Yuping **Shao**

+1 970 412 0428 | yshaoeng@gmail.com | yshao.github.io/mysite/ | yshao | yu-ping-shao-4622166

Applied data/machine learning scientist with 8+ data science/engineering experience. Additional 7 years in SW/HW integration in systems.

Education

Worcester Polytechnic Institute — *Worcester, MA*

MS Computer Engineering - Data Science, (PhD Program) — *Jan 2021 - May 2023*

Georgia Institute of Technology — *Atlanta, GA*

MS Computer Science - Interactive Intelligence — *Aug 2015 - Dec 2017*

Colorado State University — *Fort Collins, CO*

BS Electrical Engineering — *Aug 2003 - May 2008*

Projects

VISLab @ Worcester Polytechnic Institute — *Worcester, MA*

Research Project Mentor — *Jan 2021 - May 2023*

- Neuroscience research 2021: Algonauts Challenge 2021, Topic: brain fMRI coupling with videos through kmeans. Best poster award at (MIT URTC 2021). Neuromatch 2021, Topic: Prediction Model from fMRI functional connectivity.
- Research mentor for Directed research 2022-2023: Topics: Few-shot learning for vehicle identification, and keypoints detection from 3D point-cloud. Topics: Feature matching with deep learning, semi-supervised semantics segmentation, and visual instance tracking with attention.

Georgia Institute of Technology — *Atlanta, GA*

Individual Project Contributor — *Jan 2015 - Dec 2017*

- Big Data for Health Analytics/Healthcare Informatics/Human Computer Interface – Project: Predictive modeling with RNN/GRU for hospital readmission using big data. Project: Zikas virus clinical support with bayesian network. Project: design affordance of diet adherence
- Educational Technology - Project: NLP, topic modeling for creating AI reading comprehension assessment.
- Machine Learning/Machine Learning for Trading – Project: Developed a market trades simulator with ML and reinforcement learning. Other: Supervised - trees/adaboost/SVM/regression. Unsupervised - gaussian mixture/kmeans clustering.
- Computer Vision/Computational Photography – Hough transforms, disparity depth, camera calibration, RANSAC, optical flow, particle filter/kalman filter, motion history image. Project: Real-time video stylization.
- Cognitive Knowledge-based AI – Project: Agent to solve visual Raven's Progressive Matrices. Project: Weather platform with cognitive system.

Experience

VISLab @ Worcester Polytechnic Institute — *Worcester, MA*

Graduate Research/Teaching Assistant — *Jan 2021 - May 2023*

- Primary Skills: python, pytorch, opencv, pandas, sklearn, linux, latex, computer vision, jupyter
- Supervised by Ziming Zhang. Topics: multi-modal registration, deep learning/optimization, semantics segmentation. Currently 1st co-author ICCV (top machine learning conference) paper in submission. Two works in medical machine learning in submission.
- Service and teaching: Head TA for graduate course Digital Image Processing, Machine Learning for Engineers (2021-2023). Served as peer reviewer for IEEE SDS, ICCV, IEEE TPAMI, NeurIPS, and TMLR. Reviewer for capstone projects for WPI ECE Dept.

