

Sean M. Hendryx

✉ seanmhendryx@gmail.com
<https://smhendryx.github.io/>

Machine learning researcher with a love of useful models for complex problems

Experience

- Jan 2019–Present **Staff Machine Learning Research Engineer**, *Standard Cognition, Corp.*, San Francisco, CA.
- I build deep learning systems for multi-camera, machine vision problems.
 - Currently research video action recognition model architectures and training algorithms.
 - Research and engineering for human pose estimation.
 - Built end-to-end MLOps for data collection & labeling, dataset creation & versioning, and model training, validation, registration, & release for Standard's core deep model.
 - Research and service development for product change detection.
 - Developed transfer learning and domain adaptation systems for reducing generalization error in production.
 - Tracking research and algorithmic development, including development of metrics, CICD, and onboarding the tracking team to Weights & Biases and Data Version Control.
 - Optimized model training workflows to be cost effective and cloud-native when migrating off of a centralized data center.
 - Built systems for quantifying generalization error of deployed models using human in the loop feedback.
 - Lead migration of Standard's core model from TensorFlow version 1 and TensorPack to PyTorch. This migration required tens of thousands of lines of code and resulted in 6× decreased model training costs and removal of large amounts of technical debt.
 - Developed realtime production video model service components in Rust with Gstreamer, the Rust `tch` crate, and `Aten`.
 - Adapted models to new hardware arrays.
- Dec 2017–Dec 2018 **Software Engineer - Machine Learning**, *Explorer.ai*, San Francisco, CA.
- Designed, experimented, built, rapidly iterated, and launched to production machine learning and computer vision systems in the self-driving car space. Played pivotal role in the acquisition of Explorer.ai by Standard Cognition, Corp.
- Developed and deployed image segmentation systems.
 - Developed, tested, and deployed algorithms in Python for chunking massive datasets enabling efficient storage and parallelized processing.
 - Built data pipelines to generate high-definition maps for hundreds of linear miles of roads. This work included:
 - Ground-classification of point clouds, intelligent rasterization systems, and 2-D to 3-D projection of semantic features.
 - Built 3-D visualization tools for fast inspection of large 3-D datasets.
 - Trained deep neural network models and experimented with different architectures and training methodologies for image classification.
 - Developed and deployed object detection models and necessary training infrastructure including novel augmentation approaches that significantly increased accuracy.
 - Ported trained models from desktops to deployment in the cloud at scale. This work included:
 - Dependency isolation, code refactoring and optimizations, low level debug of TensorFlow model deployment, and building the pipelines to deploy the production services
 - Converted heavy, expensive ROS-based inference system to a mobile device data inference module for cloud deployment, enabling 10× scale at the same hardware cost.
 - Built API for transferring massive (multi-terabyte scale) datasets from compute instances to S3 storage.

- Aug 2017-May 2018 **Graduate Research Assistant (Computer Vision)**, *University of Arizona*, Tucson, AZ.
 Researcher and software developer of computer vision systems in the Interdisciplinary Visual Intelligence Lab. I validated and compared software for facial recognition against multiple large, annotated datasets. Specifically, this work included building C++ software, running cross-validation experiments in R and Bash and implementing equations and algorithms from the literature. I lead a group of undergraduate computer science researchers on testing and developing software for extracting information (such as pain, blink rate, and other emotions) from videos and images.
- May-Aug 2017 **Graduate Research Assistant (Natural Language Processing)**, *University of Arizona*, Tucson, AZ.
 Contributed development to natural language processing software in Scala and Python to identify cancer signaling pathways as part of the Big Mechanism, DARPA project. This technology reads research papers and hypothesizes explanatory models of cancer development. Specifically, I worked on feature engineering, implementing machine learning classifiers, and comparing their accuracy using cross-validation.
- Aug 2015-Apr 2017 **Graduate Research Assistant (Statistics)**, *University of Arizona*, Tucson, AZ.
 Applied computation to solving natural resource problems. Developed and deployed machine learning models to quantify natural resources from large-scale computer vision-derived datasets. Investigated natural resource dynamics using time-series analysis, data visualization, and other quantitative techniques as part of a National Science Foundation project.
- May-Aug 2015 **Graduate Research Technician**, *University of Arizona*, Tucson, AZ.
 Data analysis, research design, and experimental installation to determine ecosystem dynamics in response to climatic forcings. Produced regression and time series analyses. Visualized large datasets to accelerate comprehension and produce actionable knowledge.
- Jul 2011-Apr 2015 **Energy Analyst**, *Hammer & Hand*, Portland, OR.
 Energy, resource, and financial modeling for the construction industry. Project and team management. Translated quantitative, analytical insight into actionable recommendations.

Education

- 2020 **Deep Multi-Task & Meta-Learning**, *Stanford University*, Stanford, CA.
- 2015-2018 **MS, Information**, *University of Arizona*, Tucson, AZ.
 focus: Machine Learning
 gpa: 4.0/4.0
- 2015-2017 **MS, Natural Resources**, *University of Arizona*, Tucson, AZ.
 focus: Remote Sensing & Statistics
 gpa: 4.0/4.0
- 2014 **Postbaccalaureate, Computer Science**, *Portland Community College*, Portland, OR.
 focus: C++, SQL, and Databases
- 2007-2011 **BS, Applied Environmental Science**, *Western Washington University*, Bellingham, WA.
 focus: Modeling

