

NeurIPS (prev. NIPS) Papers Selection

My favorite research articles from NeurIPS (previously NIPS) 2018



Félix

Follow

Dec 4 · 7 min read



This year NeurIPS is returning to Montreal.

Sad not to attend NeurIPS (previously **NIPS**) this year? Thus I have written this **papers selection** for the **2018 edition** that I want to share with you!

. . .

NeurIPS (Neural Information Systems Processing, previously called NIPS) is more popular than a Beyoncé concert. The biggest AI conference in the world sold out in just a few minutes this year. Moreover, the number of accepted papers this year breaks all records (more than one thousand).

You will find below our selection of papers that I hope will give you a little taste of NeurIPS. My objective was to find quality papers that give an overview of different fields of AI. This selection is, of course, subjective and not exhaustive.

SING: Symbol-to-Instrument Neural Generator

Neural Audio Synthesizer

This article presents a new neural audio synthesizer: Symbol-to-Instrument Neural Generator (SING). This model can generate music from hundreds of instruments with different pitches and velocities.

SING can directly **generate a 4-second waveform** sampled at 16000 Hz and has a lightweight structure. The first part of the network is a LSTM that takes as input a concatenation of one-hot encodings: instrument used, pitch and velocity. It is used during 265-time steps. The concatenated outputs are decoded by a convolutional network that generates the waveform.

This network uses a **specific loss**: the 1-norm between the **log spectrogram** (obtained by the short-time Fourier transform) of the waveform and the target waveform.

SING has really good results (listen to audio samples [here](#)) that are better than Wavenet, the reference network so far. This network is only specialized in musical instruments. But the most remarkable result is the processing time that is **2500 times faster than that of Wavenet**.

SING: Symbol-to-Instrument Neural Generator—Alexandre Défossez (FAIR, PSL, SIERRA), Neil Zeghidour (PSL, FAIR, LSCP), Nicolas Usunier (FAIR), Léon Bottou (FAIR), Francis Bach (DI-ENS, PSL, SIERRA)

Deep Anomaly Detection Using Geometric Transformations

Anomaly Detection

This paper from the Israel Institute of Technology aims to make good use of deep learning models in the field of Anomaly Detection.

While the state of the art is **Auto-encoders** (which spot anomalies in the embedded or the reconstructed data), the paper suggests performing a set of geometrical transforms to the data and then applying a discriminative model on the transformed instances (images with bad scores will be considered as anomalies). Training a classifier to distinguish the transformed images **makes it learn salient geometrical features**, some of which are likely to **differentiate the abnormal data**. Performance-wise, the improvement brought to the metrics is sky-high: the top performing baseline AUC is **improved by 67%** compared to the state of the art algorithms on the CatsVsDogs dataset.

Deep Anomaly Detection Using Geometric Transformations—[Izhak Golan](#), [Ran El-Yaniv](#)

GLoMo: Unsupervisedly Learned Relational Graphs as Transferable Representations

Relational Graphs

This article presents a new way to perform transfer learning. Rather than transferring unary features like embeddings, this approach gives the possibility to transfer latent relational graphs that bear information about **relations between data units** (pixels, words ...) which vanish with basic embeddings.

For example, for a question answering problem, an answer predictor is trained with a **graph generator** to predict answers from question

inputs. **This network tries to generate good affinity matrix** (that bear relational information but not the values of the input), that are **injected into hidden layers** of the answer predictor. The answer predictor and the graph generator are trained together.

Once trained, the graph generator can be used with models that do different tasks (for example sentiment analysis) to improve their performances. This new approach improved performance on topics like **question answering, sentiment analysis, image classification**.

GLoMo: Unsupervisedly Learned Relational Graphs as Transferable Representations—Zhilin Yang, Jake Zhao, Bhuwan Dhingra, Kaiming He, William W. Cohen, Ruslan Salakhutdinov, Yann LeCun

Supervising Unsupervised Learning

Unsupervised Learning

One major issue with **Unsupervised learning** is that there is no straightforward way to evaluate an algorithm's performance. **This makes it very hard to select an algorithm to tune hyperparameters** or to evaluate performance.

This article tries to overcome this issue using **Meta Unsupervised Learning (MUL)**: a **classifier is trained to decide which unsupervised model to use based on the characteristics of the dataset**. For this, a collection of labeled datasets is needed.

