DCU-SAVASA
Participation and Reflections

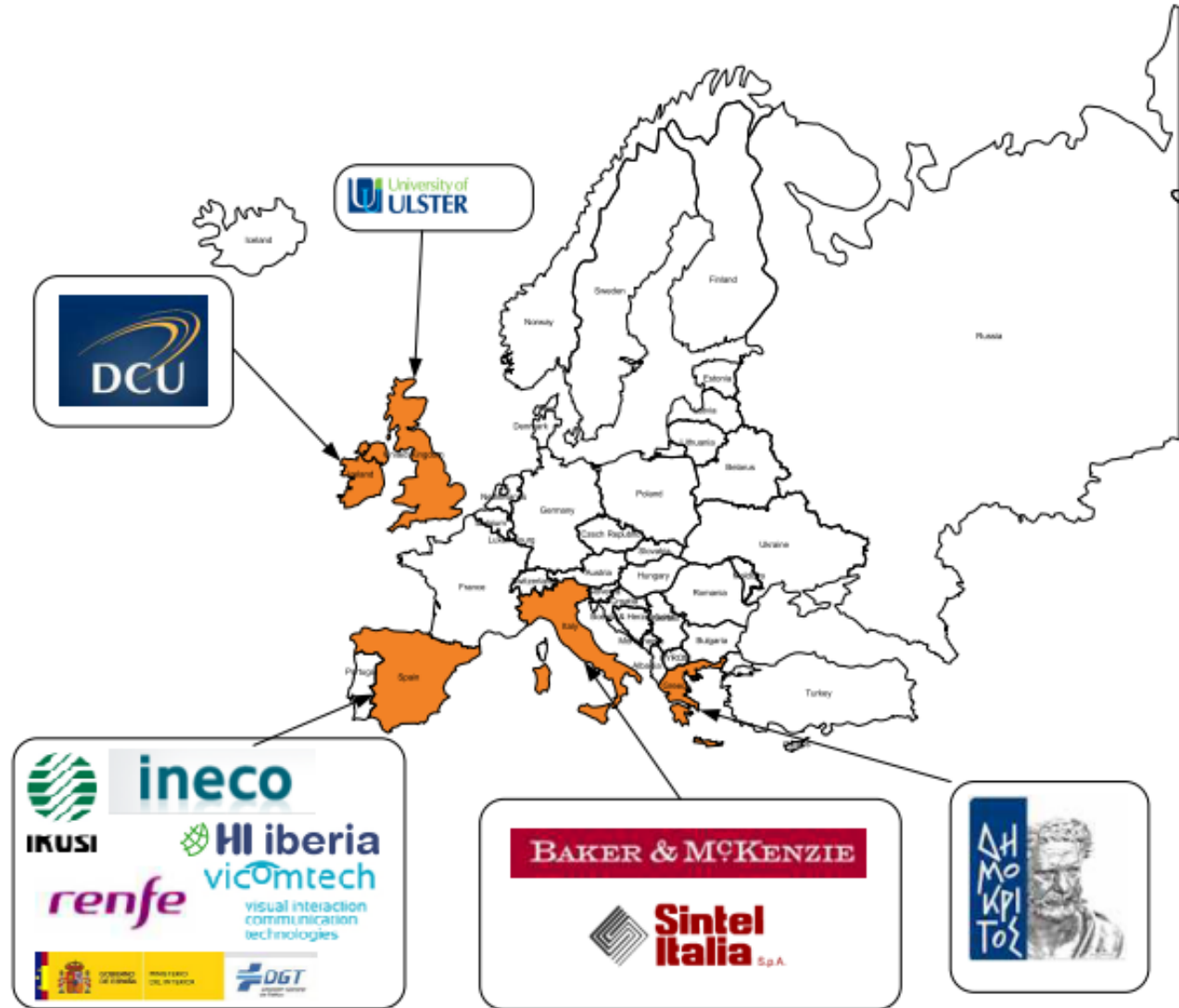Rami Albatal

rami.albatal@dcu.ie

SEVENTH FRAMEWORK
PROGRAMME

SAVASA

- FP7 project
- Started in December 2011
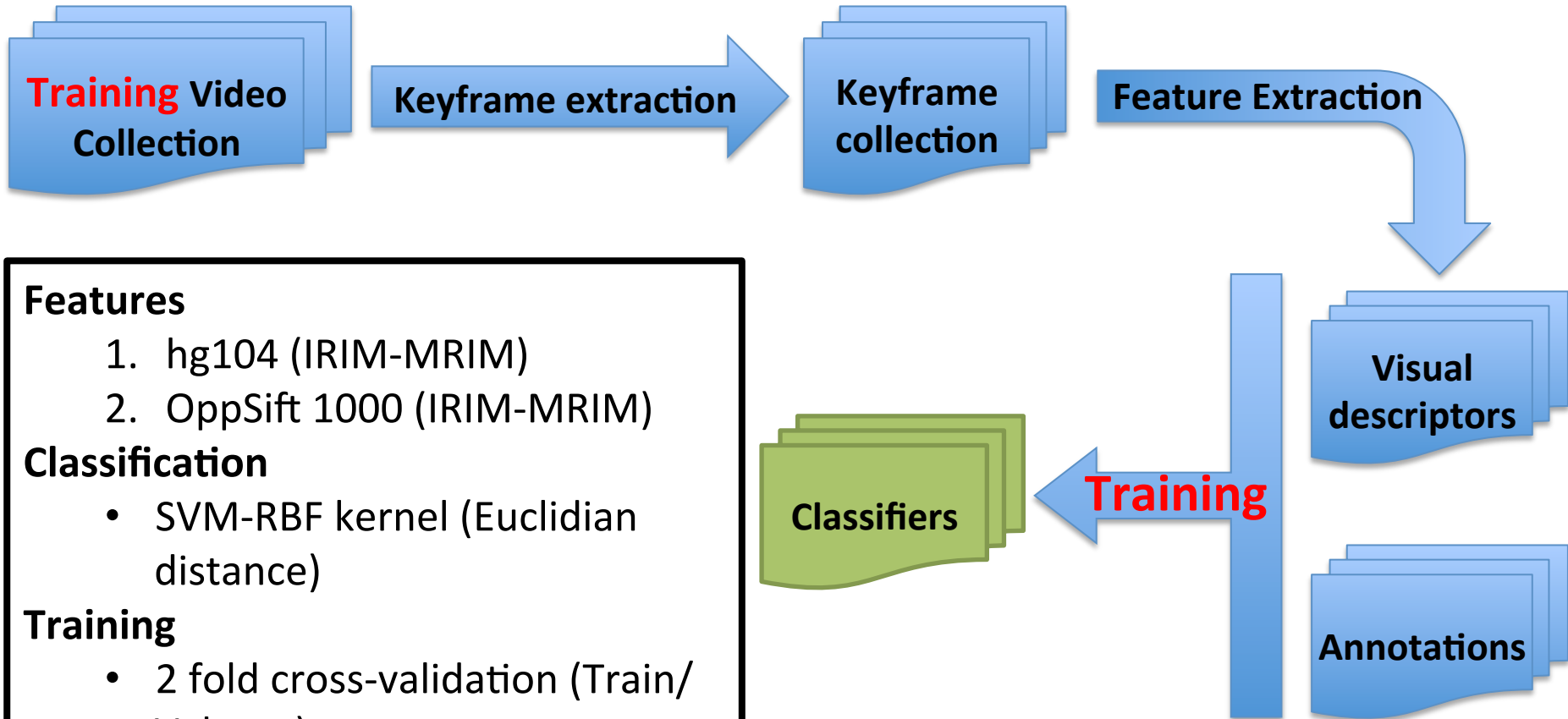- 30 months
- 11 participants
- 5 countries

The SAVASA project

- Proposes the creation of a **video archive search platform** that allows authorized users to perform **semantic queries** over various remote and non-interoperable surveillance video archives.

- Exploits the current research trends in computer vision, video retrieval and semantic video analysis.

