



# Exemplar Deep Learning Applications



# Image Classification

# Objectives



## Objective

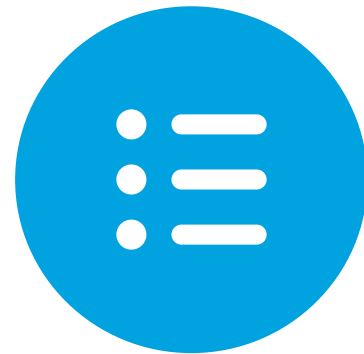
Describe an example network for image classification

.



## Objective

Explain the parameters defining the network



## Objective

Identify common tricks for improving classification performance

# Deep Learning for Image-based Recognition



| Visual recognition is an important part of human intelligence.

| ILSVRC (ImageNet Large-scale Visual Recognition Challenge) illustrates such a task.

| Many ImageNet images are difficult for conventional algorithms to classify.

# ImageNet.org Samples

## Golden retriever

An English breed having a long silky golden coat

1607  
pictures

64.99%  
Popularity  
Percentile

Wordnet  
IDs

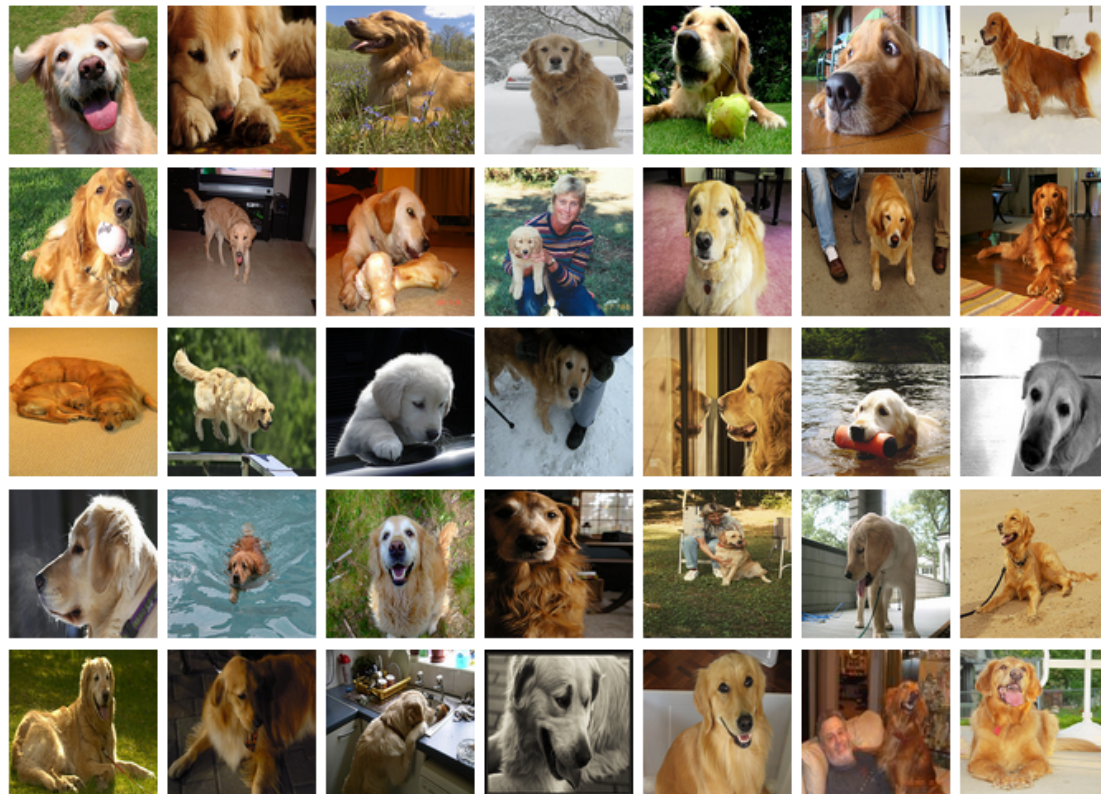
Numbers in brackets: (the number of synsets in the subtree).

