MRC
Centre for Environment & Health

Medical Research Council

Imperial College London

Pathways to Equitable Healthy Cities

IndabaX Sudan

10th September 2021

Images to understand the environment: Theory and Application

Emily Muller

PhD Candidate, Imperial College London

Supervised by Majid Ezzati, Seth Flaxman, Emily Dentone

Personal Introduction

- PhD Student in Computer Vision + Urban Analytics, Imperial College London.
- Research Assistant Pathways to Healthy Equitable Cities, ICL.
  1. Transferability: predicting socio-economic outcomes in cities around the world using street view images.
  2. Predicting pollution levels using street view images in Accra.
  3. **PhD research**
- Currently on a 6-month internship at Accenture AI Labs Dublin working on Synthetic Data Generation.

PULSE LONDON

Pathways to Equitable Healthy Cities

MRC Centre for Environment & Health

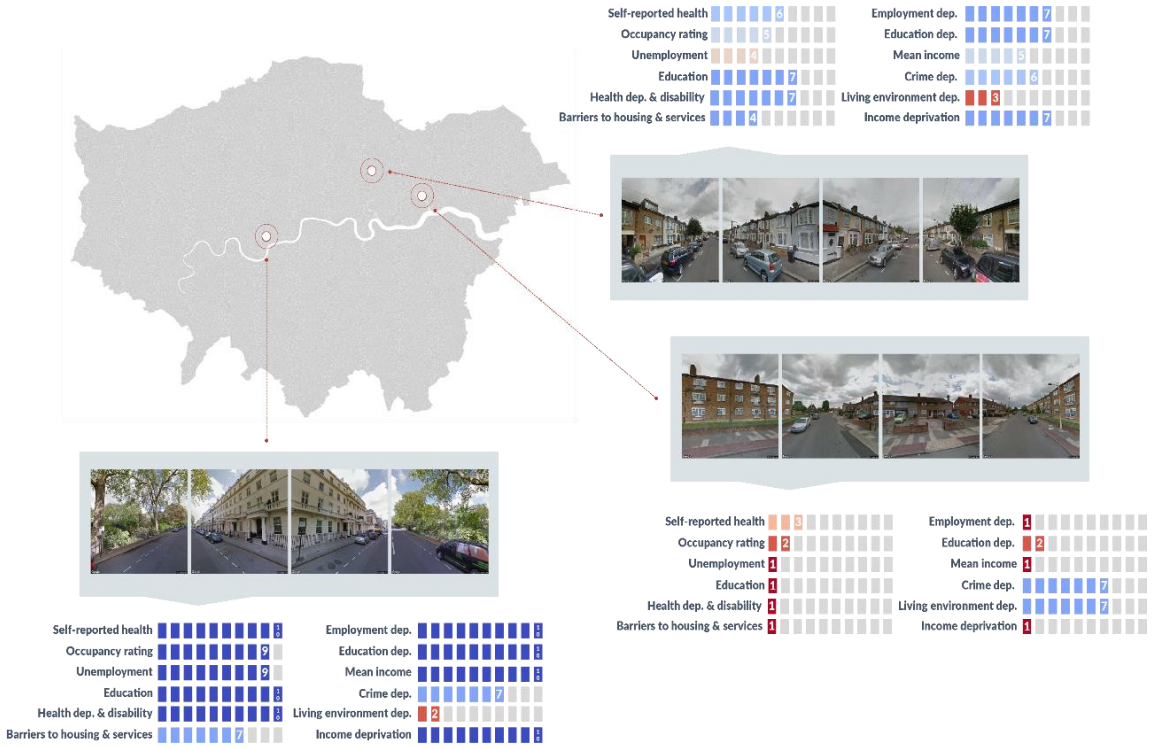
Medical Research Council

Imperial College London

accenture

My PhD - motivation

What are the social determinants of health and can these be reliably encoded with imagery?



Suel et al., 2019. Measuring social, environmental and health inequalities using deep learning and street imagery

<u>Efficacy:</u>
MAE +/-1 decile for income.

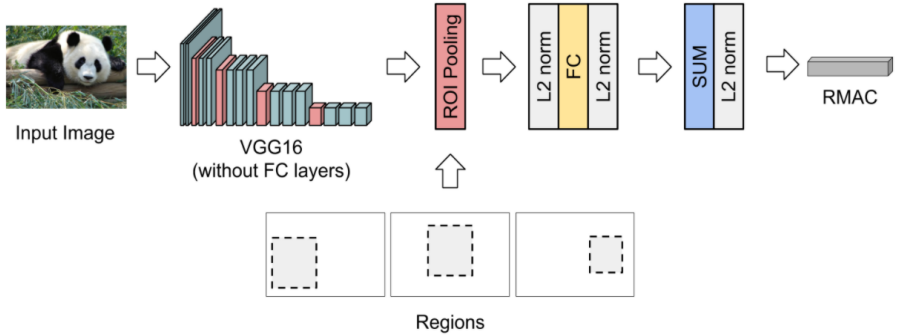
<u>Limitations:</u>
Black Box model
Aggregate outcome assignments.

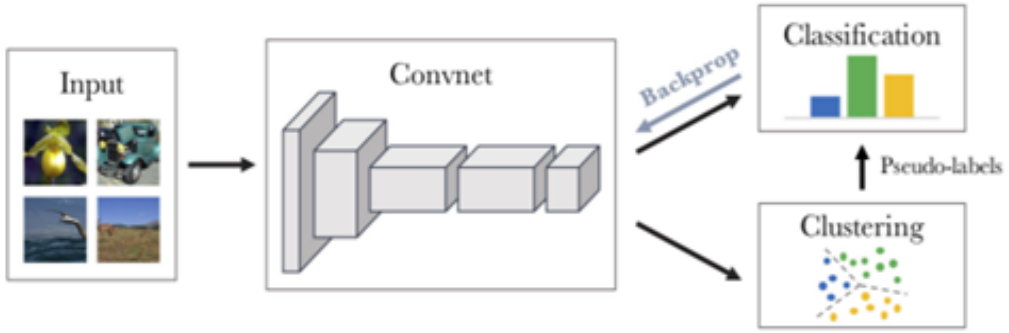
There are no outcome labels at image point level → create it.
Clustering GSV imagery

MRC
Centre for Environment & Health

UKRI Medical Research Council
Imperial College London

How to meaningfully extract information from imagery for clustering?





Top Left: Tolias et al., 2016. R-MAC Architecture uses VGG16 pre-trained on ImageNet
Rop Right: Caron et al., 2018. DeepCluster.

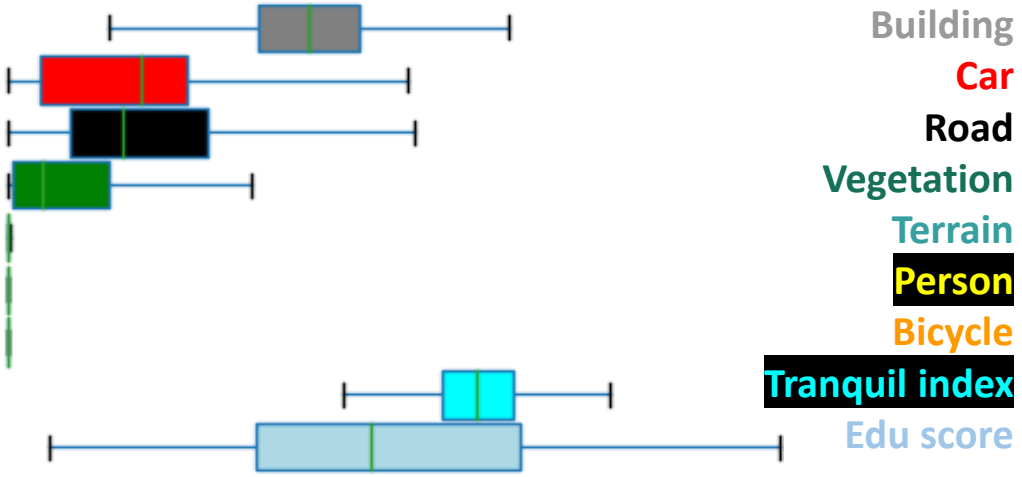
MRC
Centre for Environment & Health

UKRI Medical Research Council
Imperial College London

Interpreting Clusters

Visual inspection

Spatial Distribution + correlation with census



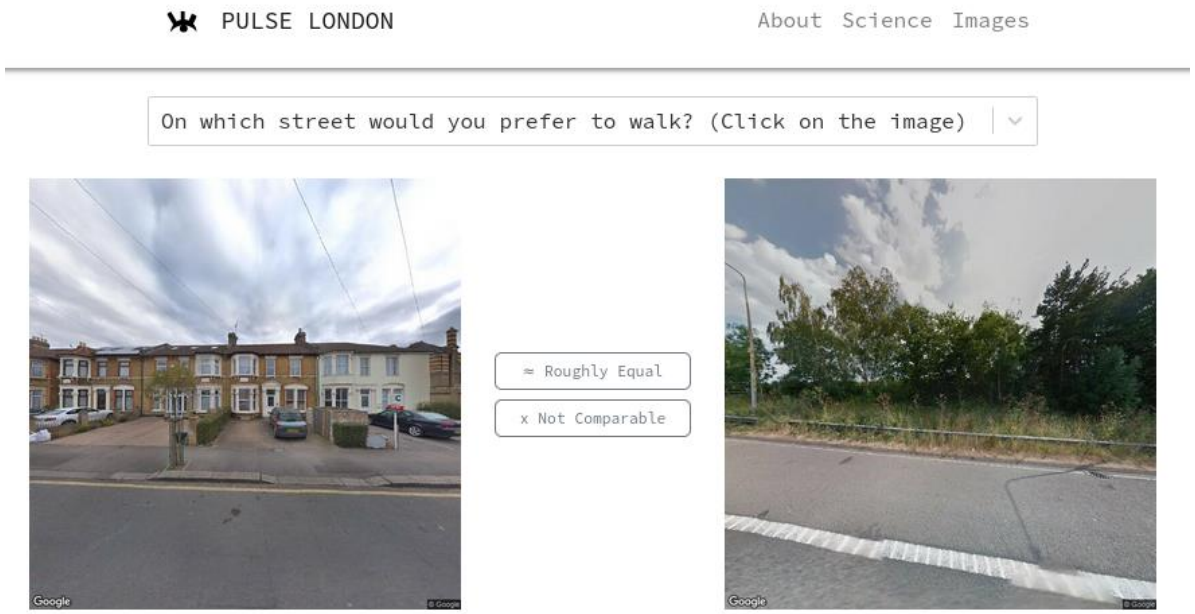
Correlation with features from imagery and external

MRC
Centre for Environment & Health

UKRI Medical Research Council | Imperial College London

Perceptions of images

Place Pulse 2.0, MIT.



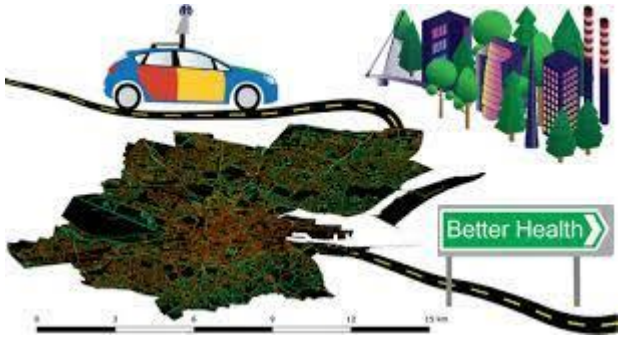
pulselondon.co.uk

- London specific
- Pre-processing
- github.com/emilymuller1991/web-app

MRC
Centre for Environment & Health

UKRI Medical Research Council | Imperial College London

Understanding environmental inequalities, *New York Times article.*

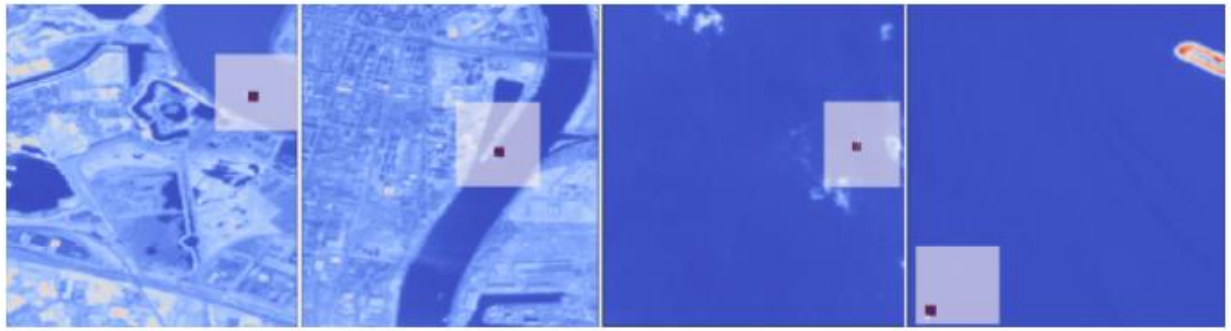


Environmental exposures, *"Biophilic cities", O,Regan 2021-06-23.*

Commercially, *Aerobotics Intelligent tools for agriculture*.

