Centre for
Data Analytics

Insight

A Neural Basis for the Implementation of Deep Learning and Artificial Intelligence

Prof. Alan F. Smeaton
Dublin City University

DCU NUI Galway OÉ Gaillimh UCC UCD DUBLIN sfi

Agenda for talk

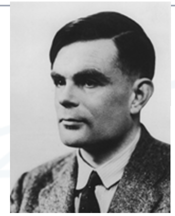
- Neurotech
- History of Computers
- The Human Brain
- Neural Networks and the AI winter
- Machine learning for computer vision
- Emergence of deep learning and CNNs
- Neural Network Components
- Are we finished with Neural Networks ?
- Implementing deep learning
- Future prospects

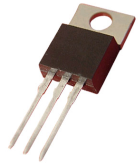
Industry Interest in “Neurotech”

- Beginning a new era of **neurotech**, brain-tech
  - **Facebook** *“working on a system that will let you type straight from your brain about 5 times faster than you can type on your phone today,”* Mark Zuckerberg in 2017
  - Elon Musk (Tesla, SpaceX) revealed **Neuralink**, building high-bandwidth BCI systems, implanted in the brain, won’t require brain surgery, using components injected into the bloodstream – neural dust
  - Bryan Johnson – Kernel – developing an implanted brain prosthetic to help people with failing memories, including dementias – now focus on a general implant to record signals from thousands of neurons at once

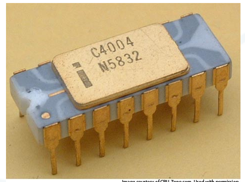
- How did it all start, where did it come from ?

How did we get here ?

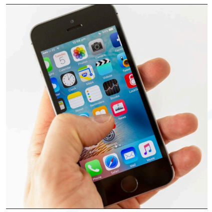




- Since WW II with ENIAC and Alan Turing
- Through to the development of the transistor
- And the 1950s with integrated circuits, magnetic memory hard disks and FORTRAN
- The 1960s with IBM mainframes and ARPANET



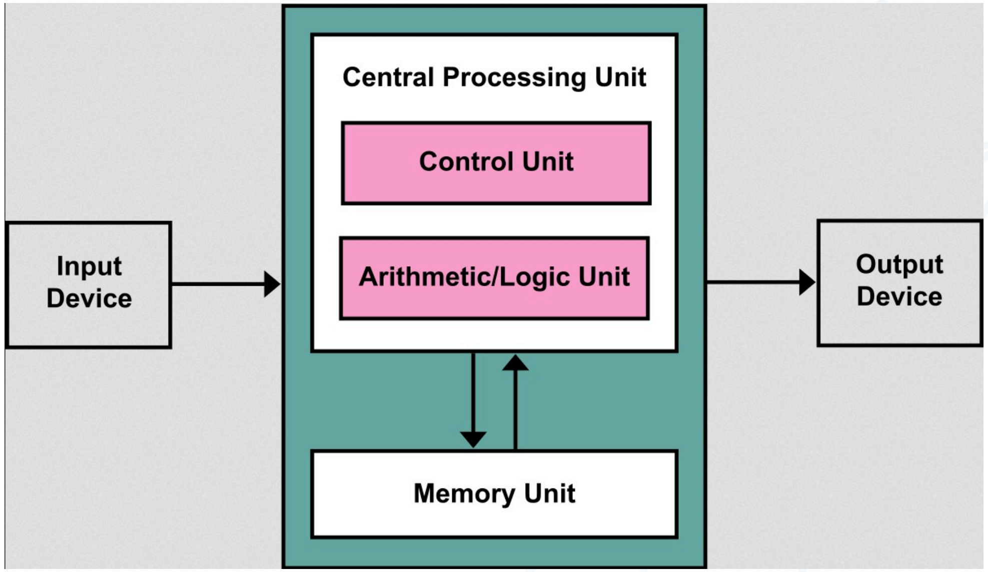
- The 1970s with ETHERNET, microprocessors and the first PC
- The 1980s with MICROSOFT, DOS and Apple MAC
- The 1990s with WWW, Google and internet companies
- The 2000s with YouTube, iPhones, wearables

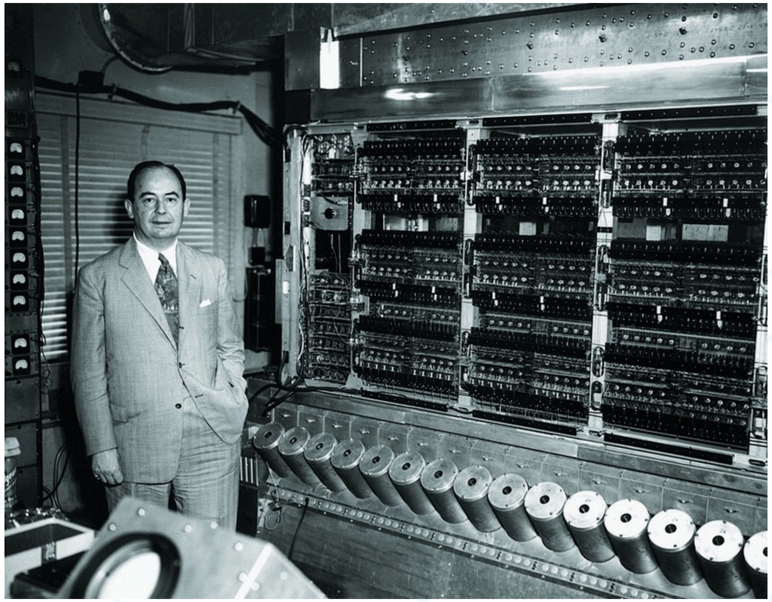


- Through to today with IoT and data analytics …

What's in common ?

- John von Neumann (1952) proposed a model for computers … that's still with us





Could computers do AI ?

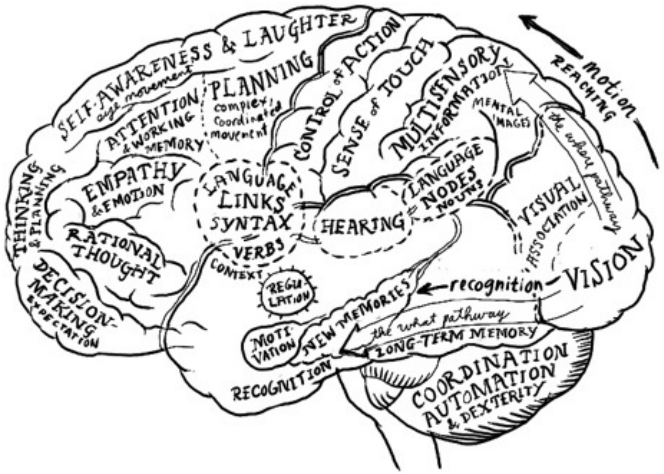
- Since the early days, always the question of could computers, with their von Neumann architecture, do artificial intelligence ?

The Human Brain

- 1.5kg, or 2% of our body weight, made of 86B neurons (grey matter) connected by trillions of connections (synapses)
- Responsible for executive functions like autonomic (heart, breathing, digestion, etc) and voluntary, in addition to executive functions like self-control, planning, reasoning, and abstract thought.
- This architecture of huge number of simple connected processors, is good for solving very complex problems, like vision, and learning

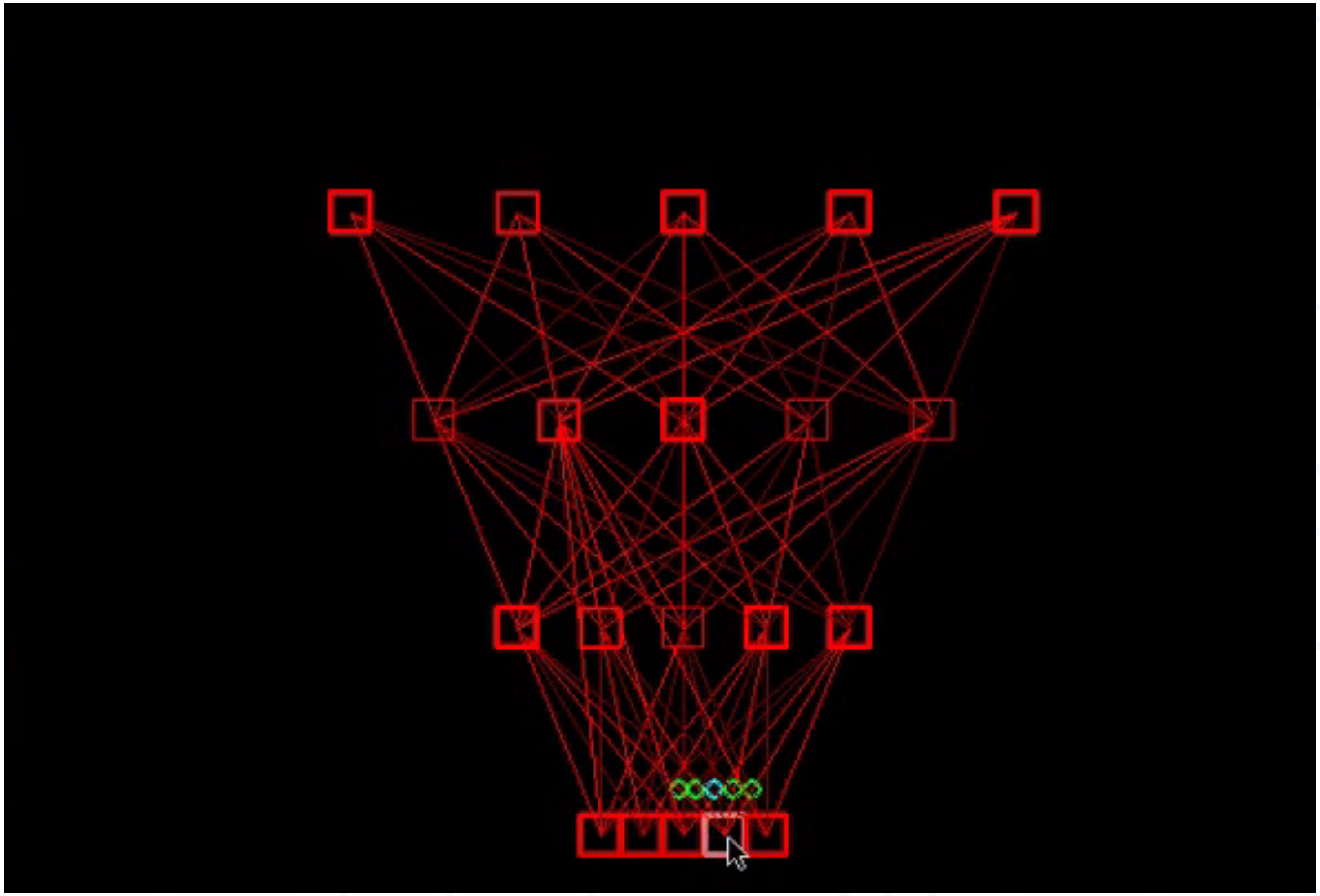
Human Memory

- The brain has been coarsely mapped



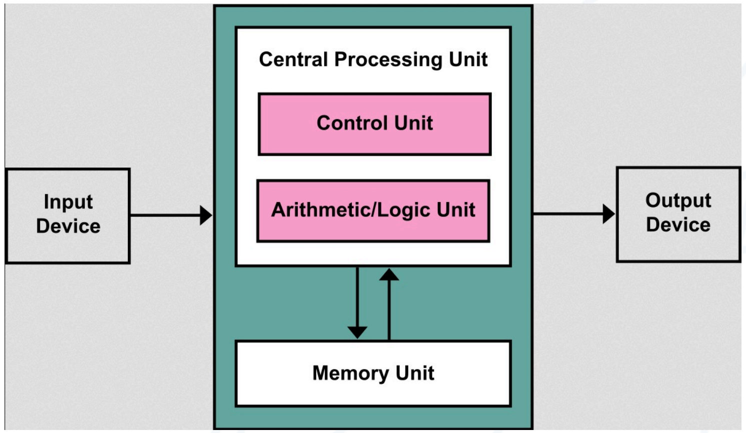
- The architecture is of simple but massively parallel processing, a form of perceptron, highly connected graph of nodes and links with signal-passing

How neurons work …



How do we do AI ?

- So how do we implement Artificial Intelligence, emulating the human brain, on the vN architecture ?



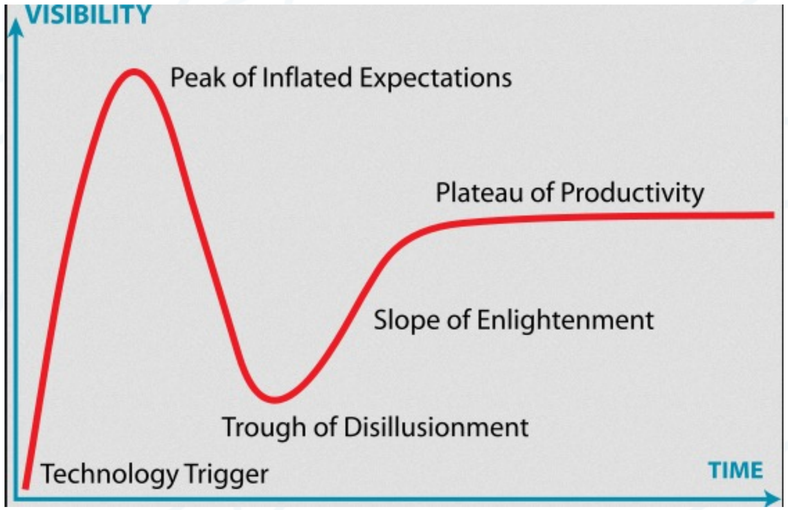
Raised expectations … AI “Winter”

