PiCoDes: Learning a Compact Code for Novel-Category Recognition

Alessandro Bergamo  Lorentzo Torresani  Andrew Fitzgibbon

Microsoft Research

PiCoDes: Learning a Compact Code for Novel-Category Recognition

Goal: Large-Scale Object Class Search

We learn a descriptor: 1) binary 2) compact 3) optimized to be used with linear models 4) suitable for novel-class recognition

Technical requirements

1) Efficiency
   • Object classifier must be learned on-the-fly
   • Search results in real-time

2) Scalability
   • Keep millions of images in memory

3) Novel class classification
   • No knowledge of the query classes

Prior work

Multiple kernel combiner [11]: Hashing [22,24,31] good accuracy expensive to test high storage

PiCoDes learning: Intuition

• Training set of K training classes

Ours: PiCoDes

Large-Scale Object Class Search

ImageNet (a subset) 10 million photos, 2565 classes 256 classes varying tr. examples / class

PiCoDes: overview

x = non-linear projection in a high-dimensional space (SIFT dimensions)

z = \( \psi(x) \)

K(x, x') = \( \langle \psi(x), \psi(x') \rangle \)

Goal: training

Query: user-provided image belonging to a new class

Want: database images of the same class

Learn our PiCoDes: \( \sum_{k=1}^{K} \sum_{i=1}^{N} \epsilon_{k,i} x_{k,i} \) \( \Omega \) (PiCoDes training)

Training set of K training classes

PiCoDes descriptor for image \( x \): C-bit vector

Linear combination of these features implements an efficient multiple-kernel classifier

PiCoDes learning: objective

The K PiCoDes Boolean functions are jointly optimized for linear classification performance on a large training set of C classes

PiCoDes learning: optimization

Alternation scheme: learn linear classifiers (while keeping \( \beta \) fixed); traditional linear SVM learning

Experimental setup

Online database

PiCoDes (PiCoDes training):

PiCoDes evaluation:

Online database

ImageNet ILSVRC 2010

1000 classes varying num. tr. ex. / class 150 test examples / class

PiCoDes: PiCoDes descriptor

\( \Phi(x; A) = 1_{[a^T z > 0]} \)

Learning A

\( h = \Phi(x; A) \)

\( \sum_{k=1}^{K} \sum_{i=1}^{N} \epsilon_{k,i} x_{k,i} \) \( \Omega \)

6 features are applied to the test data of ILSVRC2010. For bit \( c \), all images are sorted by non-binarized classifier outputs \( \hat{a} \) and the 10 smallest and largest are presented on each row.

We use a convex upper-bound:

PiCoDes learning: objective

\( E(A, w_i, b_i, k) = \sum_{k=1}^{K} \sum_{i=1}^{N} \left[ |w_i|^2 + \sum_{i=1}^{N} |\beta_{i,k}| \right] \)

PiCoDes learning: Intuition

• Training set of K images, K different classes

PiCoDes learning: optimization

Alternation scheme:

Learn linear classifiers (while keeping A fixed); traditional linear SVM learning

Evaluations

Multiclass recognition accuracy on Caltech256 using different binary codes (10 training ex. / class). Model: 1-vs-all Linear SVM.

Precision of object-class-search using 2048-bit PiCoDes on ILSVRC 2010 (150 true positive, 150K distractions per query). Linear SVM yields a training+test time of less than 1 second per query.

Visualization of 128-bit PiCoDes.

Six features are applied to the test data of ILSVRC2010. For bit \( c \), all images are sorted by non-binarized classifier outputs \( \hat{a} \) and the 10 smallest and largest are presented on each row.

The learned features encode abstract visual properties (texture, high-level shapes, color patterns etc.).