Perceptions of the built environment, *Salesses 2014*.

Illegal fishing, *Michel Deudon 2020.*



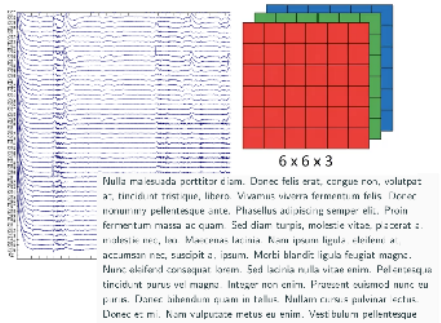
How to meaningfully extract information from imagery for clustering?

Data

Structured Data

p_i	x_1	x_2
1	27	3
2	124	58
3	17	0
⋮	⋮	⋮

Unstructured Data

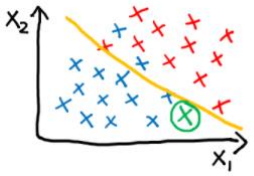


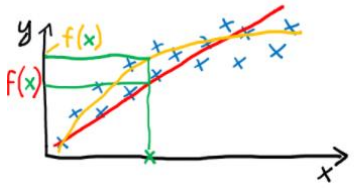
Task

Supervised Learning

Classification
predicts a qualitative output

Regression
predicts a quantitative output



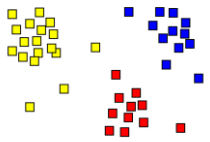


$$f : x \rightarrow y$$

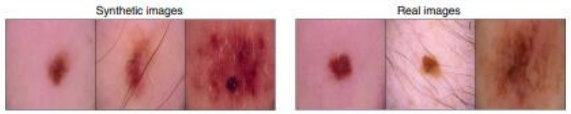
$$\hat{y}(x) = f_\theta(x)$$

Unsupervised Learning

Clustering



Synthesise New Data



Model

Traditional ML Algorithms

- Support Vector Machine
- Logistic Regression
- Linear Regression
- K-means

Deep Learning Models

- Neural Networks
- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks
- Deep Generative Models

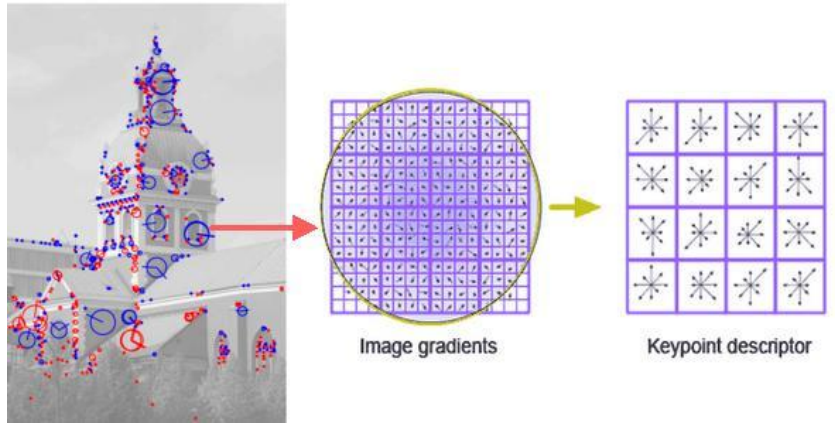
MRC
Centre for Environment & Health

UKRI Medical Research Council
Imperial College London

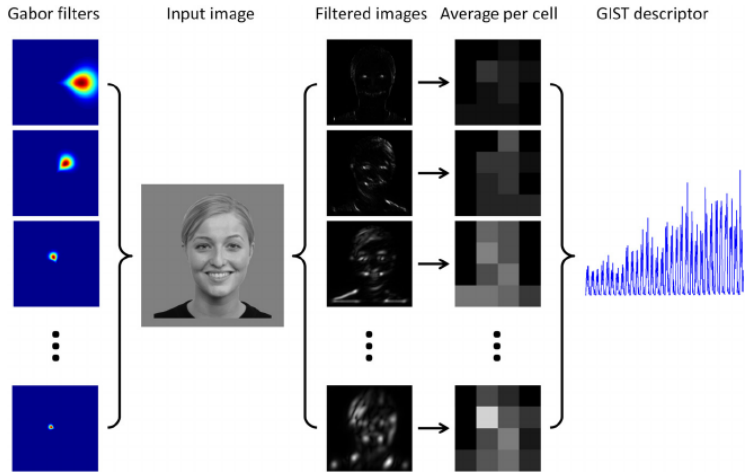
Feature Extraction – Traditional Computer Vision

Traditional Machine Learning Flow

Histogram of Gradients (HoG)



Scale invariant Feature Transformation (SIFT)

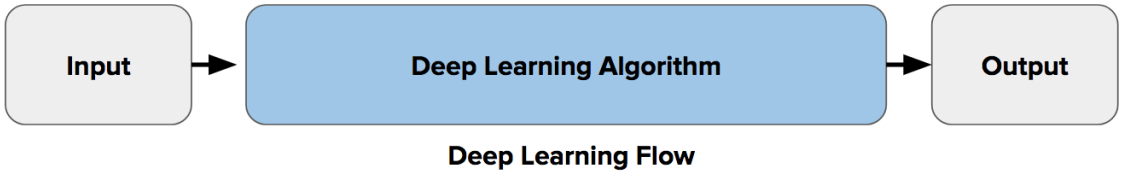


GIST

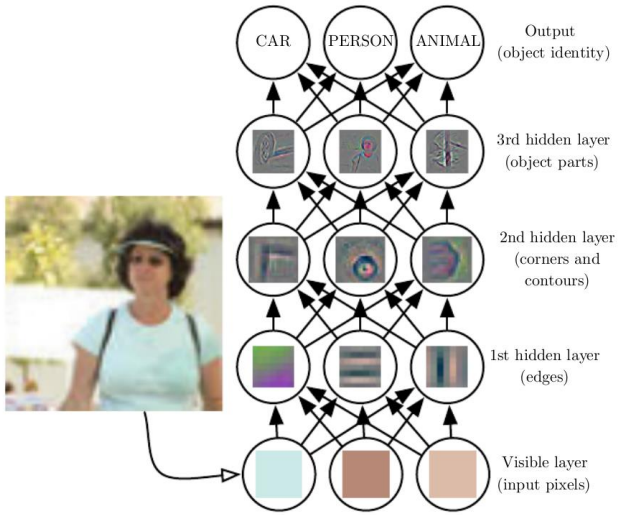
MRC
Centre for Environment & Health

UKRI Medical Research Council
Imperial College London

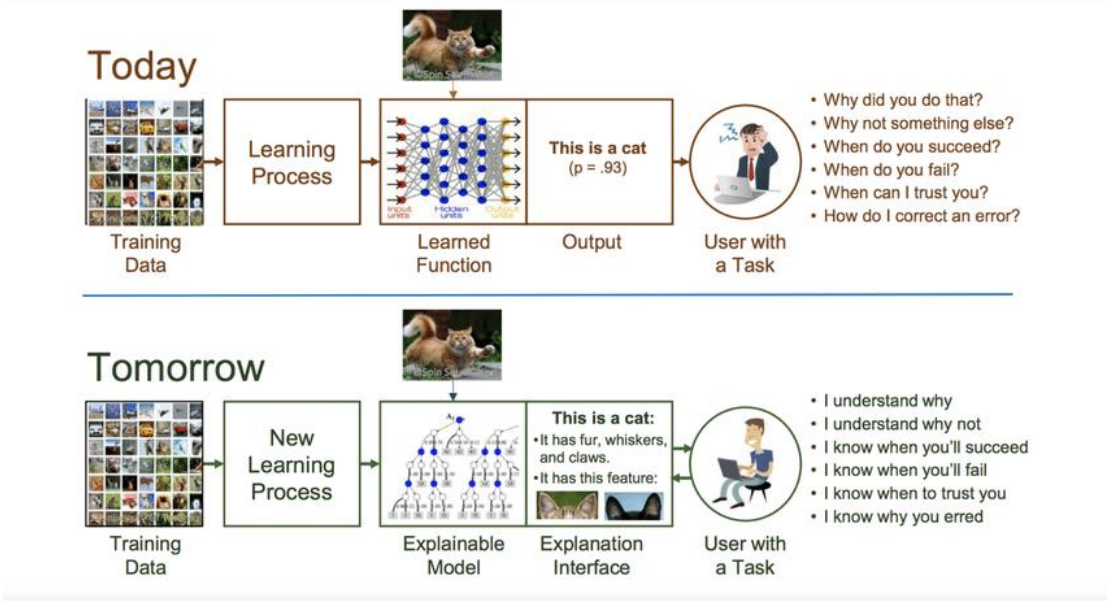
Feature Extraction – Deep Learning



Deep Learning Flow



Build complex representations on top of more simple ones / nested hierarchy of concepts.



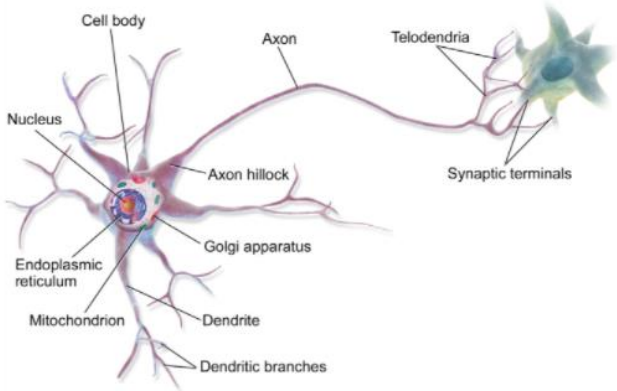
Explainable AI

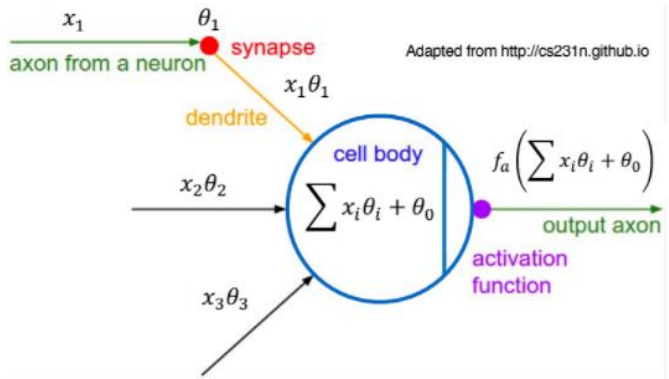
Samek, 2021.

Christopher Molnar, 2021.

MRC
Centre for Environment & Health
UKRI Medical Research Council
Imperial College London

A Brief History





[Cybernetics, 1940 - 1960]

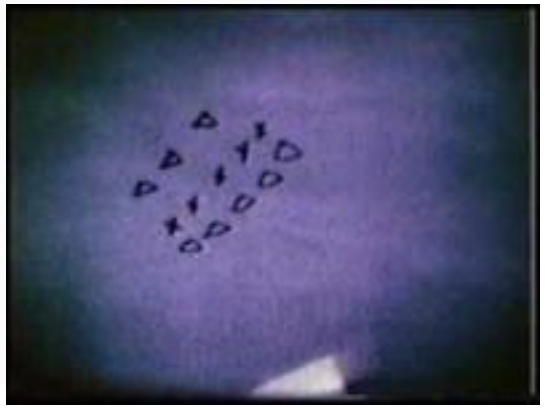
- McCulloch & Pitts, *A logical calculus of the ideas immanent in nervous activity*, 1943.
- Hebb, "Neurons that fire together wire together" - *The Organization of Behavior*, 1949.
- Rosenblatt, The Perceptron, 1958.