1974, first AI winter

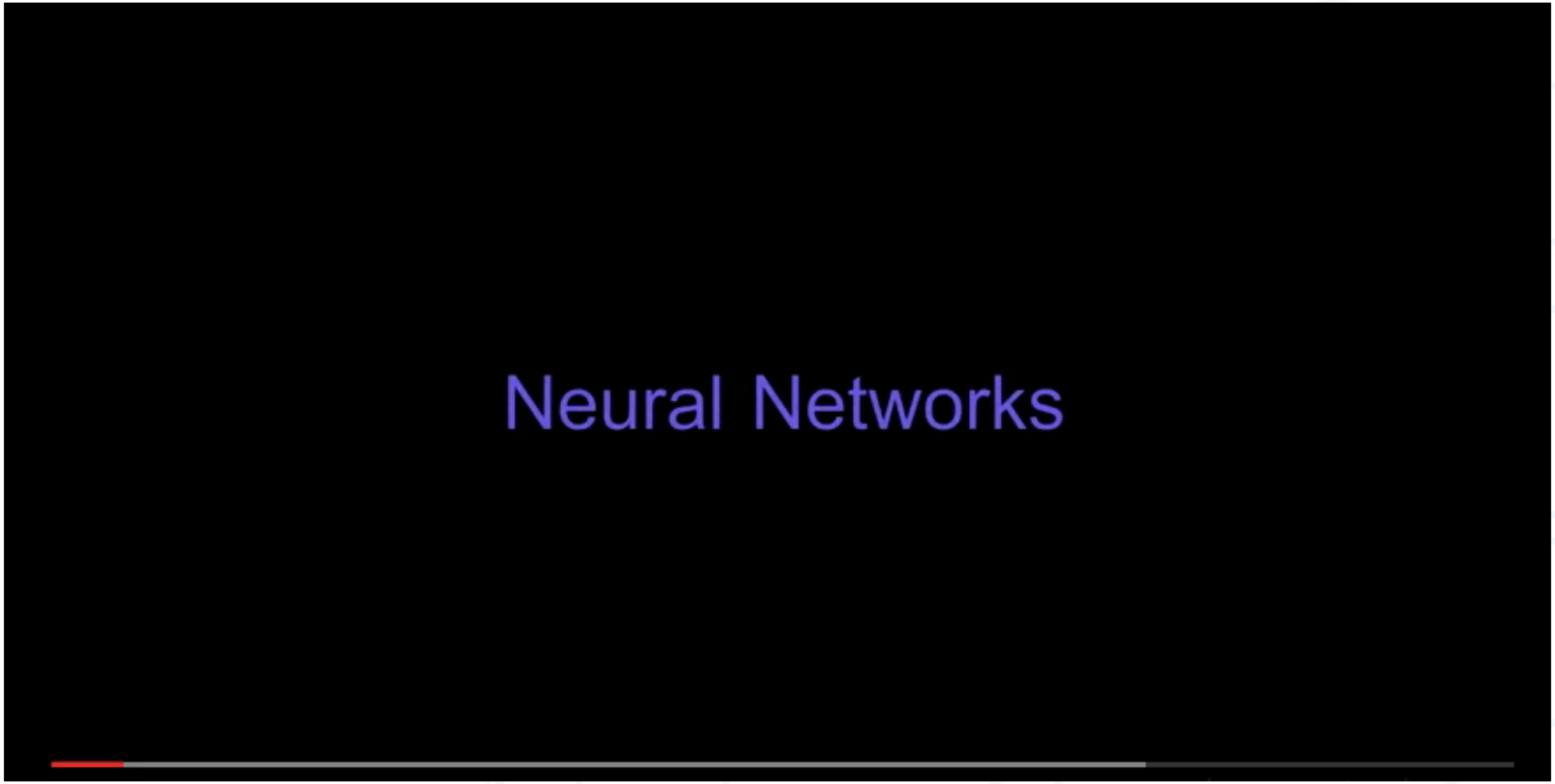
- Too ambitious / too big claims:
  - “The vodka is good, but the meat is rotten.”
    “The spirit is willing, but the flesh is weak.”
    (allusion to Mark 14:38)
  - 1966, negative report by an advisory committee, government funding of automatic translation cancelled.
- Limited knowledge of the outside world:
  - Restricted to micro-worlds (e.g. Blocks World)
  - Restricted to pattern-matching (e.g. ELIZA)
- Inherent limitations of computability:
  - Intractability, combinatorial explosion (to be discussed next week).
  - Undecidability

Neural Network research continued

- Research into neural networks – computational implementation of the brain's structure – continued but at a slow burn and an AI backwater !
- Meanwhile, hand-crafted, rules-driven AI research continued through 70s, 80s, 90s, even 00s
- AI applications were
  - Speech
  - Machine Translation
  - Computer Vision
  - Expert Systems
  - Etc.



Here's how Neural Networks developed …



But few people **really** cared about Neural Networks

Meanwhile …

Elsewhere AI and Machine Learning

- Machine learning evolved as an AI tool
- With mathematics and statistics input rather than any neural network connection
- Slow evolution of Machine Learning over decades
- Nourished by increasing availability of huge data volumes from … internet searching … social media …online transactions … etc.

- One application which pushed this was computer vision

What does this mean?

Machine Learning of Semantic Concepts

- Use Machine Learning to train a classifier to identify an object
  - Decision tree learning
  - Random forests
  - Genetic programming
  - **Support vector machines**
- Given some input data (e.g. SIFT features or colours or textures or lines or shapes or …)
- + Given a lot of + and - examples
- Let the computer figure out how to classify new examples into + or – clusters … that's the modus for machine learning

Concept Detection

	FEATURE DETETCION
0.2	Indoor
0.8	Outdoor
0.7	CityScape
0.3	Landscape
0.1	People
0.0	Face
0.8	Sky
0.2	Vegetation
0.7	Building

So how well did it work ?

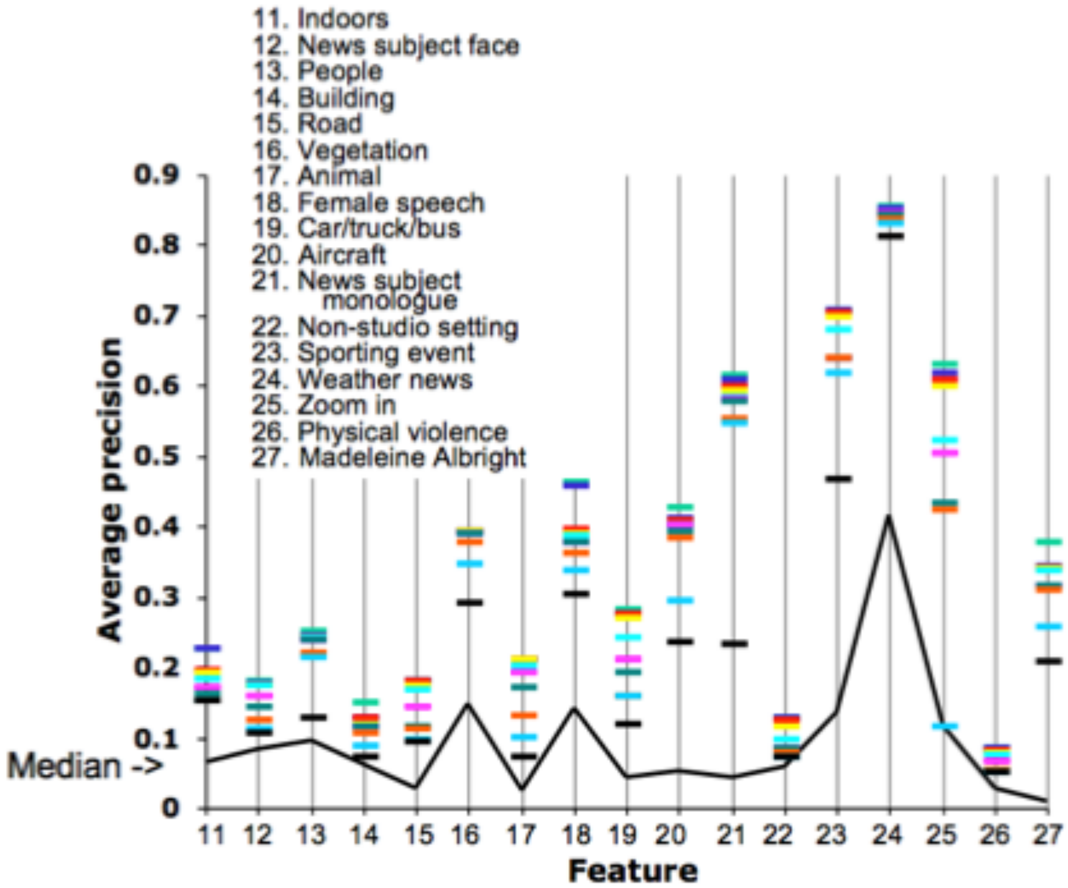
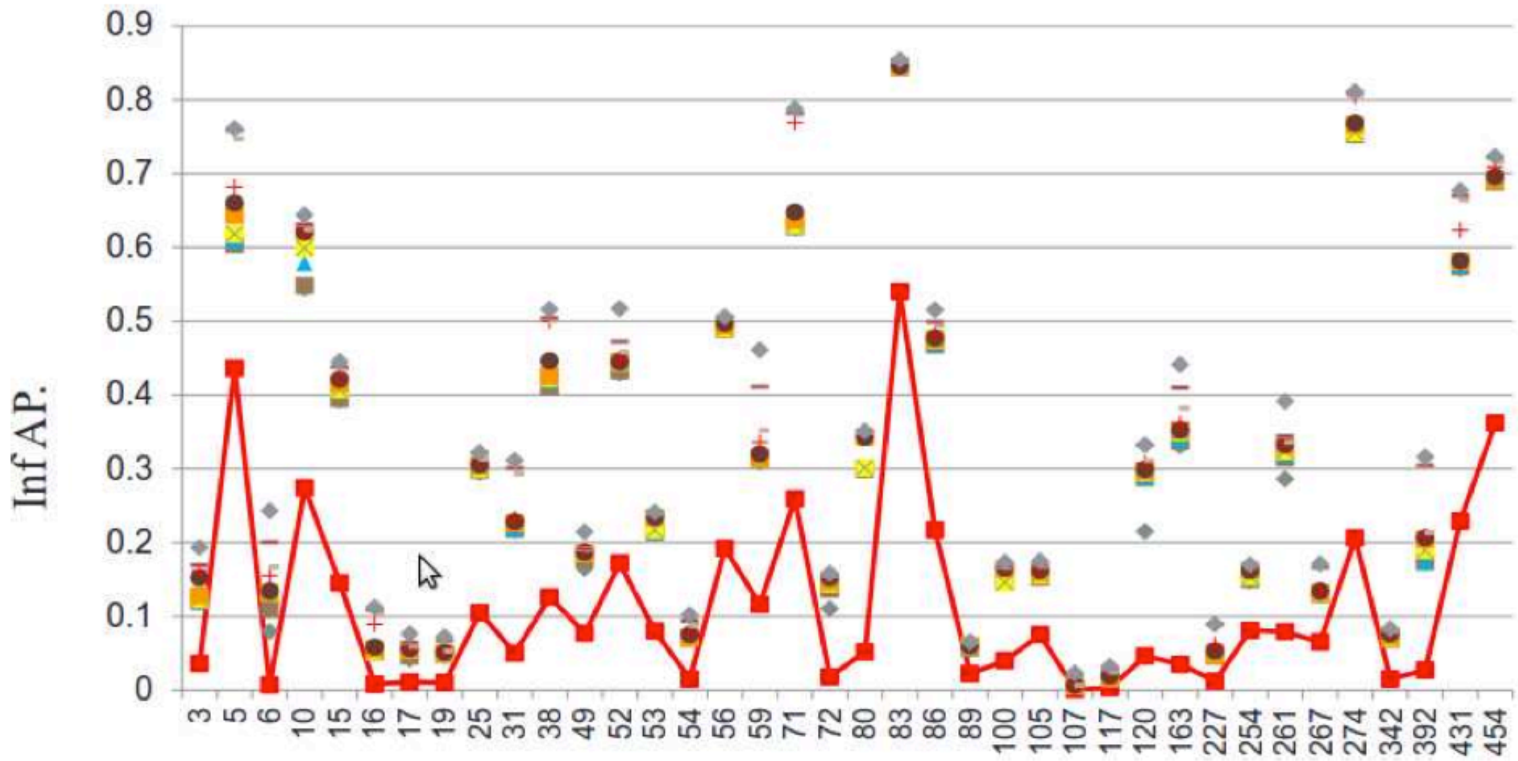


Fig. 2. Performance of Top-10 high-level Feature Detections per Feature in TRECVid 2003

And in 2013 ?

Figure 4: Top 10 runs (xinfAP) by concept



3 Airplane*
9 Basketball
10 Beach*
13 Bicycling
15 Boat_Ship*
17 Bridges*
19 Bus*
25 Chair*
27 Cheering
29 Classroom
31 Computers*
41 Demonstration_Or_Protest
59 Hand*
63 Highway
71 Instrumental_Musician*
80 Motorcycle*
83 News_Studio*
84 Nighttime
100 Running*
105 Singing*
112 Stadium
117 Telephones*
163 Baby*
261 Flags*
267 Forest*
274 George_Bush*
321 Lakes
359 Oceans
392 Quadruped*
434 Skier

Standard Machine Learning …

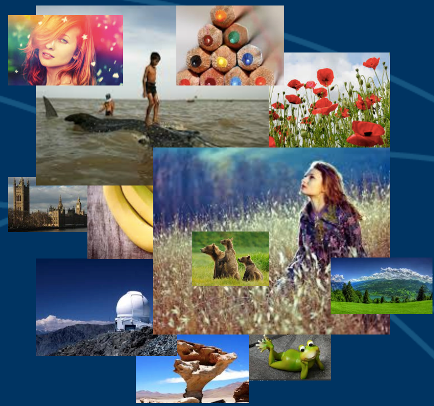
- So we're bumbling along, slow and steady progress in using ML for computer vision …
- ML is being used elsewhere also … from parole recommendations in US courts … to recommending books from Amazon … to … anything you can think of !

How does “standard” ML work ?

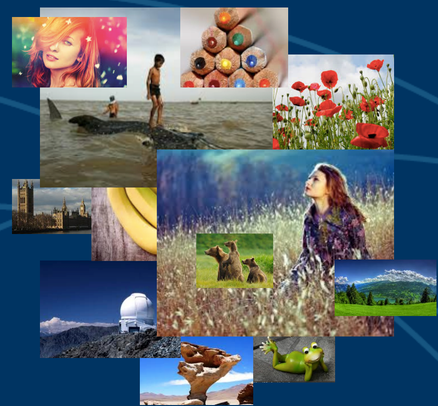
- Lots of + and – examples as training data
- Plot each data point onto an N-dimensional space
- Learn a boundary function which differentiates
- Train and test, refine, then deploy

How Does Machine Learning Work ?

