Ansh Sharma

609-375-5451 | anshsharma@gmail.com | linkedin.com/in/anshgs | anshgs.com | github.com/anshgs

Education

University of Illinois at Urbana-Champaign, GPA: 4.0/4.0 May 2024
B.S. in Computer Science, Chancellor's Scholar *Champaign, IL*

Relevant Coursework: Computer Vision (Graduate), Learning to Learn (Graduate), Reinforcement Learning, Deep Learning, Machine Learning, Bioinformatics, Real Analysis, Systems Programming, Data Structures, Algorithms

Awards: John R. Pasta Outstanding Undergraduate Award (2023), Regeneron International Science and Engineering Fair (ISEF) Finalist (2021), USA Physics Olympiad Top 50 (2021), Putnam Top 500 (2020), USAJMO Qualifier (2019)

Publications

Ansh Sharma*, Albert Xiao*, Praneet Rathi, Rohit Kundu, Albert J. Zhai, Yuan Shen, Shenlong Wang, *EarthGen: Generating the World from Top-Down Views*, In Review

Andy Zhou*, Samuel Li*, Pranav Sriram*, Xiang Li*, Jiahua Dong*, **Ansh Sharma**, Yuanyi Zhong, Shirui Luo, Maria Jaromin, Volodymyr Kindratenko, Joerg Heintz, Christopher Zallek, Yuxiong Wang, *YoutubePD: A Multimodal Benchmark for Parkinson's Disease Analysis*, NeurIPS 2023 Datasets and Benchmarks Track

Ansh Sharma*, Keerthana Nallamotu*, Mukhil Shankar, Shenlong Wang, Inki Kim, *Deep-Learning Enabled Assessment of Neurocognitive Performance in Object Following in Mixed Reality*, IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE) 2022 Workshop on XR+AI

Snehal Vadvalkar, **Ansh Sharma**, Xiaomeng Zhang, Kaneli Galiotos, Bradley Hooker, Dustin Wooten, Varsha Mohan, Quaisar Ali, Ana Basso, Eric Mohler, Abhishek Pandey, *AI based imaging biomarker development in ADPKD mouse model*, World Molecular Imaging Congress (WMIC) 2022

Research Experience

Cascaded Generative Modeling for Remote Sensing
Research Intern: Computer Vision @ UIUC (Shenlong Group) *May 2023 – Present*
Supervisor: **Shenlong Wang**, University of Illinois Urbana-Champaign

- Designing generative models that can synthesize cohesive planet-scale maps at resolutions ranging from continents to houses
- Compiled a multi-terabyte scale dataset of satellite and aerial image pyramids enabling 1024x super-resolution
- Fine tuned cascaded diffusion models on super-resolution and tiling tasks for synced multi-layer map tile generation

Early Detection & Prediction of Parkinsonism Using Multi-Modal Few-Shot Learning
Research Intern: Computer Vision @ UIUC (Yuxiong Lab) *Oct. 2022 – Present*
Supervisor: **Yuxiong Wang**, University of Illinois Urbana-Champaign

- Utilized SOTA few-shot/meta-learning techniques to detect Parkinsonism from audio/visual modalities on a novel dataset
- Developed the first publicly accessible Parkinson's video dataset and validated generalizability to private medical datasets
- Implemented audio processing pipeline and created baselines through adapting Wave2Vec and Masked Autoencoder features

Optimizing AlphaFold on the Delta Supercomputing Cluster
Visiting Student: Argonne National Laboratory *Oct. 2022 – May 2023*
Supervisor: **Eliu Huerta**, Argonne National Laboratory / University of Chicago

- Containerized and deployed AlphaFold as a service for biomedical researchers on the NCSA Delta supercomputing cluster
- Created a Singularity container and a FuncX endpoint to allow for indirect access to run model through SLURM by users
- Improved inference pipeline runtime from 75 minutes to 20 minutes through CPU parallelization and multi-GPU processing

NeuroDNA/Flightpath: Concussion Diagnosis with Mixed Reality + Deep Learning
Research Intern: Jump ARCHES | Health Care Engineering Systems Center *Jun. 2022 – Sept. 2022*
Supervisor: **Shenlong Wang & Inki Kim**, University of Illinois Urbana-Champaign

- Researched methods to provide a standardized concussion diagnosis method using augmented reality and machine learning
- Designed and trained a transformer based deep learning architecture to learn from physician-annotated time series data
- Created visualizations with Open3D to model patient trajectories during testing and assist with generating annotations

