Transfer Learning with Convolutional Neural Networks

Off the shelf top notch performances

Eric Feuilleaubois

Ph.D in Artificial Neural Networks

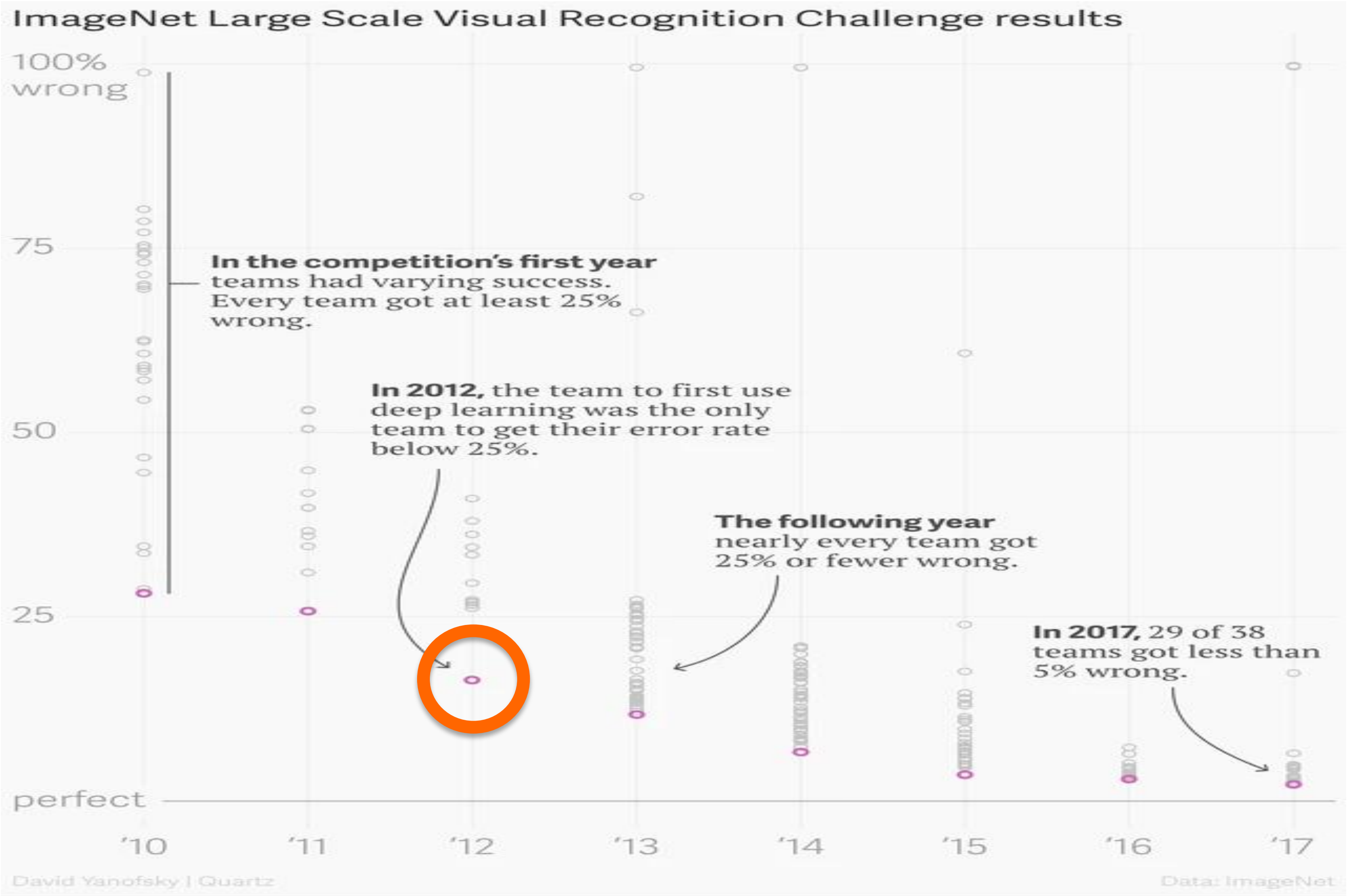
Deep Learner / Machine Learner

Curator of **Deep_In_Depth** - news feed on Deep Learning, Machine Learning and Data Science

