

SHAYAN A. AKBAR

(765) · 409 · 7789 ◊ sakbar@purdue.edu ◊ Indiana, USA ◊ Visa Status: F-1

PHD CANDIDATE AT PURDUE UNIVERSITY

- **Areas of interest:** information retrieval, machine/deep learning, software engineering, and computer vision
- **Papers published in:** ICRA, WACV, CVPRW, and MSR conferences, and JFR and JSEP journals

EDUCATION

PhD, ECE, Purdue University, IN, USA **June 2015 - present**
Thesis: Source Code Retrieval from Large Software Repositories
Advisor: Prof Avinash Kak

MS, ECE, Purdue University, IN, USA **August 2013 - May 2015**
Areas: Computer vision, 3-D modeling, image processing

TECHNICAL STRENGTHS

Languages	Python (over 7 years), Java, MATLAB, C/C++, Scheme, CUDA, SQL
Python Tools:	Numpy, Scipy, Matplotlib, Pandas, Multiprocessing, BeautifulSoup
Machine/Deep Learning Tools:	Pytorch, scikit-learn, Keras, Tensorflow
NLP and Information Retrieval Tools:	Gensim, word2vec, GloVe, AllenNLP, NLTK, WordNet
Computer Vision Tools:	OpenCV, MeshLab, PCL

EXPERIENCE

Research Assistant @ Purdue University, West Lafayette, IN, USA **June 2014 - present**

Build a search engine using modern IR and NLP techniques for bug localization: [2017-present]

- Developed a large dataset of 20000 bug reports for testing bug localization algorithms
- Constructed semantic vectors for 0.5 million software terms using 35000 GitHub repositories and 1 Billion tokens
- Improved retrieval precision by 20% using word2vec based semantic modeling and MRF based ordering
- Technologies used: Python, Java, Gensim, Terrier, word2vec, GloVe, AllenNLP, NLTK, numpy, scipy
- URL: [SCOR Word Embeddings](#)

Using GitHub data to construct a map of the world based on software activity: [hobby project]

- Technologies used: Python, Gensim, BeautifulSoup, numpy, matplotlib
- URL: [Medium Blog](#)

Create 3-D models of dormant apple fruit trees: [2014-2016]

- Published a new dataset of depth and color images of dormant apple trees from several orchards
- Developed a method for 3-D modeling of trees from depth images using Iterative Closest Point (ICP) algorithm
- Technologies used: MATLAB, C/C++, Kinect SDK, OpenCV, MeshLab, PCL
- URL: [Project Webpage](#)

Graduate Intern Technical @ Intel Corporation, Hillsboro, OR, USA **May 2017 - Dec 2017**

Model performance of SOC chips:

- Developed a tool using LSTM for modeling performance of complex multicore SOC architecture
- Used Random Forest to identify the queues which cause major bottleneck for certain types of traffic
- Technologies used: scikit-learn, Keras, Tensorflow

Computer Vision: (Graduate Level)

Fall 2016 ; Fall 2018

Key topics: SIFT, SURF, RANSAC, Gradient Descent, Levenberg-Marquardt, Image segmentation, Zhang's Algorithm, Stereo reconstruction, Machine learning, PCA, LDA, Decision trees, Adaboost

Key tools: Python, OpenCV, Numpy, Matplotlib

Computer and Network Security: (Senior Undergraduate Level)

Spring 2019 ; Spring 2020

Key topics: DES, 3DES, AES, RC4, RSA, Diffie-Hellman, ECC, Hashing and Cryptocurrency, IP spoofing, SYN flooding, DNS cache poisoning, Buffer overflow, SQL injection, cross-site scripting, DDoS, Spamming

Key tools: Python, BitVector, tcpdump, iptables, dig, gdb, nmap, socket

RELEVANT COURSEWORK

Deep Learning	Computer Vision	Artificial Intelligence
Data Mining	Image Processing	Algorithms
Computer Network Systems	Computer Network Security	Parallel Programming

SELECTED COURSE PROJECTS

- A multithreaded implementation of word2vec (Python, from scratch) (Deep Learning course)
- LeNet-5 and AlexNet implementations in Pytorch (Python) (Deep learning course)
- Comparing feature extraction from images using Harris detector and SURF (Python) (Computer Vision course)
- Modeling parameters of a camera using Zhang's Algorithm (Python) (Computer Vision course)
- PCA vs. LDA for Face Recognition (MATLAB) (Computer Vision course)
- RGB and texture-based image segmentation using Otsu's algorithm (Python) (Computer Vision course)
- Software Defined Network (SDN) using Socket Programming (Python) (Networks course)
- Traffic Engineering in Software Defined Network (Python) (Networks course)
- MapReduce on MPI and OpenMP (C) (Parallel programming course)
- CUDA-Accelerated Face Recognition Using Hidden Markov Models (C/CUDA) (Undergrad work)
- Object Recognition and Digit Recognition Using Convolutional Neural Network (MATLAB) (Undergrad work)
- Autonomous Miniature Vehicle (A white line following robot) (C) (Undergrad work)

PUBLISHED 9 PAPERS INCLUDING

Software search engine / Bug localization

- Shayan A. Akbar, and Avinash C. Kak, **A Large-Scale Comparative Evaluation of IR-Based Tools for Bug Localization**, Mining Software Repositories (MSR), 2020.
- Shayan A. Akbar, and Avinash C. Kak, **SCOR: Source code retrieval with semantics and order**, Mining Software Repositories (MSR), 2019.
- Bunyamin Sisman, Shayan A. Akbar and Avinash C. Kak, **Exploiting spatial code proximity and order for improved source code retrieval for bug localization**, Journal of Software Evolution and Process (JSEP), 2017.

3D modeling of complex objects

- Somrita Chattopadhyay, Shayan A. Akbar, Noha M. Elky, Henry Medeiros, and Avinash Kak, **Measuring and modeling apple trees using Time-of-Flight data for automation of dormant pruning applications**, Winter Conference on Applications of Computer Vision (WACV), 2016.
- Shayan A. Akbar, Noha M. Elfiky, and Avinash Kak, **A novel framework for modeling dormant apple trees using single depth image for robotic pruning application**, International Conference on Robotics and Automation (ICRA), 2016.
- Henry Medeiros, Donghun Kim, Jianxin Sun, Hariharan Seshadri, Shayan Ali Akbar, Noha M. Elfiky, Johnny Park, **Modeling dormant fruit trees for agricultural automation**, Journal of Field Robotics (JFR), 2016.