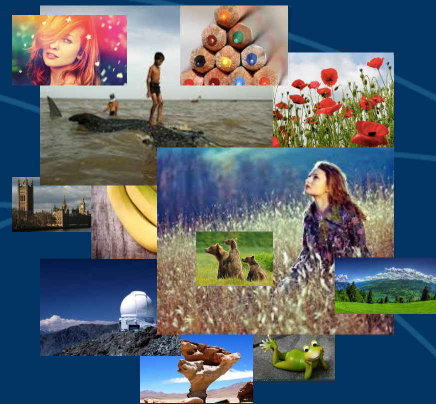
Suppose you want to build a classifier for 'boat', you need training data, + and – examples of boat images



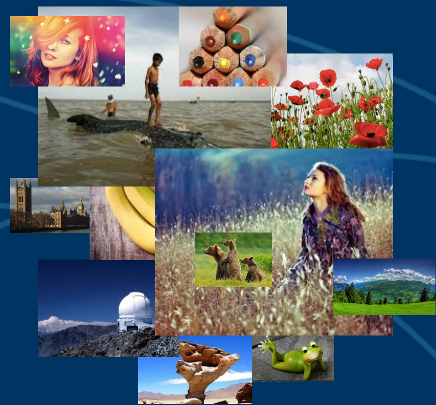
What makes a boat a boat, and a “not boat”, not a boat ?
We extract low level features from each boat/non-boat
and try to “learn” the differences



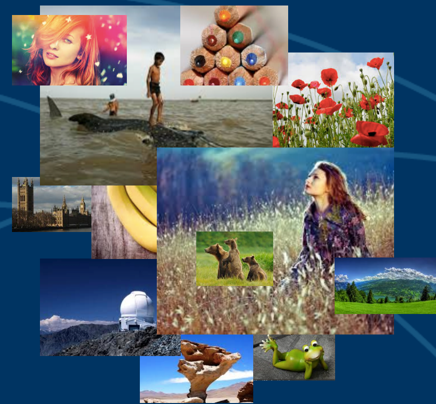
What kind of features … colours, textures, shapes, lines, across all the picture or in regions, calculated at pixel level

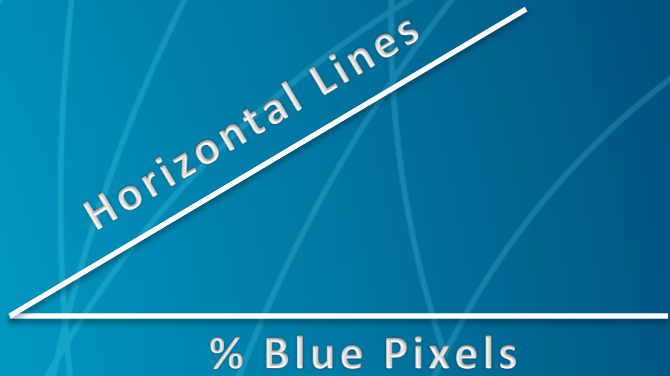


In practice, there are hundreds of such features, but lets look at just two

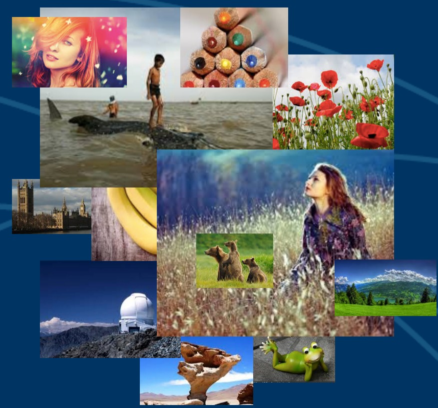


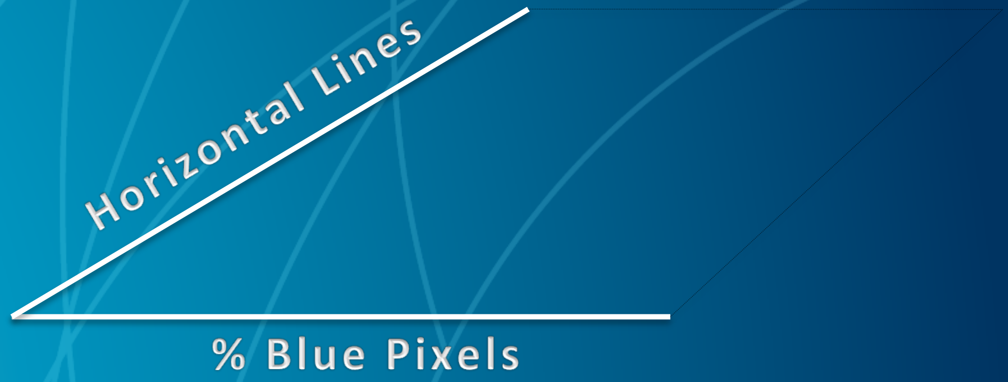
In practice, there are hundreds of such features, but lets look at just two



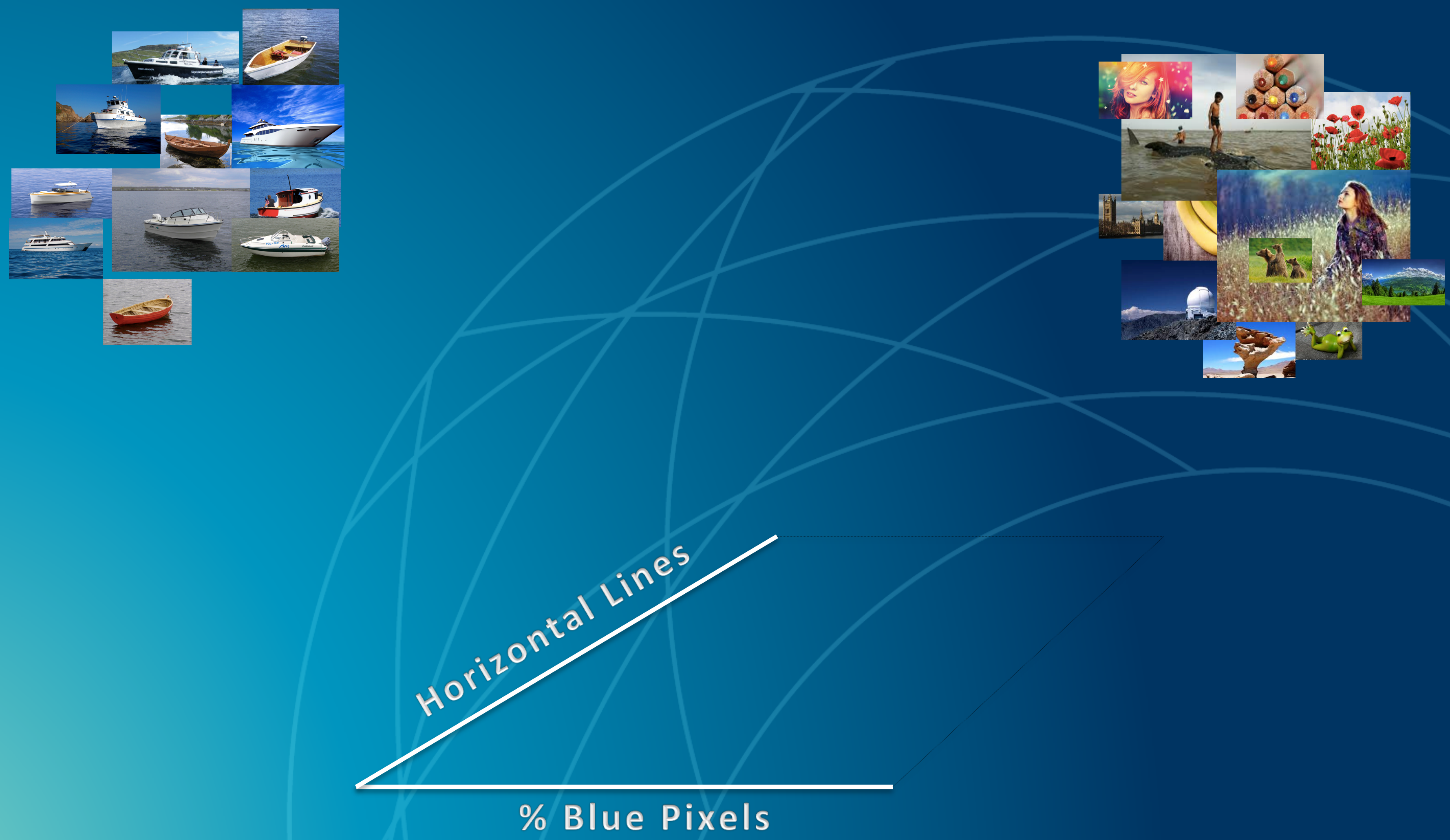


In practice, there are hundreds of such features, but lets look at just two

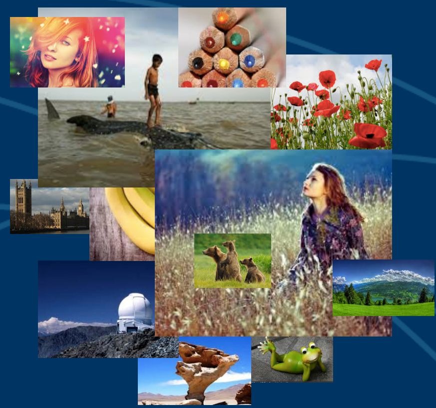


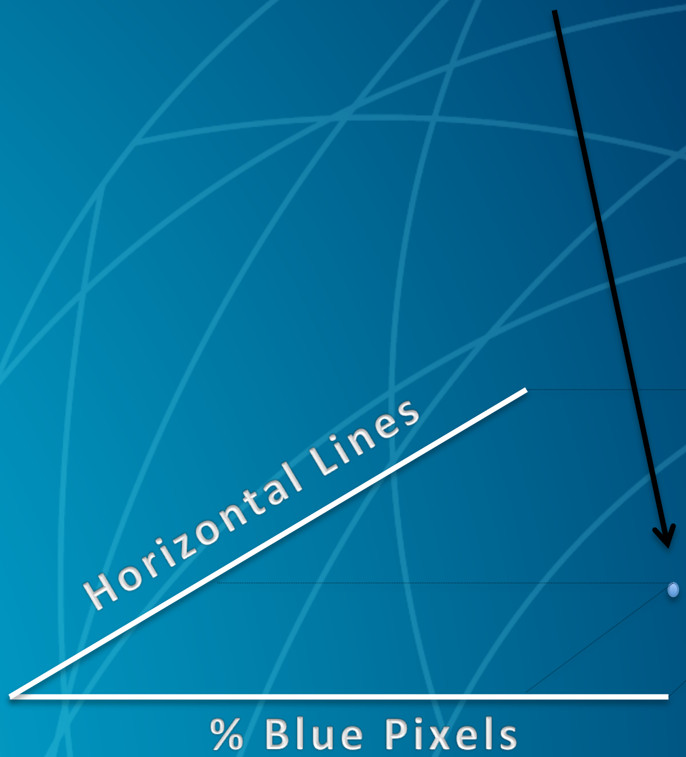


We can then take each image and “plot” it in this 2D “space”

