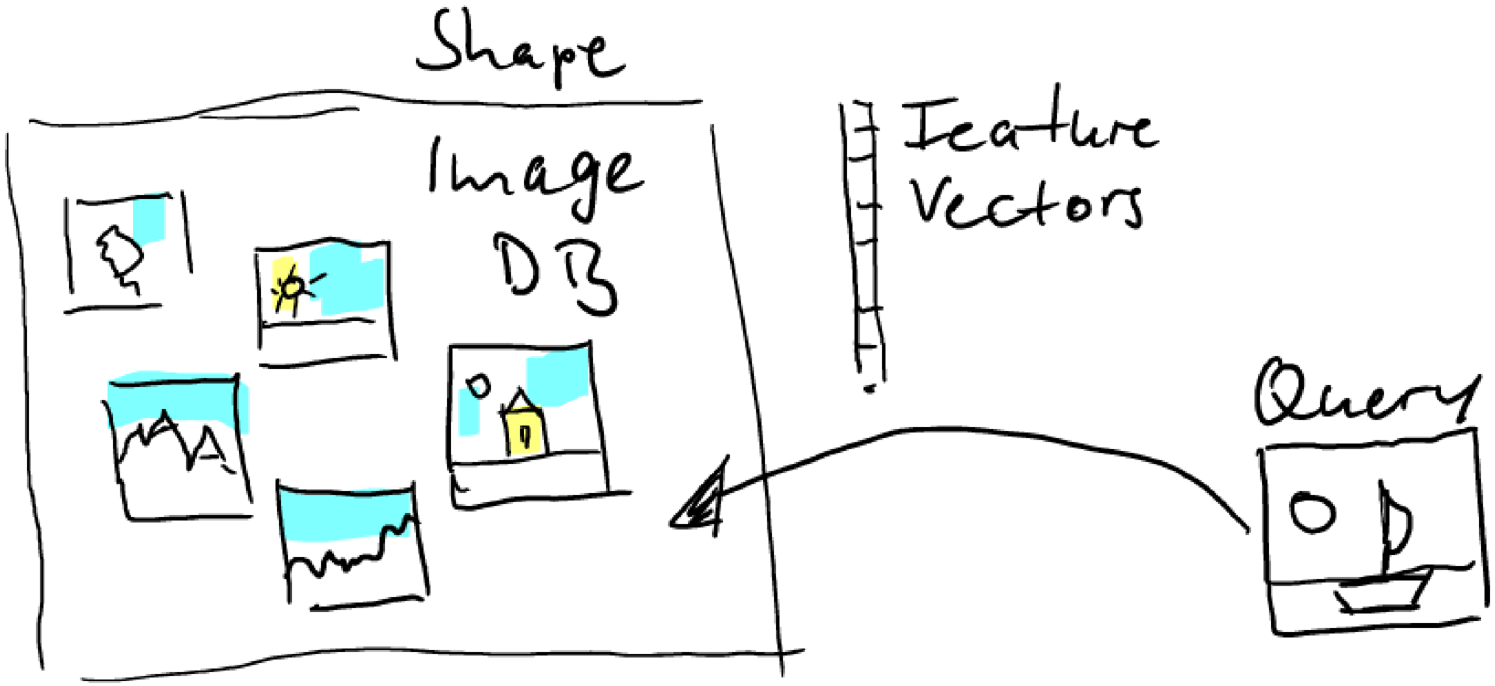
Learning from Images

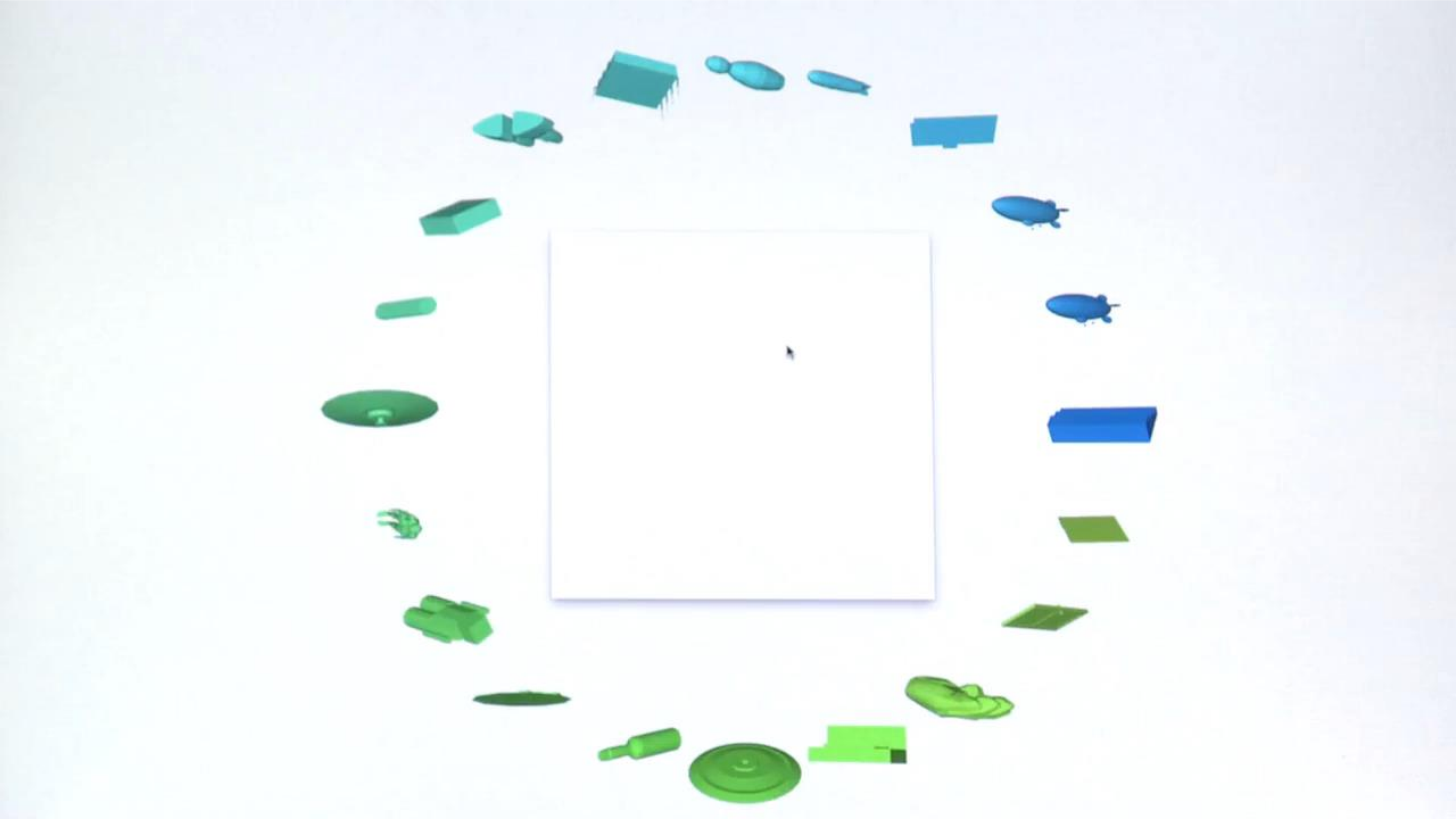
Image Classification

Master DataScience

Prof. Dr. Kristian Hildebrand
khildebrand@bht-berlin.de

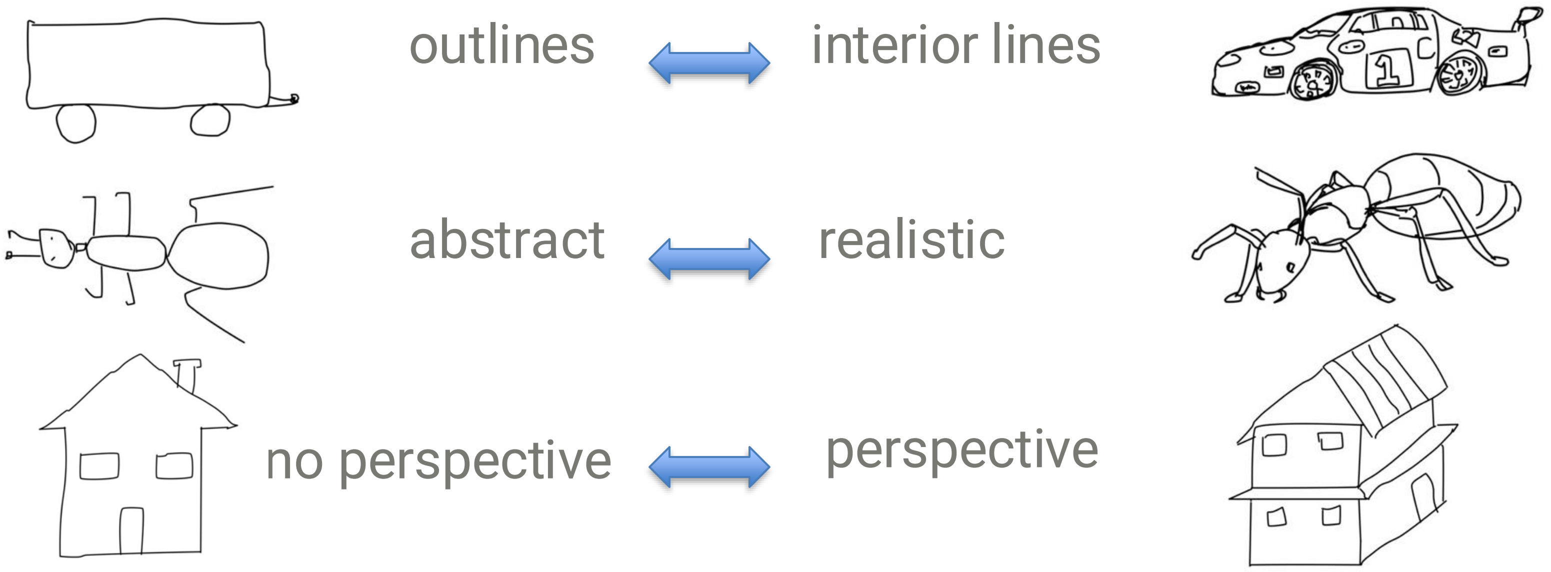
Application I: Image Retrieval



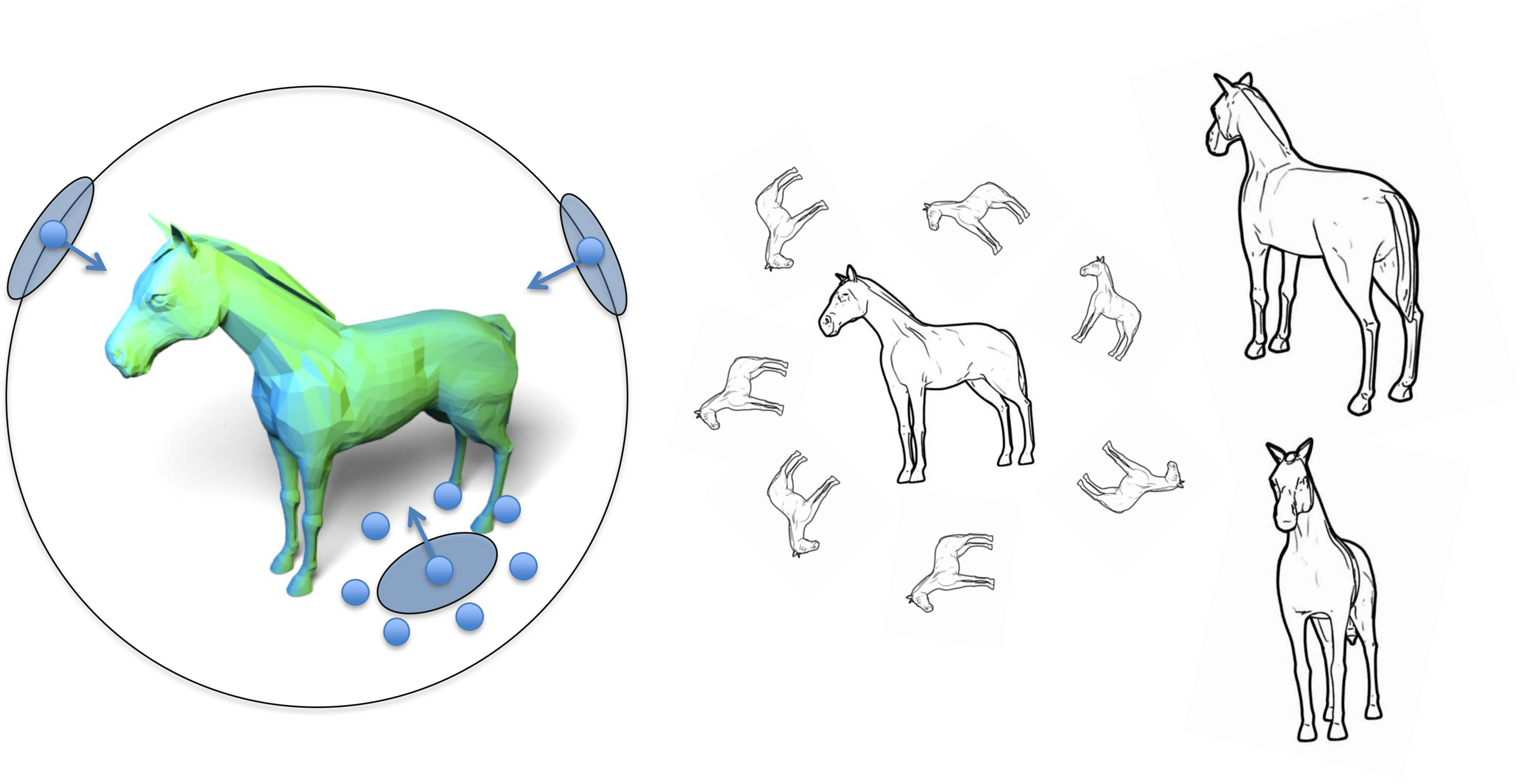


How Do Humans Sketch for Shape Retrieval?

