The Fundamentals of Deep Learning
with Applications

Jon Krohn
jon@untapt.com

Chief Data Scientist at untapt

New York Open Statistical Programming Meetup
January 17th, 2017
(slides available at jonkrohn.com/talks)
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Outline

1. Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2. Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3. Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
Outline

1. Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2. Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3. Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
Outline

1. Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2. Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3. Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
1 Antecedents

Case Study: A History of Biological & Artificial Vision
Building Blocks

2 Theory

Biological & Artificial Neurons
Neural Networks
Deep Neural Networks

3 Contemporary Applications

Convolutional Neural Networks
Long Short-Term Memory Recurrent Neural Networks
Deep Learning at untapt
Deep Reinforcement Learning
Biodiversity during the Phanerozoic

All Genera
Well-Resolved Genera
Long-Term Trend
The "Big 5" Mass Extinctions
Other Extinction Events

Millions of Years Ago
Thousands of Genera
Hubel & Wiesel (1959)
Hubel & Wiesel, 1968
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Hubel & Wiesel, 1968
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Visual Cortices
LGN
Parietal Lobe
Occipital Lobe
V5 (Motion)
V7
V3a (Motion)
V3 (Form)
V2 (Relays signals)
V1 (Catalogs Input)
VP (Relays signals)
V4 (Color and Form)
V8

Sagittal Section
V7
V3a
V3
V1
V2
VP
V4
V8
Camera Obscuro

da Vinci (15th Century)
Block World
Larry Roberts (1965)

(a) Original picture.
(b) Differentiated picture.
(c) Line drawing.
(d) Rotated view.
Neurocognitron
Fukushima (1980)
MNIST Digits & LeNet-5
LeCun, Boutou, Bengio & Haffner (1998)

Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

LeNet-5
LeCun, Boutou, Bengio & Haffner (1998)
ImageNet
Fei-Fei Li et al. (2009), 14m images, 22k categories
ImageNet Classification Error
ILSVRC: 1.4m, 1k object classes
AlexNet
Krizhevsky, Sutskever & Hinton (2012)
Deep Learning
Antecedents
Vision Case Study
Building Blocks
Theory
Neural Units
Neural Nets
Deep Neural Nets
Application
ConvNets
LSTMs
untapt
Reinforcement
Deep Learning

Antecedents

Vision Case Study

Building Blocks

Theory

Neural Units

Neural Nets

Deep Neural Nets

Application

ConvNets

LSTMs

untapt

Reinforcement

INT. EXT.

We see E pull a book from a shelf, flip through it while speaking, and then put it back.

E2

You should see the boys and shut up. I was the one who was going to be a hundred years old.

E1

I saw him again, the way you were sent to me. That was a big honest idea. I am not a bright light.

E1

Well, I have to go to the skull. I don’t know.

E1

He picks up a light screen and rights the security force of the particles of a transmission on his face.

(continued)

What do you mean?

E: (sighing)

I don’t know anything about any of this.

(to N, taking his eyes from his mouth)

Then what?

E2

There’s no answer.

(continued)

We’re going to see the money.

“Reading”

“All right, you can’t tell me that.”

Steps back. Coffey is still going through.

C

I was coming to that thing because you were so pretty.

H

I don’t know. I don’t know what you’re talking about.

C

That’s right.

H

So what are you doing?

C2

I don’t want to be honest with you.

H

He looks at him for a moment, then smiles at him.

H

You don’t have to be a doctor.

C2

I am not sure. I don’t know what you’re talking about.

H

I want to see you too.

C2

What do you mean?

H

I’m sorry, but I’m sure you wouldn’t even touch me.

C2

I don’t know what you’re talking about.

H

The principle is completely constructed for the same time.

C2

It was all about you to be true.

H

You didn’t even see the movie with the rest of the base.

C2

I don’t know.

H

I don’t care.

C2

I know that it’s a consequence. Whatever you want to know about the presence of the story, I’m a little bit of a boy in the floor.

H

I don’t know. I just have to ask you to explain to me what you say.

C2

What do you mean?

H

Because I don’t know what you’re talking about.

C2

That was all the time.

H

I know that.

C2

I don’t know.

H

(angered)

It would be a good time. I think I could have been my life.

Be starts to shake.

C2

(Contin’d)

It may never be forgotten, but that is just too bad. I love to live, but I’m not free of the world.

C

Tell, perhaps I should take it from here. I’m not going to do anything.

