Minesh Mathew

🌐 mineshmathew.github.io | 🎓 publications |
📞 +91 9642445241 | ✉ mineshmathew@gmail.com

Summary

I am a Computer Vision researcher who has primarily been working on document image understanding problems. I have 9+ years of experience in using both classical and deep-learning based models for the problem of text recognition and higher level document understanding tasks. I am looking for research positions in document intelligence or vision and language space.

Work experience

Graduate research assistant - CVIT, IIIT Hyderabad   Jan 2016 - present

Coordinating document understanding projects 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 i) Coordinating annotations efforts and the data quality review process, ii) Building RNN based segmentation-free OCR pipeline for Malayalam OCR, iii) Maintaining a web based demo for RNN-based demo OCR service for 10+ Indian languages and iv) Leading an Audio books project that make use of Indian languages OCR and TTS for creating audio books in DAISY format

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 (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

2016 - present	PhD - Computer Science -	IIIT Hyderabad	(CGPA: 8.5)
2013 - 2015	MS - Computer Science -	IIIT Hyderabad	(CGPA: 8.25)
2005 - 2009	Btech - Computer Science and Engineering -	NIT Warangal	(CGPA: 6.7)

Achievements and fellowships

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 and '23, WACV '21 and '22, ICDAR '21 and ICCV '21

- Reviewer for journals — Pattern Recognition, IEEE TNNLS, TPAMI, IJCV, Visual Computer, Concurrency n' com-

put., and Engineering science and tech.

- 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

- 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.

Skills

Programming   Python, Matlab
ML/CV libraries   Pytorch, OpenCV, Numpy

References

Available on request.

Publications

Total citations: 582; h-index=14; i10-index=14   source: Google Scholar as on January 3, 2023

Journals

1. **Minesh Mathew**, CV Jawahar, *An empirical study of CTC based models for OCR of Indian languages* - [Under review] 2022   preprint

2. **Minesh Mathew**, Lluis Gomez, Dimosthenis Karatzas, CV Jawahar, *Asking quetions on Handwritten Document Collections* - Int'l journal on Document Analysis and Recognition (IJDAR); ICDAR-IJDAR special issue 2021   Springer preprint

Conference/workshop publications

1. Soumya Jahagirdar, **Minesh Mathew**, Dimosthenis Karatzas and CV Jawahar, *Watching the News: Towards VideoQA Models that can Read* - WACV 2023

2. 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

3. **Minesh Mathew**, Viraj Bagal, Ruben Perez Tito, D Karatzas, Ernest Valveny and CV Jawahar, *InfographicVQA* - WACV 2022   link

4. **Minesh Mathew**, D Karatzas and CV Jawahar, *DocVQA: A Dataset for VQA on Document Images* - WACV 2021   link

5. Ruben Tito* , **Minesh Mathew***, CV Jawahar, Ernest Valveny and Dimosthenis Karatzas, *ICDAR 2021 Competition on Document Visual Question Answering* - ICDAR 2021 (Competition report)   link

6. Yash Khare, Viraj Bagal, **Minesh Mathew**, A Devi, UD Priyakumar and CV Jawahar, *MMBERT: Multimodal BERT Pretraining for Improved Medical VQA* - ISBI 2021   link

7. **Minesh Mathew**, Ruben Tito , Dimosthenis Karatzas and R Manmatha and CV Jawahar, *DocVQA Challenge 2020* - DAS 2020 (short paper)   link

8. Sangeeth Reddy, **Minesh Mathew**, Lluis Gomez, Marcal Rusinol, Dimosthenis Karatzas and CV Jawahar, *RoadText-1K: Text Detection & Recognition Dataset for Driving Videos* - ICRA 2020   link

9. Ali Furkan Biten, Ruben Tito, Andres Mafla, Lluis Gomez, Marçal Rusinol, **Minesh Mathew**, CV Jawahar, Ernest Valveny and Dimosthenis Karatzas, *ICDAR 2019 competition on scene text visual question answering* - ICDAR 2019 (competition report)   link

10. Deepayan Das, Jerin Philip, **Minesh Mathew** and CV Jawahar, *A Cost Efficient Approach to Correct OCR Errors in Large Document Collections* - ICDAR 2019   link

11. Kartik Dutta, **Minesh Mathew**, Praveen Krishnan and CV Jawahar, *Localizing and Recognizing Text in Lecture Videos* - ICFHR 2018   link

12. Kartik Dutta, Praveen Krishnan, **Minesh Mathew** and CV Jawahar, *Improving cnn-rnn hybrid networks for handwriting recognition* - ICFHR 2018   link

13. Kartik Dutta, Praveen Krishnan, **Minesh Mathew** and CV Jawahar, *Towards Spotting and Recognition of Handwritten Words in Indic Scripts* - ICFHR 2018   link

14. Kartik Dutta, Praveen Krishnan, **Minesh Mathew** and CV Jawahar, *Offline Handwriting Recognition on Devanagari Using a New Benchmark Dataset* - DAS 2018   link

15. **Minesh Mathew**, Mohit Jain and CV Jawahar, *Benchmarking scene text recognition in Devanagari, Telugu and Malayalam* - MOCR workshop, ICDAR 2017   link

16. Mohit Jain, **Minesh Mathew** and CV Jawahar, *Unconstrained scene text and video text recognition for Arabic script* - ASAR workshop 2017   link

17. Mohit Jain, **Minesh Mathew** and CV Jawahar, *Unconstrained ocr for urdu using deep cnn-rnn hybrid networks* - ACPR 2017   link

18. VS Vinitha, **Minesh Mathew** and CV Jawahar, *An Empirical Study of Effectiveness of Post-processing in Indic Scripts* - MOCR workshop ICDAR 2017   link

19. Kartik Dutta, Praveen Krishnan, **Minesh Mathew** and CV Jawahar, *Towards accurate handwritten word recognition for Hindi and Bangla* - NCVPRIPG 2017   link

20. **Minesh Mathew**, AK Singh and CV Jawahar, *Multilingual OCR for Indic Scripts* - DAS 2016   link