- Large variety of sketching styles:



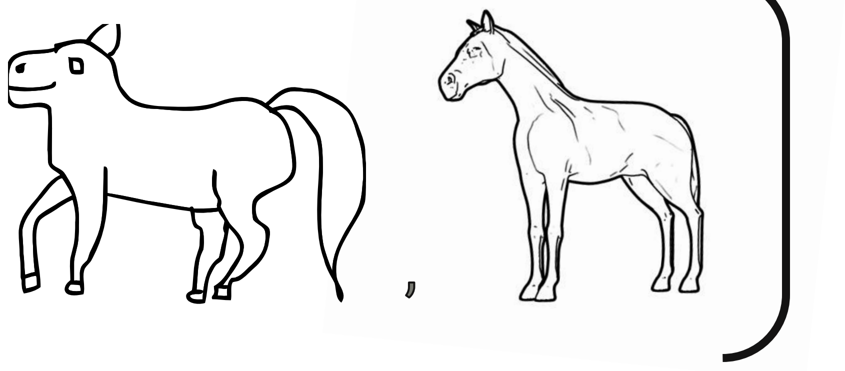
View Generation



Similarity Measure

Image-based retrieval

similarity



Requirements:

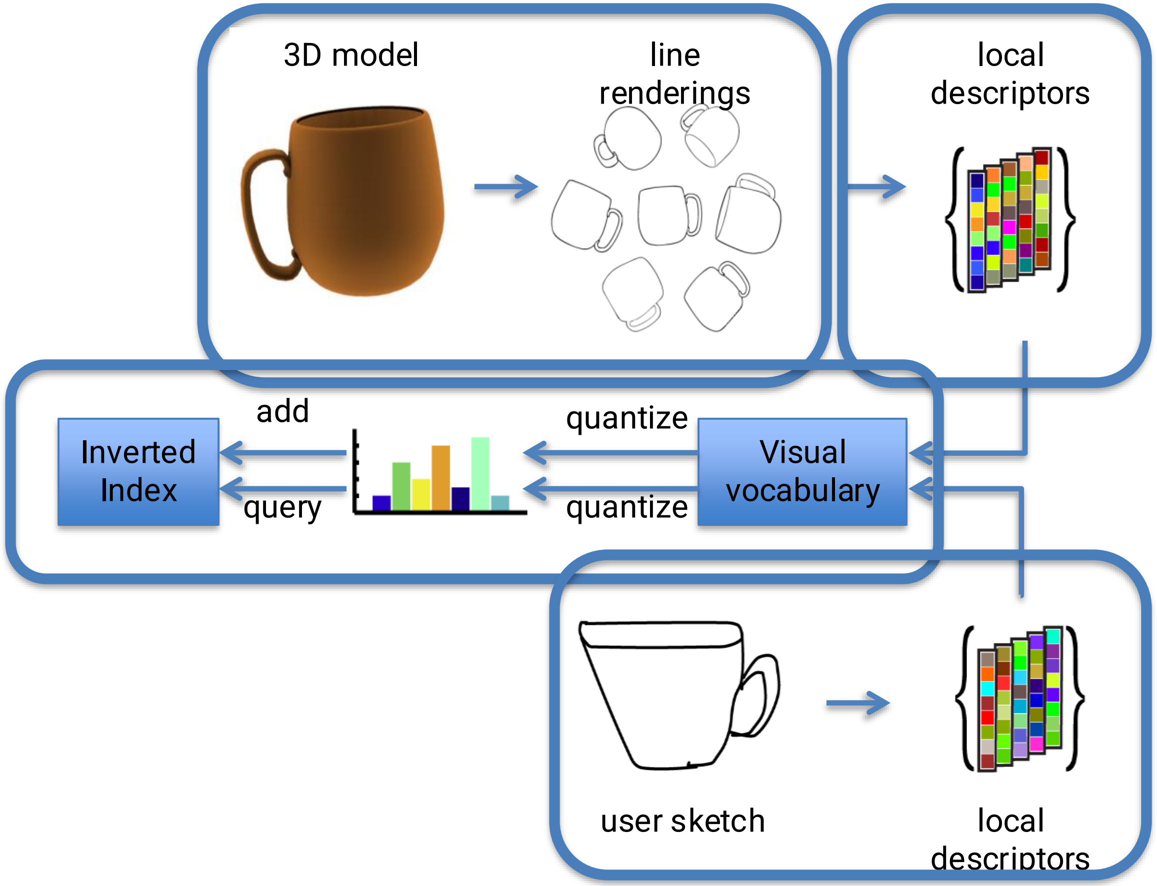
- Tolerate local and global deformations
- Support partial matching
- Fast and efficient

Need appropriate feature transform (descriptor)

Overview Computing Pipeline

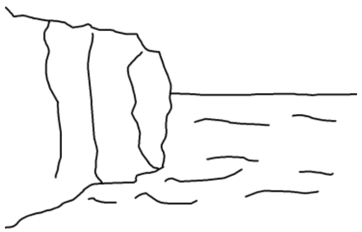
Offline Indexing

Online Search



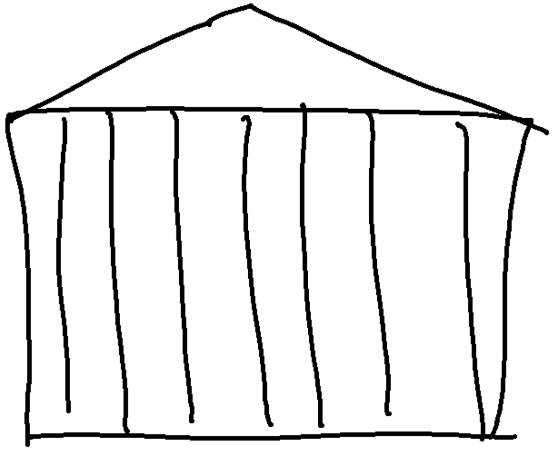
Comparing sketch and image

- Sketches/images represented as feature vector
  - encode essential image/sketch attributes
  - small memory footprint
  - fast to compare

dist( , )

$$= \left\| \begin{bmatrix} 0.21 \\ 0.13 \\ 0.75 \\ 0.31 \\ 0.41 \\ 0.06 \\ 0.93 \\ 0.17 \end{bmatrix} - \begin{bmatrix} 0.01 \\ 0.17 \\ 0.68 \\ 0.31 \\ 0.44 \\ 0.05 \\ 0.97 \\ 0.13 \end{bmatrix} \right\|_m$$

Feature extraction from sketch and image



Feature

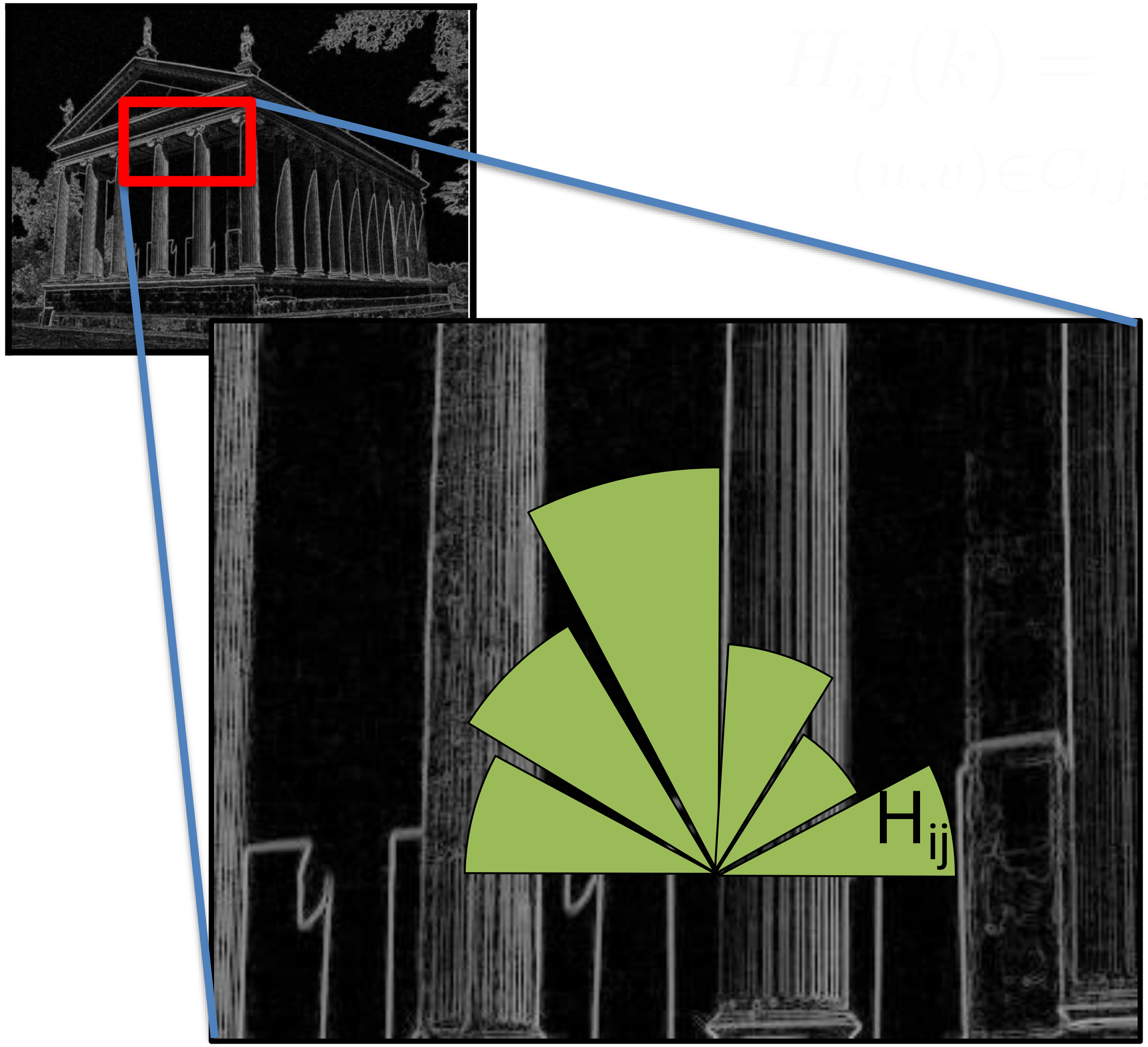
Feature

Featur

Edge Histogram Descriptor

Encodes distribution of gradients

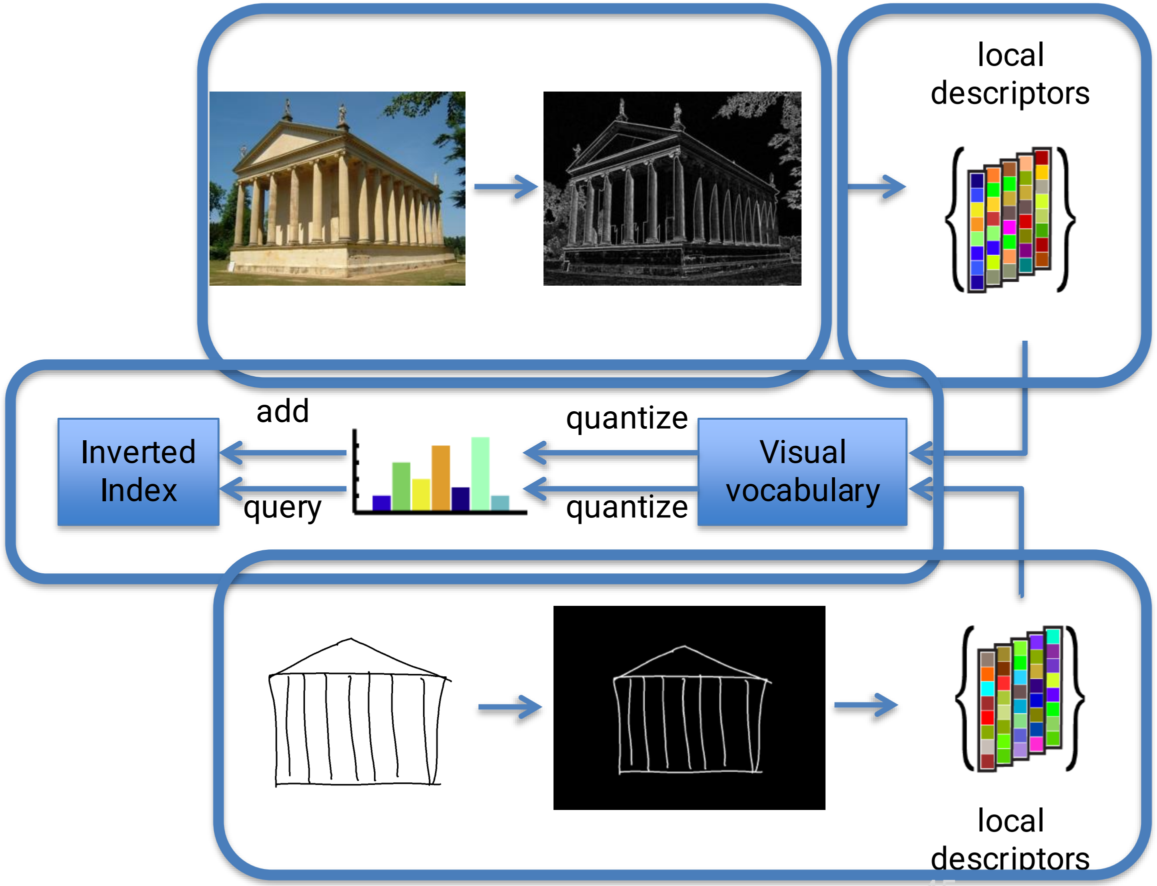
$$H_{ij}(k) = \sum_{(u,v)\in C_{ij},\, o(g_{ij})\in[k/d,k+1/d[} g_{uv}^{\top} g_{uv}$$



Computing Pipeline

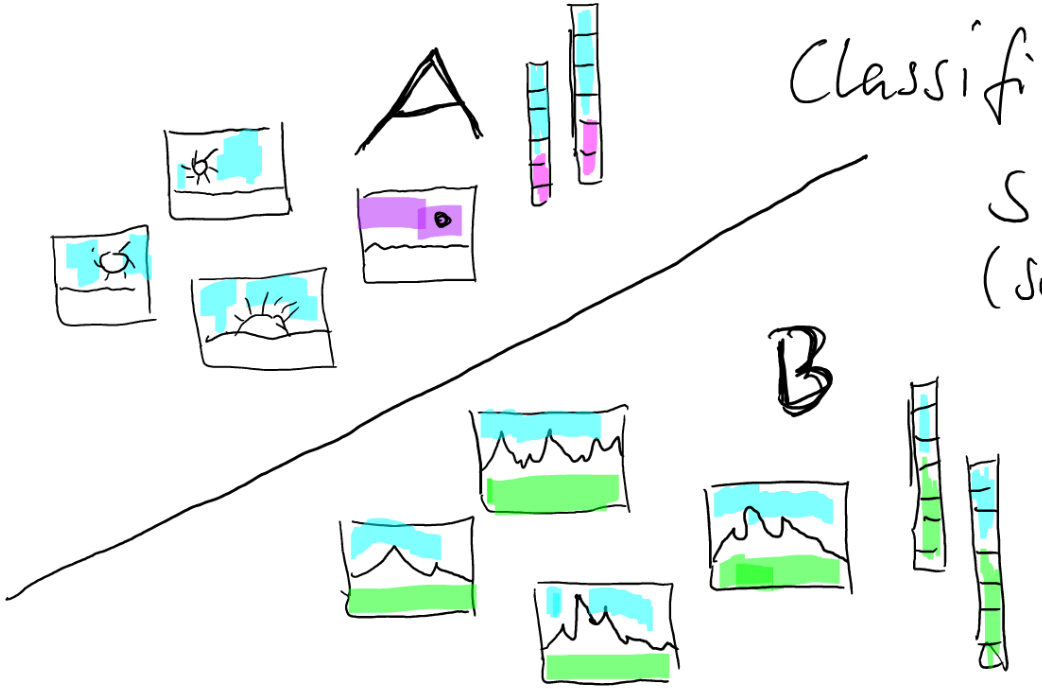
Offline Indexing

Online Search



Application II:
Image Classification

Application II: Image Classification



Classification

SVM
(Support Vector Machines)

Application II: Image Classification

How do users sketch objects?