Writer for Medium - Towards data science

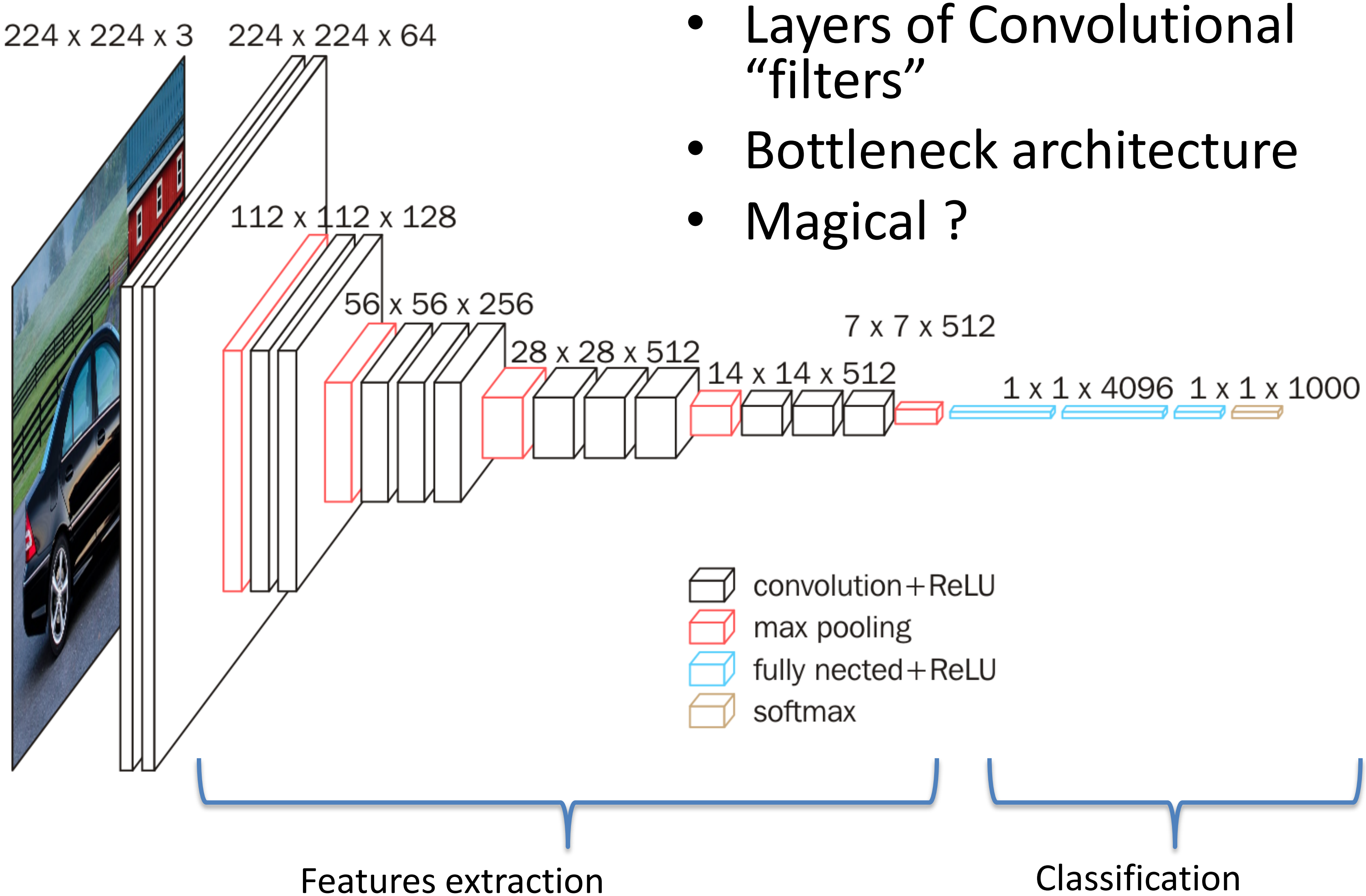
Convolutional Neural Networks

A breakthough

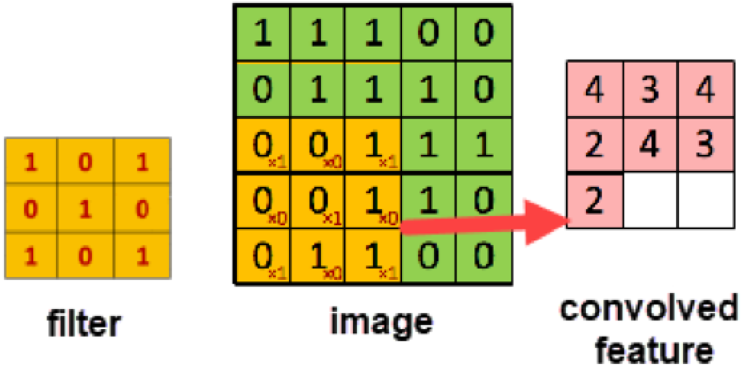


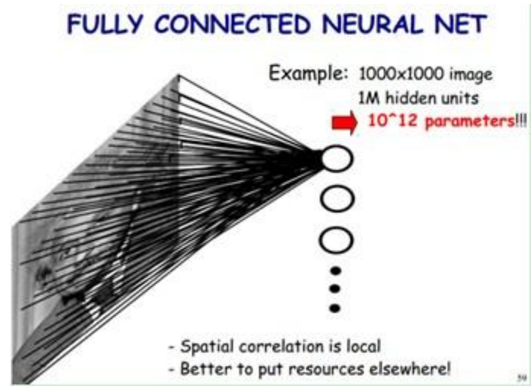
Convolutional Neural Networks
VGG-16 example

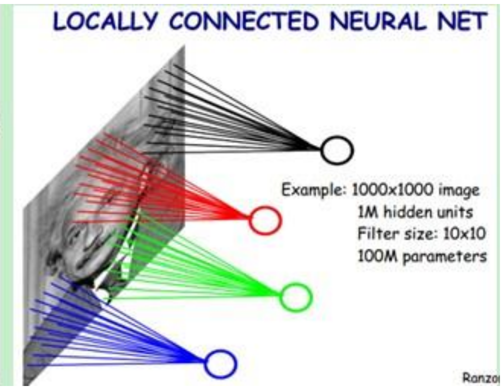
- Layers of Convolutional "filters"
- Bottleneck architecture
- Magical ?

