

ZHENAN YIN

Luddy School of Informatics, Computing, and Engineering | Indiana University Indianapolis

Email: yin10@iu.edu | Tel: +1 (317) 762-1086 | Web: zhenanyin.github.io | LinkedIn: linkedin.com/in/zhenan-yin

EDUCATION

Indiana University Indianapolis

Ph.D. in Informatics (Health & Biomedical Informatics)

Advisor: Dr. Saptarshi Purkayastha | Purkayastha Lab for Health Innovation

Aug 2025 – Present | Expected: 2030

Indianapolis, IN, USA

University of California, San Diego

Bachelor of Science in Electrical Engineering (Machine Learning & Controls)

Sep 2016 – Jun 2019

San Diego, CA, USA

SELECTED HIGHLIGHTS

- **Clinical AI Research:** Develop and evaluate medical imaging foundation-model and multimodal clinical ML pipelines across chest radiographs, EHR, EEG, IMU, discharge notes, scanned PDFs, and time-series vitals.
- **Research Impact:** First-author IEEE ICHI 2026 paper on multimodal EEG-IMU fusion; co-authored IEEE ICHI 2026 AgentDS Healthcare paper; second-author Medical Image Analysis manuscript under review on frozen CXR foundation models; contributing author on Nature Biomedical Engineering publication.
- **Deployment Readiness:** Bring 5 years of enterprise healthcare-IT experience designing, deploying, and securing clinical information systems, telemedicine platforms, EHR integrations, and high-availability medical data infrastructure.

RESEARCH EXPERIENCE

Tri-Modal Fusion for Parkinson's Disease Motor Phenotype Classification

Graduate Researcher, Indiana University Indianapolis

Apr 2026 – Present

Indianapolis, IN

- Developed a hierarchical multimodal pipeline to classify 130 participants into Healthy Control, Tremor-Dominant PD, Intermediate PD, and PIGD PD using resting EEG, walking EEG, and clinical features derived from MDS-UPDRS Part III TD/PIGD ratios.
- Built a two-stage architecture combining an XGBoost HC-vs-PD gate, EEGNet-Transformer rest/walk encoders, bidirectional EEG cross-attention, and clinical XGBoost soft fusion for within-PD subtype classification.
- Improved 4-class LOGO macro-F1 by approximately +0.10 over flat trimodal cross-attention, validated with LOSO/LOGO evaluation, multi-seed robustness testing, bootstrap 95% confidence intervals, and Wilcoxon signed-rank tests.
- Achieved 0.56 ± 0.02 macro-F1, 0.98 ± 0.01 HC-vs-PD gate AUC, and 0.43 ± 0.02 PD-only 3-class F1; used SHAP TreeExplainer to identify Color Trails Test and MoCA as dominant within-PD predictors.

TriageGuard: LLM-Driven Continuous Security Testing for Open-Source EHRs

Graduate Researcher, Indiana University Indianapolis

Jan 2026 – Present

Indianapolis, IN

- Developed TriageGuard, a PR-triggered OpenMRS security-testing pipeline that converts code diffs into executable Gherkin/Playwright tests and validates them against a Dockerized OpenMRS instance through GitHub Actions CI.
- Built a dual-LLM safety gate using Qwen-3-32B generation, GPT-OSS-120B review, SBERT semantic verification, and 13-condition approval logic to block unsafe AI-generated tests before execution.
- Integrated CodeQL static analysis with 4 custom queries covering unprotected methods, privilege annotations, HTTP entry points, and caller chains; grounded test generation with an 8-profile Gherkin scenario library.
- Operationalized CVSS 4.0 risk-band scoring and a 4-signal composite risk model combining evidence density, OWASP pattern matching, CodeQL corroboration, and model self-reporting, with SHA-256 provenance for every LLM call.