SAVASA participation in SIN 2013

- First participation of SAVASA in SIN

- Goals
  - Explore the quality of a common state-of-the-art approach for video retrieval (proof of concept)

  - Discover the requirements for building video semantic indexing tools

  - Identify the future directions that should be followed

SAVASA participation in SIN 2013

Training Video Collection

Keyframe extraction

Keyframe collection

Feature Extraction

Visual descriptors

Training

Classifiers

Annotations

Features
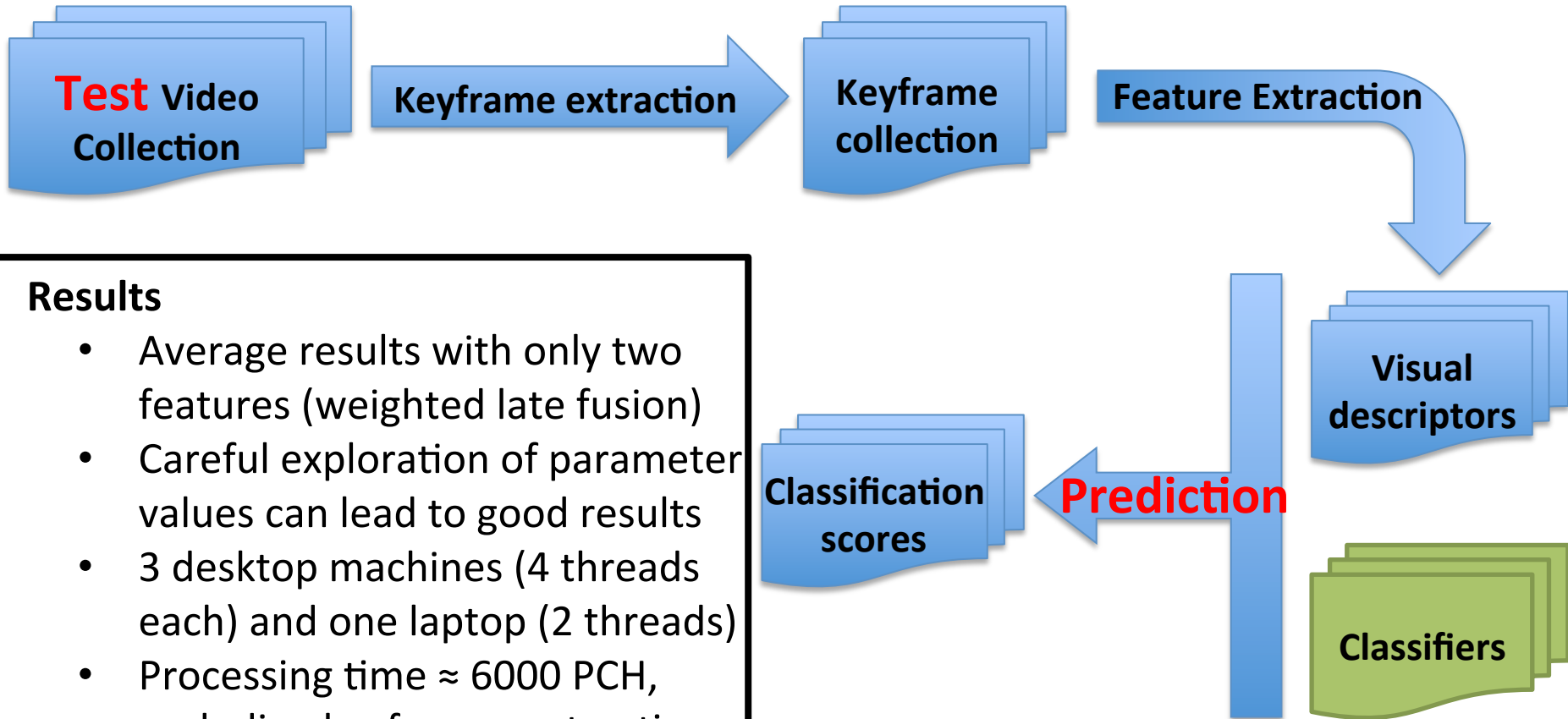1. hg104 (IRIM-MRIM)
2. OppSift 1000 (IRIM-MRIM)

Classification
- SVM-RBF kernel (Euclidian distance)

Training
- 2 fold cross-validation (Train/ Val sets)
- Careful exploration of training parameters

SAVASA participation in SIN 2013

Test Video Collection

Keyframe extraction →

Keyframe collection

Feature Extraction

Visual descriptors

Prediction

Classification scores

Classifiers

Results
- Average results with only two features (weighted late fusion)
- Careful exploration of parameter values can lead to good results
- 3 desktop machines (4 threads each) and one laptop (2 threads)
- Processing time ≈ 6000 PCH, excluding keyframes extraction and features extraction

SAVASA participation
in SIN 2013

- Overall Mean InfAP
  - IRIM-MRIM Hg104: 0.0722
  - IRIM-MRIM Opp_sift_dense_1000: 0.1165
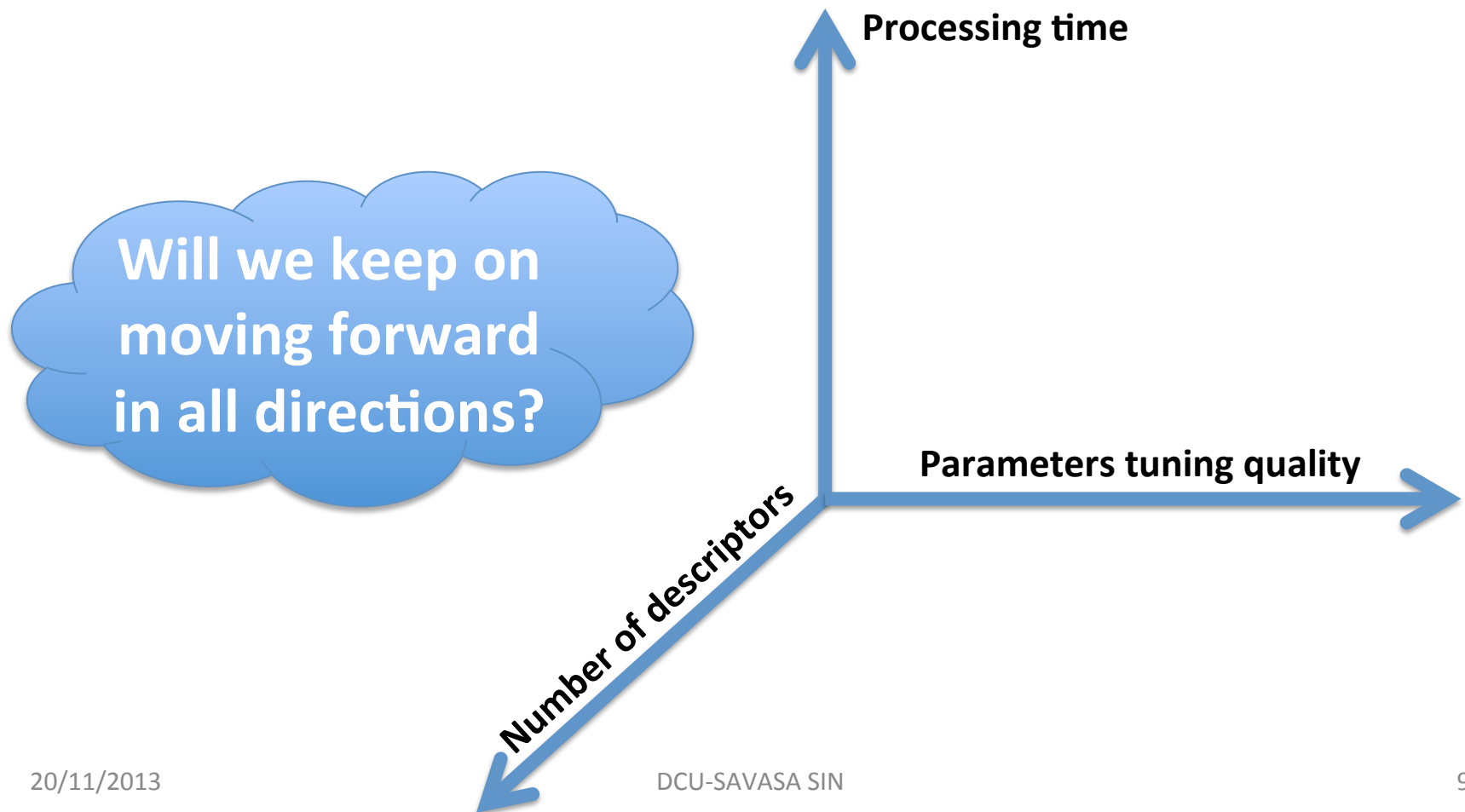 - Weighted late fusion: 0.1320
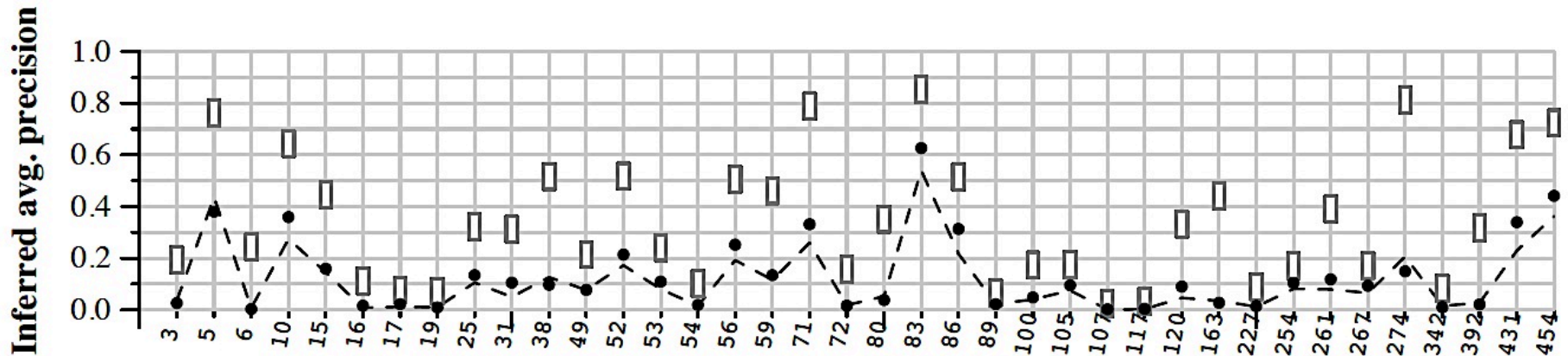
Reflections and questions

- SIN-like tasks are becoming a joint research-engineering challenges

- No magical solution!
  - The use of multiple descriptors and multiple classification methods is unavoidable! (?)

- Very high entry barrier for new participants
  - Can we lower the entry barrier?
  - Providing descriptors by IRIM was very helpful and time saving, thank you!
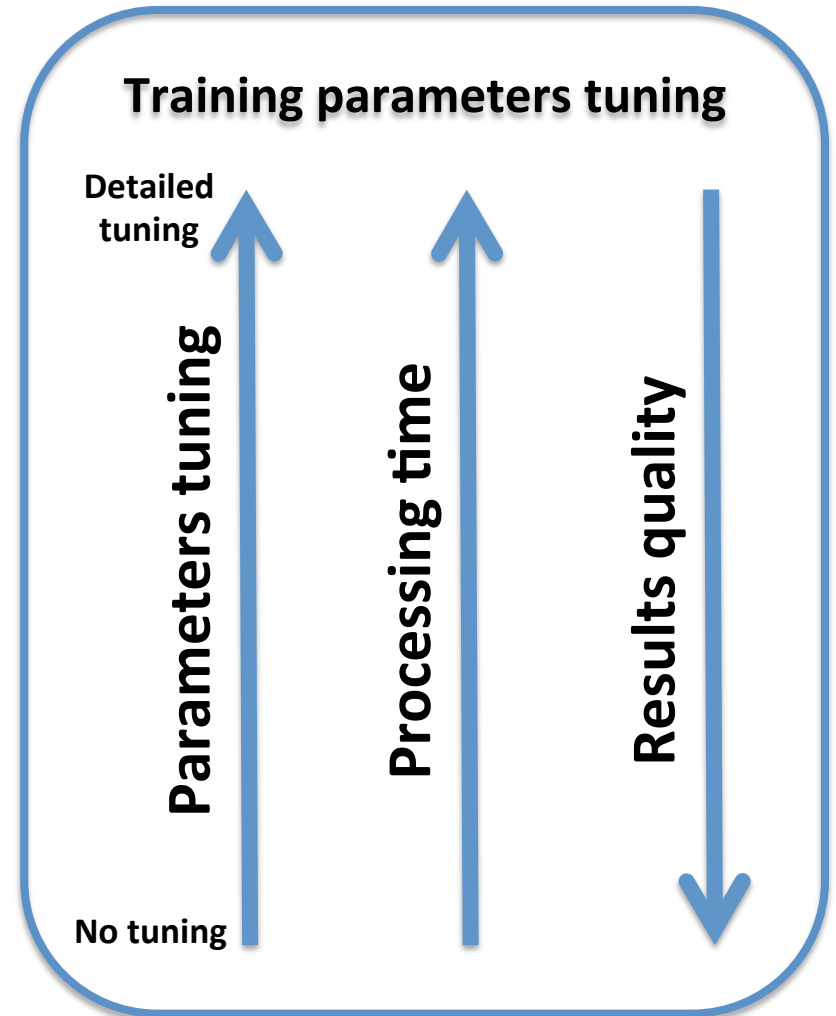 - Can we do similar initiatives for training?

Reflections and questions

- What do we need to do?

Processing time

Will we keep on moving forward in all directions?

Parameters tuning quality

Number of descriptors

Reflections and questions

Number of Descriptor

Descriptors

Processing time

Results quality

Training parameters tuning

Detailed tuning

Parameters tuning

Processing time

Results quality

No tuning

Reflections and questions

 # Descriptors

Parameters tuning
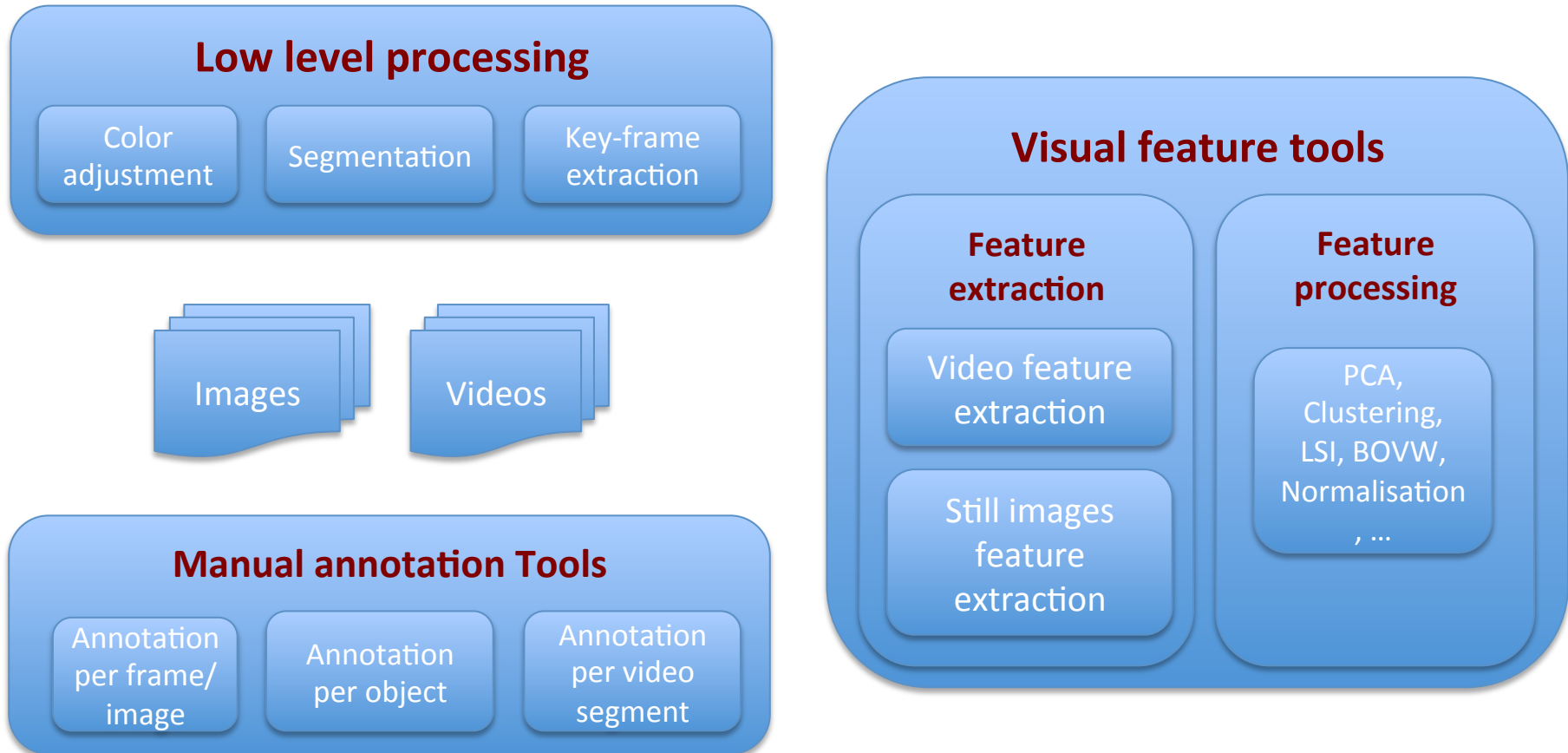
Processing time
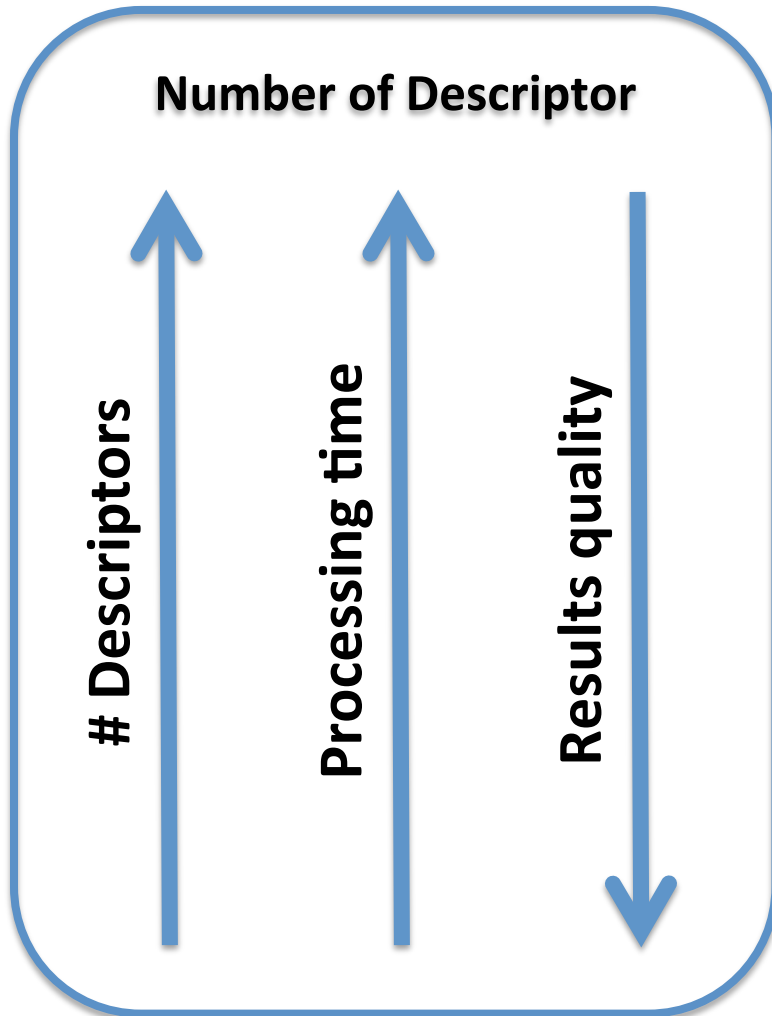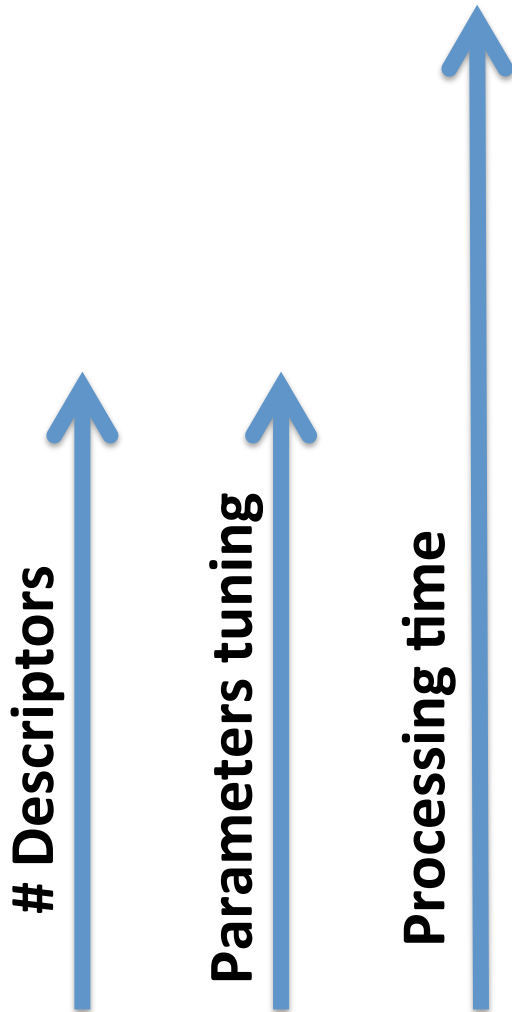
- More focus need to be done on the hardware infrastructure

