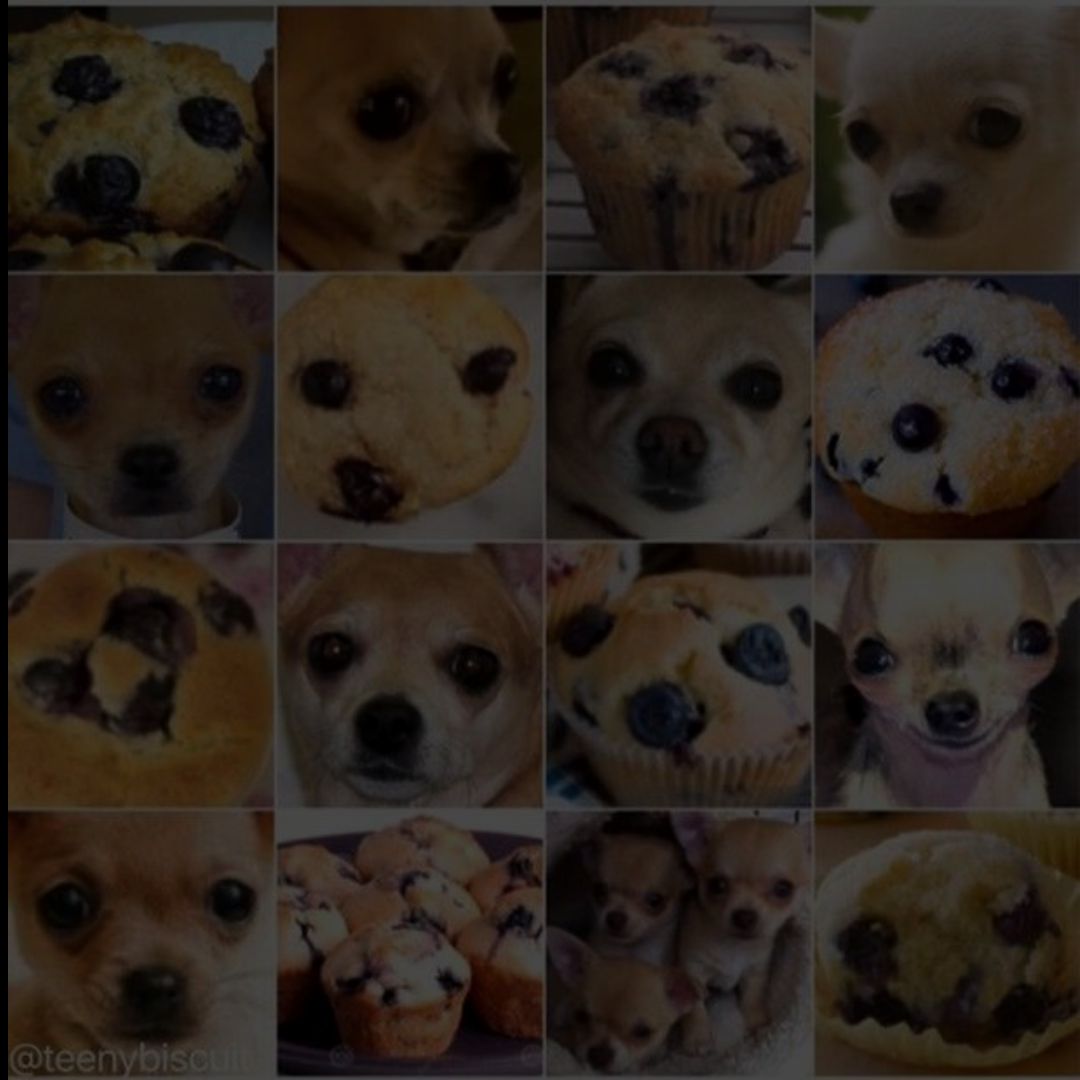


Computer Vision now and then

From counting pixels
to distinguishing
Chihuahuas from Muffins



Visual Computing

from a traditional Machine Learning Perspective

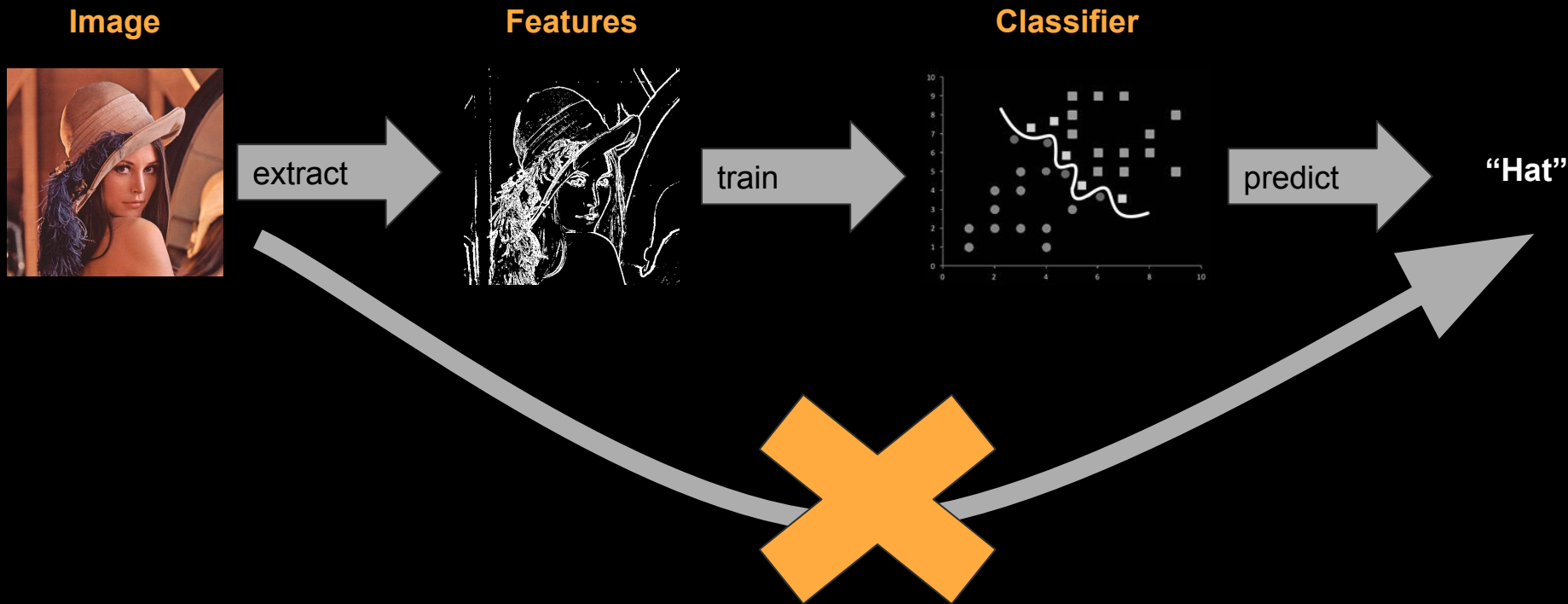


Image Processing Features

Counting Pixels



$300 \times 300 = 0.09$ Mega-Pixels



Small Image



$2048 \times 2048 = 4.2$ Mega-Pixels



Large Image

Color Histograms

Counting Pixels

in a given Color Space



=

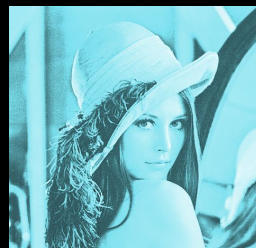
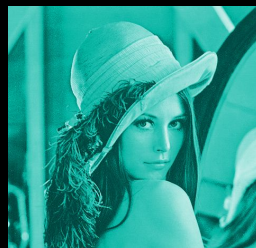
Red

+

Green

+

Blue



RGB Color Images

Colors by number

		x →				
		0	1	2	3	4
y ↓	0	0	1	2	3	4
1	5	6	7	8	9	
2	10	11	12	13	14	
3	15	16	17	18	19	
