Baijun Xie

Email: baijunxie.bx@gmail.com Mobile: +1-202-779-4168

Portfolio Curriculum Vitae Google Scholar GitHub LinkedIn Redmond, Washington

Research Areas

Emotion Recognition, Large Language Models (LLMs) Applications, Multimodal LLMs, Robotics, Reinforcement Learning (RL), Human-Robot/Computer Interaction (HRI/HCI)

Experience

- **Meta Reality Labs** — Redmond, Washington
  Research Engineer — *Apr. 2024 - Now*
  - **Algorithm Development**: Designed and implemented algorithms and software for augmented reality (AR) input prototype devices in HCI, leveraging advanced signal processing and machine learning techniques.
  - **System Integration**: Developed and maintained solutions to synchronize data and created seamless experiences across complex, multi-device wearable systems, integrating multimodal sensory inputs and hardware components.
  - **Interactive Experience**: Enhanced user experience by leveraging LLMs for personalized interactions; rapidly prototyped AR input devices to manipulate avatars, demonstrating system capabilities in a constellation.

- **Machine Learning Research - ART-MED Lab** — Washington, DC
  Machine Learning Researcher — *May. 2020 - Mar. 2024*
  - **GenAI**: Post-trained LLMs using PEFT methods (e.g., LoRA) and integrated a reward model powered by Proximal Policy Optimization (PPO) to deliver highly personalized, context-aware responses from AI agents.
    Developed a reward model for PPO that captures implicit human feedback by analyzing emotional characteristics, including discrete emotions, arousal-valence levels, and empathy, to enhance adaptive RLHF.
  - **Multimodal Learning**: Developed and presented a RoBERTa-based multimodal LLM integrating emotional and social cues from video data (visual, audio, and speech) to enhance contextual understanding and response accuracy in a video question-answering task (Social IQ 2.0 Challenge — ICCV 2023).
    Explored a contrastive loss function with emotional cues to improve performance; achieved over 75% accuracy on the social intelligence benchmark, surpassing the T5 baseline by 36.7%.
  - **Medical Imaging**: Segmented nerves from birefringence and RGB images using U-Net architecture with an attention-based fusion module; evaluated the multimodal fusion module's efficacy in nerve identification, achieving an 18% improvement in the dice coefficient and a 19.6% improvement in the F2 score over unimodal U-Net.
  - **Sensor Fusion**: Implemented a trainable Extended Kalman Filter (EKF) for ground vehicle localization by leveraging IMU inputs when SLAM is unavailable; improved the EKF by integrating a customized fusion module; obtained a 42.6% MSE reduction over the vanilla EKF baseline.

- **The George Washington University** — Washington, DC
  Doctorate Research Assistant — *Sep. 2018 - Dec. 2023*
  - **Emotion Recognition**: Developed a cutting-edge multimodal framework by fine-tuning backbone models and integrating a fusion module, improving the F1-score by 3.6% over baseline models on the MELD dataset.
  - **Robotic System**: Integrated a robotic system with real-time human pose estimation and gesture imitation for seamless HRI, achieving 92%+ accuracy on the FABO emotional gesture database utilizing the SlowFast model.
  - **Conversational AI**: Developed an AI chatbot on a humanoid robot to facilitate empathetic conversations using a speech-to-text engine and ChatGPT backend, personalizing interactions through ReAct prompting strategies.
  - **User Study Design**: Designed interactive HRI scenarios tailored for autistic users, demonstrating significant reductions in anxiety through robot-assisted interventions, as evidenced by improvements in physiological signals and questionnaire responses (p-value < 0.05).

- **The George Washington University** — Washington, DC
  Graduate Researcher — *Feb. 2017 - May 2018*
  - **Pose Estimation**: Applied a model-based least squares pose estimation method to predict the pose of a ground robot; trained custom CNN models and attained over 10 FPS for fast object detection on an embedded system.
  - **Robotic Control**: Conducted experiments and simulations with PID and Geometric controllers for a quadrotor.

Projects

- **LLM**: Fine-tuned LLMs for text summarization using PEFT, achieving a 12%+ improvement in ROUGE score.
- **Robotics Vision**: Calibrated custom IMU-camera devices and deployed VINS-Mono for SLAM using ROS.
- **Reinforcement Learning**: Employed DDPG to enable agents to control arm motions and maintain target position.
- **Stress Prediction**: Leveraged heart rate characteristics and SVM to predict human physiological stress levels.
- **Hand Gestures**: Extracted hand keypoints by OpenPose and deployed a Conv-LSTM model for contextual learning.
- **Numerical Analysis**: Derived and simulated the fully vectorized numerical results of SWEs; analyzed grid convergence using numerical schemes, and enhanced the convergence order with the Lax-Wendroff scheme.

Publications

- **Journal: Real-time teleoperation of magnetic force-driven microrobots with a motion model and stable haptic force feedback for micromanipulation**: Nanotechnology and Precision Engineering . 2025.
- **Conference: An Empathetic Social Robot with Modular Anxiety Interventions for Autistic Adolescents**: IEEE International Conference on Robot and Human Interactive Communication (ROMAN). 2024.
- **Conference: Multi-Modal Correlated Network with Emotional Reasoning Knowledge for Social Intelligence Question-Answering**: IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). 2023.
- **Conference: "Can You Guess My Moves?" Playing Charades with a Humanoid Robot Employing Mutual Learning with Emotional Intelligence**: Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI). 2023.
- **Journal: DXM-TransFuse U-net: Dual Cross-Modal Transformer Fusion U-Net for Automated Nerve Identification**: Computerized Medical Imaging and Graphics. 2022.
- **Journal: Trainable Quaternion Extended Kalman Filter with Multi-Head Attention for Dead Reckoning in Autonomous Ground Vehicles**: Sensors. 2022.
- **Conference: Empathetic Robot With Transformer-Based Dialogue Agent**: International Conference on Ubiquitous Robots (UR). 2021.
- **Journal: Robust Multimodal Emotion Recognition from Conversation with Transformer-Based Crossmodality Fusion**: Sensors. 2021.
- **Conference: Dance with a Robot: Encoder-Decoder Neural Network for Music-Dance Learning**: Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI). 2020.
- **Journal: Musical Emotion Recognition with Spectral Feature Extraction Based on a Sinusoidal Model with Model-Based and Deep-Learning Approaches**: Applied Sciences. 2020.

Education

- **The George Washington University** — Washington, DC
  Doctor of Philosophy in Biomedical Engineering; GPA: 3.81/4.0 — *Sep. 2018 – Dec. 2023*

  Honors: *Collins Distinguished Doctoral Fellowship, GW Technology Commercialization Innovation Competition*
  Dissertation: *Empathetic Robotic Companion for Autistic Adolescents With Multimodal Human-Robot Interaction*

- **The George Washington University** — Washington, DC
  Master of Science in Mechanical Engineering; GPA: 3.87/4.0 — *Sep. 2016 – May 2018*

- **Shenzhen University** — China
  Bachelor of Engineering in Photoelectric Information Engineering — *Sep. 2012 – June 2016*

  Honors: *A Hundred Outstanding Final Year Theses Prize, University Student Challenge Cup (2nd Prize)*

Technical Skills

- **Languages**: Python, C/C++, MATLAB, C#.
- **Libraries**: PyTorch, Transformers, LangChain, ROS, OpenCV, OpenPose.
- **Tools**: LaTex, Markdown, Git/Github, Jupyter Notebook, Visual Studio Code, Linux/Ubuntu.
- **Machine Learning**: Transformer, LLM, Reinforcement Learning, Multimodal Learning.