



# Searching for Images in the era of Deep Learning

STM Innovations  
December 5<sup>th</sup>, 2018

DR HARALAMBOS MARMANIS  
EVP & CTO, CCC

# “Drops splashing” on Google



drop splashing



All **Images** Videos News Maps More Settings Tools

View saved SafeSearch



Slow motion water drop splash - Yo...  
youtube.com



Why water drops splash: a non-trivial ...  
zmescience.com



High Speed and Photoinst...  
davidhazy.org



Water Drop / Splashing | HD Stock Vi...  
footage.framepool.com



High Speed and Photoinst...  
davidhazy.org



Why does water splash? We ma...  
lbtimes.co.uk



Water Drop Collision Created When ...  
shutterstock.com



Splashes, splashes and more ...  
people.fri.edu



Water Drop Liquid - Free photo on Pixabay  
pixabay.com



Drop Splash - Vector water splash...  
kispng.com



Realistic Water Drop Splash Free vector ...  
all-free-download.com



Splash (fluid mechanics) - Wikipedia  
en.wikipedia.org



Water Drop Splash - Free photo on ...  
pixabay.com



Water Drop / Splashing HD Stock V...  
footage.framepool.com



Drop splashing in water stock photo ...  
dreamstime.com



# What if?

---

- We had hand-drawn sketches for a book and wanted to search our image inventory to identify professional drawings or pictures?
- What if we wanted to organize the images for an astronomy journal so that all look-alike galaxies are grouped together?
- What if we wanted to autogenerate captions for our vast inventory of images?
- What if ... ?

# Before Deep Learning

---

- Metadata-based image indexing
- Content-based image indexing
  - Color histograms (HSV, RGB)
  - JPEG Coefficient histogram
  - Visual Bag of Words with SHIFT or SURF
  - Approximate fast search based on hashing and metric indexing

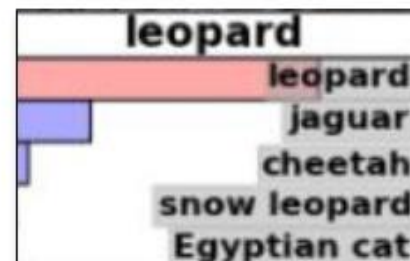
Great tool to use for the above: LIRE

# ILSVRC

- Task: Image Classification (1000 Classes)
  - 1.2m images in the training set
  - 100k images in the test set
- Goal: Predict the best (top 5) class
- Evaluation: One of the top 5 is correct



Classification



# ImageNet search for “dog”

animal, animate being, beast, brute, creature, fauna (399)

invertebrate (766)

homeotherm, homoiotherm, homotherm (0)

work animal (4)

darter (0)

survivor (0)

range animal (0)

creepy-crawly (0)

domestic animal, domesticated animal (213)

domestic cat, house cat, *Felis domesticus*, *Felis c*

dog, domestic dog, *Canis familiaris* (189)

pooch, doggie, doggy, barker, bow-wow (0)

hunting dog (101)

dalmatian, coach dog, carriage dog (1)

cur, mongrel, mutt (2)

corgi, Welsh corgi (2)

Mexican hairless (0)

lapdog (0)

Newfoundland, Newfoundland dog (0)

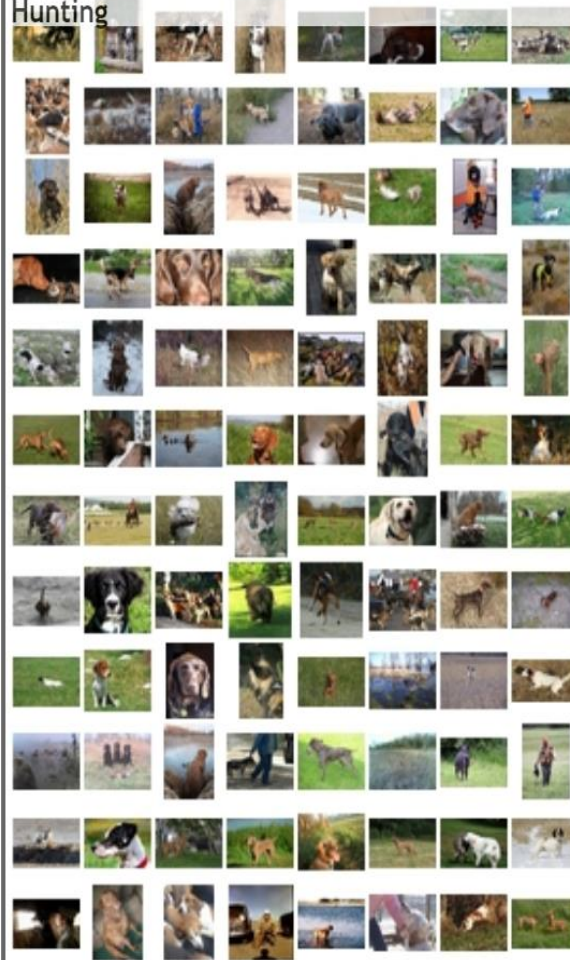
poodle, poodle dog (4)