- ImageNet 2011 Fall Release (32326)
  - plant, flora, plant life (4486)
  - geological formation, formation (1112)
  - natural object (1112)
  - sport, athletics (176)
  - artifact, artefact (10504)
  - fungus (308)
  - person, individual, someone, somebody (14198)
  - animal, animate being, beast, brute, creature, fauna (12217)
    - invertebrate (766)
    - homeotherm, homoiotherm, homeothermic (153)
    - work animal (4)
    - darer (0)
    - survivor (0)
    - range animal (0)
    - creepy-crawly (0)
    - domestic animal, domesticated (194)
      - domestic cat, house cat, Felis catus (37)
      - dog, domestic dog, Canis familiaris (1201)
        - pooch, doggie, doggy, booby (101)
        - hunting dog (101)
          - sporting dog, gun dog (101)
            - pointer, Spanish pointer (3)
            - setter (3)
            - bird dog (0)
            - spaniel (11)
            - griffon, wire-haired (0)
            - water dog (0)
            - retriever (5)
              - golden retriever (1607)

Treemap Visualization

Images of the Synset

Downloads



\*Images of children synsets are not included. All images shown are thumbnails. Images may be subject to copyright.

Prev 1 2 3 4 5 6 7 8 9 10 ... 67 68 Next

# Success Stories



# Success Stories: 2014 – Top Three

Rank	Team	Error
1	Google	0.06656
2	Oxford	0.07325
3	MSRA	0.08062

SOURCE: [ImageNet.org](http://ImageNet.org)

# Success Stories: 2015 – Top Three

Team Name	Entry Description	Description of Outside Data Used	Localization Error	Classification Error
Trimps-Soushen	Extra annotations collected by ourselves	Extra annotations collected by ourselves	0.122285	0.04581
Amax	Validate the classification model we used in DET Entry1	Share proposal procedure with DET for convenience	0.14574	0.04354
CUIImage	Average multiple models – validation accuracy is 79.78%	3000-class classification images from ImageNet are used to pre-train CNN	0.198272	0.05858

SOURCE: [ImageNet.org](http://ImageNet.org)



# Example Application 1: DR Detection

## | DR: Diabetic Retinopathy

| A recent work: Gulshan *et al.* "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs." *JAMA* 316.22 (2016): 2402-2410

- Employed large datasets
- A specific CNN architecture (Inception-v3) taking the entire image as input (as opposed to lesion/structure-specific CNNs)
- High performance:  
Comparable to a panel of 7  
board-certified  
ophthalmologists



# Example Application 2 – Visual Aesthetics



| While being subjective, computational modes are possible since there are patterns in visually-appealing pictures.

