

Interesting posters at NeurIPS'18

Yusuf Roohani

December 2018

Posters visited at NeurIPS 2018 held in Montreal, Canada. This selection reflects my interests and accounts for roughly 10% of all posters presented. Selected papers are from all 6 poster sessions. I created the section headings arbitrarily - several papers could fall under multiple headings.

1 Representation Learning

1.1 General

1. [Disconnected Manifold Learning for Generative Adversarial Networks](#)
2. [Learning a latent manifold of odor representations from neural responses in piriform cortex](#)
3. [Point process latent variable models of larval zebrafish behavior](#)
4. [ConditionalGANS robust to noisy labels](#)
5. [BourGAN: Generative Networks with Metric Embeddings](#)
6. [MiME: Multilevel Medical Embedding of Electronic Health Records for Predictive Healthcare](#)
7. [Generative modeling for protein structures](#)
8. [Manifold Structured Prediction](#)
9. [Supervising Unsupervised Learning](#)
10. [Hamiltonian Variational Auto-Encoder](#)
11. [Delta-encoder: an effective sample synthesis method for few-shot object recognition](#)
12. [Weakly Supervised Dense Event Captioning in Videos](#)
13. [Text-Adaptive Generative Adversarial Networks: Manipulating Images with Natural Language*](#)

14. [Variational PDEs for Acceleration on Manifolds and Application to Diffeomorphisms](#)
15. [Learning Latent Subspaces in Variational Autoencoders](#)
16. [Representation Learning for Treatment Effect Estimation from Observational Data](#)
17. [Representation Learning of Compositional Data](#)
18. [Learning latent variable structured prediction models with Gaussian perturbations](#)
19. [Incorporating Context into Language Encoding Models for fMRI](#)
20. [Multi-value Rule Sets for Interpretable Classification with Feature-Efficient Representations](#)- multi value rule sets for interpretable classification with feature efficient representations
21. [Communication Efficient Parallel Algorithms for Optimization on Manifolds](#)
22. [Gaussian Process Prior Variational Autoencoders*](#)
23. [Learning semantic similarity in a continuous space](#)

1.2 Graph representation learning

24. [Navigating with Graph Representations for Fast and Scalable Decoding of Neural Language Models](#)
25. [Adaptive Sampling Towards Fast Graph Representation Learning](#)
26. [Hierarchical Graph Representation Learning with Differentiable Pooling](#)
27. [Beyond Grids: Learning Graph Representations for Visual Recognition*](#)
28. [Mean-field theory of graph neural networks in graph partitioning*](#)
29. [Found Graph Data and Planted Vertex Covers](#)
30. [Graph Convolutional Policy Network for Goal-Directed Molecular Graph Generation](#)
31. [On Learning Markov Chains*](#)
32. [Robust Learning of Fixed-Structure Bayesian Networks](#)
33. [DAGs with NO TEARS: Continuous Optimization for Structure Learning*](#)
34. [Symbolic Graph Reasoning Meets Convolutions](#)

35. [Link Prediction Based on Graph Neural Networks](#)
36. [Constructing Deep Neural Networks by Bayesian Network Structure Learning*](#)
37. [Constrained Graph Variational Autoencoders for Molecule Design*](#)
38. [Cluster Variational Approximations for Structure Learning of Continuous-Time Bayesian Networks from Incomplete Data*](#)
39. [Combinatorial Optimization with Graph Convolutional Networks and Guided Tree Search](#)
40. [Bayesian Structure Learning by Recursive Bootstrap](#)
41. [Watch Your Step: Learning Node Embeddings via Graph Attention*](#)
42. <https://nips.cc/Conferences/2018/Schedule?showEvent=11271> Simple Embedding for Link Prediction in Knowledge Graphs

