

State of Deep Learning : H2 2018 Review

 medium.com/atlas-ml/state-of-deep-learning-h2-2018-review-cc3e490f1679

Ross Taylor

November 30, 2018

It's really hard to keep track of developments in a  field like deep learning.

Earlier in the year [Robert Stojnic](#) and I hacked together [Papers With Code](#) as a first step in solving this problem. The site is a community resource that connects deep learning research papers with code implementations.

It also enables us to take a birds-eye view of the field as a whole. We can see what the research trends are, which frameworks are being adopted by the community, and which techniques are gaining favour. This blog post details some of the results!

In this post we summarise some of the **key developments in deep learning in the second half of 2018** using the data we have from the site. We then briefly discuss the road ahead for the deep learning community.

Most popular official: BERT, vid2vid and graph_nets



BERT: Pre-training of Deep Bidirectional Transformers...

Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova • TensorFlow

We introduce a new language representation model called BERT, which stands for Bidirectional Encoder Representations from Transformers. BERT is designed to pre-train deep bidirectional representations by jointly conditioning on both left and right context in all layers.

★ 8327

NLP

Google AI's [BERT paper](#) made waves across the deep learning community in October. The paper proposes a deep bidirectional Transformer model that achieves state-of-the-art performance for 11 NLP tasks, including the Stanford Question Answering ([SQuAD](#)) datasets. Google AI open sourced the [code for their paper](#), and this was the deep learning repository that gained the most stars between July 2018 and the time of writing.



Video-to-Video Synthesis

Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Guilin Liu, Andrew Tao, Jan Kautz, ... • PyTorch

We study the problem of video-to-video synthesis, whose goal is to learn a mapping function from an input source video (e.g. a sequence of semantic segmentation masks) to an output photorealistic video that precisely depicts the content of the source video.

★ 5085

GENERATIVE

NVIDIA's [video-to-video synthesis paper](#) was yet another amazing result for generative modelling, which has been one of the most popular deep learning fields in the the past few years. The paper fixes the problem of temporal incoherence with a novel sequential generator architecture, as well as a number of other design features like foreground-and-background priors to improve performance. NVIDIA [open sourced their code](#), which was the second most popular implementation in the second half of 2018.



Relational inductive biases, deep learning, and graph networks

Peter W. Battaglia, Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, ... • TensorFlow

We present a new building block for the AI toolkit with a strong relational inductive bias--the graph network--which generalizes and extends approaches for neural networks that operate on graphs, and provides an interface for manipulating structured knowledge and producing structured behaviors.

★ 2211

GRAPH NETWORKS

Google DeepMind's [paper on graph networks](#) received a lot of attention in the middle of the year as a new type of structured data that deep learning could begin to attack (the majority of deep learning applications have been on vectors and sequences). Their [open source library](#) was the third most popular implementation in the second half of 2018.

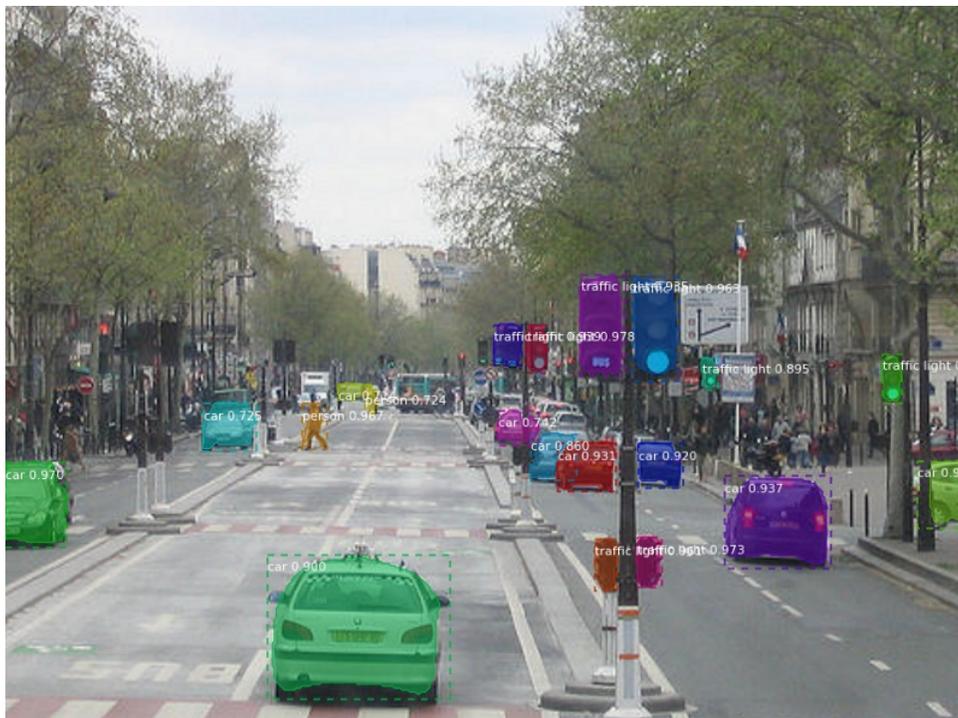
Most popular community: DeOldify, BERT and Fast R-CNNs



