Deep health: Applications of deep learning in medical imaging

Dr Greg Slabaugh
18 Oct 2017
Healthcare

[Diagram showing health expenditure per capita, 2014 (OECD stat)]

UK

https://en.wikipedia.org/wiki/List_of_countries_by_total_health_expenditure_per_capita
It’s not getting any cheaper...

https://en.wikipedia.org/wiki/List_of_countries_by_total_health_expenditure_per_capita

UK: ~£200B
AI and the Fourth Industrial Revolution

“The last 10 years have been about building a world that is mobile-first. In the next 10 years, we will shift to a world that is AI-first.” (Sundar Pichai, CEO of Google, October 2016)

It is characterized by a range of new technologies that are fusing the physical, digital and biological worlds, impacting all disciplines, economies and industries, and even challenging ideas about what it means to be human.
Let’s disrupt healthcare with AI

Challenges

• Access to data
  ◦ Ethics approval: IRB approval, Caldicott Guardian
  ◦ Privacy

• IT infrastructure

• Large datasets
  ◦ A single CT will contain hundreds, if not thousands of slices (images)
  ◦ Data wrangling

• Regulatory environment

• Risk adversion
My journey in healthcare and AI

Siemens Corporate Research

Medicsight
Strong vs weak AI

**Strong AI**
- Consciousness
- Ability to make judgements, plan, communicate, self-awareness

**Weak AI**
- Focuses on a specific task
- No self-awareness
AI... coming to *(or already in)* a device near you
The AI taxonomy (according to Greg)

- **Weak**
  - Machine Learning
  - Other

- **Strong**
  - Supervised
    - Unsupervised
    - Reinforcement

- **Deep**
  - CNN
- **Shallow**
  - Other (DBN)
The AI taxonomy (according to Greg)
What is machine learning?

• **Machine learning** is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed.

⇒ Learning from data
• Labelled data
• Learn a mapping between inputs and outputs
• Example: face detection

• Dynamic environment
• Computer gets feedback and learns to “win”
• Example: ML playing Atari 2600 games

• No labels
• Computer groups similar data to discover hidden patterns
• Example: “People who bought X also bought Y”
Neural networks
Learning

\[ f \left( \sum_{i=1}^{n} w_i x_i \right) \]

INPUT layer \quad HIDDEN layer \quad OUTPUT layer

BACKPROPAGATION

\[ \frac{\partial (\tilde{y} - y)^2}{\partial w_i} \]
Shallow vs deep

"Non-deep" feedforward neural network

Deep neural network
Big data

ImageNet
- A competition to classify images
- Running since 2010
- 1M+ images
- 1000 object categories
AlexNet (2012)

- AlexNet, a type of Convolutional Neural Network (CNN) won the ImageNet challenge by a large margin (15.4% error, compared to 26.2%). This precipitated a swell of interest in Deep Learning techniques.
Key components

1. **Convolution.** This filters an image. The weights for the filter are learned.

2. **ReLU.** This applies a non-linear transformation to the data. This way, the CNN and find a non-linear mapping between the inputs and outputs.

3. **Pooling.** This combines adjacent pixels in a filtered output. This results in abstraction. The CNN learns more “high level” features (e.g. face, instead of edges).
Deep learning frameworks

- Keras
- MatConvNet
- TensorFlow
- Caffe2
- PyTorch
# Import libraries and modules
import numpy as np
np.random.seed(123)  # for reproducibility

from keras.models import Sequential
from keras.layers import Dense, Dropout, Activation, Flatten
from keras.layers import Convolution2D, MaxPooling2D
from keras.utils import np_utils
from keras.datasets import mnist

# Load pre-shuffled MNIST data into train and test sets
(X_train, y_train), (X_test, y_test) = mnist.load_data()

# Preprocess input data
X_train = X_train.reshape(X_train.shape[0], 1, 28, 28)
X_test = X_test.reshape(X_test.shape[0], 1, 28, 28)
X_train = X_train.astype('float32')
X_test = X_test.astype('float32')
X_train /= 255
X_test /= 255

# Preprocess class labels
Y_train = np_utils.to_categorical(y_train, 10)
Y_test = np_utils.to_categorical(y_test, 10)

# Define model architecture
model = Sequential()
model.add(Convolution2D(32, 3, 3, activation='relu',
input_shape=(1,28,28)))
model.add(Convolution2D(32, 3, 3, activation='relu'))
model.add(MaxPooling2D(pool_size=(2,2)))
model.add(Dropout(0.25))
model.add(Flatten())
model.add(Dense(128, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(10, activation='softmax'))

# Compile model
model.compile(loss='categorical_crossentropy',
optimizer='adam', metrics=['accuracy'])

# Fit model on training data
model.fit(X_train, Y_train, batch_size=32, nb_epoch=10,
verbose=1)

# Evaluate model on test data
score = model.evaluate(X_test, Y_test, verbose=0)
Why is deep learning so… trendy?

Recently there has been a surge of (research, commercial) interest in Deep Learning

1. Large datasets (e.g. ImageNet)
2. New algorithms and toolkits (e.g. TensorFlow, PyTorch, MatConvNet)
3. Graphics Processing Units (GPUs)
CNNs in computer vision

- Image classification (e.g. ImageNet)

- Object detection
Current performance (2.25%) on ImageNet exceeds human ability in image classification.