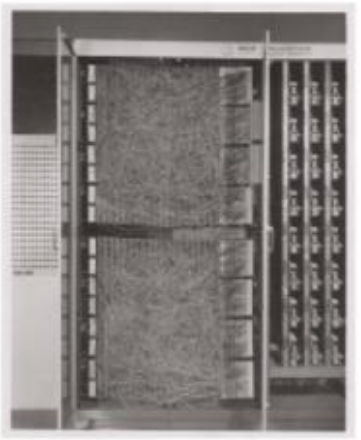
[Connectionism, 1980 - 1990]

- Rumelhart et. al, *Learning representations by back-propagating errors*, 1986.
- Artificial Neural Networks.
- Universal Approximation Theorem, 1991.

[Deep Learning, 2006 - present]

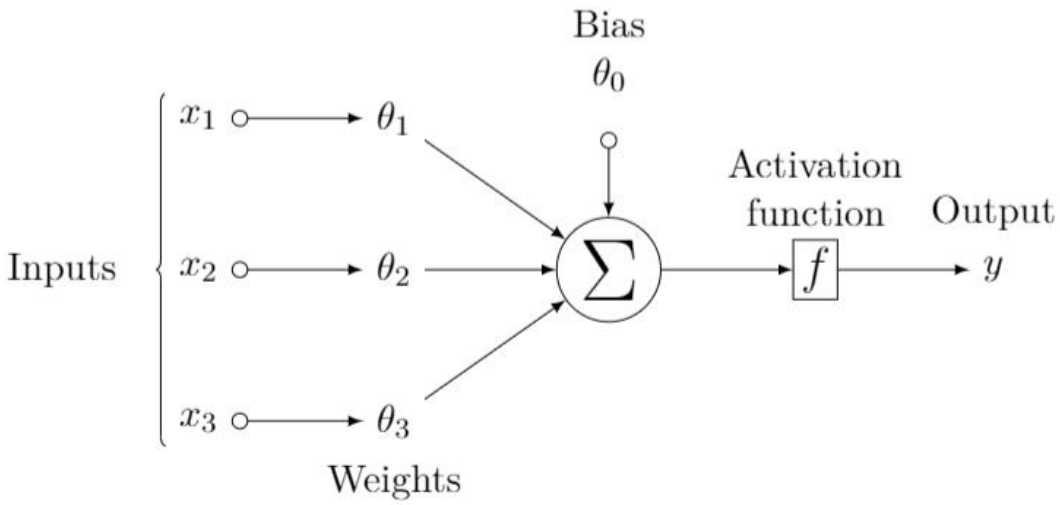
- COMPUTING, SOFTWARE, DATA + SCIENCE.





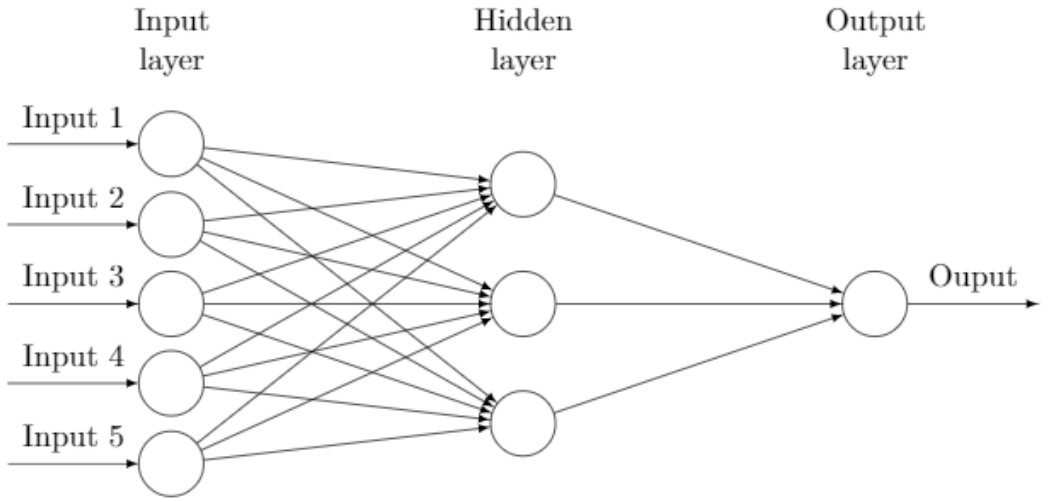
"Mark 1 perceptron" - machine designed for image recognition: it had an array of 400 photocells, randomly connected to the "neurons". Weights were encoded in potentiometers, and weight updates during learning were performed by electric motors (source Wikipedia)

Multilayer Perceptron (MLP)



A single unit (neuron)

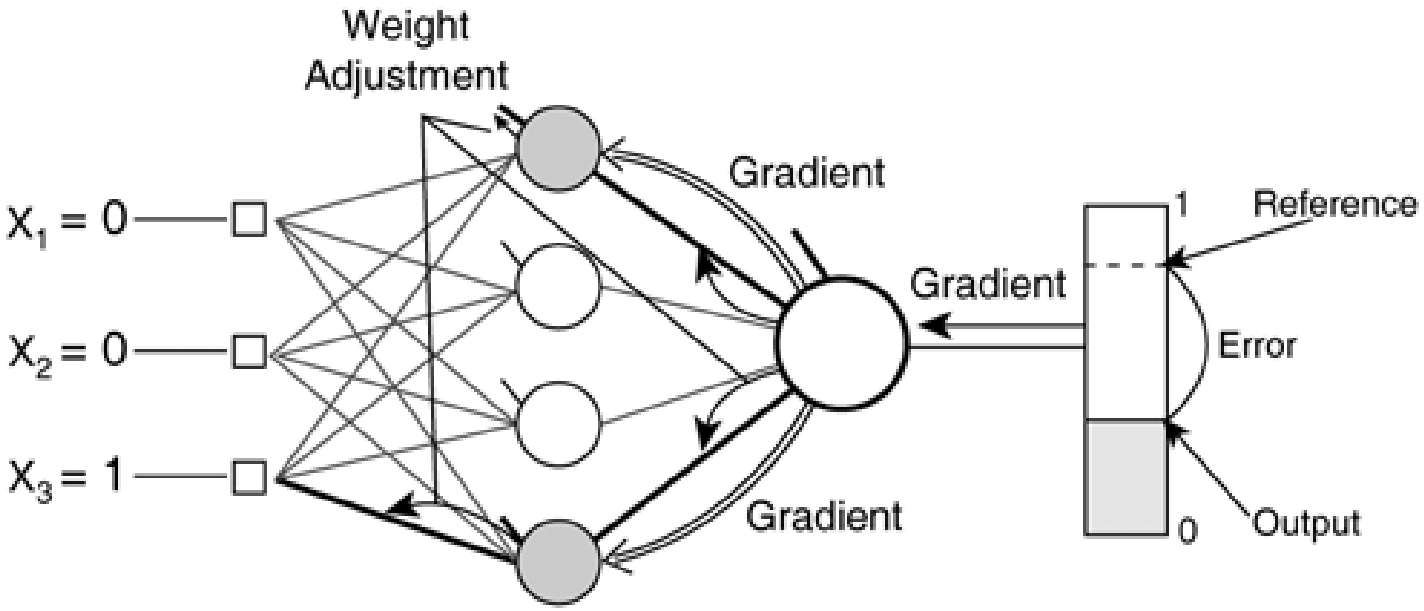
$$\hat{y}(\boldsymbol{x}) = f_a\left(\sum_i x_i\theta_i + \theta_0\right)$$

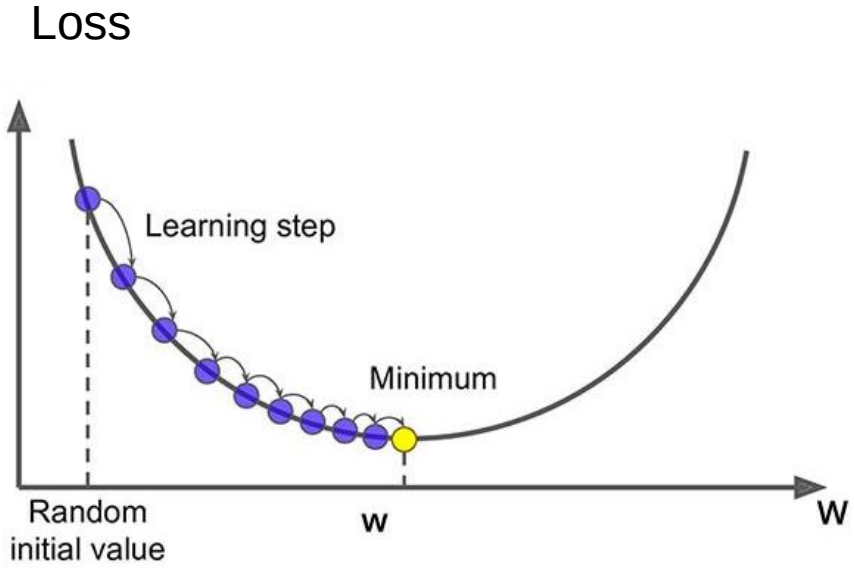


Scaling by depth/width – there exists construction which can approximate any* function → Universal Approximation Theorem.

How does the network learn?

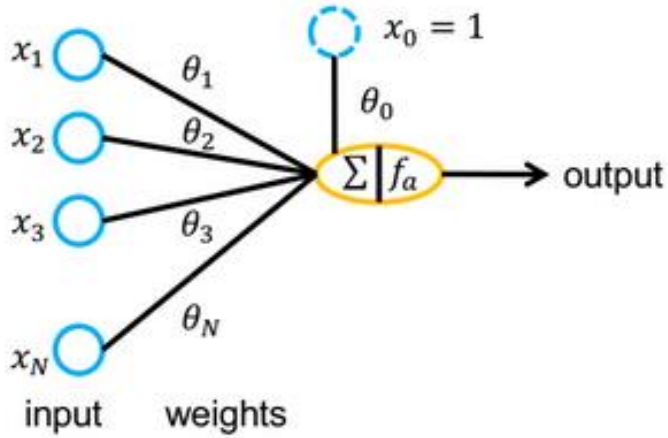
- Initialise weights
- Forward Pass
- Backward Pass





Neural Network Training

Neuron activation: $\hat{y}(\boldsymbol{x}) = f_a\left(\sum_i x_i\theta_i + \theta_0\right)$

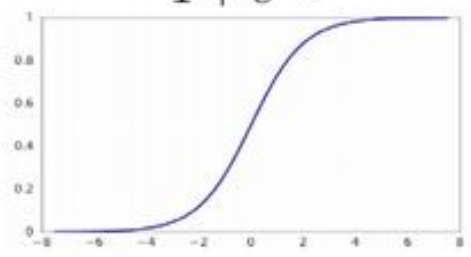


Loss: $\mathcal{L} = \frac{1}{N}\sum_i (y_i - \hat{y}_i)^2$

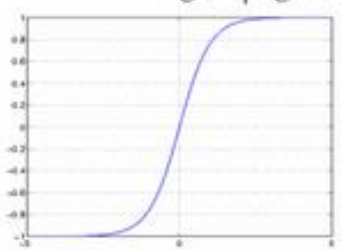
Learning: $\theta = \min_\theta \mathcal{L}(\theta, \mathbf{X}, \mathbf{Y})$ via gradient descent: $\theta_i^{t+1} = \theta_i^t + \frac{\partial \mathcal{L}}{\partial \theta_i^t}$

Activation functions:

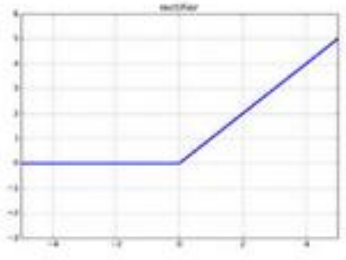
$S(x) = \frac{1}{1+e^{-x}}$



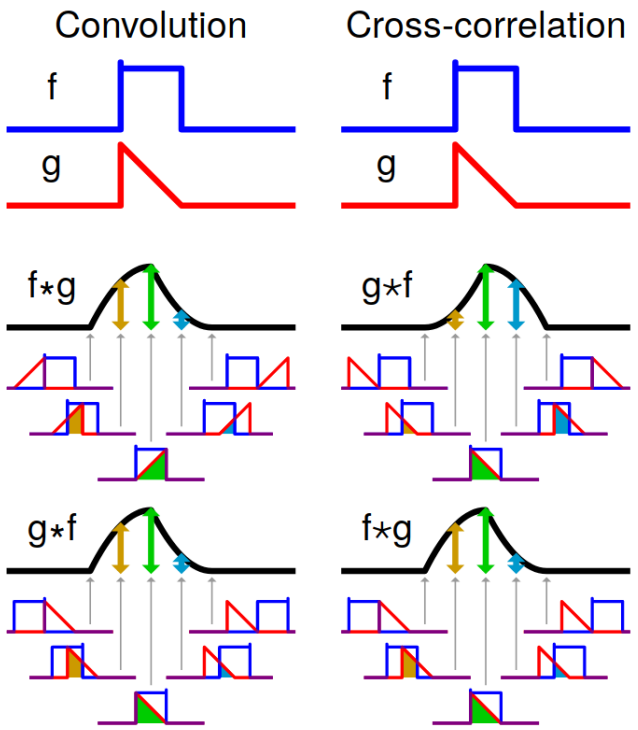
$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$



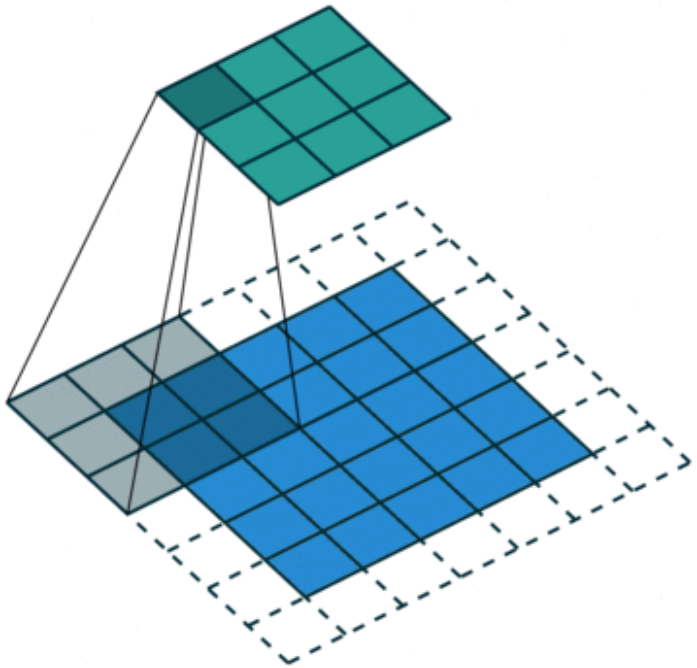
$f(x) = \max(0, x)$

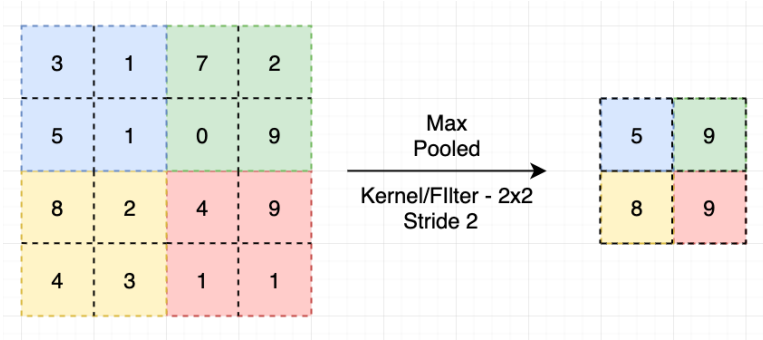


Convolutions



- Digital Signal Processing
- Quantifies amount of *f* in *g*
- ML libraries use cross-correlation no need to flip signal

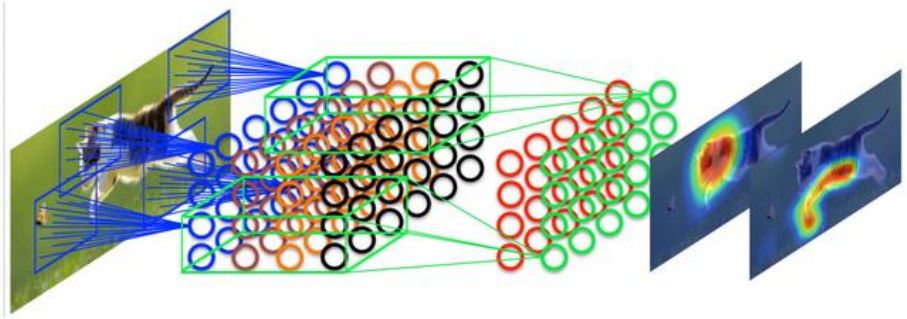


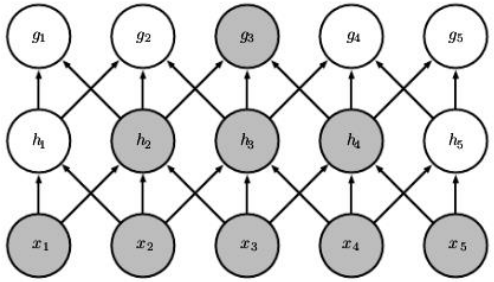


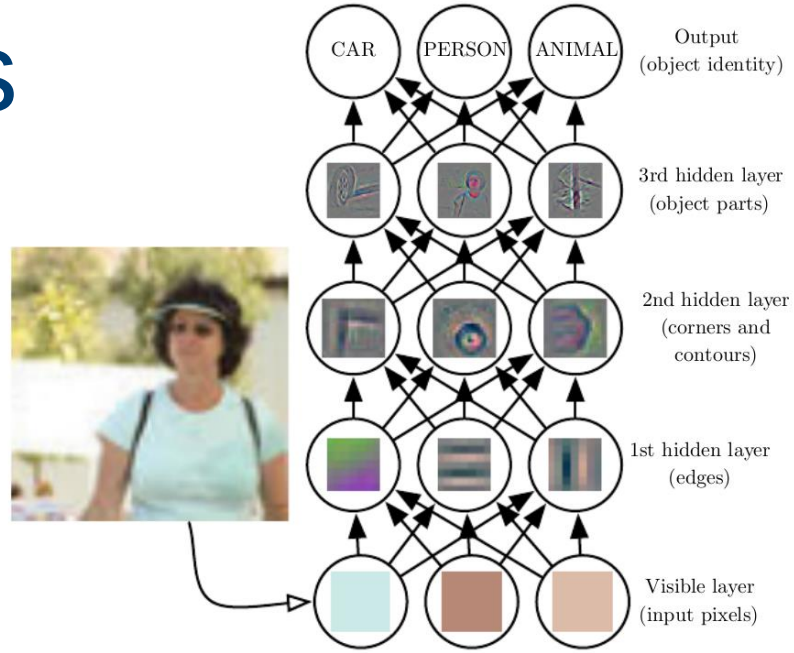
Pooling layer

MRC
Centre for Environment & Health
UKRI Medical Research Council
Imperial College London

Convolutional Neural Networks

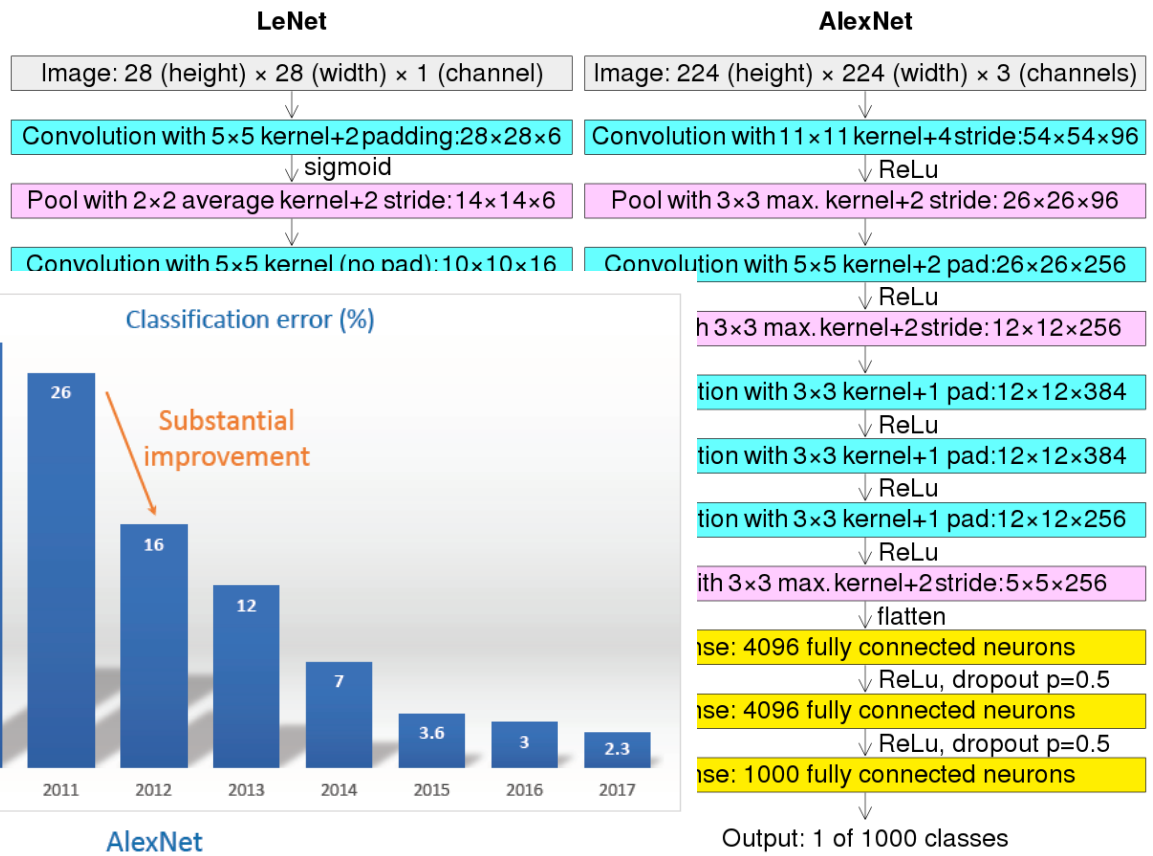






- Unlike MLPs, each neuron is typically only connected to a subset of neurons in the previous layer, known as the receptive field.
- A single CNN layer has multiple filters and therefore detects multiple features per layer.
- Deeper layers process feature maps of the previous layer, increasing the receptive field.
- As depth increases, the composition of previous feature maps extract more and more complex representations of the original image.
- The resulting sparse connectivity, along with shared weights within each layer, reduces the number of trainable parameters.

