Takuya Kurihana

tkurihana@uchicago.edu Chicago, IL

Skills Python, Fortran 90/95, C, C++, C#, MPI, OpenMP, Shell Script, MySQL, Java, HTML, SciLab, Matlab, Go, Microsoft Office, Latex, Jira, Git.
ML/DL libraries Tensorflow, Keras, Pytorch, Horovod, Scikit-learn, Spark-mllib, Weight and Biases
Libraries AWS (EC2, DynamoDB, Lambda, S3, AutoScaling), GCP (Vertex AI), Spark, Kubernetes, Docker, Parsl, Pachyderm, Dask

---

EDUCATION

University of Chicago   Chicago, IL
PhD in Computer Science   GPA 3.60   June 2023 (expected)

- Research area: Deep learning (Unsupervised-/Semi-supervised learning application in climate science), Machine learning, Clustering, High performance computing. Cloud dynamics.

    (1) Developed the rotation-invariant optimization function to improve feature representation in autoencoder. The rotation-invariant feature improves clustering of object patterns and textures regardless of their orientation of inputs without any assumptions concerning artificial categories, addressing misclassification depended on their rotations of object.

    (2) Developed an unsupervised cloud clustering framework to train and test 20TB MODIS satellite products, discovering >10 intermediate cloud categories underexplored in conventional cloud classifications.

    (3) Created the AI-driven Cloud Classification Atlas (**AICCA**), a novel classification dataset produced by applying rotation-invariant autoencoder and hierarchical clustering to identify robust and meaningful clusters of cloud patterns based on 22 years of MODIS Aqua and Terra instruments – 800 Terabytes or 198 million ocean cloud imagery.

    These research outputs are summarized into two papers accepted in Climate informatics 2019 and IEEE Transactions on Geoscience and Remote Sensing, and one paper under reviewed by MDPI Remote sensing.

- Honors: Heiwa Nakajima Global Scholarship, American Geophysical Union Outstanding Student Presentation Awards

MS in Computer Science   GPA 3.60   December 2020

- Specializations: (1) Distributed tensorflow training, (2) Neuromorphic computing, (3) Auto-scaling strategies, (4) Deep learning on IoT device

    (1) Benchmark task for distributed tensorflow over three cluster computers. Evaluate impact of TCP /IP network configuration (jitter, latency and packet loss) on the distributed tensorflow training.

    (2) Coded a spiking based unsupervised neural network on neuromorphic computing simulator. Use a hand-digits dataset (MNIST) to analyze different spiking signals on their bottleneck layer via PCA, revealing that different digit-class shows different spiking patterns.

    (3) Developed automatic scaling strategies aiming at efficient resource managements for a Python parallel scripting library (Parsl). The strategies improved the resource utilization by 30% by the reallocation of resources at unnecessary workers.

    (4) Modified tensorflow, Bazel and OS applications for ARM ODROID XU4 to deploy trained deep learning models. Code a Python wrapper for C/C++ energy consumption library, and evaluate consumed energies by different kernels.

University of Tsukuba   Tsukuba, Japan
MS in Meteorology   GPA 3.92   March 2021

- Specializations: Data assimilation, Numerical optimization, Numerical weather prediction.

    Developed a satellite-observation pre-processing algorithm that reduced the computation time by 98% than that of greedy algorithm against dense satellite observations.

Evaluated the impact of multiple pre-processing algorithms to assimilated results, showing that the proposed outperformed the error statistics in analysis.

University of Tsukuba Tsukuba, Japan
BS in Meteorology   GPA 3.60 March 2017

- Rice University, Research Internship Program, August–September, 2016
- University of California Berkeley, Summer program, June–August, 2015
- Honors: Nakatani - RIES Foundation Fellowship, Best Presentation Awards in International Student Science Forum Ho Chi Minh City, Second Presentation Prize in Urayasu City English Presentation Competition

SELECTED PUBLICATION

- Takuya Kurihana, Elisabeth Moyer, Rebecca Willett, Davis Gilton and Ian Foster, 2021; "Data-driven Cloud Clustering via a Rotationally Invariant Autoencoder", *IEEE Transactions on Geoscience and Remote Sensing*, *doi: 10.1109/TGRS.2021.3098008*.

- Takuya Kurihana, Ian Foster, Rebecca Willett, Sydney Jenkins, Kathryn Koeing, Ruby Werman, Ricardo Barros Lourenco, Casper Neo, and Elisabeth Moyer, 2019; "Cloud classification with unsupervised deep learning", *Proceedings of the 9th International Workshop on Climate Informatics: CI 2019*.

- Takuya Kurihana, Ian Foster, Rebecca Willett, Sydney Jenkins, Kathryn Koeing, Ruby Werman, and Elisabeth Moyer,2019; "Cloud classification with deep learning II", *AGU Fall Meeting 2019*.

- Takuya Kurihana, and Hiroshi Tanaka, 2018; "Analysis of cloud formation processes for arctic cyclone in the non-hydrostatic icosahedral grid model", *Fifth International Symposium on Arctic Research*.

