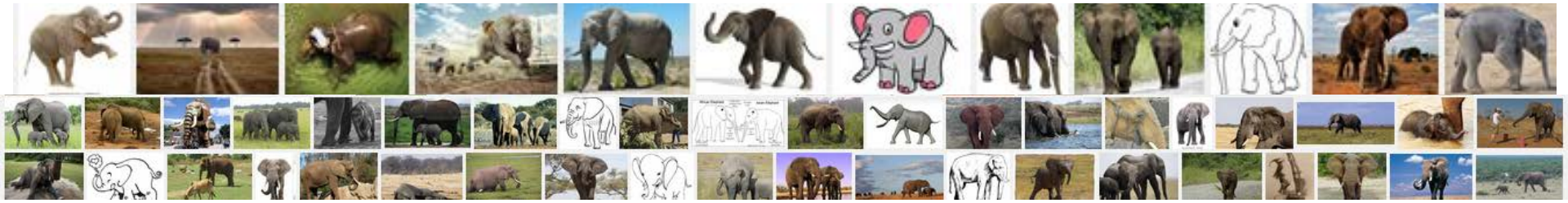


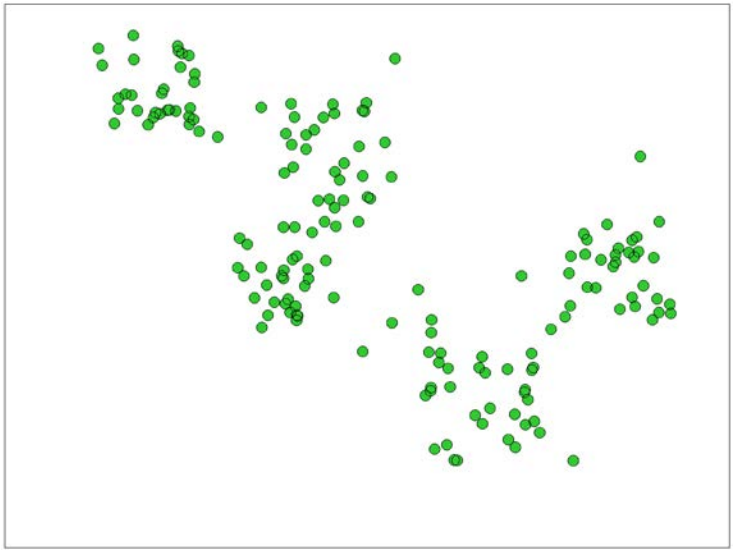
Geometric reasoning in Machine Learning

Daniel Cohen-Or

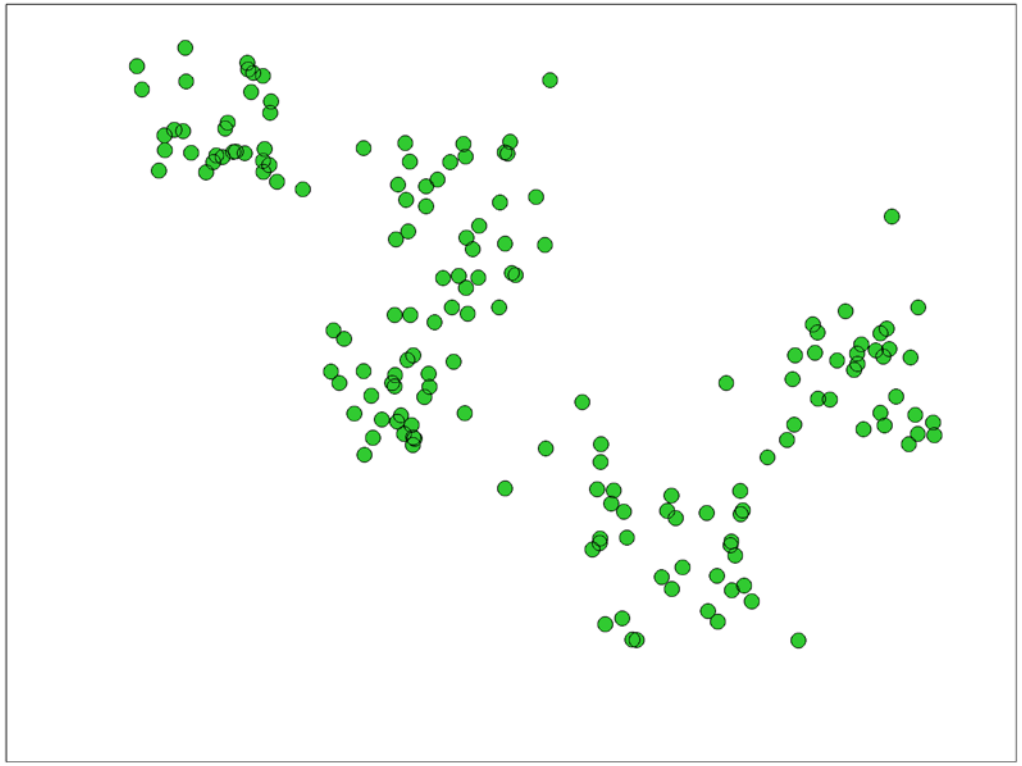
TEL AVIV UNIVERSITY
אוניברסיטת תל-אביב

Clustering is not easy…

- Clean separation to groups not always possible
- Must make “hard splitting” decisions
- Number of groups not always known, or can be very difficult to determine from data

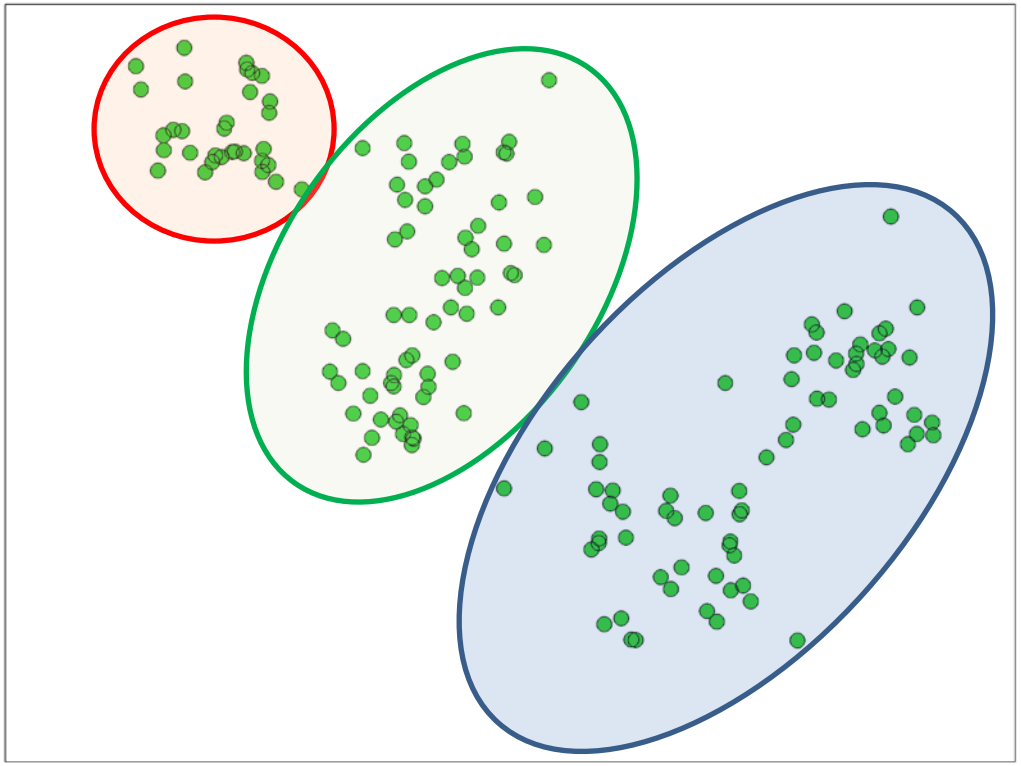


Clustering is hard!



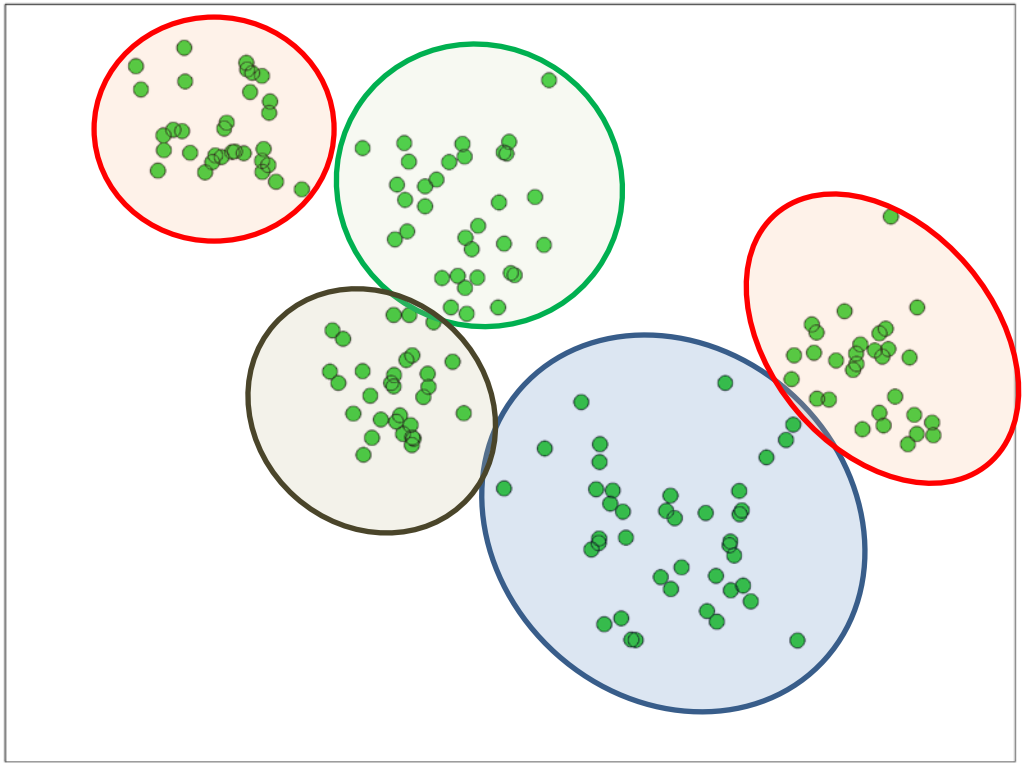
Clustering is hard!

Hard to determine number of clusters



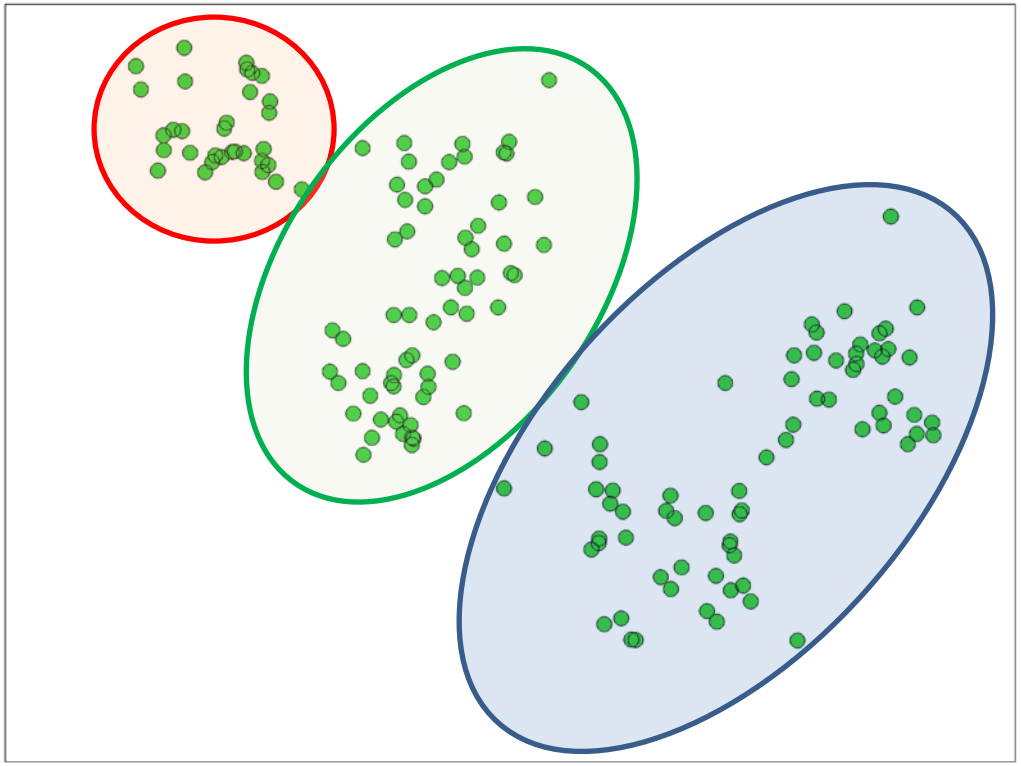
Clustering is hard!