Multimodal EEG-IMU Fusion for Robust Motor Assessment
Graduate Researcher, Indiana University Indianapolis

Sep 2025 – Apr 2026
Indianapolis, IN

- Developed a multimodal motor-assessment framework integrating 16-channel EEG and wrist-worn IMU signals across 10 motor activities to evaluate neural and kinematic modality complementarity.
- Implemented an EEGNet-Transformer model with approximately 230K parameters and a feature-engineered XGBoost IMU pipeline using 76 features per sensor across time-domain, frequency-domain, and 3D composite groups.
- Achieved $98.68 \pm 0.32\%$ activity-classification accuracy through logistic-regression late fusion, reducing EEG-only error by 81.5%, improving worst-task accuracy from approximately 87% to 96.76%, and reducing cross-task variance from approximately 3% to 1.06%.
- Characterized task-dependent modality complementarity, showing IMU dominance on 7 of 10 kinematically distinctive activities and 3x EEG error reduction for rhythmic cycling, consistent with beta-band sensorimotor coupling.

Human-Guided Agentic AI for Multimodal Clinical Prediction

Oct 2025 – Jan 2026

Graduate Researcher, Indiana University Indianapolis | AgentDS Healthcare Benchmark *Indianapolis, IN*

- Contributed to an AgentDS Healthcare benchmark submission using structured EHR data, discharge notes, scanned PDF billing receipts, and time-series vitals; placed 5th overall in the healthcare domain and 3rd on discharge-readiness prediction.
- Led 30-day readmission prediction across 5,000 admissions, engineering 871 structured-EHR and discharge-note features from admission records, demographics, 5-year ED cost history, and trigram TF-IDF notes.
- Built a 5-base-model stacking ensemble using 3 XGBoost variants, L1/L2 logistic regression, and 8-fold out-of-fold meta-learning; achieved test Macro-F1 = 0.8986, within 0.006 of the top-ranked system.
- Extracted scanned-PDF billing features with PyPDF2 and regex for 3-year ED cost forecasting, achieving test MAE = \$465.13; counterfactual ablations showed +0.065 cumulative F1 from human-guided decisions.

Frozen Foundation-Model Embeddings for Small-Lesion CXR Analysis

Aug 2025 – Mar 2026

Graduate Researcher, Indiana University Indianapolis

Indianapolis, IN

- Co-investigated how small-scale, low-contrast diagnostic signal is preserved or lost in frozen Vision Transformer foundation-model embeddings for chest radiography, addressing a pre-deployment reliability gap in medical imaging AI.
- Built an embedding-probing pipeline across 5 frozen ViTs—RAD-DINO, DINOv2-B/14, DINOv3 ViT-7B, BiomedCLIP, and MedSigLIP—plus a DINO-pretrained ResNet-50 architectural control, evaluated on NIH-CXR14, MIMIC-CXR, Emory-CXR, and ChestX-Det10 cohorts.
- Evaluated 492,724 chest radiographs and 3,543 ChestX-Det10 images with 1,462 lesion boxes across calcification, nodule, and mass categories.
- Designed a physics-motivated perturbation panel under 0.4% of image area and compared CLS, patch-mean, and bounding-box ROI patch pooling on matched frozen forward passes.
- Demonstrated that CLS pooling collapses small-scale signal to chance-level AUC = 0.500–0.524, while ROI-conditioned patch tokens recover approximately AUC = 1.0 with mean per-model gains of +0.412 to +0.488.
- Validated the mechanism on real lesions, showing small-vs-large CLS AUC gaps up to +0.243 while ROI patch pooling recovered AUC ≥ 0.899 across every model-class cell and eliminated the size-stratified disadvantage.

Host–Pathogen Metabolic Modeling of Mtb in Lung Cancer Progression

Jul 2024 – Aug 2024

Summer Research Assistant (Remote), Pembroke College, University of Cambridge *Cambridge, UK*

- Developed constraint-based systems-biology models of Mycobacterium tuberculosis-macrophage interactions to investigate infection-driven metabolic shifts and their potential contribution to lung cancer progression.
- Deployed genome-scale metabolic models, including iAB-AMØ-1410 and iAB-AMØ-1410-Mt-661, in COBRApy with Flux Balance Analysis and Flux Variability Analysis for host-pathogen flux mapping.

- Identified 223 infection-specific essential reactions and quantified nutrient dependencies through medium minimization, surfacing candidate metabolic drug targets for downstream experimental validation.
- Simulated infection-induced metabolic burden, demonstrating a 92% reduction in macrophage biomass growth rate under Mtb-driven energetic constraints.

Triton Town Autonomous Vehicles

Feb 2019 – Jun 2019

Undergraduate Researcher, University of California, San Diego

San Diego, CA

Advisor: Prof. Maurício de Oliveira & Prof. Jack Silberman

- Built a deep-learning autonomous vehicle on Raspberry Pi using TensorFlow, OpenCV, and the DonkeyCar framework for real-time lane detection, steering prediction, and obstacle avoidance with sub-150 ms inference latency on edge hardware.
- Curated and trained CNNs on 35,000+ labeled image frames while integrating sensors, motor drivers, and electronic control modules into a reliable embedded platform.
- Reduced vehicle failure rate from 25% to 3% through iterative testing and model optimization, enabling fully autonomous operation across 10 indoor and 6 outdoor laps.

Deep Learning for Plankton Image Retrieval

Nov 2017 – May 2019

Undergraduate Research Assistant, University of California, San Diego

San Diego, CA

Advisor: Prof. Nuno Vasconcelos | Statistical Visual Computing Lab (SVCL)

- Developed a deep learning-based image labeling and retrieval system to support large-scale plankton annotation for ecological research at the Scripps Institution of Oceanography.
- Applied transfer learning with ResNet-50 and InceptionResNetV2, integrating approximate nearest neighbor (ANN) search for similarity retrieval, and prototyped a web-based interface to accelerate human-in-the-loop annotation workflows.
- Boosted fine-grained plankton classification accuracy to 90% by expanding dataset diversity through targeted data-augmentation strategies.

Conformal Ultrasonic Device for Non-Invasive Central Blood Pressure Monitoring

May 2017 – Dec 2018

Undergraduate Research Assistant, University of California, San Diego

San Diego, CA

Advisor: Prof. Sheng Xu | Xu Research Group

- Fabricated a stretchable conformal ultrasonic probe using laser micro-cutting and soft elastomer encapsulation for non-invasive cardiovascular monitoring.
- Modeled ultrasound beam patterns and penetration depth in MATLAB, validating simulation results against experimental measurements.
- Evaluated device stability and acoustic coupling under mechanical deformation to support signal fidelity during motion.
- Prepared IRB documentation and risk assessments for human-subject testing and contributed to the resulting Nature Biomedical Engineering publication.

PUBLICATIONS

Peer-Reviewed Journal Publications

Wang C, Li X, Hu H, Zhang L, Huang Z, Lin M, Zhang Z, **Yin Z**, Huang B, Gong H, Bhaskaran S, Gu Y, Makihata M, Guo Y, Lei Y, Chen Y, Wang C, Li Y, Zhang T, Chen Z, Pisano AP, Zhang L, Zhou Q, Xu S. Monitoring of the central blood pressure waveform via a conformal ultrasonic device. *Nature Biomedical Engineering*. 2018 Sep;2(9):687–695. doi:10.1038/s41551-018-0287-x. PMID:30906648; PMCID:PMC6428206.

Conference Publications

Yin Z, Pulavarthy LP, Purkayastha S. Multimodal EEG-IMU Fusion for Motor Assessment: Leveraging Task-Dependent Complementarity for Robustness. *IEEE International Conference on Healthcare Informatics (ICHI)*, 2026.

Pulavarthy LP, Muthyala R, Kuruvikkattil AV, **Yin Z**, Kudamala R, Purkayastha S. Human-Guided Agentic AI for Multimodal Clinical Prediction: Lessons from the AgentDS Healthcare Benchmark. *IEEE International Conference on Healthcare Informatics (ICHI), Data Challenge Program: AgentDS Healthcare*, 2026.

Manuscripts Under Review

Muthyala R, **Yin Z**, Jilla A, Li F, Dapamede T, Khosravi B, Chavoshi M, Gichoya J, Purkayastha S. Frozen Foundation-Model Embeddings Discard Small-Lesion Signal in Chest Radiography: Implications for Pre-Deployment Evaluation. *Medical Image Analysis*. Under review, 2026.

