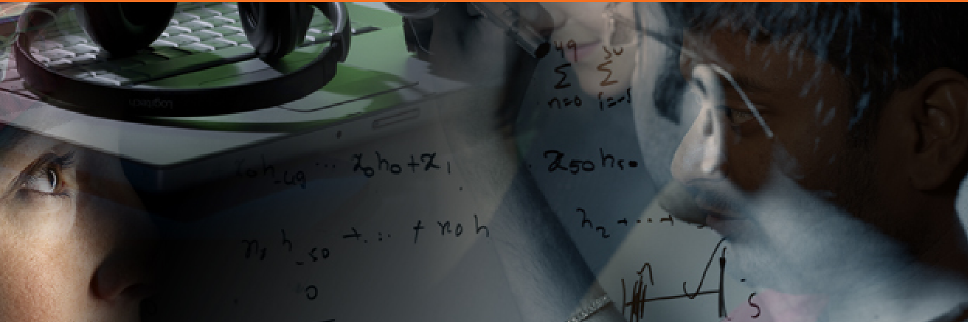
THE MASH PROJECT: TOWARD THE COLLABORATIVE DESIGN OF HYPER-COMPLEX LEARNING SOFTWARE

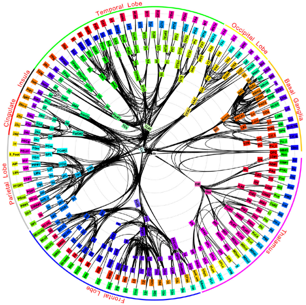
Dr. François Fleuret

International Advisory Board Meeting, Idiap, Sep. 2, 2011
Research Activities



Motivation
Complexity in Machine Learning

Artificial learning systems remain extremely simple compared to their biological counterparts.



(Macaque brain long-distance network, Modha and Singh, 2009)

Motivation
Complexity in Machine Learning

The Netflix Prize started in October 2, 2006, was won in September 2008 by “BellKor’s Pragmatic Chaos”.

> *... we use a set of diverse state-of-the-art collaborative filtering (CF) algorithms, which include: SVD, Neighborhood Based Approaches, Restricted Boltzmann Machine, Asymmetric Factor Model and Global Effects. We show that linearly combining (blending) a set of CF algorithms increases the accuracy and outperforms any single CF algorithm.*

(Jahrer et al., 2010)

Motivation

Complexity in Machine Learning

Flower classification (Gehler and Nowozin, 2009).

Single features	
Method	**Accuracy**
Colour	60.9 ± 2.1
Shape	70.2 ± 1.3
Texture	63.7 ± 2.7
HOG	58.5 ± 4.5
HSV	61.3 ± 0.7
siftint	70.6 ± 1.6

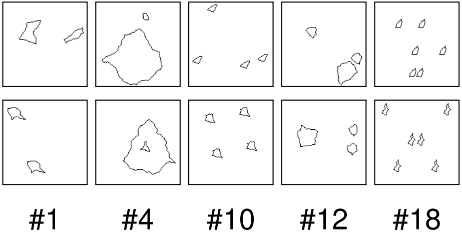
Combination methods	
Method	**Accuracy**
Product	85.5 ± 1.2
Averaging	84.9 ± 1.9
CG-Boost	84.8 ± 2.2
MKL (SILP)	85.2 ± 1.5
MKL (Simple)	85.2 ± 1.5
LP-β	85.5 ± 3.0

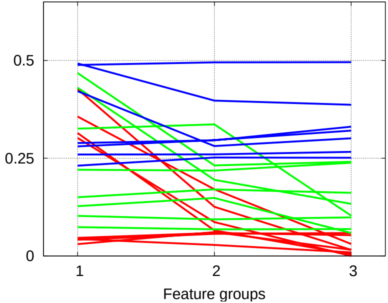
Pedestrian vs. background, INRIA dataset (MASH experiments).