CNNs existed before 2012

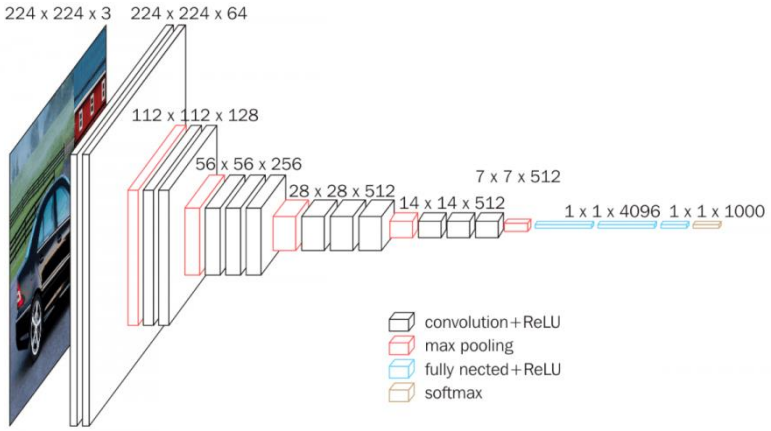


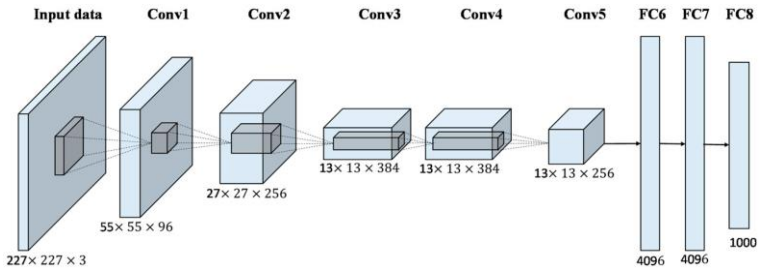
Fei-Fei Li

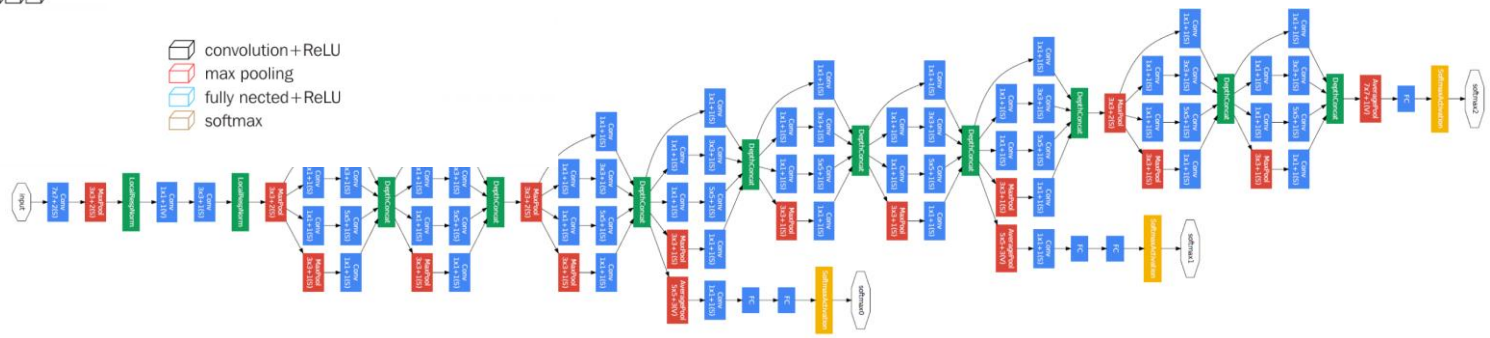
Error for image classification challenges

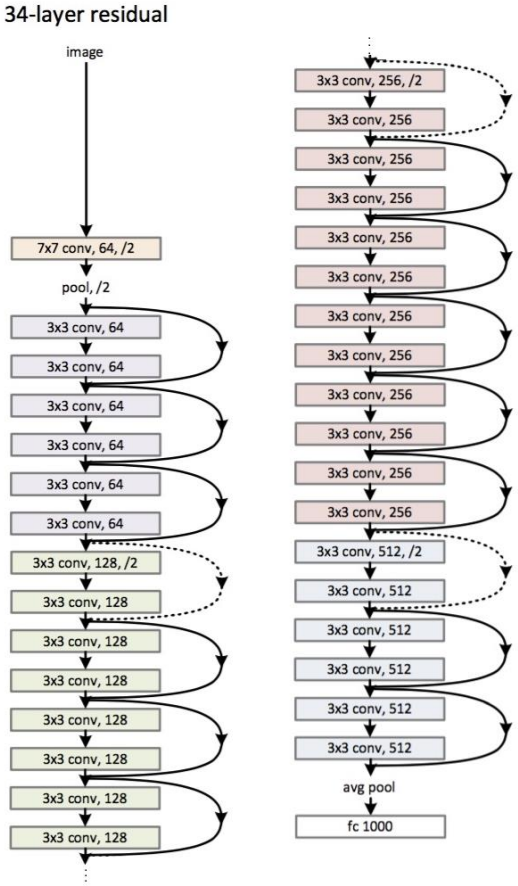
CNNs - (ImageNet Large Scale Visual Recognition Challenge)

Network	Year	ImageNet Top-1 Accuracy	Layers	Trainable Parameters
LeNet-5 [20]	1998	NA	7	60K
AlexNet [18]	2012	63.3	8	61M
VGG-16/19 [22]	2014	74.4/74.5	16/19	138M/144M
Inception V1 [23]	2014	69.8	27	5M
ResNet-50 [24]	2015	77.15	50	23M









Medical Research Council

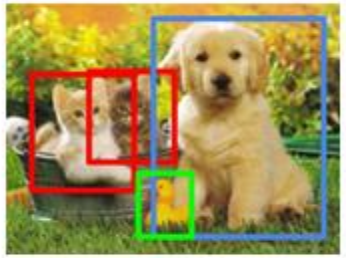
Imperial College London

Image Recognition Tasks

Classification

CAT

Object detection



CAT, CAT,
DUCK, DOG

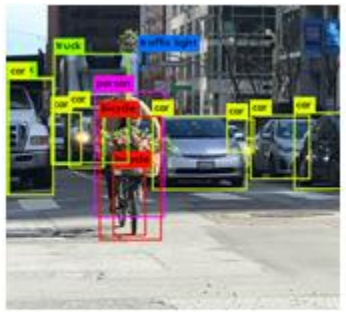
Instance segmentation

CAT, CAT,
DUCK, DOG

Semantic segmentation
Pixel classification

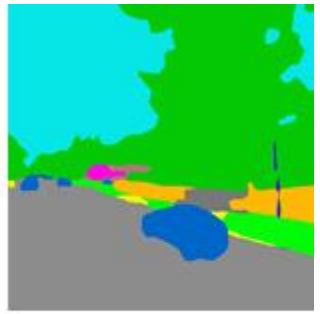
GRASS, CAT,
TREE, SKY

Deprived / poor



Street images
object locations & counts

Building footprint
extraction



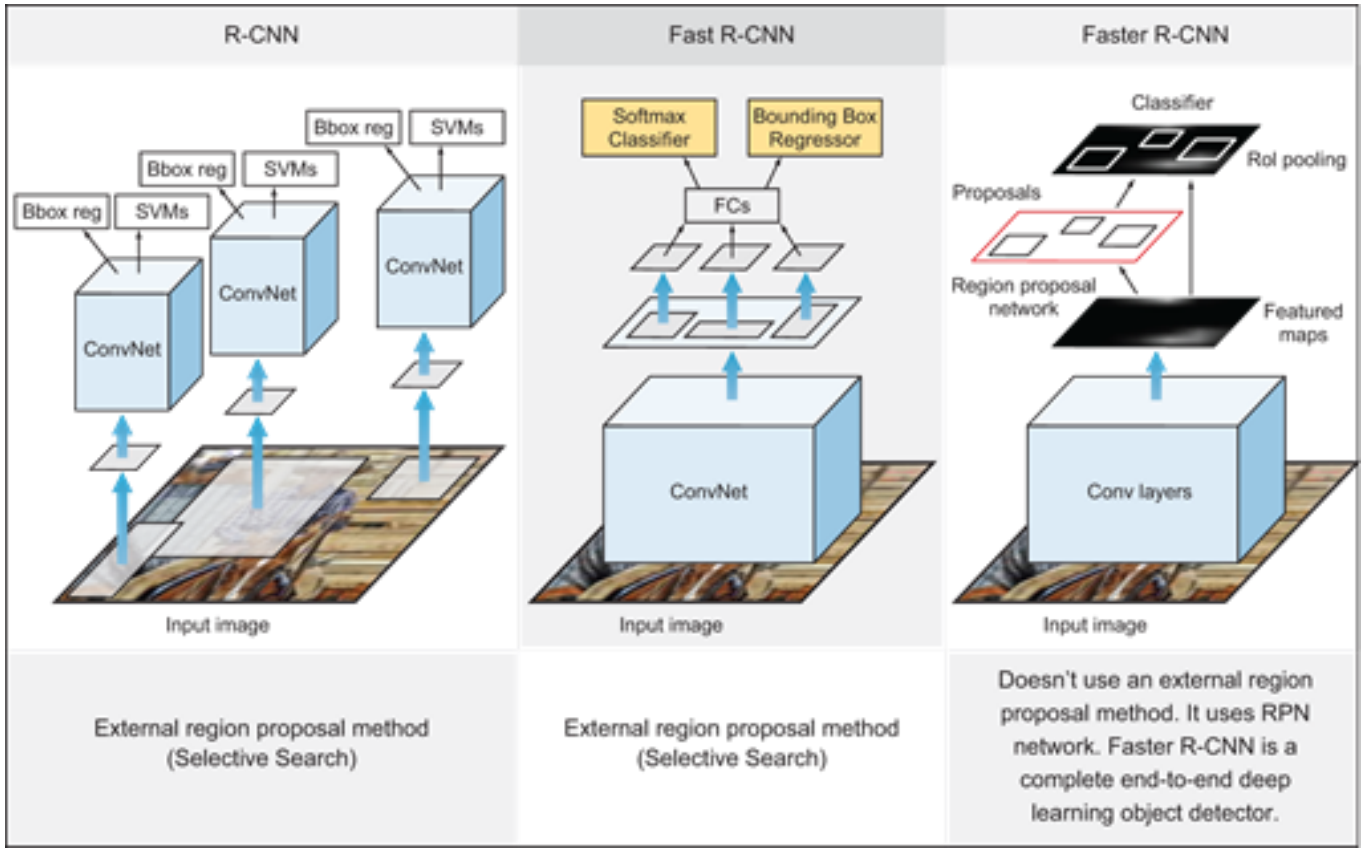
Street images
pixel classification

MRC
Centre for Environment & Health

UKRI Medical Research Council
Imperial College London

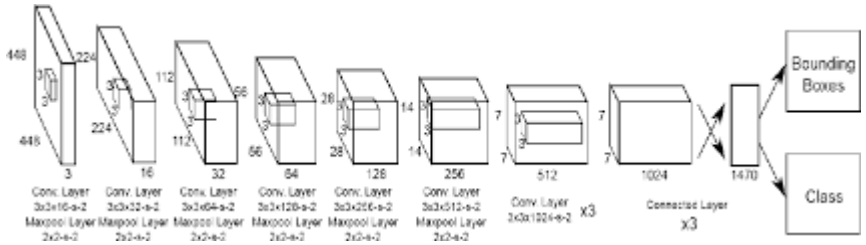
Object Detection

Two Stage Detector



R-CNN, Girschick et al., 2014. Fast R-CNN, Girshick, 2015
Faster R-CNN, Ren et al., 2016.