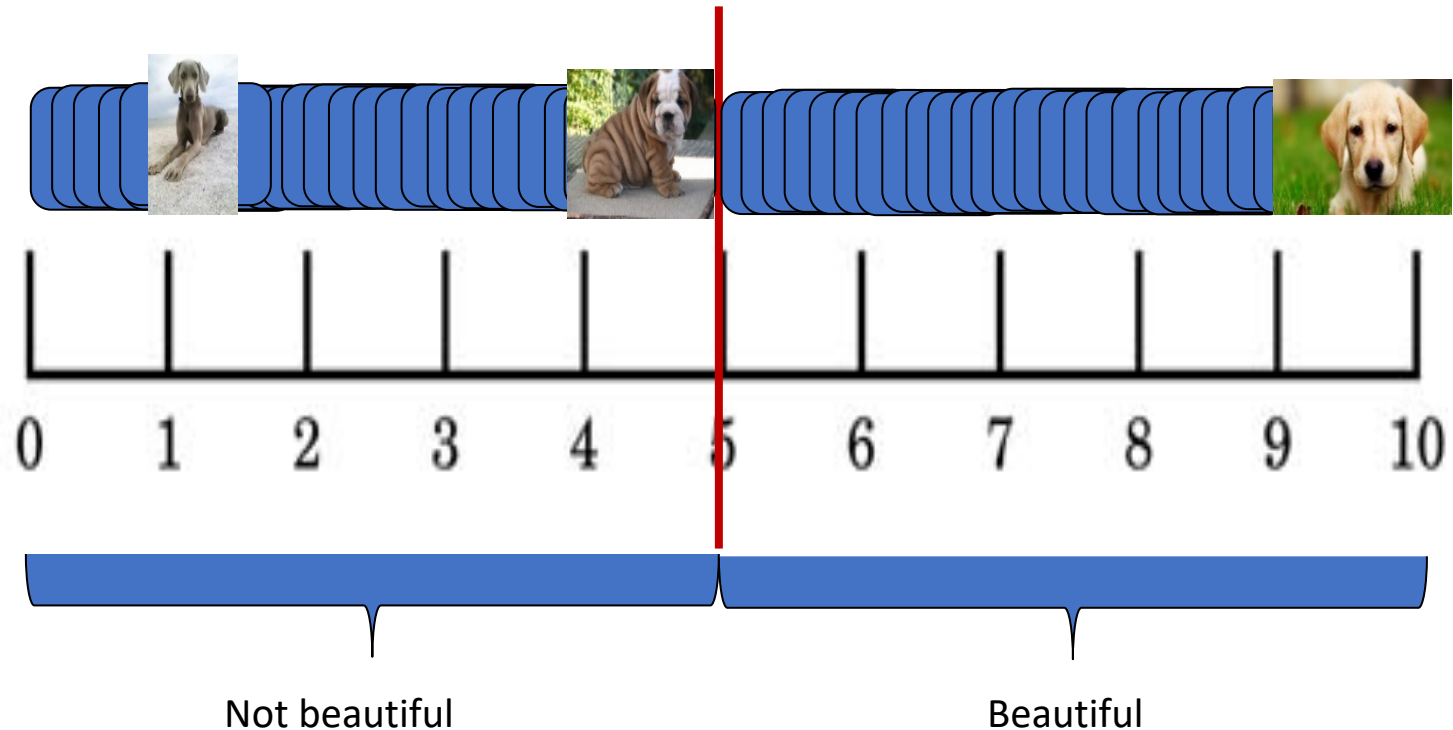
- E.g., photographic rules.

| Huge on-line datasets available. If ratings are also available, the problem becomes supervised learning.

- Conventional approaches still face the bottleneck of feature extraction.