Publications

1. Sean M Hendryx, Dharma Raj KC, Bradley Walls, and Clayton T Morrison. Federated Reconnaissance: Efficient, Distributed, Class-Incremental Learning. *arXiv preprint*. <https://arxiv.org/abs/2109.00150>, 2021.
2. Sean M Hendryx, Andrew B Leach, Paul D Hein, and Clayton T Morrison. Meta-learning initializations for image segmentation. *4th Workshop on Meta-Learning at NeurIPS 2020, Vancouver, Canada*. https://meta-learn.github.io/2020/papers/44_paper.pdf, 2020.
3. Enrique Noriega-Atala, Paul D Hein, Shraddha Satish S Thumsi, Zechy Wong, Xia Wang, Sean M Hendryx, and Clayton T Morrison. Extracting inter-sentence relations for associating biological context with events in biomedical text. *IEEE/ACM transactions on computational biology and bioinformatics*, 2019.
4. Sean M Hendryx. Assign points to existing clusters: An algorithm for labeling clustered point cloud

- datasets for predicting physical variables. In *AGU Fall Meeting Abstracts*, 2018.
5. Sean M Hendryx. Quantifying biomass from point clouds by connecting representations of ecosystem structure. Thesis published by *The University of Arizona* <https://repository.arizona.edu/handle/10150/627688>, 2017.
 6. Pratima Khatri-Chhetri, Sean M Hendryx, Kyle A Hartfield, Michael A Crimmins, Willem JD van Leeuwen, and Van R Kane. Assessing Vegetation Response to Multi-Scalar Drought across the Mojave, Sonoran, Chihuahuan Deserts and Apache Highlands in the Southwest United States. *Remote Sensing* 13, no. 6 (2021): 1103.
 7. Esther Lee, Praveen Kumar, Greg A Barron-Gafford, Sean M Hendryx, Enrique P Sanchez-Cañete, Rebecca L Minor, Tony Colella, and Russell L Scott. Impact of hydraulic redistribution on multispecies vegetation water use in a semiarid savanna ecosystem: An experimental and modeling synthesis. *Water Resources Research*, 54(6):4009–4027, 2018.
 8. Greg A Barron-Gafford, Enrique P Sanchez-Cañete, Rebecca L Minor, Sean M Hendryx, Esther Lee, Leland F Sutter, Newton Tran, Elizabeth Parra, Tony Colella, Patrick C Murphy, et al. Impacts of hydraulic redistribution on grass-tree competition vs facilitation in a semi-arid savanna. *New Phytologist*, 215(4):1451–1461, 2017.

Presentations

- Federated Reconnaissance: A New Framework for Distributed Continual Learning. Autonomy and Edge Computing session at the Fifth Annual Workshop on Naval Applications of Machine Learning (NAML 2021). <https://sites.google.com/go.spawar.navy.mil/naml>
- Balancing Machine Learning Innovation and Scalability in Fast Growing Startups. Guest lecture in the Winter 2021 graduate course Machine Learning: Special Topics taught by Dr. Paul J. Atzberger at University of California Santa Barbara.
- Meta-Learning Initializations for Image Segmentation. NeurIPS 2020 Meta-Learning Workshop. <https://meta-learn.github.io/2020/>
- Continual Progress in Cascading Model Systems by Ensuring Reproducible State. Keynote speech at Deep Learning World 2020. <https://www.deeplearningworld.com/las-vegas/2020/agenda/#session78711>

Skills

General

Research, develop, deploy, and scale solutions; object-oriented programming; functional programming; agile software development; minimum viable product; critical path method; version control; software testing; project and team management; effective communication; clairvoyant prioritization

ML, Computer Vision, & AI Methods

Implemented from scratch Deep neural networks, back propagation, meta-learning algorithms (MAML, prototypical networks, etc.), image segmentation, human pose estimation, least squares regression (homogenous and nonhomogenous), RANSAC, k-nearest neighbors, k-means, cross-validation, edge-detection, affine transformations, outlier/noise handling, search & pathfinding algorithms (such as A*), probabilistic graphical models (Bayes-nets), Expectation Maximization of Gaussian mixture models

Applied Convolutional Neural Networks (CNNs), Random forests, Support Vector Machines, Scale Invariant Feature Transform (SIFT), Structure from Motion, Iterative Closest Point, morphological filters, kernel density estimation, t-tests, time-series forecasting (with and without exogenous regressors), Bayesian Optimization with Gaussian Processes.

Technologies

Proficient Python, PyTorch, TensorFlow, Keras, NumPy, scikit-learn, pandas, matplotlib, Bash, git, Linux, OSX

Experienced	Rust, Google Cloud, Amazon Web Services, R, ggplot, data.table, lidR, Matlab, PDAL, Boto, Spark, Docker, multiprocessing, C, C++, OpenCV, PCL, Scala, \LaTeX , Markdown, snowfall for cluster computing in R, GNU Parallel, svn, Geographic Information Systems (geojson, OpenStreetMaps, Google Earth, ArcGIS, QGIS, Potree, etc.)
Knowledge Of	ROS, SQL, JavaScript

Relevant Coursework

Deep Multi-Task & Meta-Learning; Artificial Intelligence; Machine Learning; Neural Networks; Computer Vision; Bayesian Modeling & Inference; Information Research Methods; Foundations of Information; Organization of Information; Remote Sensing of Environment; Geographic Applications of Remote Sensing; Geographic Information Systems; Applied Biostatistics; Biomedical Informatics; Intro to Statistics; Calculus; Fundamentals of Computer Science & C++; C, C++, & Software Development Bootcamp through the Interdisciplinary Visual Intelligence Lab; Databases & SQL; Critical Issues in Design