 2021 [Springer preprint](#)

Conference/workshop publications

1. **Minesh Mathew**, Ajoy Mondal and CV Jawahar, *Towards Deployable OCR Models for Indic Languages* - ICPR 2024
2. Soumya Jahagirdar, **Minesh Mathew**, Dimosthenis Karatzas and CV Jawahar, *Watching the News: Towards VideoQA Models that can Read* - WACV 2023 [link](#)
3. Sergi Garcia-Bordils, George Tom, Sangeeth Reddy, **Minesh Mathew**, Marçal Rusiñol, CV Jawahar and Dimosthenis Karatzas, *Read While You Drive-Multilingual Text Tracking on the Road* - DAS 2022 [link](#)
4. **Minesh Mathew**, Viraj Bagal, Ruben Perez Tito, D Karatzas, Ernest Valveny and CV Jawahar, *Infographic VQA* - WACV 2022 [link](#)
5. **Minesh Mathew**, D Karatzas and CV Jawahar, *DocVQA: A Dataset for VQA on Document Images* - WACV 2021 [link](#)
6. Ruben Tito*, **Minesh Mathew***, CV Jawahar, Ernest Valveny and Dimosthenis Karatzas, *ICDAR 2021 Competition on Document Visual Question Answering* - ICDAR 2021 (Competition report) [link](#)
7. Yash Khare, Viraj Bagal, **Minesh Mathew**, A Devi, UD Priyakumar and CV Jawahar, *MMBERT: Multimodal BERT Pretraining for Improved Medical VQA* - ISBI 2021 [link](#)
8. **Minesh Mathew**, Ruben Tito, Dimosthenis Karatzas and R Manmatha and CV Jawahar, *DocVQA Challenge 2020* - DAS 2020 (short paper) [link](#)
9. Sangeeth Reddy, **Minesh Mathew**, Lluís Gomez, Marçal Rusiñol, Dimosthenis Karatzas and CV Jawahar, *RoadText-1K: Text Detection & Recognition Dataset for Driving Videos* - ICRA 2020 [link](#)
10. Ali Furkan Biten, Ruben Tito, Andres Mafía, Lluís Gomez, Marçal Rusiñol, **Minesh Mathew**, CV Jawahar, Ernest Valveny and Dimosthenis Karatzas, *ICDAR 2019 competition on scene text visual question answering* - ICDAR 2019 (competition report) [link](#)
11. Deepayan Das, Jerin Philip, **Minesh Mathew** and CV Jawahar, *A Cost Efficient Approach to Correct OCR Errors in Large Document Collections* - ICDAR 2019 [link](#)
12. Kartik Dutta, **Minesh Mathew**, Praveen Krishnan and CV Jawahar, *Localizing and Recognizing Text in Lecture Videos* - ICFHR 2018 [link](#)
13. Kartik Dutta, Praveen Krishnan, **Minesh Mathew** and CV Jawahar, *Improving cnn-rnn hybrid networks for handwriting recognition* - ICFHR 2018 [link](#)
14. Kartik Dutta, Praveen Krishnan, **Minesh Mathew** and CV Jawahar, *Towards Spotting and Recognition of Handwritten Words in Indic Scripts* - ICFHR 2018 [link](#)
15. Kartik Dutta, Praveen Krishnan, **Minesh Mathew** and CV Jawahar, *Offline Handwriting Recognition on Devanagari Using a New Benchmark Dataset* - DAS 2018 [link](#)
16. **Minesh Mathew**, Mohit Jain and CV Jawahar, *Benchmarking scene text recognition in Devanagari, Telugu and Malayalam* - MOCR workshop, ICDAR 2017 [link](#)
17. Mohit Jain, **Minesh Mathew** and CV Jawahar, *Unconstrained scene text and video text recognition for Arabic script* - ASAR workshop 2017 [link](#)

18. Mohit Jain, **Minesh Mathew** and CV Jawahar, *Unconstrained ocr for urdu using deep cnn-rnn hybrid networks* - ACPR 2017 [link](#)
19. VS Vinita, **Minesh Mathew** and CV Jawahar, *An Empirical Study of Effectiveness of Post-processing in Indic Scripts* - MOCR workshop ICDAR 2017 [link](#)
20. Kartik Dutta, Praveen Krishnan, **Minesh Mathew** and CV Jawahar, *Towards accurate handwritten word recognition for Hindi and Bangla* - NCVPRIPG 2017 [link](#)
21. **Minesh Mathew**, AK Singh and CV Jawahar, *Multilingual OCR for Indic Scripts* - DAS 2016 [link](#)