1.3 Domain adaptation, disentanglement

43. [Unsupervised Image-to-Image Translation Using Domain-Specific Variational Information Bound](#)
44. [Unsupervised Attention-guided Image-to-Image Translation](#)
45. [Conditional Adversarial Domain Adaptation](#)
46. [Bayesian multi-domain learning for cancer subtype discovery from next-generation sequencing count data](#)
47. [Unsupervised Cross-Modal Alignment of Speech and Text Embedding Spaces](#)
48. [Extracting Relationships by Multi-Domain Matching*](#)
49. [A Unified Feature Disentangler for Multi-Domain Image Translation and Manipulation](#)
50. [Adaptive deep embeddings - A Synthesis of Methods for k-Shot Inductive Transfer Learning](#) weight transfer + deep metric learning + few-shot learning. weight transfer is the least effective method for inductive transfer learning. histogram loss is robust regardless of the amount of label data in the target domain.
51. [Insights on representational similarity in neural networks with canonical correlation](#)
52. [Generalizing to Unseen Domains via Adversarial Data Augmentation*](#)

- 53. [Generalization Bounds for Uniformly Stable Algorithms](#)
- 54. [Revisiting \$\(\epsilon, \gamma, \tau\)\$ -similarity learning for domain adaptation](#)
- 55. [Algorithms and Theory for Multiple-Source Adaptation](#)
- 56. [Synthesize Policies for Transfer and Adaptation across Tasks and Environments](#)
- 57. [Life-Long Disentangled Representation Learning with Cross-Domain Latent Homologies*](#)
- 58. [Isolating Sources of Disentanglement in Variational Autoencoders](#)
- 59. [Image-to-image translation for cross-domain disentanglement*](#)
- 60. [Transfer Learning with Neural AutoML](#)
- 61. [Learning Deep Disentangled Embeddings With the F-Statistic Loss](#)
- 62. [Supervised autoencoders: Improving generalization performance with unsupervised regularizers](#)
- 63. [Domain-Invariant Projection Learning for Zero-Shot Recognition](#)
- 64. [A Simple Unified Framework for Detecting Out-of-Distribution Samples and Adversarial Attacks \(know when you don't know\)](#)

2 Causal graphs

- 65. [Experimental Design for Cost-Aware Learning of Causal Graphs*](#)
- 66. [Domain Adaptation by Using Causal Inference to Predict Invariant Conditional Distributions*](#)
- 67. [Causal Inference with Noisy and Missing Covariates via Matrix Factorization](#)
- 68. [Learning and Testing Causal Models with Interventions](#)
- 69. [Multi-domain Causal Structure Learning in Linear Systems*](#)
- 70. [Direct Estimation of Differences in Causal Graphs*](#)
- 71. [Identification and Estimation of Causal Effects from Dependent Data*](#)
- 72. [Submodular Field Grammars: Representation, Inference, and Application to Image Parsing](#)

3 Practical reference

- 73. [Size-Noise Tradeoffs in Generative Networks](#)
- 74. [Visualizing the Loss Landscape of Neural Nets](#)
- 75. [Bias and Generalization in Deep Generative Models: An Empirical Study](#)
If model see red car and blue bus, will it produce a red bus, built a framework to study these problems.
- 76. [Step size matters in deep learning](#)
- 77. [How Does Batch Normalization Help Optimization?](#)
- 78. [Sanity Checks for Saliency Maps](#)
- 79. [On Binary Classification in Extreme Regions](#)
- 80. [How to Start Training: The Effect of Initialization and Architecture](#)
- 81. [On GANs and GMMs](#)
- 82. [Assessing Generative Models via Precision and Recall*](#)
- 83. [When do random forests fail?](#)
- 84. [Learning to Multitask](#)
- 85. [Towards Understanding Learning Representations: To What Extent Do Different Neural Networks Learn the Same Representation*](#)
- 86. [How to tell when a clustering is \(approximately\) correct using convex relaxations](#)
- 87. [Overfitting or perfect fitting? Risk bounds for classification and regression rules that interpolate*](#)
- 88. [Realistic Evaluation of Deep Semi-Supervised Learning Algorithms*](#)
- 89. [Informative Features for Model Comparison](#)
- 90. [Understanding Batch Normalization*](#)
- 91. [How Many Samples are Needed to Estimate a Convolutional Neural Network?*](#)
- 92. [Discrimination-aware Channel Pruning for Deep Neural Networks*](#)
- 93. [Implicit Bias of Gradient Descent on Linear Convolutional Networks](#)