Hard to determine number of clusters



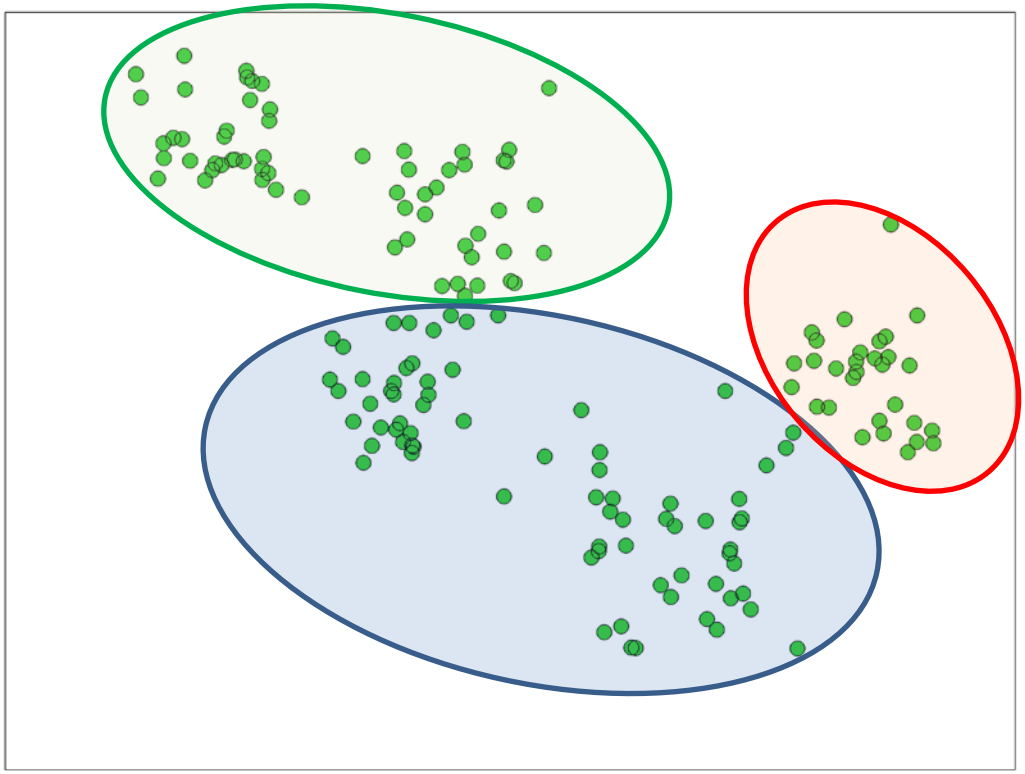
Clustering is hard!

