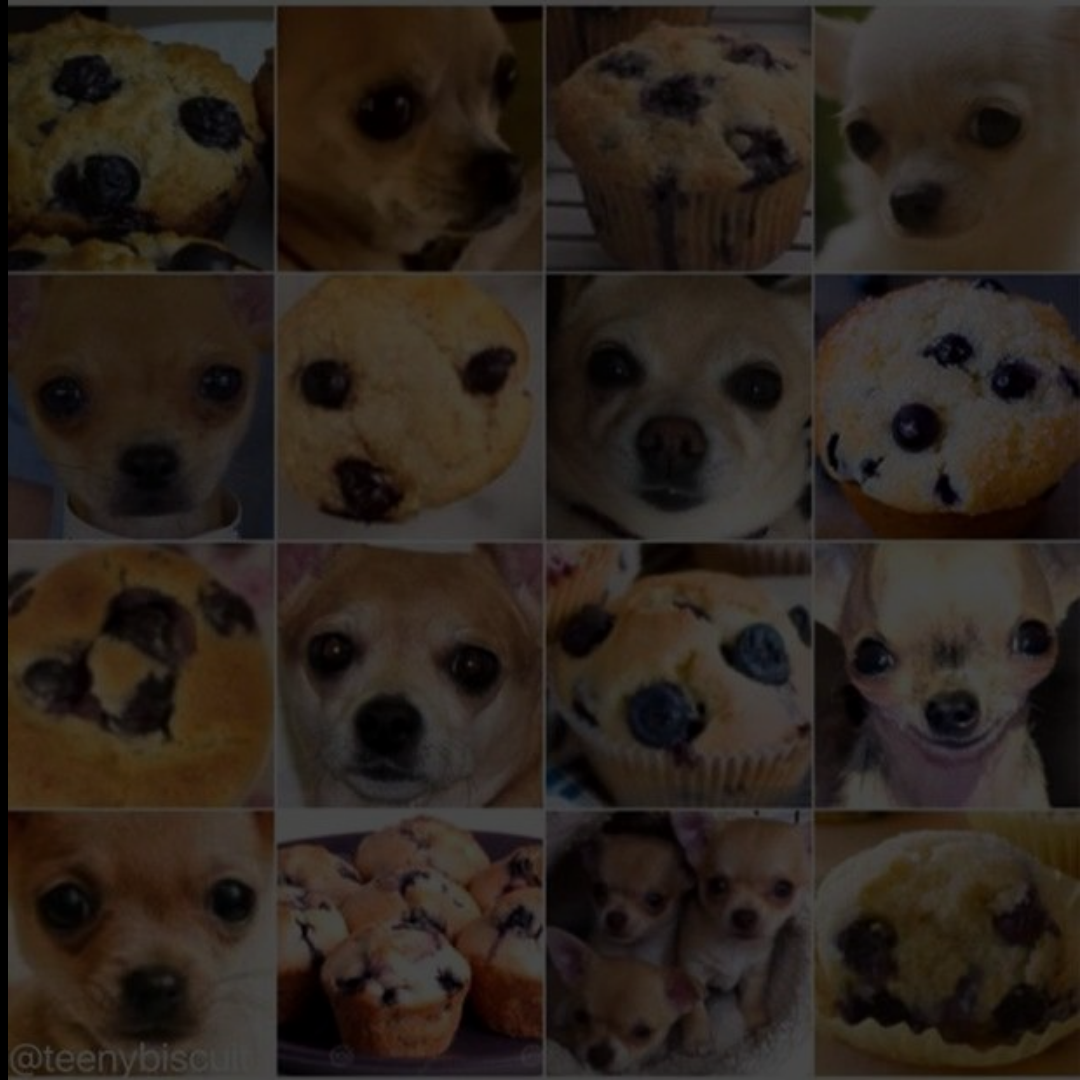
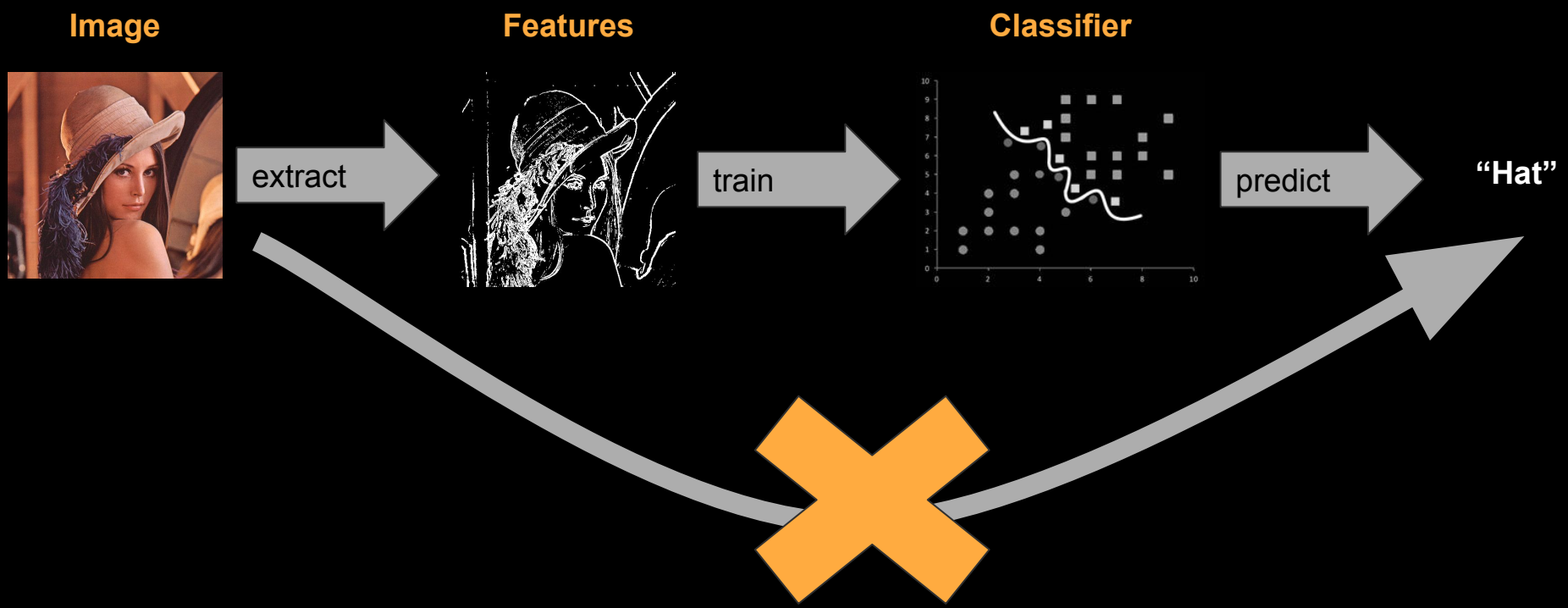
Computer Vision now and then

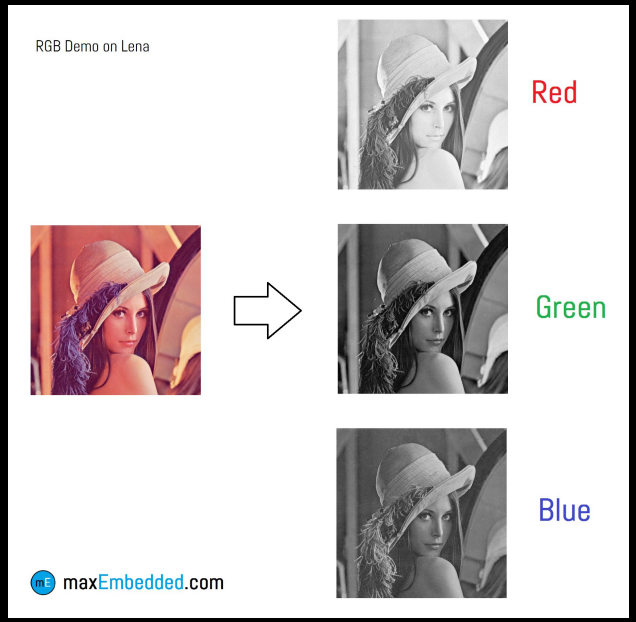
From counting pixels to distinguishing Chihuahuas from Muffins

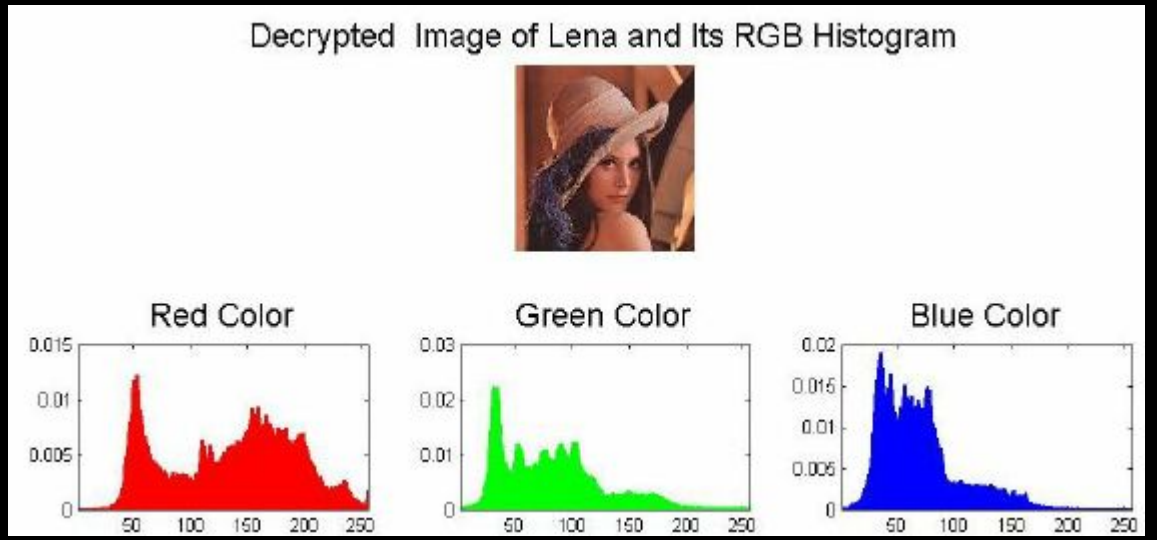


Visual Computing
from a traditional Machine Learning Perspective



Color Histograms
Counting Pixels

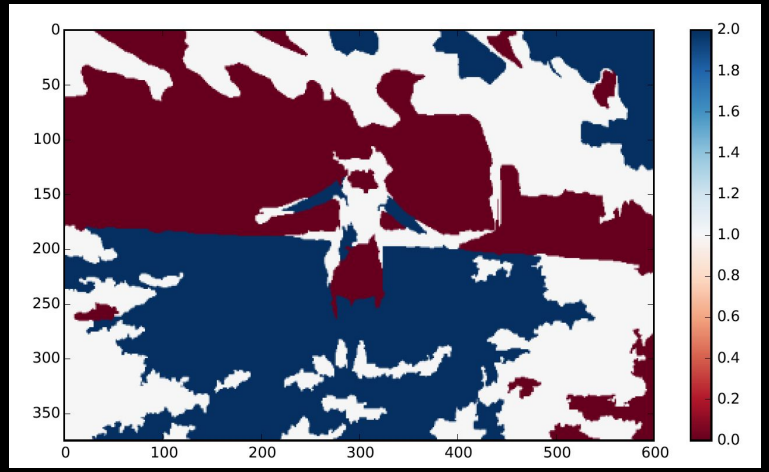


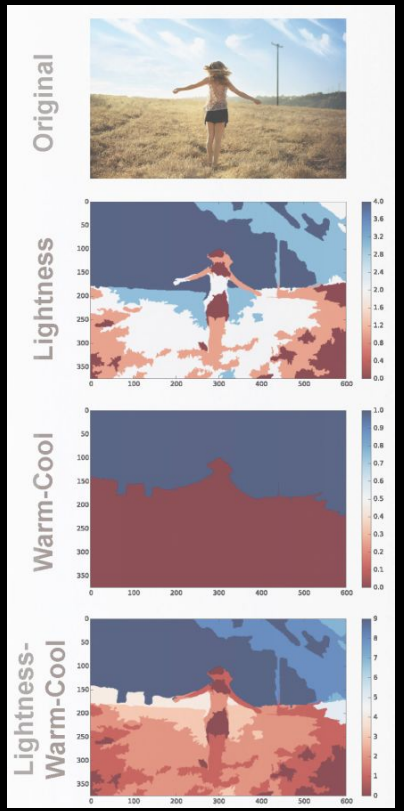


Affective Color Features

Counting Pixels smarter

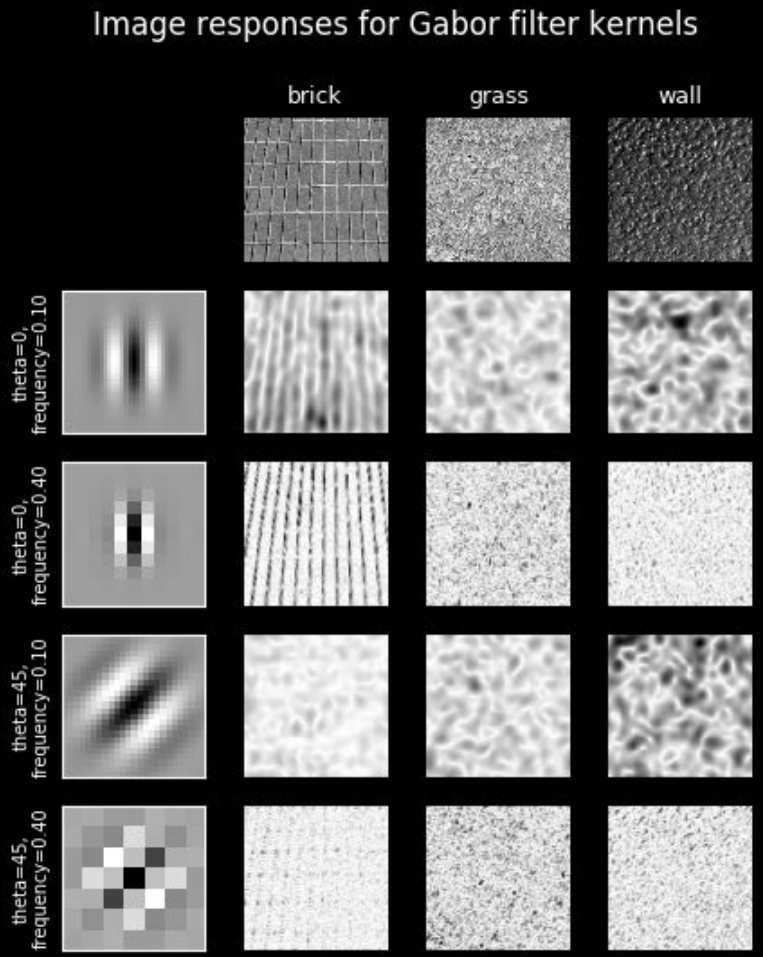
- Contrasts
  - Warm/Cool
  - Light/Dark
  - Colorfullness





Alexander Schindler and Andreas Rauber. **An audio-visual approach to music genre classification through affective color features**. In Proceedings of the 37th European Conference on Information Retrieval (ECIR'15), Vienna, Austria, March 29 - April 02 2015.

Filter
Reducing Information

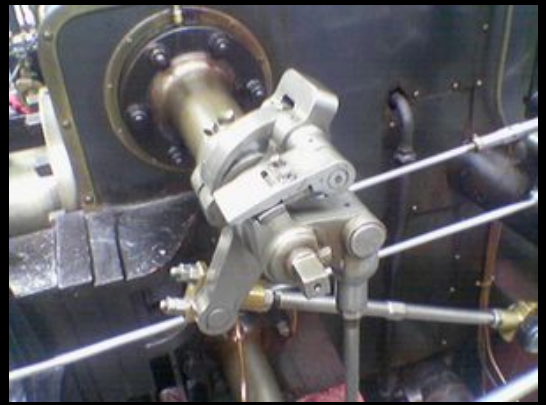


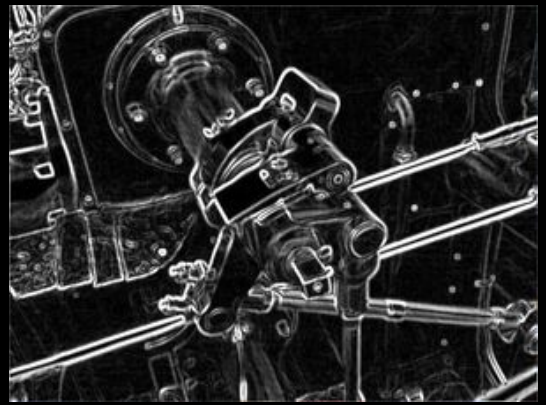
Edge Detectors

Kernel based

	0°			45°		
Sobel	-1	0	1	0	1	2
	-2	0	2	-1	0	1
	-1	0	1	-2	-1	0
Kirsch	-3	-3	5	-3	5	5
	-3	0	5	-3	0	5
	-3	-3	5	-3	-3	-3
Robinson	-1	0	1	0	1	1
	-1	0	1	-1	0	1
	-1	0	1	-1	-1	0

$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

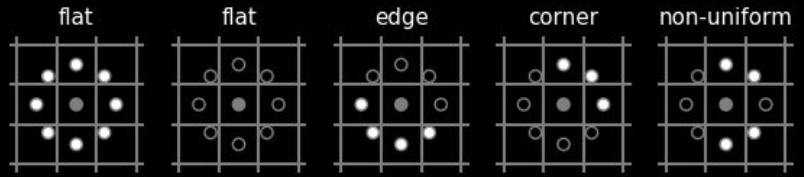


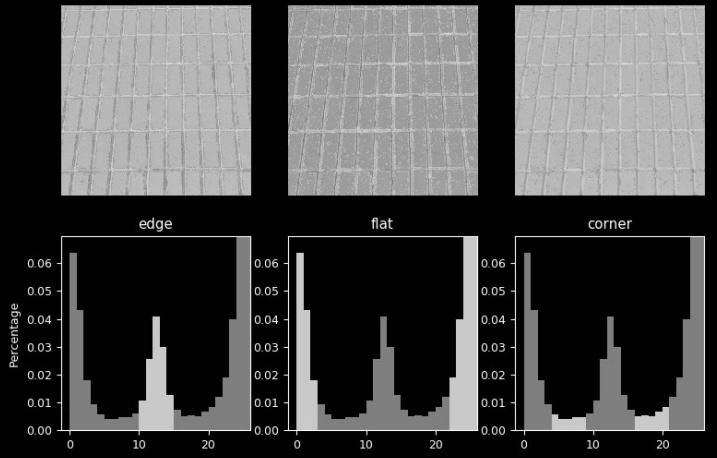


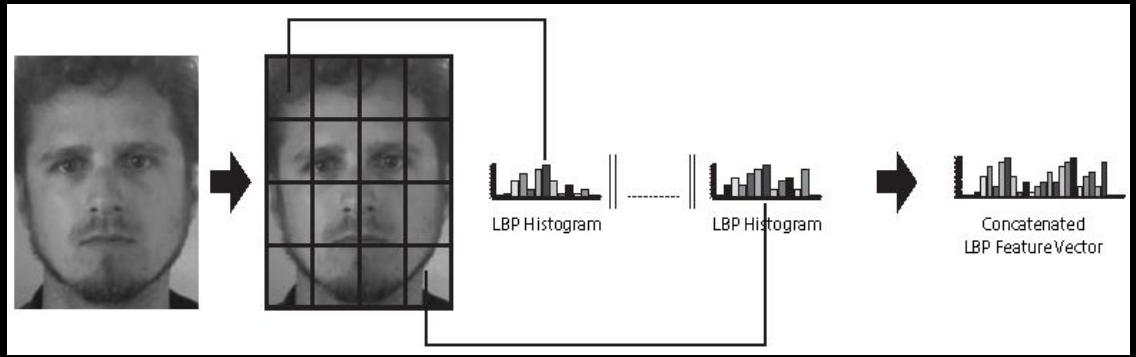
Texture

Local Binary Patterns (LBP)

Face Detection



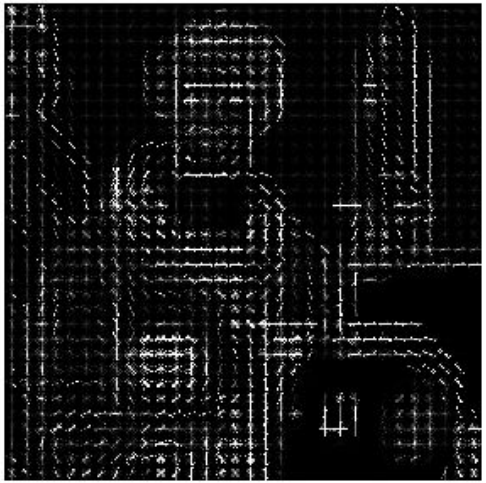




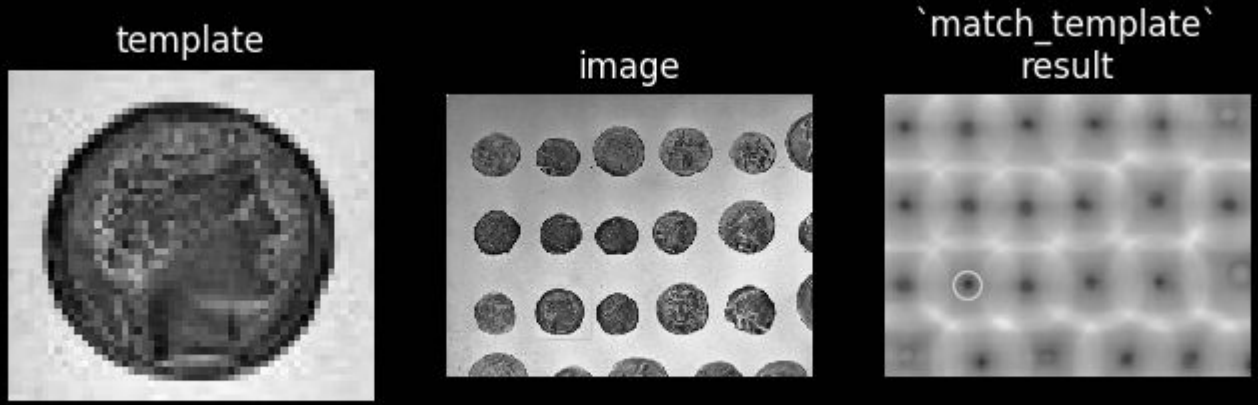
Alexander Schindler and Andreas Rauber. **A music video information retrieval approach to artist identification**. In *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR2013) to appear*, Marseille, France, October 14-18 2013.

Object Detection

Histogram of Oriented Gradients (HOG)



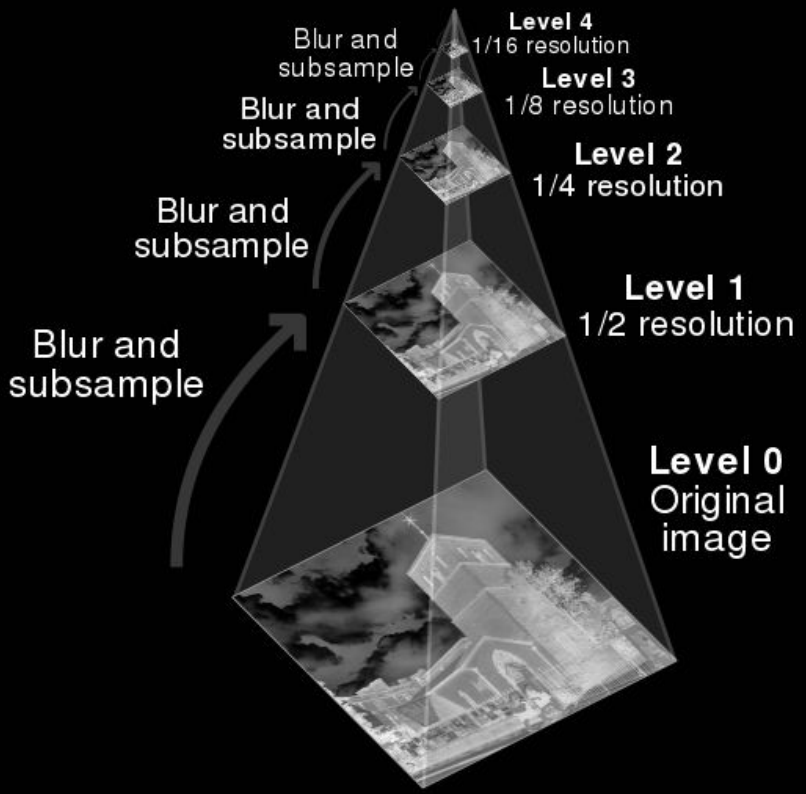
Object Detection
Template Matching



Scale Invariance

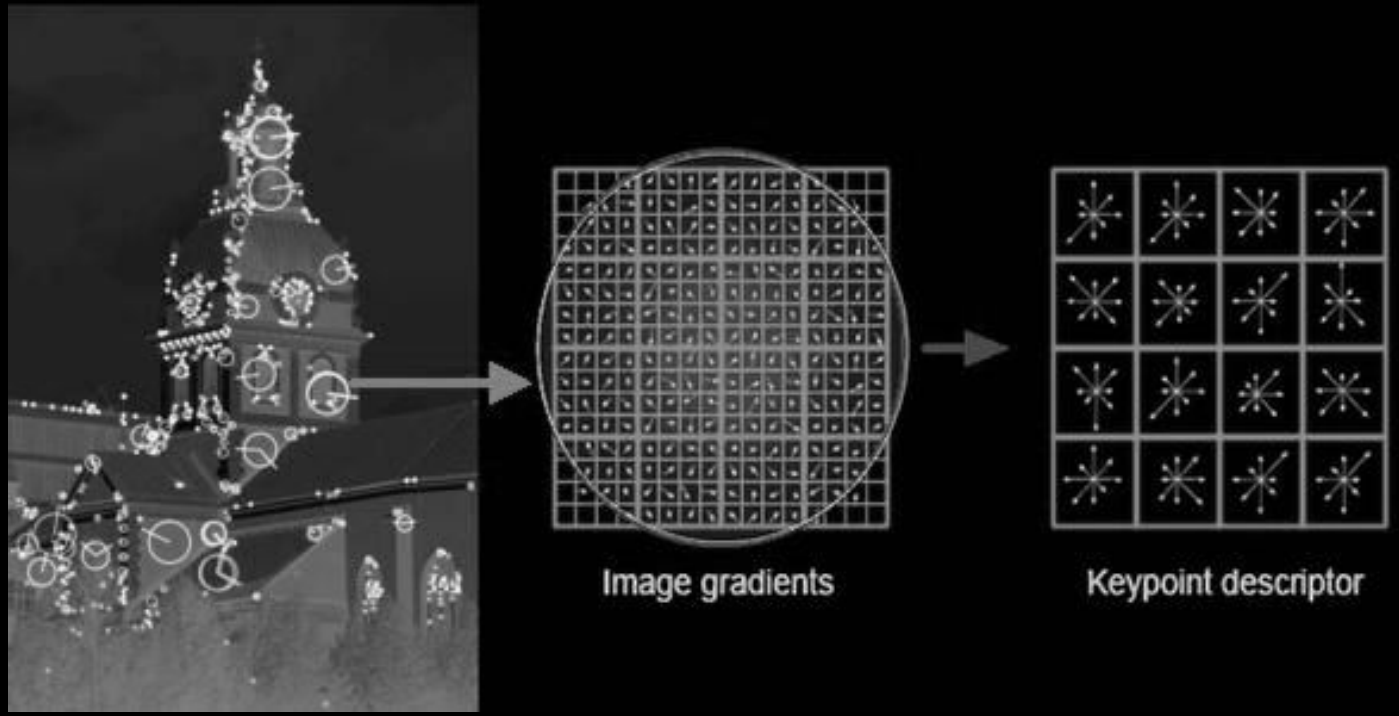
Image Pyramids

Downsample image subsequently

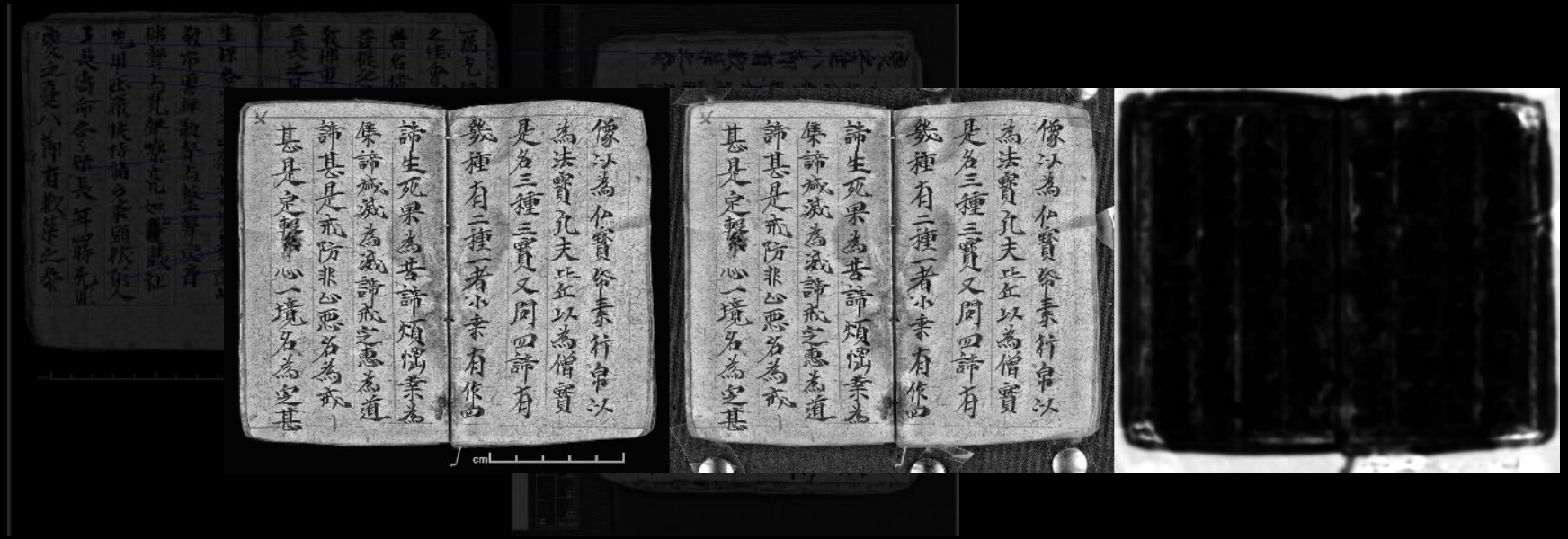


Object Detection => SIFT

Scale Invariant Feature Transforms



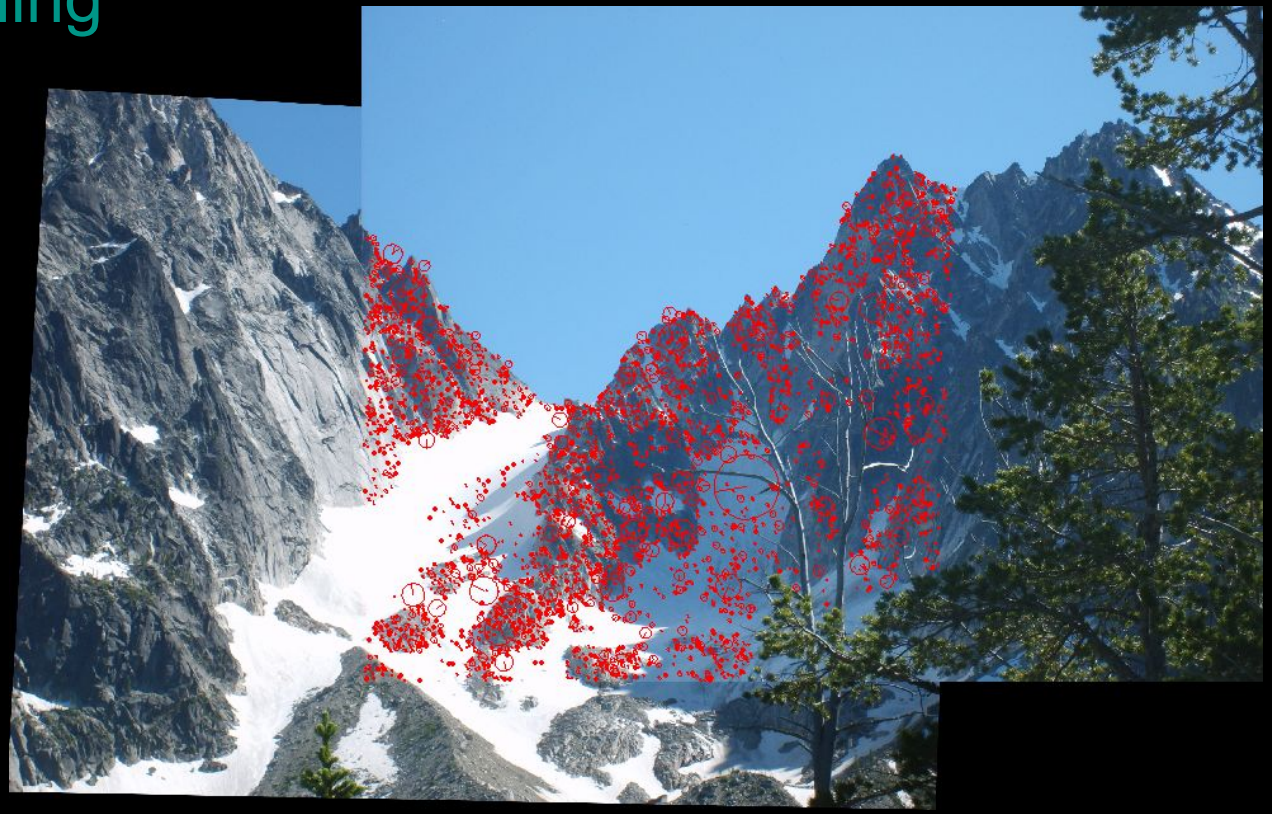
SIFT
Image Registration



Reinhold Huber-Moerk and Alexander Schindler. **Quality assurance for document image collections in digital preservation.** In *Proceedings of the 14th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS 2012)*, Lecture Notes in Computer Science, Brno, Czech Republic, September 4-7 2012. Springer.

SIFT

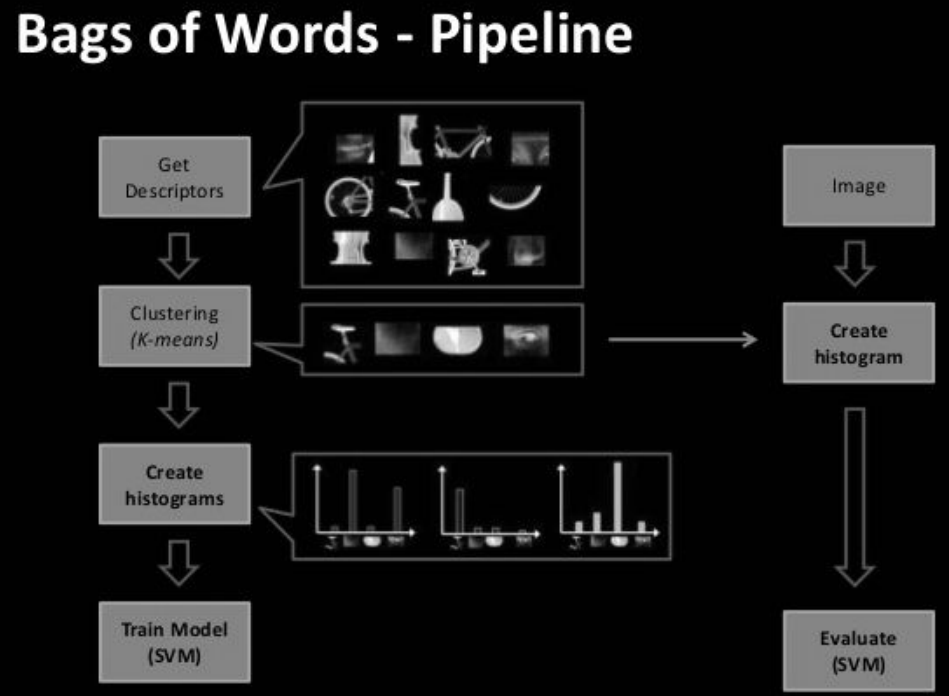
Panorama Stitching



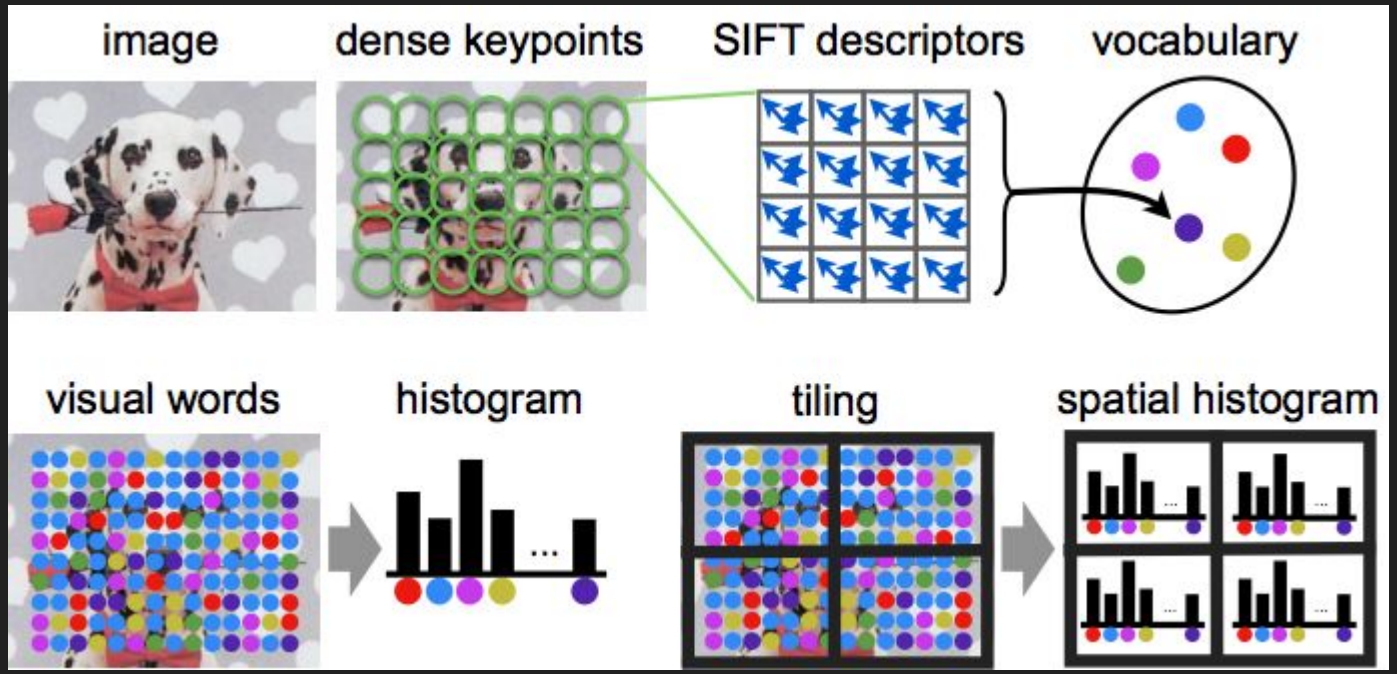
Object Detection with SIFT

Bag of Visual Words

Pre - Deep Learning
State-of-the-art
in Object Detection



SIFT - Bag of Visual Words approach

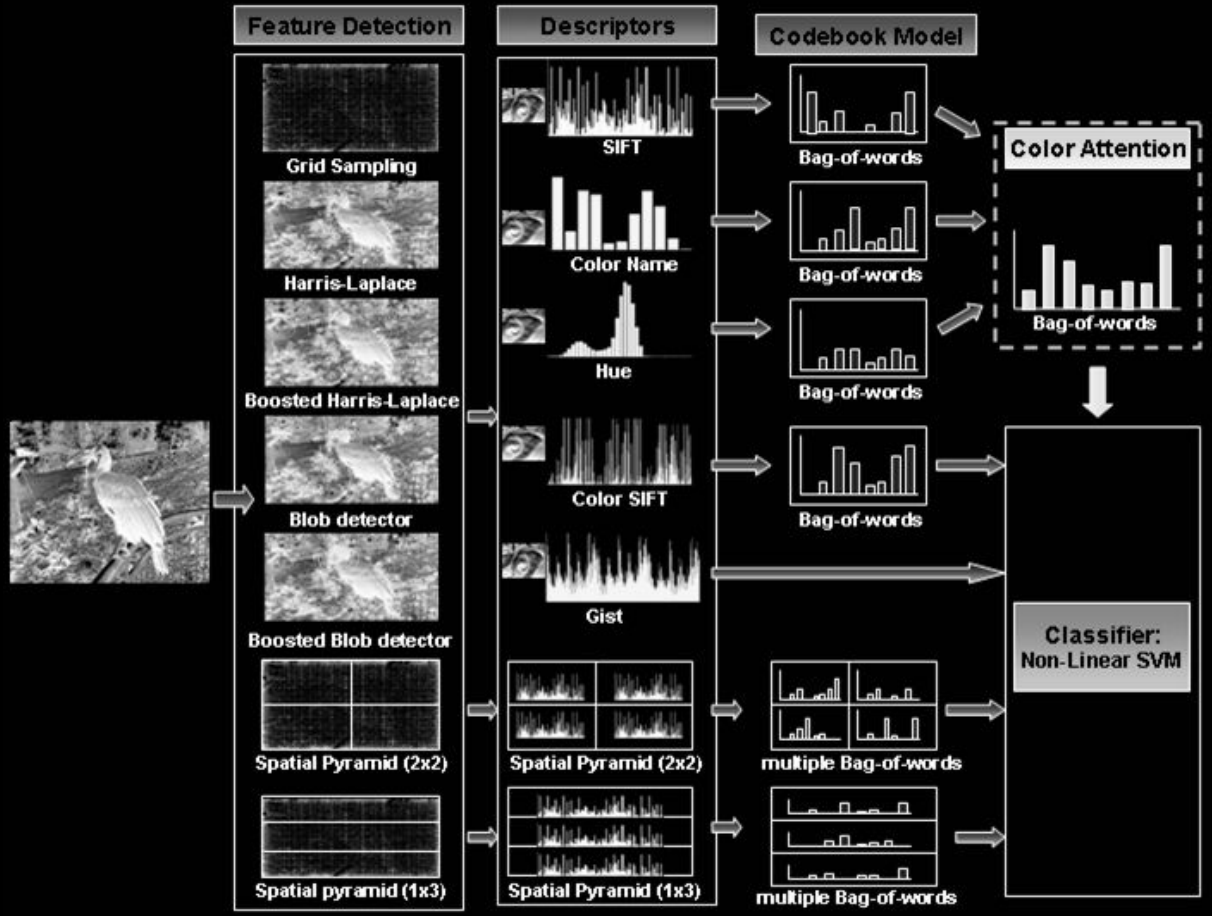


Ahuja, Sarthak, and Anchita Goel. "Scene Recognition using Bag-of-Words." Forest 100: 228.

Feature Composition

Complex Object recognition Approaches

- Early/Late fusion
- Ensemble Classifiers

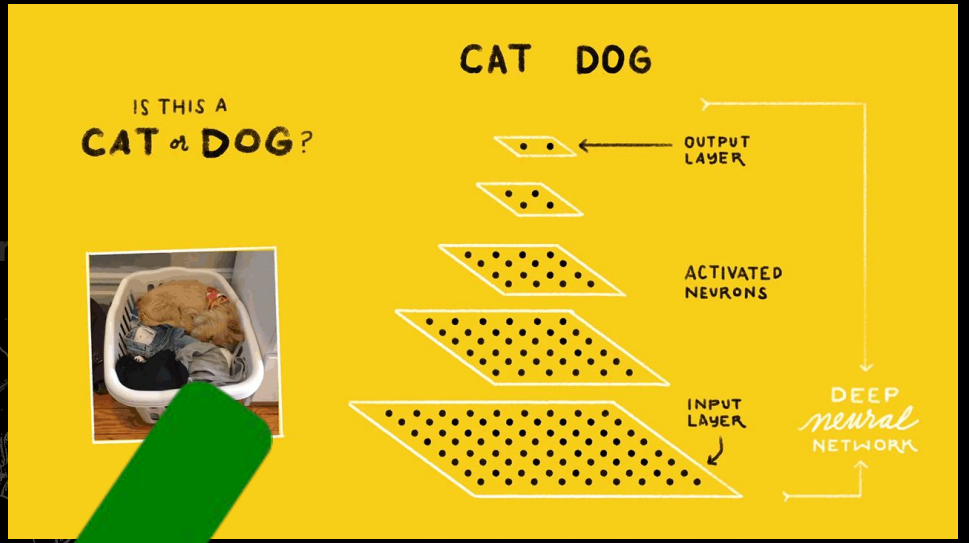


Deep Learning

Predicting input data

Image

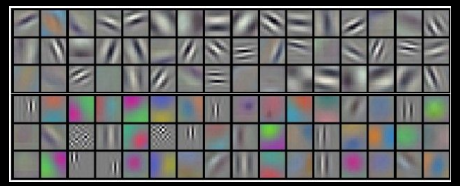
Featur

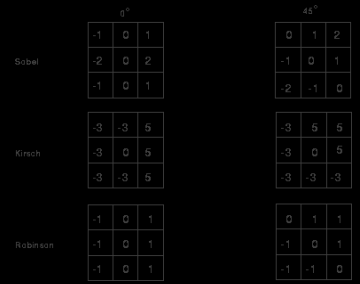


Deep Learning

Convolution Layer Properties

Filtering



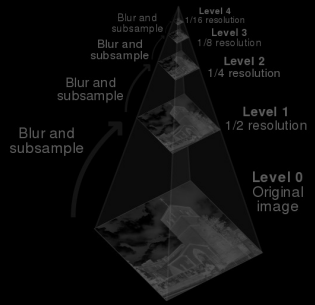


Scaling

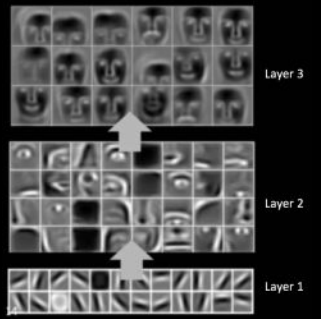
12	20	30	0
8	12	2	0
34	70	37	4
112	100	25	12

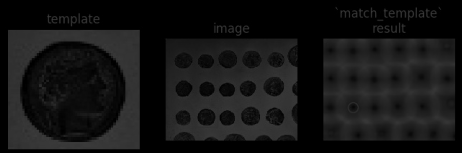
2×2 Max-Pool

20	30
112	37



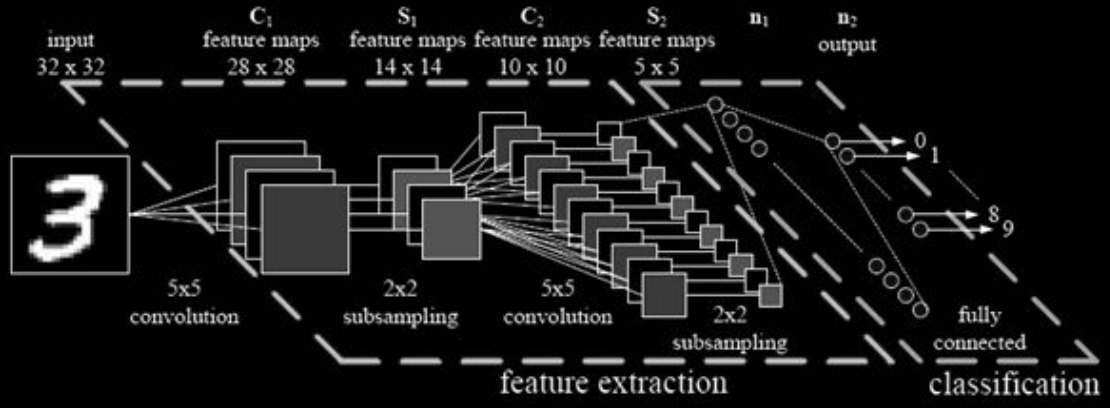
Matching





Deep Learning
Advantages

- Resembles many approved traditional methods
- Simplifies the processing chain (implicit feature extraction)
- Simplifies Multi-label Classification
- Commonly higher accuracies



Thank You
for your attention!

Alexander Schindler
alexander.schindler@ait.ac.at