

Van Phan

+1(470) 350 - 7733 | vanthienphan2004.work@gmail.com | [Personal Portfolio](#) | [Google Scholar](#) | [LinkedIn](#) | [GitHub](#)

EDUCATION

Georgia State University

Atlanta, GA

B.S. Mathematics, GPA: 4.11/4.30

Expected Graduation: May. 2027

- *Relevant Coursework:* Linear Algebra, Differential Equations, Numerical Analysis, Analysis, Modern Algebra, Optimization, Calculus, Biostatistics, Probability and Statistics, Applied Combinatorics, Computer Science

RESEARCH EXPERIENCE

Density and BI-RADS - Aware Triage and Automatic Report Generation for Mammography

Advisor: Dr. Russell Jeter

- Developed the Density and BI-RADS-Aware Triage and Report Generation (DB-ATRG) framework to automate diagnostic text generation and enable risk-based case prioritization, flagging scans with extremely dense breasts (ACR Category D) for supplemental screening.
- Fine-tuned the 4-billion parameter MedGemma 1.5 vision-language model using QLoRA on the DMID and VinDr-Mammo dataset, achieving an ACR Density Accuracy of 0.7039, a ROUGE-L score of 0.8650, and a METEOR score of 0.9001 for clinical report generation.
- Accelerated clinical review in simulations on a 100-case cohort by applying a Cumulative Urgency Score, capturing 100% of high-risk malignancies (BI-RADS 4 and 5) within the first 20% of the reading workload (compared to 40% in FIFO) and shifting the mean rank of severe cases from 42.8 down to 3.

Time Series Classification of Active Range of Motion in Robotics-Assisted Stroke Rehabilitation

Advisor: Dr. Russell Jeter

- Developed supervised machine learning models to classify active range of motion (ROM) in stroke rehabilitation, leveraging 28M+ high-resolution time-series samples from 431 users collected through robotics-assisted in-home therapy.
- Evaluated 11 ML architectures, including XGBoost, LSTM, 1D-CNN, Transformer, InceptionTime, Temporal Convolutional Networks (TCN), and TSFresh-based models, achieving 43.2% overall accuracy with TCN on a standardized 100,000-window evaluation set.
- Improved movement-pattern classification by engineering time-series features from Angle and Pressure sensor data, with TSFresh-based embeddings achieving 1.58-bin MAE and LSTM reaching 0.670 AUROC, while identifying class imbalance as a key modeling challenge.

Modeling Wound Healing Without Migration and Proliferation

Advisor: Dr. Yi Jiang

- Established a mathematical framework for Age-related Macular Degeneration (AMD) therapeutics, addressing a leading cause of vision loss by modeling Retinal Pigment Epithelium (RPE) wound healing.
- Simulated purse-string contraction and cell fusion repair mechanisms in Python-based CompuCell3D, by utilizing a Q-Potts cellular model and Monte Carlo evaluation of the Hamiltonian free energy states.
- Quantified wound closure efficiency across algorithmic iterations by analyzing texture tensors and spatial points, proving that lower cellular volume constraints, higher wound's radius and affected area yield statistically slower wound closure.

Morphological Feature Analysis of Retinal Pigment Epithelial Cell from Mice during Aging

Advisor: Dr. Yi Jiang

- Developed an automated end-to-end machine learning pipeline for classifying Retinal Pigment Epithelial (RPE) cell during aging, generating insights for potential therapeutic strategies in Age-related Macular Degeneration.
- Achieved highly accurate classification as measured by 90%+ cross-validation F1-score, by extracting 133 morphological and texture-based features from 326 RPE cell images and implementing a stacking ensemble of XGBoost, LightGBM, and CatBoost with a Logistic Regression meta-learner.
- Identified critical RPE aging biomarkers, as demonstrated by top feature importance rankings, suggesting Local Binary Patterns (LBP) and geometric shape metrics (circularity, aspect ratio) as the most influential classification variables.

Enhanced Physics-Informed Neural Networks for Collective Cancer Invasion

Advisor: Dr. Yi Jiang

- Engineered a Physics-Informed Neural Networks (PINNs) model to simulate collective cancer invasion, resulting in a data-efficient, mesh-free solution for complex Partial Differential Equations (PDEs).
- Developed a framework utilizing Tensorized Fourier Neural Operators (TFNO) and PyTorch's autograd engine to encode physical laws (PDEs, boundary, and initial conditions) directly into the model's loss function, reducing reliance on labeled data.
- Explored advanced techniques, including a sequence-to-sequence (Seq2seq) PINO approach and Augmented Lagrangian methods (ALM), to overcome model failure modes like vanishing gradients and improve solution accuracy and stability.

