

AYUSH KUMAR SHAH

5th year Ph.D. student in Computer Science

✉ as1211@rit.edu 📍 Rochester, New York ☎ (585) 471-9866 🐦 @ayushkumarshah
📧 @ayushkumarshah 📄 @ayush7 🖥 shahayush.com 🗨 Ayush Kumar Shah

EDUCATION

PhD in Computing and Information Sciences, CGPA: 3.93/4 *Aug 2020 – Present*
Rochester Institute of Technology (RIT) *Rochester, NY, USA*
Area of focus: extraction and visual parsing of graphical structures and notations from documents
Relevant Courses: Pattern Recognition, Computer Vision, Deep Learning Mathematics, NLP, Software Engineering.

Bachelors in Computer Engineering, CGPA: 3.96/4 *Aug 2015 – Oct 2019*
Kathmandu University *Kavre, Nepal*

PROFESSIONAL EXPERIENCE

Amazon - Alexa Speaker Understanding AI Sunnyvale, California
Applied Scientist Intern *May 2022 – Aug 2022*

- Applied generative AI and Large Language Models to enhance text-to-speech and synthetic speech generation.
- Improved speaker identification results in voice assistants like Alexa by reducing training time and annotation costs through semi-supervised learning, including clustering of audio embeddings for pseudo-label generation.

Fusemachines Kathmandu, Nepal
Machine Learning Engineer *June 2019 – Aug 2020*

- Optimized client's business decisions for chemical products that go unsold using boosting classifiers.
- Automated bank data extraction by building a 95% accurate handwritten text (English & Nepali) recognizer.

PUBLICATION

- A. K. Shah**, B. M. Amador, A. Dey, M. Creekmore, B. Ocampo, S. Denmark, and R. Zanibbi, "ChemScraper: Leveraging PDF Graphics Instructions for Molecular Diagram Parsing," in Document Analysis and Recognition (Journal) - **IJDAR** 2024, vol. 27, Sep. 2024, pp. 395-414, doi: 10.1007/s10032-024-00486-7.
- A. K. Shah**, and R. Zanibbi, "Line-of-Sight with Graph Attention Parser (LGAP) for Math Formulas," in Document Analysis and Recognition - **ICDAR** 2023, Cham: 2023, pp. 401-419, doi: 10.1007/978-3-031-41734-4_25.
- B. M. Amador, M. Langenkamp, A. Dey, **A. K. Shah**, and R. Zanibbi. "Searching the ACL Anthology with Math Formulas and Text" in Proceedings of the 46th International ACM **SIGIR** Conference on Research and Development in Information Retrieval, in **SIGIR '23**. ACM 2023, Jul. 2023, pp. 3110-3114, doi: 10.1145/3539618.3591803
- A. K. Shah**, A. Dey, and R. Zanibbi, "A Math Formula Extraction and Evaluation Framework for PDF Documents," in Document Analysis and Recognition - **ICDAR** 2021, Cham, 2021, pp. 19-34, doi: 10.1007/978-3-030-86331-9_2

RESEARCH EXPERIENCE

17th International Conference on Document Analysis and Recognition San José, California
Program Committee (PC) Member *2023*

- Reviewed and evaluated five research paper submissions, and provided feedback and recommendations to authors.

Document and Pattern Recognition Lab (DPRL), RIT Rochester, New York
Graduate Research Assistant *Aug 2020 – Present*

- Leveraged Large Language Models (LLMs) and generative AI to fine-tune mathematical and chemical formula recognition models, achieving a 10% increase in recognition accuracy.
- Developed a fast and accurate molecular diagrams parser, with automated annotated data generation for training visual chemical parsers, and novel graph-based evaluation metrics and error analysis tools.
- Enhanced accessibility of mathematical information through a documents search system within the ACL Anthology, integrating both text and mathematical formulas search for users with context-aware word and formula matching.

- Improved expression recognition rate of math formulas by 15% using improved attention and context features using modified graph attention network (GAT) and spatial pyramidal pooling.
- Accelerated math formula recognition by 6 times by implementing a custom dataloader with dynamic batch size for full GPU utilization in a distributed parallelization framework.
- Aided the document recognition community by introducing a valuable open-source visualization tool, facilitating the evaluation of graphical recognition results and the identification of specific errors within documents in context.

Research Interests: Pattern recognition, recognition of graphical structures, computer vision, speaker understanding, large language models, multi-modal deep learning, natural language processing

HONORS AND AWARDS

- RIT Ph.D. Merit Scholarship/Assistantship.** Financial Support for Ph.D. at RIT, which includes support via NSF Grants. *2020 – Present*
- Kathmandu University Merit-based scholarship (4x).** \$440 worth scholarship awarded for securing the highest GPA in the Computer Engineering cohort (4/7 semesters). *2015 – 2019*
- Fusemachines Artificial Intelligence Scholarship Program.** Selected among thousands of candidates nationwide for fuse.ai Artificial Intelligence Scholarship Online Course. *Nov 2018*
- American Society of Nepalese Engineers Merit Award.** A merit worth \$200, rewarded to the entrance topper of each university in Nepal, seeking admission for undergraduate degrees. *May 2016*
- 46th International Physics Olympiad (IPhO) Contestant.** One of the largest olympiads for high school Physics enthusiasts with 5 contestants, each from 100 participating countries. *June 2015*

TEACHING EXPERIENCE

- Rochester Institute of Technology** Rochester, New York
Graduate Teaching Assistant *Aug 2022 – Dec 2022*
- Course: CSCI 335: Machine Learning
- Samriddhi College** Kathmandu, Nepal
Computer Science Instructor *Jan 2020 – June 2020*
- Course: “Foundations in AI: Computer Science and Mathematics”

TECHNICAL SKILLS

- | | |
|------------------------------|--|
| Programming Languages | Python, R, Matlab, C, C++, JAVA |
| Python Packages | Pytorch, Tensorflow, Scikit-Learn, OpenCV, Nltk, Pandas, Numpy, Matplotlib, Fastapi, BeautifulSoup, Regex, NetworkX, Jupyter |
| Database | MySQL, MongoDB |
| Miscellaneous | Git, Github, Bash, L ^A T _E X, Jira, Linux, Arduino, Raspberry-pi |

TALKS

- Oral presentation** on “ChemScraper: Leveraging PDF Graphics Instructions for Molecular Diagram Parsing” at the 18th International Conference on Document Analysis and Recognition ICDAR 2024, Athens, Greece. *Sept 3, 2024*
- Poster presentation** on “ChemScraper: Extracting Molecule Diagrams from PDF Vector and Raster Images with CDXML and SMILES Output” at the Molecule Maker Lab Institute (MMLI) All-Institute Retreat at **University of Illinois Urbana-Champaign (UIUC)**. *Sept 12, 2023*
- Research Idea Ring (RIR) talk** on “Line-of-sight with Graph Attention Parser (LGAP) for Math Formulas” at RIT. *April 17, 2023*
- Poster presentation** on “Reconstructing the Structure of Molecular Diagrams in PDF Documents using a CNN-Attention-Based Parsing Model” at the Molecule Maker Lab Institute (MMLI) All-Institute Retreat at **University of Illinois Urbana-Champaign (UIUC)**. *Sept 28, 2022*
- Guest lecture** on “Bayesian Decision Theory” for RIT’s undergraduate course - Intro to Machine Learning (40 students). *Sept 5, 2022*
- Research Idea Ring (RIR) talk** on “A Fast and Interpretable Context-aware Parser for Isolated Formulas and Chemical Diagrams” at RIT. *April 7, 2022*