

The MASH Project

François Fleuret Philip Abbet
Charles Dubout Leonidas Lefakis

Idiap Research Institute, Martigny, Switzerland

July 12, 2011

Abstract

It has been demonstrated repeatedly that combining multiple types of image features improves the performance of learning-based classification and regression. However, no tools exist to facilitate the creation of large pools of feature extractors by extended teams of contributors.

The MASH project aims at creating such tools. It is organized around the development of a collaborative web platform where participants can contribute feature extractors, browse a repository of existing ones, run image classification and goal-planning experiments, and participate in public large-scale experiments and contests.

The tools provided on the platform facilitate the analysis of experimental results. In particular, they rank the feature extractors according to their efficiency, and help to identify the failure mode of the prediction system.

1 Introduction

Research in Artificial Intelligence has historically focused on two diverse approaches each with their own advantages: Symbolic methods which allow the hand-design of rich prior knowledge under the form of large sets of formal rules, and statistical methods which can cope with the unpredictability and the randomness of real-world situations.

At a crossroads between these two approaches, it has systematically been shown that increasing the complexity of learning systems, for instance by

combining algorithms developed independently by different groups of experts is a successful strategy. The Netflix challenge was ultimately won by combining multiple predictors, developed by different teams [1]. In object recognition, state-of-the-art performances are attained by combining multiple feature extractors, with simple learning techniques, e.g. [2]. Our own experiments on the INRIA pedestrian data set show that using a Boosting learning scheme in conjunction with multiple feature extractors, as opposed to solely the best, leads to a reduction in error rate from $\simeq 1\%$ to $\simeq 0.3\%$.

Despite this evidence and the admitted long-term goal of machine learning to produce systems dealing with versatile real-world challenges, no tools have been invented for the design of complex learning architectures by extended teams of people. Though tools such as version control systems or general modeling methods facilitate some classical aspects of development, for which modular and deterministic specifications can be drafted, the heart of a learning system requires a statistical and redundant approach. While it may be possible to predict from the formal definition of a learning method some crude aspects of its behavior, such as convergence or robustness to overfitting, it is extremely difficult to foresee its performance on real data. This fog of machine learning requires a systematic experimental evaluation to an extent which is not handled properly by existing software development tools and methodologies.

2 The MASH Platform

In order to address this structural complexity, we have developed a framework, centered around a web platform, to study the collaborative design of very large families of feature extractors (see figures 1 and 2).

We call “heuristic” an algorithm that computes a feature vector at any scale and location in an image, and possesses a persistent state (see figure 3). This definition is general enough to allow classical pre-processing from the computer vision world, such as edge detectors, color histograms, SIFT, HOG, LBP, etc. and leads to a clear and simple specification as a C++ class implementing a few methods. The persistent state is irrelevant to image classification, but is used for goal-planning with POMDP, one of the target applications of the project.

Participants to the project may download a multi-platform Software Development Kit (SDK) which allows them to develop and test MASH heuris-

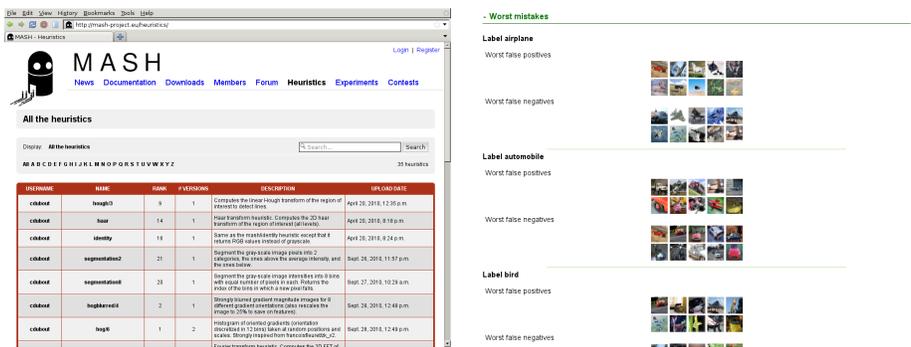


Figure 1: The MASH platform at <http://mash-project.eu>.

tics on their Linux, Microsoft Windows, or Mac OSX machine. The platform

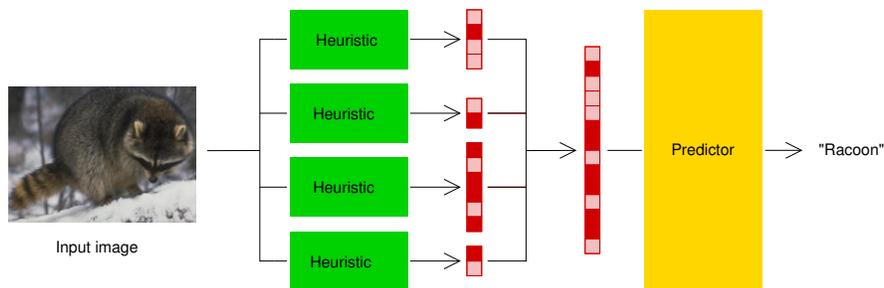


Figure 2: Feature extraction and processing of the resulting signal.

gives users access to documentation, including screen-cast tutorials, regarding the functionality of the web platform itself as well as the development of the heuristics. The platform is further equipped with a forum and private messaging which allow user interaction and communication.