	Number of weak learners		
	30	300	3000
Best family	4.06% (0.31)	1.43% (0.14)	1.04% (0.07)
19 families	3.65% (0.53)	0.69% (0.08)	0.30% (0.05)

Motivation

Complexity in Machine Learning



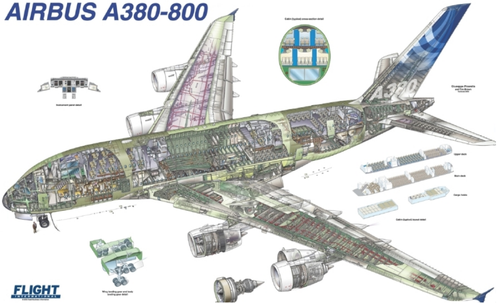


We tried to solve the Synthetic Visual Reasoning Test using Boosting and (1) Pixel counting, (2) Group 1 + Edge-like, (3) Group 2 + Fourier-like.

Motivation

Complexity in Machine Learning

So, complex learning systems are worth investigating. Should we envision ML software as engineering mega-projects ?



Motivation

The Fog of Machine Learning

Unfortunately, development of ML software involves specific engineering difficulties:

- Specifications involve a very complex object (data set, real-world POMDP)
- Limited understanding beyond rough behaviors: Convergence, over-fitting, cost, some invariance.
- Resulting algorithms combine very large numbers of (simple) cues. The emerging behavior is of a different nature.

We often have no idea why it truly works (buggy code sometime works as well ...)

Motivation
The Fog of Machine Learning

Developing ML-based methods for applications is a meta-learning algorithm.

Developers go back and forth between identifying mistakes (*High-frequencies patterns generate false alarms!*), fixing them (*Let's add features to detect high frequency blobs*), repeat (*we do not detect bald people! Let's add features to pick roundish shapes*), and repeat (*now we are over-fitting! Let's add a penalty!*)

This is similar to Boosting or SVM: At any moment, the most severe errors drive greedy changes in the constructed predictor.

Human are super-optimizers seeing (a bit) more than the gradient.

Motivation
The MASH project

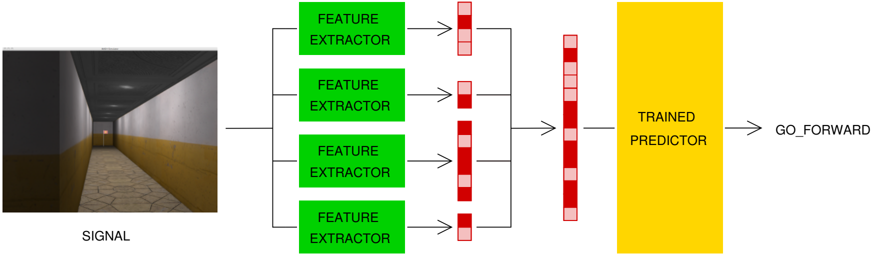
The MASH project is a European research initiative motivated by these observations.

- Machine learning lacks tools to rationalize the design of very complex architectures.
- Combining multiple feature extractors and prediction methods improves performances.
- Internet based collaborative tools allow large teams of individuals to work together.

We want to create new tools for designing complex learning systems in a collaborative manner.

Objectives
Feature extraction

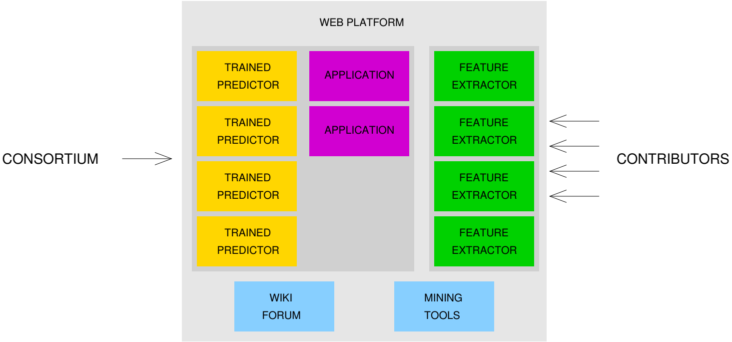
The project focuses on the design of very large families of feature extractors.



The feature vectors produced by multiple extractors are combined and fed to a machine learning algorithm.

Objectives
Collaborative platform

The collaborative design is organized on a web platform



The consortium selects the ML algorithms and experiments.
External contributors design feature extractors in C++.

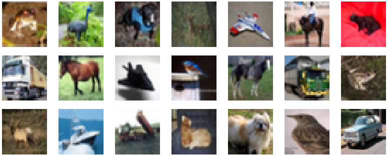
Objectives
Test applications

The initial application is image classification

MNIST

INRIA pedestrians

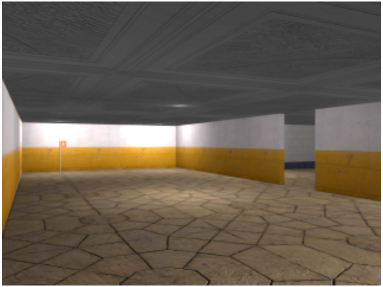
Caltech 256



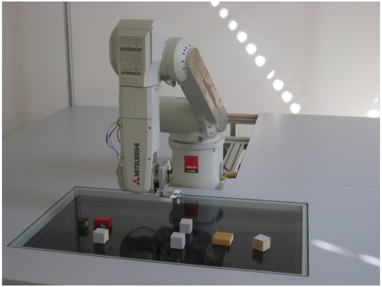
CIFAR-10

Objectives
Test applications

But the platform will soon includes goal-planning tasks.



3D simulator



Robotic arm

Objectives
Test applications

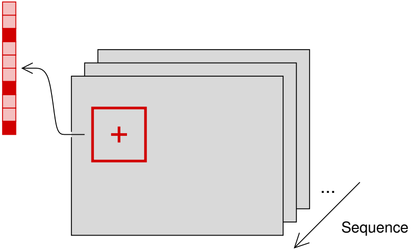
The project's objectives are to:

- Create tools to help the collaborative development of large families of feature extractors (open platform, metric, clustering).
- Develop ML methods to use them (very large dimension).
- Measure performance on image classification and goal planning.
- Deliver to the community an open platform, open-source implementations of novel algorithms, new data-sets.

The long-term goal is to create a new domain of "feature extractor mining".

Heuristics
Introduction

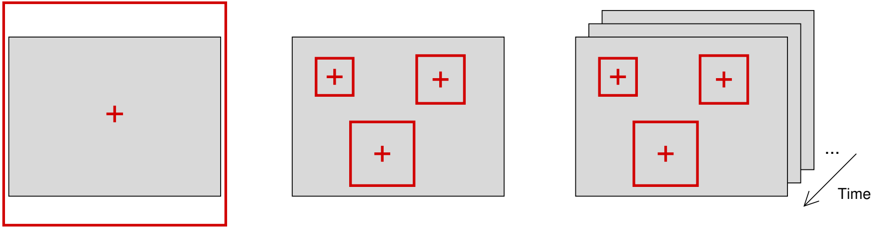
We define the concept of "Heuristic", a feature extractor with a persistent state.



Its core function is to compute a vector of features, for a certain sub-window size, at a given location in an image.

Heuristics
Introduction

This definition can be applied to multiple contexts.



Classification

Detection

Goal-planning

Contributing
Introduction

The project's web platform is operational at

http://mash-project.eu/

Contributors can create user accounts and upload as many heuristics as they want. The platform tests that they compile, runs a few experiments to evaluate their performance, and ranks them.

Each heuristic can be

- Private: The source remains undisclosed, but only small-scale experiments can be scheduled.
- Public: The source is distributed under the GPL v2, and the heuristic is enrolled in large-scale experiments and may participate to contests.