Machine Learning Guided Directed Halogenase Evolution
Research Assistant: Molecule Maker Lab Institute *Dec. 2021 – May 2022*
Supervisor: **Haiyang Cui**, University of Illinois Urbana-Champaign

- Created a dataset for training molecule prediction tasks through data mining reaction information from research papers
- Used language models for biomedical entity relationship extraction to mine reaction information from 500 research papers
- Worked with computer vision models to detect and extract structural diagrams and convert into SMILES representations

Work Experience

Amazon New York City, NY
Software Development Engineer Intern (ML) - Amazon Translate *May 2023 - Aug. 2023*

- Evaluated and fine-tuned various large language models (LLMs) for automated translation quality evaluation tasks
- Reimplemented research papers and worked with applied scientists to evaluate techniques on internal LLMs and datasets
- Created an LLM-based scoring workflow in the translation model deployment pipeline to compute automated quality metrics

AbbVie Chicago, IL
Machine Learning Intern - Pharma R&D (AbbVie - Calico) *Aug. 2022 – May 2023*

- Researched representation learning models for cell painting using self supervised contrastive learning and arcface loss
- Implemented and trained a ResNet based architecture on terabytes of molecular perturbation data using PyTorch and AWS

Amazon Seattle, WA
Software Development Engineer Intern - AWS AppFabric *May 2022 - Aug. 2022*

- Developed and deployed a full-stack prototype for an upcoming greenfield AWS product using Java, TypeScript, and React
- Implemented internal authentication protocols and multi-origin CORS handling to allow call access to a Lambda API
- Designed and integrated an interactive front end into an existing AWS product to demonstrate functionality for stakeholders

AbbVie Chicago, IL
Machine Learning Intern - Pharma R&D *Jan. 2022 – May 2022*

- Transfer learned a 3D U-NET model to segment kidneys and cysts from MRI scans for tracking PKD progression in mice
- Utilized multiomic data (MRI scans, RNA-Seq, blood/urine biomarkers) to cluster and predict response to drug treatment
- Identified potential biological pathways for a novel treatment via gene ontology analysis, enabling further drug optimization

Selected Projects

Meta-Learning For Regression Via Data Re-Weighting | *PyTorch*

- Designed a data-reweighting algorithm to alleviate test-train imbalance for regression tasks using meta-learning techniques
- Tested algorithm on an age-detection task (UTKFace) for proof of concept and outperformed proposed baselines

Neural Music Transcription with Spatiotemporal Vision Models | *PyTorch, Librosa, Mido*

- Designed and implemented U-Net, CRNN, and Transformer based architectures to transcribe piano audio files into notes
- Evaluated qualitative/quantitative performance of each model with differently weighted losses to reduce data imbalance
- Built data processing pipelines to convert 200 Hours of audio data into mel-spectrograms to train and evaluate models

NeuroTech @ UIUC: Mind Controlled RC-Car | *Tensorflow, OpenBCI*

- Collected data and implemented ML models to control an RC Car using readings from a brain computer interface
- Trained an adaptation module which efficiently fine-tunes the model to a new user after two examples of each instruction
- Reached 87% live accuracy in classifying facial expressions as instructions using engineered features from the EEG data

Style Share | *React/Redux, three.js, Tensorflow.js, Flask, Firebase Firestore*

- A website that allows users to generate 3D scenes and stylize them according to the style of another image using ML
- Used quantization and distillation to reduce TF.js model size and improve speed by 4x while running within browser
- Included authentication with Google OAuth and a gallery to upload and share photos using a cloud storage bucket

Infected & Detected | *TFLite, OpenCV, Flask, MongoDB*

- **HackIllinois 2022: Best Community & Sustainability Track Project**
- An ML based edge computing tool to help farmers get analytics on their crop health and invasive weed growth over time
- Trained an image classifier using transfer learning on MobileNetV2 and pruning/quantization to fit it on a Rasberry Pi

Technical Skills

Proficient: Java, Python, C++, LaTeX, SQL, Linux Shell, Git, PyTorch, Pandas, Numpy, MatPlotLib, OpenCV, Qiskit
Familiar: JavaScript, HTML/CSS, Node.js, React, Express, Bootstrap, Sass, Catch2, Flask, OpenGL, Tensorflow, Keras