Partitioning of Image Datasets using Discriminative Context Information

Christoph H. Lampert

Max-Planck-Institute for Biological Cybernetics, Tübingen, Germany

Overview

Summary

- **Discriminative Context Partitioning (DCP)** is a new unsupervised method to partition a dataset.
- It splits the dataset such that the resulting parts are best separated from a disjoint *context class*.
- DCP is not *clustering*. The parts are not determined by *peaks* in the sample density, but purely *discriminatively*.
- For suitable context, DCP is more robust than clustering methods.
- By varying the context, one can explore different partitionings.

How to split a unimodal dataset into meaningful parts?

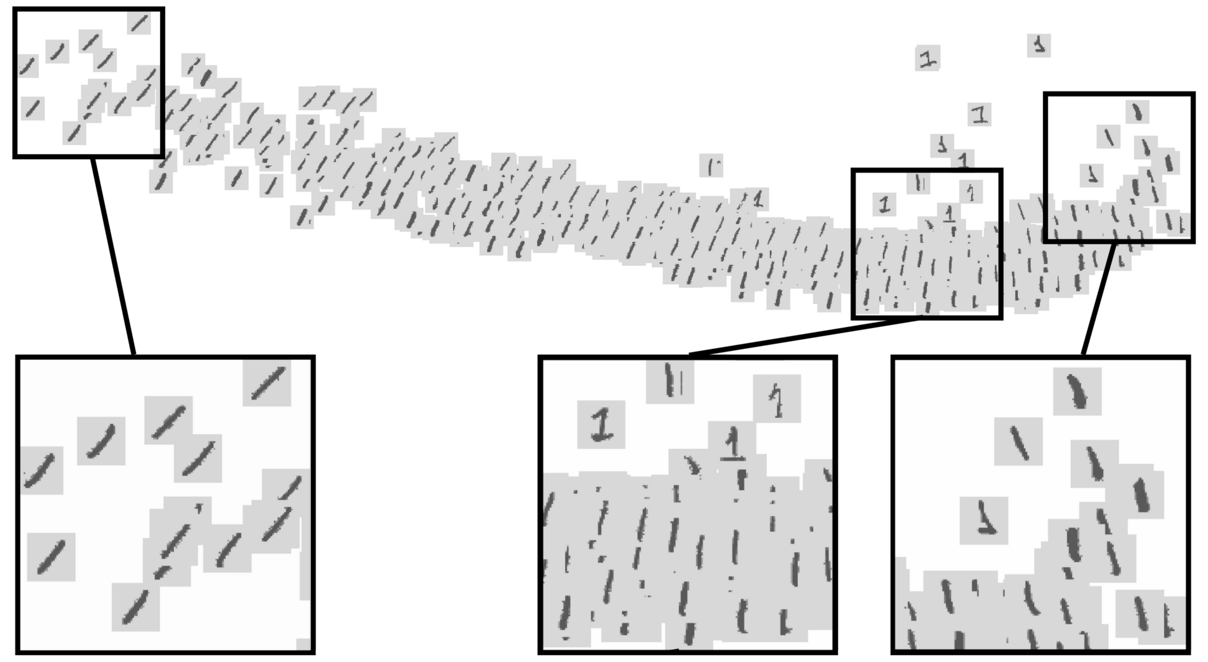
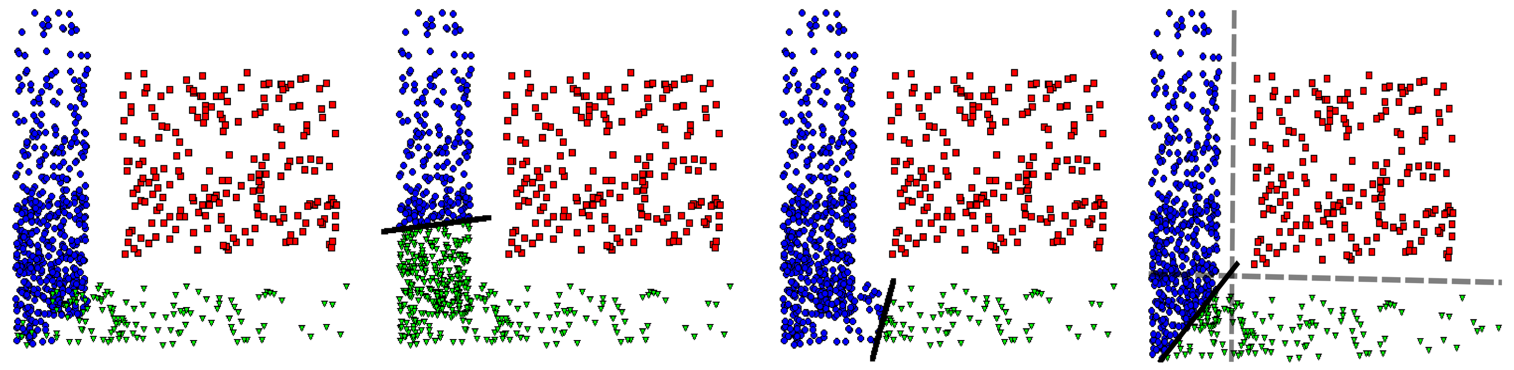