4	20	21	22	23	24	
		...				

- **Discrete intensity values**
 - **0** ... dark
 - **255** ... bright
- **Color Image**
 - **Combination of Layers**
 - **Red + Green + Blue**
- **Color Histogram**
 - **Frequencies of color value (e.g. 255) per color channel**
 - **Normalize to number of pixels per image**

Color Histograms

Counting Pixels

RGB Demo on Lena



Red



Green

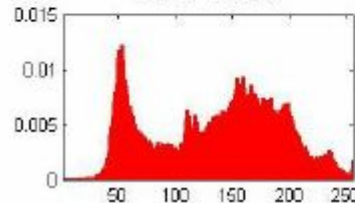


Blue

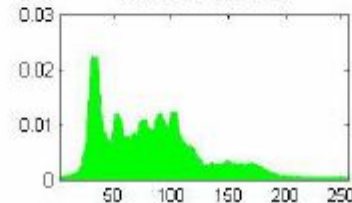
Decrypted Image of Lena and Its RGB Histogram



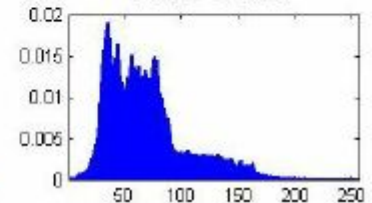
Red Color



Green Color



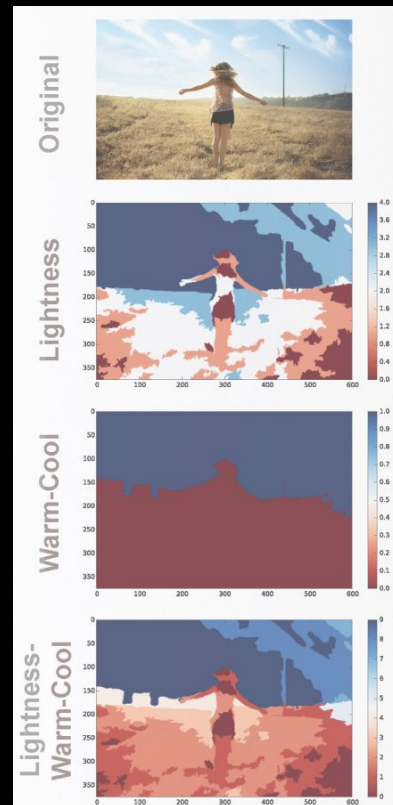
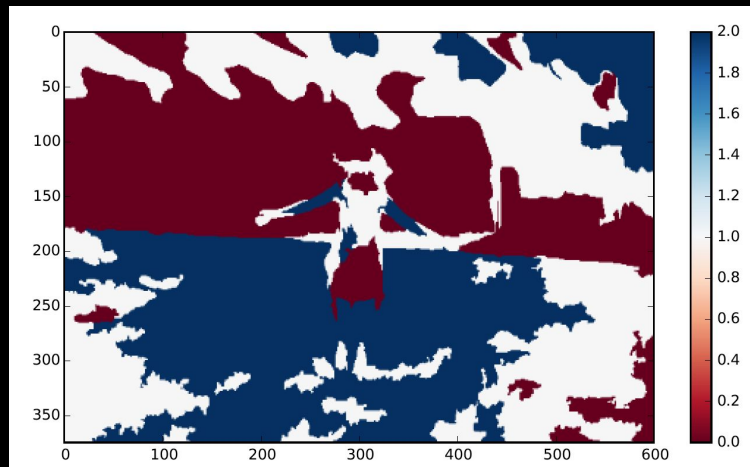
Blue Color



