Simmi Mourya

linkedin.com/in/simmi-mourya

simmimourya.github.io | *Google Scholar* | **simmimourya@gmail.com**

Seattle, WA 98121 | Mobile : +1-267-968-8186

Education

- **University of Pennsylvania** — Philadelphia, PA
  *Master of Science in Computer and Information Science, GPA: **3.81/4.0*** — *Aug 2019 - May 2021*
  - **Coursework**: OS, Networked Systems, Analysis of Algorithms, Internet & Web Systems, Advanced Machine Perception, Comp. Linguistics, Computer Vision, Machine Learning. **Teaching Assistant:** CIS-581: Computer Vision
- **Cluster Innovation Center, University of Delhi,** — New Delhi, India
  *Bachelor of Technology in Computer Science and Applied Mathematics, GPA: **8.2/10*** — *Aug. 2013 - July 2017*
  - **Minor**: Computational Biology

Experience

- **Moderna Inc.** — Seattle, WA
  Machine Learning Engineer II, ML Infrastructure team — *Present*
  - Designing and developing the compute platform that supports **Large Language Model** and **Computational Sciences** teams to run various workflows (pipeline, high-performance compute and distributed model training) on scale. This platform enables in-house scientists to perform **vital validation testing** for **vaccine development** *(AWS, Python, TypeScript).*
  - Created tools for multiple organizations to leverage Large Language Models in order to enhance their workflows, integrating custom plugins including **retrieval, OCR, data analysis, and dynamic API routing**.
  - **Enhanced model performance** significantly by designing an automated evaluation for retrieval tasks, utilizing LLMs for generating a diverse set of test cases, thereby effectively doubling the test creation process.
- **Amazon, Packaging Innovation** — Seattle, WA
  Software Developent Engineer — *Jul 2021 - March 2023*
  - Implemented, tested, and deployed a packaging automation feature to enable shipping packages in vendor provided containers. The initiative is projected to generate Amazon **$280 million** in annualized savings towards packaging, transportation and labor cost in **16** fulfillment centers across **NA & EU**. *(Java)*
  - Added functionality to create, clone and monitor custom SageMaker jobs in a React based image labeling platform that handles **100,000 images** worth of traffic **per day**. *(Typescript, React)*
- **University of Pennsylvania** — Philadelphia, PA
  Graduate Research Assistant — *May 2020 - Dec 2020*
  - **Multimodal Question Answering framework**: Worked on **intersection of NLP and Computer Vision**. Developed a novel task framework for Goal-Step inference, Step membership inference using multimodal Wikihow data. *(PyTorch)*
- **ESRI** — New Delhi, India
  Software Developer — *May 2019 - July 2019*
  - **ArcGIS Python API**: Developed framework for Multispectral (**near infrared**) support for Pixel classification in **ArcGIS** Python API. Developed Pyramid scene parsing backbone support of **object segmentation** for the API. *(PyTorch)*
  - **Spatial Dataframes**: Optimized validation checks in arcgis.geometry package using pre-compiled Cython binaries. This processes **0.1 million** entries in less than **2 ms**, which earlier took **45-55 ms**. *(Python, Cython)*
- **IIIT Delhi** — New Delhi, India
  *Research Software Engineer, **Full-time*** — *Feb 2018 - March 2019*
  - **Article**: Mourya, S., Kant, S., Kumar, P., Gupta, A. and Gupta, R., 2018. LeukoNet: DCT-based **CNN** architecture for the classification of **normal versus Leukemic** blasts in B-ALL Cancer.
  - **Accepted Challenge**: Classification of Normal versus Malignant Cells in B-ALL White Blood Cancer Microscopic Images, challenge selected at IEEE ISBI '19, Venice, Italy. *(Python, PyTorch)*
- **Predible Health** — Bangalore, India
  Software Developer — *August 2017 - December 2017*
  - **Development**: Developed **U-Net** based framework for Lung nodule segmentation from 3D CT scans. Also developed classifiers to analyze nodule level malignancy and emphysema. Built POC for identifying cancerous lung nodules from Radiomics data. Streamlined prototyping and testing via parallelization of the data pre-processing pipeline. *(Python)*
- **Google Summer of Code** — Portland State University
  Software Developer Intern — *May 2016 - August 2016*
  - **Cyvlfeat**: Designed and developed 12 new features for a **high-performance** Python/Cython wrapper of a computer vision library, VLFeat. Emulated the wrapper from **MATLAB MEX scripts**.

Software Projects

- **Penn OS**: Built a user-level UNIX-like operating system consisting of a kernel, scheduler, FAT based file-system and shell. Built a SIGALARM - based priority scheduler for context switching (using ucontext library) and a Shell with job control, stdin/out redirection and other builtins like sleep, kill, ps etc. Also handled Shell's integration with kernel. *(C)*
- **Computer Vision**: Built an **attention mechanism** in form of **Region Proposal network** (RPN) as a backbone for **Mask RCNN**. Implemented **vectorized ROIAlign** for FPN-ROI Mapping. Developed **YOLO** (end-to-end) for **object detection**, with a Non Maximum Suppression post-processing module. **NLP:** Developed **Bilingual Named Entity Recognition** module using Bi-LSTM CRF and Self Attention.