Clinical Data Animation Center @ Mass. General Hospital — *Boston, MA*

Data Scientist/Data Engineer — *Aug 2016 - Dec 2020*

- Primary Skills: python, pyspark, sql, HPC, pandas, numpy, scipy, jupyter, pytorch, NLP, postgres, matlab, aws/gcp, keras/tensorflow, csharp
- Researched sepsis retrospective study on MGH ICU Inpatients (Dept. of Health and Human Services - solving sepsis) - SOFA scoring: neurologic, blood, liver, respiratory, renal, and cardiovascular. Involing HL7, ICD, LOINC, FHIR, SNOMED, CPT, and data harmonization.
- Developed realtime clinical informatics as the technical lead on physiological predictive algorithm on adverse events in ICU/ER.
- Developed the proof-of-concept (with 100 patients) of a large-scale spatial-temporal clustering visualization (tSNE with word2vec) for seizure pattern annotation. Utilized HPC for scalable data preparation through explorative data analytics and active learning (offline RL).
- Led efforts in curating and building medical informatics data pipelines for delirium, sepsis, and cardiac waveform research
- Operationalized proof-of-concept matching algorithm of multimodal NLP topics modeling/timeseries pipeline of EEG recordings, neurologist's text, and (EHR) entity matching and fuzzy matching scoring. Published in Dell/EMC knowledge sharing. Tokenization for ElasticSearch
- Developed scalable physiological waveform data storage for bedside/EEG/respiratory signal/database for critical care clinical informatics (ccci).
- Built an automated ventilator-assisted events (VAE) pipeline from the large scale clinical timeseries database. Resulted in a journal article for the location accuracy and then the data accuracy.
- Organized and presented in series of 4 workshops on clinical SQL timeseries, datawarehouse, NLP, and big data health analytics, structured query on EHR and machine learning algorithms for EEG.

Boulder Environmental Sciences and Technology

Boulder, CO

Software/System Engineer

Feb 2014 - May 2015

- Primary Skills: sql, ms-sql, xml/json, c/c++, python, postgresql, sqlite, django(REST), i2c/spi/uart/serial, ubuntu, h.264 rtsp, numpy, pandas, pyqt, embedded linux, TCP/IP.
- Developed embedded system software control for unmanned aerial vehicle mounted radiometry DAQ system (ADC, environmental sensors, IMU/GPS, rotational encoder, 8-ports switch, onboard computers, motor control).
- Developed radiometry weather station based on the software-defined building project concept (weather transmitter, GPS, soil sensors, camera).

Spectra Logic

Boulder, CO

Software/System Engineer

Jul 2009 - Jul 2013

- Primary Skills: OOD/UML, DSP, perl, ruby, bash, matlab, c/c++, oscilloscope, digital analyzer, EDA, arduino/rasp-pi.
- Developed high-level hardware/system/software for robotics manufacturing business unit. with logging, monitoring, reporting capabilities for quality improvement analysis (deep storage tape library, robotics, environmental sensors, memory scanner, servo controllers, barcode scanner).
- Developed quality assurance, sustaining engineering database system on robotics assemblies and boards.
- Developed 15+ electrical test fixtures for robotics-controlled library parts (motor drives, DC brushless, sensors, servos, barcode readers) to conduct failure analysis.
- Configuration management and development of automation scripts. Start-up diagnostics, system status, connectivity, on-board interfaces.

Woodward Governors

Fort Collins, CO

Software/System Engineer

May 2008 - Apr 2009

- Developed control system protocol analyzer for grey-box testing on turbine controller products.