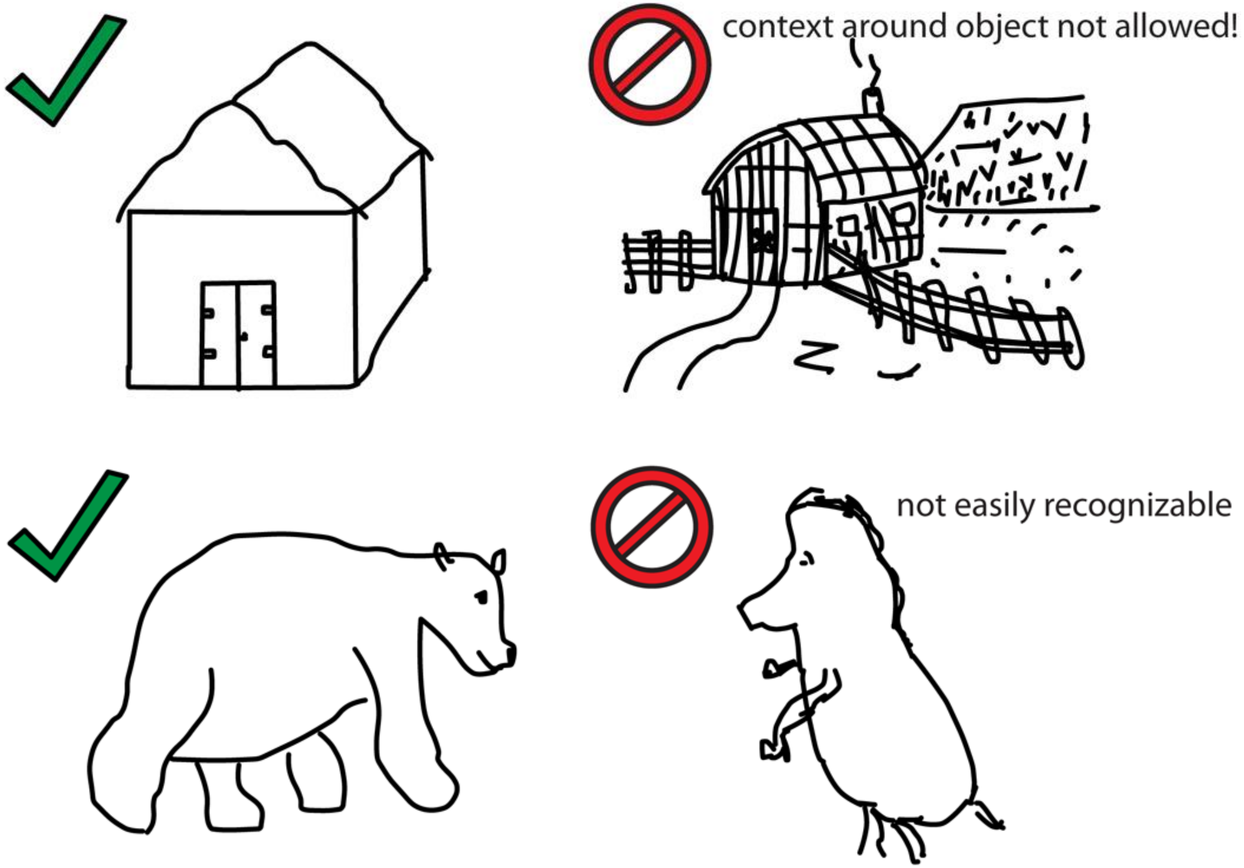
blimp

cow

How do users sketch objects?

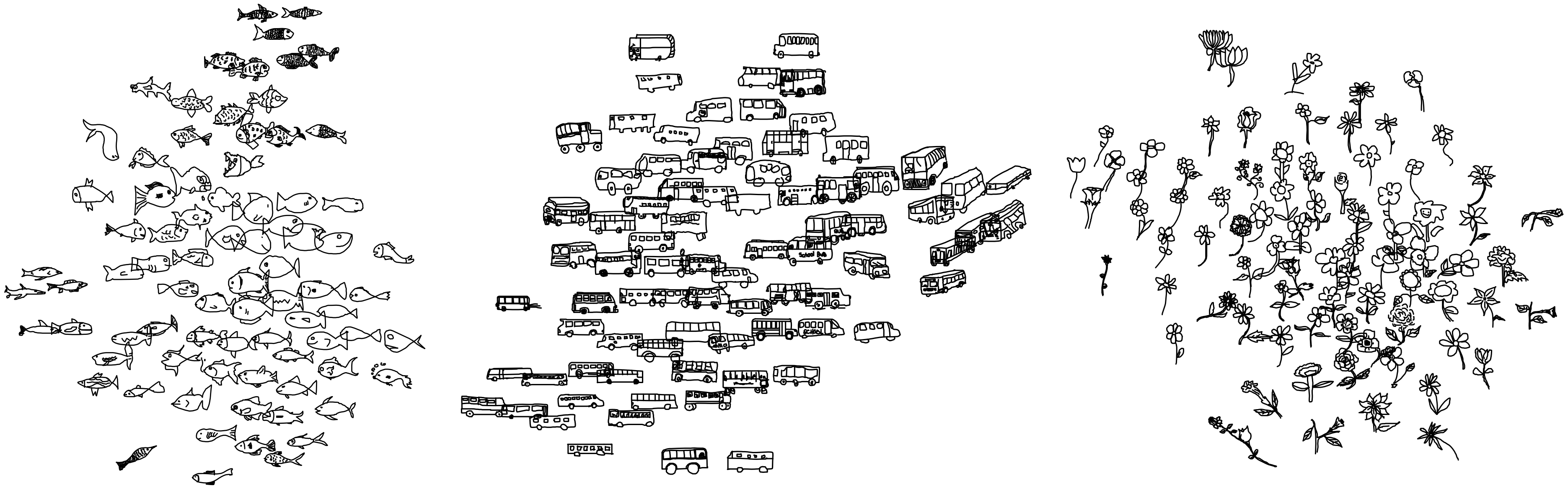
barn

bear



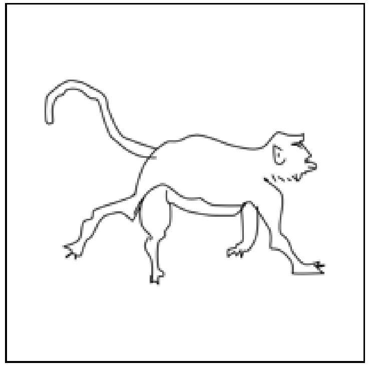
How do users sketch objects?

- 20,000 sketches in 250 categories
  - 1,350 unique participants, 741 hours drawing time



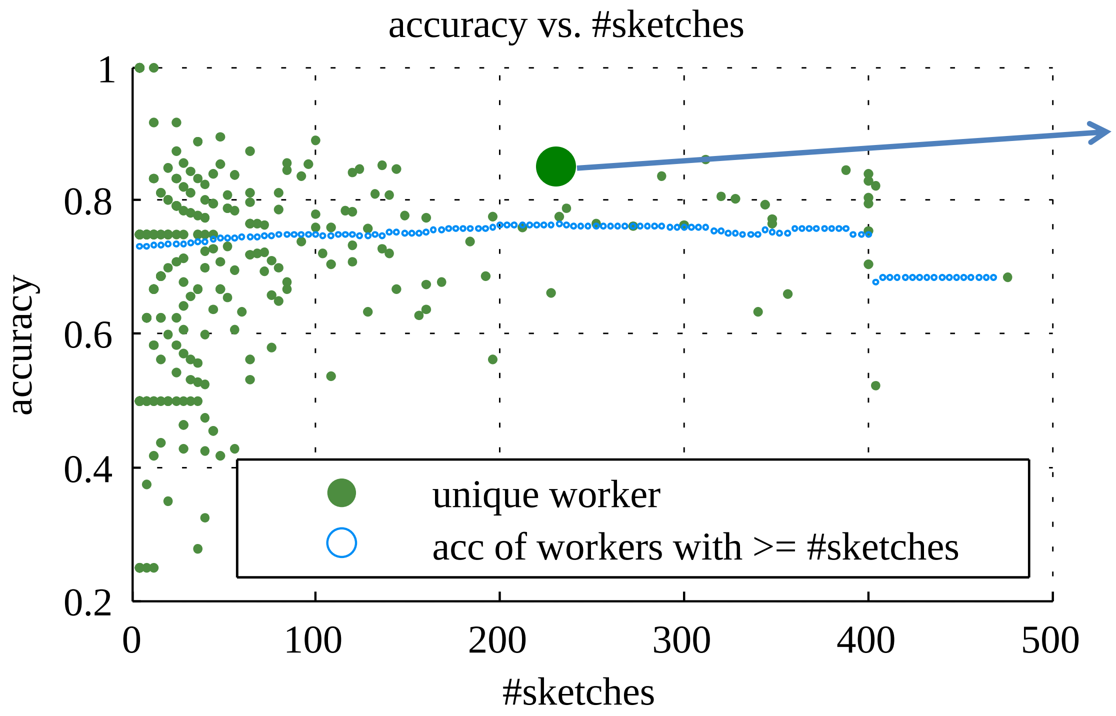
Human sketch recognition

- 2nd study on Amazon Mechanical Turk



Human sketch recognition

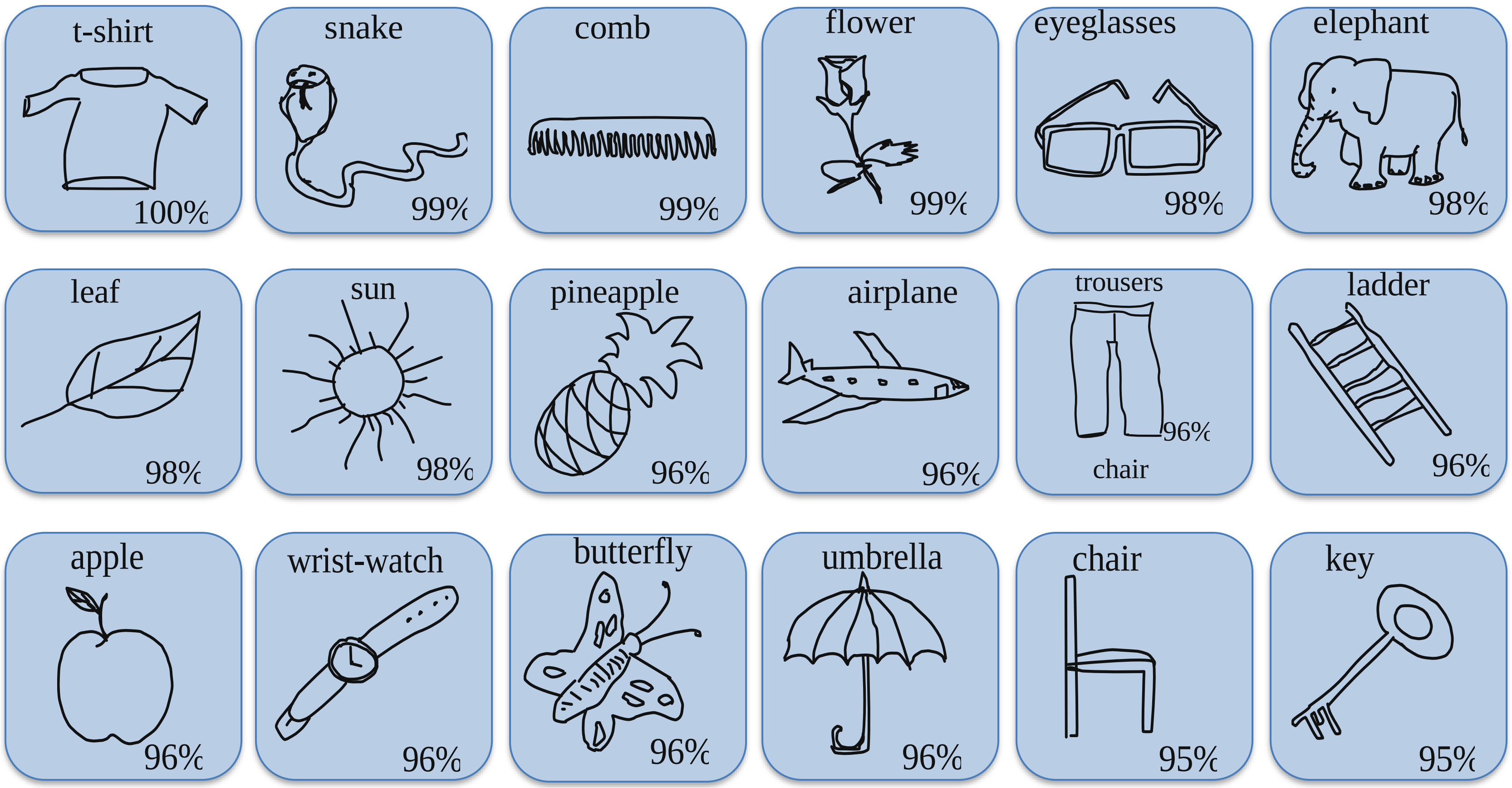
- 73% overall human recognition accuracy



