

Wei Guo

Citizen of China, Permanent Resident of the United States

Email: robustor@gmail.com

Webpage: <https://wguo.rbind.io>

Github: <https://github.com/w-guo>

LinkedIn: <https://www.linkedin.com/in/w-guo>

EXPERTISE SUMMARY

- Hands-on experience using deep learning on image segmentation under widely varying illumination conditions in manufacturing systems
- Research experience applying *topological data analysis* (TDA) methods to feature extraction, computer vision and dynamic network problems
- Extensive experience dealing with both structured and unstructured data on large datasets, and using machine learning-based predictive analytics to solve business problems

COMPUTER SKILLS

Programming Languages: Python, C/C++, R, SQL, Matlab, AMPL
Software & Tools: Spark, Hive, Keras, Pandas, Scikit-Learn, NLTK, PySpark, OpenCV, CVX

EDUCATION

University of Washington, Seattle, WA
Ph.D., Industrial and Systems Engineering *Mar '20*
Dissertation: Feature Extraction Using Topological Data Analysis for Machine Learning and Network Science Applications

University of Minnesota, Twin Cities, Minneapolis, MN
M.S., Industrial and Systems Engineering *Apr '14*
M.S., Aerospace Engineering and Mechanics *Dec '10*

Harbin Institute of Technology, Harbin, China
M.S., Control Science and Engineering *Jul '08*
B.S., Control Science and Engineering *Jul '06*

RELEVANT PROJECTS

Deep Learning for Automated In-Process Inspection of Composite Layup (US Patent Application) [\[Featured on BARC website\]](#) [\[Github\]](#)
Supervisors: Dr. Agnes Blom-Schieber (Boeing) and Prof. Ashis G. Banerjee Apr '18 - Dec '19

- Won **best presentation award** (presented by Dr. Agnes Blom-Schieber) in data analytics track at *2019 Boeing Tech Excellence Conference* as main contributor
- Developed semantic segmentation-based methods for visual inspections of tow boundaries that form the edges of the individual composite plies
- Trained a two-stage modified U-Net model to learn binary pixel-level segmentation
- Extracted tow boundaries based on the differences between predicted pixel class probabilities
- Improved tow end detection accuracy from 88% using current software to > 99%

Customer Churn Prediction for Streaming Service using PySpark [\[Post\]](#) [\[Github\]](#)

- Built an end-to-end machine learning pipeline with engineered user behavioral features using random forest classifier to identify customers at risk of churning
- Trained and evaluated large-scale model from 26M+ of log data on AWS EMR (f1 score: 0.91)

Content-Based Article Recommender for IBM Watson Studio [\[Github\]](#)

- Applied non-negative matrix factorization to a matrix of TF-IDF features obtained from combined raw documents to find latent topics
- Calculated article similarities and made recommendations for new and existing users

Real-Time Disaster Response with Figure Eight [\[Github\]](#)

- Built an ETL pipeline to store categorized emergency messages in a SQLite database
 - Created a machine learning pipeline for a web application using TF-IDF transformer and multiclass logistic regression to classify incoming messages (average f1 score: 0.94)
-

PROFESSIONAL EXPERIENCE	<p>University of Washington, Seattle, WA</p> <p><i>Research Assistant</i>, Boeing Advanced Research Center (BARC) Jan '16 - Dec '19</p> <ul style="list-style-type: none"> • Community Detection in Large-Scale Dynamic Networks Using TDA [Github] <ul style="list-style-type: none"> – Developed a unified framework for detecting communities and tracking major structural changes during network evolution based on a newly proposed <i>community tree</i> representation – Designed and implemented algorithms to build and update community trees efficiently • Sparse Realization of TDA for Multi-Way Classification [Github] <ul style="list-style-type: none"> – Presented a new method, a.k.a. <i>Sparse-TDA</i>, that incorporates sparse sampling to extract discriminative features in the presence of noisy and redundant information – Demonstrated its advantage over a state-of-the-art kernel TDA method (comparable accuracy / up to 98% training time reduction) and L_1-regularized feature selection methods (2%-8% accuracy increase / up to 73% training time reduction) on 3D meshes of synthetic and real human postures and textured images • Application of TDA in Manufacturing for Feature Selection <ul style="list-style-type: none"> – Applied <i>TDA Mapper algorithm</i> on benchmark data sets for chemical yield prediction and semiconductor fault detection – Yielded topological networks to facilitate a better understanding of casual relationships between process variables and outputs through direct visualization
HONORS & AWARDS	<ul style="list-style-type: none"> • NSF Doctoral Consortium Travel Award, IEEE CASE & ISAM, 2016 • Long March Fellowship from the First Academy of China Aerospace Science & Industry Corp. (CASIC), 2003
COURSEWORK	<ul style="list-style-type: none"> • Deep Learning: Deep Learning (Coursera specialization certificate) • Statistics: Statistical Inference, Stochastic Modeling of Scientific Data, Nonparametric Regression and Classification, Statistical Computing, Stochastic Programming & Game Theory • Operations Research: Linear Optimization Models in Engineering, Convex Optimization, Stochastic Processes in Engineering, Engineering Simulation, Decision Analysis
PUBLICATIONS	<ul style="list-style-type: none"> • W. Guo, R. Chen, Y.-C. Chen, and A. G. Banerjee. Efficient Community Detection in Large-Scale Dynamic Networks Using Topological Data Analysis. Working paper. • E. U. Samani, W. Guo, and A. G. Banerjee. Deep Learning-Based Semantic Segmentation of Microscale Objects. In <i>Proceedings of International Conference on Manipulation, Automation and Robotics at Small Scales (MARSS)</i>, Helsinki, Finland, 2019. • W. Guo, K. Manohar, S. L. Brunton, and A. G. Banerjee. Sparse-TDA: Sparse Realization of Topological Data Analysis for Multi-Way Classification. <i>IEEE Transactions on Knowledge and Data Engineering</i>, 30(7): 1403-1408, 2018. • R. Chen, Y.-C. Chen, W. Guo, and A. G. Banerjee. A Note on Community Trees in Networks. In <i>Workshop on Synergies in Geometric Data Analysis at Neural Information Processing Systems (NIPS)</i>, <i>arXiv preprint arXiv:1710.03924</i>, 2017. • W. Guo and A. G. Banerjee. Identification of Key Features Using Topological Data Analysis for Accurate Prediction of Manufacturing System Outputs. <i>Journal of Manufacturing Systems</i>, 43(2): 225-234, 2017. • W. Guo and A. G. Banerjee. Toward Automated Prediction of Manufacturing Productivity Based on Feature Selection Using Topological Data Analysis. In <i>Proceedings of IEEE International Symposium on Assembly and Manufacturing (ISAM)</i>, Ft. Worth, TX, 2016. • W. Guo, Y. J. Zhao, and B. Capozzi. Optimal Unmanned Aerial Vehicle Flights for Seeability and Endurance in Winds. <i>Journal of Aircraft</i>, 48(1): 305-314, 2011.