TECHNICAL SKILLS

Programming & Data: Python, R, MATLAB, SQL, C/C++, NumPy, pandas, LabVIEW

Machine Learning & Deep Learning: PyTorch, TensorFlow/Keras, scikit-learn, XGBoost, Hugging Face Transformers, CNNs, Vision Transformers, EEGNet, transfer learning, fusion, stacking ensembles

Clinical AI: Chest radiography foundation models, EHR data, EEG/IMU signal processing, ultrasound signal processing, MNE-Python, OpenCV

Evaluation: LOSO/LOGO CV, nested CV, bootstrap CIs, Wilcoxon tests, SHAP, occlusion analysis, calibration analysis

Engineering/MLOps: Docker, GitHub Actions, CUDA, Linux, Git, SQLite, Playwright, Pytest-BDD, CodeQL, OpenMRS

PROFESSIONAL EXPERIENCE

Healthcare IT Solution Team Lead
China Mobile Yunnan Co., Ltd.

Mar 2024 – Jul 2025
Kunming, Yunnan, China

- Led 5+ concurrent healthcare IT implementation projects across Yunnan Province, coordinating engineering teams, hospital administrators, and third-party vendors to deliver clinical information systems.
- Directed EHR system integration for client hospitals, aligning data interoperability, implementation continuity, and clinical workflow requirements.
- Improved project delivery rate from 75% to 90% by introducing milestone tracking, risk forecasting, and performance-monitoring workflows.
- Supported design and deployment of regional telemedicine and Internet Hospital platforms on China Mobile Cloud, reducing onboarding costs for smaller healthcare institutions by 55% while maintaining healthcare data-governance compliance.

Healthcare IT Project Solution Manager
China Mobile Yunnan Co., Ltd., Kunming Branch

Oct 2019 – Feb 2024
Kunming, Yunnan, China

- Managed end-to-end delivery of large-scale healthcare digital transformation projects totaling 21 million CNY, spanning medical networks, healthcare data platforms, and clinical information systems.
- Designed the VPDN Medical Insurance Account Management System, producing technical specifications and functional prototypes that improved delivery efficiency by 15% and achieved national software copyright certification.
- Led deployment of the Fangcang Shelter Hospital medical data transmission network during the COVID-19 response, completing implementation 10 days ahead of schedule while reducing labor costs by 8%.
- Collaborated with hospitals, insurers, and telecom engineering teams to deliver secure, high-availability medical data transmission infrastructure for regional healthcare operations.

PRESENTATIONS & CONFERENCES

Oral Presentation, First Author: **Yin Z**, Pulavarthy LP, Purkayastha S. “Multimodal EEG-IMU Fusion for Motor Assessment: Leveraging Task-Dependent Complementarity for Robustness.” *IEEE International Conference on Healthcare Informatics (ICHI)*, 2026.

Oral Presentation, Co-Author Presenter: Pulavarthy LP, Muthyala R, Kuruvikkattil AV, **Yin Z**, Kudamala R, Purkayastha S. “Human-Guided Agentic AI for Multimodal Clinical Prediction: Lessons from the AgentDS Healthcare Benchmark.” *Data Challenge Program @ IEEE ICHI 2026—AgentDS Healthcare*, Minneapolis, MN, Jun 2026.

Oral Presentation, Co-Author: **Yin Z**. “Construction of a Hospital-Wide Critical Illness Rescue Network Based on 5G for Severe Illness Early Warning Platform.” Finalist, Top 30, *5G+ Healthcare Special Competition Finals, 7th Blooming Cup 5G Application Competition*, Sep 19, 2024; hosted by China Academy of Information and Communication Technology (CAICT) and China Communication Standards Association (CCSA).

PATENTS & INTELLECTUAL PROPERTY

China Mobile VPDN Management System. National Software Copyright Registration, China, 2023. Certificate No. 10834368; Registration No. 2023SR0247197.

AWARDS & CERTIFICATIONS

University Fellowship, Indiana University, 2025.

Qualification of Telecommunication Professional, Ministry of Human Resources and Social Security and Ministry of Industry and Information Technology, China, 2020.

Certified LabVIEW Associate Developer (CLAD), National Instruments, 2018.