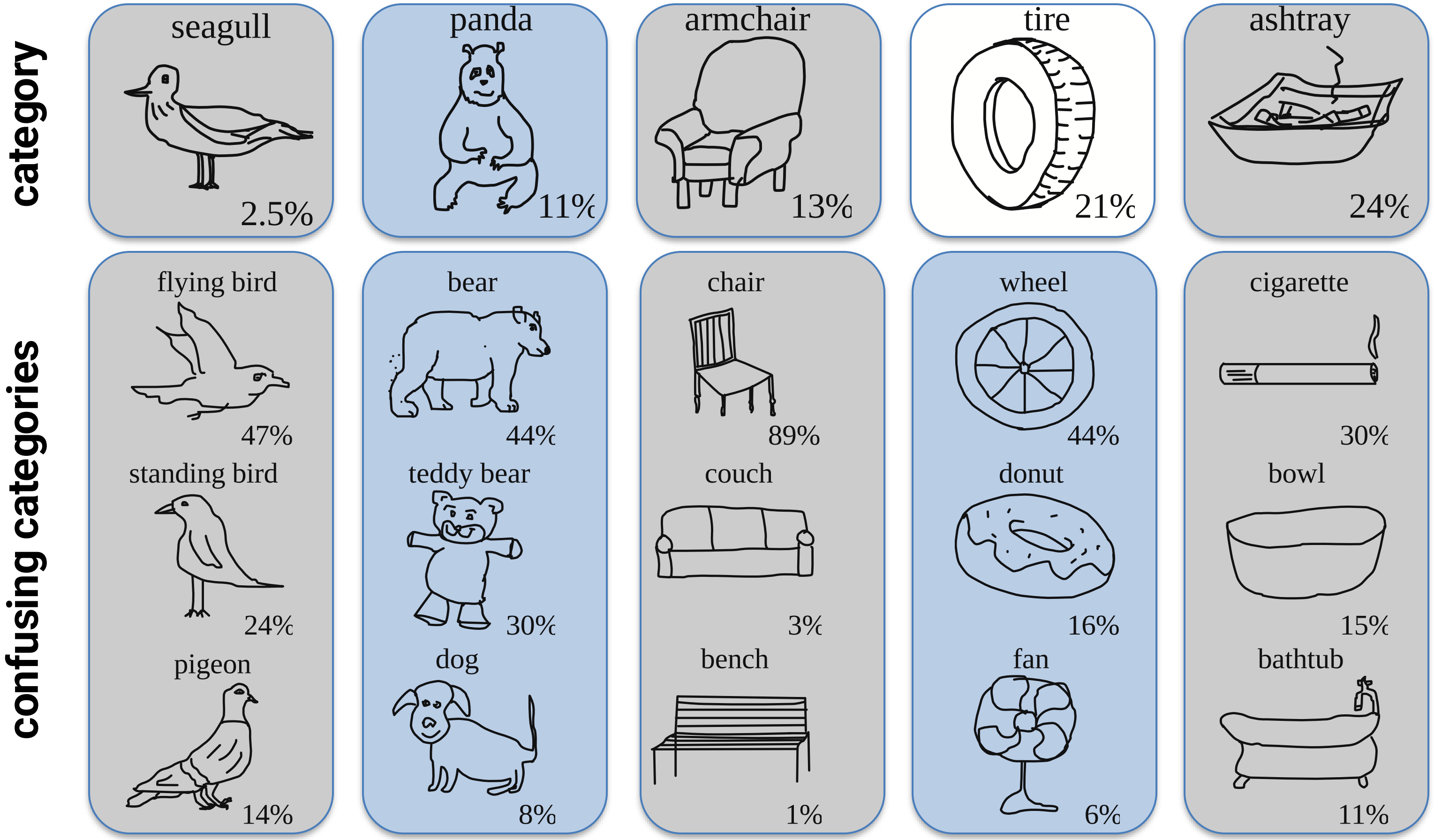
This particular worker:

- 250 sketches
- 85% accuracy

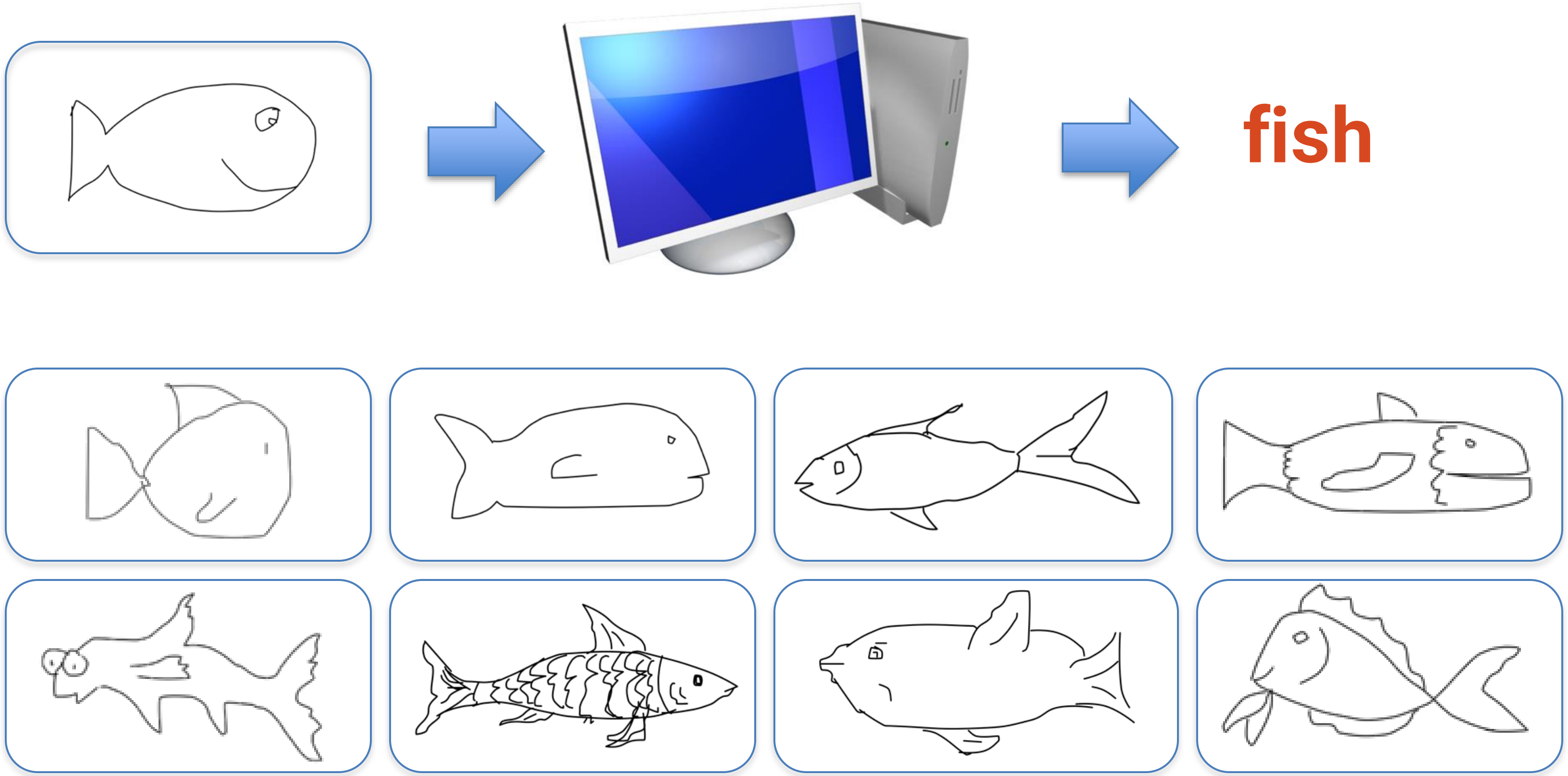
Human sketch recognition



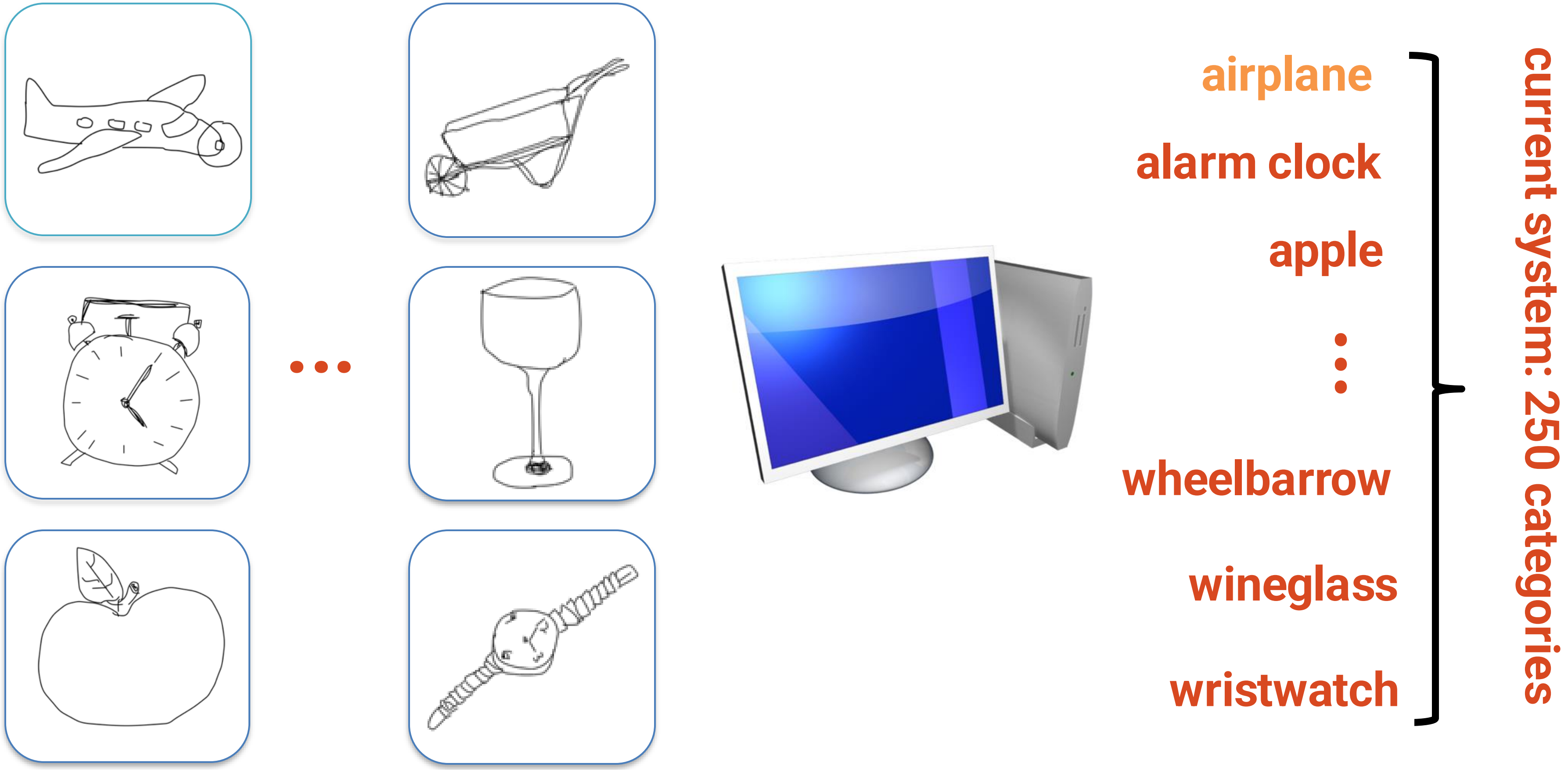
Human sketch recognition



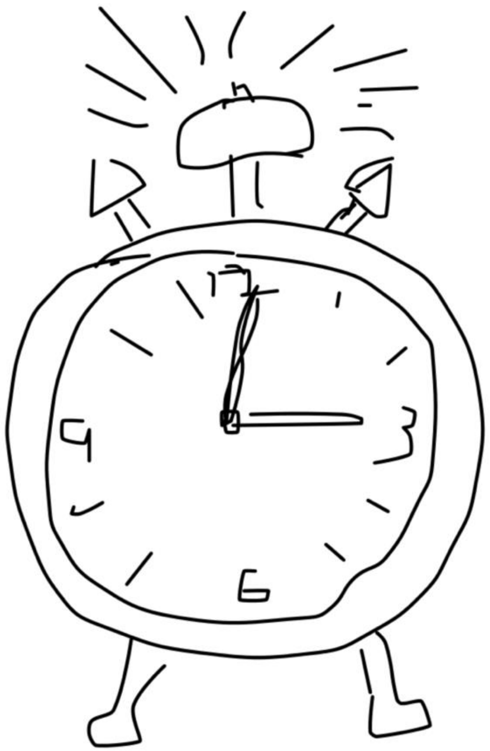
Computational sketch recognition

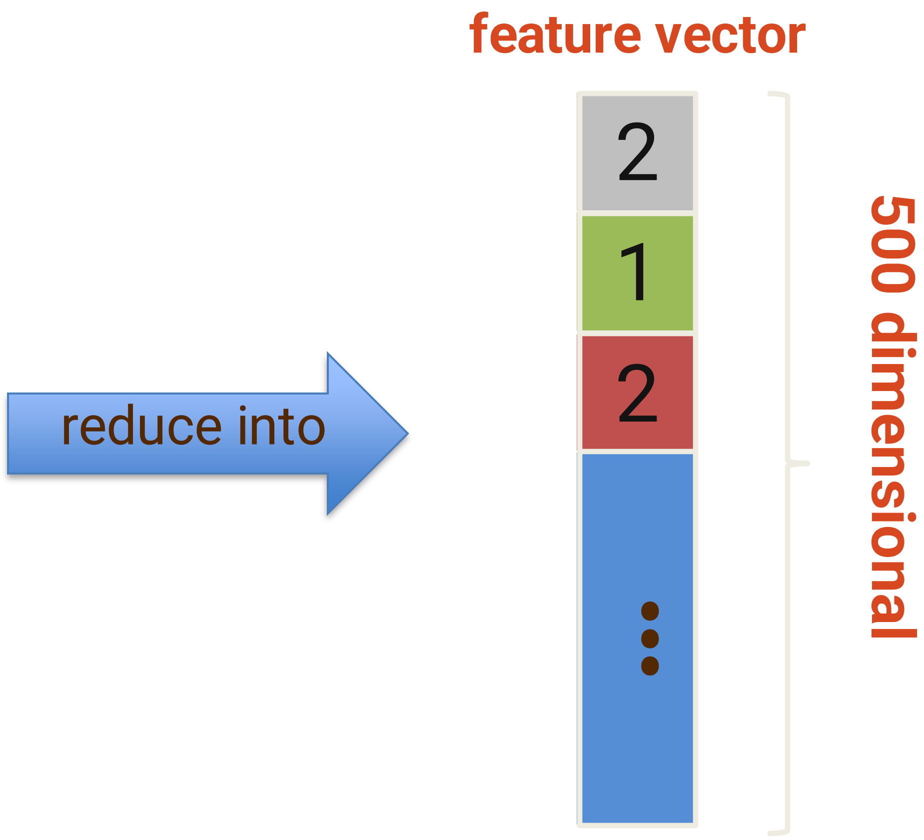


Computational sketch recognition

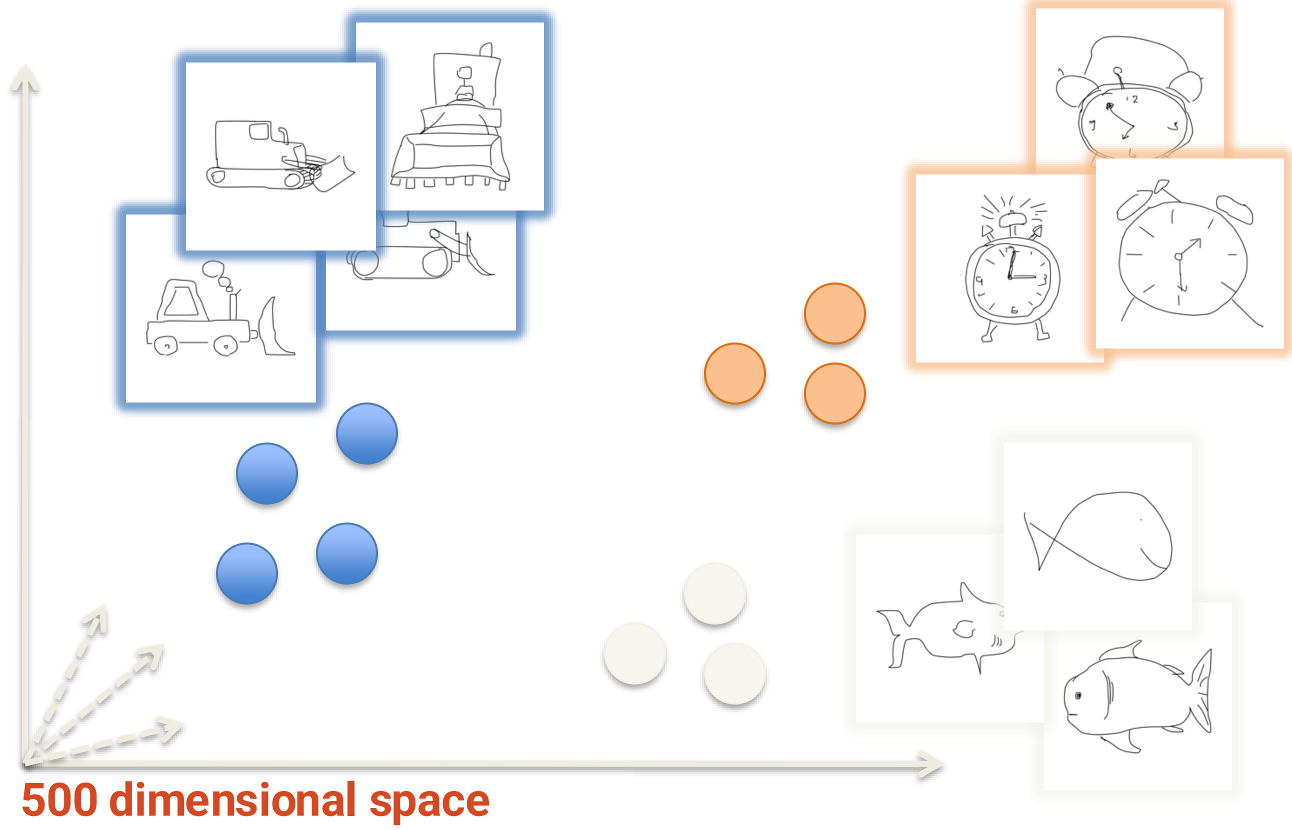


Bag-of-feature representation

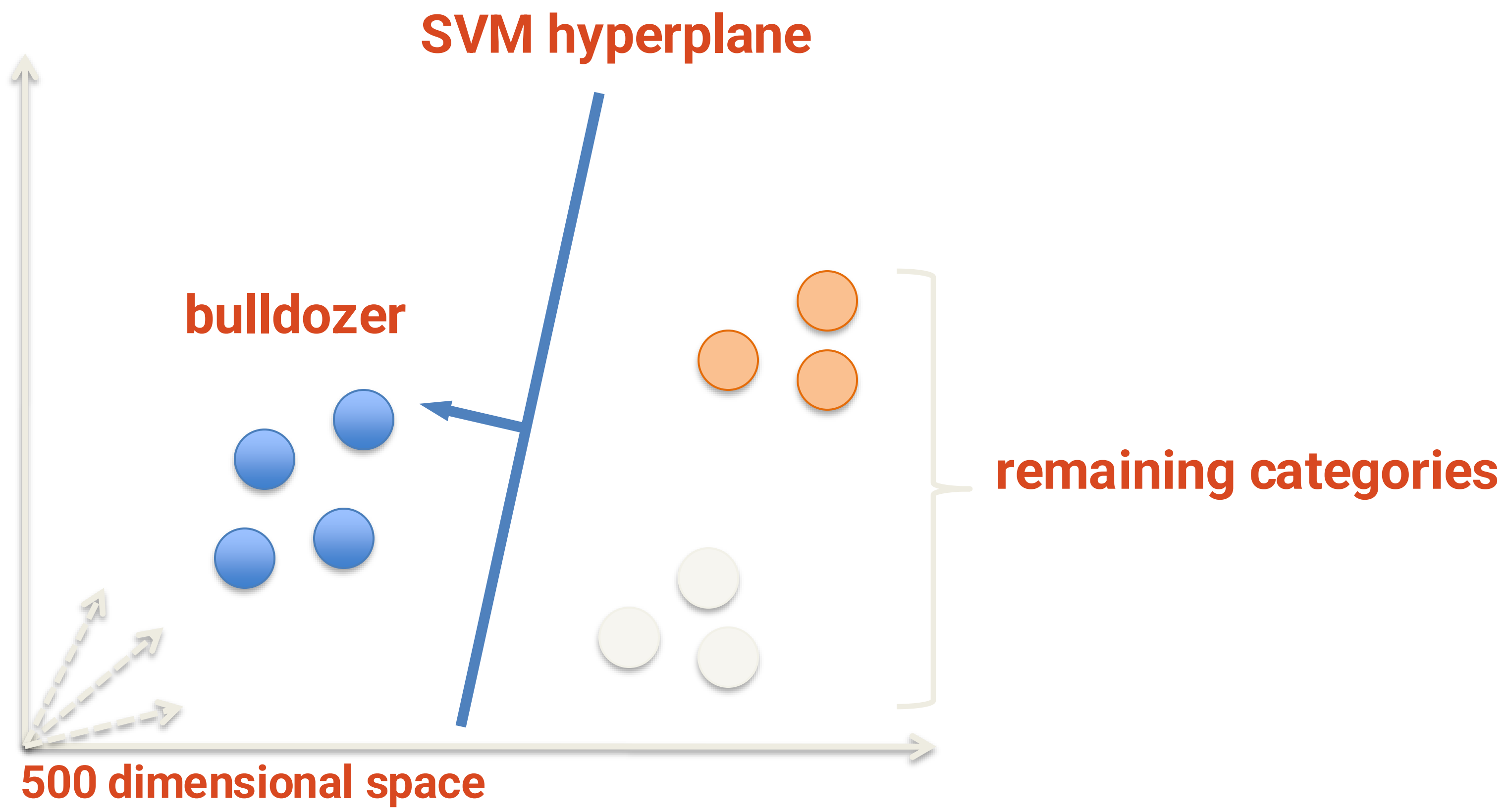




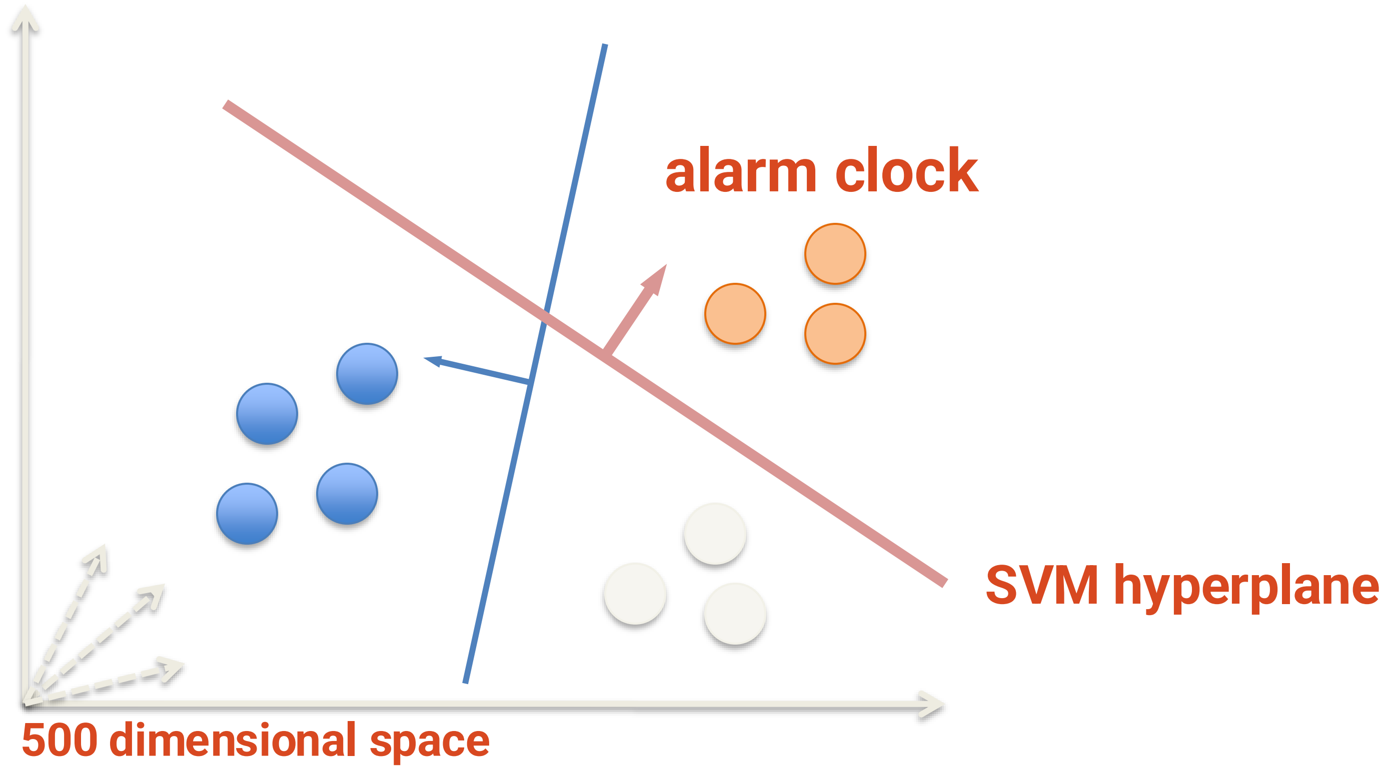
Sketch-feature space



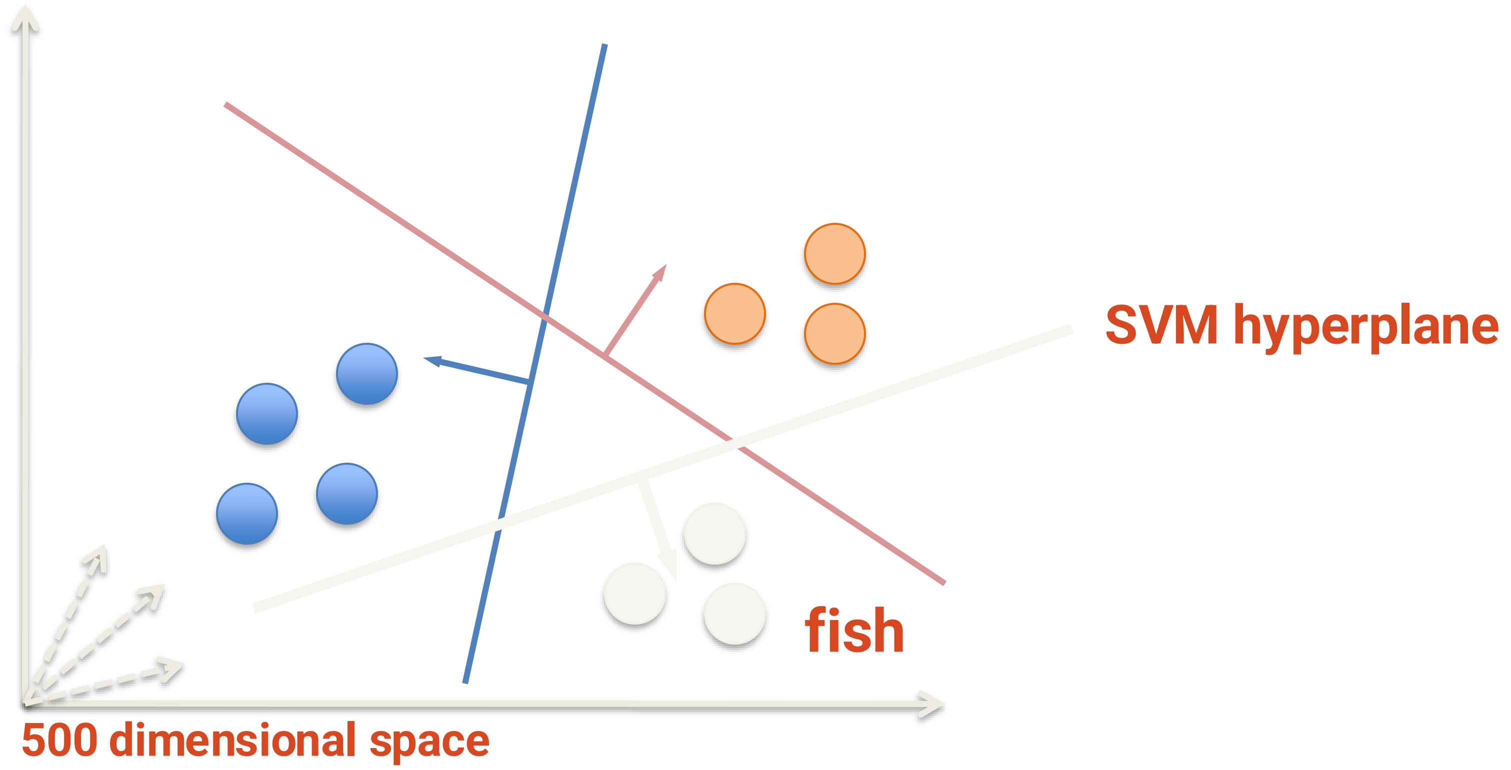
Training SVMs



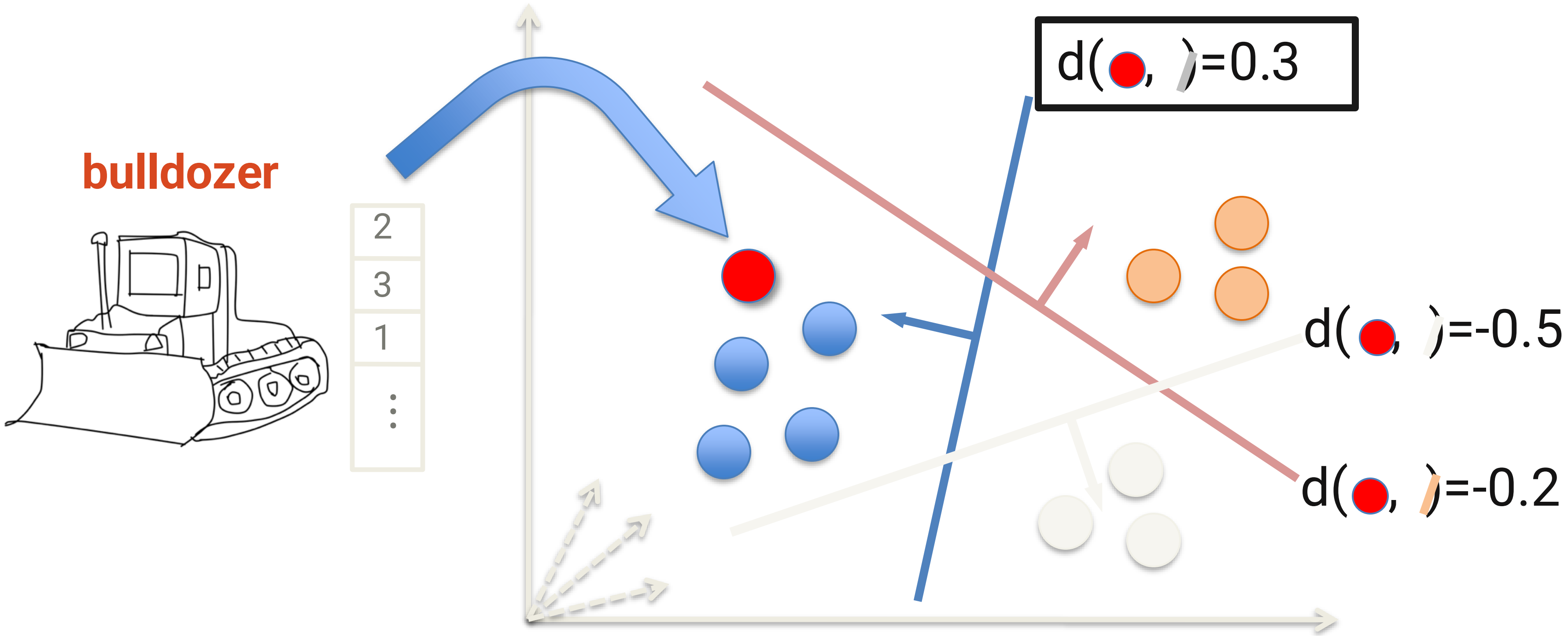
Training SVMs



Training SVMs



Classification

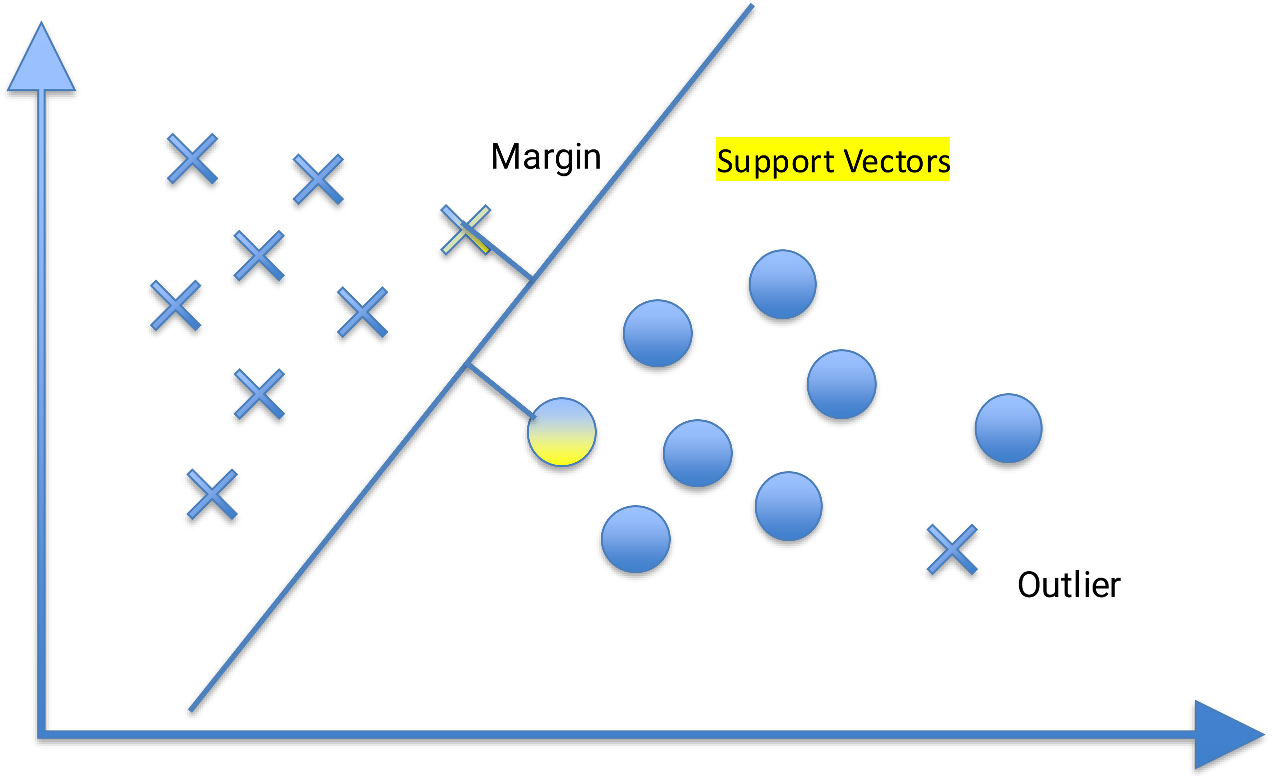