DeOldify uses SA-GANs, a PG-GAN inspired architecture and a two time-scale update rule

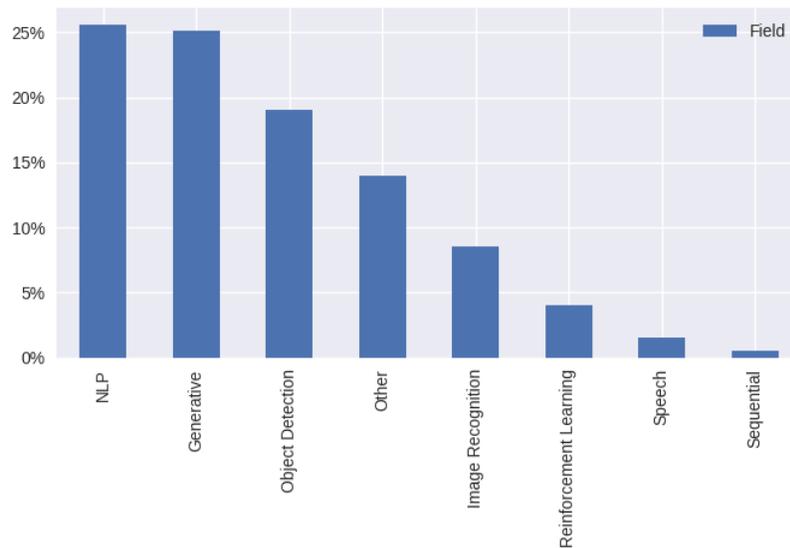
The DeOldify project got a lot of interest from the wider deep learning community. The author Jason Antic implemented techniques from a number of papers in the generative modelling field, including self-attention GANs, progressively growing GANs and a two time-scale update rule. The code for the project has over 4,000 stars on GitHub at the time of writing.

Junseong Kim's BERT implementation for PyTorch also got a lot of attention from the community. With the community increasing building implementations in one of two frameworks, there is a big need for papers to be implemented in both frameworks so the entire deep learning community can use them. Kim's work is a clear illustration of this, and her implementation enjoys over 1,500 GitHub stars at the time of writing.



Finally, [Waleed Abdulla's Keras/TensorFlow implementation of Mask R-CNN](#) was the third most popular community implementation in terms of GitHub stars gained. Architecturally, the implementation uses a Feature Pyramid Network and a ResNet101 backbone, and the library can be used for a number of applications such as 3D building reconstruction, object detection for self-driving cars, detecting types of building in maps and more. The library has over 8,000 stars on GitHub.

Most activity: NLP and GANs



Looking at the top fifty implementations, the hottest fields appear to be **generative methods** and **natural language processing (NLP)**. Within generative methods, popular implementations on GitHub included: [vid2vid](#), [DeOldify](#), [CycleGAN](#) and [faceswaps](#). Within NLP, popular GitHub repositories included [BERT](#), [HanLP](#), [jieba](#), [AllenNLP](#) and [fastText](#).

1 out of 7 new papers have code

One of our goals with the site was to encourage deep learning researchers to post code with their research. No code with your research? You don't get featured on the site: simple as that. Here is the state of play with reproducibility on our platform:

- We've processed **60,000+** machine learning papers from the **last 5 years**
- Of these 60k papers **~12% of have code implementations**
- In the last 6 months **~15% of newly published papers** (i.e. 1 in 7) have code implementations