# Related Approaches

| Solving the task as binary classification

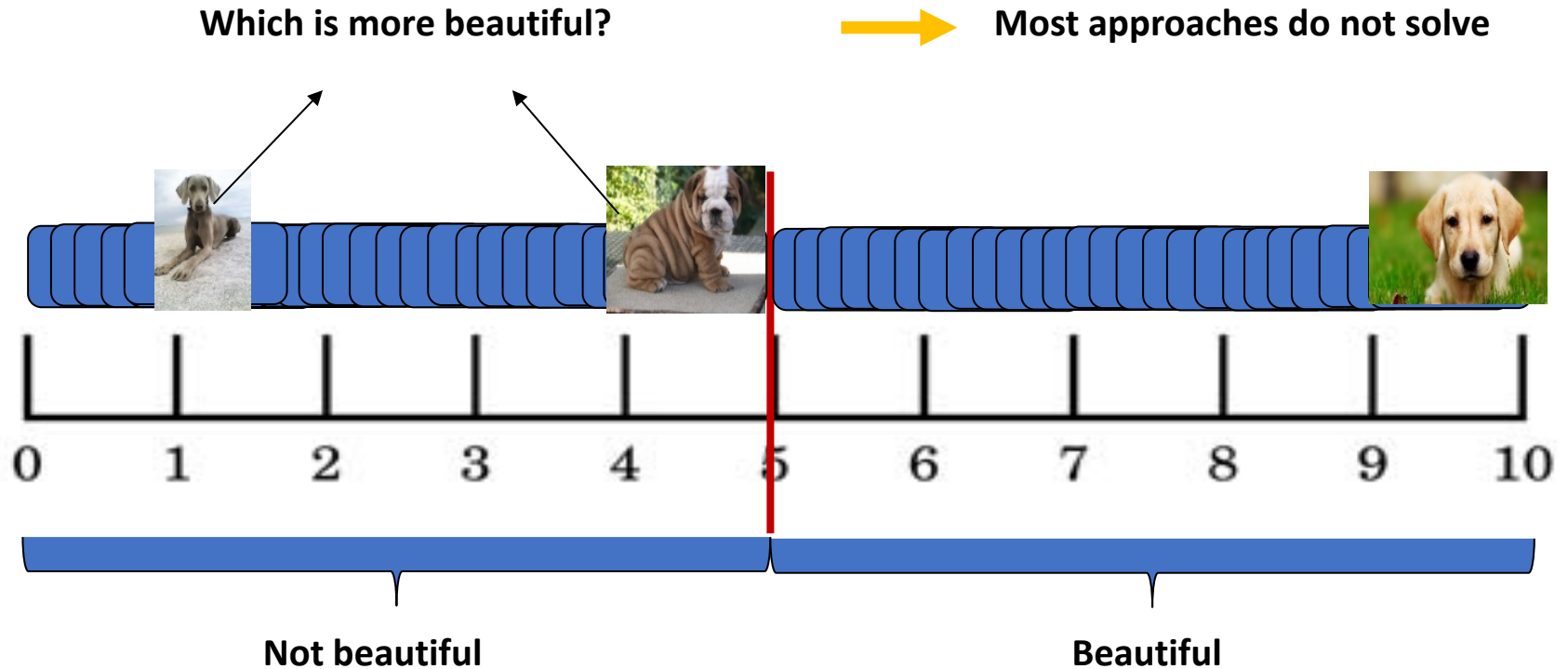


# Related Approaches: Examples



- | *RAPID: Rating Pictorial Aesthetics using Deep Learning* (Lu et al.)
- | *Deep Multi-Patch Aggregation Network for Image Style, Aesthetics, and Quality Estimation* (Lu et al.)
- | *Image Aesthetic Evaluation Using Paralleled Deep Convolution Neural Network* (Guo & Li)

# A New Task: Relative Aesthetics



| Image retrieval

| Image enhancement

# A Deep Learning Approach

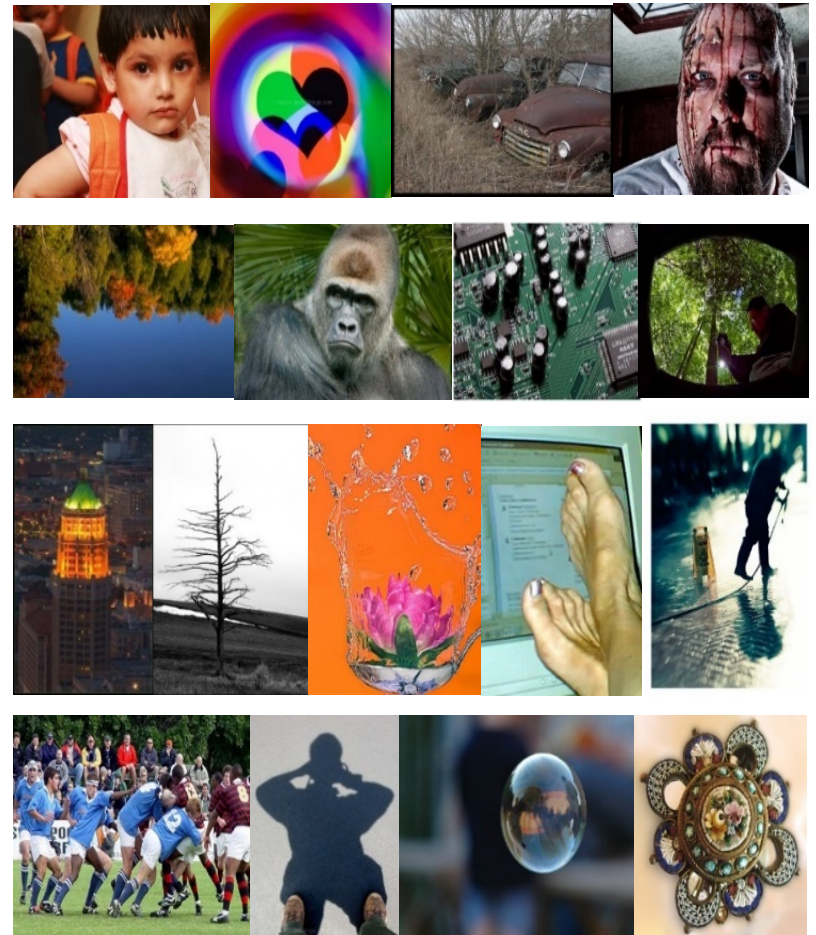


- | Dual-channelled CNN trained using relative learning
- | Siamese Network characteristics (weight sharing) and hinge-loss function
- | A custom data-set with relative labels – pairs formed based on aesthetic rating

SOURCE: Gattupalli *et al.* “A Computational Approach to Relative Aesthetics”, *International Conference on Pattern Recognition (ICPR)* 2016.

# Constructing a Useful Data Set – 1/2

- | Total of 250,000 images extracted from [dpchallenge.com](http://dpchallenge.com)
- | Challenges under which users post their submission
- | Peers rate and a final winner is selected based on the average rating
- | Belong to a wide variety of semantic categories