2017: ImageNet is shutting down – problem solved?

Segmentation

- Image segmentation partitions an image into regions (aka *segments*).

- Essentially, segmentation is a grouping problem. Depending on the image, this can be difficult! Humans are adept at visual grouping!
Deep segmentation network

- Segmentation can be achieved using a CNN, by replacing fully connected layers with an expanding (deconvolutional) path

Generative adversarial networks (GANs)

Image colourisation

Image-to-image networks
Generative adversarial networks (GANs)

- Deep De-Aliasing of MRI
The spine
Clinical challenge

- Up to 20% of cervical spine injuries are missed or receive a delayed diagnosis
- Of these, up to 67% suffer neurological deterioration
- Computer-aided detection may help
- This relies on accurate analysis of the x-ray image
Data

- 336 X-ray images (ethically sourced) from the Royal Devon and Exeter Hospital
- 16 bit unsigned integer format (0 to 65,535)
- Patient ages: 17 to 96
- Different scanners (Philips, Agfa, Kodak, GE, Carestream)
- The data...“bites” (i.e., is challenging)
  - All patients had experienced trauma to the c-spine
  - Degenerative changes
Ground truth

- C3 to C7 vertebrae have been outlined (segmented)
- Using a 20 point boundary representation (4 corners)
- For part of the data, multiple human annotations
Given an X-ray image, where is the cervical spine?

We approach this as a segmentation problem on coarse scale.

We resize an X-ray image to 100x100, along with filled-in annotation.

Goal: build a network to predict the cervical spine (blue) region, given an x-ray image of the cervical spine.
Fig. 3.5 (a) Legends (b) FCN (c) DeConvNet (d) UNet.
CSpine localisation

- We implemented three segmentation networks (FCN, DeConvNet, and UNet)
- Networks had between 1 and 6 million parameters to be learned.
- Training:

\[
\hat{\mathcal{W}}_o = \arg \min_{\mathcal{W}} \sum_{n=1}^{N} L_t(\{x^{(n)}, y^{(n)}\}; \mathcal{W})
\]

- Since the CSpine region should be a single connected component, we introduce a novel region-aware term to penalise more than one predicted region

\[
\hat{\mathcal{W}} = \arg \min_{\mathcal{W}} \sum_{n=1}^{N} L_t(\{x^{(n)}, y^{(n)}\}; \mathcal{W}) + L_r(\{x^{(n)}, y^{(n)}\}; \mathcal{W})
\]

- Training takes 30 hours using a dual NVidia Quadro M4000 workstation.
Localisation results

~96% pixel accuracy across dataset
Keypoint (landmark) detection

- Classic problem in computer vision: identify key points in an image.
- For example, pedestrian detection (Hough forest)
- Each patch classified as a pedestrian “votes” for the centre
- Output can be interpreted as a probability distribution

Inspired by Hough forest, we recast the keypoint detection problem as a regression problem.

Objective: regress a probability distribution over the image space to identify key points (such as the vertebral centre or corners)

Essentially, this converts an image into a probability distribution

Ground truth has interobserver variation:
Can model centre as a 2D Gaussian
From the localisation, we know roughly where the cervical spine is located
We form patches along the length of the cervical spine
Probabilistic spatial regression: training

- For each patch, we regress a continuous function over the 2D image.
- Various loss functions (squared difference, Bhattacharyya distance, etc.)

\[
L(\{x, P_{GT}\}; W) = \frac{1}{2|\Omega_p|} \sum_{i \in \Omega_p} \sum_{j=1}^{2} w_i (\hat{y}_i^j - P_{GT_{i,channel=j}})^2
\]

![Input patch]![GT]  
Input patch  GT  

![Input patch]![GT]  
Input patch  GT
Probabilistic spatial regression: testing

- Patches are put into the network, which predicts a probability map
- Accumulate over all patches and detect peaks × GT, + detection
Quantitative results