One Stage Detector

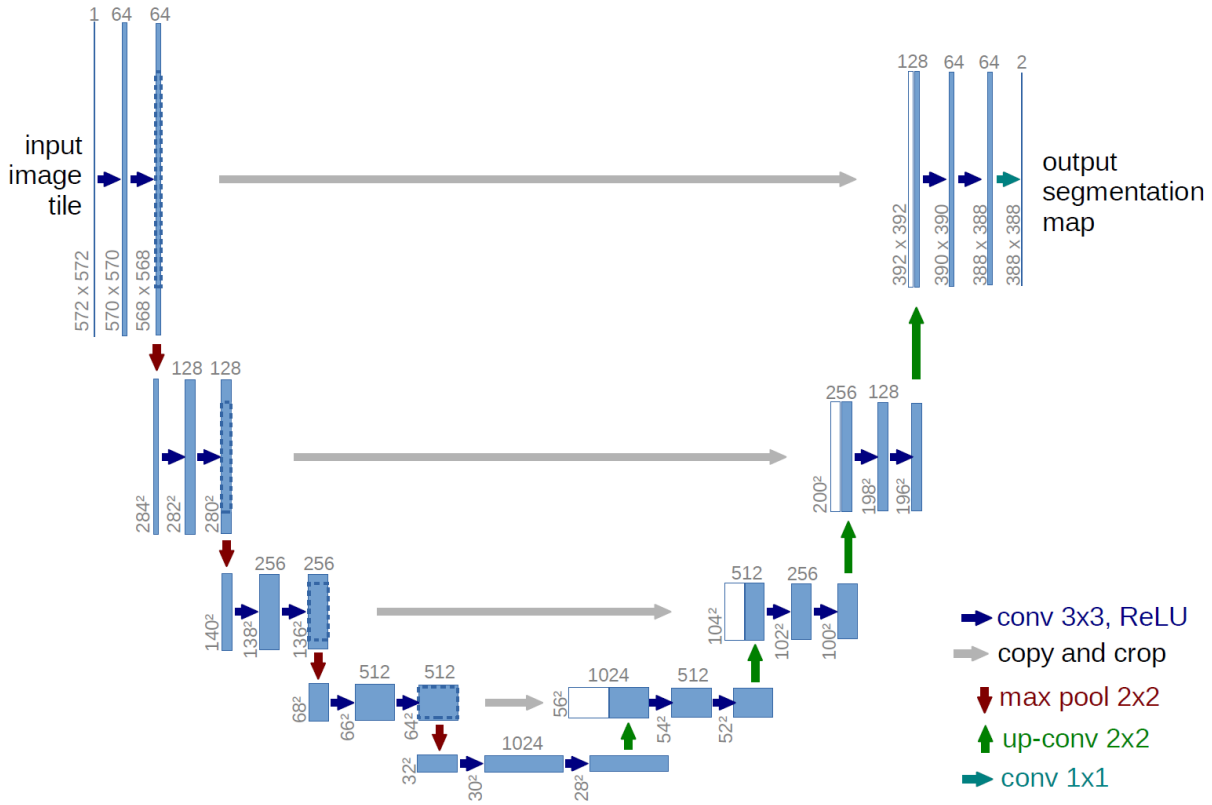


YOLO, Redmon et al, 2015.

MRC
Centre for Environment & Health

Medical Research Council
Imperial College London

Semantic Segmentation



U-Net, Ronnenberger et al., 2015.

MRC
Centre for Environment & Health

UKRI Medical Research Council | Imperial College London

Summary + Resources

- Motivation for the use of imagery
- Feature Extraction
- CNNs
  - Coursera
  - Deeplearningbook.org
  - Deep Learning Indaba Tutorials
- Object Detection + Semantic Segmentation
  - Tensorflow Deeplab API

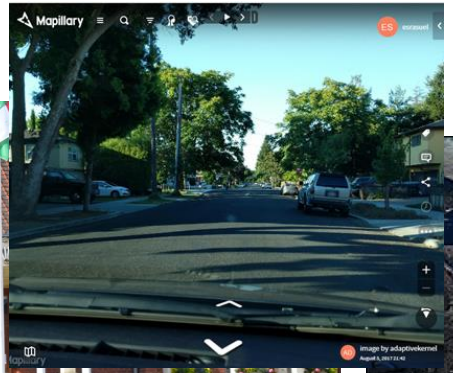
Applications

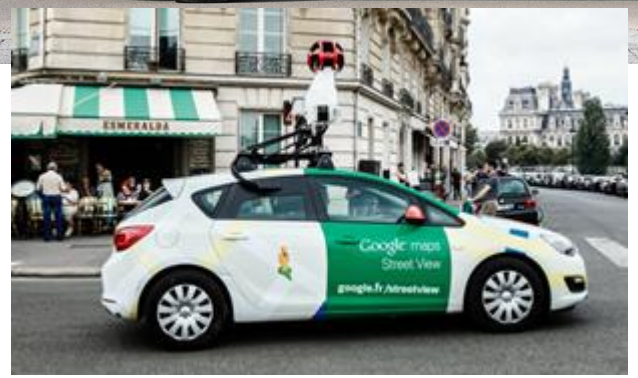
1. Street View Imagery [Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]
2. Satellite Imagery [Slides from Ricky Nathvani, Healthy Data Analytics Course ICL, 2021]
3. Predicting socio-economic status
4. Extracting features to predict pollution levels

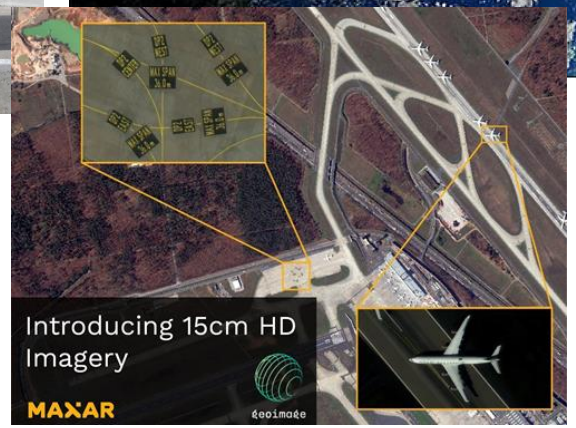
MRC
Centre for Environment & Health
UKRI Medical Research Council
Imperial College London

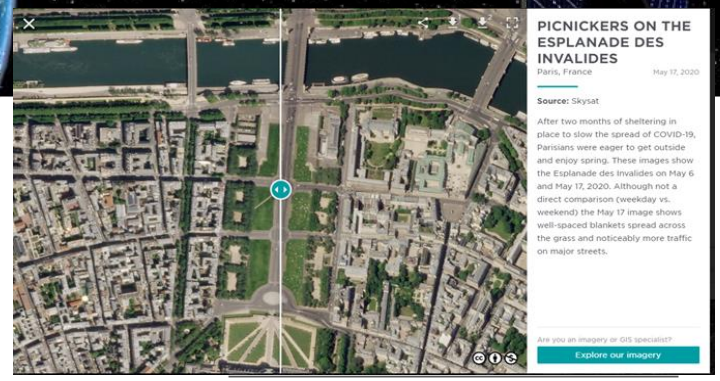
[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

Sources of imagery









UKRI Medical Research Council | Imperial College London

[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

Street View Imagery

Providers: Google Street View (GSV), OpenStreetCams, geo-tagged Flicker images, Mapillary, Bing Maps Streetside...

OpenStreetCam

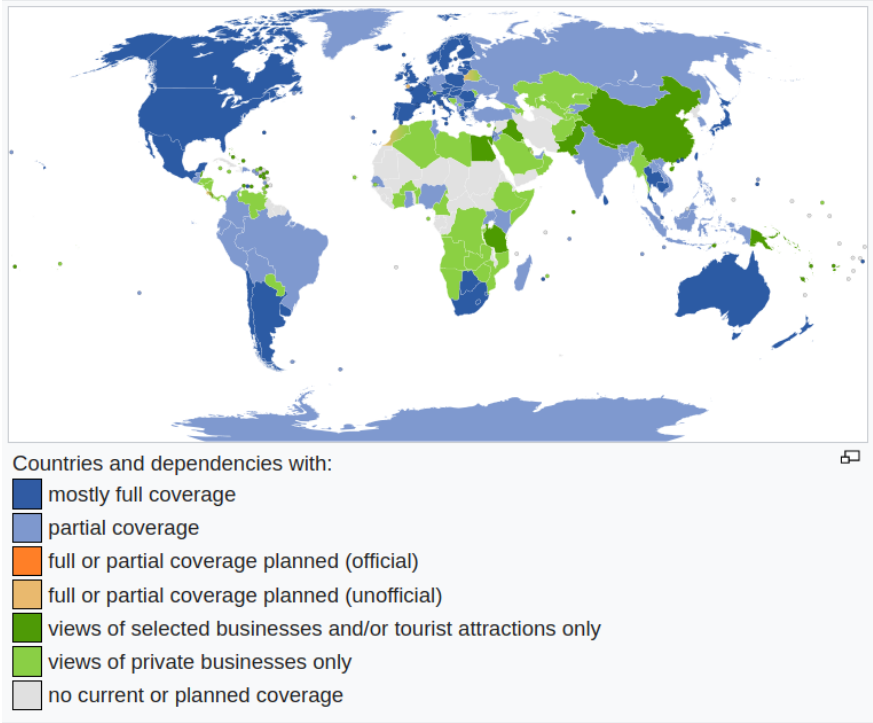
In practice, **GSV** is most commonly used in research.

MRC
Centre for Environment & Health
Medical Research Council
Imperial College London

[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

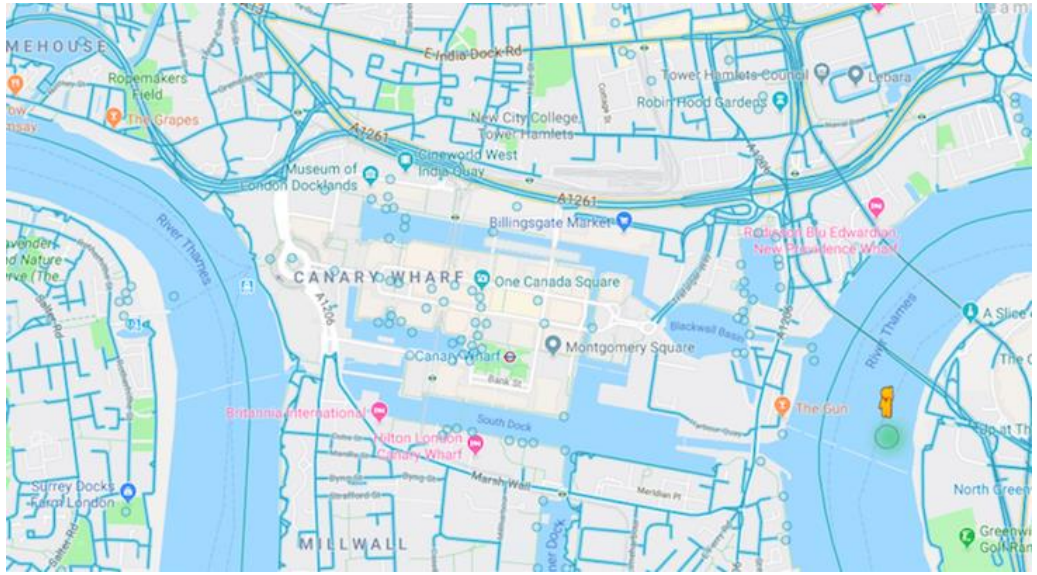
Street view advantages

- View corresponds to personal experience of streets (and hence potential exposures)
  - Small scale street characteristics obtainable from this imagery.
  - High spatial resolution.
- Coverage can be frequent for some countries
- Broadly available (at cost…)



[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

Street view disadvantages

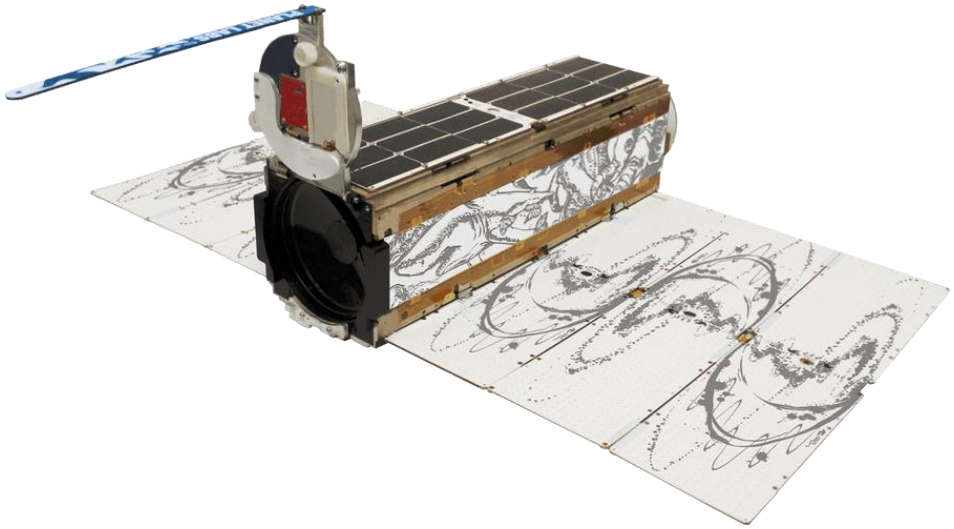


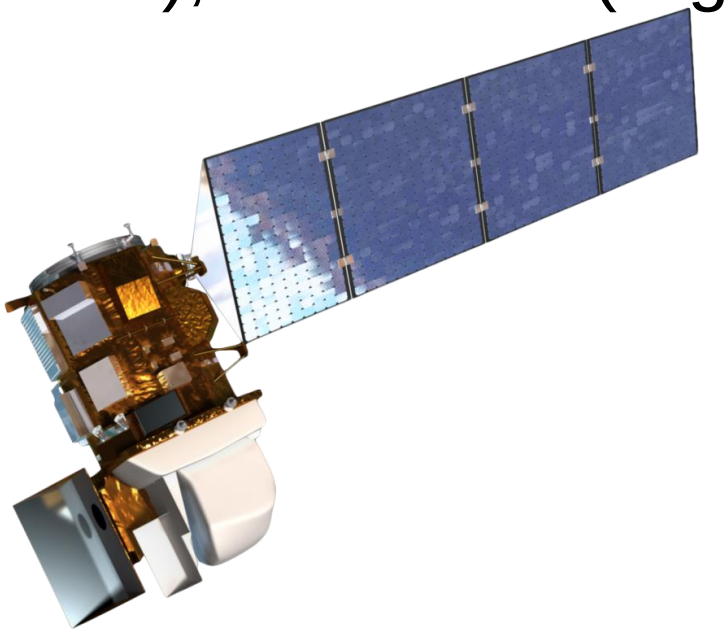
- Each image only covers a small spatial area not contiguous.
- Low temporal resolution (typically annual).
- Large, unwieldy data sets.
- Most comprehensive sources are only commercially available, restricted terms of use.
- Coverage