Hard to decide where to split clusters

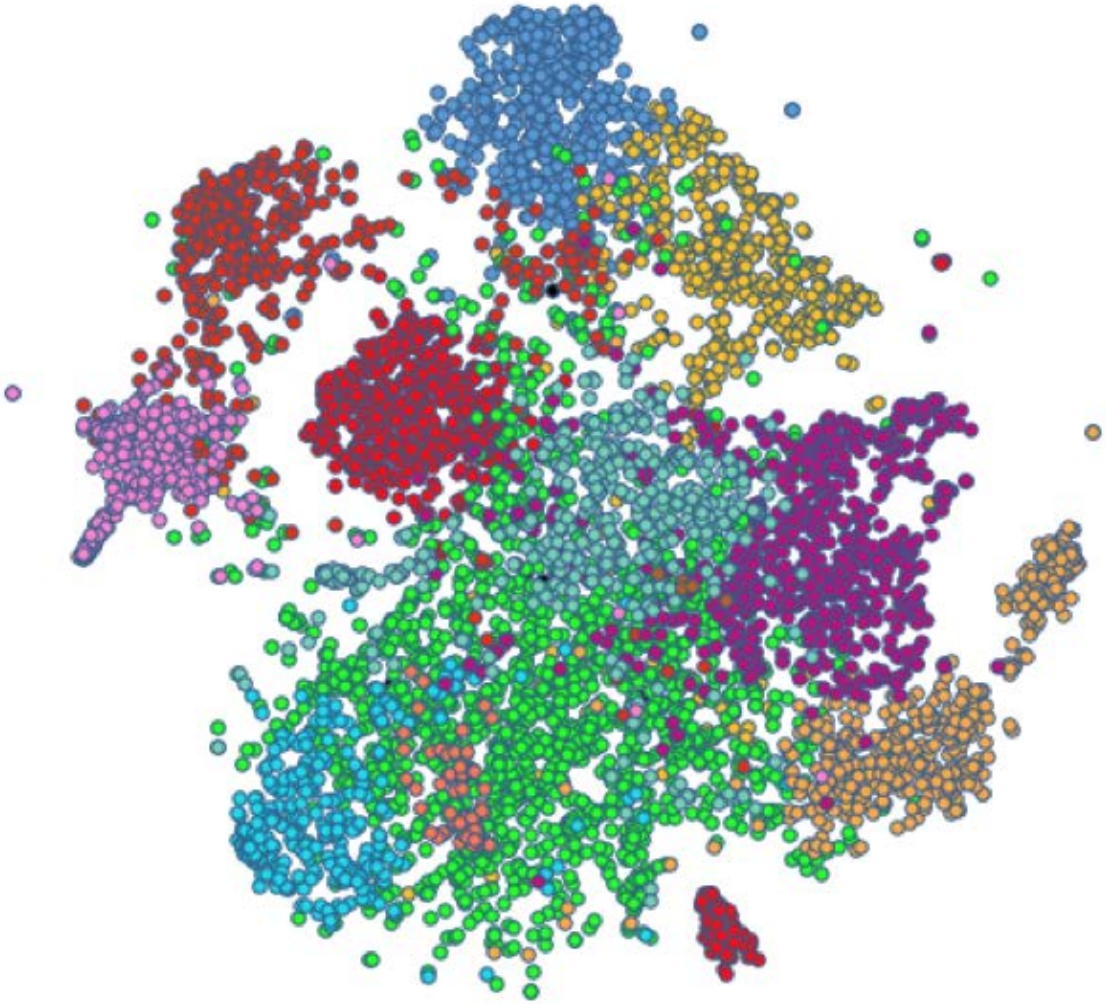


Clustering

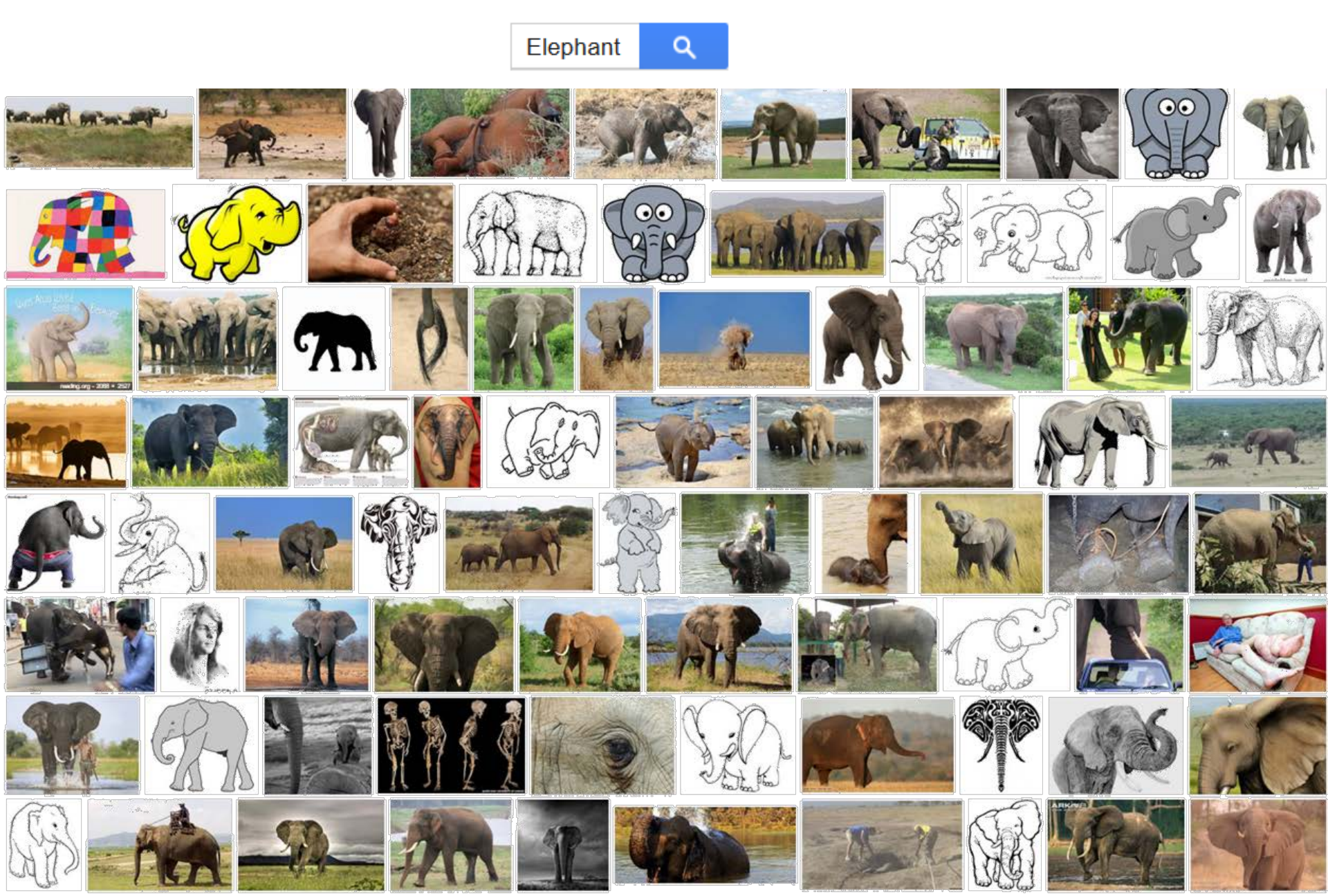
Hard to decide where to split clusters

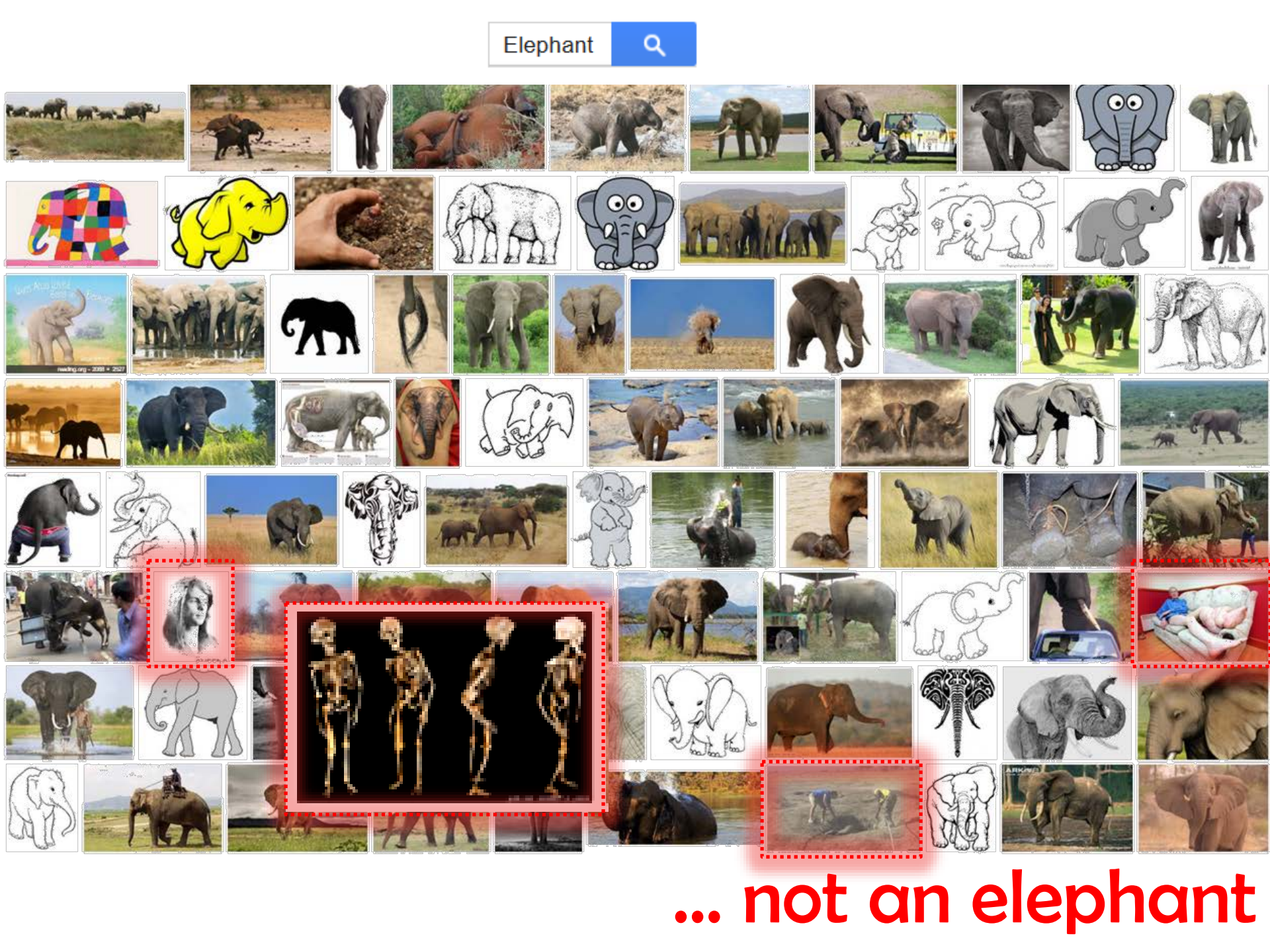


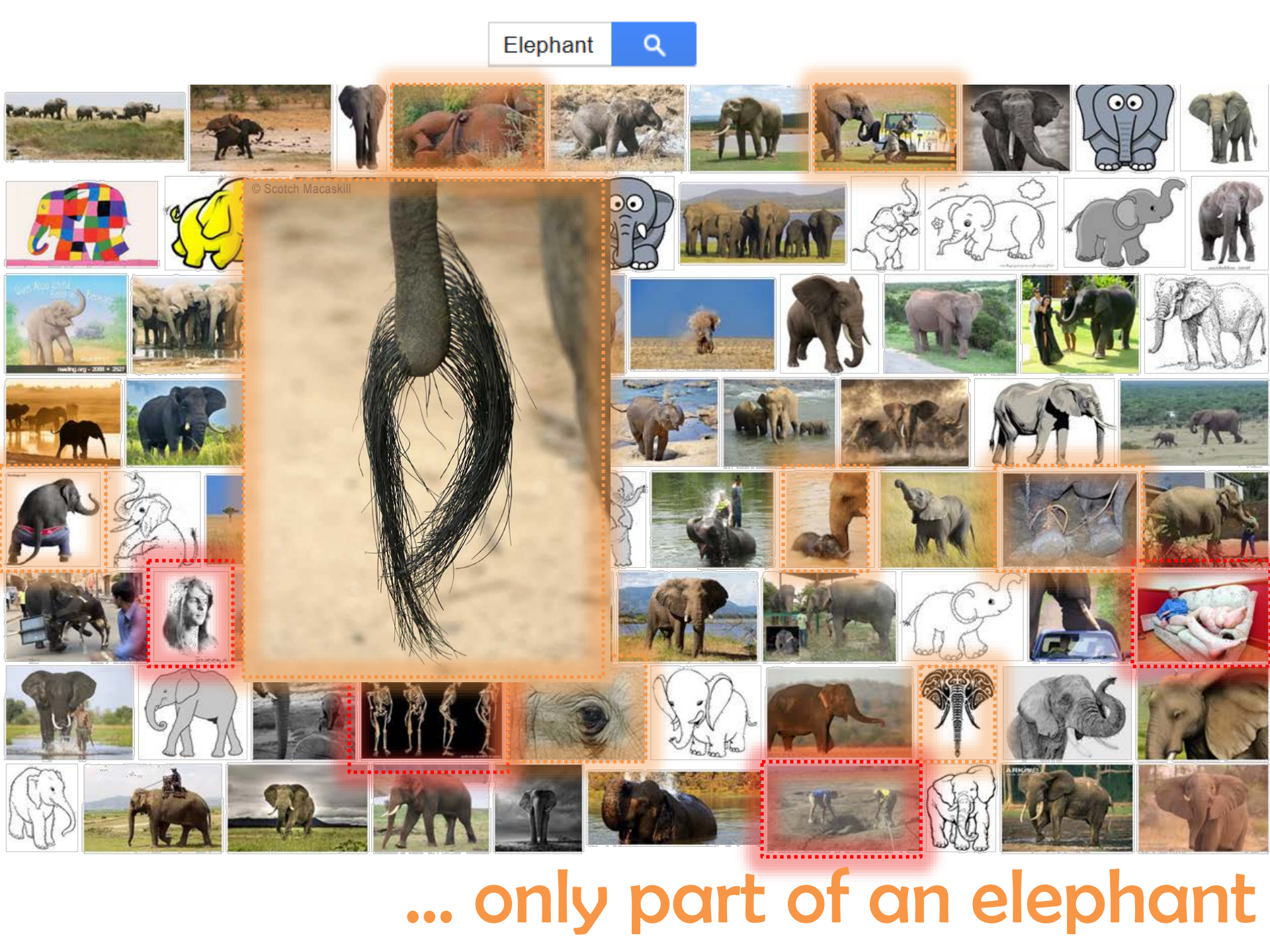
Deep Features

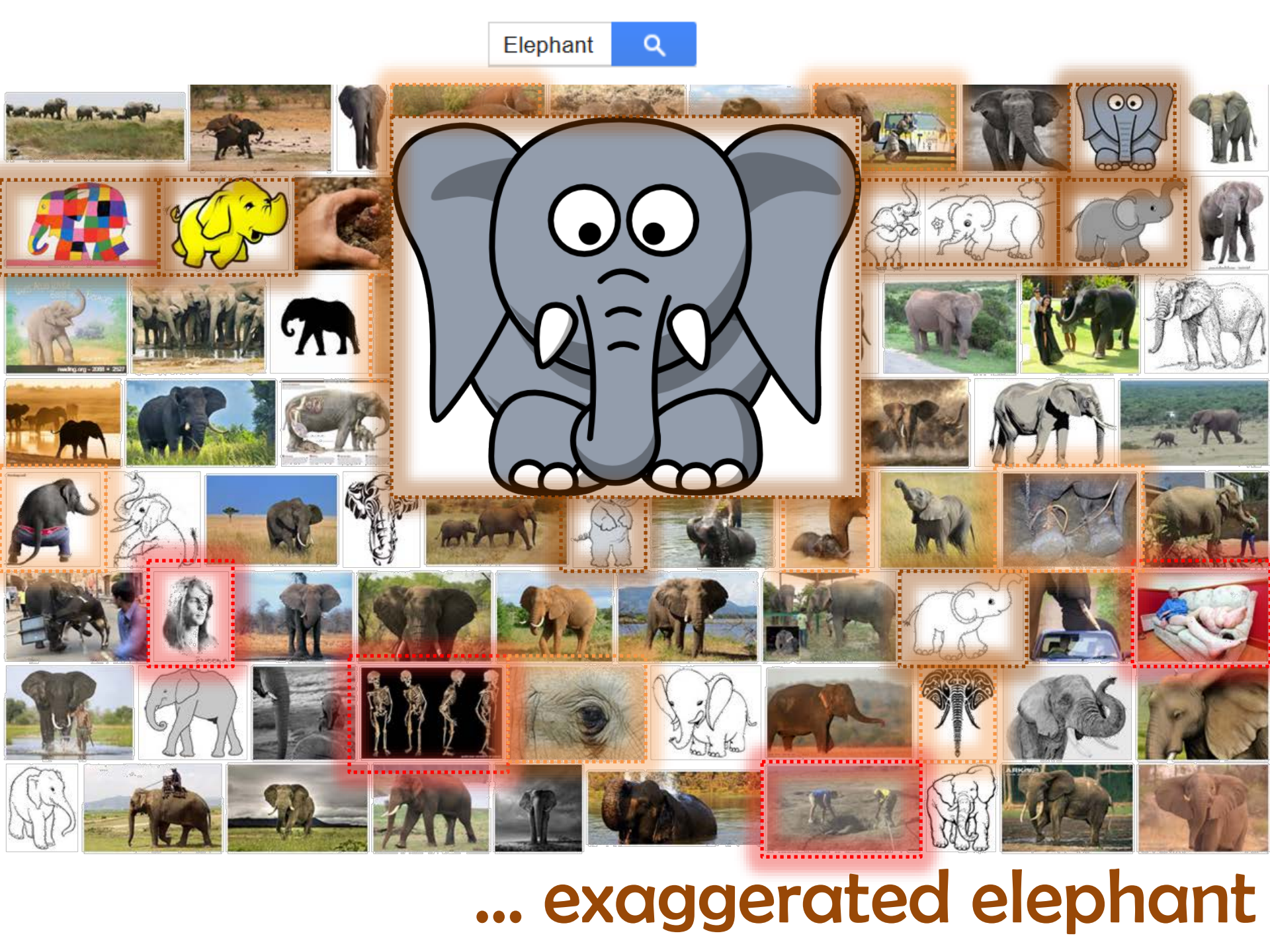


Shape-from-Text

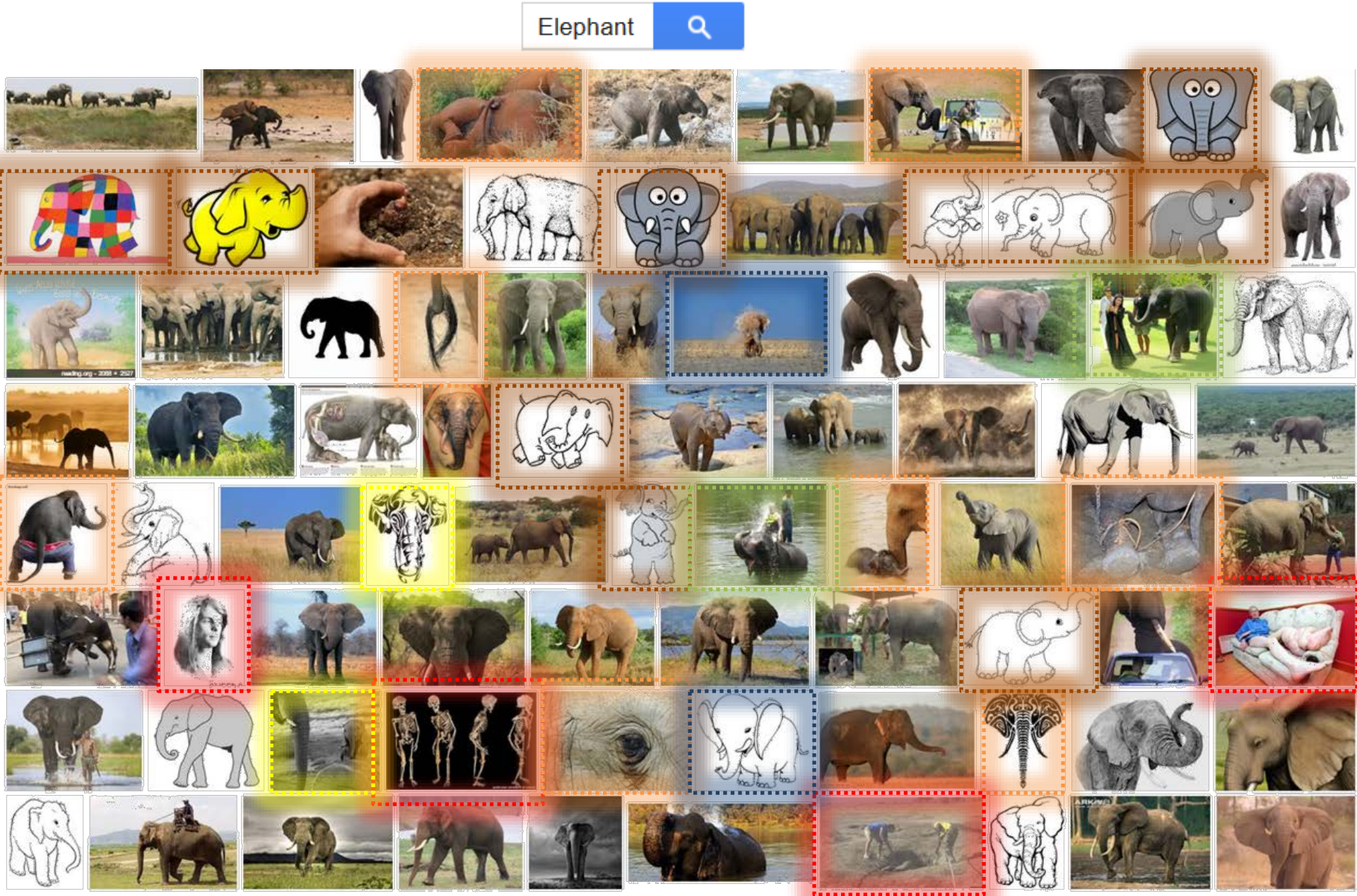

Elephant


Elephant
... not an elephant


Elephant
© Scotch Macaskill
... only part of an elephant



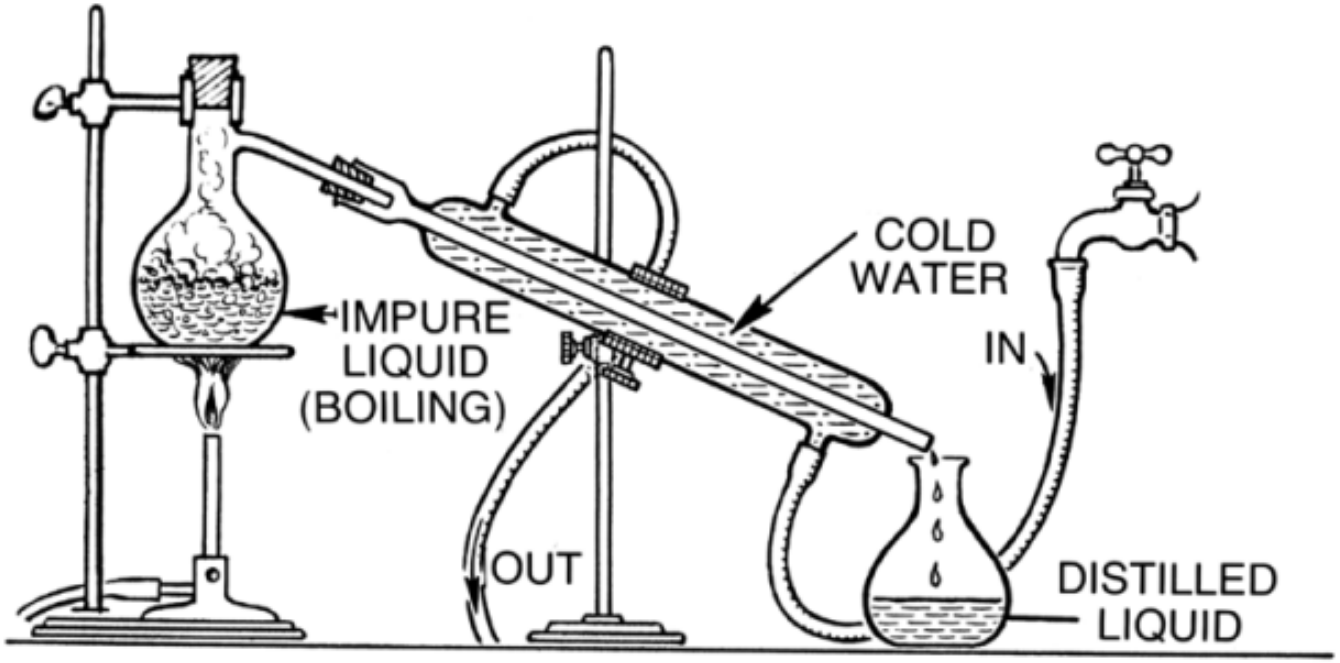
... exaggerated elephant



what is an elephant?

Distillation:

A process of separating the component substances from a liquid mixture.



Distillation:

true object

A process of separating the ~~component substances~~ from ~~a liquid mixture.~~

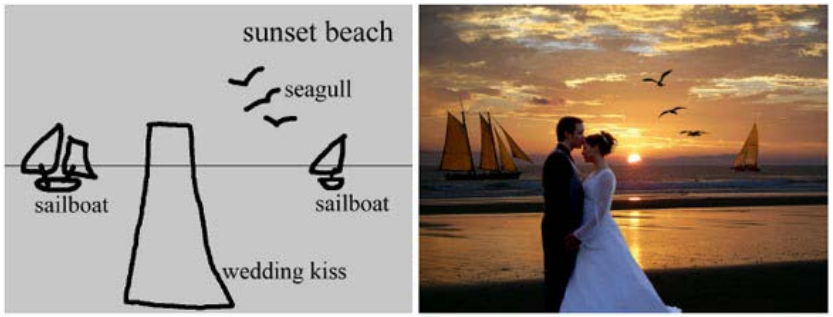