H

You can’t afford to take this personally, it’s not a dream. But I got a good time to stay there.

C

Well, I think you can still be back on the table.

C2

Mmm. It’s a damn thing scared to say, nothing is going to be a thing but I was the one that got on to the thing and then I left the other two.

He is standing inside the store and sitting on the floor. He takes a seat on the counter and pulls the camera over to his hair. He raises it up. He is on the phone. He puts the phone on the edge of the room and puts it in his mouth. He sees a black hole in the floor leading to the man on the roof.

He comes up behind him to protect him. He is still standing next to him.

He looks through the door and the door closes. He looks at the boy from his backpack and starts to cry.

Well, there’s the situation with me and the light on the ship. The guy was trying to stop me. He was like a baby, and I was... ap... I was crazy. He would have done it all. He couldn’t even stand, couldn’t do it. He was... I was crazy to take it out. It was a long time ago, he was a little late. I was going to be a moment. I just wanted to pull you up. I was crazy. He would have done it all. He wanted to stop him and I couldn’t even tell. I didn’t want to hurt him. I’m sorry, I know I don’t like him, I can’t let it be. I can’t let him be. I love him. So I get him all the way over here and find the square and go to the gate with him and she won’t show up. Then I’ll check it out. I don’t want to run him. I don’t want to run him. He’s going to see me. He looks at me and he checks me out of his eyes. Then he said he’d go to bed with me.
Sunspring
Sharp & Goodwin (2016)

[video]
Outline

1. Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2. Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3. Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
Hardware

- local machine
- build your own server
- AWS / Google Cloud Platform
- GPU(s) / TPU(s)
## Popular Libraries

based on Johnson (2016) in Stanford CS231n I.12

<table>
<thead>
<tr>
<th></th>
<th>Caffe</th>
<th>Torch</th>
<th>Theano</th>
<th>TensorFlow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td>C++, Python</td>
<td>Lua</td>
<td>Python (Lasagne)</td>
<td>Python</td>
</tr>
<tr>
<td><strong>Pretrained</strong></td>
<td>Yes++</td>
<td>Yes++</td>
<td>Yes</td>
<td>Inception</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Data</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Model</strong></td>
<td>No</td>
<td>Yes</td>
<td>Experimental</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Readable Source Code</strong></td>
<td>Yes (C++)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Good at RNN</strong></td>
<td>No</td>
<td>Mediocre</td>
<td>Yes</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Higher-Level APIs</strong></td>
<td>No</td>
<td>No</td>
<td>Keras</td>
<td>Keras and TFLearn</td>
</tr>
</tbody>
</table>
Deep Learning

[Human Learning Resources]

First Steps. For people in New York, I founded a Deep Learning Study Group. If you're further afield, you can track our progress via GitHub. Otherwise, get a lay of the land from:

- the sequence of courses suggested by Greg Brockman, or
- this (more comprehensive) introductory resource post from Ofir Press

Textbooks. Relative to viewing lectures, I prefer reading and working through problems. The stand-out resources for this, in the order they ought to be tackled are:

- Michael Nielsen's e-book Neural Networks and Deep Learning
- the in-press Deep Learning textbook by Goodfellow, Bengio and Courville

Interactive Demos. Top-drawer interactive demos you can develop an intuitive sense of neural networks from are provided by:

- Chris Olah
- the Illustrious Andrej Karpathy

Applications. Scroll down to see my recommendations for high-quality data sources as well as global issues in need of solutions. Problems worth solving with deep learning approaches in particular are curated by OpenAI.

Academic Papers. If you're looking for the latest deep learning research, bookmark:

- Flood Sung's roadmap for deep learning papers
- Adit Deshpande's list of nine key papers
- this thorough, subcategorized reading list
- Karpathy's arXiv Sanity Preserver
- GitXiv for open-source implementations of popular arXiv papers
Outline

1. Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2. Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3. Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Biological Neuron Morphology

[Diagram of a neuron with labeled parts: Dendrites, Cell Body, Axon, Terminal Bulb]
Perceptron
Rosenblatt (1957)

\[
\text{output} = \begin{cases} 
0 & \text{if } \sum_j w_j x_j \leq \text{threshold} \\
1 & \text{if } \sum_j w_j x_j > \text{threshold}
\end{cases}
\]
Biological Neuron Physiology

