Baijun Xie

Email: baijunxie.bx@gmail.com Mobile: +1-202-779-4168
Portfolio Curriculum Vitae Google Scholar GitHub LinkedIn Redmond, Washington

Ph.D. research in machine learning and human-robot interaction, including deep learning, computer vision, natural language processing, and robotics; 5+ years of Python experience; different research projects with cross-group collaborations.

Technical Skills

- **Languages**: Python, C/C++, MATLAB, C#, SQL, R.
- **Frameworks**: PyTorch, TensorFlow, OpenCV, Transformers, OpenPose, Pandas, Scikit-learn, SciPy, Numpy, Matplotlib.
- **Tools**: Latex, Markdown, Git/Github, Jupyter Notebook, Visual Studio Code, Linux/Ubuntu, AWS, ROS.
- **Machine Learning**: CNN, RNN, Transformer, Reinforcement Learning (RL), Multimodal Learning, Self-Supervised Learning.

Experience

- **Meta Reality Labs** Redmond, Washington
  Research Engineer *Apr. 2024 - Now*
  - **Algorithm Development**: Design and implement algorithms and software for AR input prototype devices, using state-of-the-art signal processing and machine learning techniques.
  - **Data Synchronization**: Develop and maintain solutions to synchronize data and build experiences across complex, multi-device wearable systems.
  - **User experiments**: Design and implement user experiments to evaluate the performance of prototype devices.
- **The George Washington University — ART-MED Lab** Washington, DC
  Machine Learning Research Assistant *Sep. 2018 - Dec. 2023*
  - **Emotion Recognition**: Developed a state-of-the-art multimodal framework by fine-tuning backbone models and integrating a fusion module, with an increment of over 2% in accuracy and F1-score compared with the baseline models.
  - **Robotic System**: Programmed a robotic system with real-time pose estimation and gesture imitation for seamless human-robot interaction (HRI); employed advanced computer vision models such as the SlowFast model to identify multimodal human behaviors; reached 92%+ accuracy in classifying emotional upper body gestures.
  - **Conversational AI**: Developed an AI chatbot on a humanoid robot for providing empathetic conversations utilizing a speech-to-text engine and LLMs (ChatGPT, LangChain); fine-tuned language models by implementing PPO and LoRA.
  - **User Study Design**: Designed interactive HRI scenarios tailored for autistic users, aiding robot-assisted intervention in alleviating anxiety with significant outcomes (p-value < 0.05) in physiological signals and questionnaire responses.

Selected Projects

- **Social IQ 2.0 Challenge — ICCV 2023 (Deep Learning, Computer Vision, Natural Language Processing, Multimodal)**: Tech: Python, PyTorch, Transformers, Scikit-learn, Pandas, Numpy.
  - Presented a multimodal RoBERTa-based model leveraging emotional social cues in a video question-answering task.
  - Achieved 75%+ accuracy on the validation set of the social intelligence benchmark, outperforming the baseline by 36.7%.
- **Medical Image Segmentation for Automated Nerve Identification — Collaboration with a Research Group at Children's National Hospital (Deep Learning, Computer Vision, Image Processing, U-Net, Multimodal, Image Segmentation)**: Tech: Python, OpenCV, PyTorch, Scikit-learn, Numpy, Pandas, Matplotlib.
  - Segmented nerves from birefringence and RGB images via a U-Net architecture with a Transformer-based fusion module.
  - Systematically assessed the efficacy of the multimodal fusion module in facilitating nerve identification; achieved 0.72 on the dice coefficient with an improvement of 0.11, and 0.76 on the F2 score, a gain of 19.6% over single-modality networks.
- **Generative AI with Large Language Models (LLMs) for Text Analysis (Natural Language Processing, Reinforcement Learning)**: Tech: Python, PyTorch, Transformers, Parameter Efficient Fine-Tuning (PEFT).
  - Employed PEFT methods to fine-tune LLMs for text summarization and gained over 12% improvement in ROUGE score.
  - Utilized a personalized reward model with Proximal Policy Optimization (PPO) to reduce the LLMs' negative outputs.

Selected Publications (3 out of 9)

- **Conference: Multi-Modal Correlated Network with Emotional Reasoning Knowledge for Social Intelligence Question-Answering**: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.
- **Journal: DXM-TransFuse U-net: Dual Cross-Modal Transformer Fusion U-Net for Automated Nerve Identification**: Computerized Medical Imaging and Graphics 99 (2022): 102090.
- **Journal: Robust Multimodal Emotion Recognition from Conversation with Transformer-Based Crossmodality Fusion**: Sensors 21.14 (2021): 4913.

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*
- **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)*