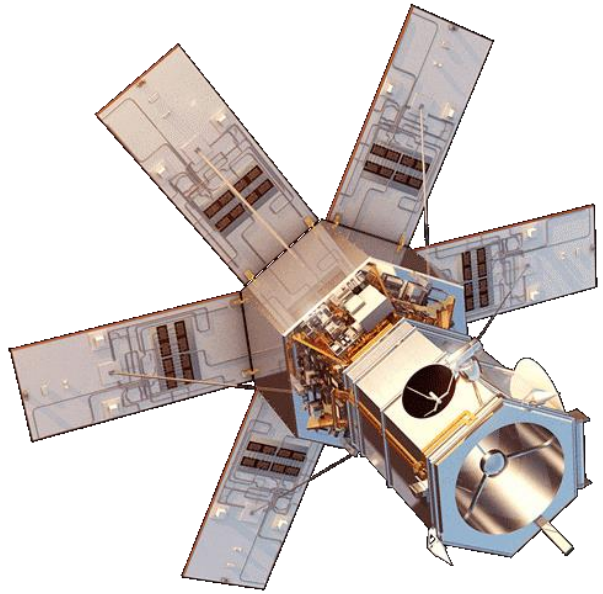
[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

Satellite Imagery

Providers: Planet, LandSat (NASA), Worldview (Digital Globe)...







Provides imagery in the **visible** range and **non-visible** spectra.

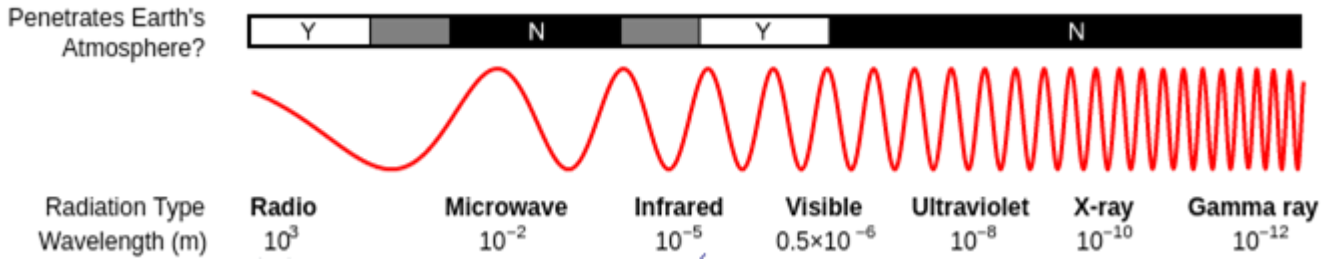
Broad range of resolution: 30m to 0.3m.

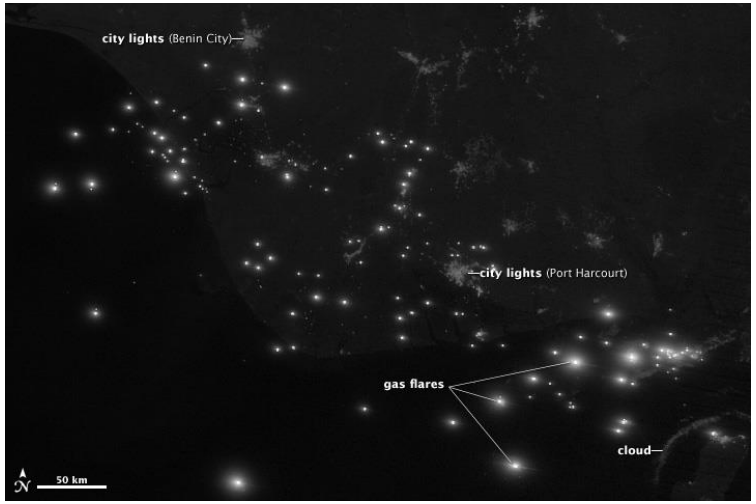
MRC
Centre for Environment & Health
UKRI Medical Research Council
Imperial College London

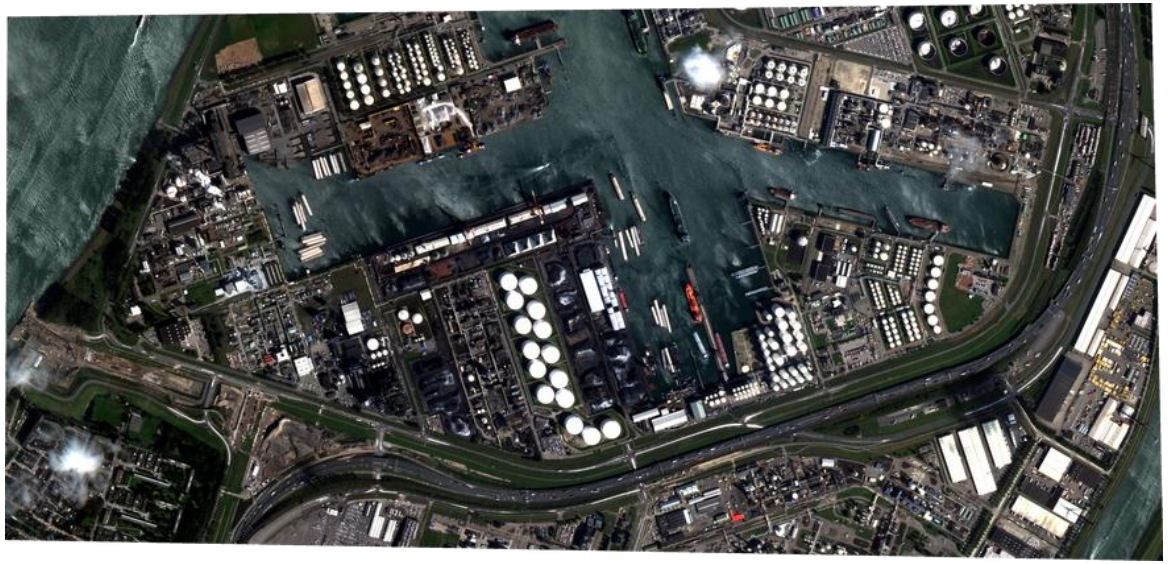
[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

Satellite imagery advantages

- Covers broad geographic areas.
- Relative spatial relationships less sensitive to field of view (e.g. distance between features will be consistent, unlike street view).
- Better availability across the electromagnetic spectrum.
- Some provide better temporal resolution (monthly/daily).



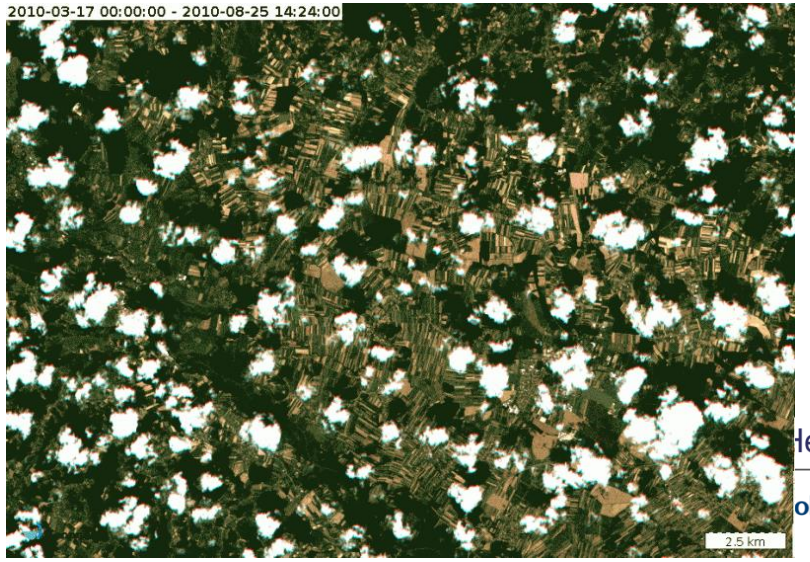




[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

Satellite imagery disadvantages

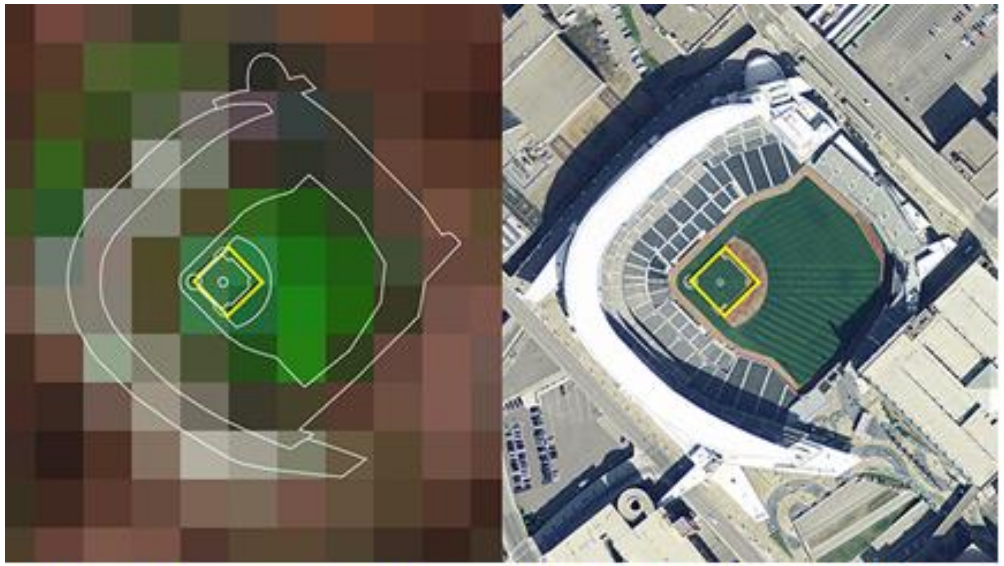
- Coarse spatial resolution (in comparison to SV imager)
- Top-down view may occlude features of interest (e.g. faces of buildings)
  - Sensitive to weather conditions (cloud coverage)
- Images often provided in large tiles that require extensive pre-processing.



[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

Comparison of Satellite providers

- Landsat by NASA/USGS (since 1972)
- Sentinel by ESA (since 2014)



Target Field, Minneapolis, Minnesota, Home of 2014 MLB All-Star Game, Landsat image (left), aerial photograph (right). A single Landsat pixel is roughly the size of a baseball diamond. *Source:* https://landsat.gsfc.nasa.gov/article/picturing-pixel

Recent, higher resolution commercial satellites

- Maxar, up to 15cm resolution
- Planet, up to 50cm resolution

Source: https://www.facebook.com/MaxarTechnologies/photos/pcb.10158466616266289/10158466615906289/

onment & Health

Imperial College London

[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

General features of imagery data

- Most comprehensive data sets are available through private providers (cost and restriction).
- Both forms of data are unstructured and require large amounts of memory storage.
- Both are “snapshots” of a given moment.
  - Sensitive to particularities in weather, daylight, visibility (sometimes of interest, other times background).