Contributing

Tools: heuristics

Logged in as johndoe | Inbox (0) | Logout

M A S H

News Documentation Downloads Members Forum **Heuristics** Experiments Contests

All the heuristics

Upload a new heuristic file: [ ] Browse... Upload

Display: **All the heuristics** **My public heuristics** **My private heuristics**

Search... Search

All A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

43 heuristics

USERNAME	NAME	RANK	# VERSIONS	DESCRIPTION	UPLOAD DATE
AndreBeinrucker	**simple_pixel_difference**	-	1	Checks whether the difference of a pixel value to the value of the neighbor pixel is higher than a threshold.	May 8, 2011, 8:57 p.m.
cdubout	**hough/3**	-	1	Computes the linear Hough transform of the region of interest to detect lines.	April 20, 2010, 12:35 p.m.
cdubout	**haar**	-	1	Haar transform heuristic. Computes the 2D haar transform of the region of interest (all levels).	April 20, 2010, 8:18 p.m.
cdubout	**identity**	-	1	Same as the mash/identity heuristic except that it returns RGB values instead of grayscale.	April 20, 2010, 8:24 p.m.
cdubout	**segmentation2**	-	1	Segment the gray-scale image pixels into 2 categories, the ones above the average intensity, and the ones below.	Sept. 26, 2010, 11:57 p.m.
cdubout	**segmentation8**	-	1	Segment the gray-scale image intensities into 8 bins with equal number of pixels in each. Returns the index of the bins in which a new pixel falls.	Sept. 27, 2010, 10:29 a.m.
cdubout	**hogblurred/4**	-	1	Strongly blurred gradient magnitude images for 8 different gradient orientations (also rescales the	Sept. 28, 2010, 12:48 p.m.

Contributing

Tools: heuristics

Logged in as johndoe | Inbox (0) | Logout

MASH

News Documentation Downloads Members Forum **Heuristics** Experiments Contests

My private heuristics

Upload a new heuristic file: Browse... Upload

Display: **All the heuristics** **My public heuristics** **My private heuristics** Search... Search

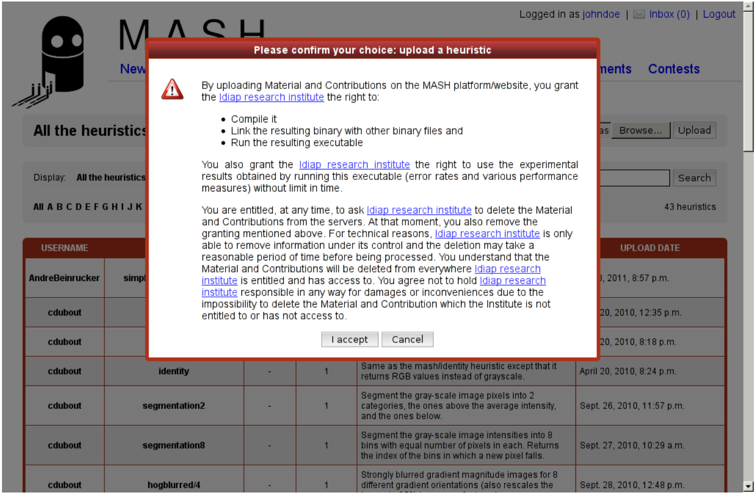
All A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

6 heuristics

NAME	RANK	# VERSIONS	DESCRIPTION	UPLOAD DATE	TOOLS
average2	-	0		Sept. 28, 2010, 3:34 p.m.	
blobs	-	0		March 1, 2011, 4:10 p.m.	
closerpixel	-	0	This heuristic looks at 9x9 blocks of grayscale pixels and returns the index of the one which is the closer (intensity-wise) of the center one.	Feb. 16, 2011, 2:45 p.m.	
identity	-	0		Dec. 9, 2010, 7:30 a.m.	
fartherpixel	-	0	This heuristic looks at 9x9 blocks of grayscale pixels and returns the index of the one which is the farther (intensity-wise) of the center one.	Feb. 16, 2011, 2:47 p.m.	
rgbhistogram	-	0	fghfghfghf	Feb. 18, 2011, 10:03 a.m.	

Contributing

Tools: heuristics



Contributing

Tools: heuristics

Logged in as johndoe | Inbox (0) | Logout

MASH

News Documentation Downloads Members Forum **Heuristics** Experiments Contests

Your heuristic is being tested...

Results:

Status: Waiting...
Compilation:
Analyze:
Test:
Evaluation:

zk

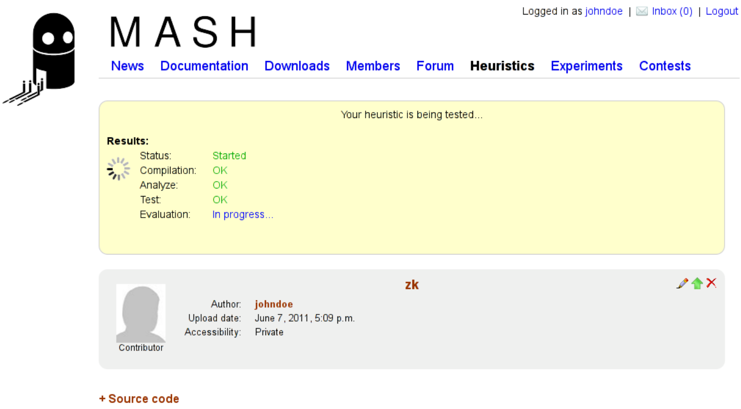
Author:	johndoe
Upload date:	June 7, 2011, 5:09 p.m.
Accessibility:	Private

Contributor

+ Source code

Contributing

Tools: heuristics



Contributing
Tools: experiments

Logged in as johndoe | Inbox (0) | Logout

MASH

News Documentation Downloads Members Forum Heuristics **Experiments** Contests

Public experiments + Schedule a new private experiment

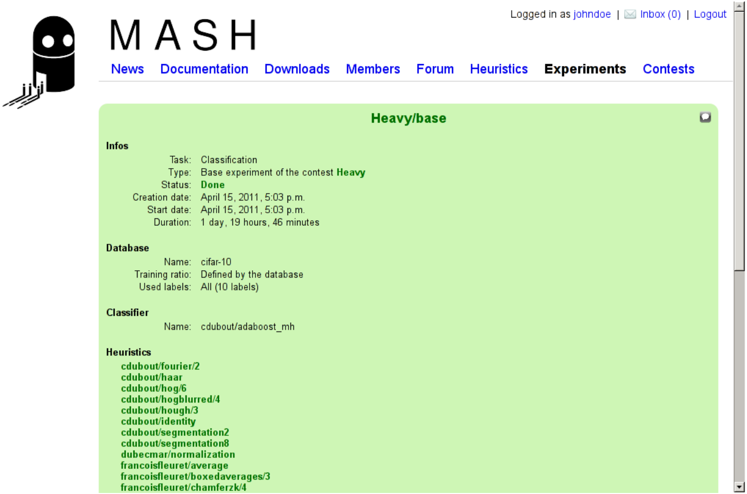
Display: The public experiments My private experiments Search... Search

All A B C D E F G H I J K L M N O P Q R S T U V W X Y Z 18 experiments

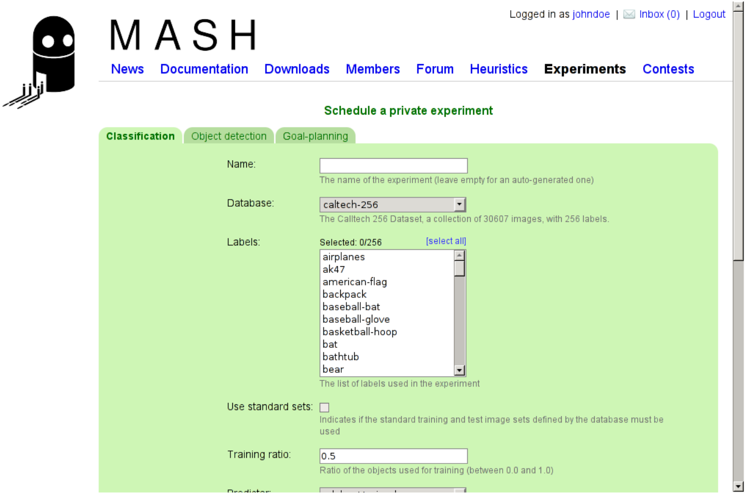
NAME	TASK	DETAILS	PREDICTOR	# HEURISTICS	RESULTS	CREATION DATE	START DATE	DURATION
public/classification/simpleadaboost-mnist/2011-05-27 00:26:18.021047	Classification	mnist (10 labels)	francoisfleuret/simpleadaboost	43	0.77% / 1.66%	May 27, 2011, 12:26 a.m.	May 27, 2011, 12:26 a.m.	15 hours, 35 minutes, 52 seconds
public/classification/simpleadaboost-mnist/2011-05-24 23:54:55.799849	Classification	mnist (10 labels)	francoisfleuret/simpleadaboost	42	0.73% / 1.45%	May 24, 2011, 11:54 p.m.	May 24, 2011, 11:54 p.m.	14 hours, 58 minutes, 40 seconds
public/classification/simpleadaboost-mnist/2011-05-24 05:14:27.878822	Classification	mnist (10 labels)	francoisfleuret/simpleadaboost	41	0.75% / 1.65%	May 24, 2011, 5:14 a.m.	May 24, 2011, 5:14 a.m.	14 hours, 31 minutes, 1 second
public/classification/simpleadaboost-mnist/2011-05-23 14:38:06.722241	Classification	mnist (10 labels)	francoisfleuret/simpleadaboost	40	0.74% / 1.63%	May 23, 2011, 2:38 p.m.	May 23, 2011, 2:38 p.m.	14 hours, 35 minutes, 21 seconds
public/classification								

Contributing

Tools: experiments

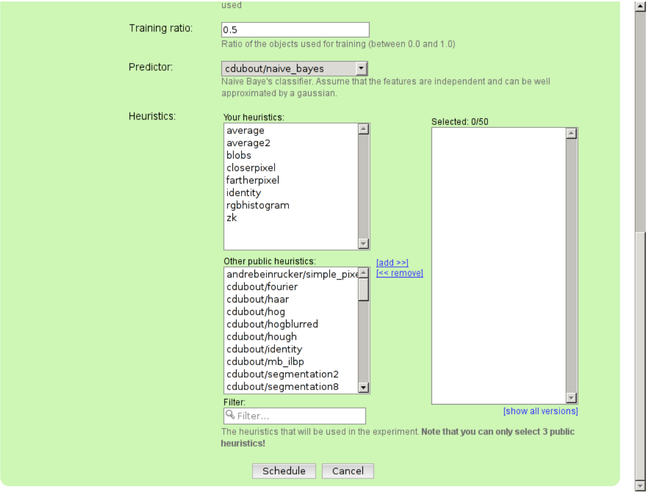


Contributing
Tools: experiments



Contributing

Tools: experiments



Contributing
Tools: instruments

Results

Training error: 1.71%
Test error: 31.96%

- Features used

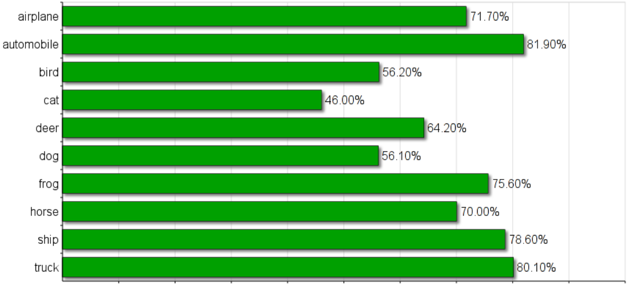
cdubout/fourier/2	**1739**/8192
cdubout/haar	**401**/4096
cdubout/hog/6	**949**/2000
cdubout/hogblurred/4	**358**/2048
cdubout/hough/3	**277**/1188
cdubout/identity	**391**/3267
cdubout/segmentation2	**7**/1089
cdubout/segmentation8	**132**/1089
dubecmar/normalization	**31**/1089
francoisfleuret/average	**1**/1
francoisfleuret/boxedaverages/3	**115**/1000
francoisfleuret/chamferzk/4	**327**/8712
francoisfleuret/zk/4	**326**/1000
gillesblanchard/blockwise	**73**/100
johndoe/average	**0**/1
kanma/blobs	**164**/1089
kanma/closerpixel	**371**/961
kanma/fartherpixel	**305**/961
kanma/grayscalehistogram	**111**/256
kanma/rgbhistogram	**105**/512
leonidas/dt_hogs10	**815**/1764
leonidas/dt_hogs6	**216**/324
leonidas/rgb_hogs	**285**/324
leonidas/simple_color_hist	**3**/3
lmerchante/edgessymmetryax	**6**/10
lmerchante/gradient	**109**/1089
lmerchante/symmetript	**4**/9
lmerchante/symmetryax/3	**10**/10
lmerchante/thresholdgradient	**1**/1089
mash/identity	**62**/1089
mash/meanthreshold	**9**/1089

Contributing
Tools: instruments

lmerchante/symmetryax/3	**10**/10
lmerchante/thresholdgradient	**1**/1089
mash/identity	**62**/1089
mash/meanthreshold	**9**/1089

- Correct classifications

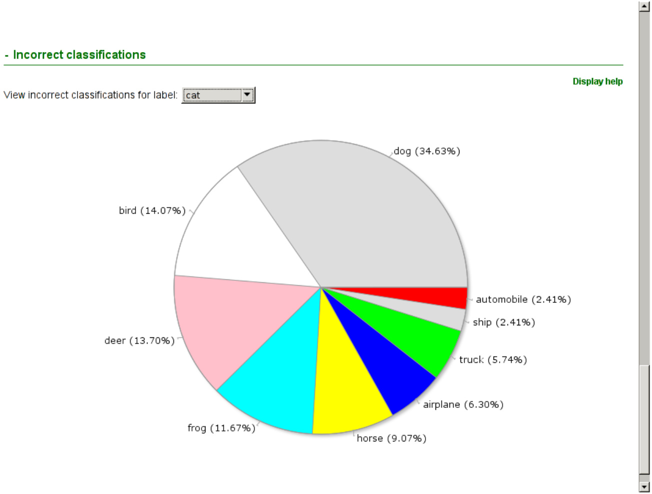
Display help



+ Incorrect classifications

+ Worst mistakes

Contributing
Tools: instruments



Contributing
Tools: instruments

- Worst mistakes

Label airplane

Worst false positives

Worst false negatives

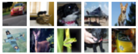
Label automobile

Worst false positives

Worst false negatives

Label bird

Worst false positives



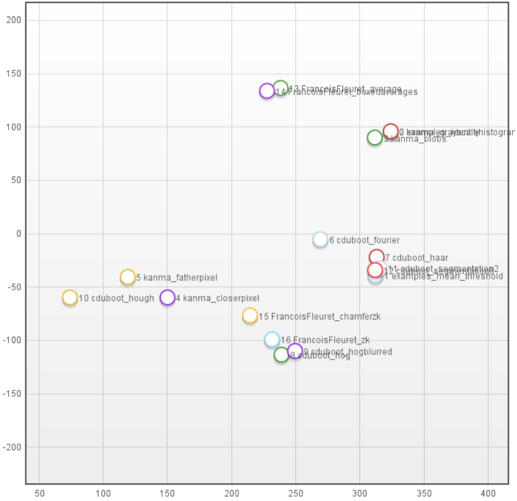
Worst false negatives

Contributing
Tools: clustering

HEURISTIC	CLUSTER ID
examples/identity	1
francoisfleuret/average	
francoisfleuret/boxedaverages	
kanma/blobs	
cduboot/fourier	2
cduboot/haar	
cduboot/hog	
francoisfleuret/zk	
kanma/closerpixel	
kanma/fatherpixel	
kanma/grayscalehistogram	
lmerchante/edgessymmetryax	
lmerchante/gradient	
lmerchante/symmetript	
lmerchante/thresholdgradient	
cduboot/hogblurred	3
cduboot/hough	
lmerchante/symmetryax	
cduboot/segmentation2	4
cduboot/segmentation8	
examples/mean_threshold	
francoisfleuret/chamferzk	

Contributing

Tools: clustering



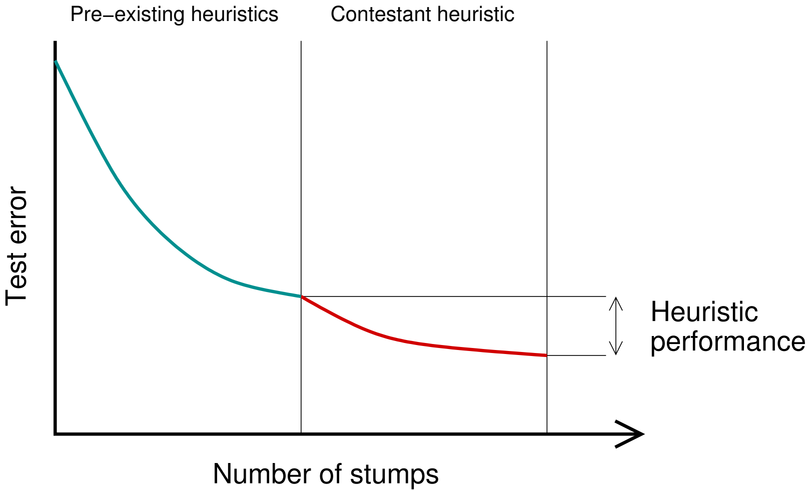
We recently launched a three-tracks heuristic design contest.

The protocol for each track is similar:

- We pre-trained a strong classifier composed of N stumps selected with Adaboost, using all the consortium heuristics.
- For each participating heuristic, we add M additional stumps using that heuristic alone.
- The final performance is the reduction of the test error.

The data-set for that first contest is CIFAR-10. The three tracks correspond to $N = 0$, 100, and $10,000$ respectively, and $M = 100$.

Contributing
Contests



Contributing
Contests

The best heuristic will be selected at the end of every month in each track, and added to the pool of heuristics used to train the strong classifier.

This process will ensure that the optimal modalities to exploit evolve over time.

To evaluate the “design over-fitting”, we kept 1/5 of the standard CIFAR training set as a validation set.

Conclusion
Future

The platform is still young, and in development on many fronts:

- Goal-planning contest.
- Deployment of new prediction algorithms (feature selection and Boosting in high-dimension).
- Development toward education.

✎ F. Fleuret, P. Abbet, C. Dubout, and L. Lefakis. The MASH project. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, pages 626–629, 2011

✎ C. Dubout and F. Fleuret. Tasting families of features for image classification. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2011. To appear

✎ C. Dubout and F. Fleuret. Boosting with maximum adaptive sampling. In *Proceedings of the Neural Information Processing Systems Conference (NIPS)*, 2011. To appear

THANK YOU