Registered users are able to upload heuristics to the platform which are then stored in those users' "private spaces" and are accessible only by them. The platform allows these contributors to run experiments which provide an evaluation of the quality of a contribution without public exposure; these experiments are run remotely on the platform's servers, thus freeing the users' personal computational resources.

A contributor can at any moment move heuristics to the public space, in which case they become visible, are included in the large-scale public exper-

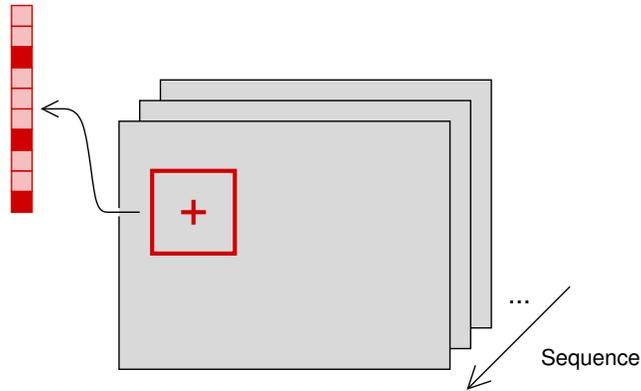


Figure 3: A heuristic is an image feature extractor with a persistent state.

iments which aim to develop complex state-of-the-art systems, and can also be re-used by any other contributor under an open-source license. Feature extractors already public include well-known methods (e.g. HoG, LBP), as well as novel methods developed on the framework. The source code of all public heuristics is available under the GPL v2 license.

The system tests performances on three families of applications : image classification, object detection, and goal-planning in a simulated environment (reminiscent of modern 3D video-games) and with a real robotic arm. Experiments are run concurrently on multiple machines and the platform aggregates the results in a synthetic manner. It highlights the strengths and weaknesses of the up-to-date resulting trained predictors, and returns to the user a number of quantitative and qualitative evaluations of the results, such as the raw accuracy of a trained classifier, or samples of the worst mistakes on the test data-set (see Fig. 1, right). The standard interaction between a contributor and the platform is to alternate between private experiments to try “new ideas”, and public experiments in which their heuristics aid machine learning modules in their identified shortcomings.

3 Contests

In addition to the usual tasks the platform addresses, the web platform also periodically holds contests where registered users can compete on a specific task and data-set.

At the moment, the website is running a three-track contest on image clas-

sification based on the CIFAR data-set [3]. We have for each track trained with Boosting a strong predictor composed of N stumps – the tracks corresponding to $N = 0$, $N = 100$ and $N = 10,000$ respectively – using all the heuristics we have already implemented. For each new heuristic participating in the contest, the system runs 100 additional iterations of Boosting, and computes the gain in test error, which is used as the performance measure for that contesting heuristic.

The best heuristic in each track will be selected every month, and added to the pool of heuristics used to train the strong classifiers, which will be subsequently re-trained. Beyond competing with fellow users and aiming to win a prize which will be offered to the “best” feature extractor, users have the opportunity to contribute to the ongoing development of a sophisticated state-of-the-art classification system. The system’s current trained classifier for the CIFAR data-set already attains state-of-the-art performance and aims to surpass this with the contributors’ help.

4 Conclusion

We advocate the need for a new research domain investigating complex learning systems. In order to develop a unified and centralized artificial intelligence, able to deal with the real world’s versatility and complexity, we will first have to develop the requisite, and novel, tools for hand-designing such architectures by hundreds or even thousands of contributors.

The MASH platform is a first initiative in this direction. It targets architectures combining large sets of feature extractors with standard learning procedures. It allows multiple contributors to combine their efforts, it hosts multiple algorithms and runs multiple experiments transparently to assess performance continuously on multiple tasks.

5 Acknowledgments

This work was supported by the European Community’s Seventh Framework Programme FP7 – Challenge 2 – Cognitive Systems, Interaction, Robotics – under grant agreement 247022 – MASH, and by the Swiss National Science Foundation under grant 200021-124822 – VELASH.

References

- [1] A. Toscher, M. Jahrer, and R. Bell. The bigchaos solution to the netflix grand prize. On-line, 2009.
http://www.netflixprize.com/assets/GrandPrize2009_BPC_BigChaos.pdf
- [2] P. Gehler and S. Nowozin. On feature combination for multiclass object classification. In International Conference on Computer Vision, 2009
- [3] <http://www.cs.toronto.edu/~kriz/cifar.html>