4 Theory, Optimization

- 94. [Generalizing Point Embeddings using the Wasserstein Space of Elliptical Distributions](#)
- 95. [On Neuronal capacity](#) Deeper architectures do not provide greater capacity but have better properties like regularization, smoother functions. New system for determining the capacity of an architecture. upper bound on the total number of bits that can be learned from the data.
- 96. [Neural Ordinary Differential Equations](#) ODE paper Traditional ODE solver takes small steps. Close connection between Resnet and Euler integrators. Taking the exact derivative through an ODE solution rather than back-prop. It has $O(1)$ memory gradients. Can now replace any set of resnet layers with ODE layers - with fewer parameters than ResNet. Can train physics style models and time series models. Have a well defined state at all times. Biggest benefit turned out to be density modeling. New family of continuous depth architectures. Adaptive computation
- 97. [Importance Weighting and Variational Inference](#)
- 98. [A theory on the absence of spurious solutions for nonconvex and nonsmooth optimization](#)
- 99. [The Limit Points of \(Optimistic\) Gradient Descent in Min-Max Optimization](#)
- 100. [A Theory-Based Evaluation of Nearest Neighbor Models Put Into Practice*](#)
- 101. [Sharp Bounds for Generalized Uniformity Testing](#)
- 102. [Scalable Robust Matrix Factorization with Nonconvex Loss](#)
- 103. [Adversarially Robust Optimization with Gaussian Processes](#)
- 104. [First-order Stochastic Algorithms for Escaping From Saddle Points in Almost Linear Time](#)
- 105. [Connecting Optimization and Regularization Paths*](#)
- 106. [The committee machine: Computational to statistical gaps in learning a two-layers neural network](#)
- 107. [Compact Representation of Uncertainty in Clustering](#)

5 Miscellaneous

108. [Leveraging the Exact Likelihood of Deep Latent Variable Models Enables quantitative comparison between architectures. Leads to the notion of structural regularization](#)
109. [ChannelNets: Compact and Efficient Convolutional Neural Networks via Channel-Wise Convolutions](#)
110. [Representer Point Selection for Explaining Deep Neural Networks](#)
111. [Multi-Task Learning as Multi-Objective Optimization](#)
112. [Revisiting Multi-Task Learning with ROCK: a Deep Residual Auxiliary Block for Visual Detection](#)
113. [Partially-Supervised Image Captioning](#)
114. [Neural Nearest Neighbors Networks](#)
115. [A Likelihood-Free Inference Framework for Population Genetic Data using Exchangeable Neural Networks](#)
116. [High Dimensional Linear Regression using Lattice Basis Reduction*](#)
117. [l1-regression with Heavy-tailed Distributions](#)
118. [A Probabilistic U-Net for Segmentation of Ambiguous Images](#)
119. [On Coresets for Logistic Regression](#)
120. [Gather-Excite: Exploiting Feature Context in Convolutional Neural Networks*](#)
121. [Coupled Variational Bayes via Optimization Embedding](#)
122. [Sparse Covariance Modeling in High Dimensions with Gaussian Processes](#)
123. [Causal Discovery from Discrete Data using Hidden Compact Representation](#)
124. [Dynamic Network Model from Partial Observations](#)
125. [Semi-supervised Deep Kernel Learning: Regression with Unlabeled Data by Minimizing Predictive Variance](#)
126. [Multitask Boosting for Survival Analysis with Competing Risks](#)
127. [Binary Classification from Positive-Confidence Data](#)
128. [The Sparse Manifold Transform](#)

- 129. [Manifold-tiling Localized Receptive Fields are Optimal in Similarity-preserving Neural Networks](#)
- 130. [Integrated accounts of behavioral and neuroimaging data using flexible recurrent neural network models](#)
- 131. [Deep Homogeneous Mixture Models: Representation, Separation, and Approximation](#)
- 132. [Model-based targeted dimensionality reduction for neuronal population data*](#)

6 From talks

- 133. [The tradeoffs of large scale learning](#) Test of time award
- 134. [Optimization methods for large-scale machine learning](#)
- 135. [Data science is science's second chance to get causal inference right: A classification of data science tasks](#)
- 136. [Clinically applicable deep learning for diagnosis and referral in retinal disease - Nature Medicine](#)