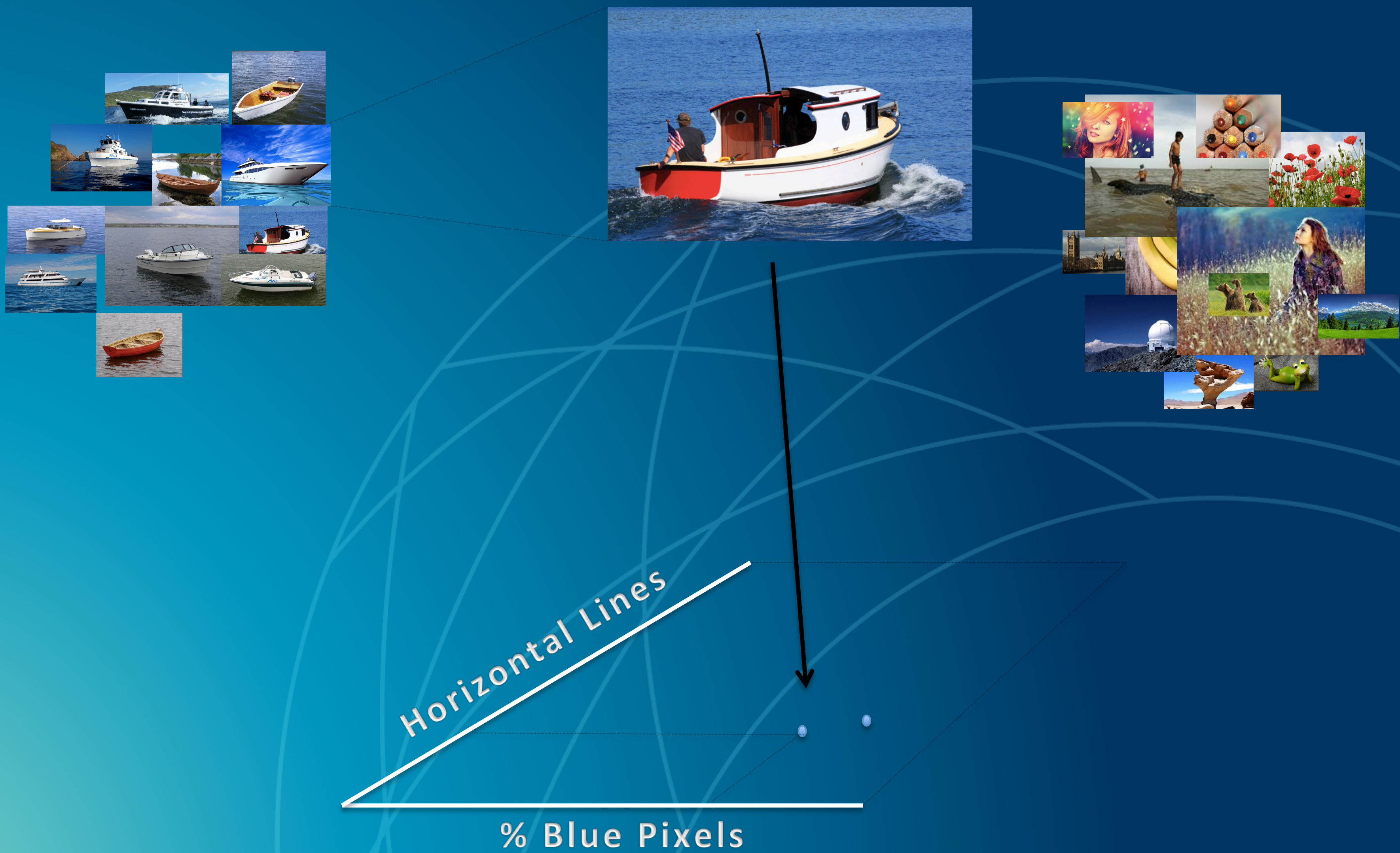


For boats …

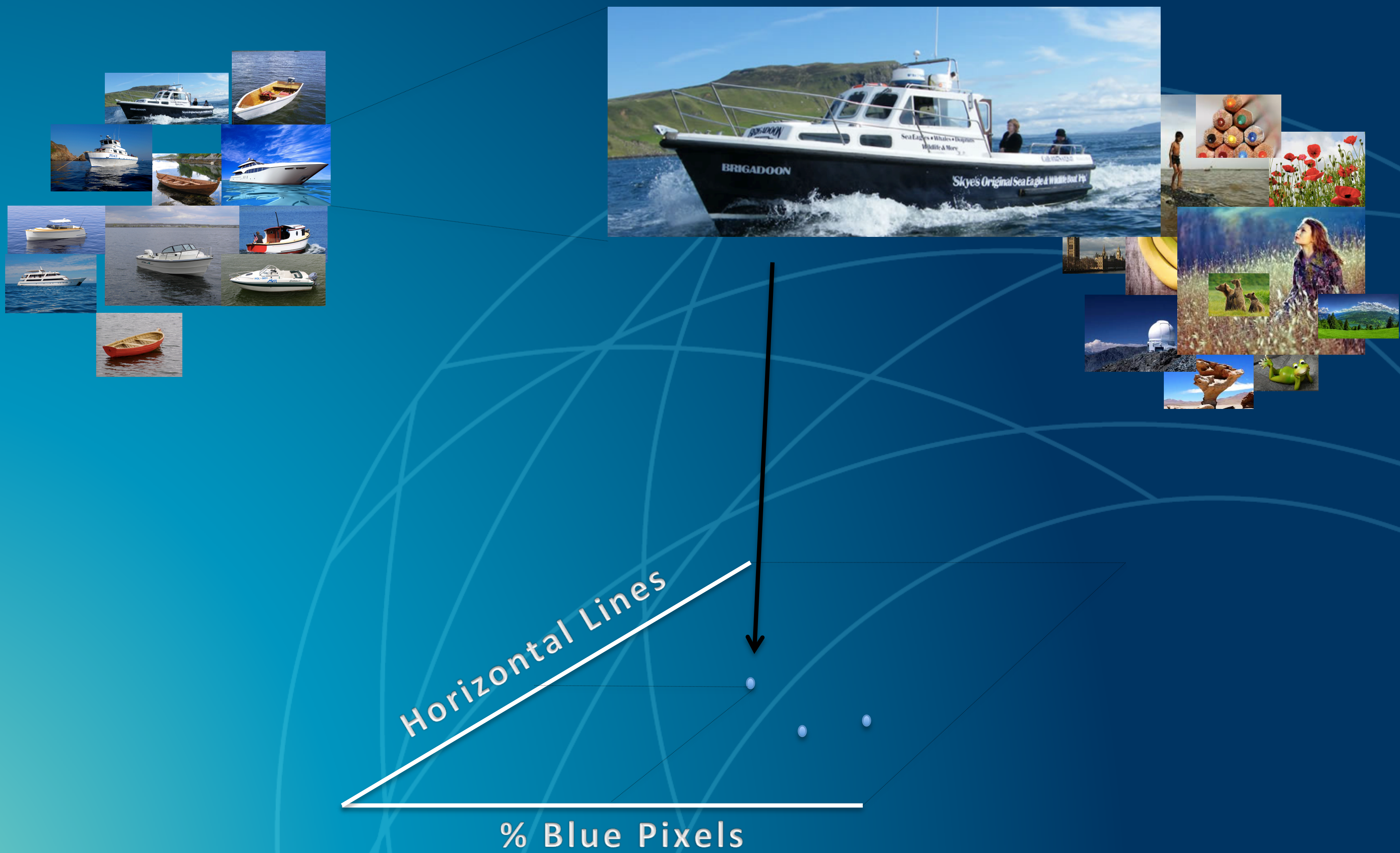




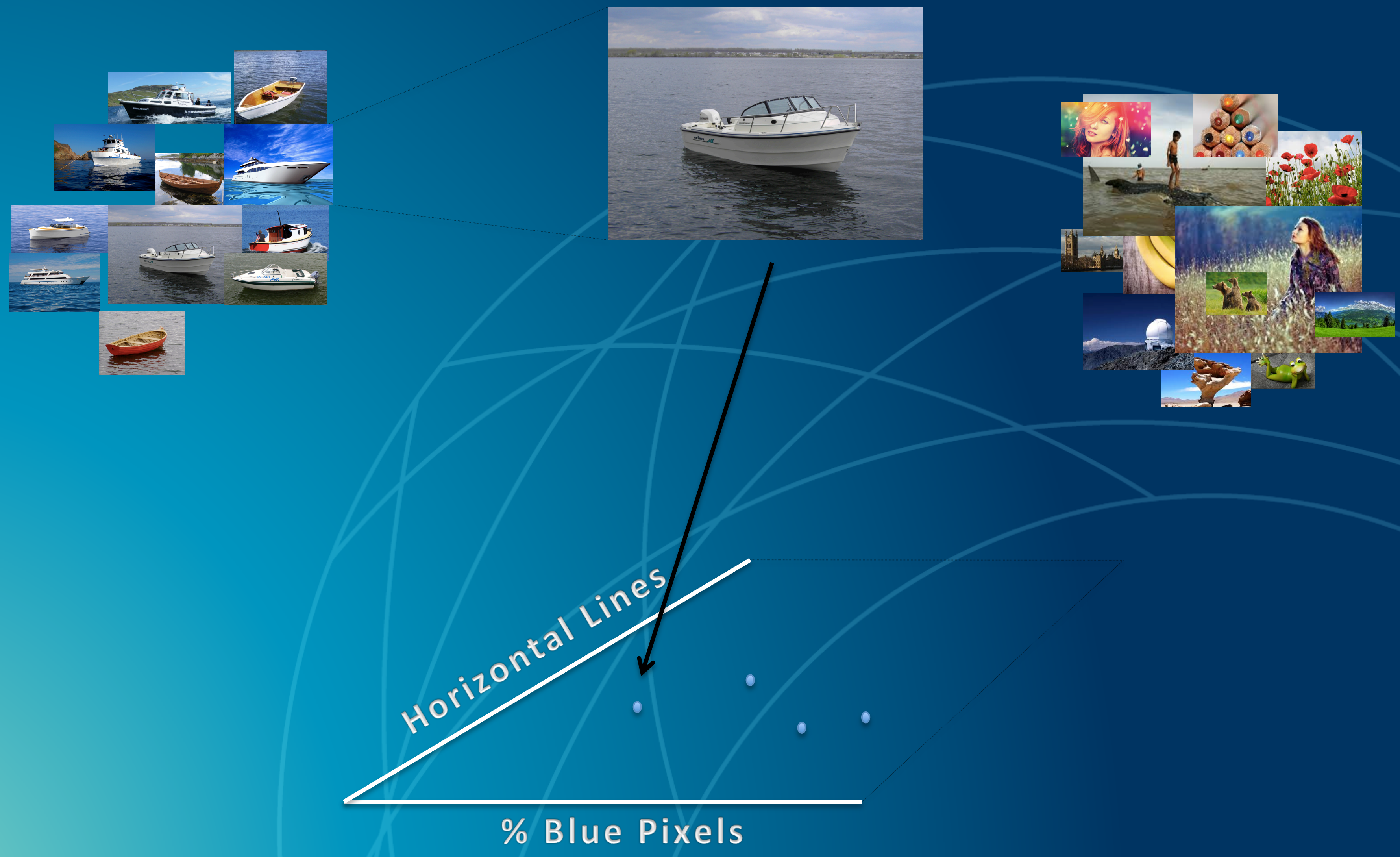
For boats …



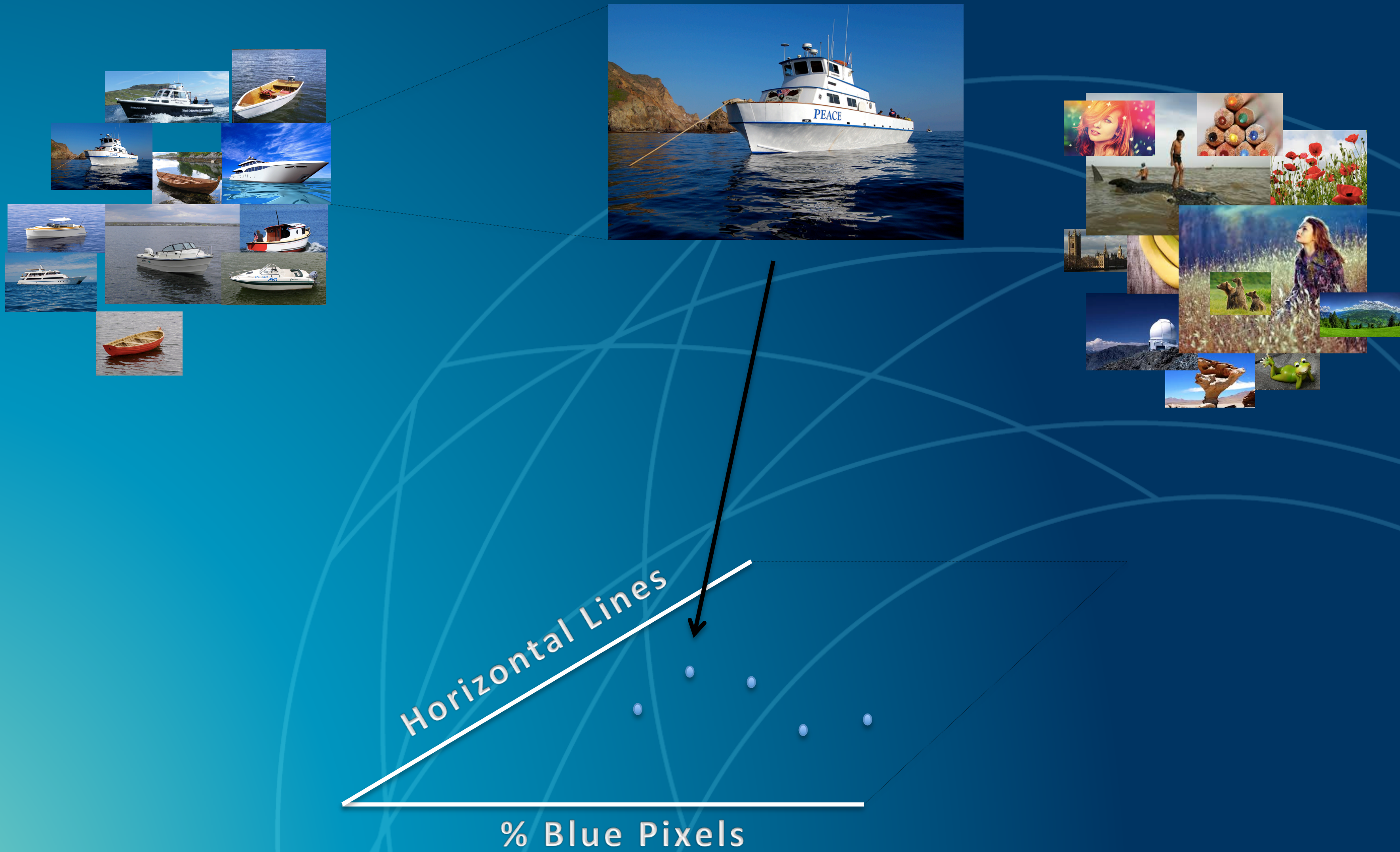
For boats …



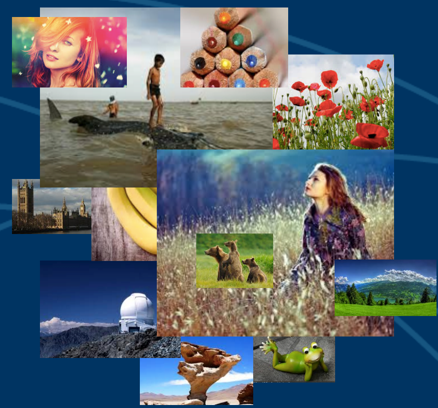
For boats …

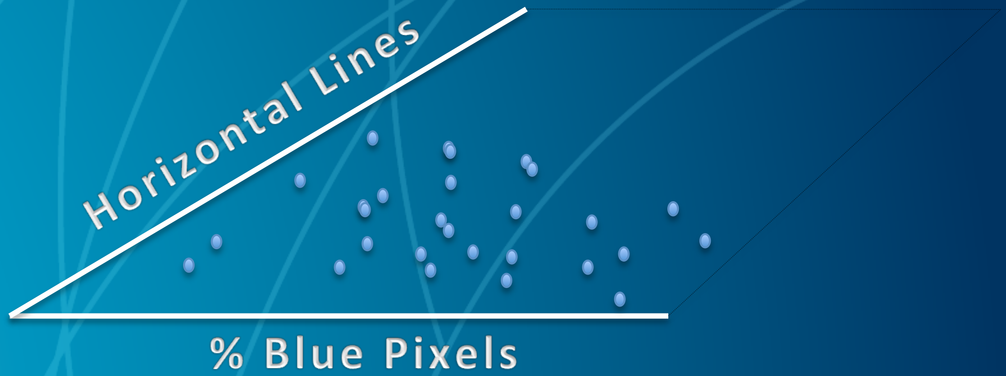


For boats …

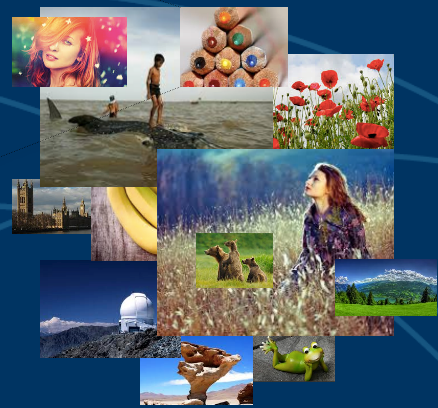


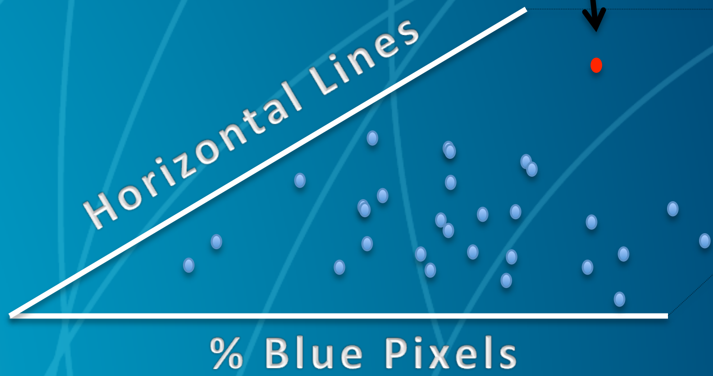
Until there are many of them



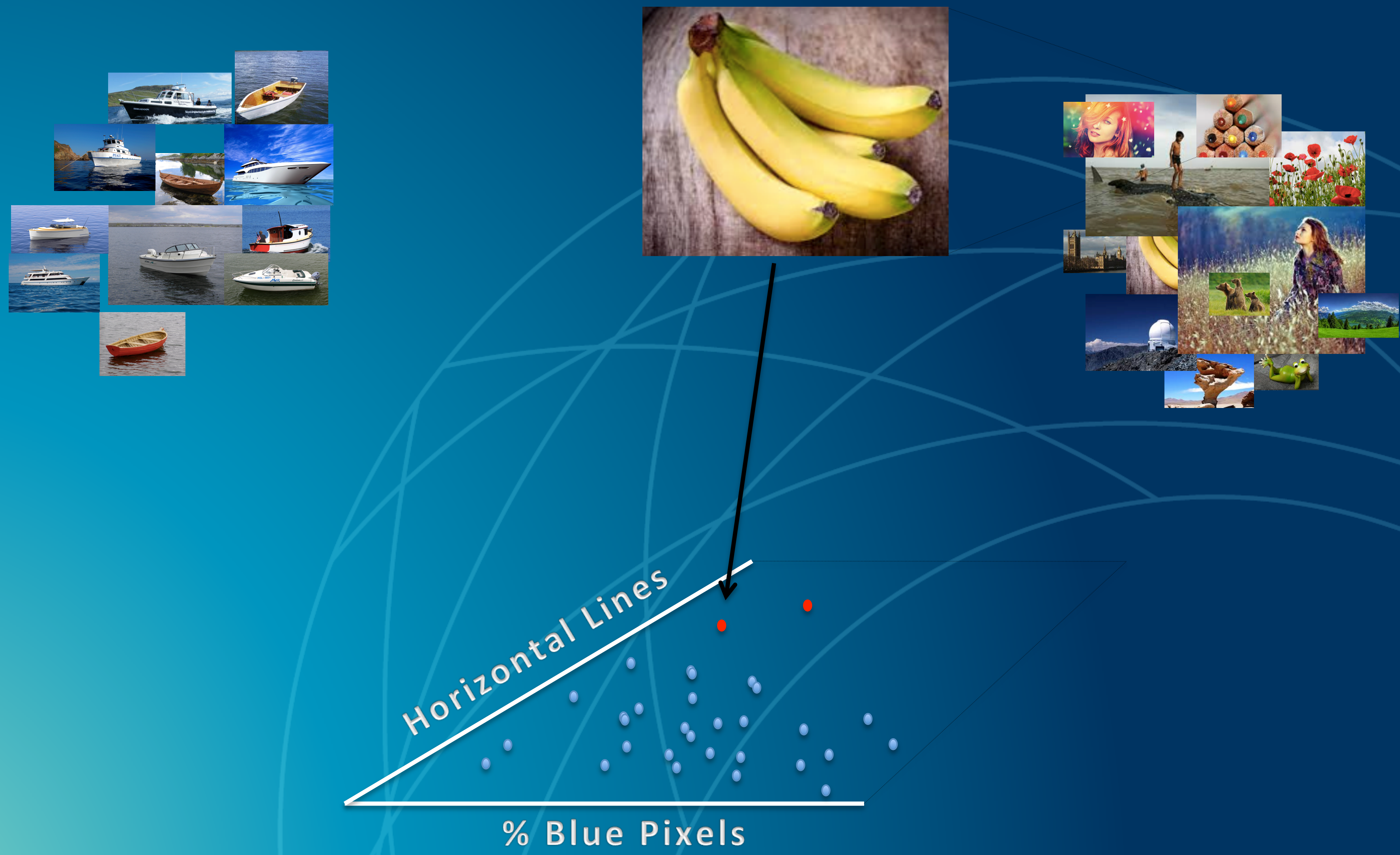


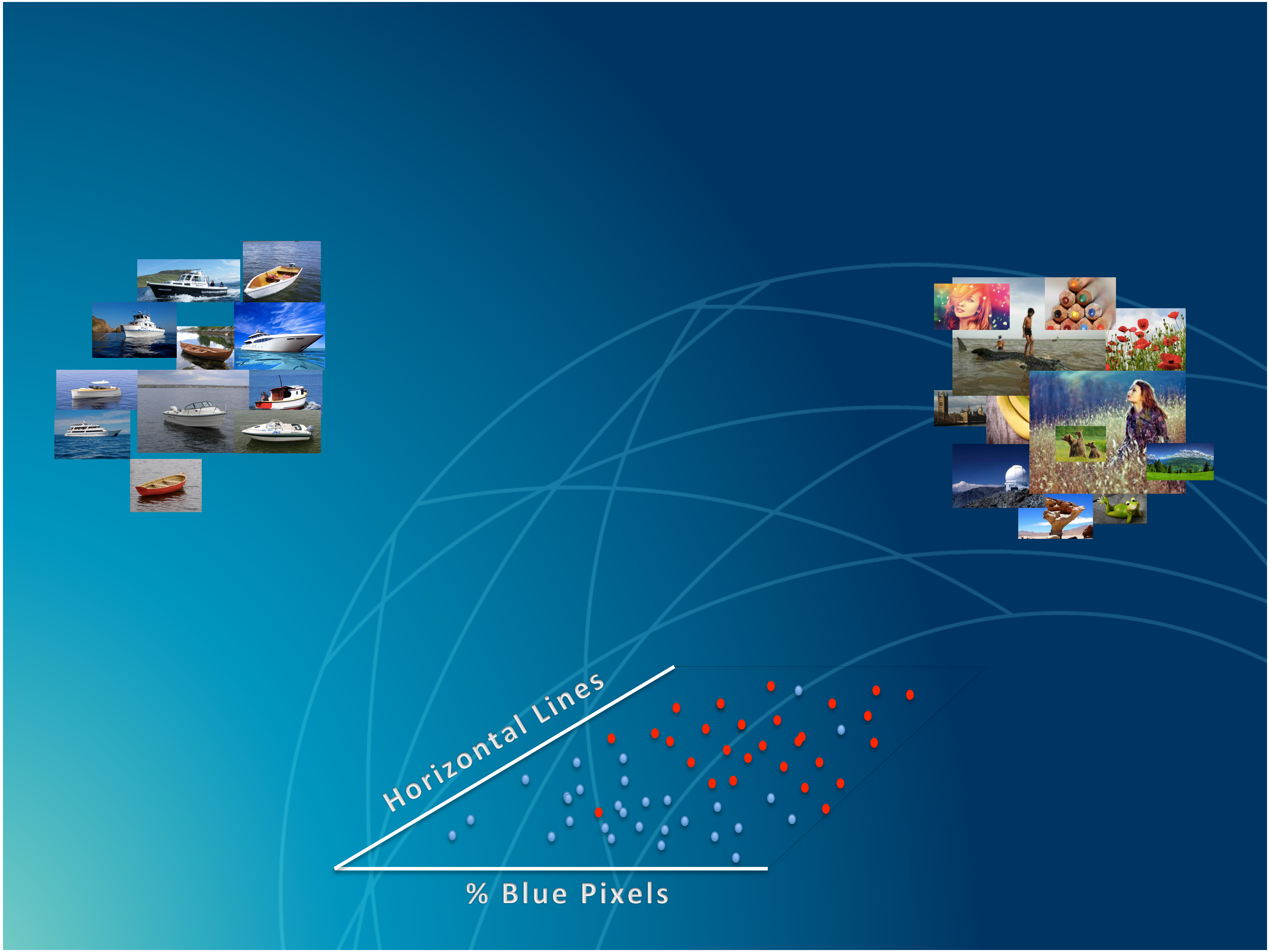
And then for non-boats …



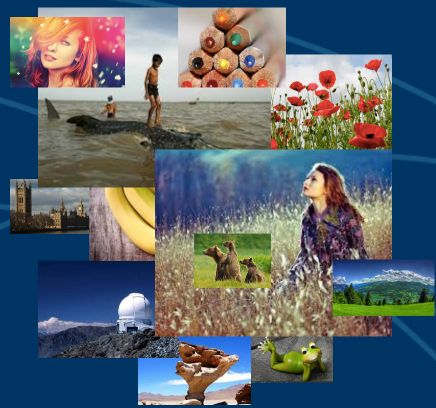