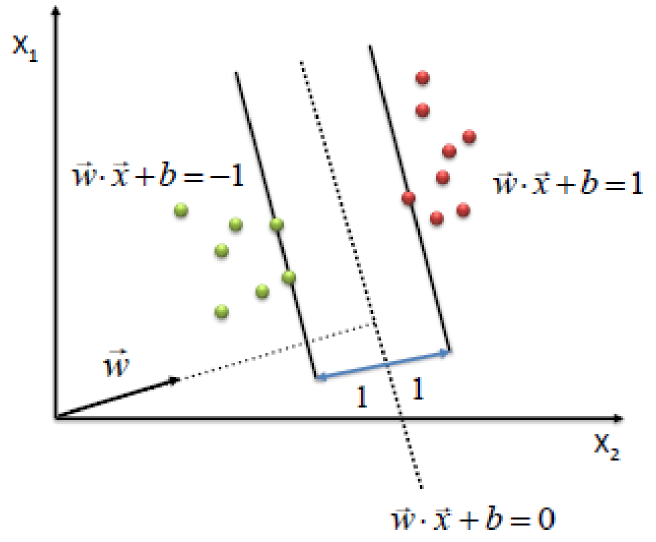
classify a sketch according to the model with largest response

Support Vector Machine

- Robust classifier
- maximises Margin and tolerates outliers

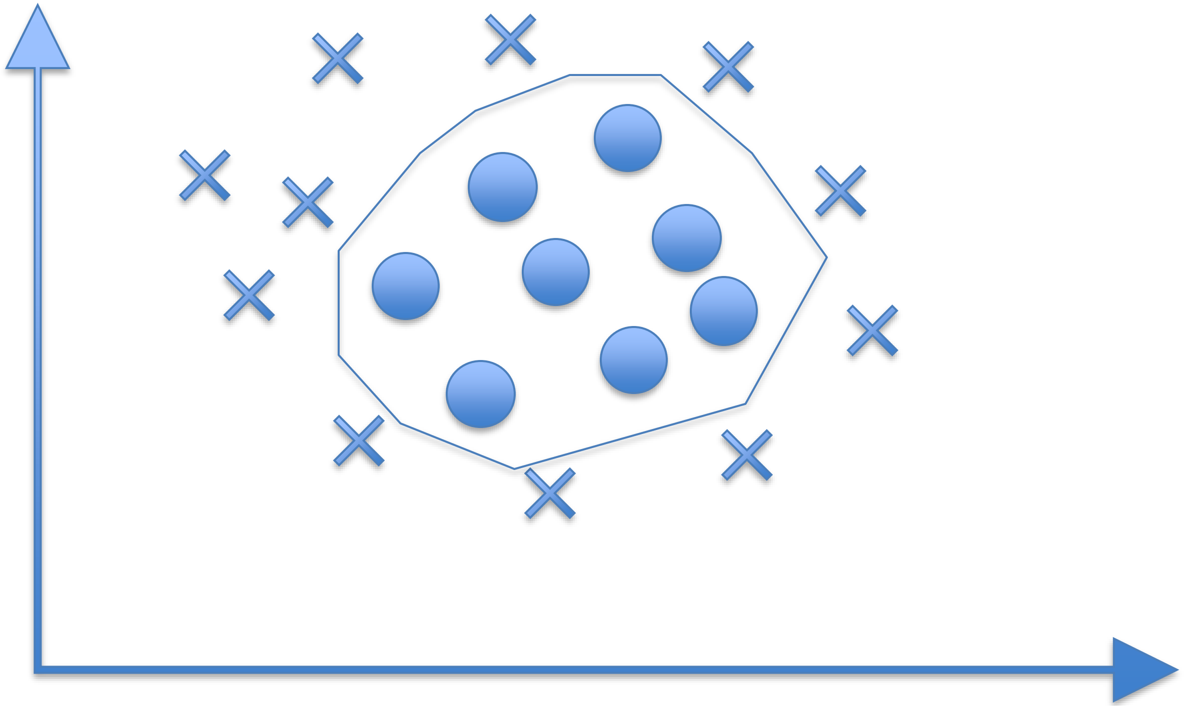
$$\text{minimize}\left(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\max(0, 1 - y_i(w \cdot x_i + b)) + \lambda\|w\|^2\right)$$





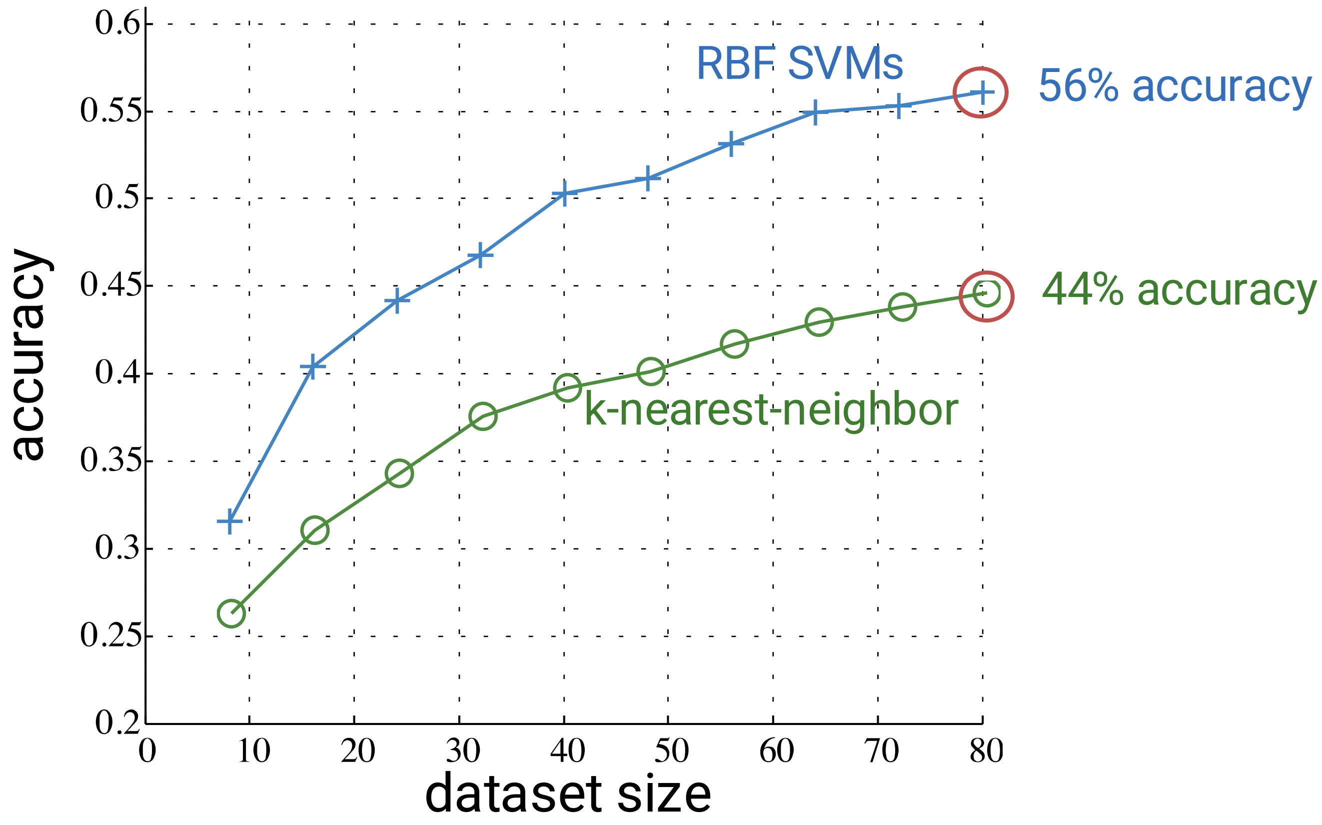
Non-linear Support Vector Machine

- non-linear separation by kernel trick
- Kernel function transforms low dimensional features into higher dimensional space
  - Separation can be found in higher dimensions
  - Non-linearity is achieved by back-transformation