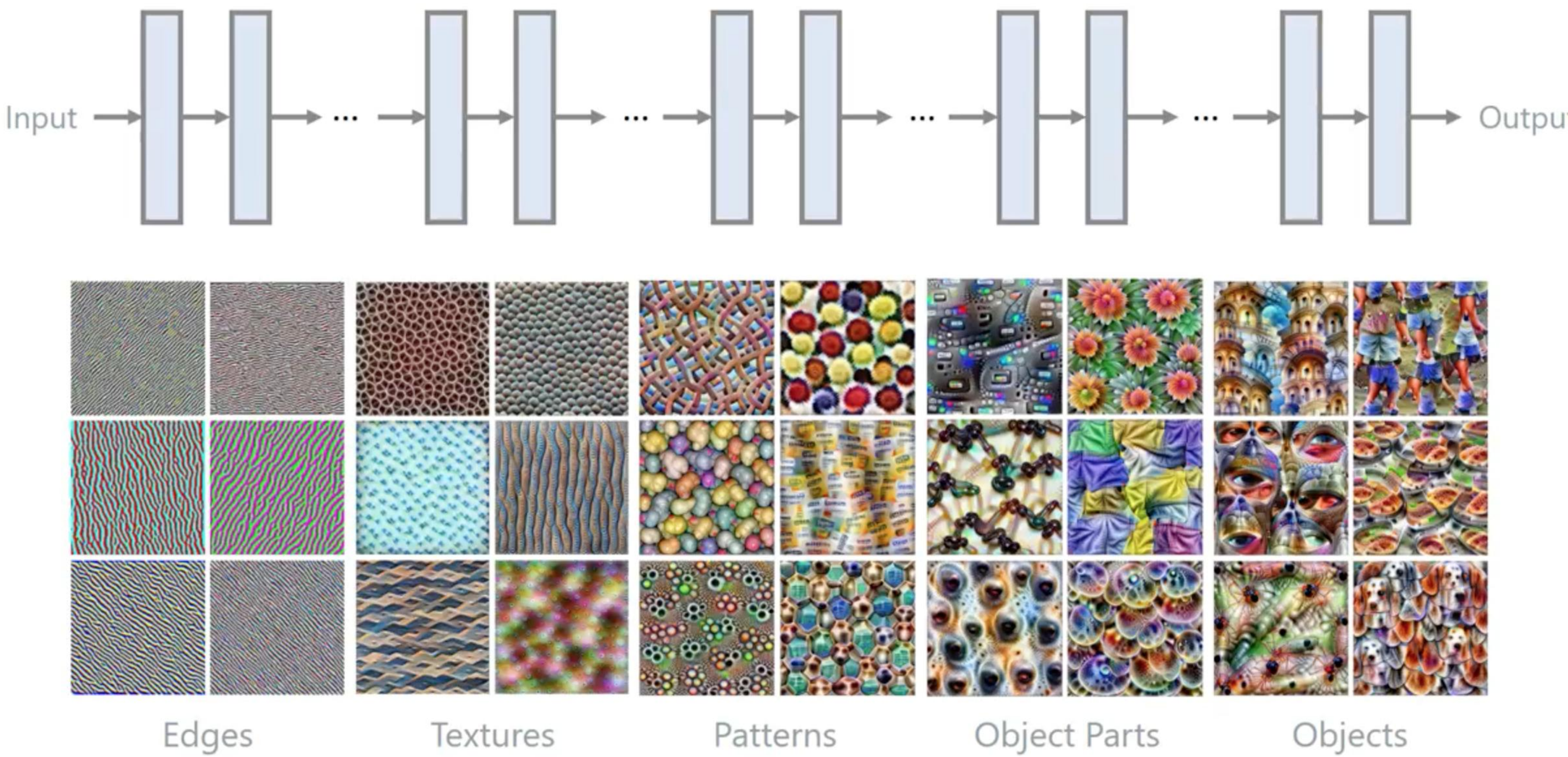


CNNs – Inner working

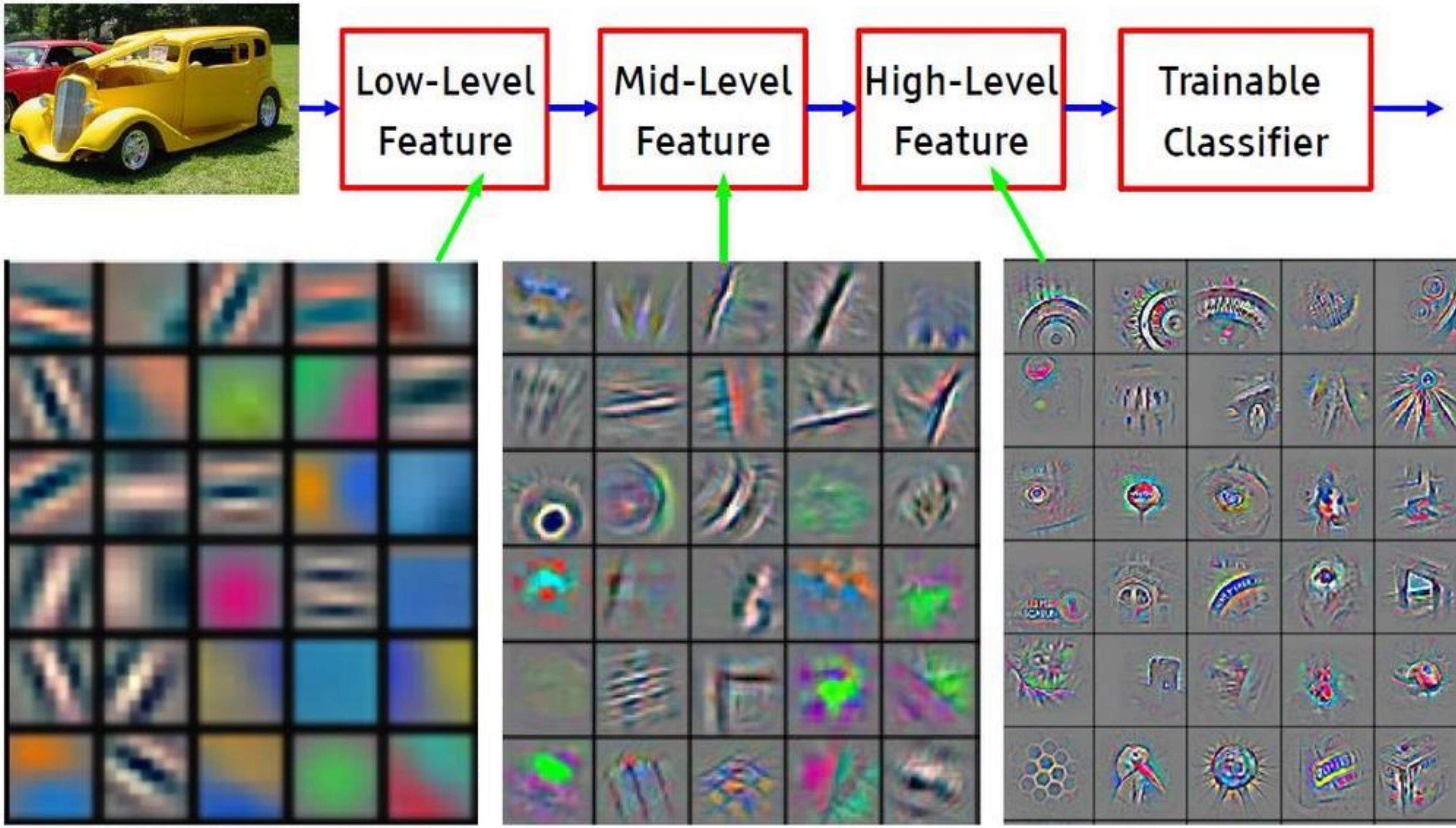








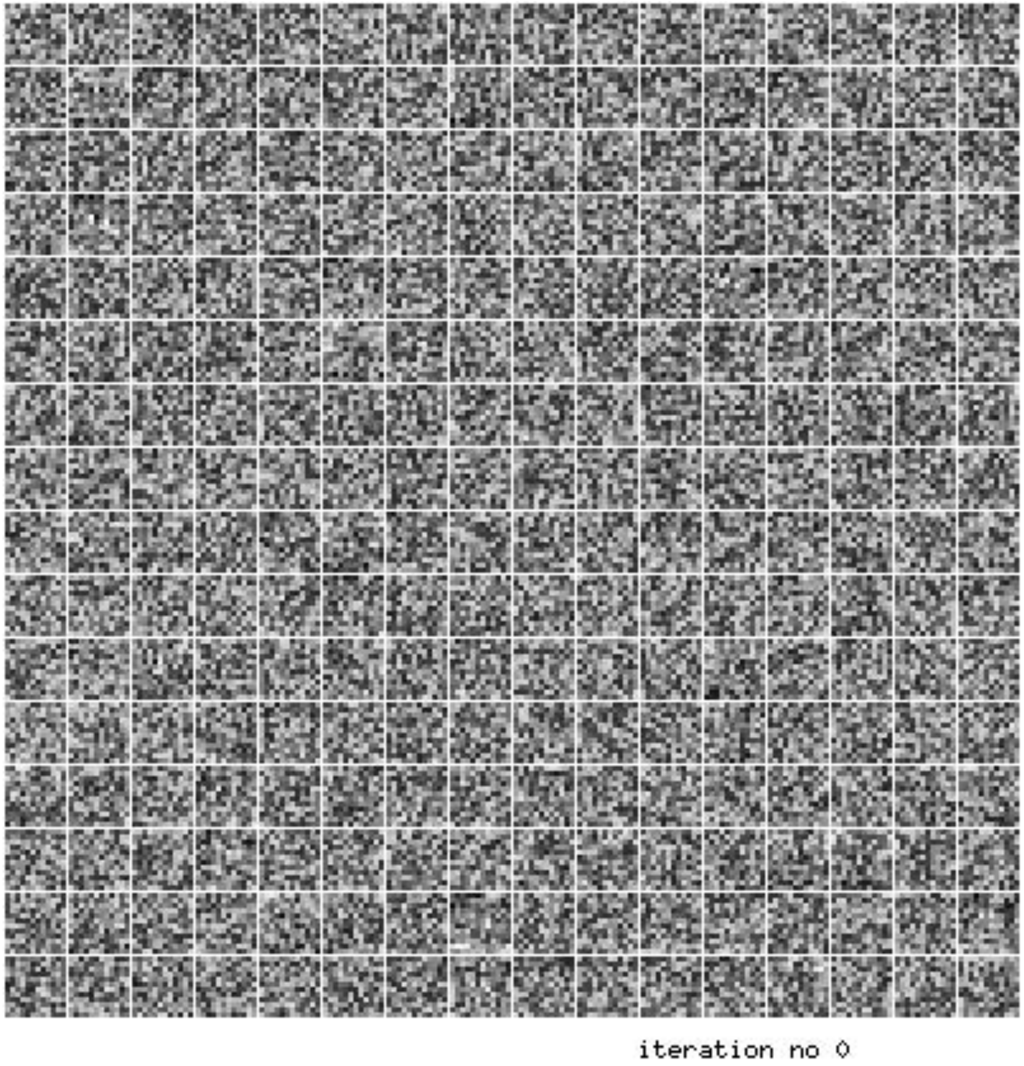
CNNs – Feature Extraction



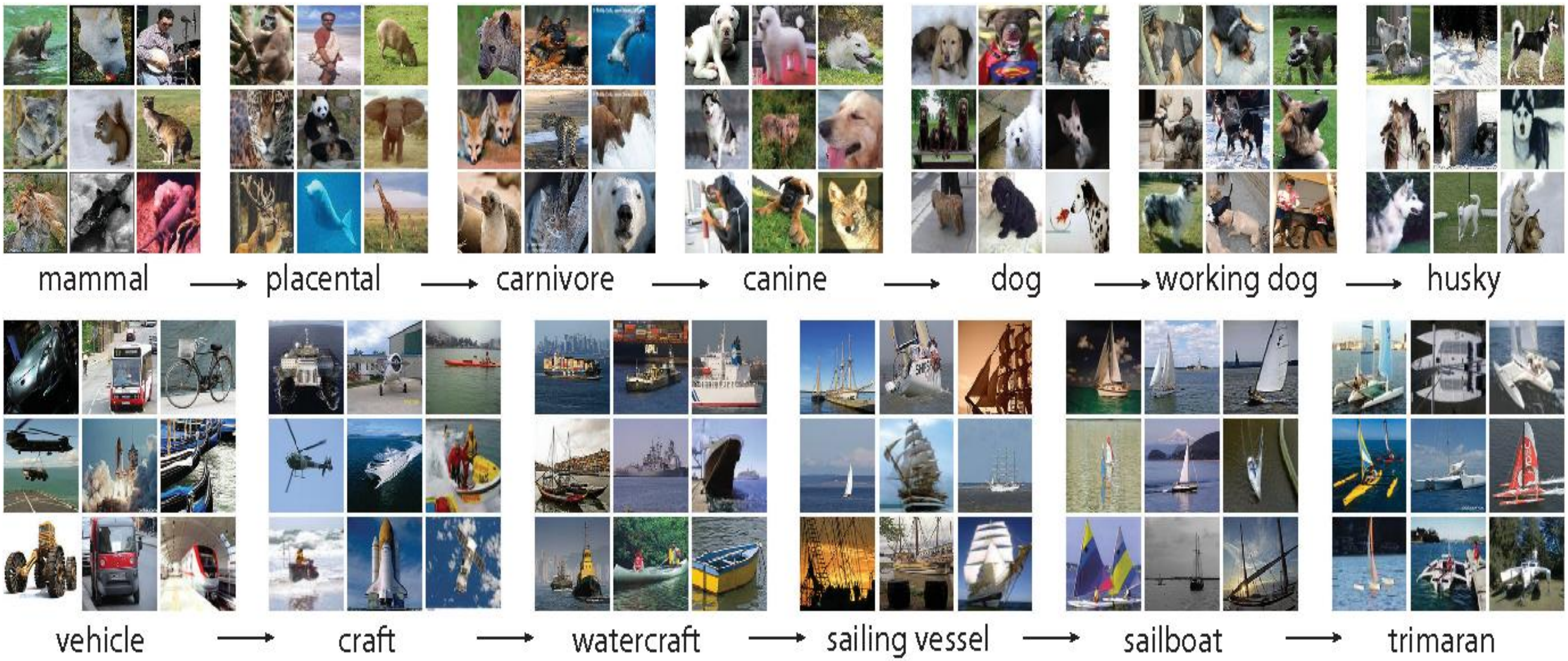
Training CNNs

- Big Dataset of 100 thousands images – usually millions
- Labelled data → Takes time to build
- Try out many different Network architecture
- Hyperparameter value :
  - training method, rate
  - Weights initial values
- No feature engineering

Training CNNs – Live view



ImageNet Database



ImageNet Dataset

- 1,000 image categories
- Training - 1,2 Million images
- Validation & test - 150,000 images

n07714571 - head cabbage
n07714990 - broccoli
n07715103 - cauliflower
n07716358 - zucchini, courgette
n07718472 - cucumber, cuke
n07718747 - artichoke, globe artichoke
n07720875 - bell pepper
n07730033 - cardoon
n07734744 - mushroom
n07742313 - **Granny Smith**
n07745940 - strawberry
n07747607 - **orange**
n07749582 - **lemon**
n07753113 - fig
n07753275 - **pineapple, ananas**
n07753592 - banana
n07768694 - **pomegranate**

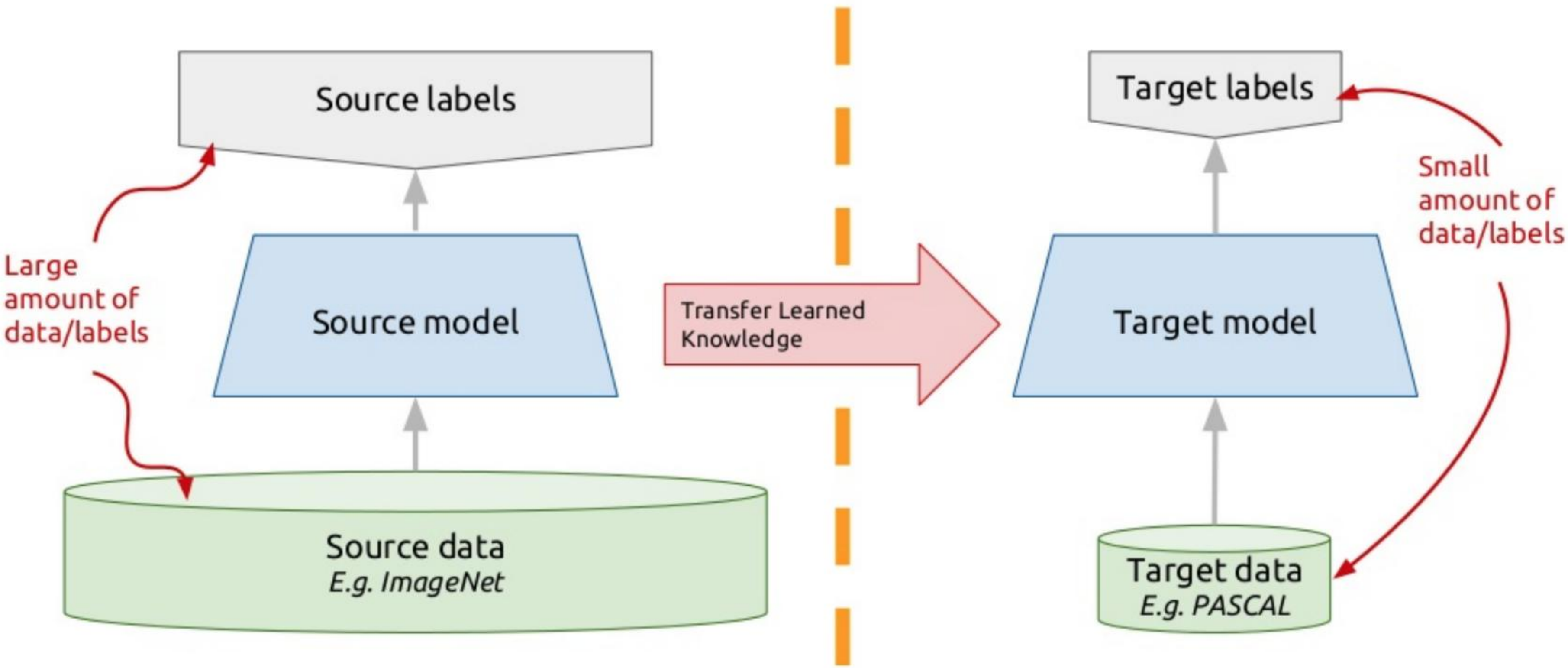
Veg Dataset - 4124 records

Pumpkin

Tomato

Watermelon

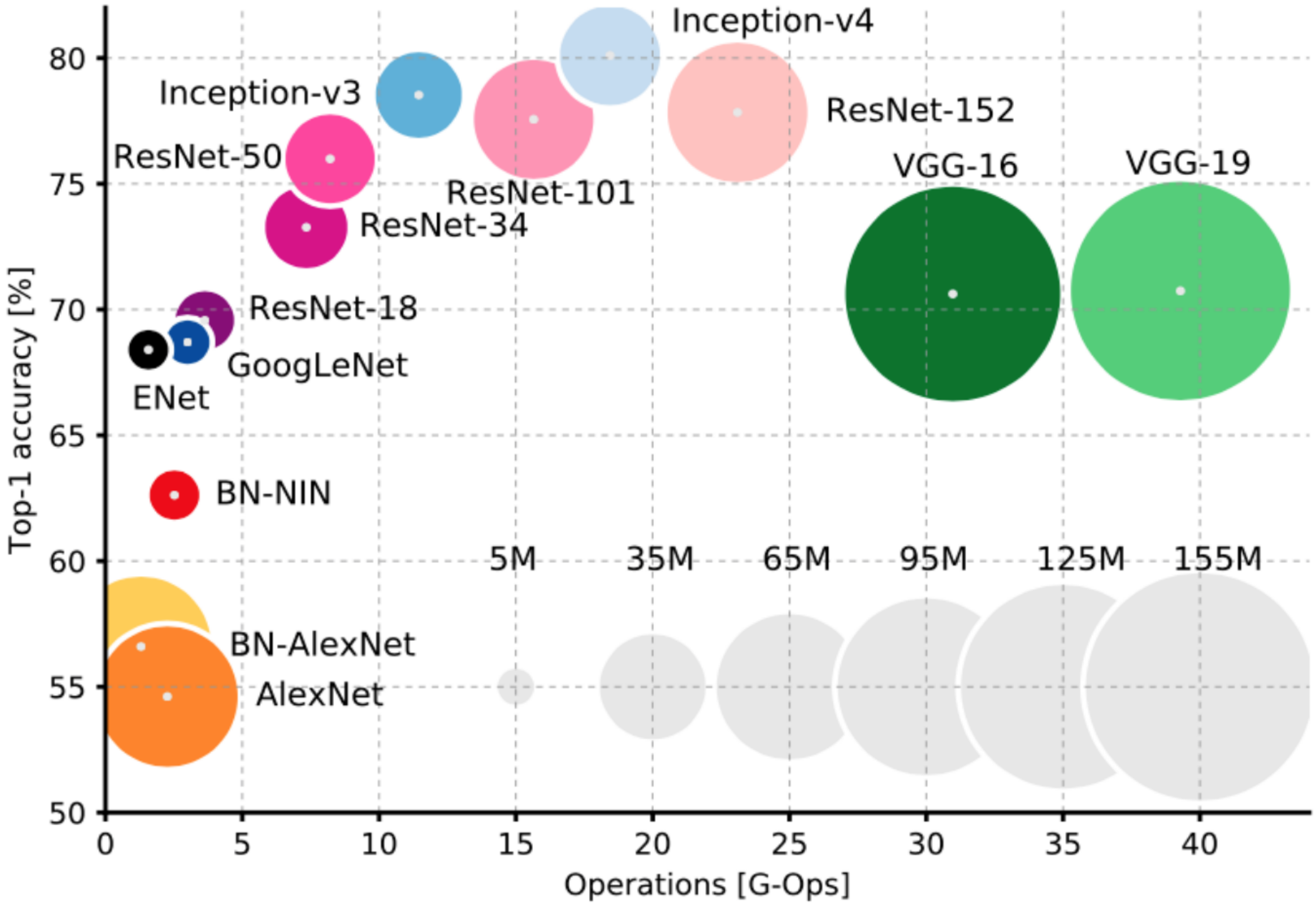
Principle of Transfert Learning



Aim: Predict classes (labels) that have not been seen by the source (pre-trained) model

Motivations for CNN TL

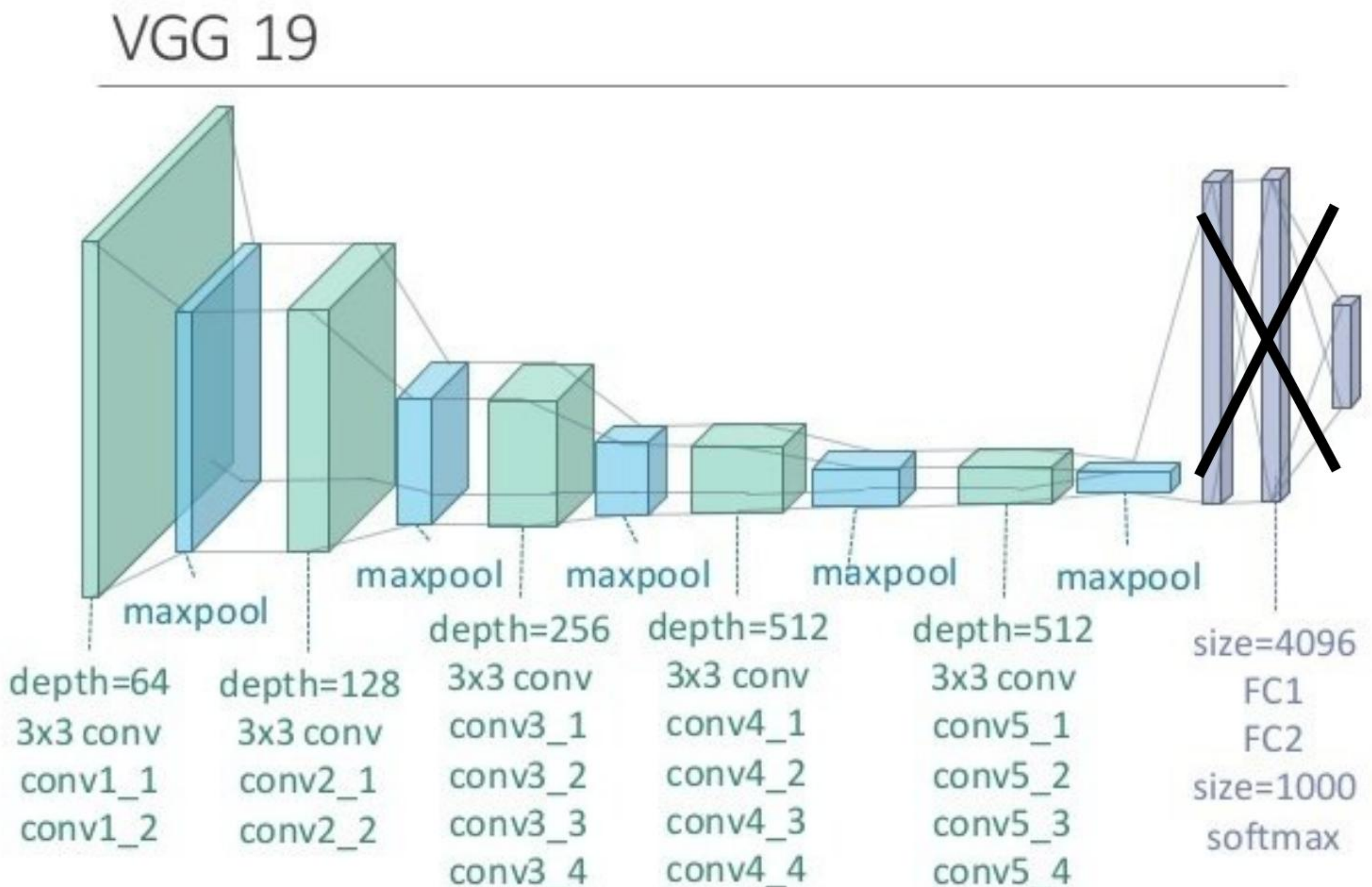
- Availability of Open CNN Models with Top class performance



From: AN ANALYSIS OF DEEP NEURAL NETWORK MODELS FOR PRACTICAL APPLICATIONS
Alfredo Canziani & Eugenio Culurciello & Adam Paszke

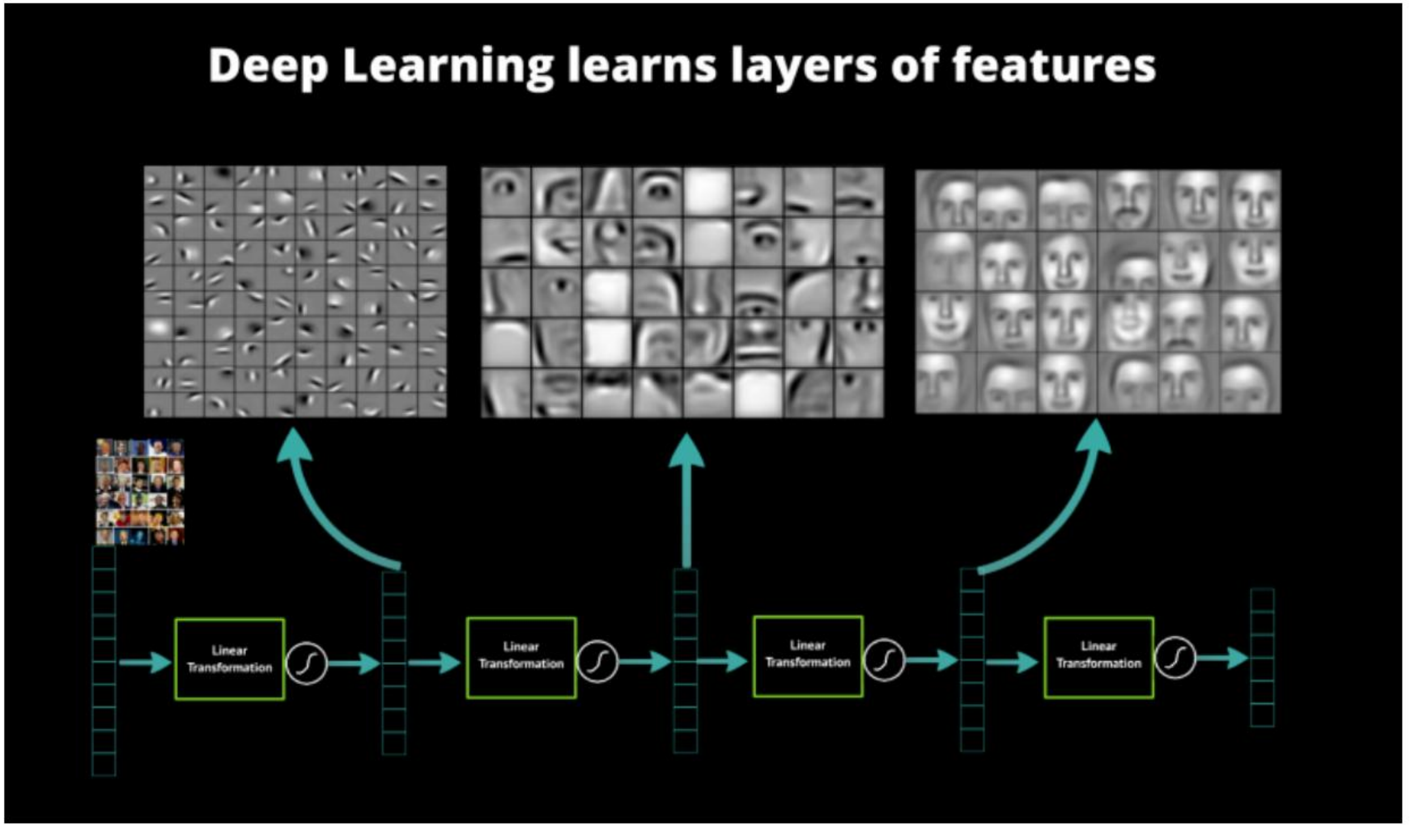
CNN Transfert Learning

Trained with ImageNet



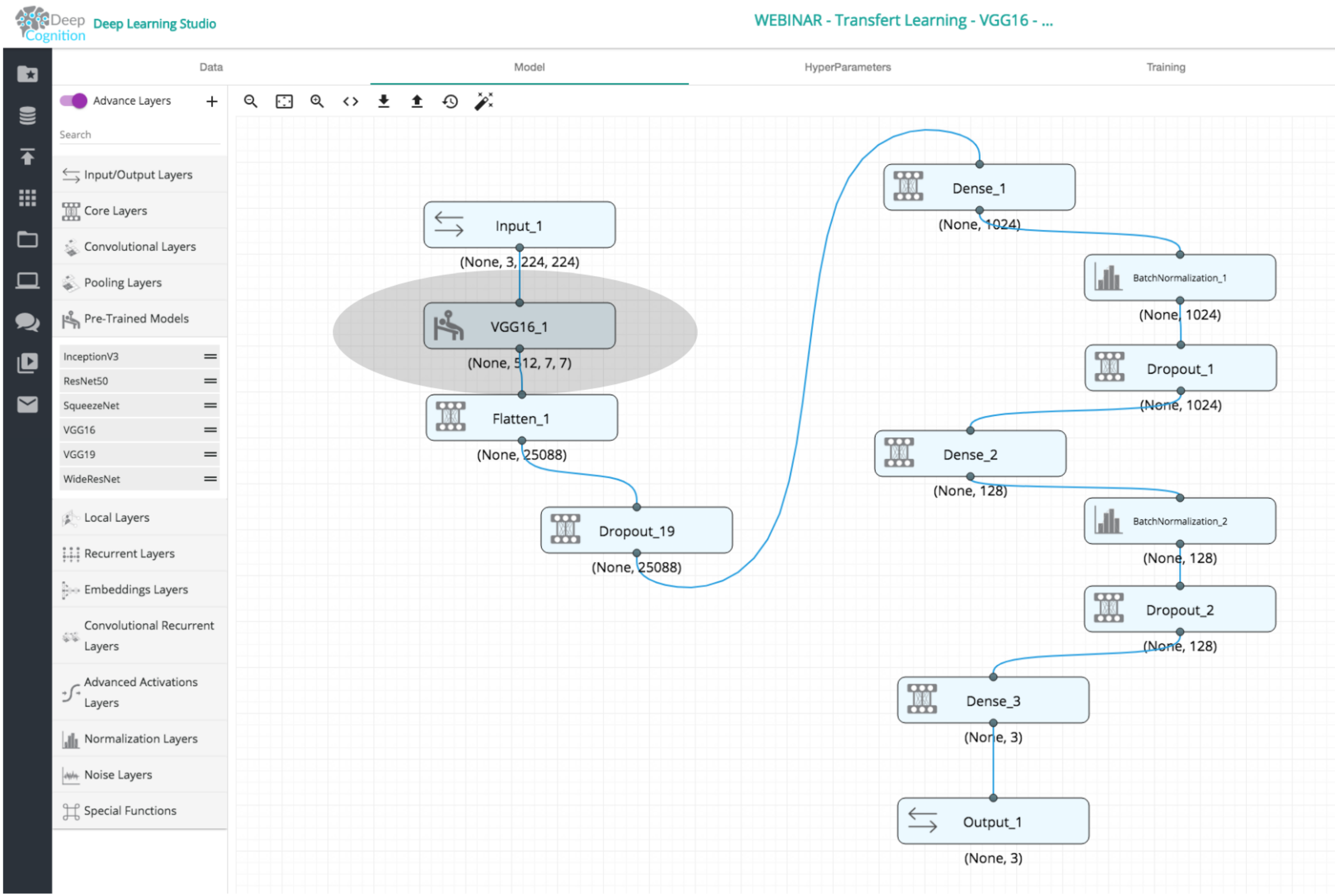
Proof by Features

- CNNs detect visual features (patterns) in images
- First layers learn "**basic**" features
- Last layers learn "**advanced**" features
- Top layers **classify**