Figure: 2D-PCA projection of MNIST digit 1

Use separation from a geometric context to distinguish parts



(a) ground truth (b) *K*-means (c) spectral clst. (d) DCP

Figure: *L*-shaped dataset consisting of 2 overlapping parts (○,▽). *DCP* with access to a *context class* (□) identifies the intended parts. *K*-means and spectral clustering without context yield worse results.

Method

A measure of separation between sets

For two disjoint sets X, Z and a decision hyperplane $f \in \mathcal{H}$, we measure the *separation* between the sets by the *negative of the SVM objective* function:

$$sep_f(X,Z) := -\frac{1}{2}\|f\|^2_{\mathcal{H}} - \sum_{x\in X}\ell(1-f(x)) - \sum_{z\in Z}\ell(1+f(z)), \qquad (1)$$

where ℓ is a monotonous convex loss function that penalizes margin violations, e.g. the hinge loss or the quadratic loss.

The most discriminative split of a sets

Let X be the dataset that we want to split. Let Z be is a disjoint context set. For $K \in \mathbb{N}$, let $X_1 \cup \dots \cup X_K = X$ be a decomposition of X. Then the *total separation* score of this split is

$$sep(X_1,\dots,X_K\,;\,Z) \quad := \quad \sum_{k=1}^{K} \max_{f\in\mathcal{H}}\; sep_f(X_k, Z). \qquad (2)$$

A decomposition $X_1^* \cup \dots \cup X_K^* = X$ is called a **most discriminative *K*-split of *X* with respect to *Z***, if it maximizes the total separation over all possible decompositions of X.

Theorem: Finding the most discriminative partitioning

The most *discriminative partitioning* of X with respect to Z is given by

$$X_k^* := \{\, x \in X \;:\; \operatorname*{argmax}_{k'=1,\dots,K}\; f^*_{k'}(x) = k \,\}, \qquad (3)$$

for $k = 1,\dots,K$, where $f_k^* \in \mathcal{H}$ minimizes

$$J(f_1,\dots,f_K) = \frac{1}{2}\sum_{k=1}^{K}\|f_k\|^2_{\mathcal{H}} + \sum_{z\in Z}\sum_{k=1}^{K}\ell(1+f_k(z)) + \sum_{x\in X}\ell(1-\max_k f_k(x)).$$

Numeric Solution

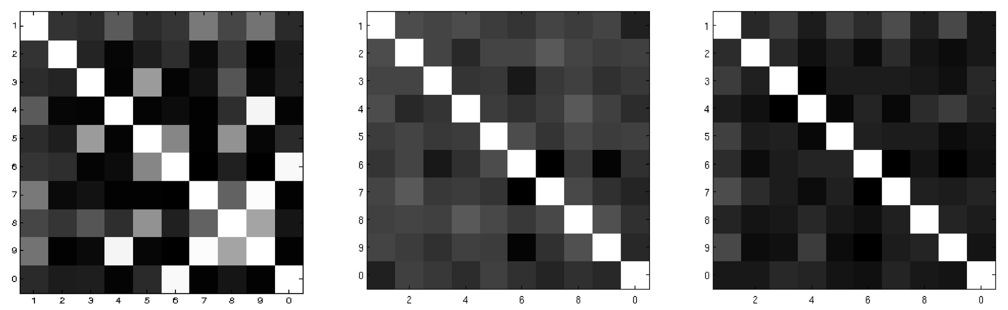
Several techniques are applicable to solve the optimization problem (3):

- *(Stochastic) gradient descent*
- Convex-Concave Procedure (CCCP)
- Deterministic Annealing