Publications

1. Ganglberger, W., Krishnamurthy, P. V., Quadri, S. A., Tesh, R. A., Bucklin, A. A., Adra, N., Da Silva Cardoso, M., Leone, M. J., Hemmige, A., Rajan, S., Panneerselvam, E., Paixao, L., Higgins, J., Ayub, M. A., Shao, Y.-P., Coughlin, B., Sun, H., Ye, E. M., Cash, S. S., … Westover, M. B. (2023). Sleep staging in the ICU with heart rate variability and breathing signals. An exploratory cross-sectional study using deep neural networks. *Frontiers in Network Physiology*, *3*. https://doi.org/10.3389/fnetp.2023.1120390
2. Bucklin, A. A., Ganglberger, W., Quadri, S. A., Tesh, R. A., Adra, N., Da Silva Cardoso, M., Leone, M. J., Krishnamurthy, P. V., Hemmige, A., Rajan, S., Panneerselvam, E., Paixao, L., Higgins, J., Ayub, M. A., Shao, Y.-P., Ye, E. M., Coughlin, B., Sun, H., Cash, S. S., … Westover, M. B. (2022). High prevalence of sleep-disordered breathing in the intensive care unit - a cross-sectional study. *Sleep & Breathing = Schlaf & Atmung*. https://doi.org/10.1007/s11325-022-02698-9
3. Ganglberger, W., Velpula Krishnamurthy, P., Quadri, S., Tesh, R., Bucklin, A., Adra, N., Cardoso, M., Leone, M., Hemmige, A., Rajan, S., Panneerselvam, E., Paixao, L., Higgins, J., Ayub, M., Shao, Y.-P., Coughlin, B., Sun, H., Ye, E., Cash, S., & Westover, M. B. (2021). *Sleep in the intensive care unit through the lens of breathing and heart rate variability: A cross-sectional study*. https://doi.org/10.1101/2021.09.23.21264039
4. Shenoy, E., Rosenthal, E., Shao, Y.-P., Biswal, S., Ghanta, M., Ryan, E., Suslak, D., Swanson, N., Moura Junior, V., Hooper, D., & Westover, M. B. (2018). Real-time, automated detection of ventilator-associated events: Avoiding missed detections, misclassifications, and false detections due to human error. *Infection Control & Hospital Epidemiology*, *39*, 1–8. https://doi.org/10.1017/ice.2018.97
5. Mukerji, S. S., Das, S., Alabsi, H., Brenner, L. N., Jain, A., Magdamo, C., Collens, S. I., Ye, E., Keller, K., Boutros, C. L., Leone, M. J., Newhouse, A., Foy, B., Li, M. D., Lang, M., Anahtar, M. N., Shao, Y.-P., Ge, W., Sun, H., … Westover, M. B. (2021). Prolonged intubation in patients with prior cerebrovascular disease and COVID-19. *Frontiers in Neurology*, *12*. https://doi.org/10.3389/fneur.2021.642912
6. Shashikumar, S. P., Wardi, G., Paul, P., Carlile, M., Brenner, L. N., Hibbert, K. A., North, C. M., Mukerji, S. S., Robbins, G. K., Shao, Y.-P., Westover, M. B., Nemati, S., & Malhotra, A. (2021). Development and prospective validation of a deep learning algorithm for predicting need for mechanical ventilation. *Chest*, *159*(6), 2264–2273. https://doi.org/https://doi.org/10.1016/j.chest.2020.12.009
7. Sun, H., Jain, A., Leone, M. J., Alabsi, H. S., Brenner, L. N., Ye, E., Ge, W., Shao, Y.-P., Boutros, C. L., Wang, R., Tesh, R. A., Magdamo, C., Collens, S. I., Ganglberger, W., Bassett, I. V., Meigs, J. B., Kalpathy-Cramer, J., Li, M. D., Chu, J. T., … Westover, M. B. (2020). CoVA: An Acuity Score for Outpatient Screening that Predicts Coronavirus Disease 2019 Prognosis. *The Journal of Infectious Diseases*, *223*(1), 38–46. https://doi.org/10.1093/infdis/jiaa663
8. Zafar, S., Rosenthal, E., Jing, J., Bordbar, E., Kassa, S., Tabaeizadeh, M., Nour, H. A., Sun, H., Javed, F., Edhi, M., Gallagher, J., Ghanta, M., Shao, Y.-P., Shoukat, M., Ge, W., Cole, A., & Westover, M. (2020). 36: BURDEN OF EEG ICTAL-INTERICTAL CONTINUUM ACTIVITY PREDICTS POOR OUTCOME IN CRITICALLY ILL PATIENTS. *Critical Care Medicine*, *48*(1).
9. Tabaeizadeh, M., Nour, H. A., Shoukat, M., Sun, H., Jin, J., Javed, F., Kassa, S., Edhi, M. M., Bordbar, E., Gallagher, J., Moura, V., Ghanta, M., Shao, Y.-P., Cole, A. J., Rosenthal, E. S., Westover, M. B., & Zafar, S. F. (2020). Burden of epileptiform activity predicts discharge neurologic outcomes in severe acute ischemic stroke. *Neurocritical Care*, *32*, 697–706.
10. Shenoy, E. S., Rosenthal, E. S., Biswal, S., Ghanta, M., Ryan, E. E., Shao, Y.-P., Suslak, D., Swanson, N., Valdery, M. J., Hooper, D. C., & Westover, M. B. (2017). Real-Time Automated Surveillance for Ventilator Associated Events Using Streaming Electronic Health Data. *Open Forum Infectious Diseases*, *4*(suppl_1), S633–S633. https://doi.org/10.1093/ofid/ofx163.1681