JUNWEN "JASON" WANG

jwang240@vt.edu | +1 312-792-4798 |
github | linkedin

Education

Ph.D in Physics — **Aug 2019 - Present**
Virginia Tech — *Blacksburg, VA*

Master of Engineering in Computer Science — **Aug 2022 - May 2024**
Virginia Tech — *Blacksburg, VA*

Skills

Programming Languages Python(Numpy, Pandas, scikit-learn) | C++ | Java | SQL | R | HTML/CSS | JavaScript
Skill Deep learning | TensorFlow | Pytorch | OpenMP | CUDA | Computer Vision | SLURM

Experience

Virginia Tech & Sandia National Laboratories — **Aug 2022 – Present**
Research Assistant — *Blacksburg, VA / Albuquerque, NM*

- Applied **machine learning models (NNPs, GNNs)** to improve force and energy predictions for complex molecular systems, achieving a **15-25%** increase in accuracy for rod-shaped particle interactions.
- Designed and implemented custom **geometric descriptors** and **feature extraction pipelines**, reducing feature dimensionality by **30%** while maintaining model accuracy, leading to **20%** faster training times on DFT-based datasets.
- Developed and optimized ML models using **TensorFlow** and **Python** libraries, integrating them into the LAMMPS simulation framework, which led to a **30%** improvement in computational efficiency.
- Processed and curated over **1,000 high-quality training samples** from quantum mechanical simulations, focusing on enhancing generalization to diverse configurations using techniques like **transfer learning** and **hybrid force fields**.

Virginia Tech(Video-Based Evaluation System for Surgical Procedures) — **July 2024 - Present**
Research Assistant — *Blacksburg, VA*

- Processed over **100 hours of surgical videos**, creating an annotation framework for identifying tools and predicting surgical steps, **improving model labeling efficiency by 30%**.
- Trained **computer vision models** to detect anatomical structures and surgical instruments in video frames, **achieving over 90% accuracy in instrument detection and 85% accuracy in surgical step prediction**.
- Applied advanced image processing techniques, including segmentation, tracking, and object detection, **reducing manual annotation time by 25%**.
- Leveraged libraries such as **OpenCV** and **vision transformers** for object detection and segmentation modeling, **improving object detection performance by 15%**.

Notable Projects

Marine Corps Community Services (MCCS) Dashboard — **Jan 2024 - June 2024**

- Designed and implemented the backend infrastructure using the **LAMP stack (Linux, Apache, MySQL, PHP) on AWS Lightsail**, resulting in a **25% reduction in deployment time** for the web application.
- Managed database configurations in **MySQL**, optimizing performance and storage capacity for handling over 1.07GB of data, leading to a **40% improvement in query response times**.
- Developed a responsive, user-friendly interface using **Bootstrap** and implemented **D3.js** for interactive data visualizations, allowing users to easily compare over **50 stores** and analyze vendor ratings.
- Contributed to the design and integration of **machine learning models**, achieving an **85% accuracy in forecasting inventory needs and detecting shrinkage**. Improved data quality through data cleaning techniques, driving data-driven decision-making and process improvements for MCCS.

Real Time Gender and Age Detection with OpenCV — **Aug 2023 - Dec 2023**

- Employed **OpenCV**'s Haar cascades and Dlib's HOG-based face detector, achieving **95%** detection accuracy.
- Trained **deep learning** models(CNN) on the Adience dataset with over **26,000 images**, covering 8 age groups and 2 gender classes.
- Implemented real-time prediction with webcam integration, processing at 15 frames per second (FPS).
- Performed image normalization, resizing, and augmentation, improving model accuracy by 12%.
- Used **VGG16** and **ResNet50** for transfer learning, resulting in **90% accuracy** for gender and **85% for age predictions**.

Prediction Molecular Properties using Graph Neural Networks — **Jan 2024 – June 2024**

- Led a project to leverage **Graph Neural Networks (GNNs)** for accurate prediction of crucial molecular properties (Absorption, Toxicity, Solubility) essential for drug development, achieving a **predictive accuracy of over 92%** across TOX21, BBBP, and ESOL datasets.
- Engineered and compared four GNN architectures—MPNN, GCN, GraphSAGE, and GAT—alongside MOL2VEC embedding techniques, **improving model performance by 20% over baseline models**.
- Conducted comprehensive analyses to visualize and identify critical molecular substructures influencing drug properties, **reducing prediction error by 15%**, and providing actionable insights for chemical data interpretation in drug discovery.

Selected publications

- **Wang, J** and Cheng, S. ”Integrating Machine Learning Potentials with Geometric Descriptors for Enhanced Molecular Dynamics of Rod-Shaped Particle Systems.”
- **Wang, J** Seidel, G. and Cheng, S. ”Analytical Interaction Potential for Lennard-Jones Rods.”
- **Wang, J** Seidel, G. and Cheng, S. ”Simulation of Rod using LAMMPS: An Implementation of the Rod Model and Its Applications.”