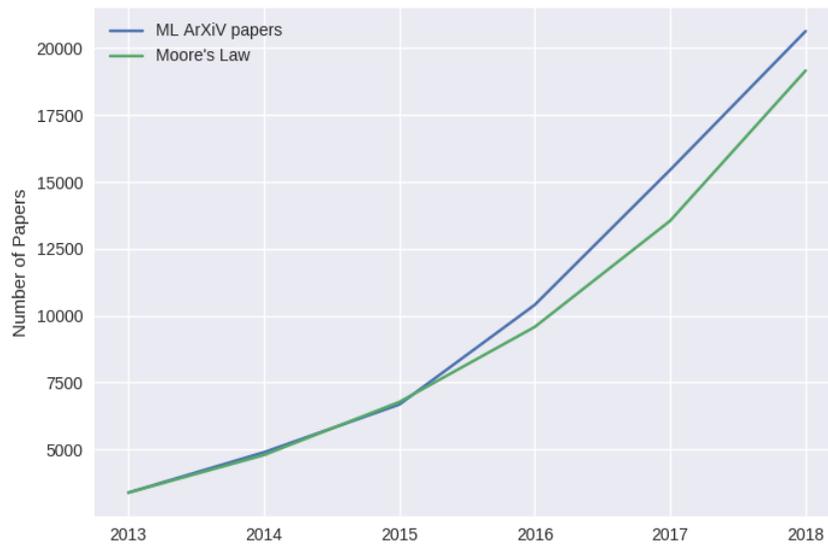
Clearly there is more work to be done, but things are moving in the right direction!

Every 20 minutes, a new ML paper is born

The growth rate of machine learning papers has been around 3.5% a month since July – which is around a 50% growth rate annually. This means around **2,200 machine learning papers a month** and that **we can expect around 30,000 new machine learning papers next year**.

As a reference speed we can compare to Moore's Law. Moore's Law shows progress at the first level of computing – the speed at which we can process. Machine Learning can be thought of achieving *learnt functionality*, and can be thought of an abstraction on top of hardware and software.

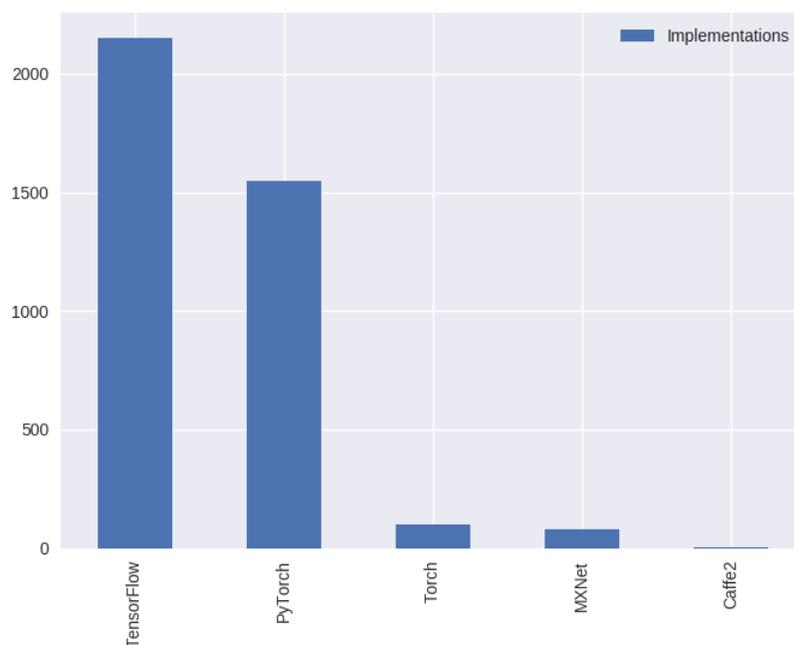
The number of ML papers on our site appears to be growing faster than Moore's Law in the past 3 years, which gives you a sense that people believe that this is where the future value in computing is going to come from.



The credit for this comparison idea comes from Jeff Dean et al <https://ieeexplore.ieee.org/document/8259424>

Framework duopoly: TensorFlow and PyTorch

The majority of implementations on the site appear to be in TensorFlow, although PyTorch isn't far behind. The remaining frameworks (MXNet, Torch and Caffe2) have a much smaller presence in the ecosystem. Given the changes happening in both frameworks – TensorFlow moving towards eager execution and a new Keras-inspired API and PyTorch looking to make it easier to productionise models – it will be interesting to see how this balance changes within the next year.



The Road Ahead

The community is making progress with reproducibility, but the fact that we only have 1/7 coverage means we have more work to do. We feel there will be a big role for the wider community in making reusable open source ML code—rather than just relying on researchers as we've done in the past.

In particular, we think the **"indie ML community"**—who are outside the big companies like Google and Facebook—will help push forward the coverage of code for research, and also help us verify whether paper results hold up.

To achieve the prize, the research community and developer community need to reach out to each other so we can create useful machine learning artefacts that can be used in real-world applications. Only then will deep learning reach its full potential. Full steam ahead!