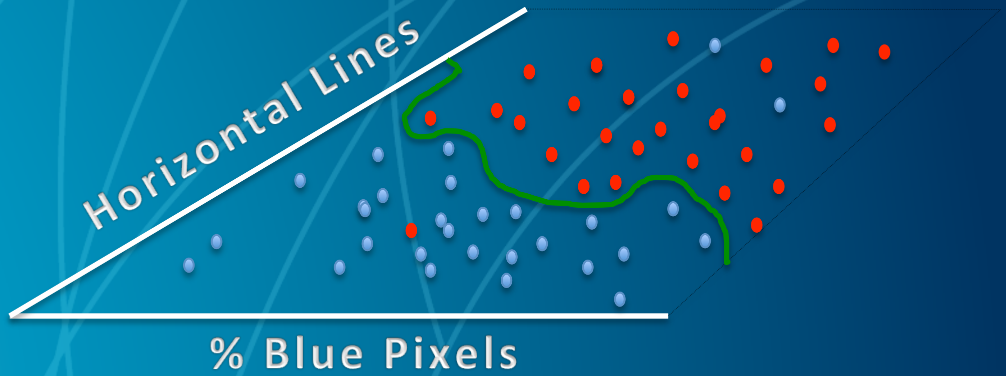
And then for non-boats …



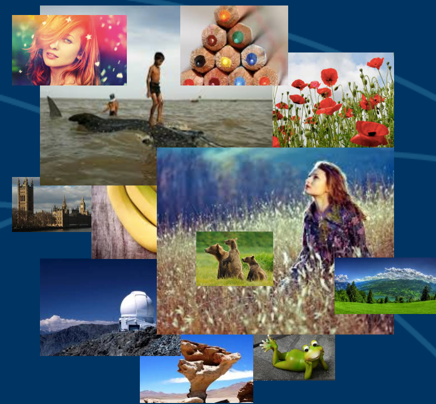

Horizontal Lines
% Blue Pixels

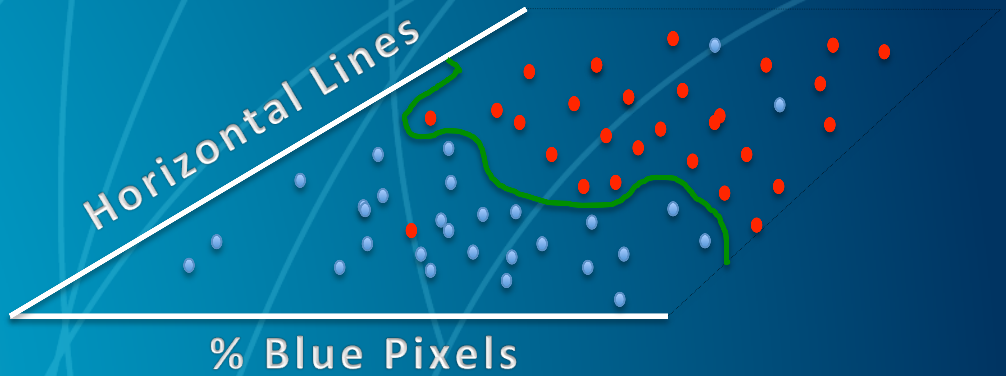
We then “learn” the differences between a boat and a non-boat, in terms of %Blue pixels/Horizontal Lines



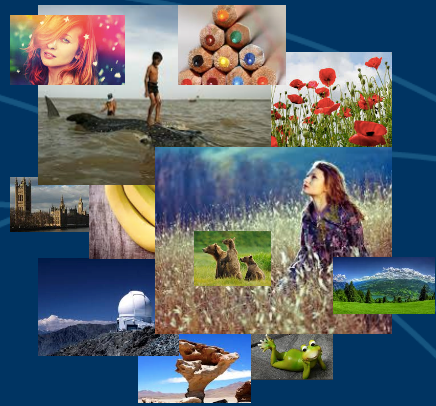


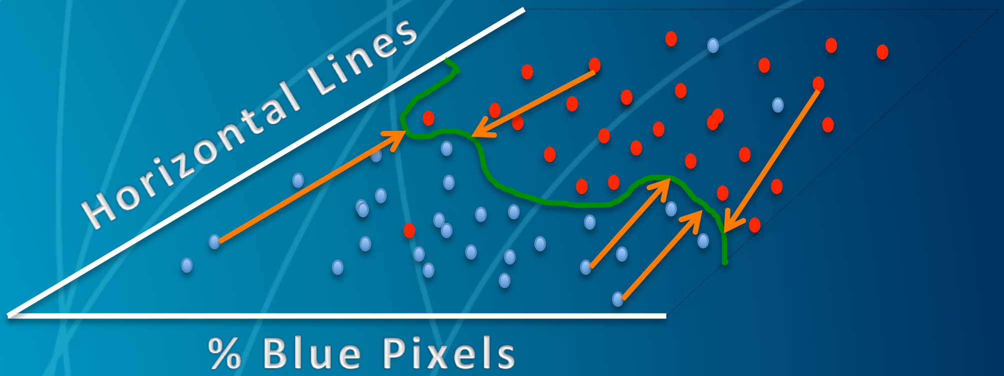
There are outliers, but mostly its correct



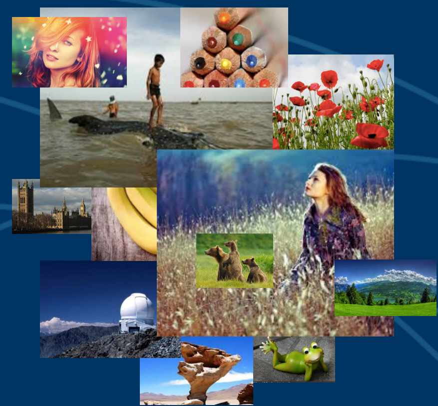


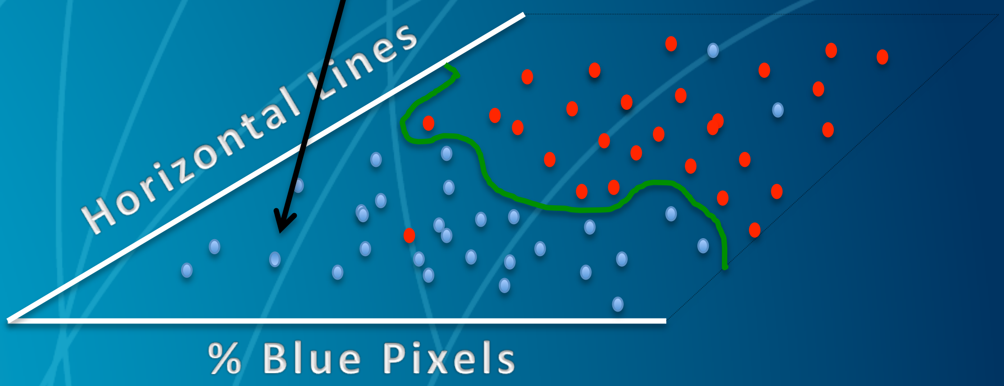
The “distance” from this “hyperplane” is a measure of confidence in boat/non-boat



