AYUSH KUMAR SHAH

5^{th} year Ph.D. candidate 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*

- Improved speaker identification results in voice assistants like Alexa by reducing training time and annotation costs through semi-supervised learning.

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.
- Prepared course materials for Fusemachines AI Education Programs.

PUBLICATION

- **A. K. Shah**, et al., "Multimodal Search in Chemical Documents and Reactions", *accepted at* **SIGIR** Conference on Research and Development in Information Retrieval, in SIGIR '25.
- **A. K. Shah**, et al., "ChemScraper: Leveraging PDF Graphics Instructions for Molecular Diagram Parsing," in Document Analysis and Recognition - **IJDAR 2024**, vol. 27, 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**, doi: 10.1007/978-3-031-41734-4_25.
- B. M. Amador, M. Langsenkamp, 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**, doi: 10.1007/978-3-030-86331-9_2

RESEARCH EXPERIENCE

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

- Developed a fast and accurate parser for molecular diagrams, including automated generation of annotated data for training visual chemical parsers, along with novel graph-based evaluation metrics and error analysis tools.
- Improved access to mathematical content by designing a search system for the ACL Anthology that integrates both textual and mathematical formula search, using context-aware word and formula matching.
- Increased math formula recognition accuracy by 15% through enhancements in attention mechanisms and contextual features, using a modified Graph Attention Network (GAT) combined with spatial pyramidal pooling.
- Achieved a 6× speedup in math formula recognition by implementing a custom data loader with dynamic batch sizing, fully utilizing GPU resources in a distributed parallel training framework.

- Contributed to the document recognition community by developing an open-source visualization tool to support the evaluation of graphical recognition results and enable detailed, context-aware error identification.

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

REVIEW EXPERIENCE

- **Program Committee Member**, ICDAR 2023 (5 papers), ICDAR 2025 (6 papers)
- **Journal Reviewer**, Pattern Recognition (Elsevier), 4 manuscripts (2024–2025)

HONORS AND AWARDS

RIT Ph.D. Assistantship. Full funding via NSF-supported research projects. *2020 – 2025*

Kathmandu University Merit Scholarship (4x). Awarded $440 total for highest GPA in the Computer Engineering cohort across 4 of 7 semesters. *2015 – 2019*

Fusemachines AI Scholarship. Selected from a nationwide pool for the Fuse.ai Artificial Intelligence Scholarship Program. *Nov 2018*

American Society of Nepalese Engineers Merit Award. $200 award for top university entrance rank in Nepal. *May 2016*

46$^{\text{th}}$ International Physics Olympiad (IPhO) Contestant. Selected among Nepal's top 5 to compete internationally with participants from 100+ countries. *June 2015*

TEACHING EXPERIENCE

Graduate Teaching Assistant, RIT, Rochester, NY *Aug 2022 – Dec 2022*
Course: *CSCI 335: Machine Learning*

Instructor, Samriddhi College, Kathmandu, Nepal *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, LaTeX, Jira, Linux, Arduino, Raspberry-pi

TALKS

Poster presentation on "ChemScraper: Pipeline for Parsing Raster and Vector Molecule Diagrams from PDFs" at the at Molecule Maker Lab Institute (MMLI) Symposium 2025 at **University of Illinois Urbana-Champaign (UIUC)**. *April 15, 2025*

Oral presentation on "ChemScraper: Leveraging PDF Graphics Instructions for Molecular Diagram Parsing" at the 18^{th} 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*