Affective Color Features

Counting Pixels smarter

- Contrasts
 - Warm/Cool
 - Light/Dark
 - Colorfulness

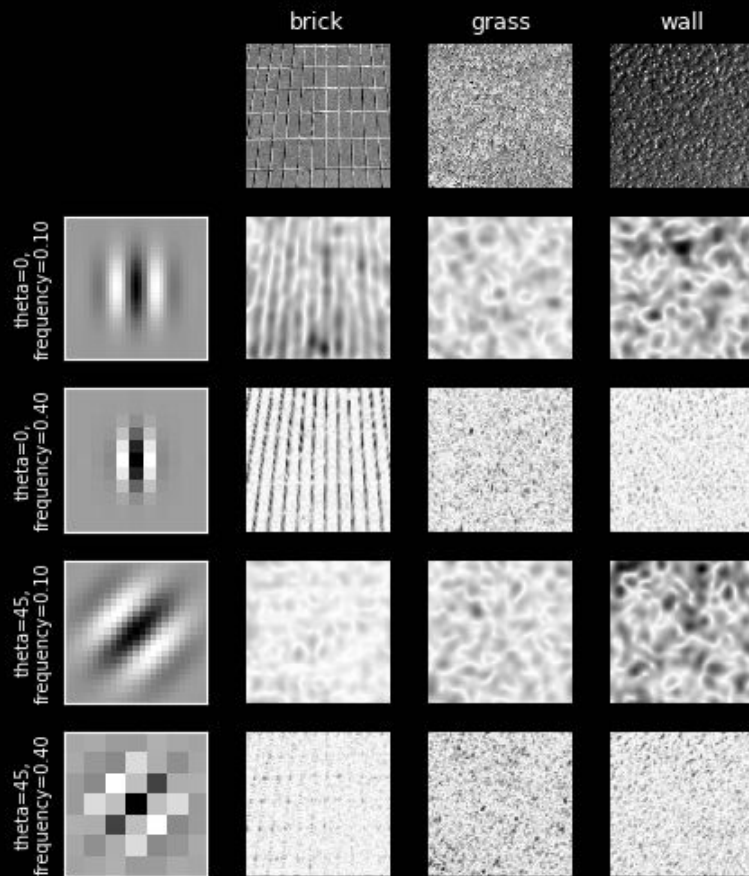


Alexander Schindler and Andreas Rauber. **An audio-visual approach to music genre classification through affective color features.** In Proceedings of the 37th European Conference on Information Retrieval (ECIR'15), Vienna, Austria, March 29 - April 02 2015.

Filter

Reducing Information

Image responses for Gabor filter kernels



Edge Detectors

Kernel based

Sobel

0°		
-1	0	1
-2	0	2
-1	0	1

45°		
0	1	2
-1	0	1
-2	-1	0

Kirsch

-3	-3	5
-3	0	5
-3	-3	5

-3	5	5
-3	0	5
-3	-3	-3

Robinson

-1	0	1
-1	0	1
-1	0	1

0	1	1
-1	0	1
-1	-1	0

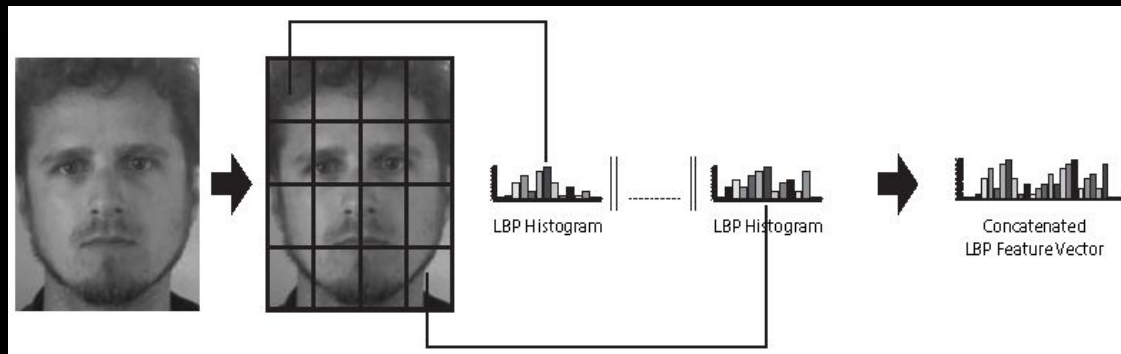
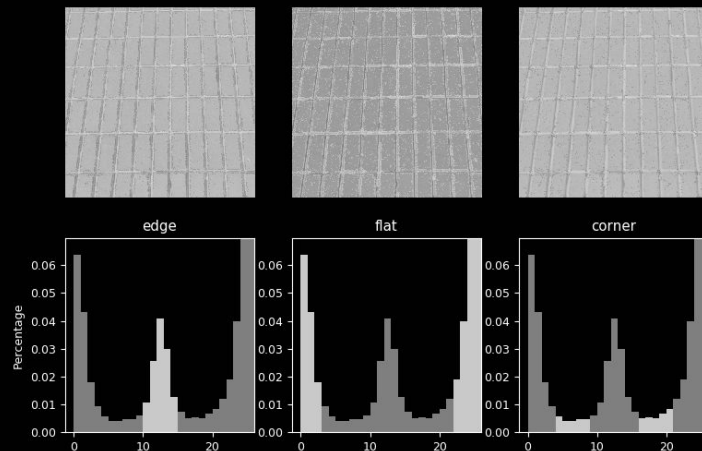
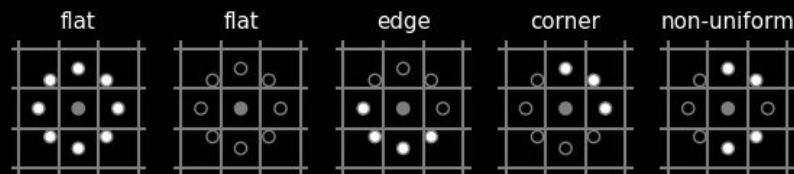
$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$



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.