<table>
<thead>
<tr>
<th>Test patch creation</th>
<th>Fully automatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>True positive rate (TPR)</td>
<td>93.10%</td>
</tr>
<tr>
<td>False discovery rate (FDR)</td>
<td>9.40%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distance error (mm)</th>
<th>Median</th>
<th>Mean</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.54</td>
<td>1.72</td>
<td>0.99</td>
</tr>
</tbody>
</table>
Corner detection

- Spatial probability produced using a spatial normalisation layer at the end of the network.
- Results: 1.54 mm accuracy (average), 0.99 (median), 38% improved
Boundary detection
Boundary detection – full CSpine
Shape

- The boundary of a segment provides a *shape*.
- We can collect and align a large number of shapes and do interesting things, such as
  - Population statistics: mean, median, etc.
  - Principal component analysis (PCA) to study variation in the data
  - (others)

Registered (aligned) C4 vertebra shapes
Objective: perform a precise segmentation of vertebral bodies in x-ray images given a collection of manually segmented images.

We implemented a deep segmentation approaches.

For training, we introduce a novel shape aware term that penalises differences between the ground truth shape and the predicted shape.

$$\hat{W} = \arg \min_{W} \sum_{n=1}^{N} L_t(\{x^{(n)}, y^{(n)}\}; W) + L_s(\{x^{(n)}, y^{(n)}\}; W)$$

Results show 35% improvement over non-deep learning methods.
Shape regression network
Results

Image  
UNet  
Our method
Functional Magnetic Resonance Imaging (fMRI) provides opportunity to observe neural activity in brain.

- Active brain regions need nutrients, thus drawing these from oxygen-enriched blood.
- fMRI can differentiate between oxygen-rich and deoxygenated blood.

Applications

- Understanding cognition
- Identification of brain function altered by brain disorders
The brain as a network

• Brain – a complex network controlling all body functions
• Studies have revealed that the functional connectivity of brain networks are altered in a particular neurological disorders
Functional connectivity

AAL template

Functionally connected regions
Attention Deficit Hyperactive Disorder (ADHD)

- One of the most common childhood disorder
- 3-5% children affected
- Underlying mechanism not clearly understood
- No single diagnostic test
- Diagnostics typically based upon symptoms observed for months
ADHD-200

- Resting state fMRI
- Multiple institutions
  - NYU
  - NeuroImage
  - Peking University
- Brains parcelleated into 90 regions, each with a time series signal
We introduce FCNet, which is a convolutional neural network that measures functional connectivity.

FCNet has a Siamese architecture.
• We then put FCNet with shared weights into an architecture with a classification network, which makes predictions on the patient level from predicted functional connectivity.
Results

<table>
<thead>
<tr>
<th></th>
<th>NI</th>
<th>Peking</th>
<th>NYU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average accuracy [8]</td>
<td>56.9%</td>
<td>51.0%</td>
<td>35.1%</td>
</tr>
<tr>
<td>Highest accuracy [11]</td>
<td>--</td>
<td>58%</td>
<td>56%</td>
</tr>
<tr>
<td>Clustering method [1]</td>
<td>44%</td>
<td>58.8%</td>
<td>24.3%</td>
</tr>
<tr>
<td>Correlation</td>
<td>52.0%</td>
<td>52.9%</td>
<td>56.1%</td>
</tr>
<tr>
<td>FCNet [7]</td>
<td>60.0%</td>
<td>62.7%</td>
<td>58.5%</td>
</tr>
<tr>
<td>DeepFMRI</td>
<td><strong>67.9%</strong></td>
<td><strong>62.7%</strong></td>
<td><strong>73.1%</strong></td>
</tr>
</tbody>
</table>

- We have observed differences frontal lob functional connectivity is altered the most in this dataset.
- Frontal lobe controls important cognitive skills, including emotional expression, problem solving, memory, and behaviour.
Deep learning

Hype, or hope?
Deep learning limitations

- Requires large datasets
- Methods described in this talk also require *labelled* data
- Algorithms are complex
  - Slow to train (but fast at test time)
  - Difficult to interpret results (Explainable AI)
  - Black-box
- Biologically inspired, but don’t capture the biological mechanisms of the brain
- Limited theoretical understanding
- Hyperparameters
Thanks

- City’s Computer Vision Group: [http://computervision.city.ac.uk/](http://computervision.city.ac.uk/)
- In particular,

![Arif](image1.jpg)  
Arif

![Atif](image2.jpg)  
Atif
References

8. Image-to-Image Translation with Conditional Adversarial Networks, Isola et al., CVPR 2017
Interobserver error - segmentation
CSpine CAD

Get Vertebra Centres

Perform Segmentation

Segmentation Saved

Alignment Curves From Segmentation