

Shashi Bhushan Jha

(386) 341-4293 | jhas1@my.erau.edu | <https://www.linkedin.com/in/shashi-jha/>

SHORT BIOGRAPHY

- **Assistant Professor** of Computer Science at **University of West Florida**
- **Research interests:** Deep learning, Machine Vision Inspection, Defect Detection, Sustainable Computing, Cybersecurity, Optimization, and Graph ML.
- **8+** years of research experience as a Graduate Research Assistant or higher (**Google scholar citation: 210**).
- **4** years of teaching experience, responsible for preparing contents, **delivering lectures** in-person/remote, holding office hours 2 times per week, and grading assignments.
- **5** years of academic and industry experience in Machine Learning, Deep learning, and Data Science.
- Proficient in Machine Learning and Optimization.
- Recently designed and developed a deep CNN-based hybrid model for an Automated Optical Inspection System for Aerospace Components.
- ML Tools and Frameworks: Tensorflow, Keras, PyTorch, PyG, PBG, Scikit-learn, Numpy, Pandas, Google Cloud Platform, Vertex AI, Container Registry, Cloud Storage, Snowflake, DataGrip, Visual Studio, GitHub, Docker.
- Programming Language and OS: Python, SQL, C/C++, Java, Matlab; Mac OS, Linux, Windows.

EDUCATION

Embry-Riddle Aeronautical University (ERAU) Daytona Beach, FL
Ph.D. in **Computer Science** Jan. 2019 – June. 2023

- GPA: 4.0/4.0

Indian Institute of Technology Kharagpur (IIT Kharagpur) Kharagpur, India
M.S. in **Industrial and Systems Engineering** Sept. 2015 – Sept. 2018

- GPA: 3.6/4.0

Nagpur University Nagpur, India
B.Eng. in **Computer Science** Aug. 2009 – July 2013

- GPA: 3.0/4.0

Relevant Coursework in PhD: CS 595A: Machine Learning, MA 505: Statistics and Probability, CS 450: Database and Data Retrieval, MA 510: Optimization, CS 800: Deep Learning, Dissertation (Deep CNN Based automated optical inspection system for aerospace components).

RESEARCH/TEACHING EXPERIENCE

Assistant Professor, University of West Florida Aug. 2024 – Present

- Teaching and Research (tenure-track)

Ajunct Faculty, University of West Florida June 2024 – Aug. 2024

- Taught graduate course (Data Mining)

Visiting Professor, College of Charleston Sept. 2023 – May 2024

- Extending the dissertation work of deep CNN-based automated optical inspection for aerospace components to address the imbalanced and limited dataset problem.
- Unsupervised pixel-level segmentation techniques to detect defects in images.

Doctoral Researcher, ERAU Jan. 2019 – Aug. 2023

Dissertation work (Deep CNN-Based Automated Optical Inspection for Aerospace Components)

- Conducted a literature review of machine vision techniques for automated optical inspection systems.
- Acquired **novel aerospace composite components image dataset** from the aerospace research lab, and labeled the image dataset.
- Designed and implemented a **new hybrid deep CNN-based models**, utilizing MobileNetV2 and Random Forest, for an automated optical inspection system to identify defects for aerospace components.
- Trained and evaluated **pre-trained, fine-tuned, and end-to-end deep CNN supervised and unsupervised models** using an image dataset of aerospace composite materials.
- The defect detection experiments are examined using **five-fold cross validation techniques** considering **different set of datasets**.
- Addressed the **imbalanced and limited dataset** problem of aerospace composite components using augmenting the dataset with **improved DCGAN model**.

Projects Involved in PhD:

- Improved CNN Model Performance by Incorporating Image Gradients as Secondary Input:
 - **Proposed a new CNN architecture** that utilizes multiple forms of input, including original images and gradients, by sharing layers across all input forms.
 - **Incorporated image gradient as a secondary input** associated with the original input image.
 - Trained and tested the proposed model using diverse datasets (e.g., MNIST, CIFAR10, CIFAR100) and **achieved better performance than benchmark models.**
- Flight Delay Prediction Using Hybrid Machine Learning Approach: A Case Study of US Airlines:
 - Collected and preprocessed **27 months of recent flight data of US Airlines.**
 - Conducted **data analysis** and **visualization** on the dataset.
 - **Proposed a novel hybrid machine learning approach** that combines deep learning and random forest/XGBoost to predict flight delay in terms of departure, arrival, and total delay.
- Housing Market Prediction
 - Collected and preprocessed housing data from the Florida Volusia County Property Appraiser database, including socio-economic factors such as GDP, CPI, PPI, HPI, EFR.
 - Developed a housing price prediction model using XGBoost, Lasso, Decision Tree, and Random Forest.
 - **Proposed a prediction model** that classifies whether a closing price is greater or less than the listing price using XGBoost, Random Forest, Voting Classifier, SVM, and Logistic Regression.
- **Stock Market Price Prediction Using Multiple Linear Regression Model** (in a Statistics Course Project).
- Formal Modeling of **Cyber-Physical Resource Scheduling** in IIoT Cloud Environments:
 - Used discrete state-machine diagrams to model resource reliability and availability status, and logistics timing purposes.

Graduate Research Assistant, IIT KHARAGPUR

Sept. 2015 – Sept. 2018

Thesis: A Multi-objective Meta-heuristic Approach for TNDFS Problem

- Studied a multi-objective transit network design and frequency setting (TNDFS) problem with the aim of determining a set of routes and frequency on each route for public buses.
- Generated a set of routes using a **novel initial route set generation (IRSG) procedure combined with genetic algorithm.**
- Formulated a **multi-objective model** to assign the frequency on each route with the objectives to minimize the passenger time and the operating cost simultaneously.
- Developed several **multi-objective algorithms**, such as NSGA-II and MOPSO, to solve the problem.

TEACHING EXPERIENCE

Course Instructor	University of West Florida
<i>CAP 4770: Data Mining, CS Department</i>	Fall 2024
<ul style="list-style-type: none">▪ Preparing course contents, delivering lectures in-person and online, and class size: 40▪ Holding office hours once per week	
<i>COP 3530: Data Structures and Algorithms, CS Department</i>	Fall 2024
<ul style="list-style-type: none">▪ Preparing course contents, delivering lectures in-person, and class size: 40▪ Holding office hours once per week	
Course Instructor	University of West Florida
<i>CAP 6771: Data Mining, CS Department</i>	Summer 2024
<ul style="list-style-type: none">▪ Preparing course contents, delivering lectures online, and class size: 16▪ Holding office hours once per week	
Course Instructor	College of Charleston
<i>CSCI 340: Operating Systems, CS Department</i>	Fall 2023
<ul style="list-style-type: none">▪ Preparing course contents, delivering lectures in-person, and class size: 25▪ Holding office hours 2 times per week▪ Received high scores in teaching evaluation survey.	
<i>CSCI 221: Computer Programming II, CS Department</i>	Fall 2023
<ul style="list-style-type: none">• Two sections: 25 students in each section (Fall 2023)▪ Preparing course contents and delivering lectures in-person and holding office hours	

- DATA 221: **Applied Data Analytics** Spring 2024
- One section: 25 students
 - Preparing course contents, delivering lectures in-person and holding office hours
- CSCI 221: **Computer Programming II** Spring 2024
- **Three sections:** 25 students in each section
 - Preparing course contents and delivering lectures in-person and holding office hours

Course Instructor

SYS 301: Introduction to Systems Engineering, EECS Department (Fall course) ERAU
Aug. 2019 – Dec. 2021

- Taught **46 undergraduate** engineering students in a section
- Prepared course contents, **delivered lectures in-person/remote**, and held **office hours 2 times per week**

SYS 303: Optimization in Systems Engineering, EECS Department (Spring course) Jan. 2020 – May 2022

- Taught undergraduate engineering students
- Prepared course contents, **delivered lectures in-person/remote**, and held **office hours 2 times per week**
- **Received high scores** in the teaching evaluation survey

Grader

CEC 500/SYS 560: Engineering Project Management/Systems Engineering Management ERAU
Jan. 2019 – May 2019

- Grader for 35 students (interacted with students on request)

INDUSTRY EXPERIENCE

Machine Learning Modeling/Engineer, Cash App @ Block Sept. 2022 – Dec. 2022

- Collaborated within a workstream alongside a Machine Learning Engineer/Modeler and cross-functional partners.
- Conducted analysis on **large datasets** using **Python/MySQL** and scripting languages to **uncover significant insights**.
- Provided **quantitative assessment** of the advantages and disadvantages of different model classes of Graph ML
- Managed a massive dataset (**billions of records**) for GML using **machine type of 4 TB RAM and 160 CPUs**.
- Worked on PyTorch BigGraph (**PBG**) using a distributed system.
- Methodically tested and validated advanced ML models.
- Utilized statistical and mathematical modeling techniques to study and comprehend **customer behavior**.
- **Developed code for efficient data processing, cleansing, and integration** across multiple sources, and **deployed in a production environment**.

