

Modern Deep Learning Architectures

MLRG Winter term 2

Deep Learning is taking over the world

Deep Learning is taking over the world

DÉTECTER LA LANGUE **ANGLAIS** FRANÇAIS ▾ ↔ **FRANÇAIS** ANGLAIS ▾

WaveNet is a deep neural network for generating raw audio. It was created by researchers at London-based artificial intelligence firm DeepMind. The technique, outlined in a paper in September 2016,[1] is able to generate relatively realistic-sounding human-like voices by directly modelling waveforms using a neural network method trained with recordings of real speech. ✕

WaveNet est un réseau neuronal profond pour générer de l'audio brut. Il a été créé par des chercheurs de la société d'intelligence artificielle basée à Londres DeepMind. La technique, décrite dans un article en septembre 2016, [1] est capable de générer des voix de type humain relativement réalistes en modélisant directement des formes d'onde en utilisant une méthode de réseau neuronal formé avec des enregistrements de parole réelle.

370 / 5000 ▾

Écouter

Deep Learning is taking over the world

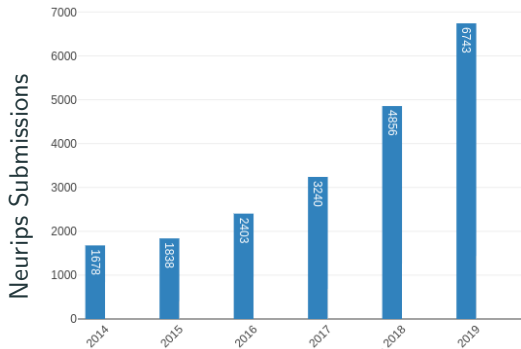
Title: Star's Tux Promise Draws Megyn Kelly's Sarcasm

Subtitle: Joaquin Phoenix pledged to not change for each awards event

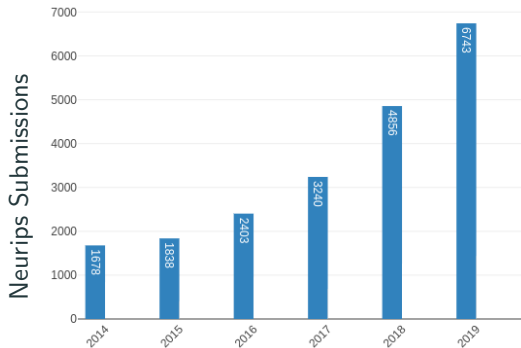
Article: A year ago, Joaquin Phoenix made headlines when he appeared on the red carpet at the Golden Globes wearing a tuxedo with a paper bag over his head that read, "I am a shape-shifter. I can't change the world. I can only change myself." It was a promise to not change to fit into the Hollywood mold: "I think that's a really special thing, to not change yourself. I think it's a really special thing to say, 'This is what's inside of me, I'm proud of it, and I'm not going to be ashamed because of the way that someone else thinks I should be.'" Now, it's the Oscars, and Phoenix is at it again. But this time, his publicist is saying he'll be wearing a tux no matter what.

Megyn Kelly was not impressed, and she let him have it on The Tonight Show. "You know, I feel like, I feel like you could have worn the tux," she says. "But you're saying you're a shape-shifter. I don't know if you can change your tux, but you can change your mind. You can change your mind. You can change your mind." Phoenix says he did, but it didn't stick. "I was like, 'Okay, I'm going to wear a tuxedo to this thing.' And then I thought, 'I don't want to wear a tuxedo to this thing.'" Kelly goes on to encourage him to change his mind again, but Phoenix says it's too late: "I'm committed to wearing this."

And it's still growing



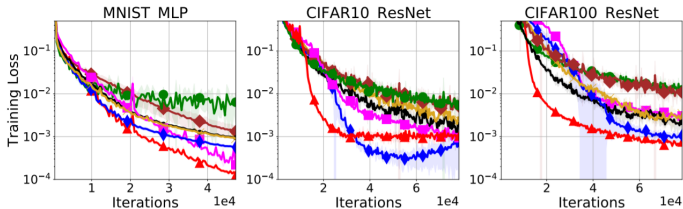
And it's still growing



Bidirectional Encoder Representations
Length-Direction Decoupling
Depthwise Separable Convolutional Neural Network
Highway Networks
Deep Convolutional GAN
Dissentangled Wasserstein Auto-Encoder
Multi Head Self Attention