White-board example, Code

Classification



Important!!! - Metrics for the evaluation of classification tasks

- **Accuracy:** Proportion of correctly classified examples

$$\text{Genauigkeit} = \frac{\text{Anzahl der korrekten Vorhersagen}}{\text{Gesamtanzahl der Beispiele}}$$

- **Precision:** Proportion of correct positive predictions out of all positive predictions

$$\text{Präzision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

- **Recall:** Proportion of correctly identified positive examples out of all actual positive examples

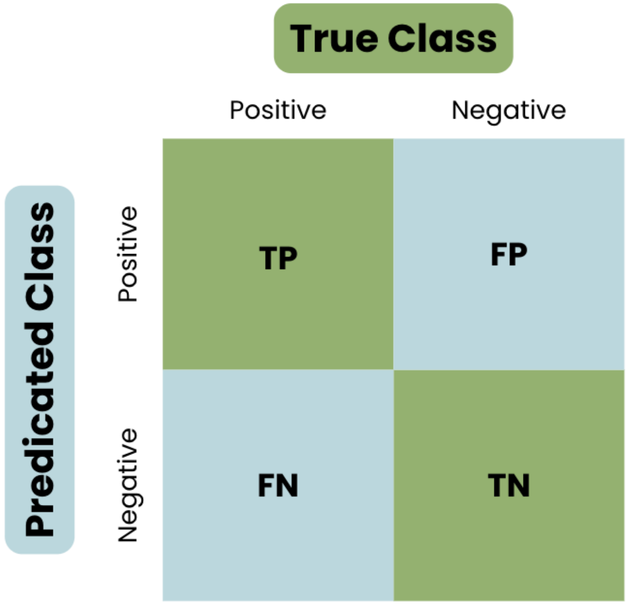
$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

Important!!! - Metrics for the evaluation of classification tasks

- **F1-Score:** harmonic mean of precision and recall

$$\text{F1-Score} = 2 \cdot \frac{\text{Präzision} \cdot \text{Recall}}{\text{Präzision} + \text{Recall}}$$

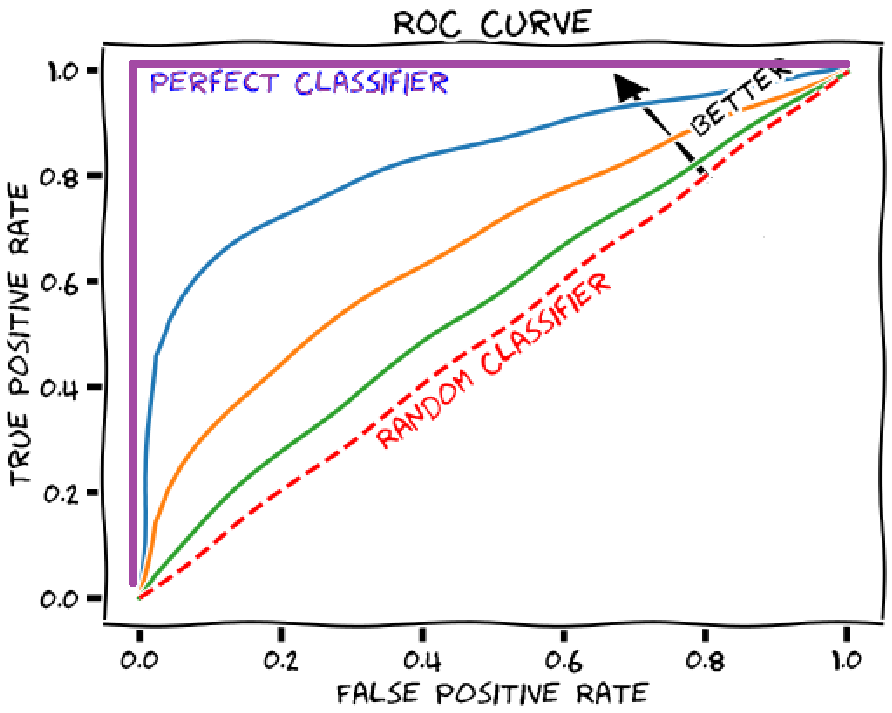
- **Confusion Matrix:** Table shows True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN)



https://www.datacamp.com/tutorial/what-is-a-confusion-matrix-in-machine-learning

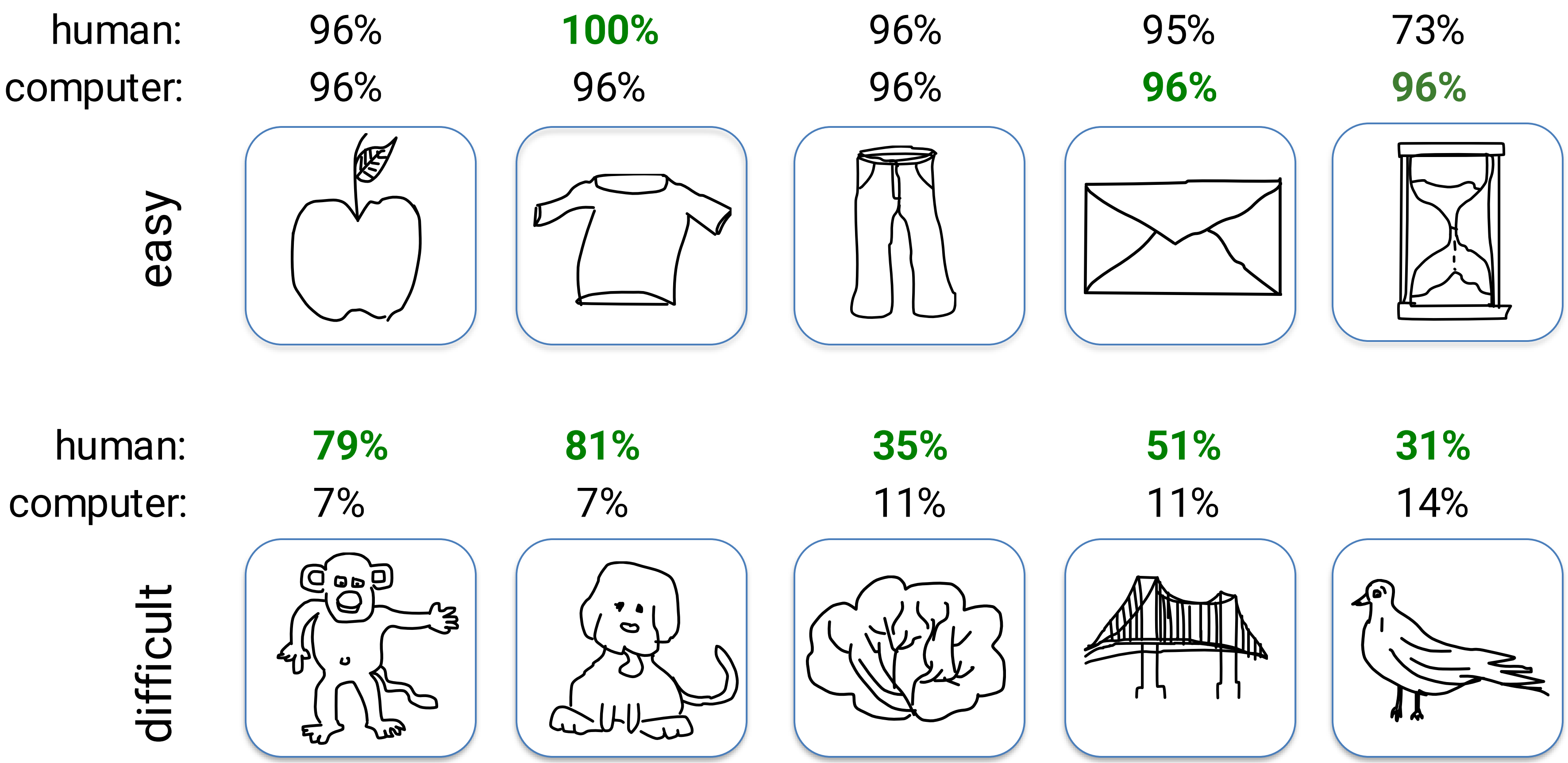
Important!!! - Metrics for the evaluation of classification tasks

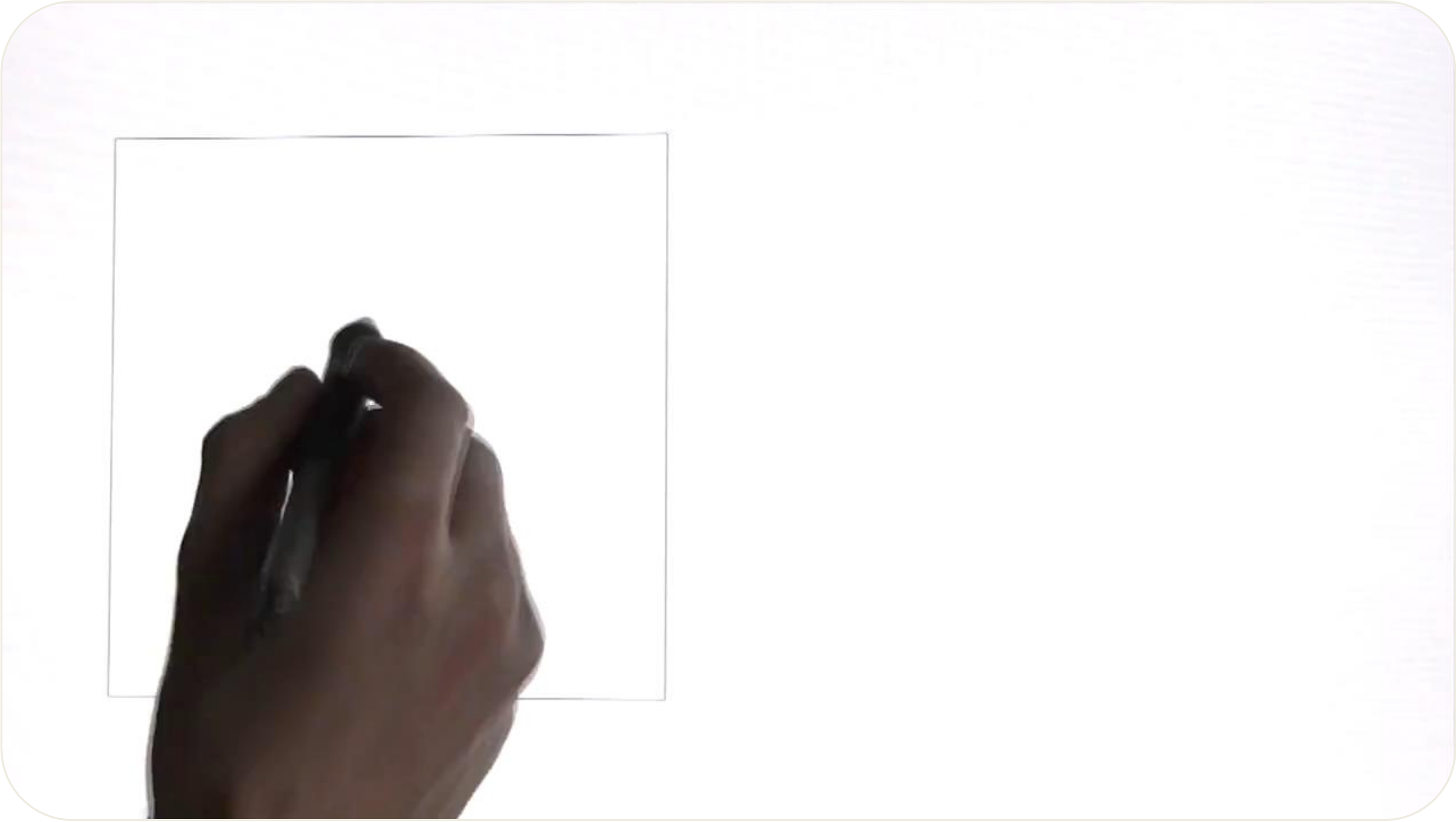
- **AUC-ROC (Area Under the Receiver Operating Characteristic Curve):** Grafische Darstellung eines Klassifikationsmodells bei verschiedenen Schwellenwerten.
  - AUC (Area Under the Curve) misst gesamte zweidimensionale Fläche unter der ROC-Kurve
  - AUC von 1 repräsentiert ein perfektes Modell, AUC von 0.5 Modell nicht besser als Zufall