The *Binary* Action Potential

![Diagram of Membrane Potential vs Time](image)
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Perceptron
Rosenblatt (1957)
Multi-Layer Perceptron
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Multi-Layer Perceptron

\[ w + \Delta w \rightarrow \text{output} + \Delta\text{output} \]
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Sigmoid Neuron

\[
\frac{1}{1 + \exp(-\sum_j w_j x_j - b)}
\]
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

$tanh$ Neuron

\[
\sigma(z) = \frac{1 + \tanh(z/2)}{2}
\]
ReLU: Rectified Linear Units
Nair & Hinton (2010); Maas, Hannun & Ng (2014)
1. Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2. Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3. Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
MNIST
LeCun, Cortes & Burges
Fully-Connected Neural Net

Single Hidden Layer
TensorFlow Playground
Deep Learning

Outline

1. Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2. Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3. Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
Deep Fully-Connected Net

3 (or more) Hidden Layers
<table>
<thead>
<tr>
<th>Feature</th>
<th>Caffe</th>
<th>Torch</th>
<th>Theano</th>
<th>TensorFlow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td>C++, Python</td>
<td>Lua</td>
<td>Python (Lasagne)</td>
<td>Python Inception</td>
</tr>
<tr>
<td><strong>Pretrained</strong></td>
<td>Yes++</td>
<td>Yes++</td>
<td>Yes</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Data</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Model</strong></td>
<td>No</td>
<td>Yes</td>
<td>Experimental</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Readable Source Code</strong></td>
<td>Yes (C++)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Good at RNN</strong></td>
<td>No</td>
<td>Mediocre</td>
<td>Yes</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Higher-Level APIs</strong></td>
<td>No</td>
<td>No</td>
<td>Keras</td>
<td>Keras and TFLearn</td>
</tr>
</tbody>
</table>
A Simple Deep Net in TFLearn

[notebook]
Synaptic Pruning
(Stochastic) Gradient Descent

Adam = AdaGrad + RMSprop
Backpropagation

computes error & gradient of cost function

\[ \delta^L = \nabla_{\alpha} C \odot \sigma'(z^L) \]  
\[ \delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l) \]  
\[ \frac{\partial C}{\partial b_{j}^l} = \delta_{j}^l \]  
\[ \frac{\partial C}{\partial w_{jk}^l} = a_{k}^{l-1} \delta_{j}^l \]
Deep Learning

Overfitting
...and avoiding it

- L1/L2 regularization
- dropout
- artificial data set expansion
Improving Neural Networks
Mostly Hyperparameter Tuning

- problem simplification
- number and width of layers
- cost fxn: quadratic, cross-entropy, log-likelihood, &c.
- more epochs, early stopping
- clever initialization of weights and biases
- learning rate $\eta$, variable schedule
- regularization parameter $\lambda$
- mini-batch size
- automation, e.g., with Spearmint

[Summary Blog Post]
Universality
Solve Any Continuous Function (Nielsen, 2015)
Unstable Gradient

Typically Vanishes (but can Explode)
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
Reinforcement

---

**Revolution of Depth**

152 layers

ImageNet Classification top-5 error (%)

- **ILSVRC’15 ResNet**
  - 3.57

- **ILSVRC’14 GoogleNet**
  - 6.7

- **ILSVRC’14 VGG**
  - 7.3

- **ILSVRC’13**
  - 11.7
  - 8 layers

- **ILSVRC’12 AlexNet**
  - 16.4
  - 8 layers

- **ILSVRC’11**
  - 25.8
  - Shallow

- **ILSVRC’10**
  - 28.2
Classic Deep Architectures

...introducing Convolutional Layers

Diagram of AlexNet and VGGNet architectures, showing layers such as Conv, Pool, FC, and Softmax.
Outline

1 Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2 Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3 Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
Hubel & Wiesel (1959)
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Hubel & Wiesel, 1968
AlexNet
Krizhevsky, Sutskever & Hinton (2012)

Conv 1: Edge+Blob
Conv 3: Texture
Conv 5: Object Parts

Fc8: Object Classes

Numerical
Data-driven
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement
ConvNet Visualisation
Yosinski et al. (2015)

[video]
Network Architectures
AlexNet: ILSVRC ‘12 winner
Krizhevsky et al. (2012)