Basic features are very similar from one CNN model to another --> **No need to re-learn them**, best to re-use them

DL model with pre-trained VGG16



Training results

Dataset	Train Acc. (%)	Validation Acc. (%)
300 - 7%	95	**88**
600 - 14%	96	89
900 - 21%	97	90
3917 - 95%	97	91
300 - with Data Aug.	98	88
600- with Data Aug.	98	90
900- with Data Aug.	98	90

Wrong results

		Training Dataset
	WaterMelon	300 - 600 - 900
	Pumpkin	300 - 600 - 900
	WaterMelon	300 - 600

Result for "difficult" images

	Pumpkin	300 - 600 - 900
	Tomato	300 - 600 - 900
	WaterMelon	300 - 600 - 900

Fine-tuning TL models

Two approaches:

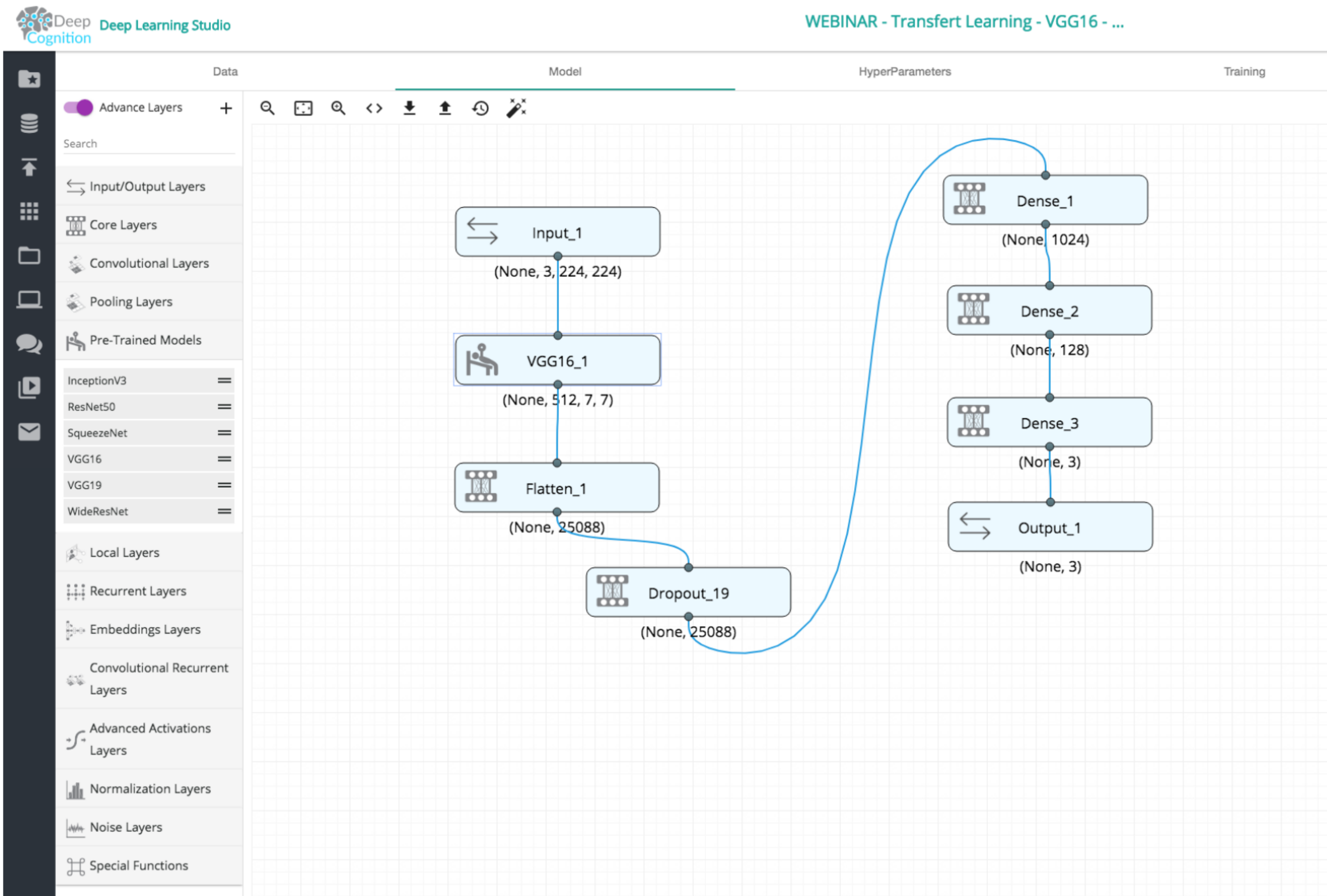
1) Fine tune the **convolutional** part of the CNN

- Retrained the last convolutional layer (Deep Learning Studio provide easy way to do it)
- Freezing and training specific layers

2) Fine tune the **classification** part of the CNN

- Classification architecture optimisation
    - number of neuron in the classification layers
- Hyper-parameter optimization

Simpler DL model with pre-trained VGG16



Training results

Dataset	**Train Acc. (%)**	**Validation Acc. (%)**
300 - Simpler	99	**88**
300 - 512 - 64	99	86
300 - MSimpler-no BN	98	84
300 - Simpler VGG 10% trainable	99	**89**
3917 - 95% - Simpler	98	**93.5**
95% - MSimpler-no BN	40	35
95% - MSimpler- BN	97	**92**

Benefits of CNN Transfert Learning

- Dataset does not need to be huge
- Save time and effort
  - no CNN architecture search
  - training phase faster
  - less computing power needed
- Produce very good models that can then be fine tune
- Can be applied to very diverse classification problem
- Less likely to overtrain

And drawbacks:

- need a fair amount of memory on GPU
- not fully tunable with Keras

Alternatives

- CNN architecture & hyperparameter "automatic" tuning
  - AutoML (Architecture) – driven by "AI"
  - IBM Watson Suite (Hyperpameters)
  - Microsoft Custom Vision Services

- Drawbacks
  - Limited exploration of the space of possibilities
  - Black box inside a "black API"