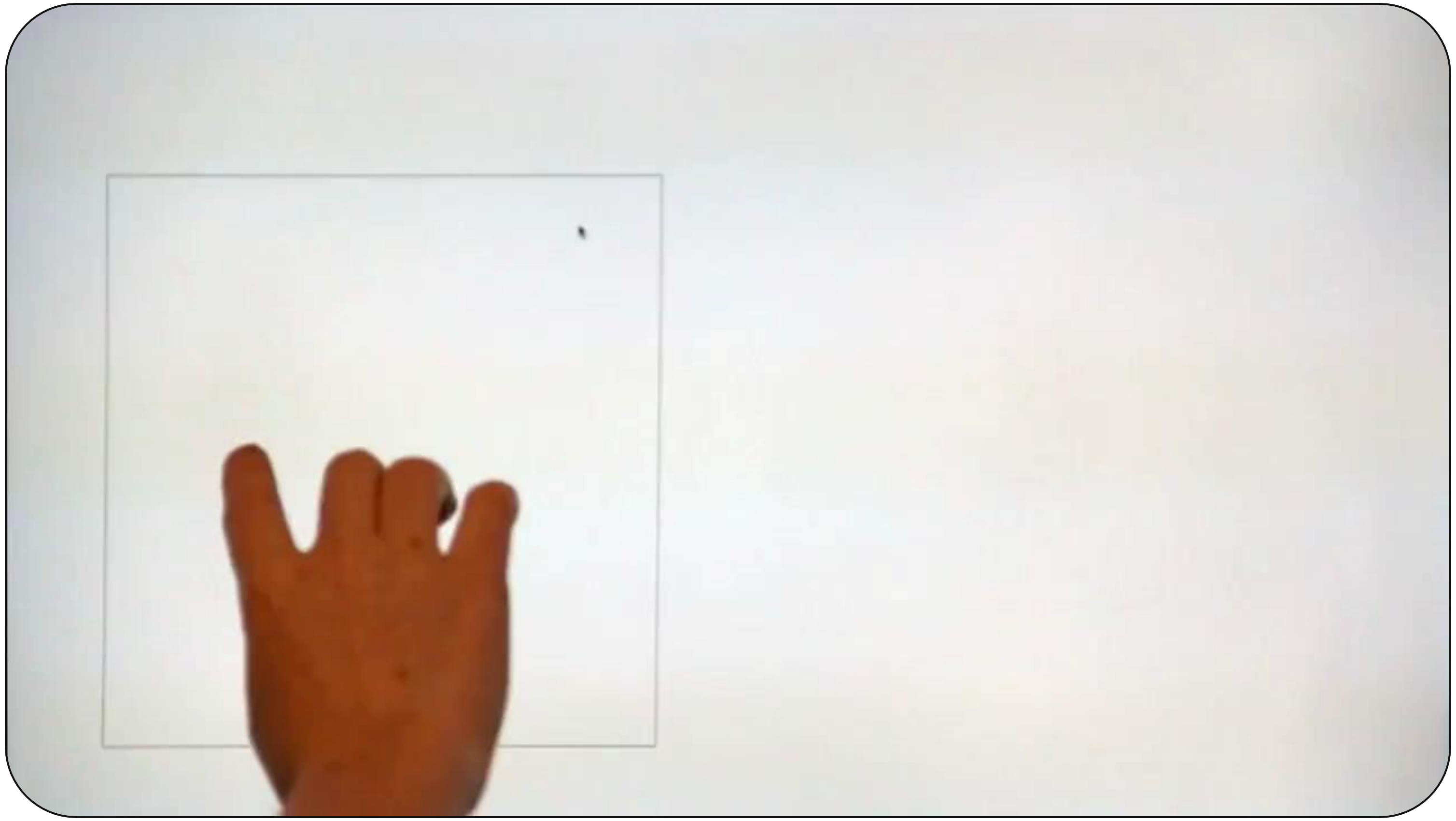


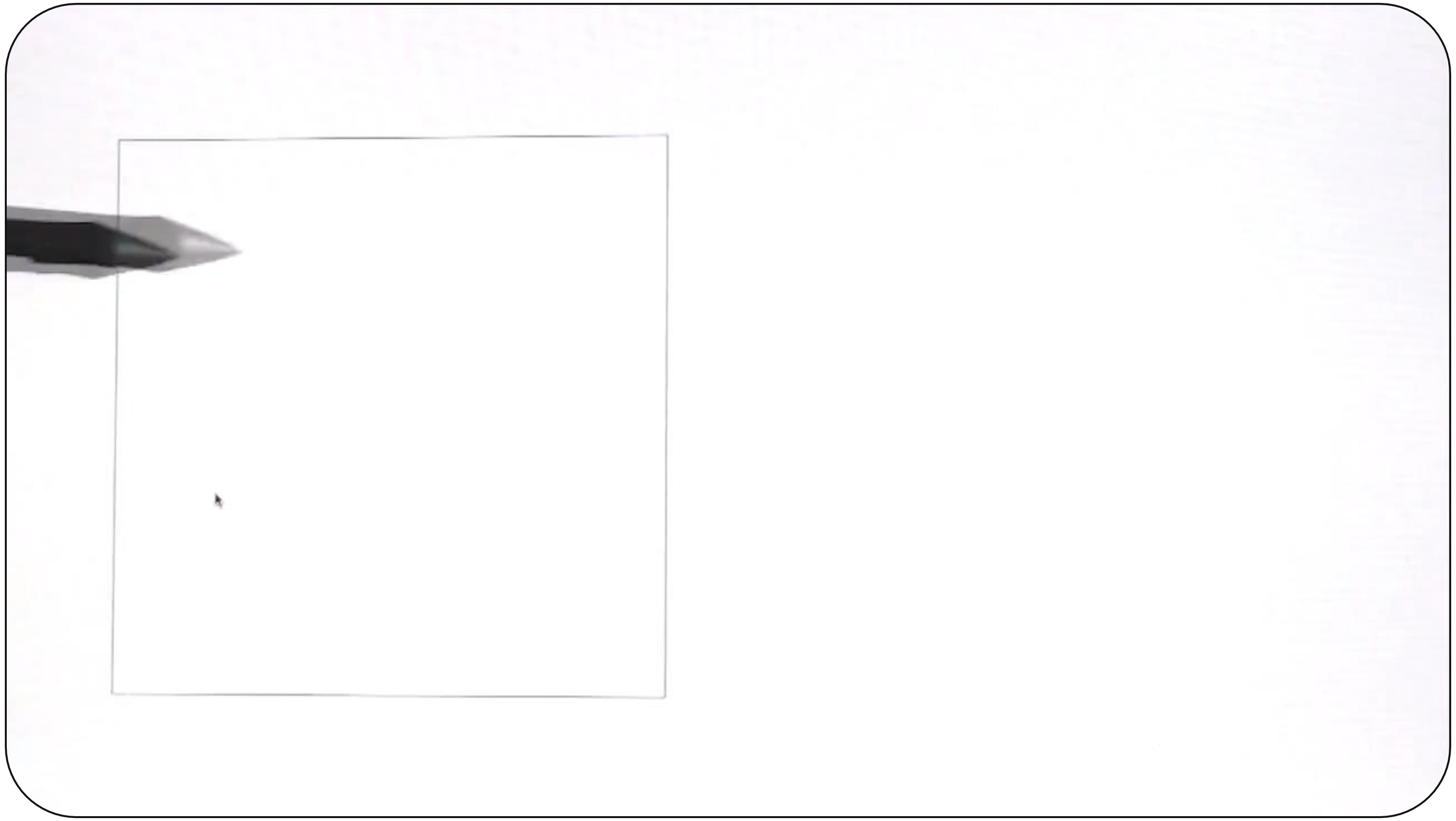
https://glassboxmedicine.com/2019/02/23/measuring-performance-auc-auroc/

Computational Sketch recognition

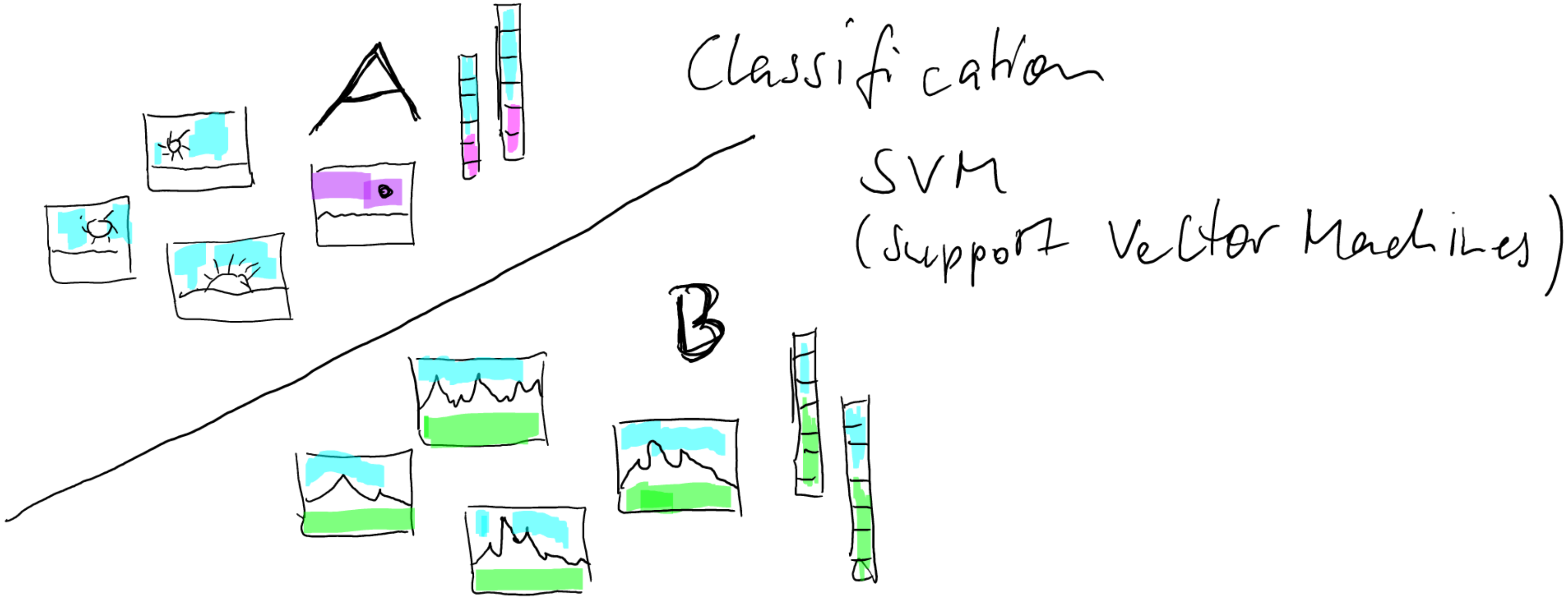








Support Vector Machines



Code a short demo

What if?

We could train a ***neural network*** to create image features that we can easily match.

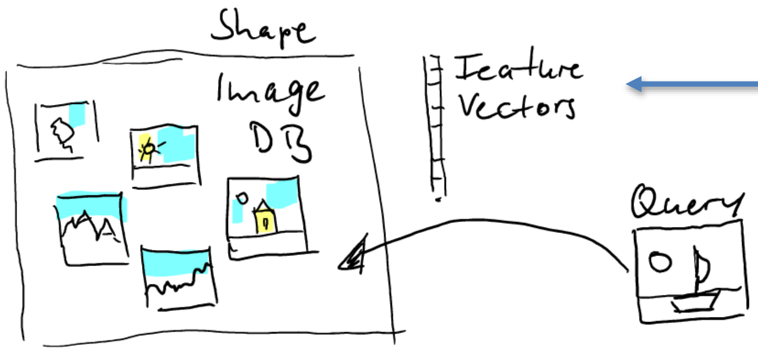
Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Neural Information Processing Systems (2012)

Draw me something

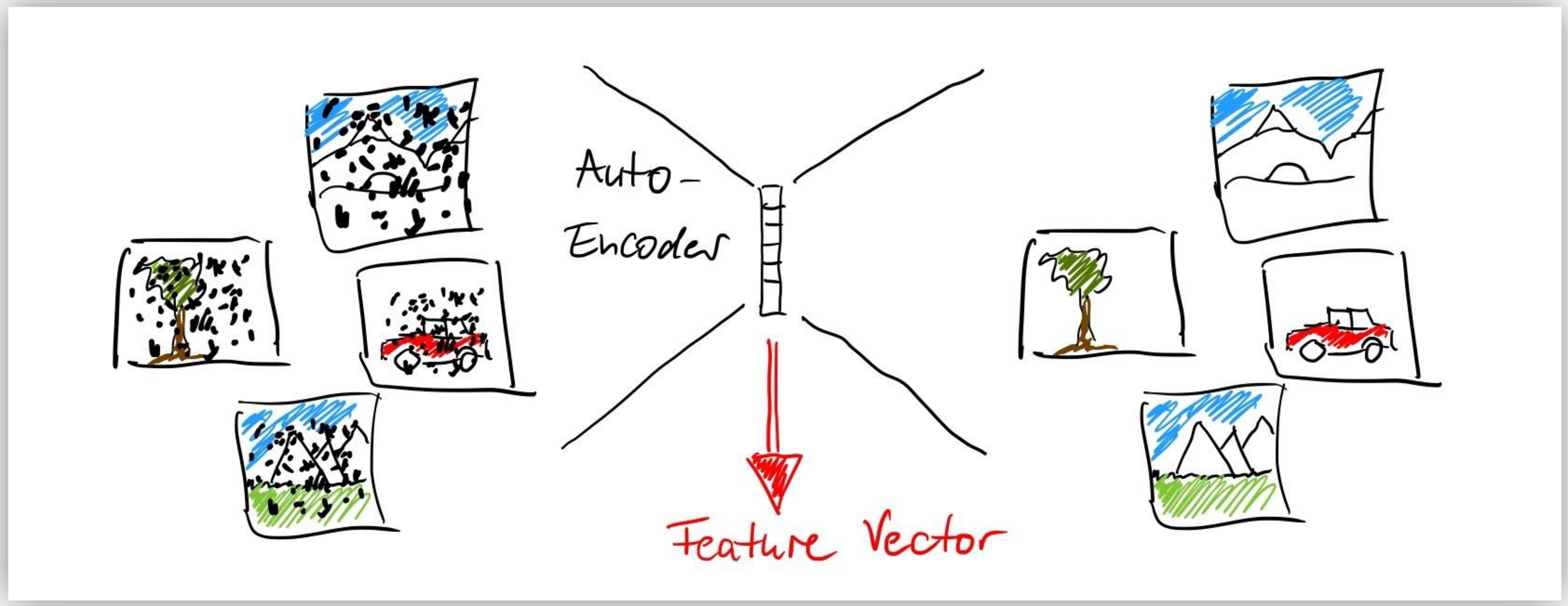
and I'll tell you, what I think it is!

Start again!

Content-based Image Retrieval using Convolutional Neural Networks



Same idea BUT:
Can we learn features?



Neural network introduction

- Watch all four videos! https://www.youtube.com/playlist?list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi