

Pascal Challenges

Michèle Sebag

EuCOGIII, April 11th, 2013

Vision 1

- ▶ Structure *ab ovo* a research roadmap
- ▶ Define core issues and measures of progress

Vision 2

- ▶ Designed by organizers
- ▶ Reviewed
- ▶ Proposed to the community
- ▶ Defining actionable research questions
- ▶ Structuring *a posteriori* the community

The website of the challenge remains open for post-challenge submissions... the challenge and workshop websites [plus a survey paper/book] offer a complete teaching toolkit and a valuable resource for engineers and scientists.

Requisite

relevant

fun

doable

a catalyst for both research and application development

Three categories of challenges

- ▶ Core enabling skills for cognitive systems:
Vision; Speech and Language
- ▶ Advancing ML
Representation (transfer, unsupervised); Active learning;
Causality; Scalability, ...
- ▶ Reaching out
e-Science, Health, CHI, CRM, Content selection, ...

Co-evolution of organizers/participants

- ▶ Needed: rigorous evaluation procedure
If someone can cheat, someone will...
- ▶ Comprehensive comparison of approaches
efficiency / originality / insights
- ▶ Be prepared for unexpected results

Challenges set the trend

- ▶ Build resources
- ▶ Definition of tasks
- ▶ Definition of metrics
- ▶ Citations, BP Awards
- ▶ Super-methods (ensemble of submitted methods)

→ mldata.org

Guyon, 2013

Challenges are a powerful tool to perform empirical research in Machine Learning and get (relatively) unbiased answers to questions of general interest. Often, the answers go against common opinions established with largely overfitted studies.

- ▶ *Should one perform active learning? Most published studies: yes. Active learning challenge : NO unless you really need; once you start selecting examples your can easily distort your input distribution and learn the "wrong" problem.*
- ▶ *Does unsupervised learning help as a preprocessing? Opinions are split. Unsupervised challenge: YES!*

- ▶ *Should you use causal discovery methods to find more causally relevant input features? Studies assuming an oracle is available to give you conditional independence relationships: yes. Causality challenges: NO unless you really know what you are doing. It is very hard to beat non-causal feature selection even to solve causal problems.*
- ▶ *Can you learn from just one training examples difficult tasks like recognizing gestures from video data? Opinions are split. Gesture recognition challenge: YES!*

1. Setting the trend: Recognizing Textual Entailment
2. Pushing the state of the art: Visual Object Challenges
3. Reaching out:
 - ▶ Gesture recognition
 - ▶ Causality studies

Ido Dagan, Bernardo Magnini et al.

Goal Given H and T, does H implies T ?

Example

H Reagan attended a ceremony in Washington to commemorate the landings in Normandy.

→ ?

T Washington is located in Normandy.

	TEXT	HYPOTHESIS	TASK	ENTAIL- MENT
1	<i>Regan attended a ceremony in Washington to commemorate the landings in Normandy.</i>	<i>Washington is located in Normandy.</i>	IE	False
2	<i>Google files for its long awaited IPO.</i>	<i>Google goes public.</i>	IR	True
3	<i>...: a shootout at the Guadalajara airport in May, 1993, that killed Cardinal Juan Jesus Posadas Ocampo and six others.</i>	<i>Cardinal Juan Jesus Posadas Ocampo died in 1993.</i>	QA	True
4	<i>The SPD got just 21.5% of the vote in the European Parliament elections, while the conservative opposition parties polled 44.5%.</i>	<i>The SPD is defeated by the opposition parties.</i>	IE	True

A building block for:

- ▶ Reading comprehension
- ▶ Question answering
- ▶ Information extraction
- ▶ Paraphrase Acquisition
- ▶ Summarization
- ▶ Machine Translation

- ▶ Slot filling (Knowledge Population Base)

Last four RTEs

- ▶ Track of the Text Analysis Conference (TAC) at NIST, the U.S. National Institute of Standards and Technology since 2008
- ▶ Plus summarization, novelty detection, slot filling (Knowledge Base Population)

Next: toward educational technology

- ▶ Student Response Analysis task at SemEval 2013 (<http://www.cs.york.ac.uk/semeval-2013/task7/>).
- ▶ Given a question, a known correct *reference answer* and a *student answer*, **classify student answer** as correct; partially correct but incomplete; contradictory; irrelevant; not in the domain.

1. Setting the trend: Recognizing Textual Entailment
2. Pushing the state of the art: Visual Object Challenges
3. Reaching out:
 - ▶ Gesture recognition
 - ▶ Causality studies

Dining Table



Dog



Horse



Motorbike



Person



Potted Plant



Sheep



Sofa



Train



TV/Monitor



Visual Objects Challenge, tasks

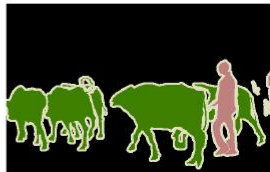
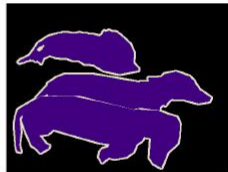
Image



Object segmentation



Class segmentation



State of art

- ▶ Features: Dense SIFT, HOG, colour
- ▶ Encodings: spatial pyramid, BOW, Fisher vector
- ▶ Detectors: DPM
- ▶ Classifier: SVM

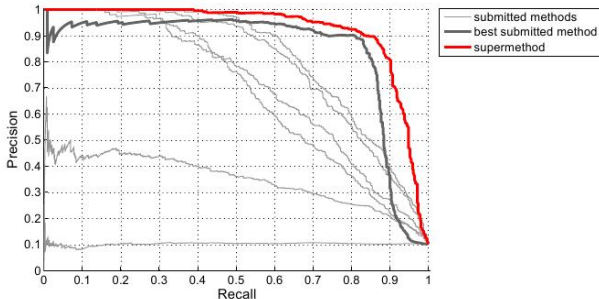
This year

- ▶ Complex log-normal features
- ▶ Sub-clusters for classes
- ▶ Combinations of clusterings and projections

- ▶ New metrics: timeliness (AP vs time).

Super-methods

Cascade: use methods score as attributes; train using linear SVMs.



Continued

ImageNet Large Scale Visual Recognition Challenge

1. Setting the trend: Recognizing Textual Entailment
2. Pushing the state of the art: Visual Object Challenges
3. Reaching out:
 - ▶ Gesture recognition
 - ▶ Causality studies



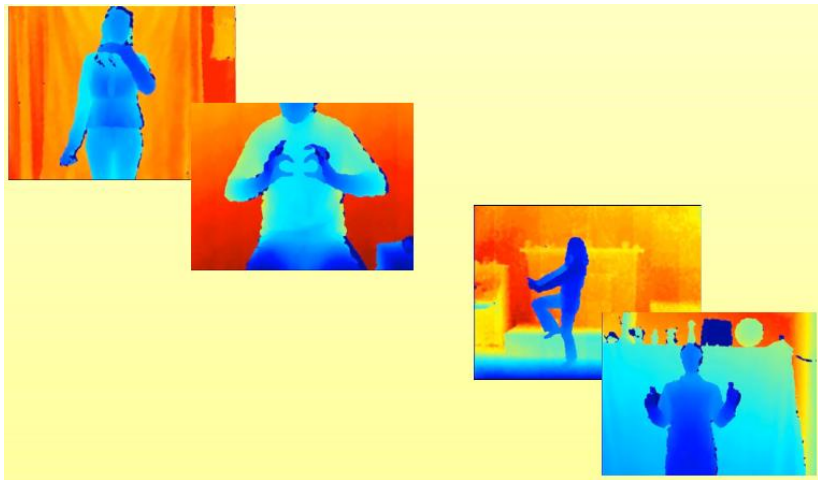
A building block for:

- ▶ Action and activity recognition
- ▶ Recognition of sign languages
- ▶ Human action recognition
- ▶ Teaching communication
- ▶ Video indexing/retrieval
- ▶ Emotion recognition
- ▶ Video surveillance
- ▶ Performing arts
- ▶ Games

- ▶ Fixed camera
- ▶ Depth available
- ▶ Within a batch, single user, small vocabulary, homogeneous recording conditions
- ▶ Mostly arm and hand gestures
- ▶ Camera framing upper body

Challenging

- ▶ A single example of each gesture within a batch
- ▶ Skeleton tracking data not provided
- ▶ Variation in background, clothing, skin color, lighting, resolution among batches
- ▶ Some parts of body occluded
- ▶ Some users more skilled than others



1st

novel technique inspired by the neural mechanisms underlying information processing in the visual system.

2nd

HOG/HOF features. Recognition and temporal segmentation with DTW. Quadratic-chi kernel similarity.

3rd

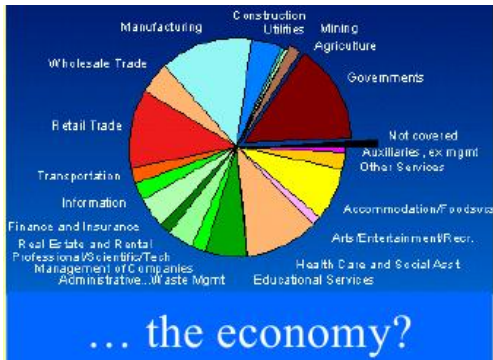
DTW temporal segmentation. Bag of features from 3D motion SIFT from RGB-D data. KNN classifier.

1. Setting the trend: Recognizing Textual Entailment
2. Pushing the state of the art: Visual Object Challenges
3. Reaching out:
 - ▶ Gesture recognition
 - ▶ Causality studies

Societal impact



Societal impact



Features

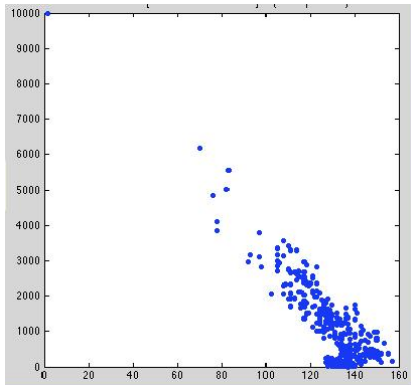
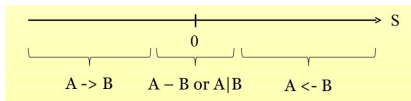
- ▶ Observations available
- ▶ Correlation \neq causality
- ▶ Experiments would serve: but ! (costly, infeasible, unethical)

Setting

- ▶ A simple benchmark: financial series, possibly reverted;
predict whether time has been reverted.
- ▶ No feedback-loop, no time
- ▶ With an independence oracle
ind. tests require data and simplifying assumptions (causal
sufficiency, Gaussian noise...)

Modelling causality

$$A = F(B, \text{noise}); B = F(A, \text{noise});$$



Vision 1

provide a common plan and vision, a consensus on what should be done first and what counts as success...

Vision 2

- ▶ Proposed by the organizers to the community
platform heterogeneity ?
- ▶ Defining actionable research questions
e.g. Knowledge: innate / acquired / generality ?

Challenge: an operational/conceptual way of structuring the community.