For instance, suppose we want to choose between several unsupervised classification algorithms for a given problem where we have no labels. We run each algorithm on many labeled datasets, on which we can compute a classification score. We then train a model to predict the best algorithm using a **mix of dataset characteristics** (dimension, eigenvalues, ...) **and unsupervised metrics on the output of the classifier** (spread within clusters, ...). This model can **be used to choose an algorithm for the dataset of interest**.

This approach seems to beat fully-unsupervised methods **even in cases when the labeled datasets are not closely related to the one we are studying**.

Supervising Unsupervised Learning—Vikas K. Garg, Adam Kalai

Generative Adversarial Network

This article introduces Banach Wasserstein Generative Adversarial Networks (BWGANs) extending Wasserstein GANs that are themselves an improvement of GANs (here is a [good introduction to GANs](#)).

For a basic GAN, assuming the discriminator is perfectly trained, the generative net actually minimizes the Jensen-Shannon distance (JSD, a symmetric version of the [Kullback–Leibler divergence](#)) between the distribution of generated images and the true distribution. But the **JSD distance is not adapted to measure the distance between image distributions.**

In WGANs, the loss is modified so as to minimize the [Wasserstein distance](#) instead of the JSD distance. In order to do so, a Lipschitz constraint is softly enforced on the network by adding an L2 penalty term on the gradient to the loss function. One main advantage of the Wasserstein distance is that it can be applied to arbitrary norms on the image space.

However, **WGANs force the use of the L2 norm** due to the penalty term on the loss, hence **it loses the ability to use norms more adapted to image** e.g., **Sobolev norms that emphasis not only on pixels but also on edges.**

The article proposes a **generalization of the penalty term** so as to overcome the L2 norm limitation.

The authors use the $W[-3/2, 2]$ Sobolev norm. It achieves beyond state of the art results on the CIFAR-10 dataset.

A very mathematical article: it contains exhaustive proofs (not only proof sketches), it recalls basics notions such as [Banach spaces](#) and [Sobolev spaces](#).

Learning to Decompose and Disentangle Representations for Video Prediction

Video Prediction

Video prediction is the task of predicting the next K frames of an image from the previous T ones. Solving the problem of video prediction could mean understanding how the world works.

More specifically, **understanding the physics of an object**, such as how a rope behaves differently from a metal bar, is natural in our everyday life but make **video prediction a complicated task**.

Videos have a high dimensionality and are unregular. This paper introduces the **Decompositional Disentangled Predictive Auto-Encoder (DDPAE)**, which finds the **lightest possible way to describe objects in a video**. It makes the assumption that every video consists of several objects. Each of them could be described using a **content vector** (a constant descriptor of the object itself), and a **pose vector** (the position that should be found and predicted).

This solution learns to find such descriptions and disentangle all of its elements. It combines VAE, RNN, and seq2seq. The **results look promising as they surpass the baseline on the Moving MNIST dataset**.

Learning to Decompose and Disentangle Representations for Video Prediction—Jun-Ting Hsieh, Bingbin Liu, De-An Huang, Li Fei-Fei, Juan Carlos Niebles

Unsupervised Learning of Artistic Styles with Archetypal Style Analysis

Style Transfer

This paper brings a new step towards unsupervised learning and Deep Learning interpretability. Especially it addresses the issue of style learning with **root styles** explanation and manipulation (here is a good introduction to [style learning](#) if you are new to this topics).

The main idea is to project the input image into a low dimensional **archetypes space** where each base archetype is interpretable. Doing so, one is able to: **attach some features to an image in an unsupervised manner** (e.g. adding a tag about texture, style, age, etc. coming from the interpretation of the archetypes) and **manipulate the coefficient over each style to influence and transfer style** to the original image.

Furthermore, the projection of the encoded image onto the archetypes is done with an optimization in the simplex in a two-sided manner: minimizing the distance of the images to their projections while enforcing the archetypes to be a linear combination of the images. So the **archetypes are easily interpretable**.

In the end, it is possible to describe any image with *base style ingredients*, learning then a sort of a **style dictionary**. The style transfer can finally be precisely managed by the **coefficients in the archetypes space**.

*Unsupervised Learning of Artistic Styles with Archetypal Style Analysis—
Daan Wymen, Cordelia Schmid, Julien Mairal*

. . .

Want more articles like this one? Don't forget to click the Follow button!

. . .

Need help on your data projects?

Email

Get in touch!

☐

I agree to leave Medium and submit this information, which will be collected and used according to [Upscribe's privacy policy](#).