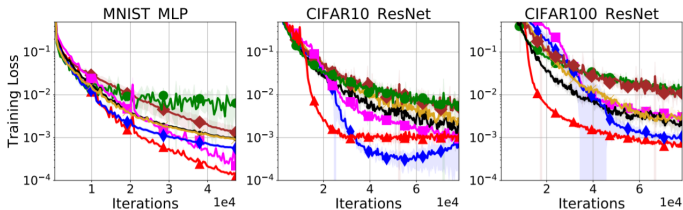
but not yet ubiquitous

Meanwhile, in overlapping fields:

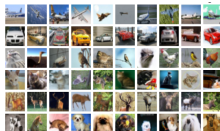


but not yet ubiquitous

Meanwhile, in overlapping fields:

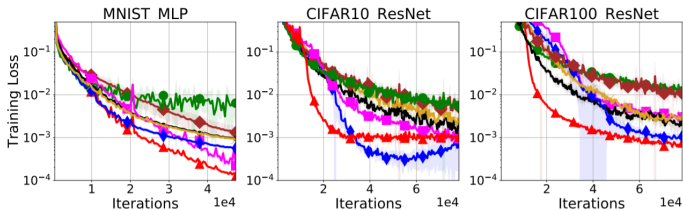


airplane
automobile
bird
cat
deer
dog

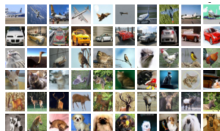


but not yet ubiquitous

Meanwhile, in overlapping fields:

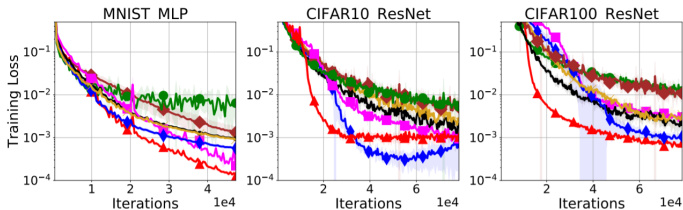


airplane
automobile
bird
cat
deer
dog

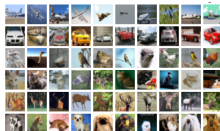


but not yet ubiquitous

Meanwhile, in overlapping fields:



airplane
automobile
bird
cat
deer
dog



IMAGENET

Goal: Overview of significant papers in deep learning

Different data types, architectures and problems

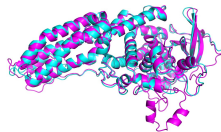
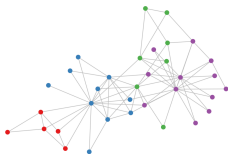
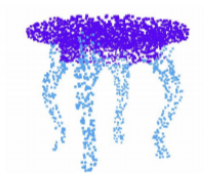
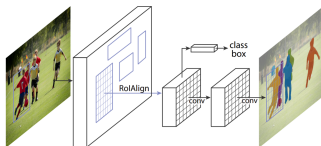


Figure 1: A second of generated speech.

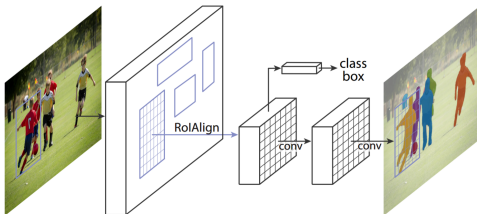


Overview of the papers

Feb 3	Segmentation
Feb 10	3D point clouds
Feb 17	Attention and Transformers
Feb 24	Audio data
Mar 3	Batch Normalization
Mar 10	GANs and optimal transport
Mar 17	Autoregressive models and VAEs
Mar 24	Neural ODE
Mar 31	Graph Neural Networks
Apr 7	AlphaFold

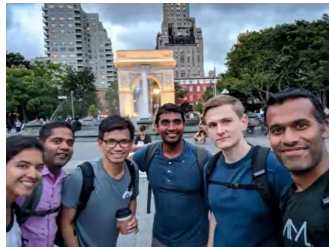
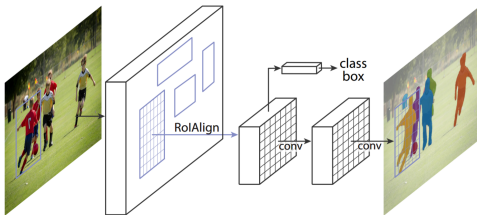
Signups