the raw image collection in an unsupervised fashion.

Precision-Recall

- In distillation, we care just for precision!
- We aim at being outlier-free.

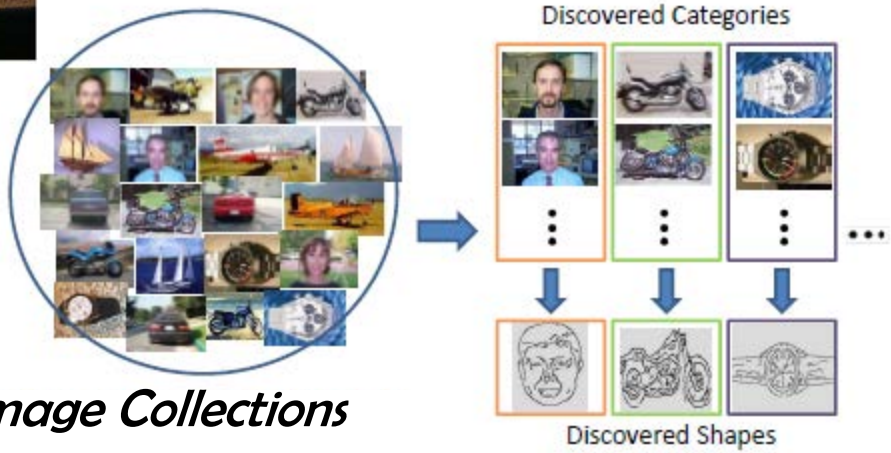
Related works

Supervised Filtering Methods



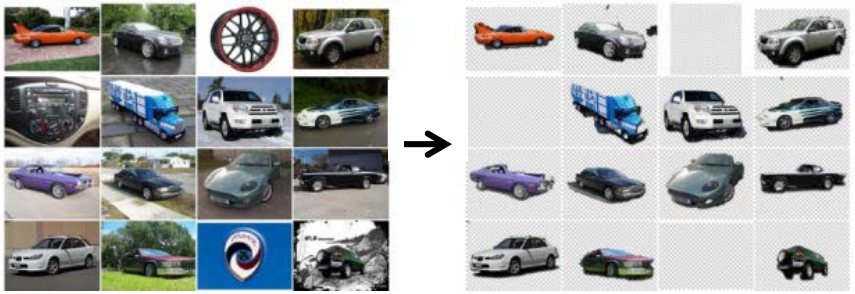
Sketch2Photo: Internet Image Montage
T. Chen, M.M. Cheng, A. Shamir, S.M. Hu

Unsupervised Object Discovery



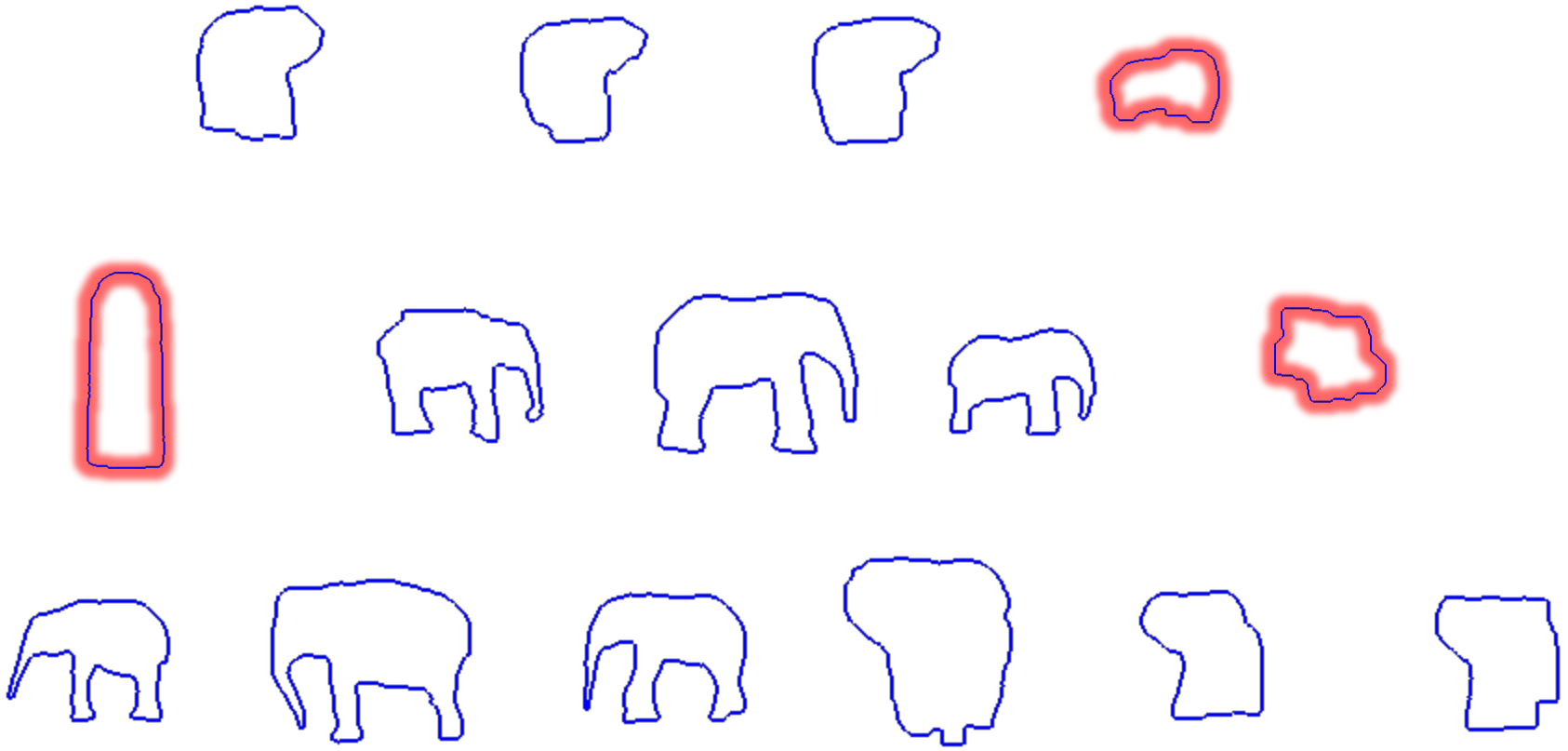
Shape Discovery from Unlabeled Image Collections
Y. J. Lee, K. Grauman

Co-segmentation



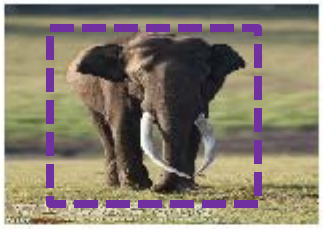
Unsupervised Joint Object Discovery and Segmentation in Internet Images
M. Rubinstein, A. Joulin, J. Kopf, C. Liu

The Guiding Principal



Outlier shapes are random in nature, while *inlier* shapes tend to be well supported by similar shapes from other images.

