Image Classification

Anna Lindahl
LT2318 Artificial Intelligence: cognitive systems
Gothenburg University

(Adapted from Wafia Adouane HT18)

Introduction

- Simple definition: grouping images into semantic classes based on some features.
- (Also: identifying patterns in images)
- Usage:
  - Robotics, together with NLP, information retrieval etc.
- Challenges:
  - Image quality
  - Definitions (What is a chair?)
  - Bias in data/annotations

Image features

- NLP vs CV features
- Visual features
  - Color, size, center, orientation etc.
  - Invariant to transformations
- Lexical features/Semantic classes
  - Labels, context
- Learned vs. pre-engineered features
  - We need to choose how to represent an image

Representing images & extracting features

- Color
  - RGB
  - CMYK
  - HSV (Hue, saturation, value)
- Structure - based on image gradients
  - SIFT(Scale invariant image features)
  - SURF (Speeded up robust features)
  - HOG (Histogram of oriented gradients)

SIFT

- SIFT: scale invariant image transformation
- Applied to grayscale images
- Used for feature matching
- Keypoints - points of interest, each keypoint has a descriptor, needed to match keypoints between images.
- Each keypoint has a 128-dimensional feature vector.

Clustering features

- “Visual bag of words”
- Clustering over all descriptors in images to build a “dictionary” of visual words, features.
- For example: k-nearest neighbours cluster similar features in n clusters. Input vector in classification/comparison is vector with integer counts of the number of each clusters present in the image.
- Slides with description of bag of visual words

Feature extraction

Notebook

Pre-trained image models

- Used for both extracting features and classifying images
- Most models trained on ImageNet http://www.image-net.org/
- Some models:
  - Vgg16 (image)
  - Vgg19
  - Resnet (50) Example in Keras
- All models available in Keras: https://keras.io/applications/
- What do the models learn?
  - Adversarial images

Examples of misclassifications

- http://www.evolvingai.org/files/DNNsEasilyFooled_cvpr15.pdf
- https://spectrum.ieee.org/cars-that-think/transportation/sensors/slight-street-sign-modifications-can-fool-machine-learning-algorithms
- https://www.theverge.com/2019/4/23/18512472/fool-ai-surveillance-adversarial-example-yolov2-person-detection
- https://algorithm.data61.csiro.au/why-were-vaccinating-algorithms-against-adversarial-attacks/

Image classification with Keras

Notebook

Guess who?

- In the game “Guess who?” players describe a face for the other player to guess.
- Idea: Ground visual features of faces in language.
  - Classify facial attributes and use them for the game.
- Dataset: Labeled faces in the wild (Huang et al.)
  - 13233 images of faces of 5749 persons
  - Faces annotated/classified with 73 different attributes from Kumar et al.

OpenFace

- Pretrained models originally for facial recognition.
- Uses dlib and OpenCV to preprocess images.