basenji (0)

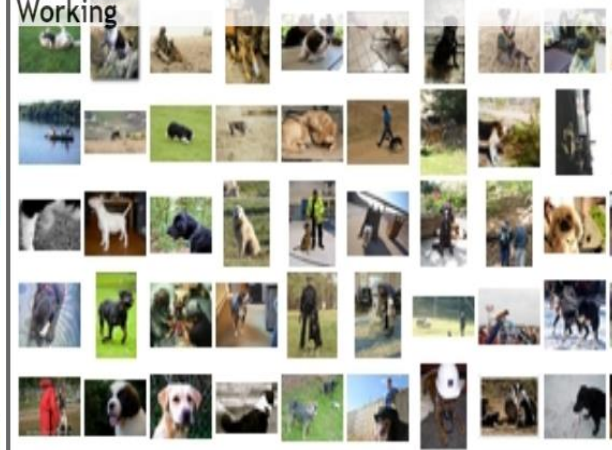
Leonberg (0)

griffon, Brussels griffon, Belgian griffon (1)

Hunting



Working



Toy



Leonberg



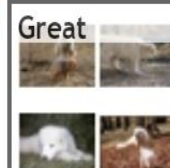
Pooch



Pup



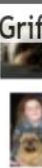
Great



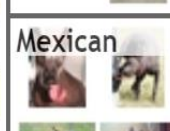
Dalmatian



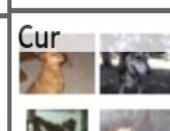
Griff



Mexican



Cur



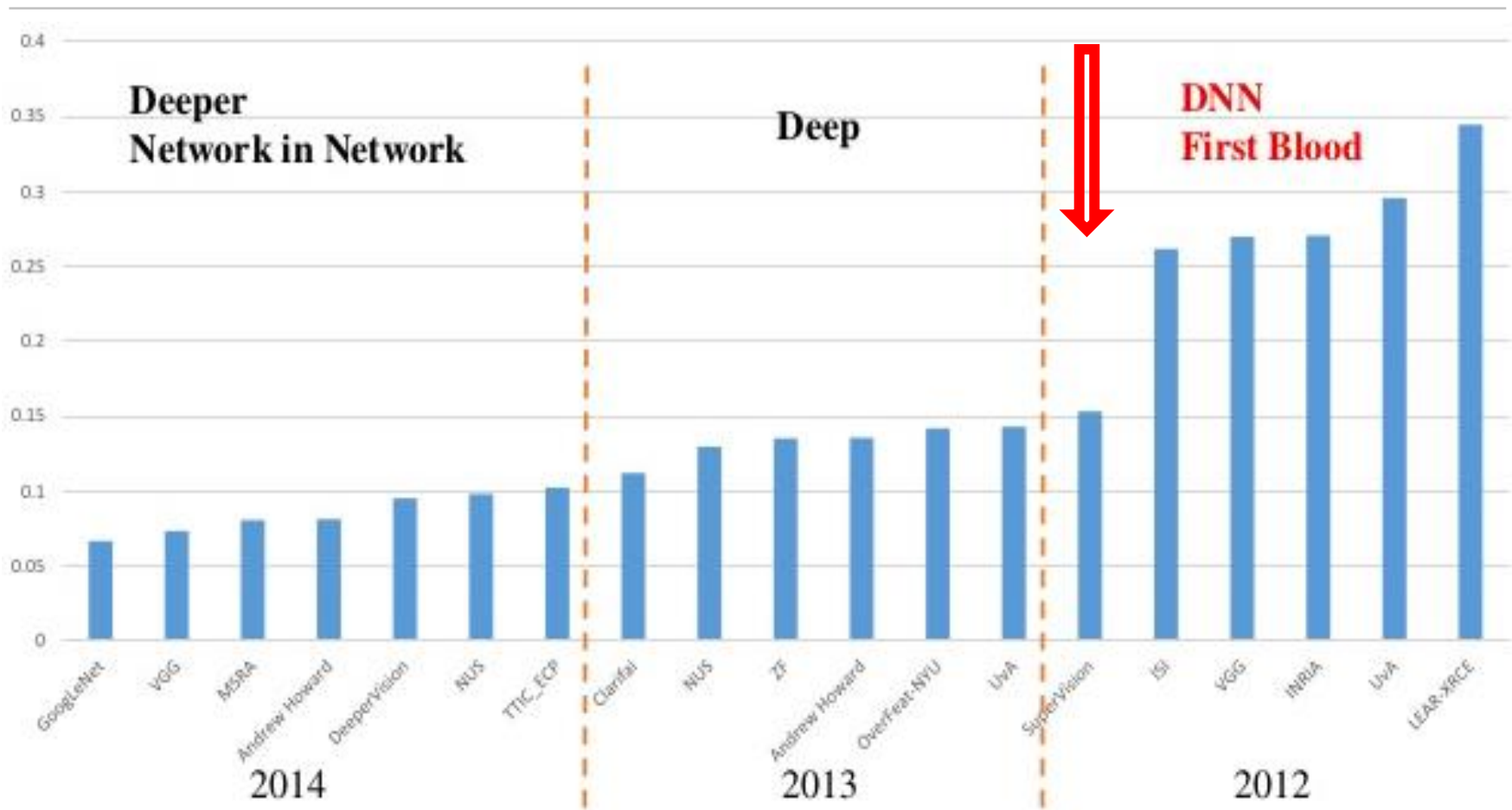
Newfoundland



Poodle



# Let the party begin!



Li Fei-Fei: ImageNet Large Scale Visual Recognition Challenge, 2014 <http://image-net.org/>

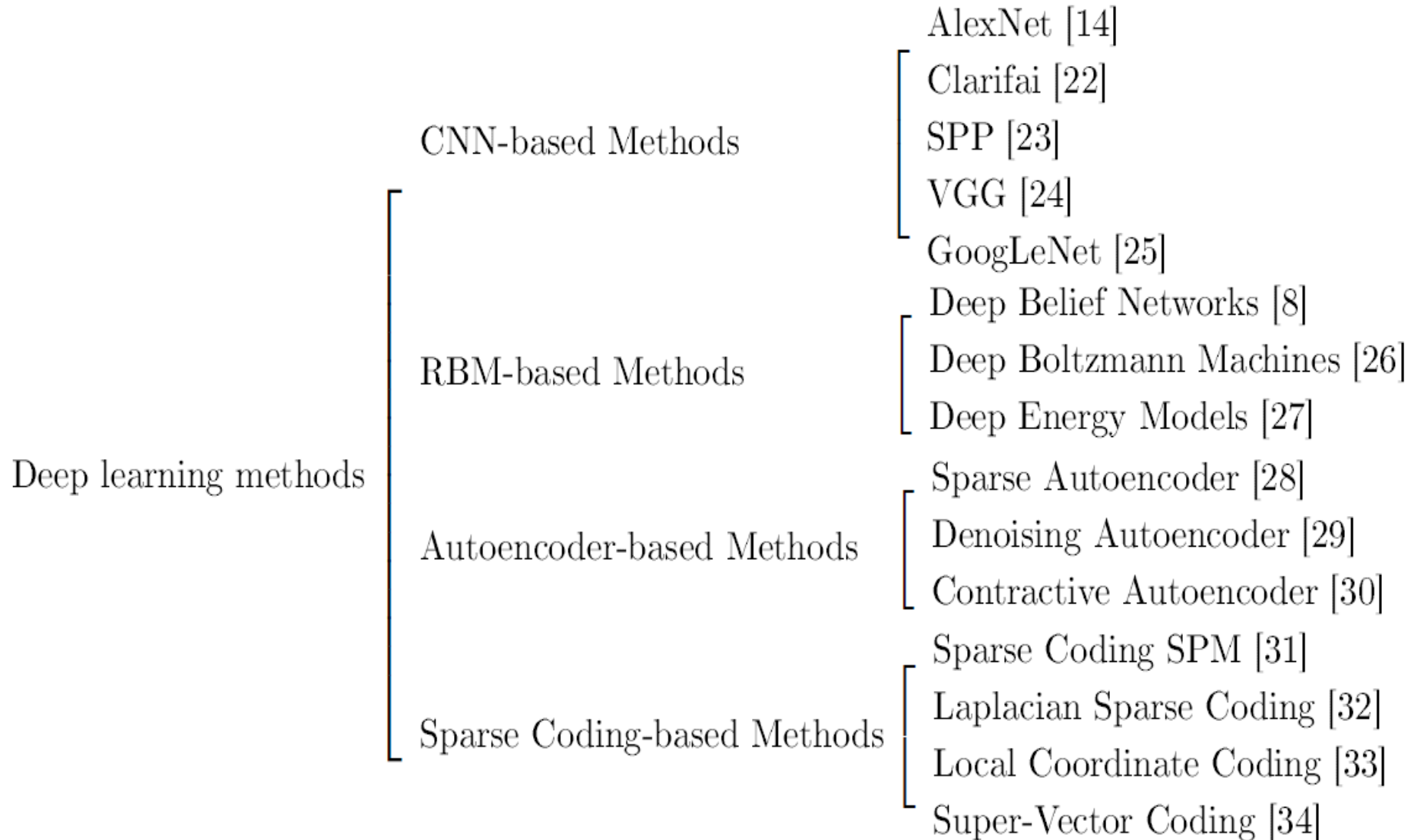
# ILSVRC Progress

Year	Winner	Error	Depth	Filter	C	T (Days)
2012	AlexNet	15.4%	7	11x11	M	6
2013	ZFNet	11.2%	7	7x7	M	12
2014	GoogLeNet	6.7%	100	1x1, 3x3, 5x5	H	7*
2015	ResNet	3.6%	152	3x3	L	14-21



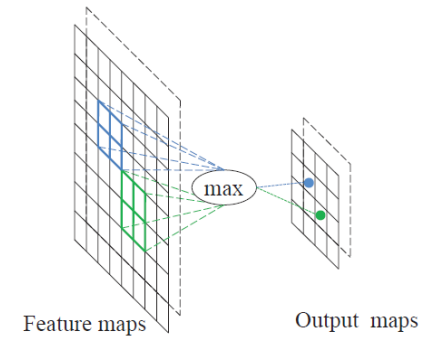
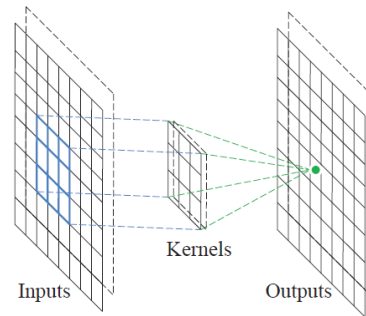
# Deep Learning is not just CNNs

---



# Typical CNN for image processing

- Input layer
- Convolution layer
- Pooling layer
- Convolution layer
- Pooling layer
- .....



• Dense layer

← Image Representation

• Output layer

← Image Classification

# Enhanced Search with CNN models

---

- Build a CNN model
- Train the model
- Store the model
  - Get the CNN feature vectors and index them in Lucene as Points
- Perform a feedforward pass with our image
- Use the `FloatPointNearestNeighbor.nearest(...)`
- Select the top-k results

# The art of the trade

---

- Network architecture
  - How many layers? What type of layers? How many nodes in each layer for a certain task?
- Training strategy
  - Dropout & DropConnect
  - Data Augmentation
    - Generate image translations and horizontal reflections
    - Alter the values of the RGB channels in training images
  - Pre-Training and Fine-Tuning
- Training with limited data
- Time complexity