AWARDS AND ACTIVITIES

- **Outstanding Bachelor Student Award**, *Department of Computer Science*, Nagpur University Apr. 2010
- **Outstanding Contribution Award** in reviewing by CAIE Journal Dec. 2016
- I have **reviewed 100 research articles of renowned journals** like, CAIE, Information Sciences, IEEE Transactions on SMCS, etc.
- I have **editorial assistance experience** of high impact factor journals such as Information Sciences (Elsevier), International Journal of Systems Sciences (Taylor & Francis), Computers & Industrial Engineering (Elsevier), International Journal of Production Research (Taylor & Francis), etc.

SKILLS

- **Machine Learning Expertise:** Computer Vision, Deep Learning, Machine Learning, Transfer Learning, GML
- **ML Frameworks & Tools:** Scikit-learn, Tensorflow, Keras, PyTorch, PyG, Numpy, Pandas, Jupyter Notebook
- **Programming:** Python, Matlab, SQL, PostgreSQL, C, C++
- **Data Analysis:** Statistical and Probability Models, Neumerical Methods for Optimization.
- **Languages:** English (Professional Proficiency) , Hindi (Professional Proficiency), Maithili (Native)

COURSERA CERTIFICATE:

Deep Learning Specialization by Prof. Andrew Ng, which includes 5 deep learning courses:

- (1) Neural Networks and Deep Learning.
- (2) Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization.
- (3) Structuring Machine Learning Projects.
- (4) Convolutional Neural Networks.
- (5) Sequence Models

PUBLICATION

Journal papers (Peer-reviewed articles):

- **Jha, S. B.**, Babiceanu, R. F.(2023). Deep CNN-Based Visual Defect Detection: Survey of Current Literature. *Computers In Industry*, 148, 103911. (**Impact Factor: 11.24**)
- **Jha, S. B.**, Babiceanu, R. F., Shekhar, P., Namilae, S.(2024). Deep CNN-Based Automated Optical Inspection for Aerospace Components. (Under Review).
- **Jha, S. B.**, Jha, J. K., & Tiwari, M. K. (2019). A multi-objective meta-heuristic approach for transit network design and frequency setting problem in a bus transit system. *Computers & Industrial Engineering*, 130, 166-186. (**Impact Factor: 7.18**)
- **Jha, S. B.**, Babiceanu, R. F., & Seker, R. (2019). Formal modeling of cyber-physical resource scheduling in IIoT cloud environments. *Journal of Intelligent Manufacturing*, 1-16. (**Impact Factor: 7.13**)
- Mogale, D. G., Lahoti, G., **Jha, S. B.**, Shukla, M., Kamath, N., & Tiwari, M. K. (2018). Dual market facility network design under bounded rationality. *Algorithms*, 11(4), 54. (**Impact Factor: 3.3**)
- **Jha, S. B.**, Pandey, V., Jha, R. K., & Babiceanu, R. F. (2020). Machine Learning Approaches to Real Estate Market Prediction Problem: A Case Study. *arXiv preprint arXiv:2008.09922*.
- **Jha, S. B.**, Babiceanu, R. F., Pandey, V., & Jha, R. K. (2020). Housing Market Prediction Problem using Different Machine Learning Algorithms: A Case Study. *arXiv preprint arXiv:2006.10092*.
- **Jha, S. B.**, Babiceanu, R. F., Pandey, V., & Jha, R. K. (2024). To Predict the Flight Delay using Novel Hybrid Machine Learning Approach: A Case Study of US Airlines. (To be submitted).
- Pandey, V., & **Jha, S. B.** (2020). Incorporating Image Gradients as Secondary Input Associated with Input Image to Improve the Performance of the CNN Model. *arXiv preprint arXiv:2006.04570*.

Conferences:

- **Jha, S. B.**, Jha, J. K., Tiwari, M. K., (2016) “A multi-criteria Round-Based algorithm for Public Transit Routing – RAPTOR”, International conference on E-Business and Supply Chain Competitiveness, IIT Kharagpur, India.