



Ahmad Maroof Karimi

Email: karimiahmad@ornl.gov

 <https://orcid.org/0000-0002-7270-8847>, <https://bitbucket.org/axk962>

<https://linkedin.com/in/ahmad-marooof-karimi/>,  [Google Scholar](#)

RESEARCH INTEREST

Data Science, High Performance Computing, and Distributed Computing: Computer vision, deep learning, graphical modeling, Hadoop, HPC, pattern recognition, slurm, spark, spatio-temporal analysis

EDUCATION

Case Western Reserve University, Cleveland, OH 08/2016- 10/2020

Ph.D, Computer Science

Area: Data Mining and Machine Learning

Dissertation: *Data science and machine learning to predict degradation and power of photovoltaic systems: convolutional and spatiotemporal graph neural networks*

Advisor: Prof. Mehmet Koyutürk and Prof. Roger H. French

GPA: 3.7/4.0

The University of Toledo, Toledo, OH 08/2014 - 07/2016

M.S., Computer Science

Dissertation: *Distributed machine learning based intrusion detection system*

Advisor: Prof. Weiqing Sun

GPA: 3.9/4.0

Aligarh Muslim University, Aligarh, India 08/2007 - 06/2011

Bachelor of Technology (B.Tech) in Electrical Engineering

Project: *Developed collision avoidance system*

RESEARCH EXPERIENCE

Oak Ridge National Lab 11/2020 - present

HPC Operational Data Scientist, AAIMS Group, CCSD Directorate

- Enabling AI solutions for Pre-exascale HPC facility
- Investigating I/O behavior of HPC application on large Scale storage system
- Leveraging generative models for generating I/O dataset for characterizing HPC application I/O pattern

Case Western Reserve University (CWRU) 08/2016-10/2020

Research Assistant, Solar Durability Lifetime Extension Research Center

Funded Research Projects

- Reliability & Power Degradation of PERC Modules, *Department of Energy (DOE), EERE, SETO*
- Towards 50 Year Lifetime PV Modules: Double Glass vs. Glass/Backsheet, *DOE, EERE, SETO*
- Module Level Exposure & Evaluation Test, *DOE, EERE, SETO*
- Energy Diagnostic Investigator for Efficiency Savings, *DOE ARPA-E*

Research Works

- Built framework for pre-processing and segmentation of thousands of PV modules electroluminescence images from wide variety of sources as requisite for machine learning algorithms
- Developed deep learning models for over 20K photovoltaic (PV) cell image classification to automate identification of sub-optimal PV modules
- Discovered clusters in PV cells based on textural patterns exhibited by the cell images corresponding to degradation modes

- Proposed novel way to estimate PV modules' power estimation using features extracted from electroluminescence images
- Quantified evolution of busbar corrosion and series resistance for PV modules as they degrade
- Unveiled the effect of climate zone on commercial PV power plant by clustering 350 large scale commercial power plant timeseries based on their degradation
- Published unique method for virtual energy audit of commercial buildings' to detect losses and recommend retrofits based on data driven statistical techniques
- Built distributed computational framework (Hadoop, HBase and Spark) and developed python and R packages that provide interface to NoSQL database and enable our research group to perform large scale data analytics as a team.
- Developing a network graph based PV power plant forecasting model for predicting power output using similarly behaving plants

The University of Toledo, Toledo, OH 08/2014 - 07-2016

Research Assistant, Advanced Computing Research Lab

- Developed real-time distributed intrusion detection system for network monitoring
- Built Hadoop scalable storage and computing platform

WORK EXPERIENCE

The University of Toledo, Toledo, OH 01/2015 - 05/2016

Graduate Assistant, Enrollment Management

- Maintained backend application and database for enrollment management
- Developed reports of enrollment forms

Technology Control Company, Riyadh, SA 03/2014 - 07/2014

Full stack developer

- Developed an application for security system and database design
- Languages: Delphi, C++ and MySQL