SOURCE: AVA Dataset

# Constructing a Useful Dataset – 2/2

| The minimum gap between the average rating of the two images is one

–e.g., 3.4 and 4.5, 6.3, and 7.8

| The maximum variance allowed between the ratings of different voters is 2.6

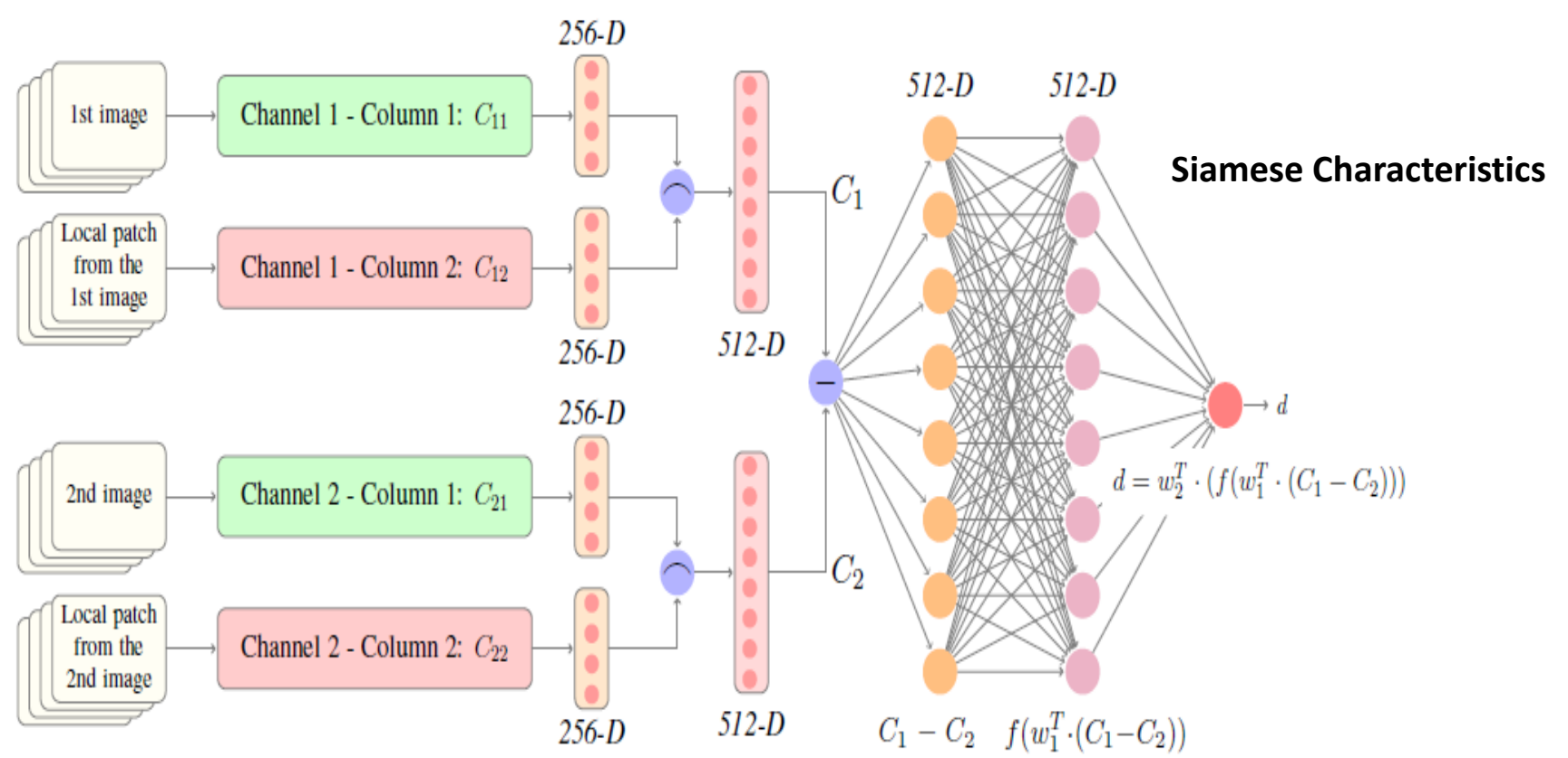
| Pick pairs from the same category only

–e.g., cannot compare an image of a car and a building





# The Network Architecture & Other Characteristics



Padded Input	Conv	Max-pooling	Conv	Max-pooling	Conv	Conv	Dropout	Dense	Dropout	Dense	Dropout
$3 \times 230 \times 230$	2, 64, 11, 2	$2 \times 2$	1, 64, 5, 1	$2 \times 2$	1, 64, 3, 1	—, 64, 3, 1	0.5	1000	0.5	256	0.5

# Further Implementation Details



| Each channel contains two streams of processing: column 1 for global, and column 2 for local

| Global Patch

- e.g., rule of thirds, golden ration

| Local Patch

- e.g., smoothness/graininess

# The Loss Function

$$L = \max(0, \delta - y \cdot d(I_1, I_2)) \quad \longrightarrow \quad \text{Hinge Loss}$$

$$d(I_1, I_2) = f(C_1 - C_2)$$

| where,

$y$  = True label of the image pair,

i.e., 1 if  $I_1 > I_2$  and

-1 otherwise

$C_1, C_2$  = Outputs of channel 1 and channel 2 respectively

# Sample Results



## | Two ways of training

- Using binary labels
- Using relative labels

## | Tested for two tasks

- For Binary Classification task
- For Ranking task

# Eight Experiments Total

	Ranking (custom test-set)	Ranking (standard test-set)	Classification (custom test-set)	Classification (standard test-set)
Base-line	62.21	65.87	<b>59.92</b>	69.18
Relative aesthetics	<b>70.51</b>	<b>76.77</b>	59.41	<b>71.60</b>