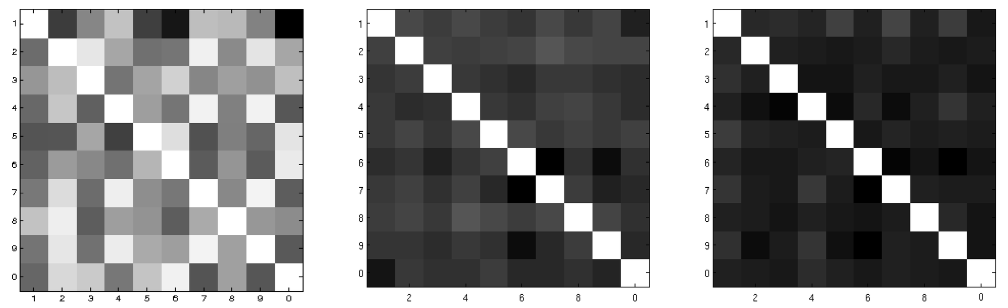
Experimental Results

Unsupervised separation of USPS digits

We create an image dataset by mixing two USPS digit classes in different ratios. How well can we recover the original partitioning?



(a) balanced mixture: KM (left), DCP-avg (center), DCP-best (right).



(b) unbalanced mixture: KM (left), DCP-avg (center), DCP-best (right).

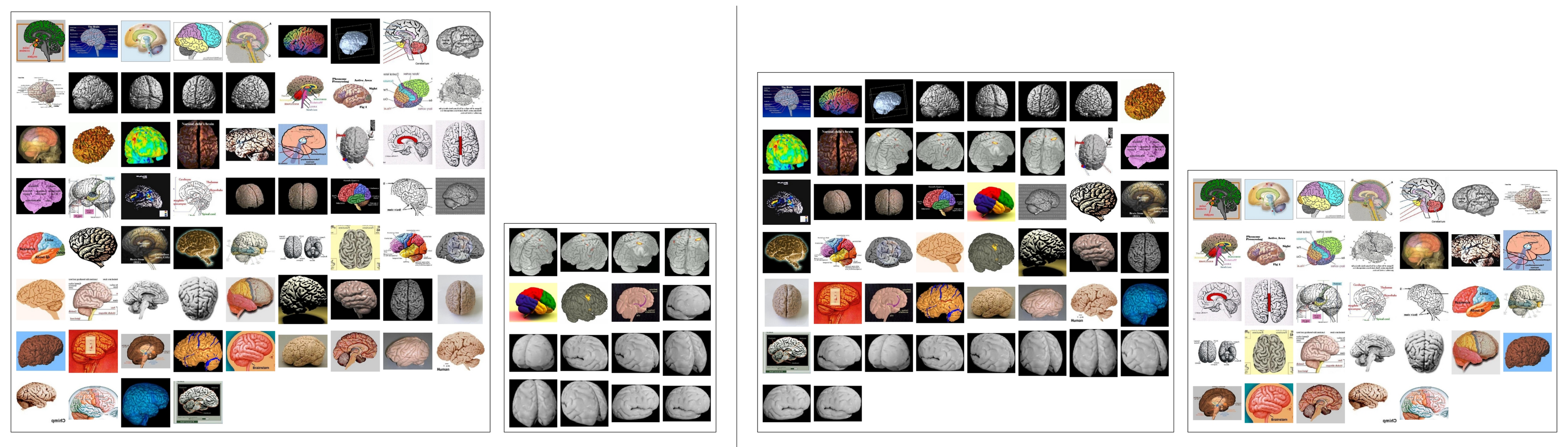
Figure: Errors rate in unmixing USPS digits for *K*-means (KM) clustering, *DCP* with best context class (DCP-best) and average of *DCP* over all context classes (DCP-avg). Black indicates 0% and White 50% error rate.

	2 vs. 0	
	1 : 1	1 : 10
KM	5.9±0.0	36.3±0.6
DCP-*1*	22.6±2.1	22.5±2.6
DCP-*2*	—	—
DCP-*3*	23.4±1.7	27.5±2.1
DCP-*4*	13.9±3.2	12.7±2.1
DCP-*5*	20.4±2.1	19.5±2.2
DCP-*6*	21.1±1.4	20.7±2.1
DCP-*7*	**5.8±2.6**	**6.0±1.8**
DCP-*8*	22.0±3.1	23.9±3.4
DCP-*9*	8.0±2.9	6.6±2.5
DCP-*0*	—	—
avg.	17.2±2.4	17.4±2.4

Figure: Numeric errors rate in unmixing USPS digits 2 vs. 0 with varying context class.

Finding substructures within Caltech256 classes

By varying the context class, we can browse through different splits of a single-label dataset. To humans, such splits can be *interpretable*:



(a) Split of `020.brain` with `021.breadmaker` context (b) Split of `020.brain` with `053.desk-globe` context

Figure: Explorative use of *DCP*: the *brain* class in Caltech256 is split using two different context classes. A human could interpret the first split as *structured vs. smooth*, and the second as *natural vs. schematic*.