Dong Wang

dwoanngg@gmail.com, (408)930-2463, San Jose, CA, Website, LinkedIn, Google Scholar

EDUCATION

- Ph.D., Electrical and Computer Engineering 2003
  Carnegie Mellon University, Pittsburgh, PA
- M.S., Electrical and Computer Engineering 1999
  Carnegie Mellon University, Pittsburgh, PA
- B.S., Computer Science and Engineering 1995
  Beijing Institute of Technology, Beijing, China

WORK EXPERIENCES

Chief Architect, Deep Learning for Computer Vision, **Matroid Inc.** 06/2016-11/2018

- Train age gender Keras CNN classification model with 1.4 million parameters. Clean and create balanced dataset with 33k faces from age 1 to 70. Gender accuracy 97.5%, age MAE 4.9 years. Model runs under 250ms on a mobile camera and is also integrated in an Android App using Tensorflow Mobile.
- Improve insightface using data augmentation, face verification accuracy on CFP-FP for sideface is increased from 94% to 96.95%.
- Implement a C++ GRPC server based on Tensorflow serving to dynamically load and serve models. Debug production issues, eg Tensorflow GPU multiple session bug [GitHub].
- Implement a fault tolerant distributed real time task scheduler in Python. Match task resource requirement with worker resources. Maintain task dependencies. Dynamically scale up and down workers through Kubernetes autoscaler on AWS.

Research Scientist, Machine Learning for e-commerce, **Houzz Inc.** 12/2015-05/2016

- Integrate Spark with Luigi job scheduler in YARN.
- Use Spark to collect impression, click and purchase logs, built machine learning models for bidding on Google product search. Feature engineering to deal with data sparseness.

Software Engineer, Cloud infrastructure, **Databricks Inc.** 06/2014-12/2015

- Spark as a Service cluster manager involving resource allocation, health monitoring, auto scaling, security, 3rd party integration.
- Implement enhanced RDA L1 updater for Spark MLlib. [GitHub]
- Implement async execution and cancellation for Spark SQL JDBC thrift server. [GitHub]

Senior Software Engineer, Search and Relevance, **Twitter Inc.** 11/2009-06/2014

- Implement cosine similarity computation for >100M Twitter users on Hadoop using weighted sampling and two-hop random walk. Create ML models for user similarity prediction.
- Hadoop pipeline to predict social connection strength, model is used for personalization and recommendation.

Senior Software Engineer, Display Advertising, **Yahoo! Inc.** 06/2008-10/2009

- Rightmedia exchange Ad serving and a real-time Ads bidding engine**.**

Staff Engineer, Advanced Technology Group, **Synopsys Inc.** 09/2004-05/2008

- Formal verification based on symbolic simulation.

Search Technologist, Model Checking, **0-In Design Automation** 05/2003-08/2004

- Build Model checker based on Boolean satisfiability proofs and assume guarantee reasoning.

GITHUB PROJECTS

Conditional Imitation Learning in Carla Self-driving Simulator [GitHub]

- Train a deep learning model based on ResNet50 to predict steering, gas and throttle conditioned on driving commands from local planner. Improve overall task completion rate from 87% to 93% compared with published result.

Small Object Detection using RetinaNet [GitHub]

- Train an efficient RetinaNet model with 1.2M parameters to identify teacher correction marks. Achieve mAP 0.94 at IOU 0.2, 0.88 at IOU 0.5. SSD can only achieve mAP 0.72 at IOU 0.2.

Handwriting English Line Recognition using GRU and CTC Loss [GitHub]

- Train OCR bidirectional GRU seq2seq model using IAM dataset. Use curriculum learning to train from character, word, short phrase to full sentence for model to converge.

Space Carving for 3D Head Reconstruction [GitHub]

- Collect multiple view images, estimate camera extrinsic parameters. Perform space carving for each view. Compare results with strong priors, such as 3dmm.

SELECTED PUBLICATIONS

- K. Kamath, A. Sharma, D. Wang and Z. Yin, "RealGraph: User Interaction Prediction at Twitter", 2nd *workshop on User Engagement Optimization*, KDD 2014. [PDF]
- A. Goel, A. Sharma, D. Wang and Z. Yin, "Discovering Similar Users on Twitter", *11th workshop on Mining and Learning with Graphs*, KDD 2013. [PDF]
- P. Gupta, A. Goel, J. Lin, A. Sharma, D. Wang, R. Zadeh, "WTF: The Who To Follow Service at Twitter", *22nd International World Wide Web Conference*, WWW 2013. [PDF]
- E. M. Clarke, O. Grumberg, M. Talupur, D. Wang, "Making Predicate Abstraction Efficient: How to Eliminate Redundant Predicates", *15th International conference on Computer Aided Verification (CAV)*, 2003. [PDF]

BOOKS READ

- Machine Learning: A Probabilistic Perspective. [Notes]
- Probabilistic Robotics. [Notes]
- Multiple View Geometry in Computer Vision. [Notes]