Tata Consultancy Services, Mumbai, India 07/2011 - 12/2013

Software engineer

- Developed an application for trading and settlement platform for government bonds, securities and other money market instruments
- Designed database on Oracle 10g
- Languages: C++, shell script, SQL query and procedure

INVENTION DISCLOSURE

pvimage:A Python Package for Image Processing and Machine Learning for PV Module Performance and Degradation <https://pypi.org/project/pvimage/>

PROFESSIONAL SOCIETIES & SERVICES

- Member, Association for Computing Machinery (ACM)
- Member, Institute of Electrical and Electronics Engineers (IEEE)
- Research presentation judge at undergraduate research symposium, CWRU, 2018
- VIP data science consultant, American Statistical Association Data Fest at CWRU, 2017 and 2018
- Organized and instructed at summer *Tea Time* program organized by SDLE and RCCI at CWRU
 - Introduction to data science and distributed computing summer 2017
 - Python and R programming summer 2018
 - Machine Learning summer 2018

HONORS & AWARDS

- Best presentation award, *Data Science Symposium*, Tohoku University, Sendai, Japan, 2018
- International travel award to attend symposium at Tohoku University, Sendai, Japan, 2018
- Full tuition waiver at the University of Toledo for MS program
- Employee of the month award three times at Tata Consultancy Services

RESEARCH PROPOSAL

- Contributed to the proposals by Prof. Roger H. French and Prof. Mehmet Koyutürk*
- Robust Photovoltaic Performance Loss Rate Determination and Power Forecasting: Using Spatio-temporal Graph Neural Network Models in a Reliable System-Topology-Aware Learning Framework, *Funding Agency*: Department of Energy, Energy Efficiency and Renewable Energy, Solar Energy Technology Office (FY2020)
 - Advanced Data Analytics to Increase Resilience, Reliability and Security of Power Systems with High PV Penetration, *Funding Agency*: Department of Energy, Energy Efficiency and Renewable Energy, Solar Energy Technology Office (FY2019)
 - Spatio-temporal Evolution of Soft/Anisotropic Polymeric Nanoparticles Under External Stimuli, *Funding Agency*: Department of Energy, Basic Energy Science (FY2019)
 - Hammers & Nails: Optimizing Machine Learning for Diverse & Challenging Environments (Hs&Ns:OptML-DCE), *Funding Agency*: Air Force Office of Sponsored Research, Center of Excellence (FY2018)

TEACHING EXPERIENCE

Case Western Reserve University, Cleveland, OH

Teaching Assistant,

- *Introduction to Database Systems (EECS 341)*: Relational model, ER model, relational algebra and calculus, SQL, OBE, views, query processing, normalization theory, concurrency control, object relational systems, Oracle SQL server, Microsoft SQL server.
- *Data Structure (EECS 233)*: pointers, files, and recursion. representation and manipulation of data: one way and circular linked lists, doubly linked lists, different representations of stacks and queues, representation of binary trees, trees and graphs, hashing; searching and sorting.
- *Introduction to Java Programming (EECS 132)*: Introduction to computer programming and problem solving with the Java language. Computers, operating systems, and Java applications; software development; conditional statements; loops; methods; arrays; classes and objects; object-oriented design; unit testing exception handling; comparison of Java to C, C++, and C#.

Research Mentorship:

- Masters theses
 - Suk Hyun Hwang, *Quantify the performance loss rate for large scale commercial PV power plants using distributed computing*, expected 05/2020
 - Shuo Xin, *Machine learning model to track the clouds over large PV power plants*, expected 05/2020
 - Benjamin G. Pierce, *Developed feature extraction methods for clustering of PV cell degradation*, expected 05/2021
- Under graduate research
 - Nicholas A. Parilla, *Developed crack detection method and planar indexing on PV modules' images* fall 2017 - fall 2018
 - Andrew J. Loach, *Measured nucleation growth on the surface of aluminum-nickel alloy by analyzing over 100K images* spring 2017 - spring 2018
 - Jonathan K. Ligh, *Contributed to data cleaning and data ingestion pipeline for real-world timeseries data* spring 2017 - fall 2017

- Kevin J. Nash, *Contributed to building an R package for collecting weather information through a REST API* spring 2017- summer 2017
- Outreach program, high school summer internship
 - Russell Krofta, *Statistical analysis of PV images and ingesting images to database* summer 2019
 - Precious Flanders, *Labeled global horizontal irradiance timeseries dataset for clear sky detection.* summer 2018

The University of Toledo, Toledo, OH

Teaching Assistant,

- *Database System I (EECS 4560)*: Relational database modeling, query languages, normalization, design issues and implementation issues of databases

LAB EXERCISE DEVELOPED

Case Western Reserve University, Cleveland, OH

- *Data Science - Statistical Learning, Modeling and Prediction (DSCI 353)*: Developed lab exercises and tutorials for deep learning modeling

Tohoku University, Sendai, Japan

- *IT Fundamentals: Applied Data Science with R*: Developed lab exercises for deep learning modeling in R using GPUs and Tensorflow

PUBLICATIONS

1. **Karimi, A. M.**, Fada, J. S., Parrilla, N. A., Pierce, B. G., Koyutürk, M., French, R. H., Braid, J. L., “Generalized and Mechanistic PV Module Performance Prediction From Computer Vision and Machine Learning on Electroluminescence Images,” *IEEE Journal of Photovoltaics*, pp. 1–10, 2020. DOI: 10.1109/JPHOTOV.2020.2973448, [karimi_1*]
2. Pierce, B. G., **Karimi, A. M.**, Liu, J., “Identifying Degradation Modes of Photovoltaic Modules Using Unsupervised Machine Learning on Electroluminescence Images,” en, Jun. 2020, p. 6
3. Whitaker, C. M., Pierce, B. G., **Karimi, A. M.**, “PV Cell Cracks and Impacts on Electrical Performance,” en, Virtual, Jun. 2020, p. 6
4. Khalilnejad, A., **Karimi, A. M.**, Kamath, S., Haddadian, R., French, R. H., Abramson, A. R., “Automated pipeline framework for distributed processing of large-scale building energy time series data,” *Accepted: PLOS one*, p. 23, 2020
5. Liu, J., Wang, M., Curran, A. J., **Karimi, A. M.**, Huang, W.-h., Schnabel, E., Kohl, M., Braid, J. L., French, R. H., “Real-world PV Module Degradation across Climate Zones Determined from Suns-Voc, Loss Factors and I-V Steps Analysis of Eight Years of I-V , Pmp Time-series Datastreams,” in *2019 46th IEEE PVSC, Chicago IL, USA, Jun. 2019*, p. 7
6. **Karimi, A. M.**, Fada, J. S., Hossain, M. A., Yang, S., Peshek, T. J., Braid, J. L., French, R. H., “Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification,” *IEEE Journal of Photovoltaics*, pp. 1–12, 2019. DOI: 10.1109/JPHOTOV.2019.2920732
7. **Karimi, A. M.**, Fada, J. S., Liu, J., Braid, J. L., Koyutürk, M., French, R. H., “Feature Extraction, Supervised and Unsupervised Machine Learning Classification of PV Cell Electroluminescence Images,” in *2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC)*, Jun. 2018, pp. 0418–0424. DOI: 10.1109/PVSC.2018.8547739
8. **Karimi, A. M.**, Niyaz, Q., Sun, W., Javaid, A. Y., Devabhaktuni, V. K., “Distributed network traffic feature extraction for a real-time IDS,” in *2016 IEEE International Conference on Electro Information Technology (EIT)*, May 2016, pp. 0522–0526. DOI: 10.1109/EIT.2016.7535295

PRESENTATIONS

7. Adachi, M., Pierce, B. G., **Karimi, A. M.**, Wilson, L. G., French, R. H., Carter, J. L. W., Fukuyama, H., “Nucleation and Growth of AlN: A Case Study of the Challenges in Blending Materials Science and Data Science in an International Collaboration,” presented at the Symposium BI01—Materials Data Science—Transformations in Interdisciplinary Education, Materials Research Society, 2019 Fall Meeting (Boston MA), Dec. 2019
8. Adachi, M., Sonoko, S., Kanbara, A., Wilson, L. G., Pierce, B. G., **Karimi, A. M.**, French, R. H., Carter, J. L., Fukuyama, H., “AlN growth behavior on Ni-Al liquid solutions,” Invited, presented at the 4th International Workshop on Ultraviolet Materials and Devices (Saint Petersburg, Russia), Sep. 8–13, 2019
9. **Karimi, A. M.**, French, R. H., “Invited: Computer Vision and Machine Learning in a Distributed Computing Environment: Photovoltaic Degradation Quantified Using Electroluminescent Images,” presented at the DCMS Materials 4.0 Summer School (TU-Dresden, Dresden Germany), Sep. 10–15, 2018
10. **Karimi, A. M.**, Fada, J. S., French, R. H., “Supervised and Unsupervised Machine Learning of Electroluminescent Images of Photovoltaic Modules,” presented at the PyCon, Python Software Foundation (Cleveland, OH), May 13, 2018
11. **Karimi, A. M.**, Fada, J. S., Liu, J., Braid, J. L., Koyutürk, M., French, R. H., “Supervised and Unsupervised Machine Learning Methods on Photovoltaic Electroluminescence Images to Characterize Degradation,” presented at the Data Science Symposium (Tohoku University, Sendai Japan), Aug. 1, 2018
12. Curran, A. J., Liu, J., **Karimi, A. M.**, Hwang, S. H., Morrison, S. M., Haddadian, R., Braid, J. L., French, R. H., “Inverter level time series analysis of real-world PV performance data,” Poster, presented at the Kyocera Materials Data Science Symposium (CWRU), Jul. 30, 2018
13. **Karimi, A. M.**, Khalilnejad, A., Curran, A. J., Liu, J., Nash, K. J., Fada, J. S., French, R. H., “Energy-CRADLE: A Scalable Infrastructure For Large Scale Distributed Database & Computational Analytics,” (NSF Midwest Big Data Hub Meeting, Omaha, Nebraska), Jun. 2017
14. Kamath, S., Khalilnejad, A., Haddadian, R., **Karimi, A. M.**, Koehrsen, W. J., Blincoe, D. R., Kennedy, Q. C., French, R. H., Abramson, A. R., “A Data Analytics Approach to Identifying Saving Opportunities and Inefficiencies,” Poster Presentation, presented at the Intersections: SOURCE Symposium and Poster Session (Cleveland, OH), Apr. 20, 2018
15. “Module Level Exposure and Evaluation Test (MLEET) for Real-World & Laboratory-Based PV Modules: Common Data and Analytics for Quantitative Cross-Correlation and Validation,” Poster Presentation, presented at the Department of Energy Sun-Shot Poster Crawl, Jun. 2017
16. Curran, A. J., **Karimi, A. M.**, Hu, Y., Haddadian, R., Peshek, T. J., French, R. H., “Month-by-Month Analysis of the Power Change Rate of RealWorld, Outdoor, Photovoltaic Systems Across Multiple ClimateZones,” presented at the CWRU Cyberinfrastructure Day, Apr. 2017
17. Curran, A. J., Liu, J., **Karimi, A. M.**, French, R. H., “Rate Of Change Determination Of Real-world Commercial PV Power Plants using Data-driven Modeling,” presented at the CWRU Research Showcase (CWRU, Cleveland OH), May 10, 2017
18. **Karimi, A. M.**, Ligh, J. K., French, R. H., “E-CRADLE’s Scalable Computing Infrastructure For Large Scale Advanced Computational Analytics,” presented at the CWRU/Tohoku University Data Science Symposium in Life Sciences and Engineering 2017 (Case Western Reserve University, Cleveland OH), Aug. 2017