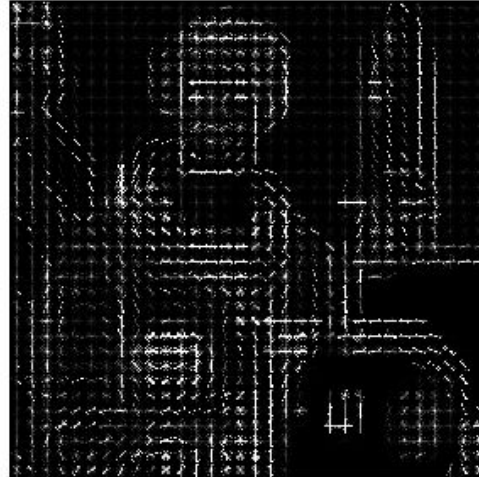
Object Detection

Histogram of Oriented Gradients (HOG)

Input image



Histogram of Oriented Gradients



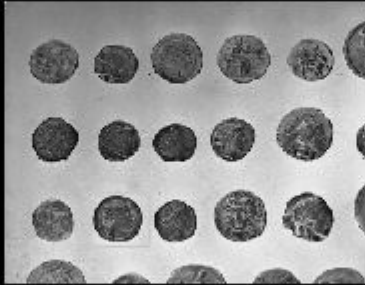
Object Detection

Template Matching

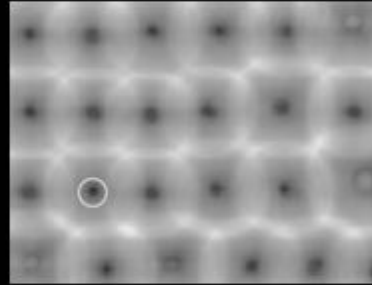
template



image