Shape Representation



Grabcut: Interactive foreground extraction using iterated graph cuts

C. Rother, V. Kolmogorov, A. Blake

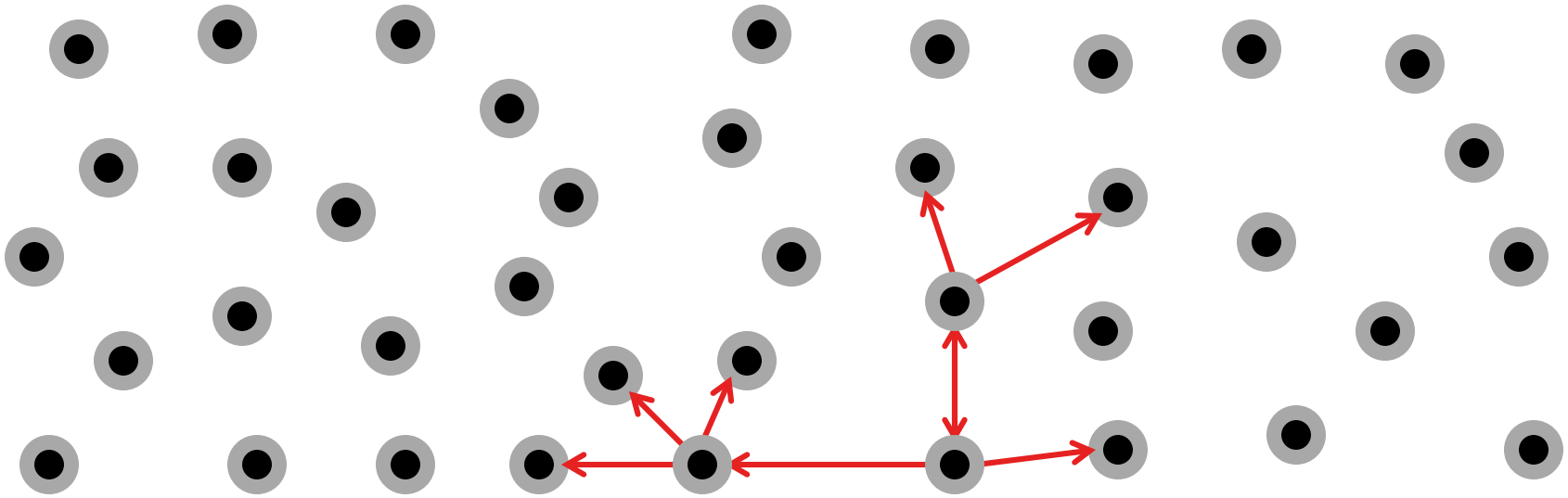
Measuring the objectness of image windows

B. Alexe, T. Deselaers, V. Ferrari

Mutual KNN Graph

An *inlier* shape belongs to a tight, non-trivial cluster.

Mutual KNN Graph

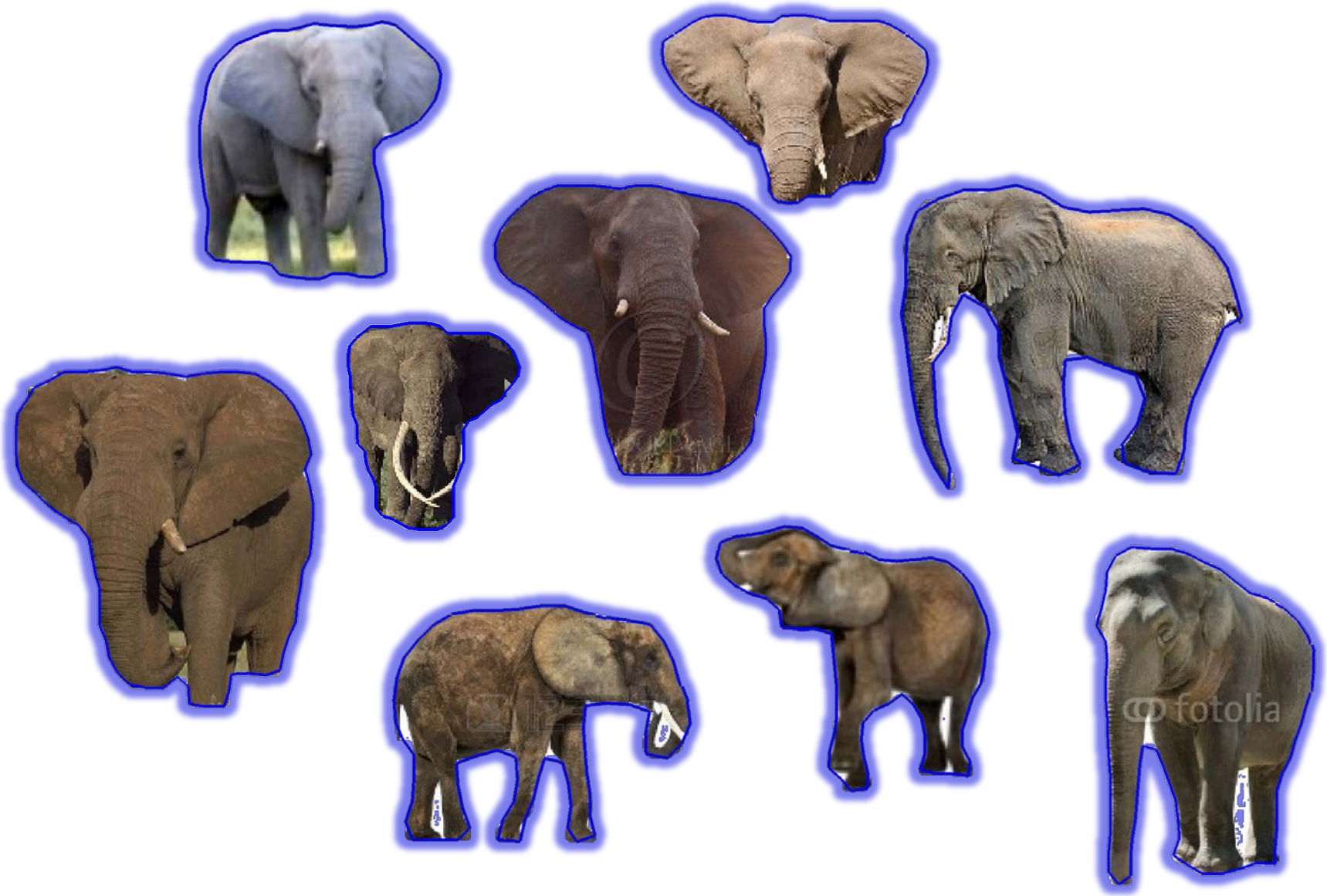


Mutual KNN Graph

Choose clusters that are

tight and visually informative.

Distilled Collection

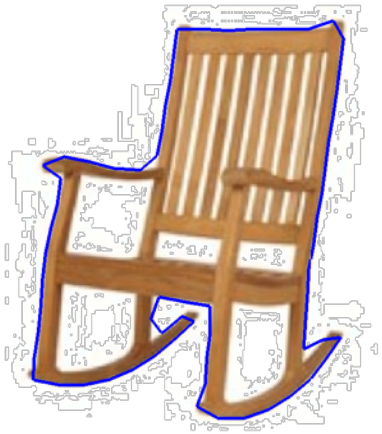


Input Collection (“rubber duck”)

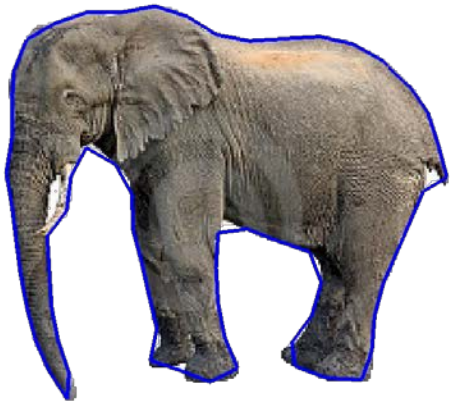
Distilled Collection (“rubber duck”)

Input Collection (“rocking chair”)

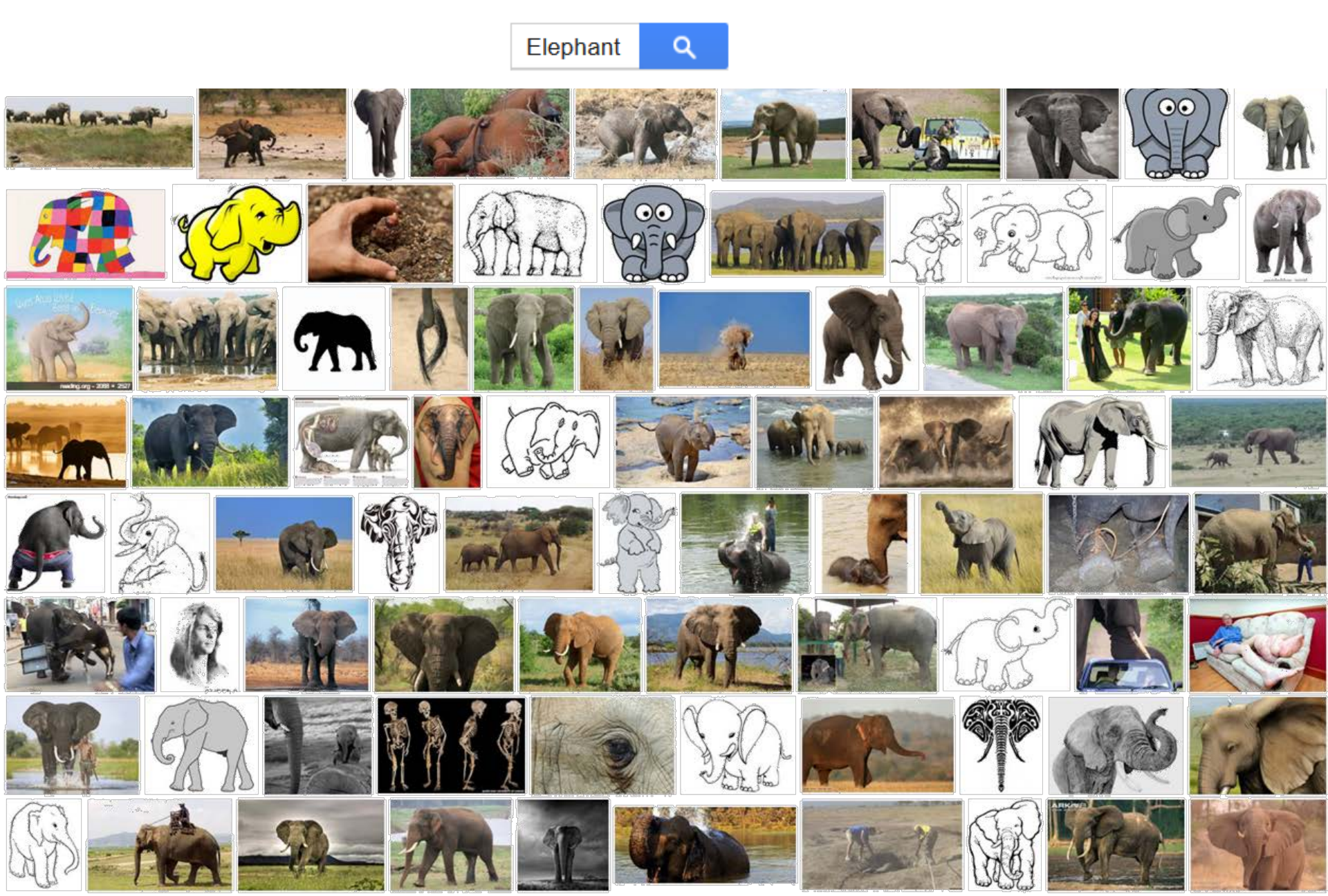
Distilled Collection (“rocking chair”)



Outliers are random, inliers are not!