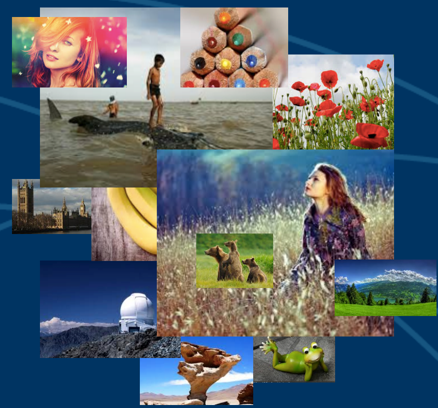


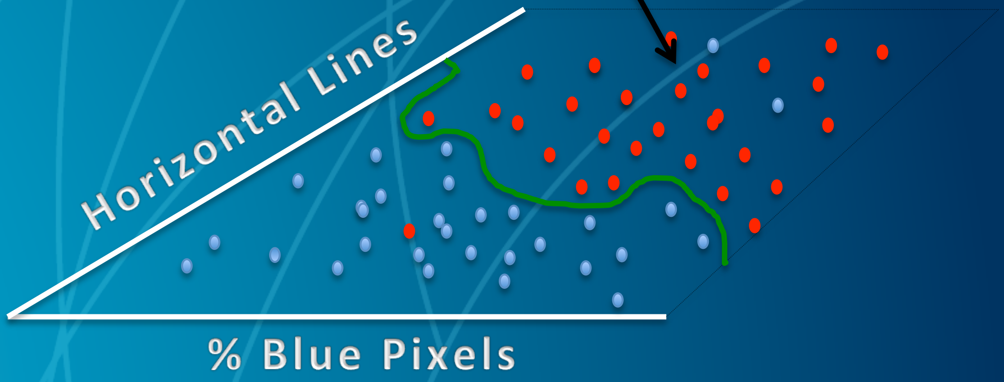
Then take new (untrained) images …



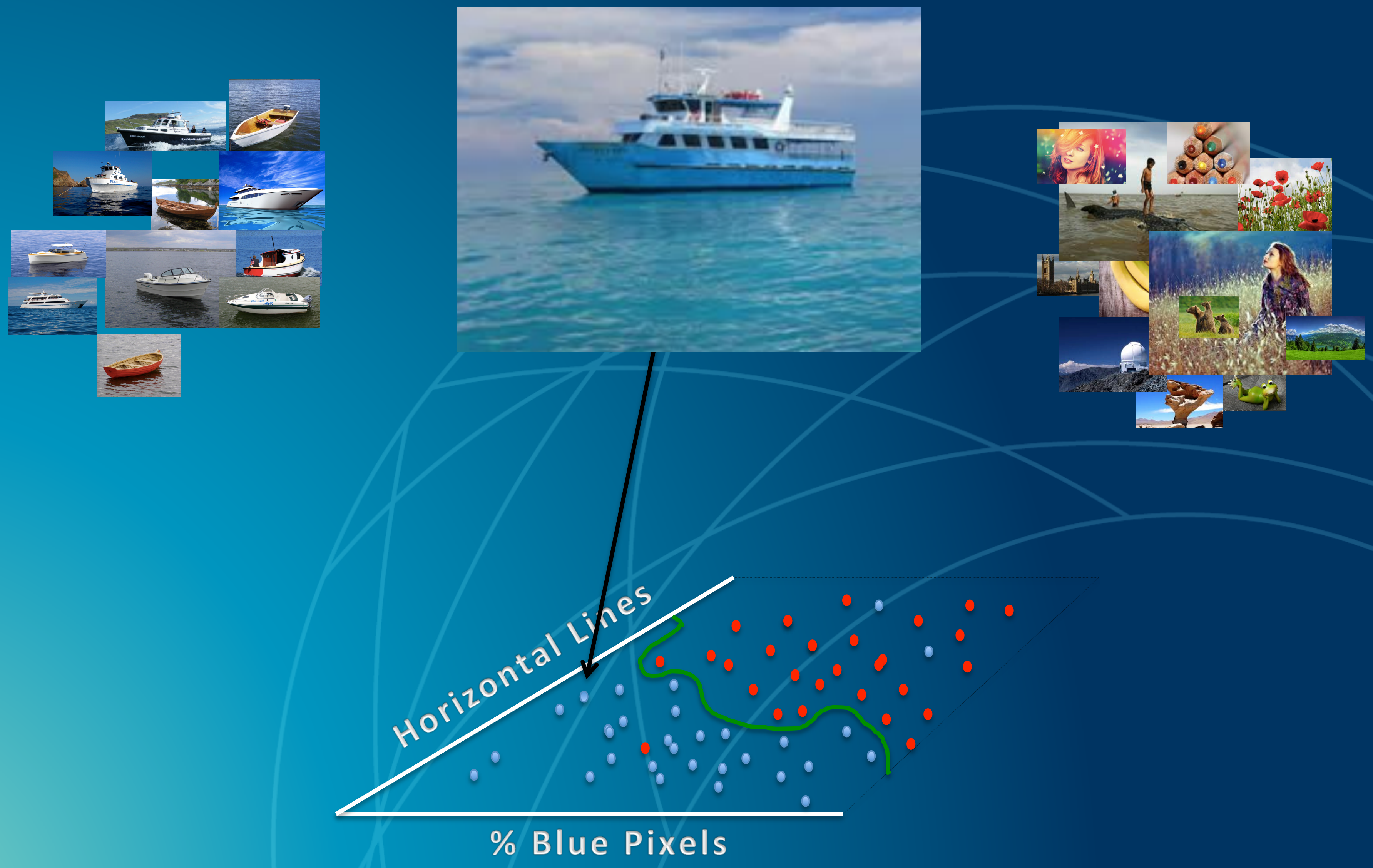


Then take new (untrained) images …

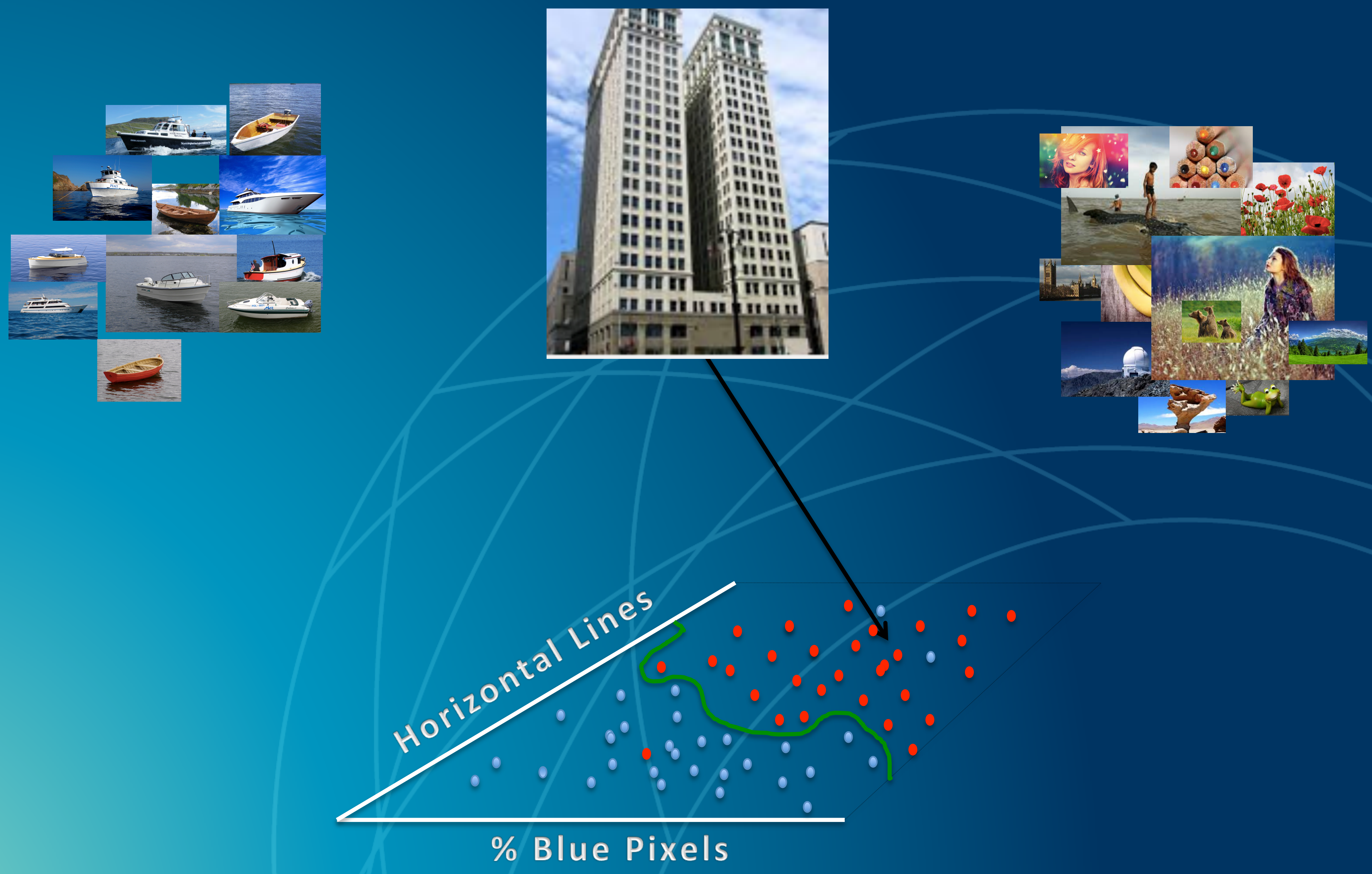




Then take new (untrained) images …



Then take new (untrained) images …



So that’s machine learning … building classifiers

- Training set (positive and negative examples)
- Balanced numbers of each
- Features for each
- Lots of computing needed to extract features and learn the classifier

Very fast to run new examples through the classifier

Which learning functions, which kernal, which features … all that is a black art !

One of the pluses of “standard ML” is that it can always conjure up some explanation for a result


Horizontal Lines
% Blue Pixels

- Another plus is that we can rate the relative importance of each of the axes .. i.e. features

- A downside is we have to do feature engineering to define the axes, and that's a black art

- Lots of applications across domains … great … everybody happy

And in 2012, this happened !

- Krizhevsky, Sutskever and Hinton at U Toronto, "won" the ImageNet large scale visual recognition challenge with a *"convolutional neural network"*

- 6 months later, they all work at Google

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

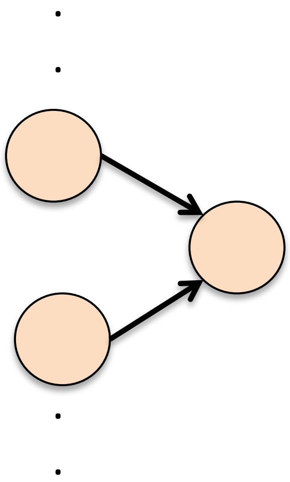
Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

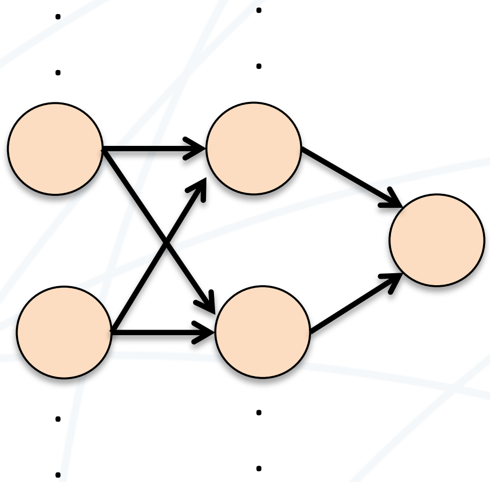
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

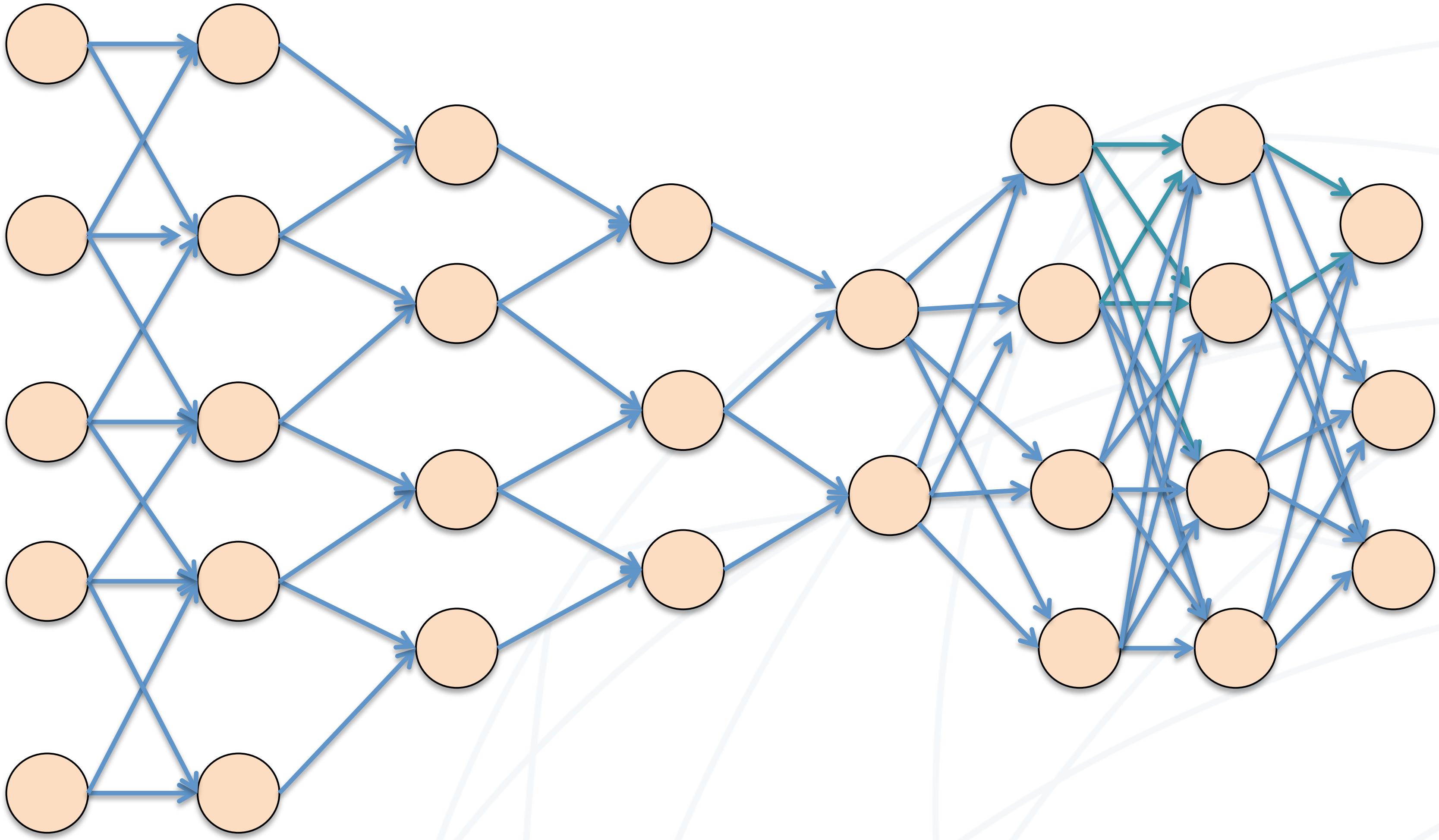
We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called "dropout" that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.

Hinton started playing with NN Architectures





Hinton started playing with NN Architectures



How does this happen ?

They “won” the ImageNet large scale visual recognition challenge with a convolutional neural network

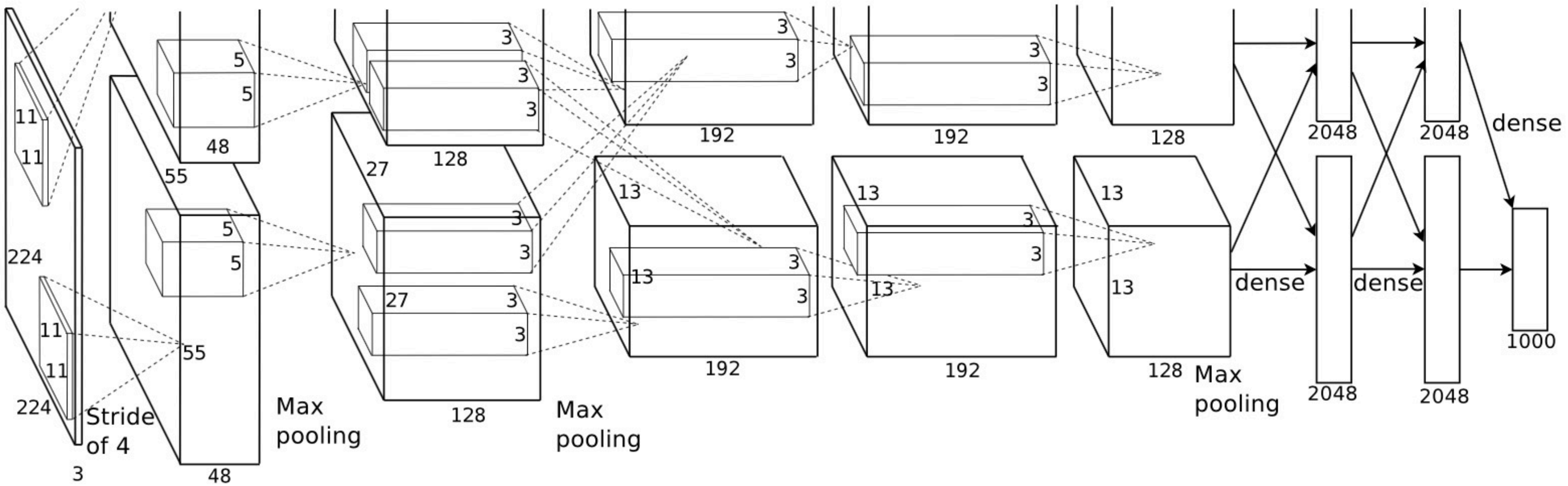
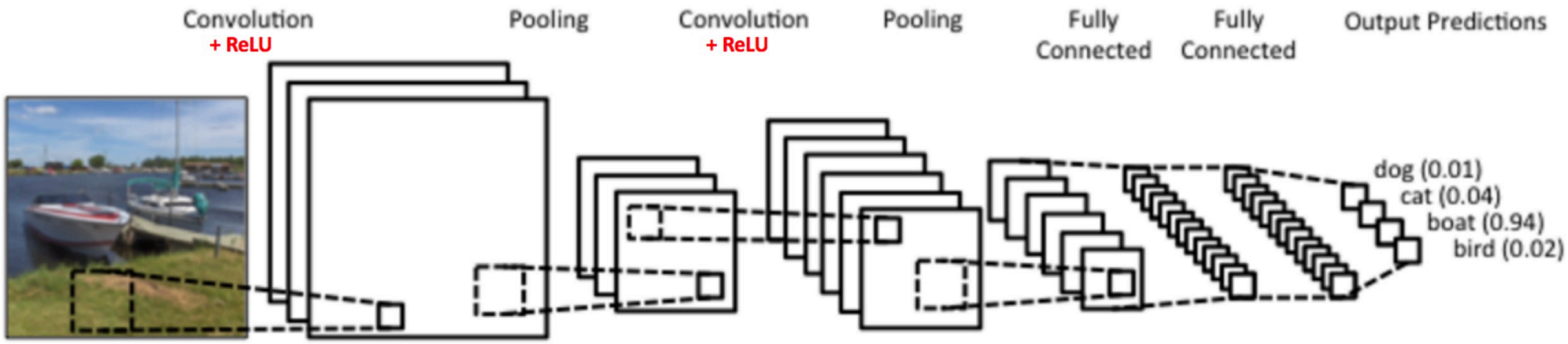


Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network’s input is 150,528-dimensional, and the number of neurons in the network’s remaining layers is given by 253,440–186,624–64,896–64,896–43,264–4096–4096–1000.