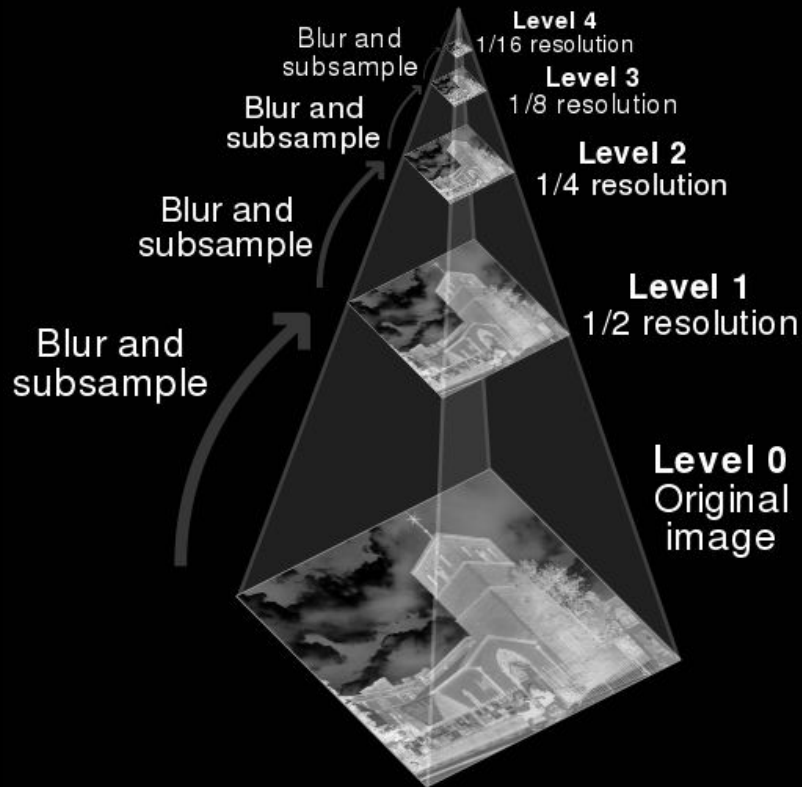
`match_template`
result



Scale Invariance

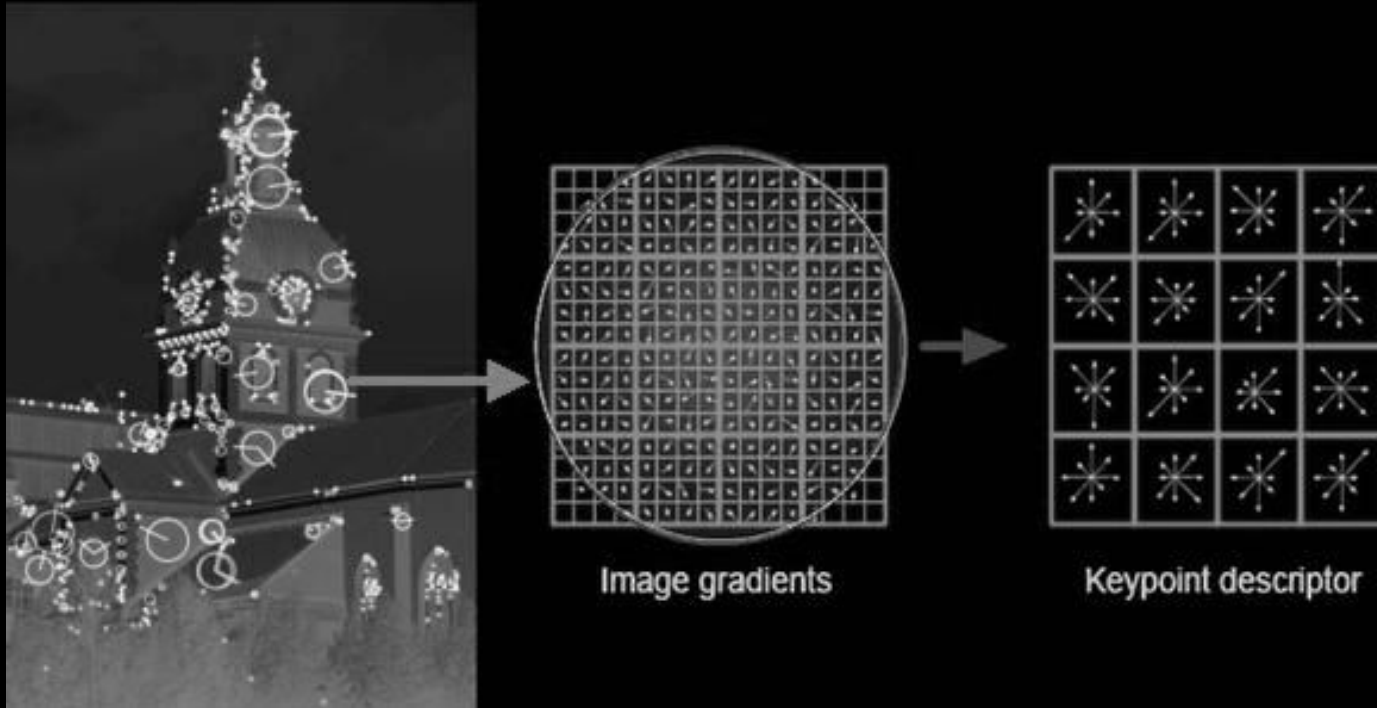
Image Pyramids

Downsample image subsequently



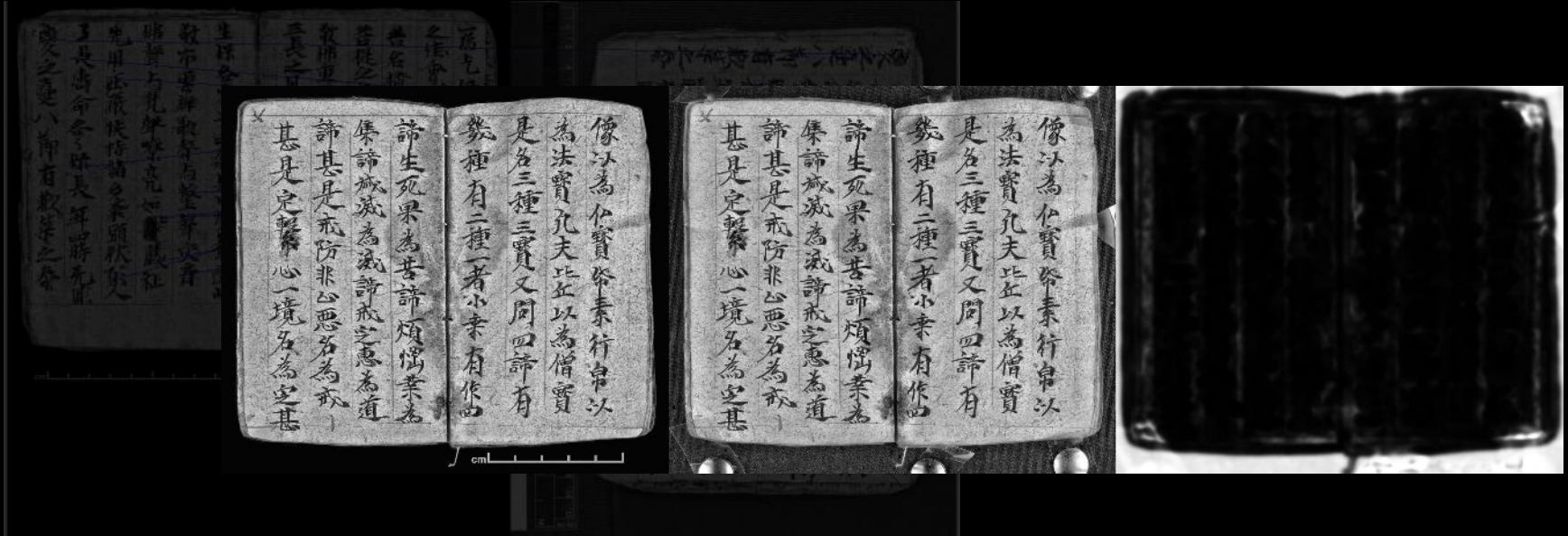
Object Detection => SIFT

Scale Invariant Feature Transforms



SIFT

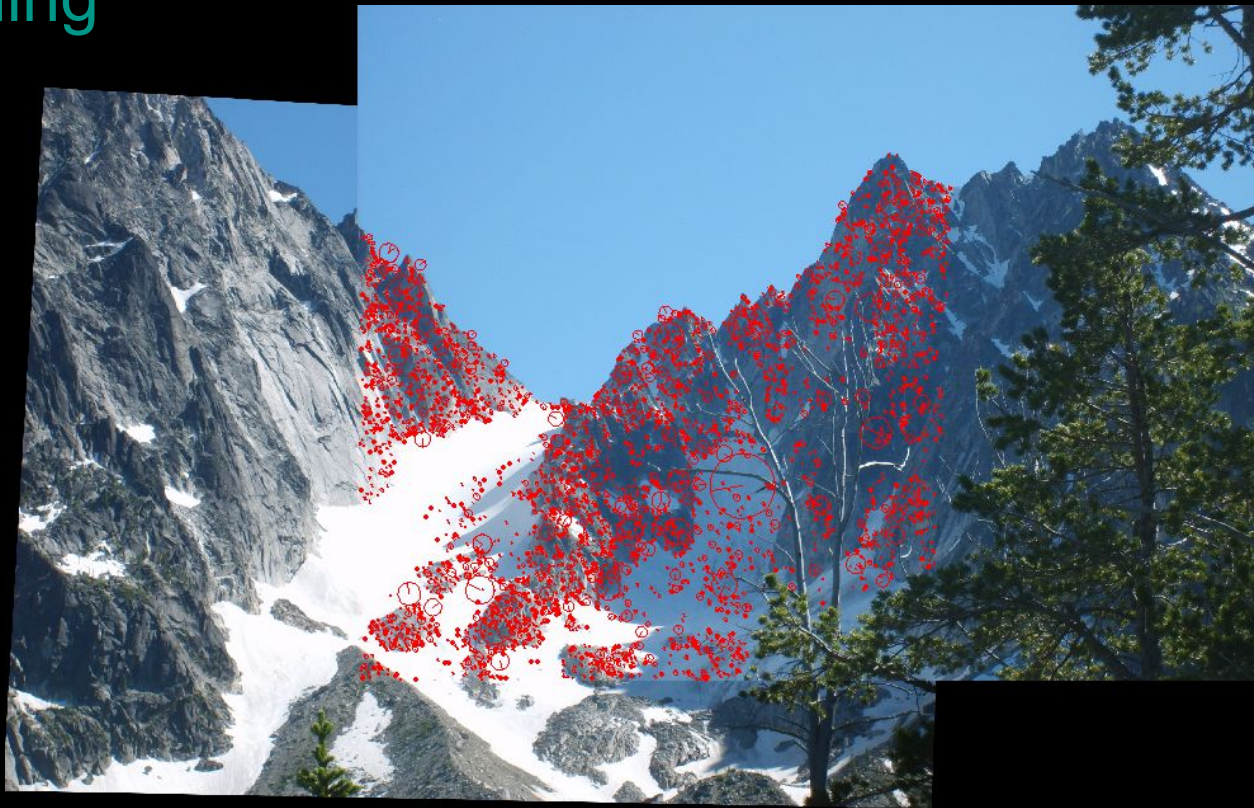
Image Registration



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

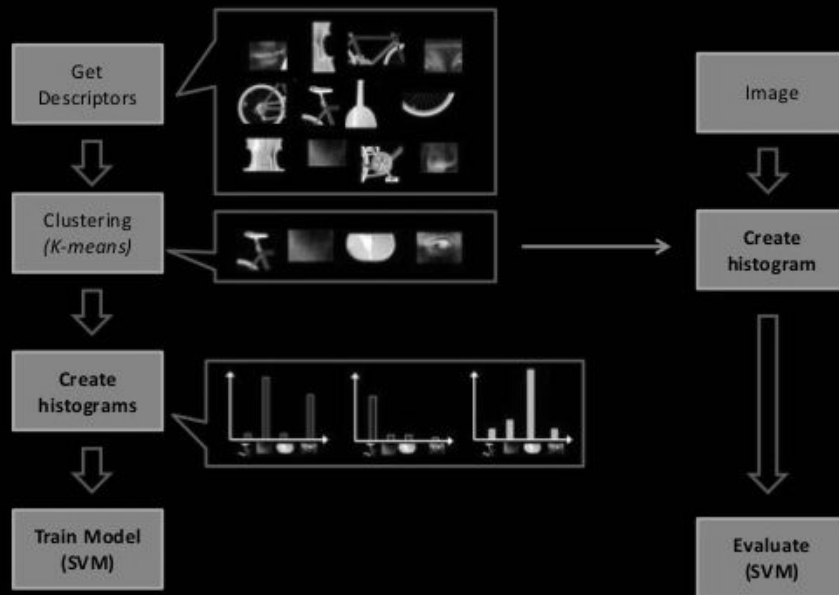


Object Detection with SIFT

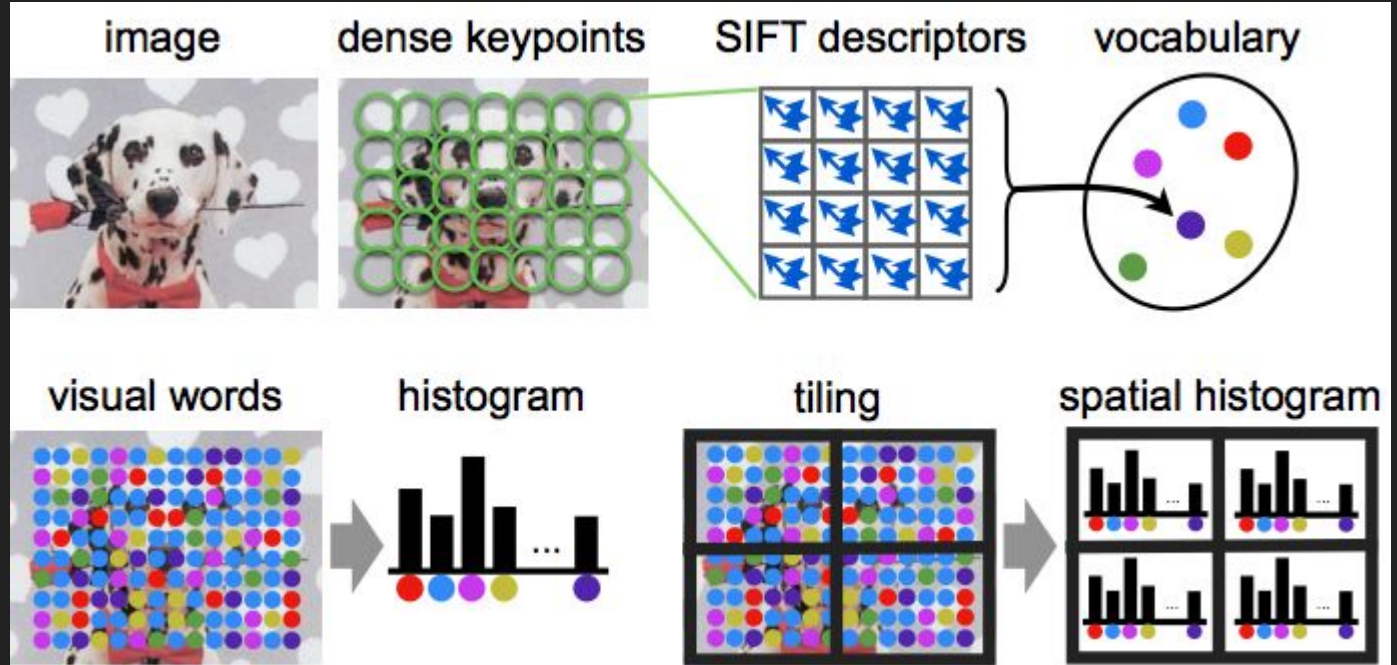
Bag of Visual Words

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

Bags of Words - Pipeline



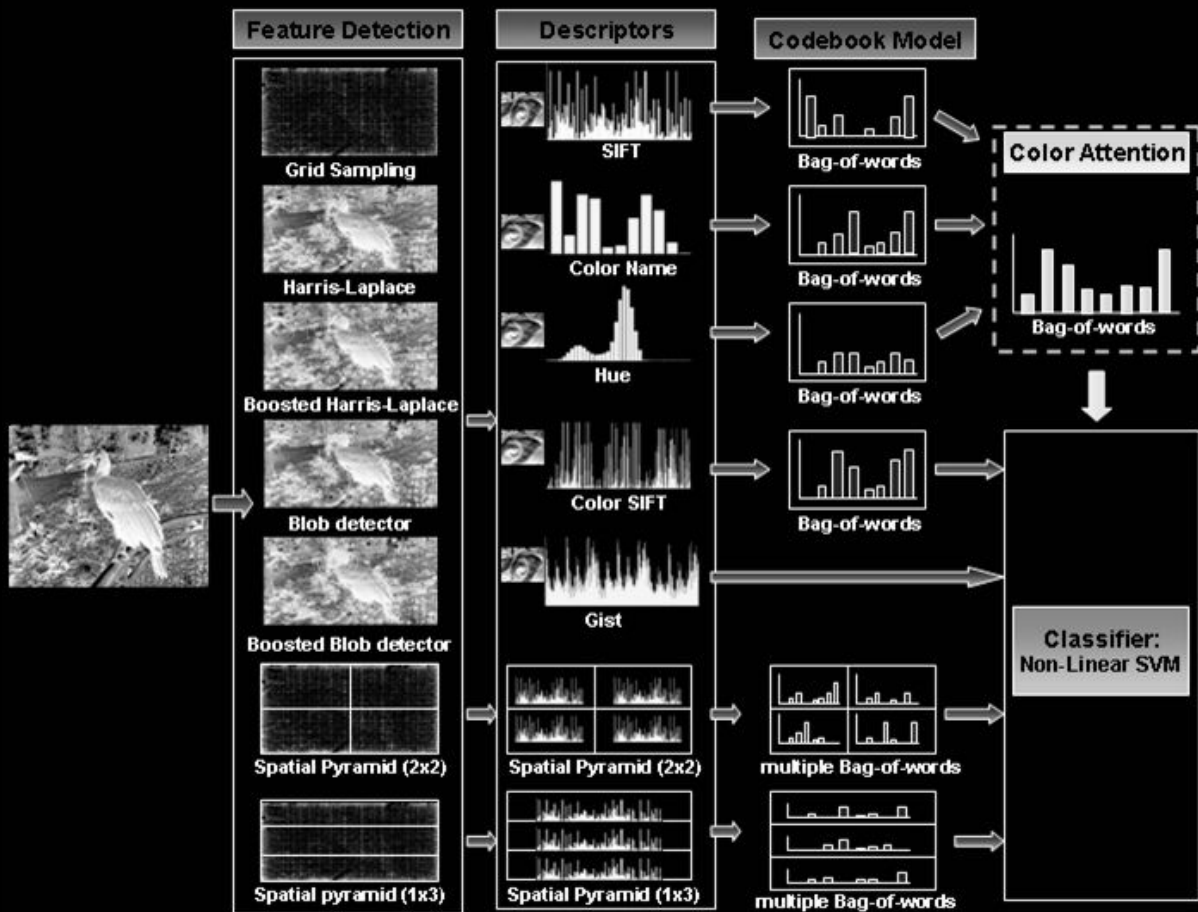
SIFT - Bag of Visual Words approach



Feature Composition

Complex Object recognition Approaches

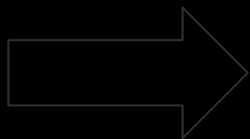
- Early/Late fusion
- Ensemble Classifiers