Laboratory Rearing Model of *Lissachatina fulica* for Acharan Sulfate Extraction & Evaluation

- Established a laboratory rearing model of *Lissachatina fulica*, using garlic essential oil and *Lactobacillus casei* for inhibition of pathogenic bacteria (*E. coli*, *S. aureus*, and *Salmonella spp.*).
- Extracted acharan sulfate of snail flesh and mucus from *Lissachatina fulica*.
- Investigated anticoagulant and thrombolytic properties of extracted acharan sulfate, achieving 89.40% anticoagulant activity within 30 minutes via clotting time assay and 32.09% thrombolytic activity within 1 hour via clot lysis assay.

PUBLICATIONS

- [In Review] DB-ATRG: The Density and BI-RADS-Aware Triage and Automatic Report Generation System for Mammography, doi: 10.64898/2026.07.22.26358655.

POSTERS & PRESENTATIONS

Density and BI-RADS - Aware Triage and Automatic Report Generation System for Mammography,

- Poster presented at **SIAM Conference on Mathematics of Data Science**, 2026.
- Poster presented at Georgia State Undergraduate Research Conference, 2026.

Enhanced Physics-Informed Neural Networks for Collective Cancer Invasion,

- **Best Project Award** at Summer Undergraduate Research Symposium, 2025.

Morphological Feature Analysis of Retinal Pigment Epithelial Cell from Mice during Aging,

- Poster presented at Georgia State Undergraduate Research Conference, 2025.

Modeling Wound Healing Without Migration and Proliferation,

- Poster presented at Summer Undergraduate Research Symposium, 2024.

WORK EXPERIENCE

Viet Home Care LLC

Jun. 2024 – Sep. 2026

Data Scientist

Atlanta, GA

- Developed data pipelines integrating Vision-Language Models (VLMs) to extract and structure information from health records, health coaching sessions, and home health visits, transforming unstructured healthcare data into analysis- and model-ready datasets.
- Built machine learning models to optimize personal care aide matching based on client needs and compatibility across a pool of 1,000+ clients and aides.
- Analyzed service utilization and operational data across several services, generating actionable insights to improve service processes, quality, and outcomes for 1,000+ households.

Georgia State University

Aug. 2025 – Aug. 2026

Chemistry Laboratory Teaching Assistant

Atlanta, GA

- Mentored 50+ undergraduate students in Survey of Chemistry I and II, improving understanding of fundamental principles and reducing procedural errors through one-on-one guidance and demonstration of proper laboratory techniques.

Center for the Advancement of Students and Alumni

Jun. 2024 – Aug. 2025

Near-peer Mentor

Atlanta, GA

- Provided near-peer mentorship and research support to students in the Math Path Program, guiding research technique implementation and project development for the Enhanced Physics-Informed Neural Networks for Collective Cancer Invasion project.

HONORS & AWARDS

- **Best Project Award** at Summer Undergraduate Research Symposium (2025)
- **President's List** at Georgia State University (2023, 2024, 2025, 2026)
- **Prospect Award (National Level) & First Place (City Level)** at Vietnam Science and Engineering Fair (ViSEF) (2022)
- **Consolation Prize** at Engineering and Science Fair of Ho Chi Minh City University of Technology (2022)
- **Bronze Prize** in Physics at the Ho Chi Minh City Excellent Student Contest (2019)

TECHNICAL SKILLS

- **Programming & Tools:** Python, Julia, R, MATLAB, SQL, Git/GitHub, Linux/CLI, LaTeX, CompuCell3D
- **Scientific & Math Modeling:** Physics-Informed Neural Networks, PDE Modeling, Tensorized Fourier Neural Operators, Augmented Lagrangian Method, Q-Potts Cellular Modeling, Monte Carlo Evaluation
- **ML, LLM & AI:** Supervised Learning, Unsupervised Learning, Ensemble Methods, LLMs, Fine-Tuning, Quantization, Transformers, TRL, Hugging Face
- **Libraries:** PyTorch, TensorFlow, Scikit-learn, XGBoost, LightGBM, CatBoost, TSFresh, OpenCV, Pandas, NumPy, SciPy, Scikit-image, Pillow, Mahotas, Matplotlib, Seaborn

REFERENCES

Russell Jeter, Ph.D.

Assistant Professor of Math. & Stat. Foundations of Big Data
Department of Mathematics and Statistics, Georgia State University
Email: rjeter@gsu.edu

Yi Jiang, Ph.D.

Frady Whipple Professor
Department of Mathematics & Statistics, Georgia State University
Email: yjiang12@gsu.edu