This is a CNN variant



CNNs and deep learning

- Convolutional Neural Networks are end-to-end solutions in which both the feature extraction and the classifier training are performed at once.
- The first layers extracts a type of information which is similar to the features/descriptors extracted in the classical approaches.
- These, called "*deep features*", are designed to preserve spatial image features and turn out to be significantly more efficient that the classical "*engineered*" ones, even when used with classical machine learning for classifier training
- Once a model for a concept is built, it can be packaged and released and easily run in a hosted environment

Online systems now available

Insight

Upload your photo

You can upload your photo or paste any URL to an image

Generated tags

Concepts

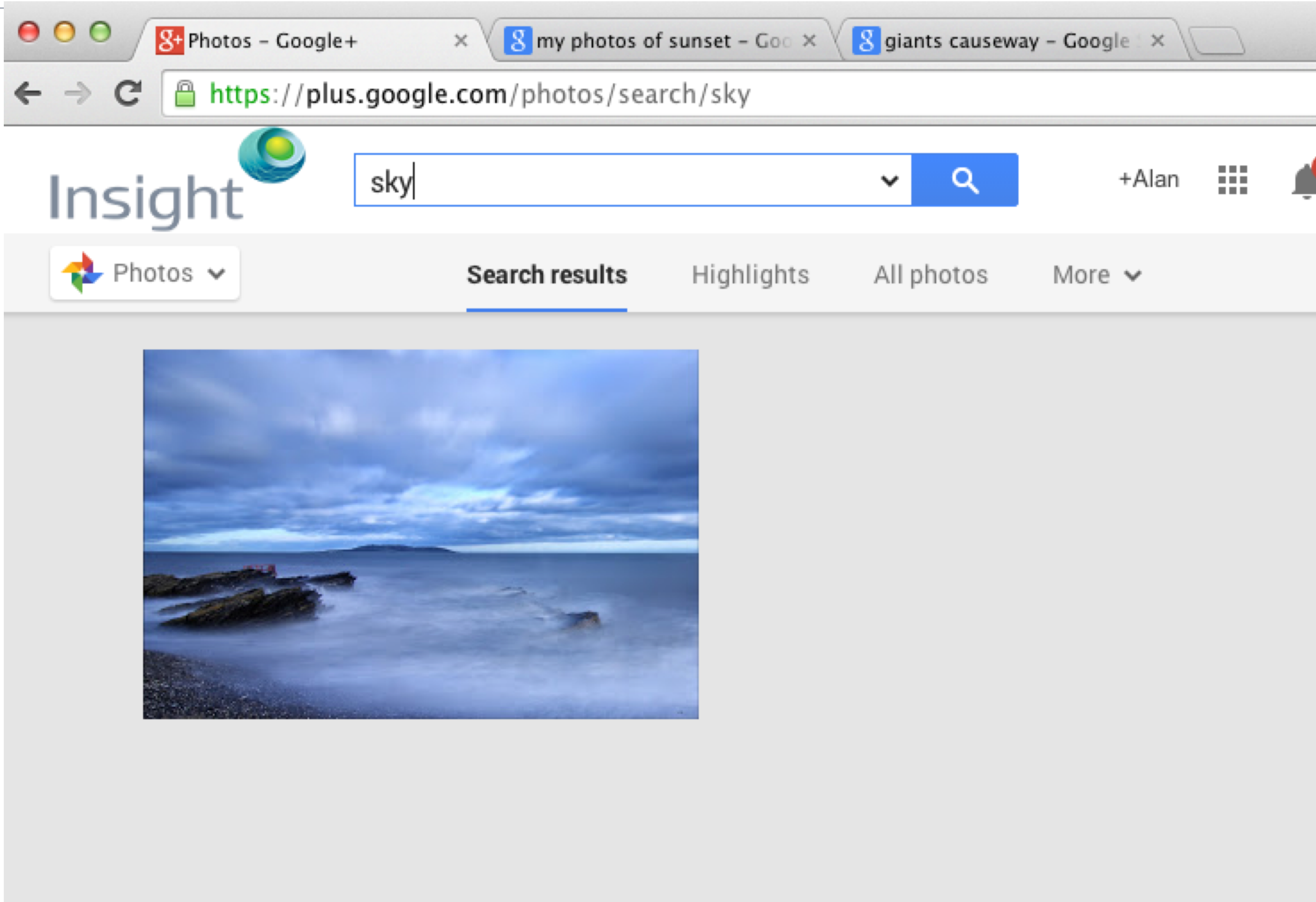
deer	79.76%
bighorn	38.25%
caribou	34.53%
mountain sheep	30.61%
bovid	25.83%
animal	23.11%
wild sheep	22.96%
wilderness	18.55%
goat	17.59%
mammal	15.33%

↪ show me more tags

Concepts in image search

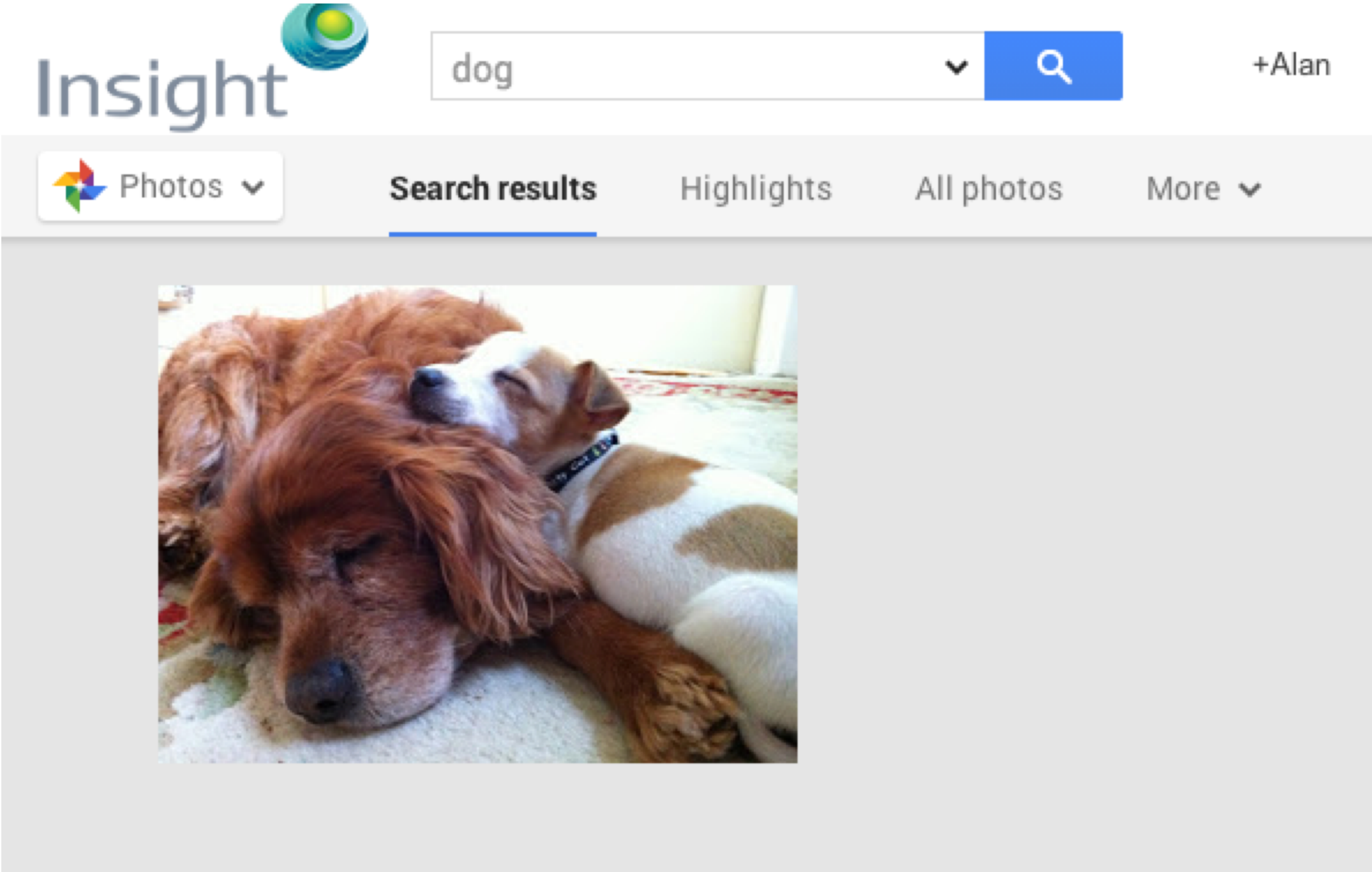
- Google+ photos now uses computer vision and machine learning to identify objects and settings in your uploaded snapshots
- FaceBook uses it to tag photos

My uploads

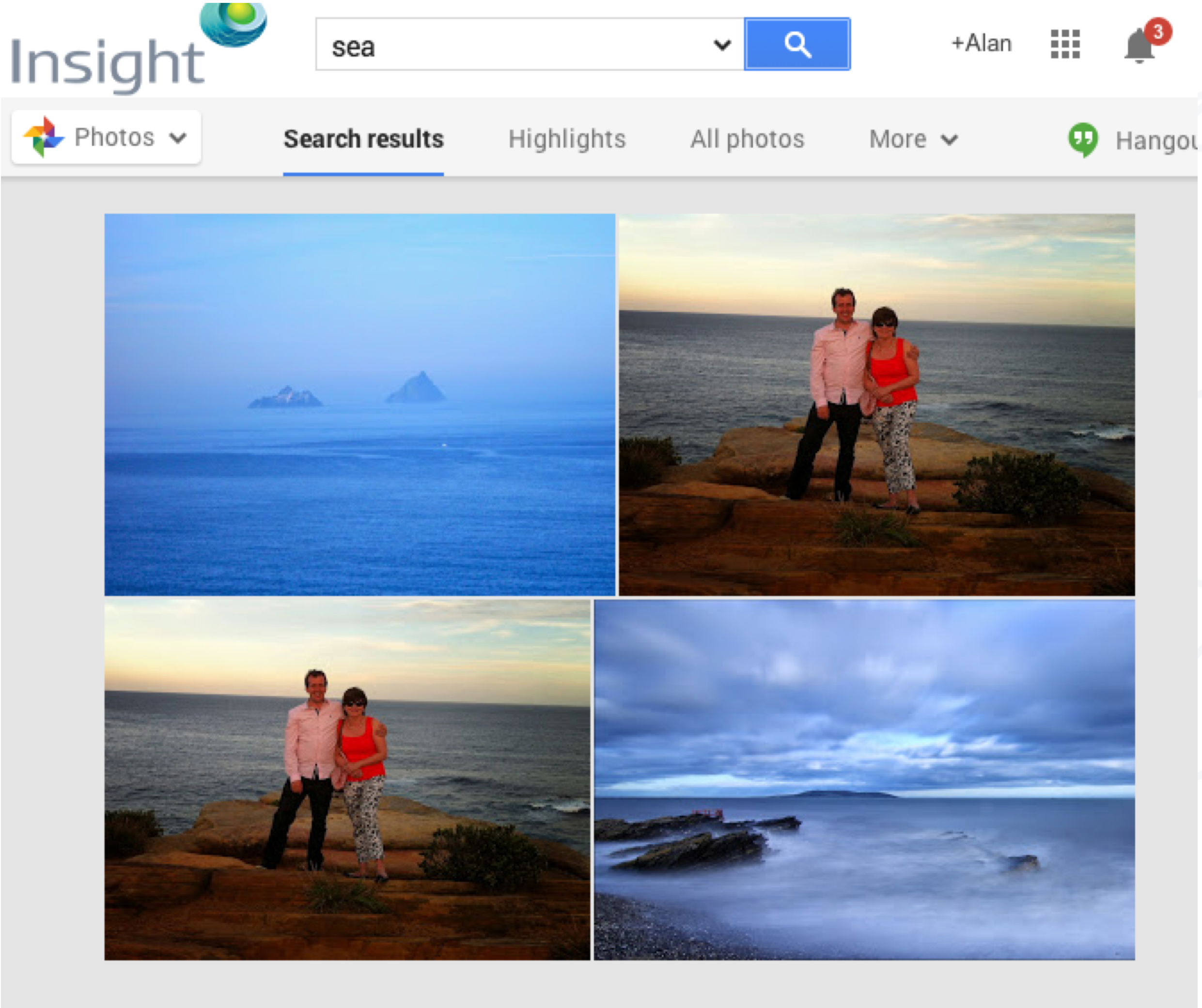


Photos – Google+
my photos of sunset – Goo
giants causeway – Google
https://plus.google.com/photos/search/sky
Insight
sky
+Alan
Photos
Search results
Highlights
All photos
More

My uploads


Insight
dog
+Alan
Photos
Search results
Highlights
All photos
More

My uploads


Insight
sea
+Alan
3
Photos
Search results
Highlights
All photos
More
Hangou

How good can a CNN get at automatically captioning images without using any context at all ?

Captioned by Human and by Google's Experimental Program

Human: "A group of men playing Frisbee in the park."
Computer model: "A group of young people playing a game of Frisbee."

Captioned by Human and by Google's Experimental Program

Human: "A young hockey player playing in the ice rink."
Computer model: "Two hockey players are fighting over the puck."

