

# AYUSH KUMAR SHAH

5<sup>th</sup> 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

Rochester Institute of Technology (RIT)

Aug 2020 – Present

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

Kathmandu University

Aug 2015 – Oct 2019

Kavre, Nepal

**Relevant Courses:** Artificial Intelligence, Data Structures and Algorithms, Algorithm and Complexity, Software Engineering, Probability and Statistics, Machine Learning, Speech and Language Processing, C, C++.

## PROFESSIONAL EXPERIENCE

**Amazon - Alexa Speaker Understanding AI**

*Applied Scientist Intern*

Sunnyvale, California

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**

*Machine Learning Engineer*

Kathmandu, Nepal

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 Fusemachines AI Education Programs course materials for AI Democratization.

## 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. 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, Cham, 2021, pp. 19-34, doi: 10.1007/978-3-030-86331-9\_2

## RESEARCH EXPERIENCE

**17th International Conference on Document Analysis and Recognition**

*Program Committee (PC) Member*

San José, California

2023

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

**Document and Pattern Recognition Lab (DPRL), RIT**

*Graduate Research Assistant*

Rochester, New York

Aug 2020 – Present

- 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

---

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

## TEACHING EXPERIENCE

---

<b>Rochester Institute of Technology</b> <i>Graduate Teaching Assistant</i>	Rochester, New York Aug 2022 – Dec 2022
<ul style="list-style-type: none"> <li>• Course: CSCI 335: Machine Learning</li> </ul>	
<b>Samriddhi College</b> <i>Computer Science Instructor</i>	Kathmandu, Nepal Jan 2020 – June 2020
<ul style="list-style-type: none"> <li>• Course: “Foundations in AI: Computer Science and Mathematics”</li> </ul>	

## TECHNICAL SKILLS

---

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

## TALKS

---

<b>Oral presentation</b> on “ChemScraper: Leveraging PDF Graphics Instructions for Molecular Diagram Parsing” at the 18 <sup>th</sup> International Conference on Document Analysis and Recognition ICDAR 2024, Athens, Greece.	Sept 3, 2024
<b>Poster presentation</b> 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 <b>University of Illinois Urbana-Champaign (UIUC)</b> .	Sept 12, 2023
<b>Research Idea Ring (RIR) talk</b> on “Line-of-sight with Graph Attention Parser (LGAP) for Math Formulas” at RIT.	April 17, 2023
<b>Poster presentation</b> 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 <b>University of Illinois Urbana-Champaign (UIUC)</b> .	Sept 28, 2022
<b>Guest lecture</b> on “Bayesian Decision Theory” for RIT’s undergraduate course - Intro to Machine Learning (40 students).	Sept 5, 2022
<b>Research Idea Ring (RIR) talk</b> on “A Fast and Interpretable Context-aware Parser for Isolated Formulas and Chemical Diagrams” at RIT.	April 7, 2022