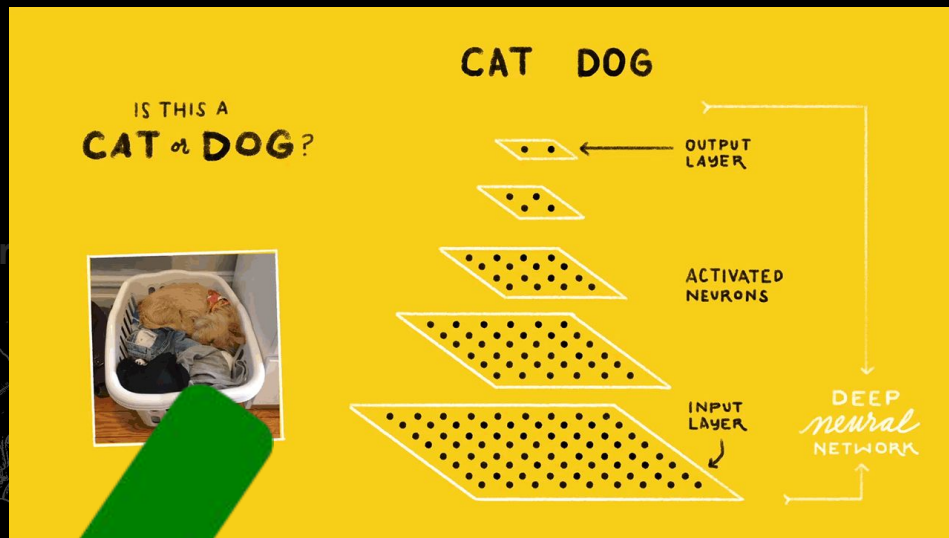
Deep Learning

Predicting input data

Image



Feature



Deep Learning

Convolution Layer Properties

Filtering



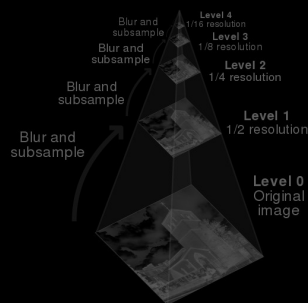
	0°	45°																		
Sobel	<table border="1"> <tr><td>-1</td><td>0</td><td>1</td></tr> <tr><td>-2</td><td>0</td><td>2</td></tr> <tr><td>-1</td><td>0</td><td>1</td></tr> </table>	-1	0	1	-2	0	2	-1	0	1	<table border="1"> <tr><td>0</td><td>1</td><td>2</td></tr> <tr><td>-1</td><td>0</td><td>1</td></tr> <tr><td>-2</td><td>-1</td><td>0</td></tr> </table>	0	1	2	-1	0	1	-2	-1	0
-1	0	1																		
-2	0	2																		
-1	0	1																		
0	1	2																		
-1	0	1																		
-2	-1	0																		
Kirsch	<table border="1"> <tr><td>-3</td><td>-3</td><td>3</td></tr> <tr><td>-3</td><td>0</td><td>3</td></tr> <tr><td>-3</td><td>-3</td><td>3</td></tr> </table>	-3	-3	3	-3	0	3	-3	-3	3	<table border="1"> <tr><td>-3</td><td>3</td><td>3</td></tr> <tr><td>-3</td><td>0</td><td>3</td></tr> <tr><td>-3</td><td>-3</td><td>-3</td></tr> </table>	-3	3	3	-3	0	3	-3	-3	-3
-3	-3	3																		
-3	0	3																		
-3	-3	3																		
-3	3	3																		
-3	0	3																		
-3	-3	-3																		
Robinson	<table border="1"> <tr><td>-1</td><td>0</td><td>1</td></tr> <tr><td>-1</td><td>0</td><td>1</td></tr> <tr><td>-1</td><td>0</td><td>1</td></tr> </table>	-1	0	1	-1	0	1	-1	0	1	<table border="1"> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>-1</td><td>0</td><td>1</td></tr> <tr><td>-1</td><td>-1</td><td>0</td></tr> </table>	0	1	1	-1	0	1	-1	-1	0
-1	0	1																		
-1	0	1																		
-1	0	1																		
0	1	1																		
-1	0	1																		
-1	-1	0																		

Scaling

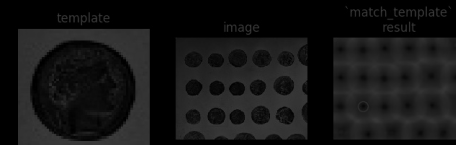
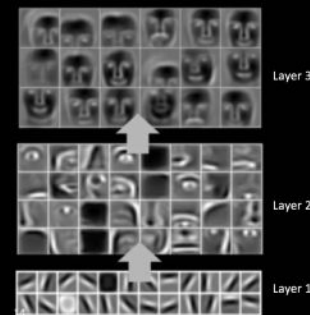
12	20	30	0
8	12	2	0
34	70	37	4
112	100	25	12

 $\xrightarrow{2 \times 2 \text{ Max-Pool}}$

20	30
112	37



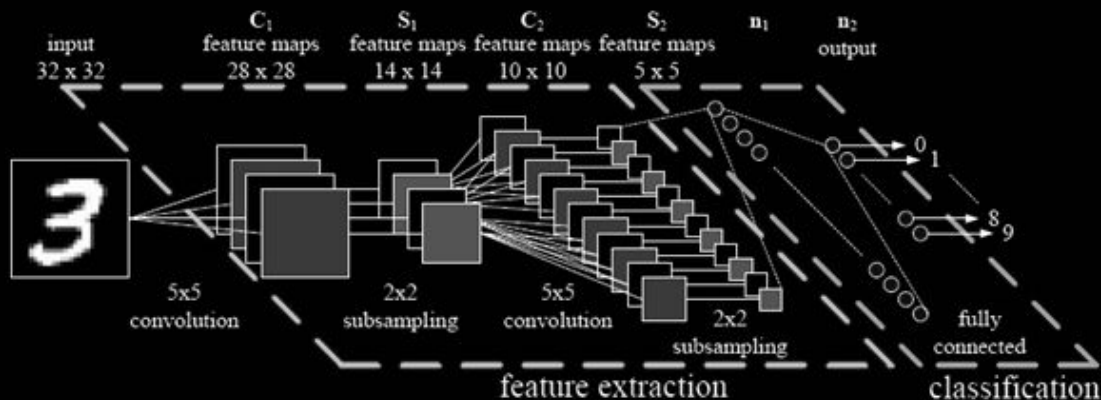
Matching



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

