Simmi Mourya

linkedin.com/in/simmi-mourya-34406886 | simmimourya.github.io | **simmim@seas.upenn.edu**
Philadelphia, PA 19104

Education

- **University of Pennsylvania** — Philadelphia, PA
  Master of Science in Computer and Information Science, GPA: ***3.74/4.0*** — ***Graduating May 2021***
- **Coursework**: Internet & Web Systems, Advanced Machine Perception, Comp. Linguistics, Computer Vision, Machine Learning
  Independent Research: *Video Object Segmentation* (advised by **Prof. Jianbo Shi**)
  Teaching Assistant: Computational Linguusitics, C++ Programming
- **Cluster Innovation Center, University of Delhi** — Delhi, India
  Bachelor of Technology in Information Technology, GPA: ***8.2/10*** — *Aug. 2013 – July 2017*

Skills

- **Research**: Python, PyTorch, FastAI, Keras, Scikit-Learn, Numpy, Pandas, Caffe, Cython, Python/C API
  Software: Java, Apache Spark, Apache Storm, Apachebench, Oracle BDB, PHP, HTML/CSS, SQL, Nose, Jenkins, ArcGIS

Experience

- **University of Pennsylvania** — Philadelphia, PA
  Graduate Research Assistant — *May 2020 - Present*
  - **Multimodal Question Answering framework**: Advised by Prof. Chris Callison Burch. Developing a novel task framework for Goal-Step inference and Step membership inference using multimodal Wikihow data.
- **ESRI** — Delhi, India
  Data Scientist — *May 2019 - July 2019*
  - **ArcGIS Python API**: Developed framework for **Multispectral support for Pixel classification** in **ArcGIS** Python API. This achieved segmentation improvements on **near infrared** imagery of Delaware county. Developed Pyramid scene parsing backbone support of object segmentation for the API using PyTorch and FastAI.
  - **Spatial Dataframes**: Optimized validation checks in arcgis.geometry package using pre-compiled Cython binaries. Now processes **0.1 million entries in less than 2 ms**, which earlier took **45-55 ms**.
- **IIIT Delhi** — New Delhi, India
  Research Associate — *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.
- **Predible Health** — Bangalore, India
  Deep Learning Developer — *August 2017 - December 2017*
  - **Development**: Developed U-Net based framework for **Lung nodule segmentation** from 3D CT scans (LIDC-IDRI dataset) using PyTorch and Python Scientific Stack. 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 (patch extraction and clean-up from CT scans).
- **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 computer vision library, VLFeat. (Added algorithms specializing in image understanding and local features extraction and matching such as LBP, SIFT, hierarchical k-means, SLIC). Built **unit and integration tests** using Python's Nose test suite.

Research Projects

- **Computer Vision**: Built an **attention mechanism** in form of **Region Proposal network** (RPN) as a backbone for **MaskRCNN**. Implemented **vectorized ROIAlign** for FPN-ROI Mapping.
  Developed **YOLO** (end-to-end) for **object detection**, with a **Non Maximum Suppression** post-processing module. Built a semi-automated **optical flow based tracker** for videos.
- **NLP**: Developed a **Bilingual Named Entity Recognition** using Bi-LSTM CRF, Self Attention.
- **Learning Visual control for Car Racing**: Implemented a **Fully connected Deep Q-network** and achieved an average reward of 210.92 for 10 evaluation steps. The best performing model had 70,475 parameters and trained for **only 570 episodes**.

Software Projects

- **Multi-threaded web server and Service framework**: A Java based web **HTTP 1.1 compliant web server** developed with **custom** implementations of underlying Blocking Queue and Thread Pool. Later merged it with a custom-built web service framework which emulates the behaviour of **Java Spark**.
- **Web crawler and Search Engine**:
  Developed a **multithreaded web crawler** with a custom XPath Parser and to query and store matched HTML, XML documents into a persistent data store. Also developed a **Map-Reduced Based Indexer.** Responsibilities: Developing and scaling Hadoop based Indexer. DevOps for Gradle, EMR, Hadoop, EMRFS. Minor Hadoop DevOps for PageRank.