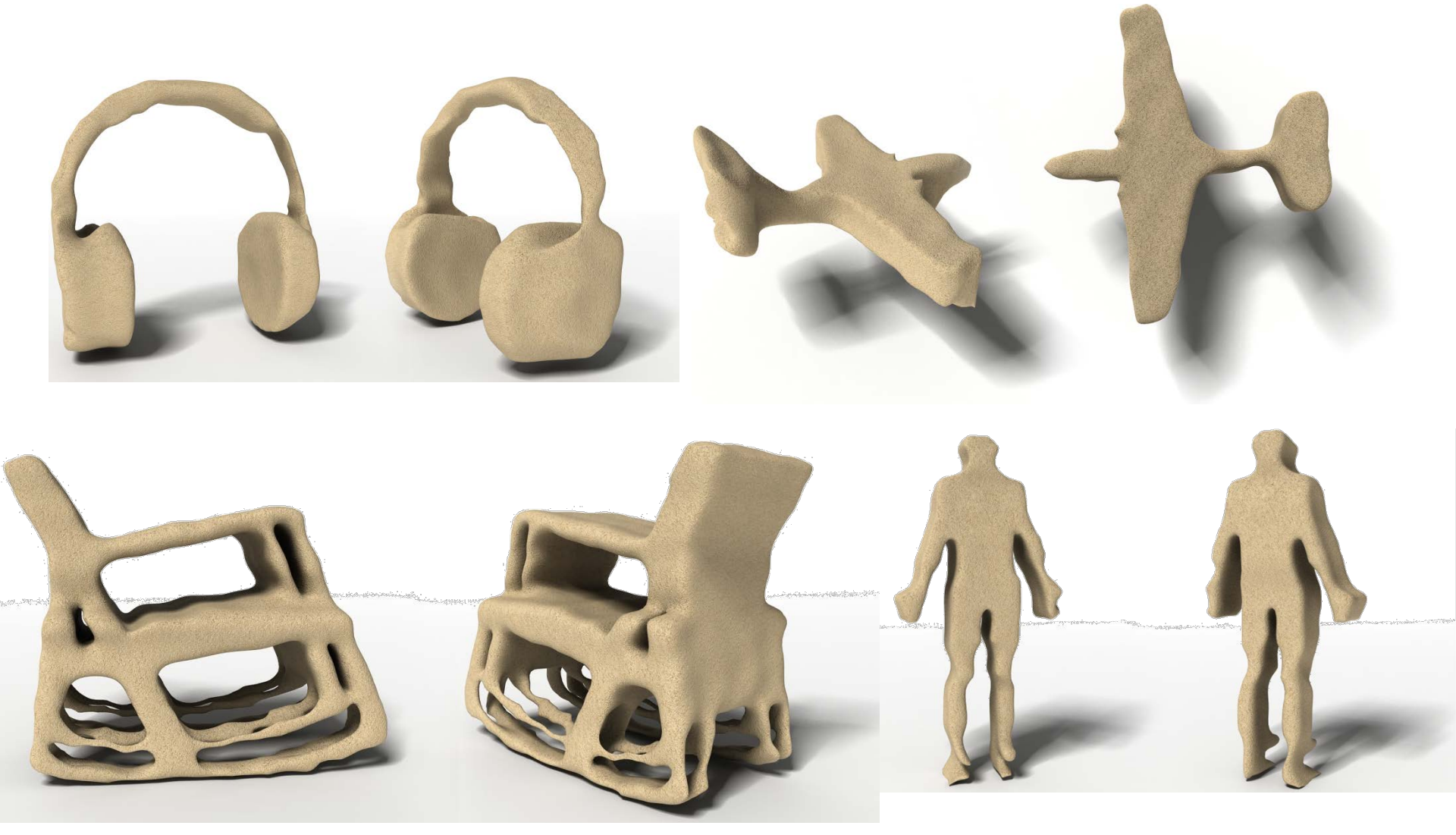


Shape-from-Text

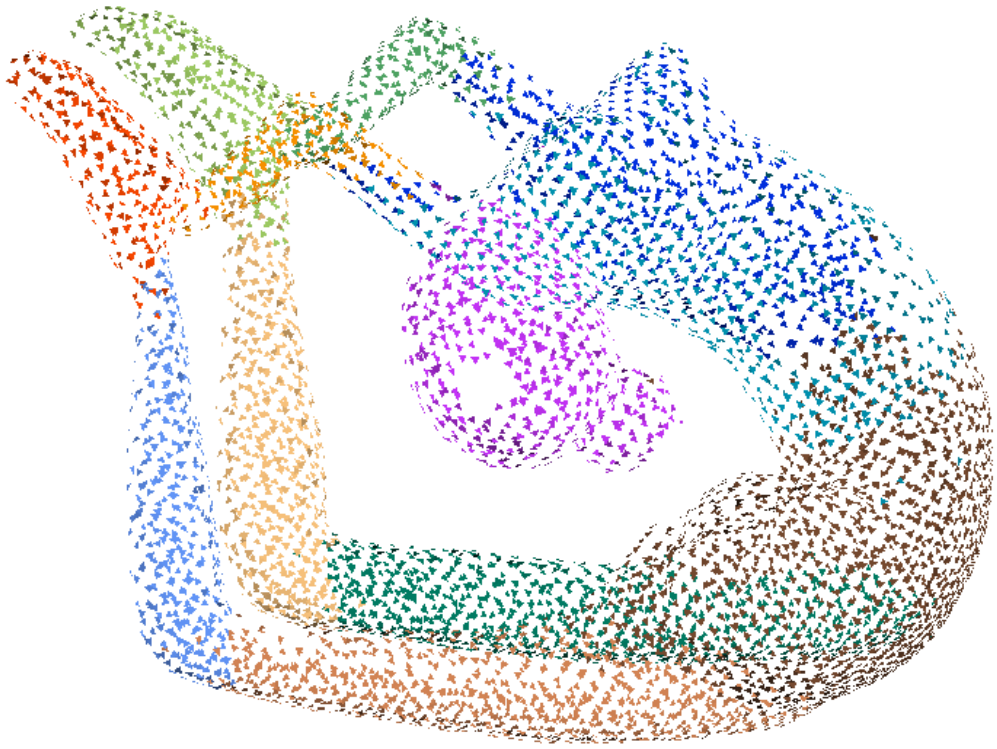

Elephant

Shape-from-Text

Shape-from-Text

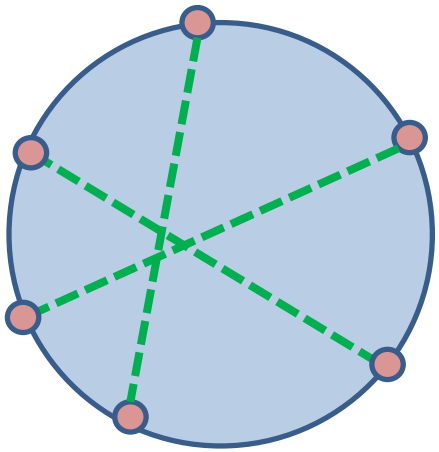


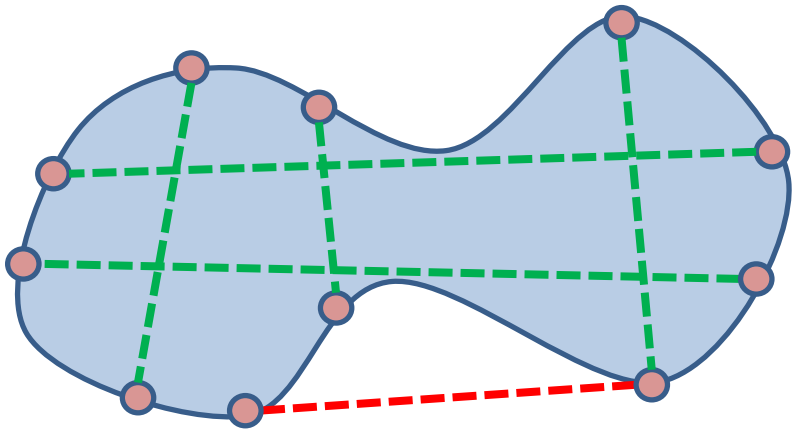
Convex Decomposition by Clustering [SGP13]



Recap of (Exact) Convexity

A convex set of points contains all line segments between each pair of points

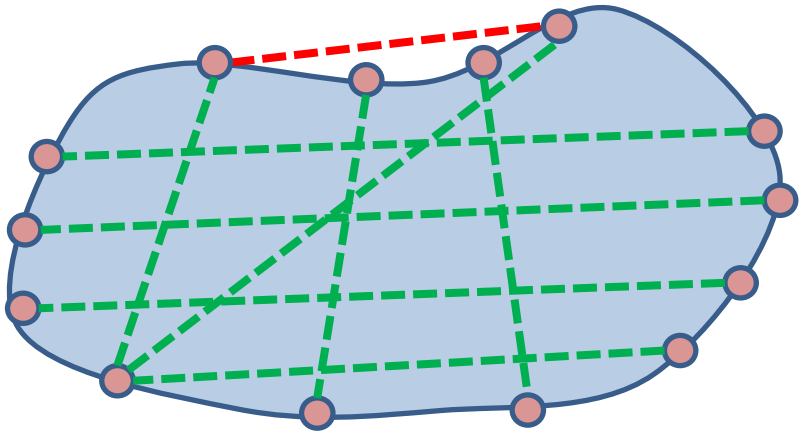




No line of sight
(within the shape),
not convex

Almost Convex..

