ZI DENG

dengzikang@gmail.com | 925-408-3088 | github.com/Zi-Deng | San Diego, CA

PhD candidate in Electrical Engineering with expertise in deep learning and computer vision. Experienced in developing ML systems for image classification, species identification, and large-scale data pipelines. Seeking a Summer 2026 internship in machine learning, data science, or applied AI research.

Education

Ph.D. in Electrical Engineering, University of Arizona **2021 – Present**

Minor in *Data Science and Statistics* — Advisor: Dr. Jeffrey J. Rodriguez
Research Focus: *Machine Learning*, *Data Science*, *Computer Vision*

M.S. in Electrical Engineering, University of California, San Diego **2019 – 2021**

Depth in *Signal and Image Processing*

B.S. in Electrical Engineering, University of California, San Diego **2015 – 2019**

HKN Engineering Honors Society

Experience in Academia

University of Arizona

Researcher, Signal and Image Laboratory **Spring 2022 – Present**

- Designed a deep learning framework for **cost-sensitive image classification** that implements configurable N×N cost matrices to optimize for asymmetric misclassification costs in safety-critical domains such as medical imaging (breast cancer detection) and dangerous species identification. The framework introduces a novel loss function utilizing dynamic alpha-scaled **regularization**, and **logit adjustment** and can be integrated into existing vision architectures such as **ViT**, **ResNet** and **ConvNeXt** with ease.
- Developed **FLOW-DC**, a distributed downloader with policy-aware **congestion control** tailored for large-scale acquisition of machine learning datasets under the guidance of Dr. Nirav Merchant. It leverages **TaskVine** with a **BBR** inspired state machine to allow users to utilize a heterogeneous resource pool for machine learning dataset acquisition and processing. It is designed to be an improvement upon existing distributed downloaders such as **img2dataset**.
- Created a set of machine learning models to support applications related to species classification: (1) **AraNet**, which utilizes a two-stage **ConvNeXtV2** design, combining intermediate fine-tuning and downstream distillation to perform species identification, and (2) **OspideR**, an **ensemble OOD detector** for new species detection that constructs a **Mahalanobis distance** based score for each in-distribution class and takes the minimum of that pool to construct an output anomaly score.

Graduate Consultant, Arizona Institute for AI and Society **Winter 2025 – Present**

- Worked as an AI/ML consultant for the **Innovation Pilot** program for University of Arizona's AI Institute. Graduate students, faculty, and staff apply to the program to receive machine learning consultation, expertise and resources for their specific projects. Additionally aided in the planning of methods in which AI could be integrated into the University of Arizona campus at large.

Researcher, Data Science Institute **Spring 2022 – Winter 2025**

- Collaborated with **AI Institute for Resilient Agriculture (AIIRA)** as part of a cross-institutional team to create the following deep vision models: (1) **InsectNet**, which utilizes **SSL pre-training** with a **RegNetY32** architecture for insect and pest classification, and (2) **WeedNet**, which utilizes **Masked Autoencoder ViT** architecture for weed species classification.
- Worked with Dr. Arti Singh and Dr. Baskar Ganapathysubramanian to construct **BioTrove**, an open-access benchmark dataset for machine learning/artificial intelligence applications in biodiversity (161.9 million images), and **TerraIncognita**, a dynamic benchmark dataset to support **OOD detection** for species discovery.

University of California, San Diego

Researcher, Adaptive Computing and Embedded Systems Lab **Sum. 2020 – Spring 2021**

- Evaluated the secure computation library **emp-toolkit** for **multi-party garbled circuit (GC)** in both malicious and semi-honest settings using C. Then, performed investigation into adapting the emp-toolkit's multi-party GC to allow segmented computation of memory-intensive modules.
- Worked with Dr. Karla Esquerre to examine and test different machine learning models for water end-use classification for low-income households, settling on **Random Forest**.

Programmer, Air Sea Interaction Lab — Spring 2017 – Spring 2019

- Part of the Waveglide team, which worked on full-stack development in **Python** for a **Wave Glider**, an unmanned research vehicle designed for ocean wave analysis. Further developed a **MATLAB** script to apply phase unwrapping algorithms to ocean velocity data from an **acoustic doppler current profiler (ADCP)** attached to the vehicle.

Experience in Industry

ViaSat - Carlsbad, CA

Data Science and Machine Learning Intern — Summer 2019

- Utilized **unsupervised learning** techniques to investigate resource optimization possibilities for the **ViaSat-2** satellite network to achieve maximum internet availability. This culminated in the development of a **DiANA** clustering analysis pipeline on **ViaSat-2** historical data using tidyverse philosophy in **R**.

Software Engineering Intern — Summer 2018

- Developed a tool to automatically generate configuration files for hardware components of the **ViaSat-2 SurfBeam2** satellite network. Additionally, created a **Flask** web application for visualization of satellite beam information, which allowed for simplified information sharing between internal groups.

Bbot Robotics - San Mateo, CA

Engineering Intern — Summer 2017

- Worked on the development of **Bbot**, a robotic drink delivery system that operates along a track on the ceiling. Utilized a **Raspberry Pi** to communicate with a **BQ76920** battery balancer chip to control a **PIC** microcontroller to control a prototype charging circuit for the robot's battery using embedded programming in **Python** and **C**.

Key Skills

Languages: *Python, C++, R, SQL, Bash, MATLAB* **Libraries:** *NumPy, Pandas, OpenCV, Matplotlib*
Tools: *AWS, Git, Docker, Linux* **Frameworks:** *PyTorch, scikit-learn, Hugging Face, Kaggle*
Methods: *Computer Vision, VLMs, OOD Detection, Transfer Learning, SSL, Distributed Computing*

Publications

[1] Yang, C.H., Feuer, B., Jubery, Z., Deng, Z.K., Nakkab, A., Hasan, M.Z., Chiranjeevi, S., Marshall, K., Baishnab, N., Singh, A.K., Singh, A., Sarkar, S., Merchant, N., Hegde, C., & Ganapathysubramanian, B. (2024). **BioTrove: A Large Curated Image Dataset Enabling AI for Biodiversity**. *Advances in Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track*.

[2] Deng, Z.K., & Rodriguez, J.J. (2024). **Look Out for Dangerous Spiders: Araneae Classification Using Deep Learning Methods**. *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, 134–137.

[3] Deng, Z.K., Merchant, N., & Rodriguez, J.J. (2026). **FLOW-DC: a Large-Scale Distributed Downloader Utilizing Policy-Aware Adaptive Rate Control**. *IEEE Transactions on Big Data*. (In Preparation)

[4] Feuer, B., Joshi, A., Cho, M., Jani, K., Chiranjeevi, S., Deng, Z.K., Balu, A., Singh, A.K., Sarkar, S., Merchant, N., Singh, A., Ganapathysubramanian, B., & Hegde, C. (2024). **Zero-Shot Insect Detection via Weak Language Supervision**. *The Plant Phenome Journal*.

[5] Chiranjeevi, S., Saadati, M., Deng, Z.K., Koushik, J., Jubery, T.Z., Mueller, D.S., O'Neal, M., Merchant, N., Singh, A., Singh, A.K., Sarkar, S., Singh, A., & Ganapathysubramanian, B. (2025). **InsectNet: Real-Time Identification of Insects Using an End-to-End Machine Learning Pipeline**. *PNAS Nexus*.

[6] Deng, Z.K., & Rodriguez, J.J. (2026). **Spider Species Identification Using Deep Learning Methods**. *Arthropoda*. (Submitted)

[7] Oliveira-Esquerre, K., Mello, M., Botelho, G., Deng, Z.K., Koushanfar, F., & Kiperstok, A. (2021). **Water End-Use Consumption in Low-Income Households: Evaluation of the Impact of Preprocessing on the Construction of a Classification Model**. *Expert Systems with Applications*, 185, 115623.

[8] Chiranjeevi, S., Zaremehrjerdi, H., Deng, Z.K., Jubery, T.Z., Grele, A., Singh, A., Singh, A.K., Sarkar, S., Merchant, N., Greeney, H.F., Ganapathysubramanian, B., & Hegde, C. (2025). **TerraIncognita: A Dynamic Benchmark for Species Discovery Using Frontier Models**. *arXiv preprint*, arXiv:2506.03182.

[9] Shen, Y., Ayanlade, T.T., Boddepalli, V.N., Saadati, M., Rairdin, A., Deng, Z.K., Arshad, M.A., Balu, A., Mueller, D., Singh, A.K., Everman, W., Merchant, N., Ganapathysubramanian, B., Anderson, M., Sarkar, S., & Singh, A. (2025). **WeedNet: A Foundation Model-Based Global-to-Local AI Approach for Real-Time Weed Species Identification and Classification**. *Nature Communications*, (Submitted)

[10] Deng, Z.K., & Rodriguez, J.J. (2026). **Look Out for New Spiders: Araneae Species OOD Detection Using a Pooling Ensemble**. *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, (Submitted)