2 of 6

Captioned by Human and by Google's Experimental Program

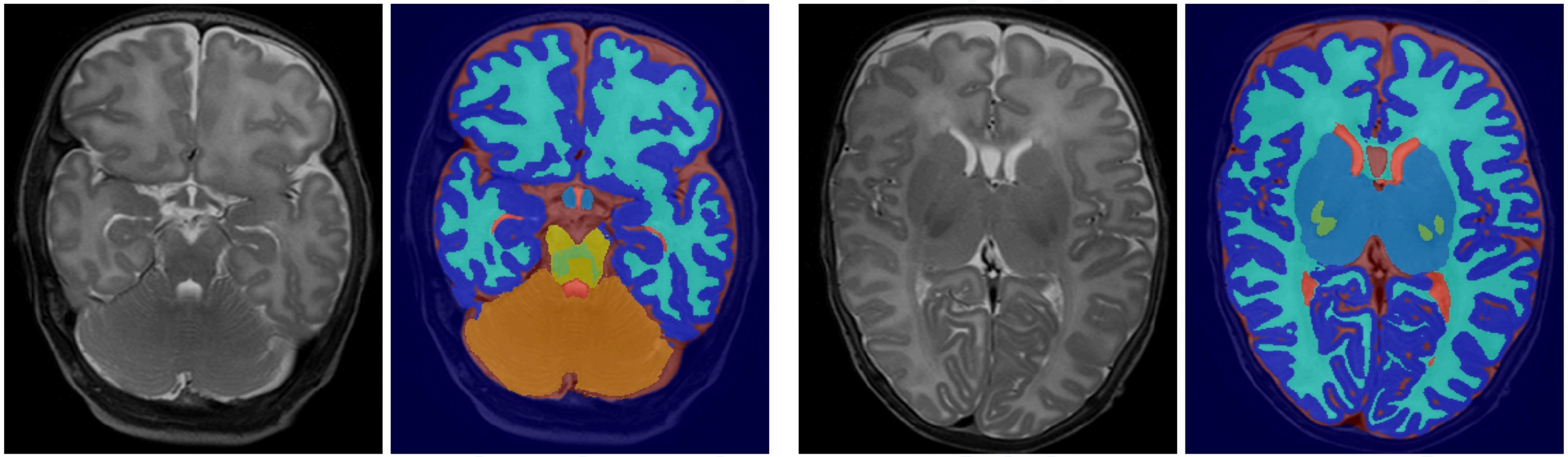
Human: "Elephants of mixed ages standing in a muddy landscape."
Model: "A herd of elephants walking across a dry grass field."

Now, everybody is using deep learning

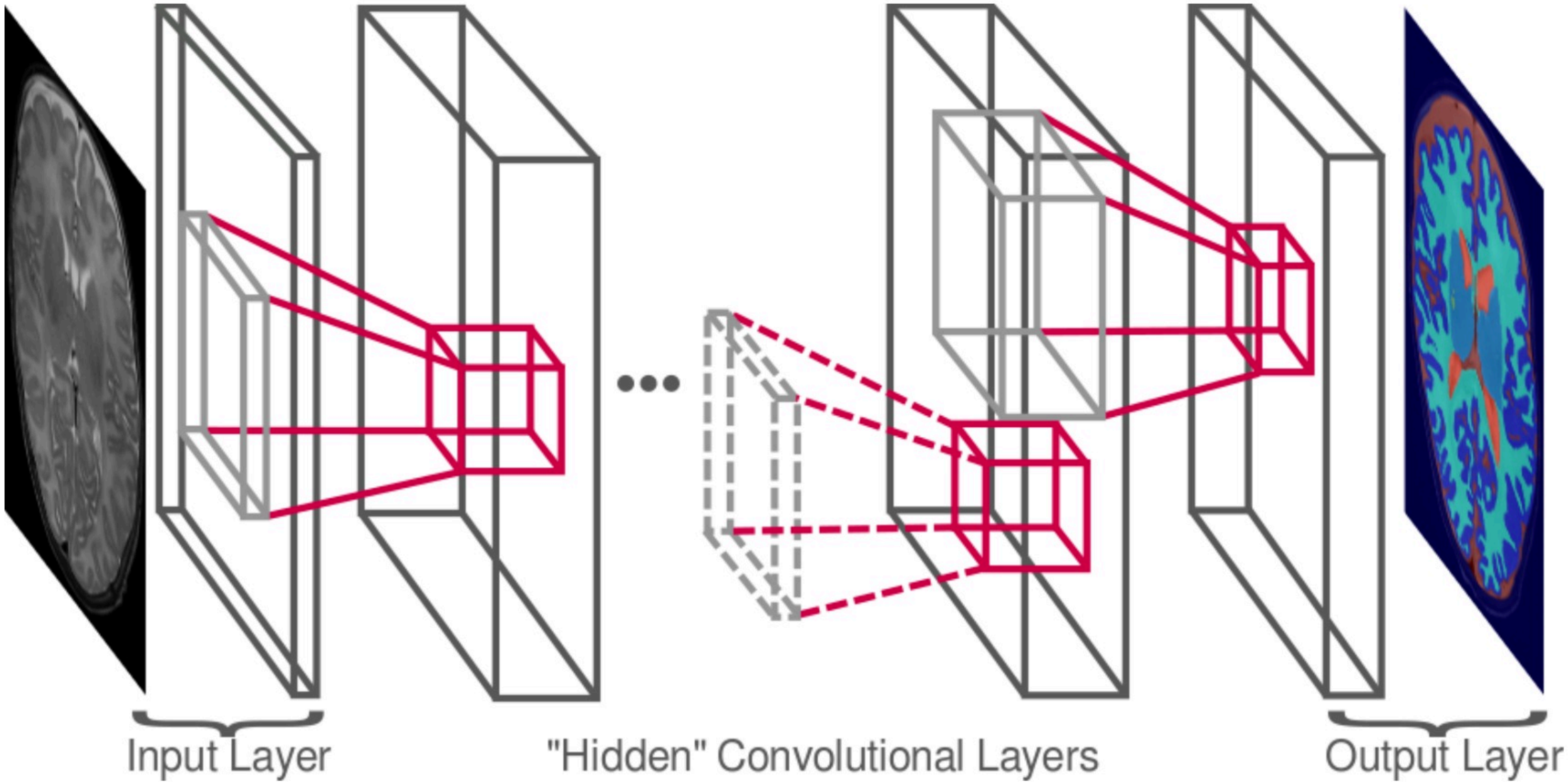
We used CNNs to classify MRI scans

Sample network segmentation results on neobrains test data.

Tissues are cerebellum (orange), cortex (darkest blue), white-matter (lightest blue), cerebro-spinal fluid (dark red), ventricles (light red), brain-stem (yellow), basal-ganglia/thalamus (mid blue), myelination (light green)



Our NN architecture …



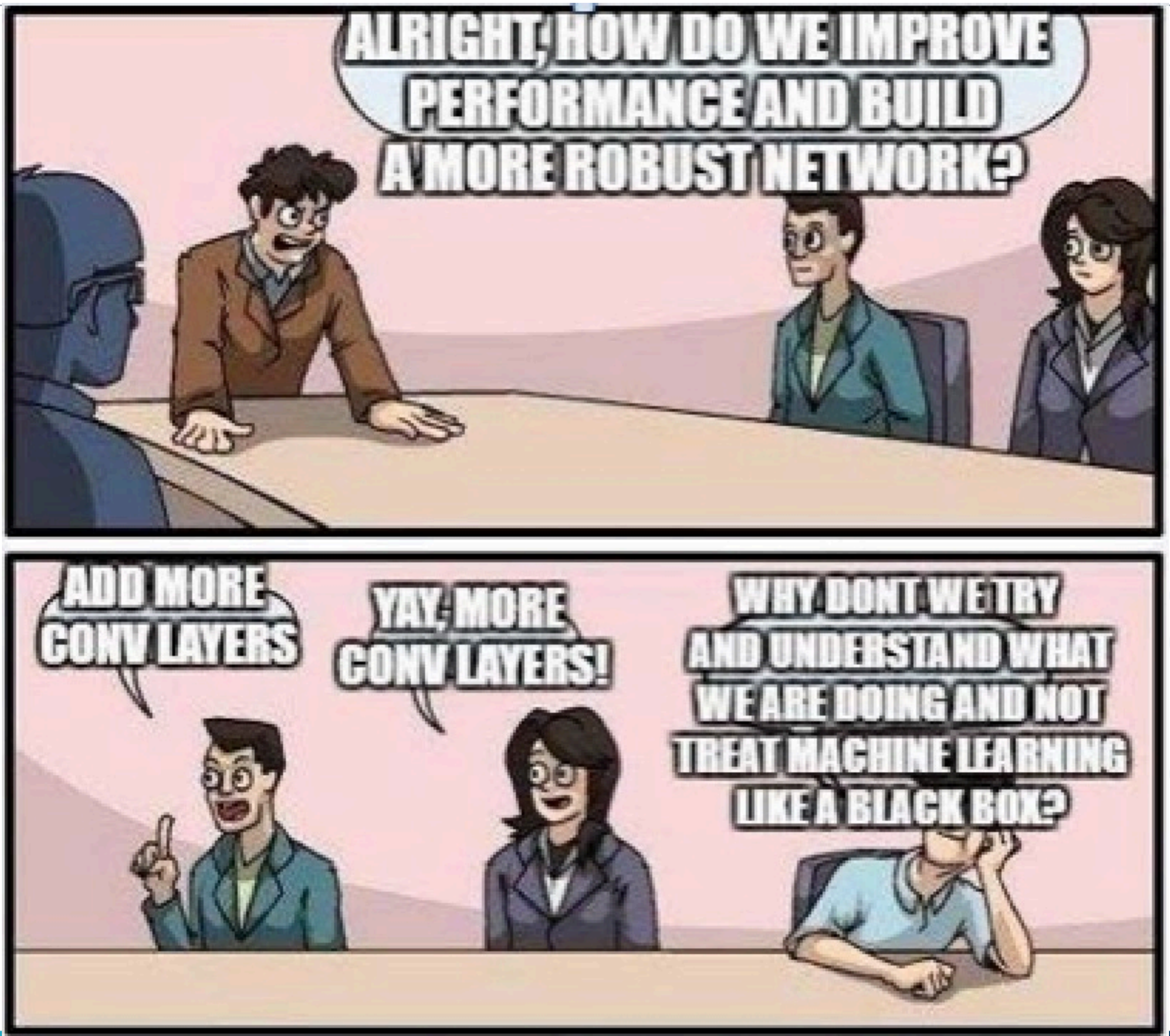
Neural Network components

- Neural Networks are a very **versatile model** for ML applications because the layers allow complex structures in the hidden layers and the activation functions
- For example, limiting the connections between layers to those which are in a receptive field from the previous layer, yields a **convolutional neural network**, good for image classification
- Allowing the output of a layer to feed back into itself yields a **recurrent neural network**, good for handling sequences of data and preserving temporal coherence

Neural Network components

- Hidden layers can themselves have configurations with:
  - **Dropout layers** where weights of some random percentage of nodes in a layer (5% to 15%) are set to 0, meaning they die
  - **Pooling layers** do dimensionality reduction, mapping N x N to n x n by choosing max or min or average from a window
  - **Binarising layers** act as gates (AND, OR, XOR) and simplify a network
  - … and others …
- Some of these are **biologically inspired**, from study of the brain, most are just **guesswork**

This is what we do now …



Neural Networks – are we finished ?

- October 2017 Hinton introduced a twist on neural nets to make machines better able to understand images or video
- **Capsule networks** group neurons whose activity represents parameters of a specific type of entity such as an object or object part
- Capsules form a hierarchy ... capsules at one level make predictions for the activation of higher-level capsules .. So they form abstractions ... when multiple predictions agree, a higher level capsule is activated

Neural Networks – are we finished ?

- Why ? Image-recognition needs a **large number of examples to** reliably recognize objects – not good at generalizing, e.g. different viewpoints
- Capsules build more knowledge into Neural Nets by tracking different parts of an object and their positions in space, giving in-built geometry
- A network of understands when a new scene is a different viewpoint of something it has seen before
- Already, achieves state-of-the-art performance on MNIST, considerably better than a convolutional net at recognizing highly overlapping digits, and as good as CNNs at recognising objects

Neural Networks – are we finished ?

- Neural networks/deep learning is fixated on humans encoding as little knowledge as possible, letting the network figure things out from scratch using large numbers of examples
- Capsule networks are a departure from this
- Hinton’s latest shows AI researchers doing more by mimicing how the brain has built-in, innate machinery for learning vision and language
- This is **really important**

How to implement CNNs ?

- Its definitely not von Neumann architecture
- So we throw massive parallelism … what's the cheapest massive parallelism … GPUs
- New role for Nvidia … not necessarily fast, but many of them !

- An alternative is to design new chips …
- Intel deep learning chips code named Lake Crest followed by Knight's Crest, under development, 2017 arrival
- Samsung handset chips and Qualcomm chips, allow deep learning on devices, and Movidius (now part of Intel) specialising in computer vision using deep learning on silicon … we hope its flexible deep learning supporting capsule networks ;-)

How to implement CNNs ?

- Google was starting to use deep networks for NLP applications like speech, but needed more horsepower, like x2 !
- Tensor Processing Unit (TPU) designed from scratch, very efficient
- Used for executing neural networks rather than training .. saved building x15 data centres and can run neural networks x30 times faster than conventional vN chip
- Downsides … need lots and lots and lots of training data, lots and lots of compute resources … and no facility for explaining why a decision or categorisation or output, is made

CADE METZ BUSINESS 04.05.17 12:03 PM

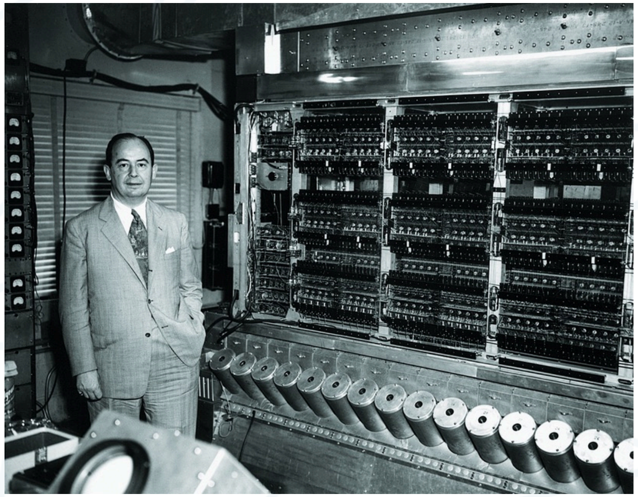
BUILDING AN AI CHIP SAVED GOOGLE FROM BUILDING A DOZEN NEW DATA CENTERS

Google server racks loaded with TPUs GOOGLE

What of von Neumann architecture ?

Insight

- We used to only have people, writing algorithms, encoded as programs which were stored and run
- We will always have such people doing this
- Now we also have data, used to train networks and develop models (sets of weights and network configurations) which are stored and run



Future Prospects ?

- Neural Networks are the new Computer Science, not replacing but creating new applications for data science
- Deep Learning is a temporary fixation, based on CNNs but new NNet architectures will materialise for different problems – vision, language, planning, decision-making ..
- Neuroscience will have the largest influence on data science and they will grow closer – symbiotically
- Specialist hardware will follow making NNets fast

- Will this be the computer model of the human brain ?
  - Sorry, my telescope doesn't see that far !