Demos and more:
http://cmusatyalab.github.io/openface/

OpenFace 2:
https://github.com/TadasBaltrusaitis/OpenFace

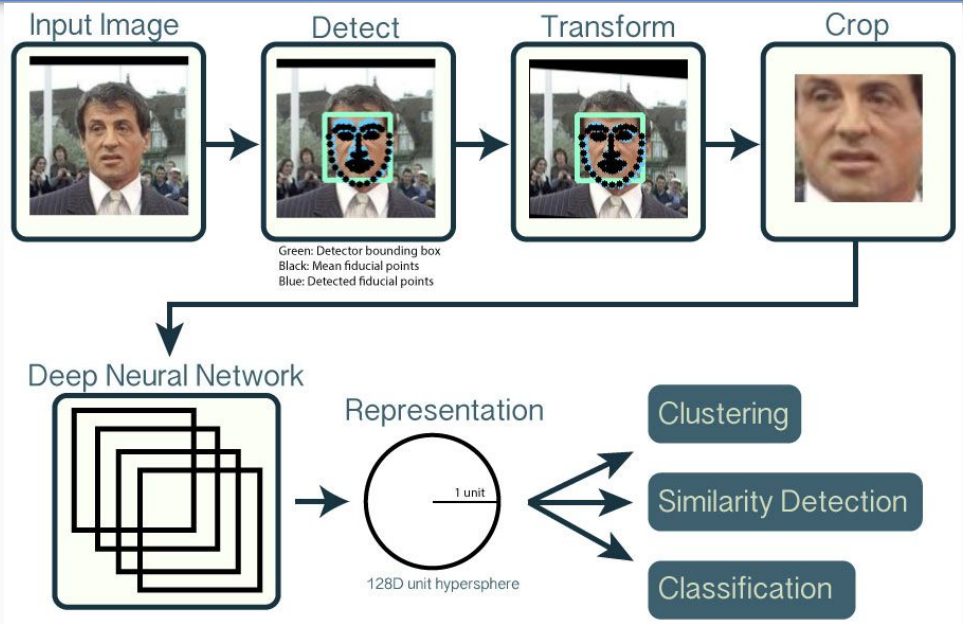


Image from https://cmusatyalab.github.io/openface/

128-dimensional embeddings + Attributes from Kumar

73 binary classifiers, presence of each attribute

Faces classified with attributes based on image data

OpenFace

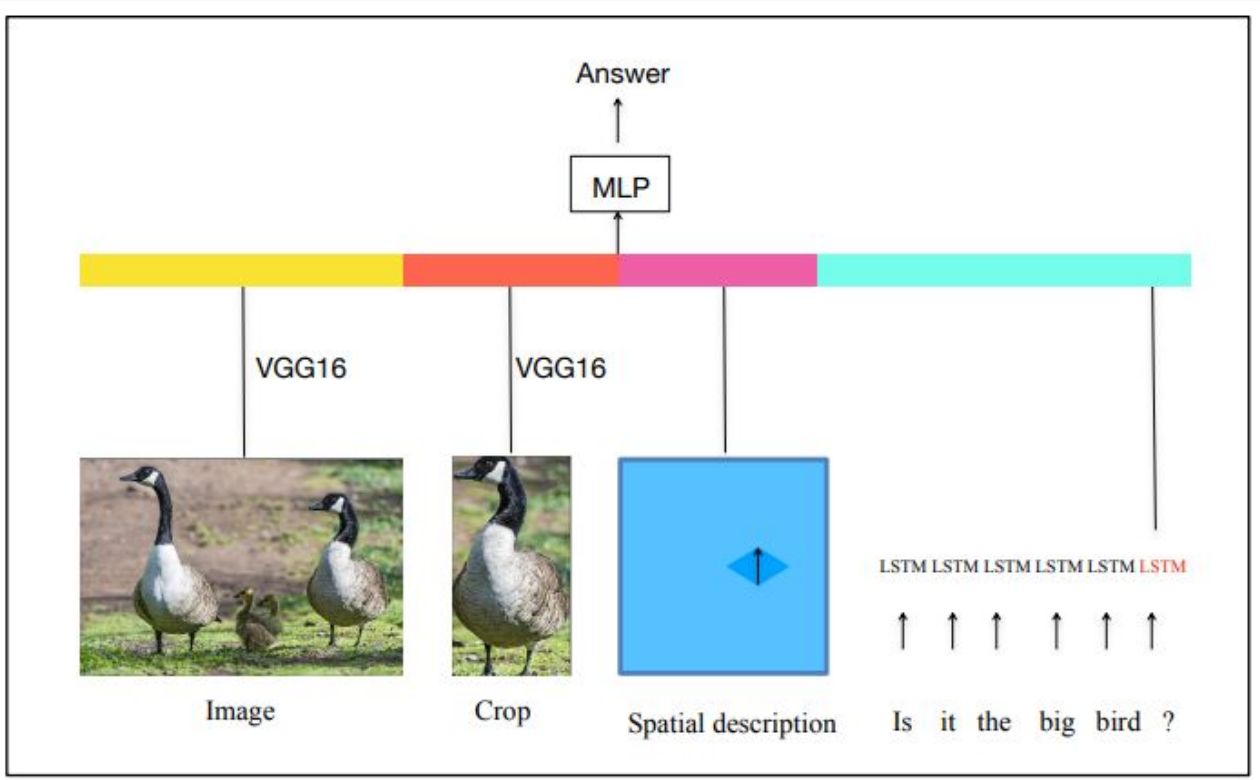
Game demo

Guess what?

- Choose one object in the image,
  the other person guesses which one it is.
- Training data: image, identified object
  and accompanying dialogue.

Questioner	***Oracle***
Is it the baby animal ?	*Yes*
Is it the smallest animal in picture ?	*No*
Is it covered by an another animal ?	*No*

Guess what?


Answer
MLP
VGG16
VGG16
LSTM LSTM LSTM LSTM LSTM LSTM
Image
Crop
Spatial description
Is it the big bird ?

Resources

OpenCv tutorials: https://docs.opencv.org/master/d9/df8/tutorial_root.html

Hog explanation: https://www.learnopencv.com/histogram-of-oriented-gradients/

Keras image classification:
https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html

References

Huang, G. B., Ramesh, M., Berg, T., & Learned- Miller, E. (2007). Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. Technical Report 07-49, University of Massachusetts, Amherst.
Link

Kumar, N., Berg, A. C., Belhumeur, P. N., & Nayar, S. K. (2009). Attribute and simile classifiers for face verification. In 2009 IEEE 12th International Conference on Computer Vision (pp.365–372).: IEEE. Link

Amos, B., Ludwiczuk, B., & Satyanarayanan, M. (2016). OpenFace: A general-purpose face recognition library with mobile applications. Technical report, CMU-CS-16-118, CMU School of Computer Science
Link