→ Researchers are also beginning to collect primary imagery data as well.

[Slides from Ricky Nathvani & Esra Suel, Healthy Data Analytics Course ICL, 2021]

Goals in computer vision for understanding the environment

What are we aiming to do with a CNN?

Classify? Regression/prediction?
Identify a characteristic?
Find certain features?
Cluster images?
Generate similar images?

→ Depends on research question!

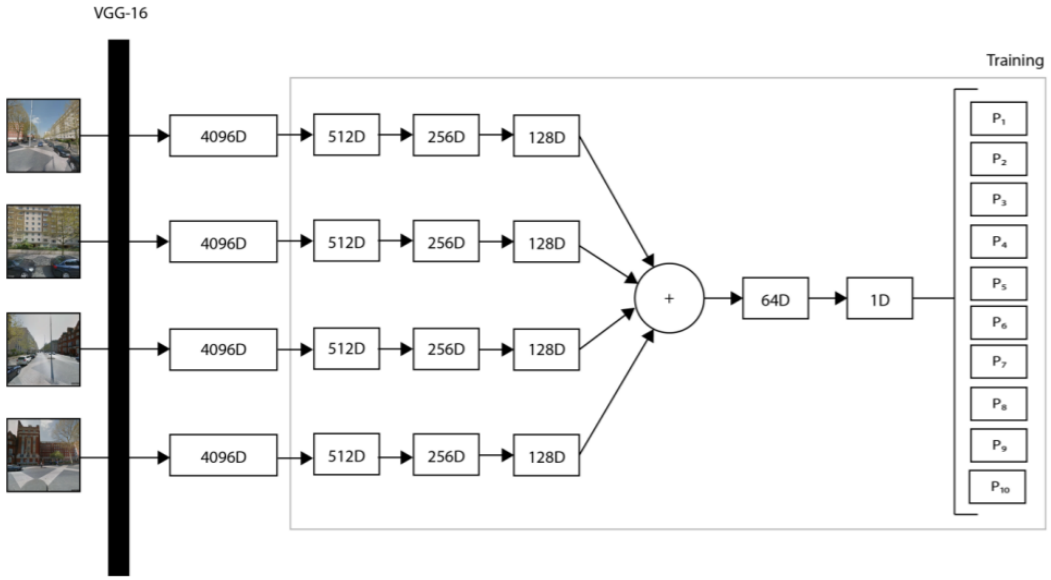
MRC
Centre for Environment & Health
UKRI Medical Research Council
Imperial College London

Do poverty and wealth look the same the world over? A comparative study of 12 cities from five high-income countries

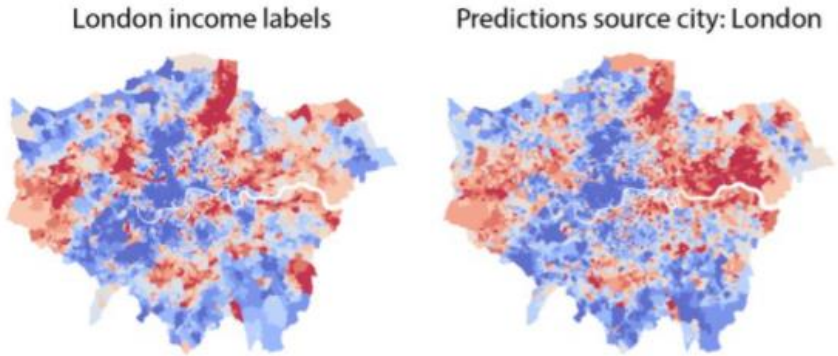
Esra Suel

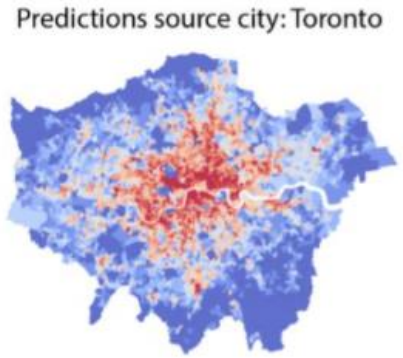
HDR UK Fellow

Imperial College London

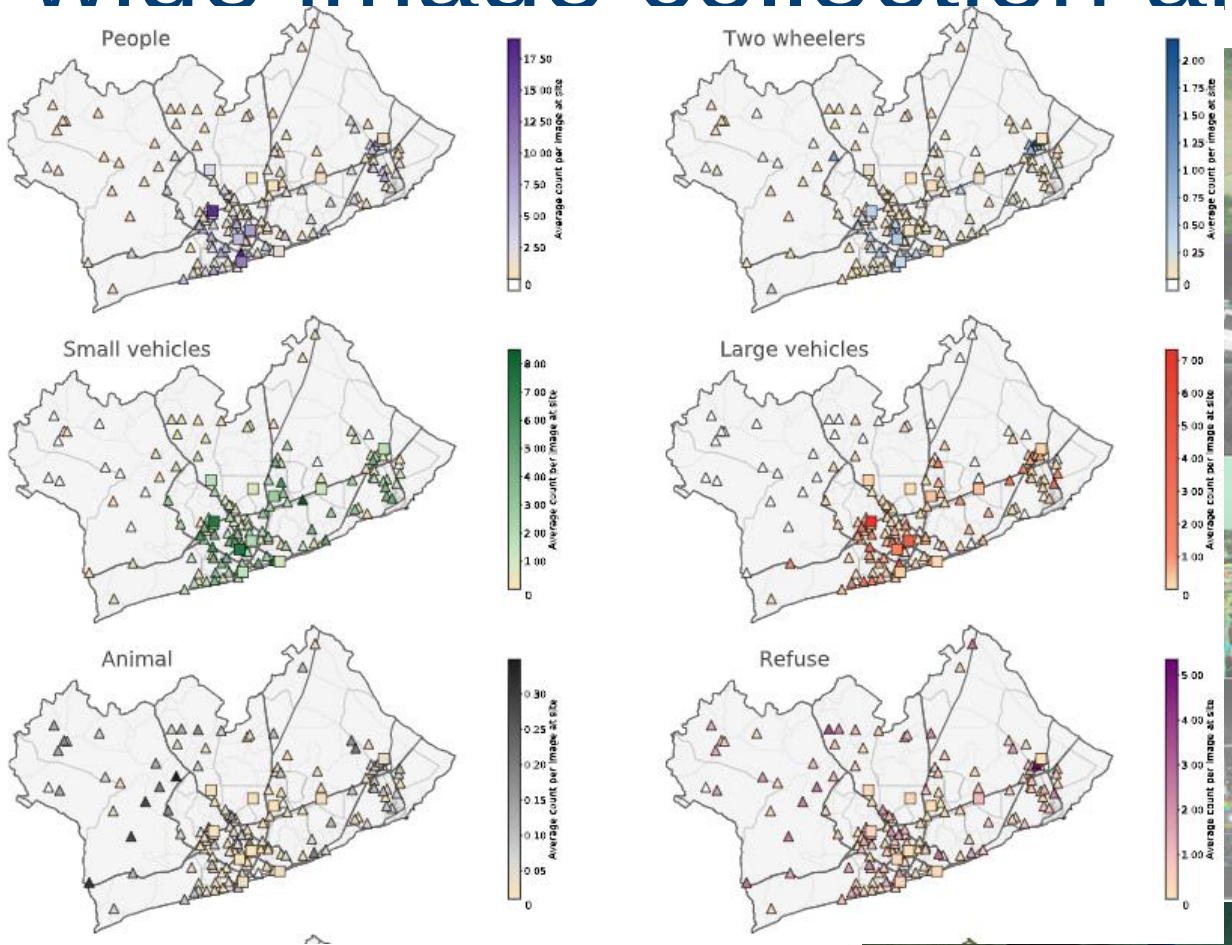


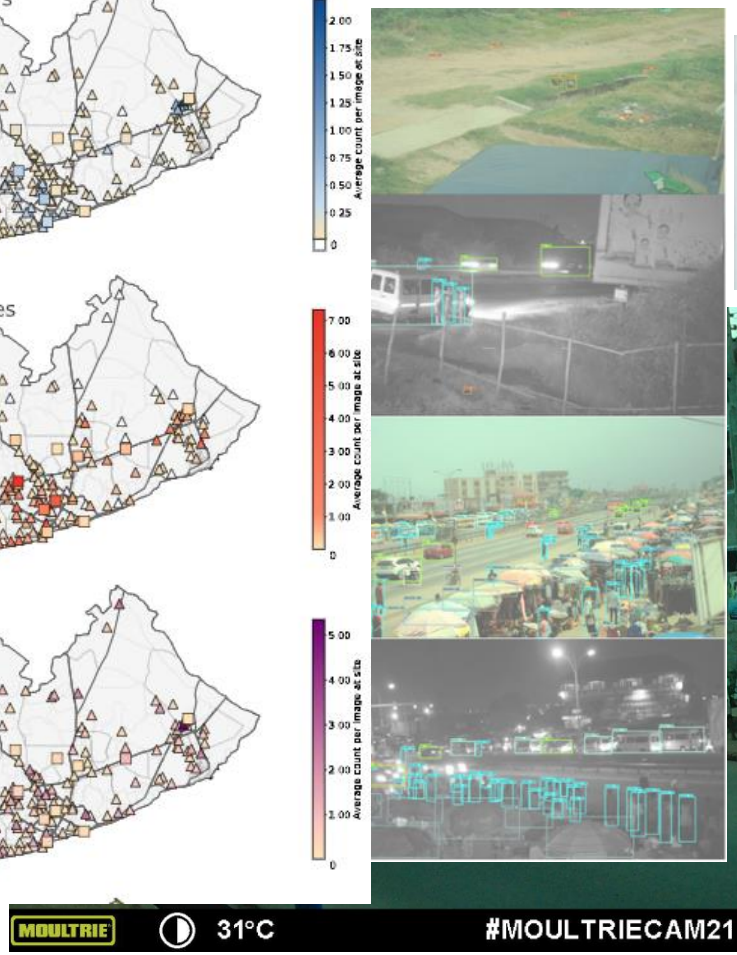
Transfer Learning using pre-trained VGG16 (ImageNet) with outcome: London income deciles from census statistics.

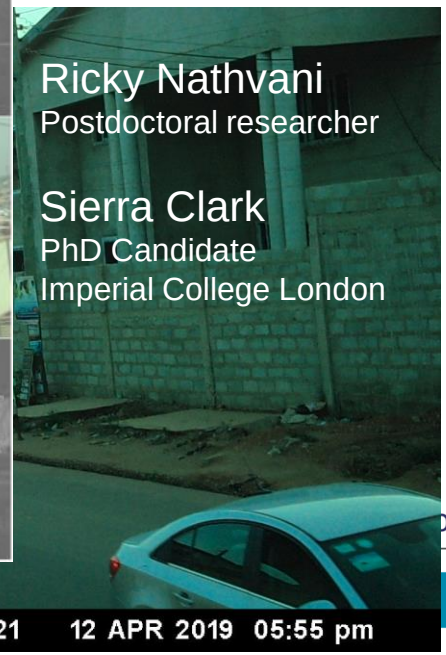




Spatiotemporal characterisation of the human activity and environment in Accra with city-wide image collection and deep learning







Ricky Nathvani
Postdoctoral researcher

Sierra Clark
PhD Candidate
Imperial College London

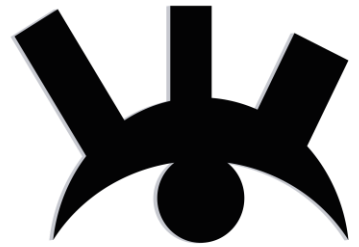
or Environment & Health

Medical Research Council

Imperial College London

More in this field: Sustainability (Stanford AI Lab)
Urban Form (Urban Grammar)
Security (Osint)
Climate (NASA, climatechangeai)
Food security

Thank you for listening



PULSE LONDON

Questions?

pulselondon.co.uk to participate in my research