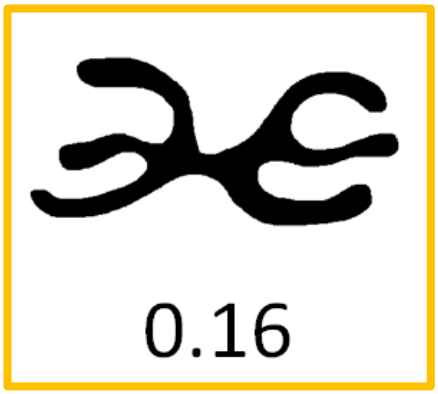
- Let's relax the definition of convexity, and define "Weak convexity" as:
  - **Most** of the point pairs on a segment boundary have a line of sight between them

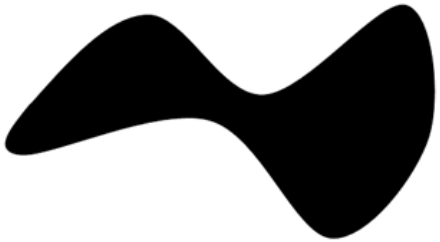


98% lines inside

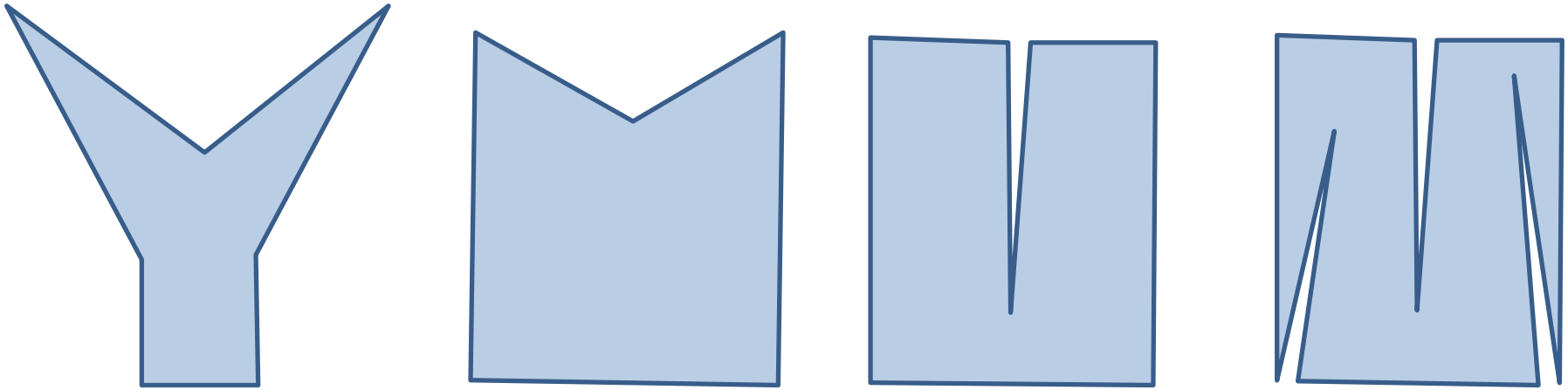
Weak Convexity

- Convexity rank provides an intuitive weak convexity measure:



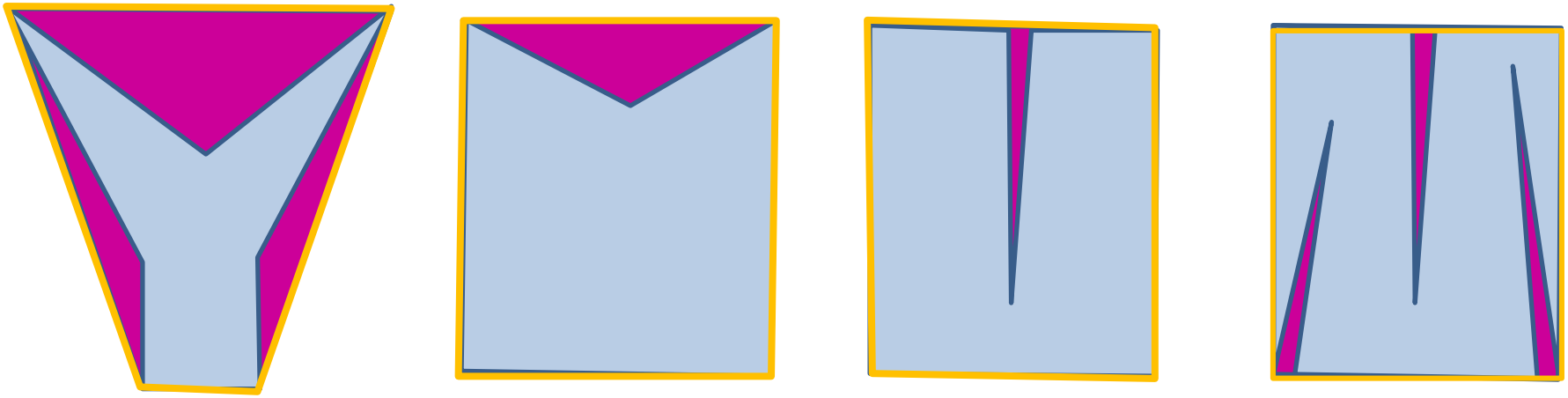


Previous Work



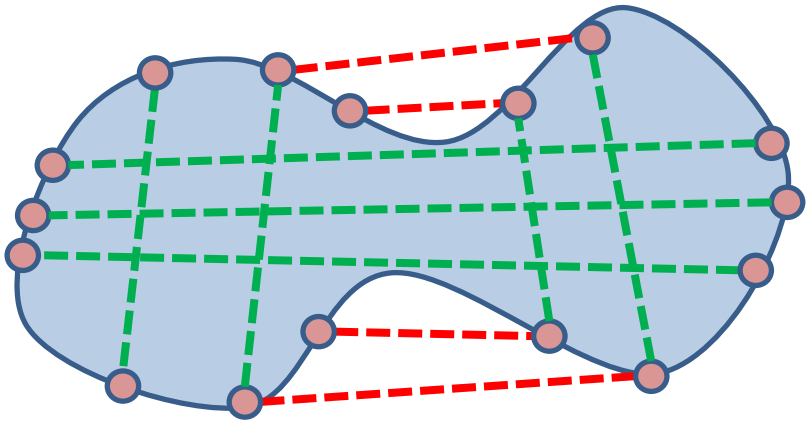
Previous Work

- Measure convexity with a convex hull

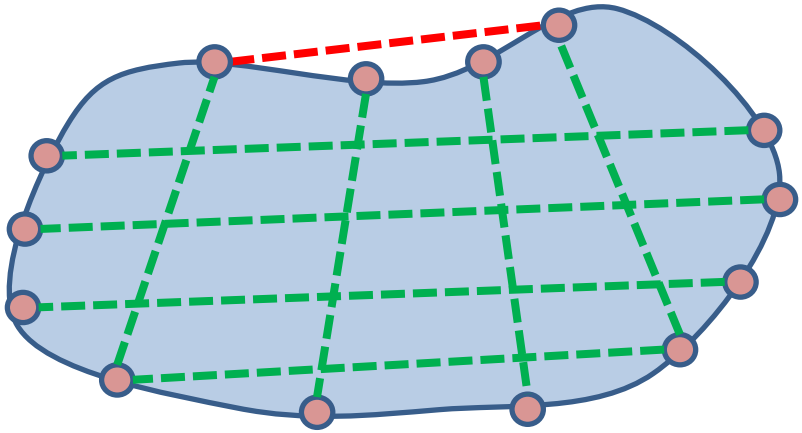


Weak Convexity

Measure the relative number of mutually visible points!



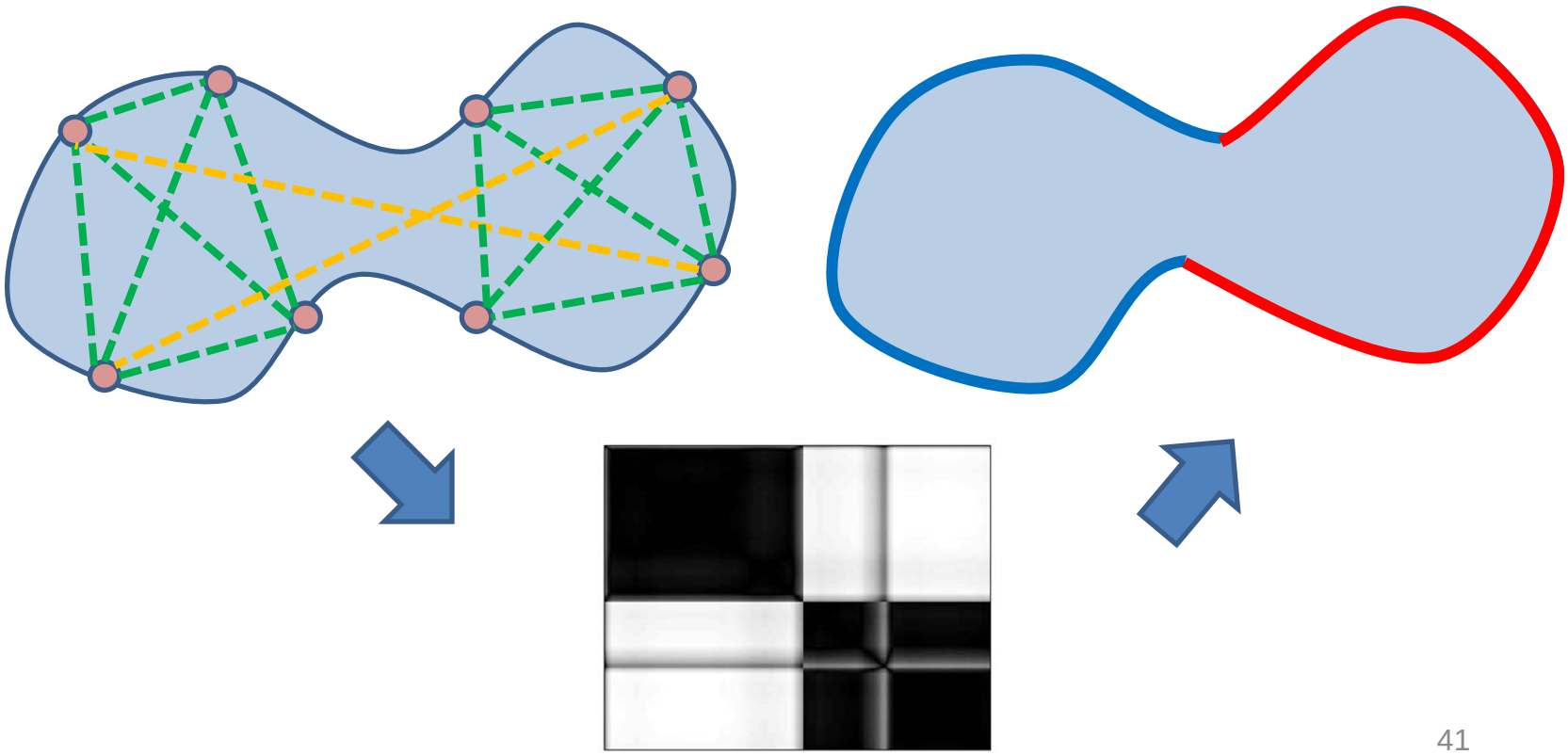
60% lines inside



90% lines inside

Weak Convex Decomposition

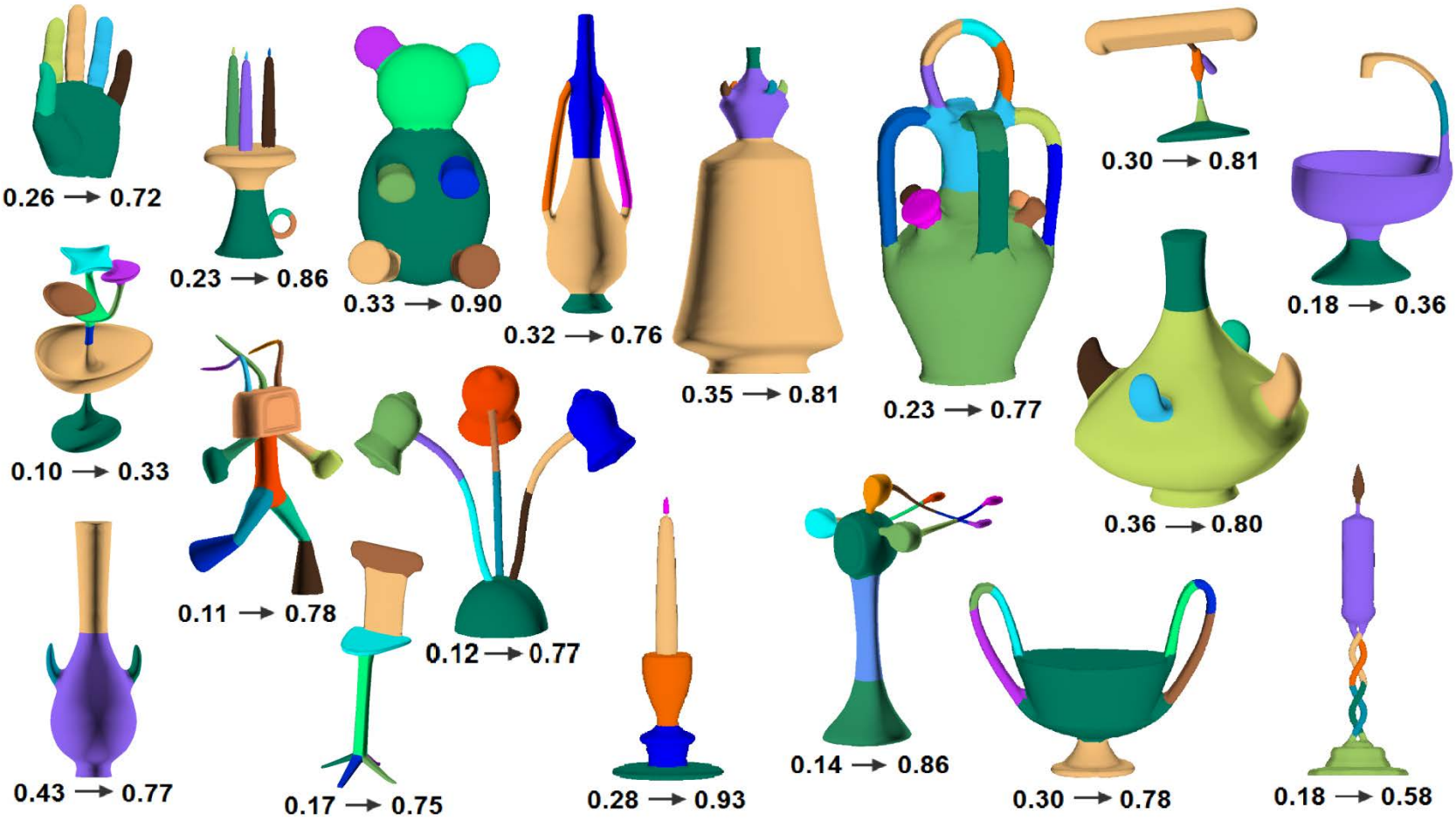
A cluster of mutually visible points is by definition a weakly convex component