[TFLearn notebook]
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

VGGMNet: ILSVRC ‘14 runner-up

[TFLearn notebook]
## ConvNet in TensorFlow

<table>
<thead>
<tr>
<th></th>
<th>Caffe</th>
<th>Torch</th>
<th>Theano</th>
<th>TensorFlow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td>C++, Python</td>
<td>Lua</td>
<td>Python</td>
<td>Python</td>
</tr>
<tr>
<td><strong>Pretrained</strong></td>
<td>Yes++</td>
<td>Yes++</td>
<td>Yes (Lasagne)</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Data</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Model</strong></td>
<td>No</td>
<td>Yes</td>
<td>Experimental</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Readable Source Code</strong></td>
<td>Yes (C++)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Good at RNN</strong></td>
<td>No</td>
<td>Mediocre</td>
<td>Yes</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Higher-Level APIs</strong></td>
<td>No</td>
<td>No</td>
<td>Keras</td>
<td>Keras and TFLearn</td>
</tr>
</tbody>
</table>
ConvNet in TensorFlow

[notebook]
### ConvNet in Theano

<table>
<thead>
<tr>
<th></th>
<th>Caffe</th>
<th>Torch</th>
<th>Theano</th>
<th>TensorFlow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td>C++, Python</td>
<td>Lua</td>
<td>Python</td>
<td>Python</td>
</tr>
<tr>
<td><strong>Pretrained</strong></td>
<td>Yes++</td>
<td>Yes++</td>
<td>Yes (Lasagne)</td>
<td>Inception</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Data</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Model</strong></td>
<td>No</td>
<td>Yes</td>
<td>Experimental</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Readable Source Code</strong></td>
<td>Yes (C++)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Good at RNN</strong></td>
<td>No</td>
<td>Mediocre</td>
<td>Yes</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Higher-Level APIs</strong></td>
<td>No</td>
<td>No</td>
<td>Keras</td>
<td>Keras and TFLearn</td>
</tr>
</tbody>
</table>
ConvNet in Theano

[Demo]
## ConvNet in Keras

calls TensorFlow or Theano

<table>
<thead>
<tr>
<th></th>
<th>Caffe</th>
<th>Torch</th>
<th>Theano</th>
<th>TensorFlow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td>C++, Python</td>
<td>Lua</td>
<td>Python</td>
<td>Python</td>
</tr>
<tr>
<td><strong>Pretrained</strong></td>
<td>Yes++</td>
<td>Yes++</td>
<td>Yes (Lasagne)</td>
<td>Inception</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Data</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Model</strong></td>
<td>No</td>
<td>Yes</td>
<td>Experimental</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Readable Source Code</strong></td>
<td>Yes (C++)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Good at RNN</strong></td>
<td>No</td>
<td>Mediocre</td>
<td>Yes</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Higher-Level APIs</strong></td>
<td>No</td>
<td>No</td>
<td>Keras</td>
<td>Keras and TFLearn</td>
</tr>
</tbody>
</table>
ConvNet in Keras

calls TensorFlow or Theano

[notebook]
“2.5-dimension” CT Scans

Roth et al. (2015)
Computer-Aided Detection
Shin et al. (2016); Roth et al. (2016)

Experimental Results (~100% sensitivity but ~40 FPs/patient at candidate generation step; then 3-fold CV with data augmentation)

- Mediastinum
  71% @ 3 FPs (was 55%)

- Abdomen
  83% @ 3 FPs (was 30%)
Kaggle
Data Science Bowl 2017
## Transfer Learning

### Caffe

<table>
<thead>
<tr>
<th>Feature</th>
<th>Caffe</th>
<th>Torch</th>
<th>Theano</th>
<th>TensorFlow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td>C++, Python</td>
<td>Lua</td>
<td>Python</td>
<td>Python</td>
</tr>
<tr>
<td><strong>Pretrained</strong></td>
<td>Yes++</td>
<td>Yes++</td>
<td>Yes (Lasagne)</td>
<td>Inception</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Data</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Parallel GPUs: Model</strong></td>
<td>No</td>
<td>Yes</td>
<td>Experimental</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Readable Source Code</strong></td>
<td>Yes (C++)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Good at RNN</strong></td>
<td>No</td>
<td>Mediocre</td>
<td>Yes</td>
<td>Yes (best)</td>
</tr>
<tr>
<td><strong>Higher-Level APIs</strong></td>
<td>No</td>
<td>No</td>
<td>Keras</td>
<td>Keras and TFLearn</td>
</tr>
</tbody>
</table>
Transfer Learning
Caffe
Transfer Learning

1. Introduction

TensorFlow is an open source library for numerical computation, specializing in machine learning applications. In this codelab, you will learn how to install and run TensorFlow on a single machine, and will train a simple classifier to classify images of flowers.

What are we going to be building?

In this lab, we will be using transfer learning, which means we are starting with a model that has been already trained on another problem. We will then be retraining it on a similar problem. Deep learning from scratch can take days, but transfer learning can be done in a short order.

We are going to use the Inception v3 network. Inception v3 is a trained for the ImageNet Large Visual Recognition Challenge using the data from 2012, and it can differentiate between 1,000 different classes, like Dalmatian or dishwasher. We will use this same network, but retrain it to tell apart a small number of classes based on our own data.

What you will learn
Video Classification

[video]
Antecedents
Case Study: A History of Biological & Artificial Vision
Building Blocks

Theory
Biological & Artificial Neurons
Neural Networks
Deep Neural Networks

Contemporary Applications
Convolutional Neural Networks
Long Short-Term Memory Recurrent Neural Networks
Deep Learning at untapt
Deep Reinforcement Learning
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Sunspring
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt

Reinforcement

Sunspring

We see H pull a book from a shelf, flip through it while speaking, and then put it back.

H

In a future with mass unemployment, young people are forced to wait
a hundred years old.

E

You should see the boys and shut up. I was the one who was going to
be a hundred years old.

E

I saw him again, the way you were

H

You look at him for a moment, then smiles at him.

E

You don't have to be a doctor.

H

I am not sure. I don't know what

E

You want to see you too.

H

I want to see you too.

E

What do you mean?

H

I don't know, but I don't think you

E

(continuing)

H

I was coming to that thing because

E

I don't know. I don't know what

H

That's right.

E

What are you doing?

H

I don't want to be honest with you.

E

He looks at him for a moment, then smiles at him.

H

You don't have to be a doctor.

E

I am not sure. I don't know what

H

You want to see you too.

E

I want to see you too.

H

What do you mean?

E

I'm sorry, but I don't think you

H

I don't know what you're talking

E

The principle is completely

H

(smiling)

It was all about you being true.

E

You didn't even see the movie with

H

I don't know.

E

I don't care.

H

I know that it's a consequence.

E

Whatever you want to know about

H

I don't know. I Just have to ask

E

You have to explain to me what you say.

H

What do you mean?

E

Because I don't know what you're

H

That was the whole time.

E

I know that.

H

I don't know.

E

It would be a good time. I think I
could have been my life.

H

He starts to shake.

E

(smiling)

It may never be forgiven, but that

H

I have to leave, but I'm not doing

E

(smiling)

I was all about you being true.

H

You can't afford to take this

E

I have to leave, but I'm not doing

H

(smiling)

You can't afford to take this

E

We'll think you can still be back

H

Win a round. It's a damn thing to

E

Win a round. It's a damn thing to

H

(smiling)

You can't afford to take this

E

(smiling)

I was all about you being true.

H

You can't afford to take this

E

(smiling)

I was all about you being true.

H

You can't afford to take this

E

(smiling)

I was all about you being true.

H

You can't afford to take this
A history of language technologies

- Scientists from IBM and Georgetown demonstrate a limited machine-translation system in 1954.
- John Pierce's highly critical report on language technologies published. Funding languishes for decades.
- Dawn of "common task" method. Researchers share data, agree on common methods of evaluation in 1970.
- No US government research funding for machine translation or speech recognition in 1980.
- Siri debuts on iPhone "Hey Siri" in 2010.
RNNs; *LSTM* RNNs

Hochreiter & Schmidhuber (1997)
Graves, … & Schmidhuber (2009)
Vector Space Embedding

Word2Vec: Mikolov, ... & Dean (2013)

- Male-Female
- Verb tense
- Country-Capital
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Hinton & van der Maaten (2008)
Word2Vec + t-SNE
Word2Vec + t-SNE
‘Understand’ Language
with Word2Vec features in your model

```python
model.most_similar(positive=['angular'])
```

```
[(
    'angularjs', 0.9534549117088318),
    ('backbonejs', 0.9315043687820435),
    ('ember', 0.905410647392273),
    ('emberjs', 0.902979103736877),
    ('reactjs', 0.896049439907074),
    ('requirejs', 0.8759748339653015),
    ('coffeescript', 0.8645504713058472),
    ('bootstrap', 0.8554328083992904),
    ('nodejs', 0.8515532612800598),
    ('backbone', 0.8443130254745483)]
```

```python
model.most_similar(positive=['managed'])
```

```
[(
    'oversaw', 0.8659406900405884),
    ('directed', 0.8491166234016418),
    ('supervised', 0.8058902621269226),
    ('coordinated', 0.7858685851097107),
    ('led', 0.7539615035057068),
    ('orchestrated', 0.7211644649505615),
    ('supported', 0.7198437452316284),
    ('comanaged', 0.6774874925613403),
    ('encompassing', 0.6726169586181641),
    ('administered', 0.6706446886665344)]
```

[even with small corpora]
Quick, Draw!
ConvNet + LSTM

[link]
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

Outline

1 Antecedents
Case Study: A History of Biological & Artificial Vision
Building Blocks

2 Theory
Biological & Artificial Neurons
Neural Networks
Deep Neural Networks

3 Contemporary Applications
Convolutional Neural Networks
Long Short-Term Memory Recurrent Neural Networks
Deep Learning at untapt
Deep Reinforcement Learning
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement

untapt
Digital Recruitment Platform
Untapt

Candidate-Side Feedback
Deep Learning
Antecedents
Vision Case Study
Building Blocks
Theory
Neural Units
Neural Nets
Deep Neural Nets
Application
ConvNets
LSTMs
untapt
Reinforcement

untapt
Client-Side Feedback

Your Candidates Care About Most

<table>
<thead>
<tr>
<th>Factor</th>
<th>Your Candidates</th>
<th>All untapt Candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>9.6</td>
<td>9.0</td>
</tr>
<tr>
<td>Work/Life Balance</td>
<td>8.3</td>
<td>8.1</td>
</tr>
<tr>
<td>Compensation</td>
<td>6.2</td>
<td>6.0</td>
</tr>
<tr>
<td>Level of Responsibility</td>
<td>5.5</td>
<td>5.4</td>
</tr>
<tr>
<td>Type of Company</td>
<td>3.8</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Where Your Applicants Live

<table>
<thead>
<tr>
<th>Location</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJ</td>
<td>3</td>
</tr>
<tr>
<td>CA</td>
<td>2</td>
</tr>
<tr>
<td>NY</td>
<td>1</td>
</tr>
<tr>
<td>PA</td>
<td>1</td>
</tr>
</tbody>
</table>

Responsiveness

<table>
<thead>
<tr>
<th>Measure</th>
<th>You</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Time, in Days</td>
<td>8.3</td>
<td>9.6</td>
</tr>
</tbody>
</table>

Experience

<table>
<thead>
<tr>
<th>Measure</th>
<th>You</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Number of Years</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Average</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Previous Roles

<table>
<thead>
<tr>
<th>Measure</th>
<th>You</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Count</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Roles Applied To

<table>
<thead>
<tr>
<th>Measure</th>
<th>You</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>14.1</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Interviewing Elsewhere

<table>
<thead>
<tr>
<th>Measure</th>
<th>You</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent</td>
<td>57%</td>
<td>29%</td>
</tr>
</tbody>
</table>
Multi-Stage Bayesian Regression with PyMC3

Krohn, Rives-Corbett & Donner (2016)
Area Under Curve of ROC for Individual Jobs

Before Fitting Job-Specific Bayesian Model

Krohn, Rives-Corbett & Donner (2016)
Give me one bullet-point from your resume:

- Sat around all day checking my Facebook feed
  I predict a 0.0% chance of interview

Give me one bullet-point from your resume:

- Developed trading applications in Python
  I predict a 24.6% chance of interview

Give me one bullet-point from your resume:

- Developed python solution for Monte Carlo risk calculation using numpy, scipy and pandas, with a Javascript frontend in AngularJS and React
  I predict a 98.1% chance of interview
Outline

1. Antecedents
   Case Study: A History of Biological & Artificial Vision
   Building Blocks

2. Theory
   Biological & Artificial Neurons
   Neural Networks
   Deep Neural Networks

3. Contemporary Applications
   Convolutional Neural Networks
   Long Short-Term Memory Recurrent Neural Networks
   Deep Learning at untapt
   Deep Reinforcement Learning
AlphaGO
Silver et al. (2016)
Deep Q-Learning
Mnih et al. (2015)

[Atari Games]
[OpenAI Universe]

[Google DeepMind Lab]
Deep Learning

Antecedents
Vision Case Study
Building Blocks

Theory
Neural Units
Neural Nets
Deep Neural Nets

Application
ConvNets
LSTMs
untapt
Reinforcement