Caleb Jones Shibu

Email : calebshibu@gmail.com
Linkedin: https://www.linkedin.com/in/caleb-jones-shibu
Mobile : +1-646-226-2199
Github: https://github.com/CalebUAz
Work Github: https://github.com/calebshibu
Github bio: https://calebuaz.github.io

Education

- **University of Arizona** — Tucson, Az
  Masters in Computer Science (Fully-funded) — *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
- **Tools**: Pytorch, TensorFlow, langchain, Plotly, IPython, Excel, Keras, Scikit-learn, streamlit, pygame, Bash, SSH, GCP, Nextflow

Experience

- **Allen Institute** — Onsite
  Software engineer 2 — *March 2024 - present*
  - **Cellpose**: Boosted Cellpose model performance by 22% (AUC-ROC) for real-time neuron segmentation in SLAP2 microscopy using PyTorch-based transfer learning, automating voltage imaging analysis and reducing manual annotation time by a few hours/week."
  - **iGluSnFR Simulation Pipeline**: Engineered simulation pipeline for iGluSnFR-expressing dendrites, optimizing 12+ parameters (motion amplitude, synapse density) to replicate Bergamo microscope imaging conditions, enabling ground-truth validation of motion correction algorithms.
  - **iGluSnFR Motion Correction**: Engineered motion correction pipeline for dendritic imaging that maintains subcellular tracking accuracy (1-2.5 RMSE) during extreme motion amplitudes (>50% frame displacement), outperforming Suite2p (5-6 RMSE), CaImAn (10-20 RMSE), and Patchwarp (6-7 RMSE) in whole-cell locomotion datasets.
  - **iGluSnFR Source Extraction**: Contributed to a super-resolution microscopy-based source extraction pipeline to identify synapses within dendrites.
  - **iGluSnFR Image Processing Methods**: Co-authored an upcoming paper detailing novel image processing methods for iGluSnFR indicator data analysis.
- **Biotronics** — Remote
  Machine learning engineer — *Jan 2024 - Feb 2024*
  - **IMF 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 - ToMCAT — *Aug 2021 - Dec 2023*
  - **Baseline Task**: Developed network-based finger tapping and imaging rating application using PyGame for the baseline task for Theory of Mind Computer Architecture for Teams (ToMCAT).
  - **Data Acquisition**: Acquired fNIRS & EEG data from multiple subjects.
  - **Real-Time Physio Visualization**: Developed a tool using PyQT5 that plots EEG and fNIRS signals which are multicasted over the network using LSL in real-time.
  - **Data conversion and labeling**: Developed a script that converts XDF files (contains EEG, fNIRS, and Eye-tracking data) into labeled CSV files. Calculated channel quality and filtered out the motion artifacts from the fNIRS signals
  - **Conference paper**: Conducted classification experiments and authored a paper that was submitted to and accepted by NeurIPS 2023 conference.
- **Sree Chitra Tirunal Institute for Medical Sciences & Technology** — Kerala, India
  Project Scientist — *January 2021 - July 2021*
  - **Neurofeedback game**: Using Python built an application that filtered fNIRS signals in real-time, predicted brain state in real-time using a deep learning model and model predictions were input to a PyGame.
  - **xAI fNIRS system**: Using DeepSHAP developed an explainable AI application for fNIRS signal classification .

- **St. Jude Children's Research Hospital** Memphis, TN
  Research intern *September 2020 - January 2021*
  - **Active vs Passive brain activation**: Developed Deep learning-based classification of Active and Passive brain state associated with single trial lower limb motor preparation for stroke patients.
- **Sree Chitra Tirunal Institute for Medical Sciences & Technology** Kerala, India
  Research intern *June 2019 - September 2020*
  - **Machine Learning Based Classification of fNIRS signals**: Created handcrafted features for fNIRS signals using PCA and ICA for improving the classification accuracy of Machine learning classifiers like SVM and KNN for fNIRS signal classification.
  - **Deep Learning Based Classification of fNIRS signals**: Developed a sliding window-based CNN and LSTM Deep Learning model for fNIRS signals by treating signals as an image which boosted the classification from 55% to 97% and solved the issue of data scarcity by increasing the dimension of the dataset. A part of this work was presented at IEEE conference in Japan.

Academic Projects

- **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**: Currently working 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)

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