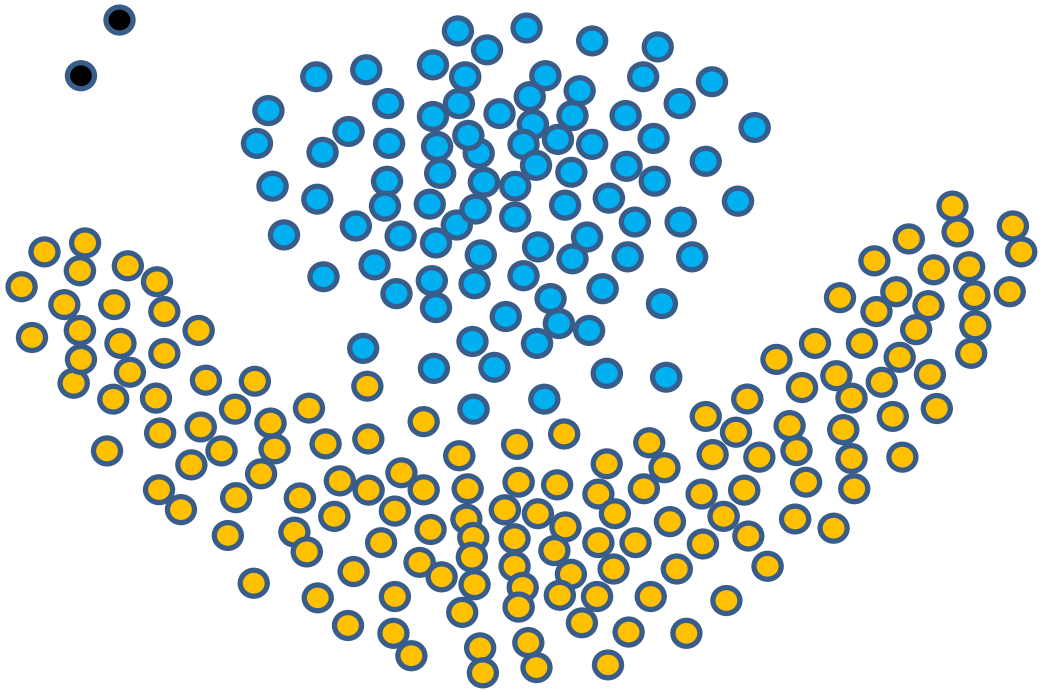
Convex Decomposition by Clustering

- We define the weak convexity rank as the degree of mutual visibility in the set
- The convex decomposition is solved by a clustering algorithm
- We avoid over-clustering using the notion of normalized cuts.

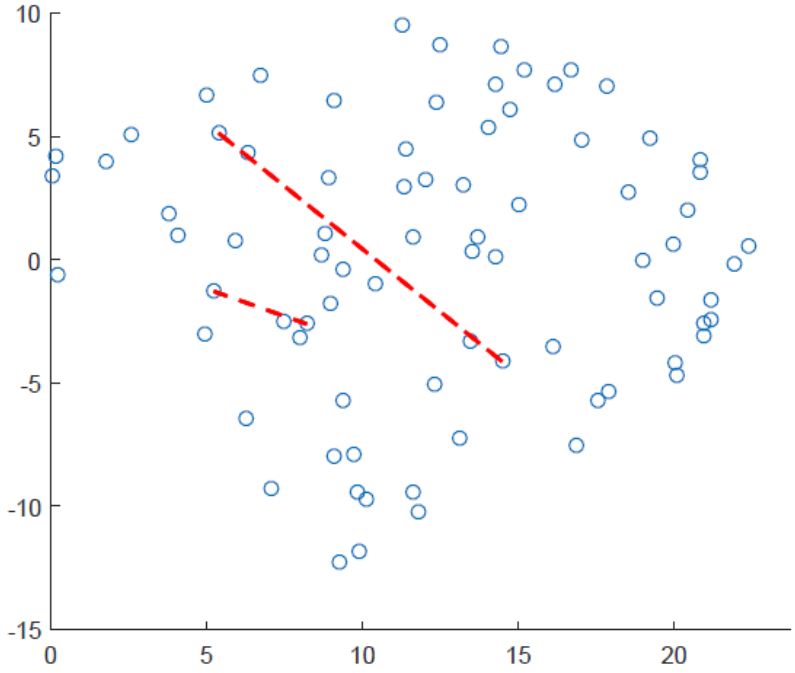
Results

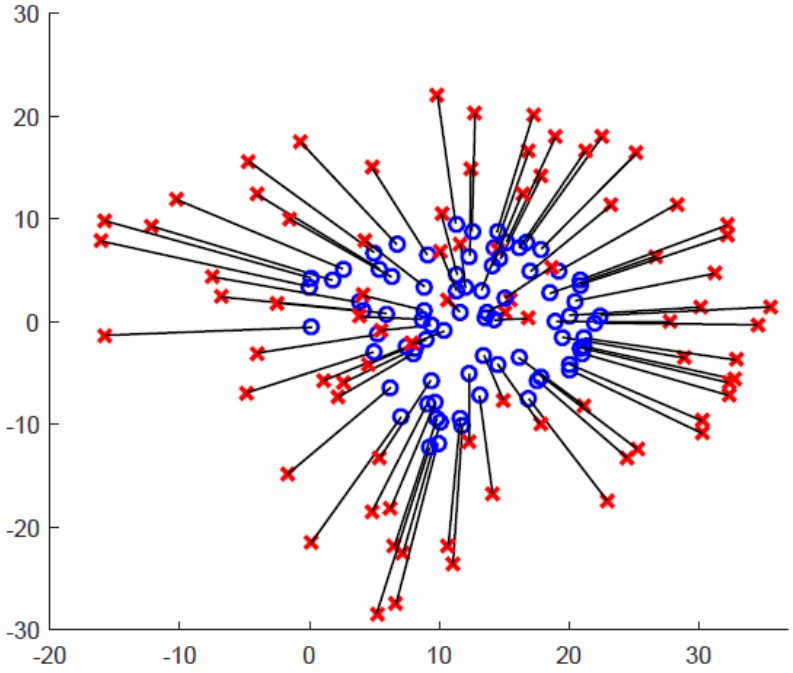

0.26 → 0.72
0.23 → 0.86
0.33 → 0.90
0.32 → 0.76
0.35 → 0.81
0.23 → 0.77
0.30 → 0.81
0.18 → 0.36
0.10 → 0.33
0.11 → 0.78
0.12 → 0.77
0.36 → 0.80
0.43 → 0.77
0.17 → 0.75
0.28 → 0.93
0.14 → 0.86
0.30 → 0.78
0.18 → 0.58

How to cluster?

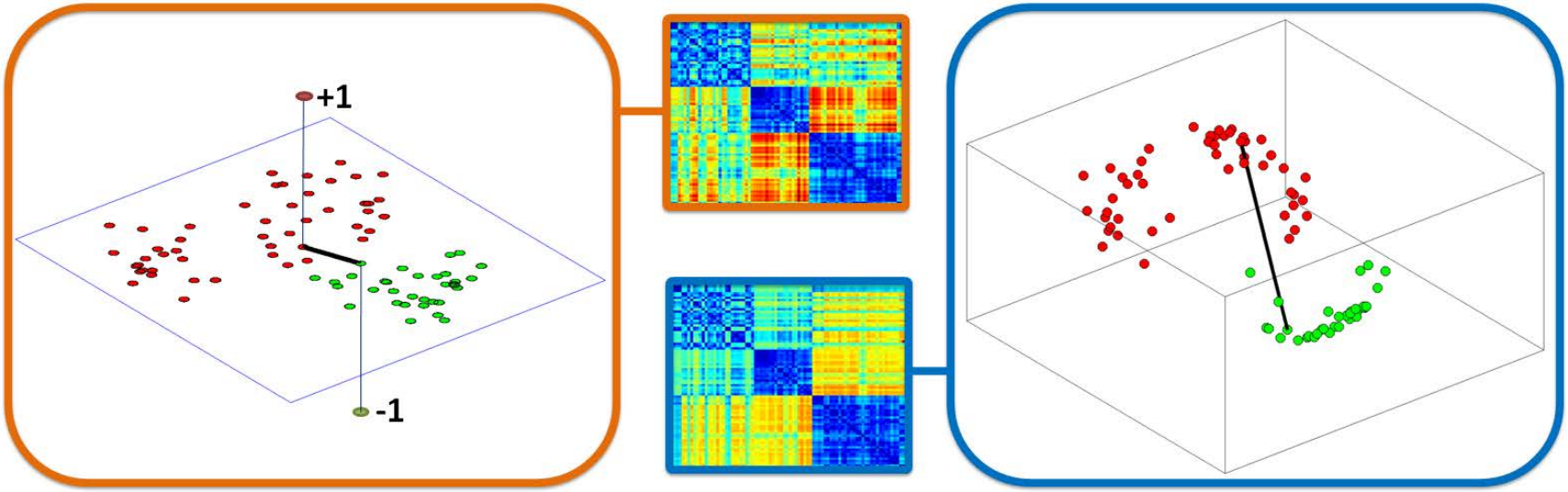


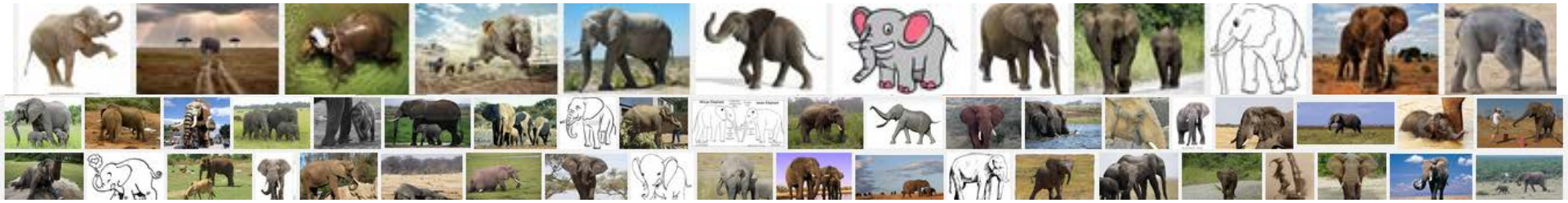
Robust Multi-dimensional Scaling





Constraints as Features





Geometric reasoning in Machine Learning

Daniel Cohen-Or

TEL AVIV UNIVERSITY
אוניברסיטת תל-אביב