He et al. (2017), "Mask R-CNN"



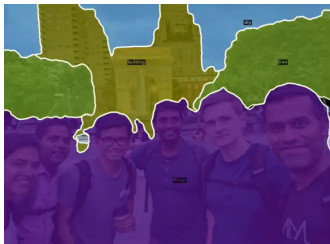
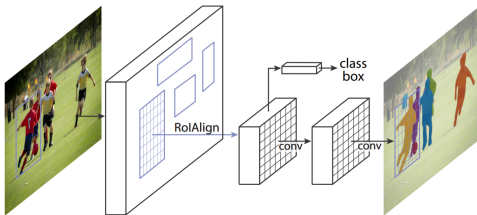
Segmentation

He et al. (2017), "Mask R-CNN"



Segmentation

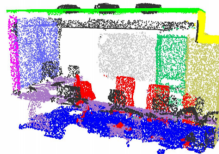
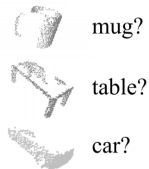
He et al. (2017), "Mask R-CNN"



FAIR's Detectron

3D point clouds

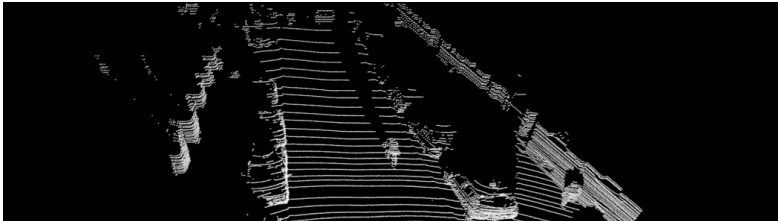
Zhou and Tuzel (2018), "VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection"



Apple Research

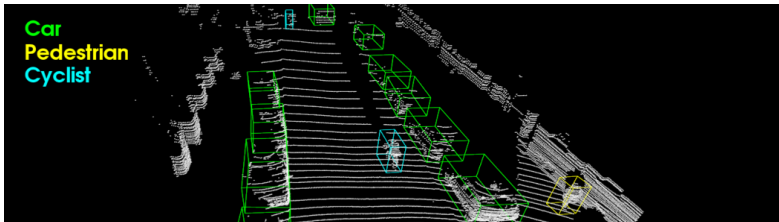
3D point clouds

Zhou and Tuzel (2018), "VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection"



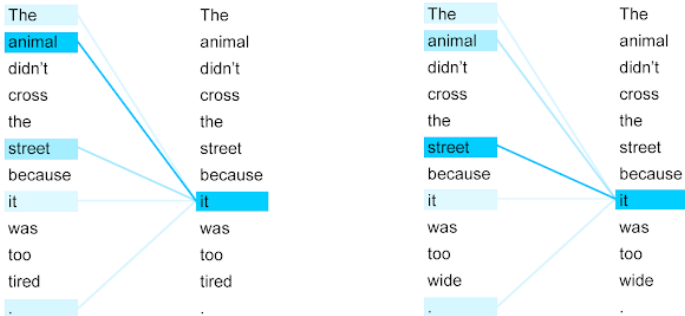
3D point clouds

Zhou and Tuzel (2018), "VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection"



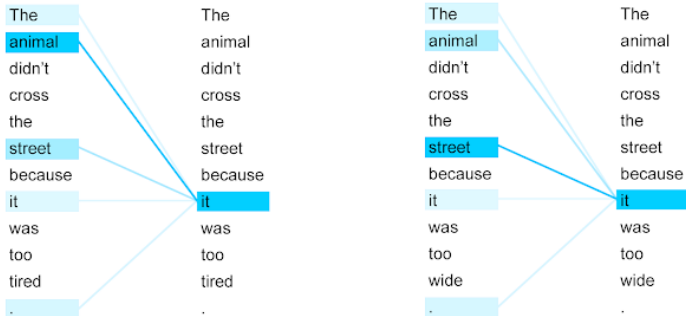
Attention and Transformers

Vaswani et al. (2017), "Attention is All you Need"



Attention and Transformers

Vaswani et al. (2017), "Attention is All you Need"



BERT and GPT3

Oord et al. (2016), "WaveNet: A Generative Model for Raw Audio"



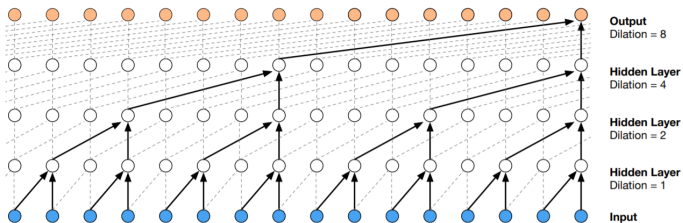
Figure 1: A second of generated speech.

Audio data

Oord et al. (2016), "WaveNet: A Generative Model for Raw Audio"



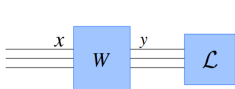
Figure 1: A second of generated speech.



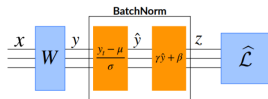
Google's audio generation

Batch Normalization

Santurkar et al. (2018), “How Does Batch Normalization Help Optimization?”



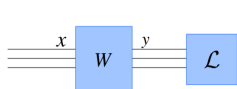
(a) Vanilla Network



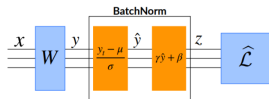
(b) Vanilla Network + BatchNorm Layer

Batch Normalization

Santurkar et al. (2018), “How Does Batch Normalization Help Optimization?”



(a) Vanilla Network

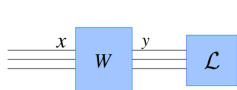


(b) Vanilla Network + BatchNorm Layer

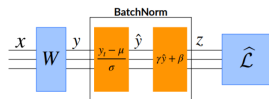
*It works amazingly well. But we know almost nothing about it.
Don't you want to know what internal covariate shift is? - Ali Rahimi*

Batch Normalization

Santurkar et al. (2018), “How Does Batch Normalization Help Optimization?”

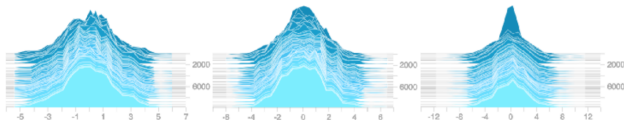


(a) Vanilla Network



(b) Vanilla Network + BatchNorm Layer

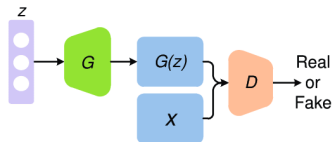
It works amazingly well. But we know almost nothing about it. Don't you want to know what internal covariate shift is? - Ali Rahimi



Building block of many modern models

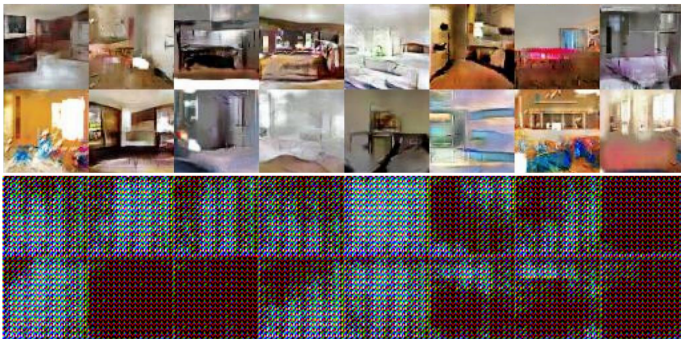
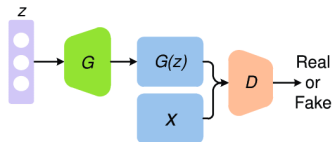
GANs and optimal transport

Arjovsky et al. (2017), "Wasserstein Generative Adversarial Networks"



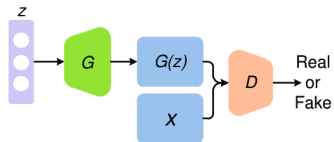
GANs and optimal transport

Arjovsky et al. (2017), "Wasserstein Generative Adversarial Networks"



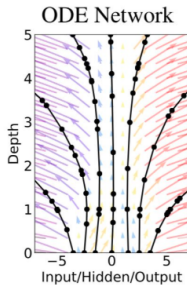
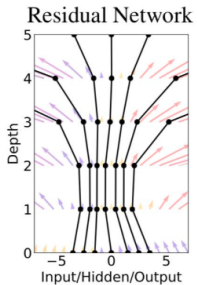
GANs and optimal transport

Arjovsky et al. (2017), "Wasserstein Generative Adversarial Networks"



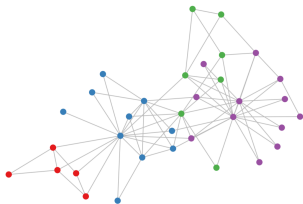
Neural ODE

Chen et al. (2018), "Neural Ordinary Differential Equations"



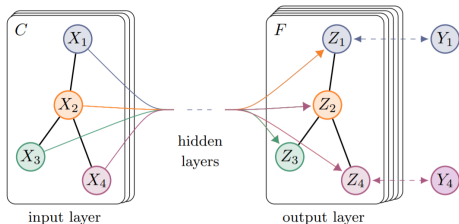
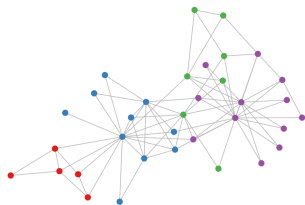
Graph Neural Networks

Kipf and Welling (2017), "Semi-Supervised Classification with Graph Convolutional Networks"



Graph Neural Networks

Kipf and Welling (2017), "Semi-Supervised Classification with Graph Convolutional Networks"



Senior et al. (2020), “Improved protein structure prediction using potentials from deep learning”

Signups

Feb 3	Segmentation
Feb 10	3D point clouds
Feb 17	Attention and Transformers
Feb 24	Audio data
Mar 3	Batch Normalization
Mar 10	GANs and optimal transport
Mar 17	Autoregressive models and VAEs
Mar 24	Neural ODE
Mar 31	Graph Neural Networks
Apr 7	AlphaFold

References

- Arjovsky et al. (2017)
Martin Arjovsky, Soumith Chintala, and Léon Bottou (2017). "Wasserstein Generative Adversarial Networks". In: *ICML*, pp. 214–223.
- Chen et al. (2018)
Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud (2018). "Neural Ordinary Differential Equations". In: *NeurIPS*.
- Chen et al. (2017)
Xi Chen, Diederik P. Kingma, Tim Salimans, Yan Duan, Prafulla Dhariwal, John Schulman, Ilya Sutskever, and Pieter Abbeel (2017). "Variational Lossy Autoencoder". In: *ICLR*.
- He et al. (2017)
Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross B. Girshick (2017). "Mask R-CNN". In: *ICCV*, pp. 2980–2988.
- Kipf and Welling (2017)
Thomas N. Kipf and Max Welling (2017). "Semi-Supervised Classification with Graph Convolutional Networks". In: *International Conference on Learning Representations*.
- Oord et al. (2016)
Aäron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew W. Senior, and Koray Kavukcuoglu (2016). "WaveNet: A Generative Model for Raw Audio". In: *ISCA*, p. 125.
- Santurkar et al. (2018)
Shibani Santurkar, Dimitris Tsipras, Andrew Ilyas, and Aleksander Madry (2018). "How Does Batch Normalization Help Optimization?" In: *NeurIPS*, pp. 2483–2493.
- Senior et al. (2020)
Andrew W. Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Zidek, Alexander W. R. Nelson, Alex Bridgland, Hugo Penedones, Stig Petersen, Karen Simonyan, Steve Crossan, Pushmeet Kohli, David T. Jones, David Silver, Koray Kavukcuoglu, and Demis Hassabis (2020). "Improved protein structure prediction using potentials from deep learning". In: *Nat.* 577.7792, pp. 706–710.
- Vaswani et al. (2017)
Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin (2017). "Attention is All you Need". In: *NeurIPS*, pp. 5998–6008.
- Zhou and Tuzel (2018)
Yin Zhou and Oncel Tuzel (2018). "VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection". In: *CVPR*, pp. 4490–4499.