EXPERIENCE

Frontier Development Laboratory Mountain view, California
Research scientist June – August, 2022

- Build a fast and reliable contaminant plume prediction under future climate scenarios

  Developed (1) an end-to-end deep Siamese autoencoder via online clustering assignments that optimizes reconstruction and clustering loss terms simultaneously to generate novel climate patterns from CMIP5 simulations, and (2) improve an U–Net integrated Fourier neural operator (U-FNO) by integrating physics-informed loss terms for groundwater transport and flow problems, to achieve multi-scale digital twin application. I worked with domain scientists to design two different deep learning algorithms, and employed data ingestion and training pipeline with tensorflow and horovod to scale them on Google Cloud Platform.

- Conference submission

  Physics-informed surrogate modeling for supporting climate resilience at groundwater contamination sites, AGU Fall Meeting 2022.

Physics-informed surrogate modeling for supporting climate resilience at groundwater contamination sites, Machine Learning and the Physical Sciences Workshop, NeurIPS 2022 (in preparation).

Woven Planet Holdings   Tokyo, Japan
Software engineer   July – October, 2021

- Conduct benchmarks and optimization of the high-definition map pipeline

  Developed a benchmark tool running in a Kubernetes with Prometheus that measured the throughput from the high-definition map (HD map) creation pipelines. Conducted extensive amount of experiments to investigate bottleneck of pipelines, optimizing the configuration of production workflow using ~200TB of satellite imagery.

- Develop dynamic autoscaling operator

  Developed and deployed an autoscaler as a Kubernetes operator that adjusts the number of k8s pods based on performance of pipelines. This autoscaler addressed suboptimal pipeline performances especially when multiple pipelines are competing the cloud resources. There is no conflict with the cloud native autoscale policy, e.g., AWS AutoScaler, as the scope of proposed autoscaler dynamically allocate optimal number of k8s pods of different pipelines running concurrently on cluster computers. The operator increased the efficiency of resource allocation by at most 500%.

Japan Meteorological Research Institute   Tsukuba, Japan
Research Assistant   June, 2016–July, 2018

- Study on uncertainty of cumulonimbus initiation and development using particle filter

  Developed parallelized Python analytical tools for dual-polarization doppler radar observation for statistical tests and assimilated simulation results.

  Researched the observation error correlation in doppler radar products and simulations; Presented the results at Japan Meteorological Society 2018.

- Study on the optimal perturbation method for ensemble data assimilation

  Researched impacts of off-diagonal elements in transform matrix under the ensemble transformation of ensemble Kalman filter via conducting observing system simulation experiments (OSSE) on SPEEDY-LETKF for 50 ensemble members. I contributed to present the results for several peer-reviewed conferences and published into a book.

  Developed Fortran95/bash-script tools for spectrum analysis, perturbation technique, and ensemble analysis for OSSE results.

Mitsubishi UFJ Morgan Stanley Securities   Tokyo, Japan
Financial Engineer   February–March, 2017

Researched weak-/strong-scaling experiments on 10 000 cores to optimize performance of a bond option pricing model

Developed an computation-time efficient algorithm for derivative price computation, and reduced the computation time by 50%.

Mizuho Securities   Tokyo, Japan
Quantitative Analyst   August, 2017

Researched case studies of volume weighted average price (VWAP) for derivative option pricing.

RIKEN Advanced Institute for Computational Science   Kobe, Japan
Graduate Research Assistant   July–August, 2017

Developed a super-observation (aggregation of dense observation to the nearest grid point of numerical simulation) algorithm to accelerate aggregation of NASA Global Precipitation Measurement product for the Nonhydrostatic ICosahedral Atmospheric Model (NICAM) i.e., find a closest grid location against a satellite observation among $O(100\,000)$ and more number of grid points. The adaption of computing efficient superobservation approach allows NICAM to simulate more realistic precipitation estimation.

Rice University   Houston, TX
Research Assistant   August–September, 2016

Developed Kalman filter and expectation-maximization algorithm for a diffusion-advection nonlinear partial differential equation

<div align="center">

ACTIVITIES

</div>

Teaching Assistant   Chicago, IL

- Teaching assistant for Inclusive technologies   January–March, 2022

  Mentor 15 undergraduate students to advise their weekly class-project proposals and inclusive technologies applications. One of my mentored team won the best award in class.

- Teaching assistant for Computer Science with Applications 2   January–March, 2021

  Manage two laboratory classes up to 40 undergraduate students to instruct python (padas, matplotlib), web-scraping, regular expression, bashscript, database and algorithms. Hold office hours and question dashboard to assist students about class materials.

- Teaching assistant for Cloud computing   April–June, 2020

  Instruct Amazon web service (EC2, Lambda, EBS, DynamoDB) and open office hours to assist up to 60 master students.

Internship mentor   Chicago, IL
Summer student supervisor

- Student mentor in METCALF program   June–September, 2020

  Supervised one undergraduate student to build an image annotation tool based on CVAT library for MODIS calibrated radiance product..

- Student mentor in Center for Data and Computing program   June–September, 2019

  Supervised one master and two undergraduate students to develop automation libraries of MODIS satellite produces' pre-processing for unsupervised cloud clustering.

  Supervised an undergraduate student to compare clustering analysis (HAC, K Means++, Spectral, and DBSCAN) for outputs from our developed cloud classification framework.