Caleb Jones Shibu

Linkedin: https://www.linkedin.com/in/caleb-jones-shibu
Github: https://github.com/CalebUAz
Work Github: https://github.com/calebshibu
Github bio: https://calebuaz.github.io

Email : calebshibu@gmail.com
Location: Denver, Colorado
Mobile : +1-646-226-2199

Machine Learning - Software Engineer

Python • Pytorch •TensorFlow • OpenCV • LLM

Experience

- **Page Technologies** — Boulder, CO
  Full Stack AI Engineer — *Sep 2025 – Present*
  - **AI Product Development**: Design, implement, and deploy end-to-end AI solutions, integrating modern ML/DL models with scalable cloud-based APIs and user interfaces.
- **Computer Connection Wisconsin** — Reedsburg, WI
  Web developer (LLM Engineer) — *July 2025 – Present*
  - **Web Maintenance**: Maintained and updated multiple CCWIS websites.
  - **RAG Chatbot**: Developed and deployed a CCWIS chatbot to provide customer tech support, with integrated human-agent escalation and a ticket-based support system.
- **Allen Institute** — Seattle, WA
  Software Engineer II (Machine Learning) — *March 2024 – March 2025*
  - **Cellpose**: Improved Cellpose model performance by 22% (AUC-ROC) for real-time neuron segmentation by implementing PyTorch-based transfer learning, resulting in automated voltage imaging analysis and reduced manual annotation time by 2 hours per week.
  - **iGluSnFR Simulation Pipeline**: Developed a simulation pipeline for iGluSnFR-expressing dendrites, optimizing 12+ parameters to replicate microscope imaging conditions, enabling ground-truth validation of motion correction algorithms.
  - **iGluSnFR Motion Correction**: Engineered a motion correction pipeline for dendritic imaging that maintained subcellular tracking accuracy during extreme motion, outperforming existing methods like Suite2p, CaImAn, and PatchWarp.
  - **iGluSnFR Source Extraction**: Contributed to an OpenCV-based super-resolution microscopy source extraction pipeline to identify synapses within dendrites, advancing iGluSnFR indicator data analysis techniques.
  - **iGluSnFR Image Processing Methods**: Co-authored an upcoming paper detailing novel image processing methods for iGluSnFR indicator data analysis.
- **Biotronics** — Ames, IA
  Machine Learning Engineer — *Jan 2024 – Feb 2024*
  - **ML Model Explainability**: Enhanced ResNet-50 model interpretability for livestock fat prediction by integrating SHAP (SHapley Additive exPlanations) and Grad-CAM, improving stakeholder trust in predictions by quantifying feature importance.
- **University of Arizona** — Tucson, AZ
  Graduate Research Assistant — *Aug 2021 – Dec 2023*
  - **Baseline Task**: Developed a PyGame-based finger-tapping and imaging rating application for the baseline task in the Theory of Mind Computer Architecture for Teams (ToMCAT) project.
  - **Data Acquisition**: Built ETL pipelines for acquiring and preprocessing Eye-tracking, fNIRS, and EEG data from over 100 subjects.
  - **Real-Time Physio Visualization**: Developed a PyQT5 tool to plot EEG and fNIRS signals multicasted over network using LSL in real-time.
  - **Data Conversion and Labeling**: Automated conversion of XDF files (EEG, fNIRS, Eye-tracking) into labeled CSV files; implemented channel quality assessment and motion artifact filtering for EEG and fNIRS signals, streamlining analysis preparation.
  - **Conference Paper**: Conducted machine learning–based classification experiments using a multimodal dataset and authored a paper accepted at NeurIPS 2023.

- **Sree Chitra Tirunal Institute for Medical Sciences & Technology** Kerala, India
  Project Scientist *Jan 2021 – July 2021*
  - **Neurofeedback Game**: Built a neurofeedback game application that filtered fNIRS signals in real-time, predicted brain states using a deep learning model, and incorporated predictions into a PyGame interface, advancing brain-computer interface research.
  - **xAI fNIRS System**: Developed an explainable AI application for fNIRS signal classification using DeepSHAP, enhancing interpretability of brain state predictions; published in *Frontiers in Human Neuroscience*.
- **St. Jude Children's Research Hospital** Memphis, TN
  Research Intern *Sep 2020 – Jan 2021*
  - **Active vs Passive Brain Activation**: Developed deep learning–based classification of active vs. passive brain state associated with single-trial lower limb motor preparation in stroke patients.
- **Sree Chitra Tirunal Institute for Medical Sciences & Technology** Kerala, India
  Research Intern *June 2019 – Sep 2020*
  - **ML Classification of fNIRS Signals**: Created handcrafted features for fNIRS signals using PCA and ICA, improving classification accuracy of SVM and k-NN models.
  - **Deep Learning Classification of fNIRS Signals**: Developed CNN and LSTM models using sliding windows to treat fNIRS signals as images, boosting accuracy from 55% to 97%. Addressed data scarcity by augmenting dataset dimensionality. Part of this work was presented at an IEEE conference in Japan.

Academic Projects

- **RL-Bowling**: Developed a reinforcement learning based bowling game using OpenAI Gym environment.(Feb '25)
- **Resume-GPT**: A Streamlit-based chatbot application that utilizes GPT-3.5-turbo, enabling users to interact and ask questions about a person's background based on their resume. (Oct '23)
- **Grocify**: A web application, developed using Django and React and deployed on the Google Cloud Platform, enhances the grocery list-making experience and is accessible on both PC and mobile devices. (May '23)
- **ToMCAT-offline-Viz**: PyQt5-based GUI not only visualizes the TOMCAT dataset but also presents detailed views of EEG and fNIRS signals, providing a thorough understanding of brain activities. Moreover, it offers a topological perspective, mapping these neural signals in a spatial context. This enhances the user's comprehension of cognitive processes and aids researchers in observing how team interactions evolve across various tasks and environments. (May '23)
- **Data poisoning in Machine Learning**: Explored an approach to poison a Machine Learning Model by attaching a Trojan Net which makes the model misclassify with high accuracy (Dec '22)
- **Multi-Modal emotion recognition**: Worked on a diffusion-based deep neural network that classifies emotions from a multimodal dataset. (Aug '22 - Dec '23)
- **fNIRS data augmentation with GANs**: Using Python implemented a GAN model which generated synthetic fNIRS data to solve issues related to fNIRS data acquisition and model training. (Nov '21)
- **Explainable medical image classification**: Developed a model which was able to classify Covid19 chest X-ray with an accuracy of 90% and ISIC Skin cancer dataset with an accuracy of 70%. The model classification was explained using LIME, GradCam HeatMap, and Saliency Maps. (Nov '21)

Education

- **University of Arizona** Tucson, Az
  Masters in Computer Science (Full Academic Scholarship) *Aug 2021 - Dec 2023*
  Courses: Advance Topics in Artificial Intelligence, Principles of Machine Learning, Design and Analysis Of Algorithms, Computer Vision, Computer Security, Software engineering, Advanced-Data Visualization
- **Goa University** Goa, India
  Bachelor of Computer Science *Aug 2014 - May 2018*

Skills Summary

- **Languages**: (Proficient) Python and MATLAB; (Familiar) C, C++, C#, SQL, Django, PostgreSQL, BASH, TS
- **Tools**: Pytorch, TensorFlow, langchain, Plotly, IPython, Excel, Keras, Scikit-learn, streamlit, pygame, Bash, SSH, GCP, Nextflow, MLflow, OpenAI Gym, AWS

Publications

Caleb Jones Shibu, Sujesh Sreedharan, Arun KM, and Chandrasekharan Kesavadas. "Comparison of classification performance of handpicked, handcrafted, and automated-features for fNIRS-BCI system". In: *2020 5th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*. IEEE, 2020, pp. 152–157

Caleb Jones Shibu, Sujesh Sreedharan, Arun KM, Chandrasekharan Kesavadas, and Ranganatha Sitaram. "Explainable artificial intelligence model to predict brain states from fNIRS signals". In: *Frontiers in Human Neuroscience Brain-Computer Interfaces* (2023)

Adarsh Pyarelal, Eric Duong, **Caleb Jones Shibu**, Paulo Soares, Savannah Boyd, Payal Khosla, Valeria Pfeifer, Diheng Zhang, Eric S Andrews, Rick Champlin, Vincent Paul Raymond, Meghavarshini Krishnaswamy, Clayton Morrison, Emily Butler, and Kobus Barnard. "The ToMCAT Dataset". In: *Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track*. 2023. URL: https://openreview.net/forum?id=ZJWQfgXQb6

Caleb Jones Shibu. "Decoding Emotional Responses: A Comparative Study of fNIRS and EEG Neuroimaging Techniques". In: (2023). URL: https://repository.arizona.edu/handle/10150/670846