- Feature extraction can be a centralised task?
  - To promote classification research

- Classifiers can be provided?
  - To promote fusion research

Reflections and questions

- If the hardware infrastructure is a solved problem in the future, on what we have to focus now?

- Using very advanced infrastructure can solve some problems but is it feasible in real-life scenarios (e.g. CCTV analysis companies)
    - Reformulating some questions: How well can we achieve with X core, and Y hours?

DCU's framework for visual big data processing (2014)

Low level processing

Color adjustment

Segmentation

Key-frame extraction

Images

Videos

Manual annotation Tools

Annotation per frame/image

Annotation per object

Annotation per video segment

Visual feature tools

Feature extraction

Video feature extraction

Still images feature extraction

Feature processing

PCA, Clustering, LSI, BOVW, Normalisation, …

DCU's framework for visual big data processing (2014)

Machine learning

Discriminative learning

SVM, Neural networks, logistic regression, Boosting

Generative learning

GMM, Hidden Markov model, LDA, Naive Bays

High level semantic extraction

Object/Scene classification
Object recognition/ localisation
Face recognition
Person identification ….

Retrieval and matching

Query by keyword or by content

Visual features/ descriptor matching RANSAC, distance metrics

DCU's framework for visual big data processing (2014)

- The framework will be hosted in the servers of The Irish Centre for High-End Computing (ICHEC)

	Class A "High Impact"	Class B "Regular"	Class B "GPU"	Class C "Discovery"
Max Processor Core Hours	4,000,000	600,000	300,000	30,000
Max Storage	1,500 GB	500 GB	500 GB	50 GB
Max Project Duration	24 Months	18 Months	18 Months	12 Months

Thank you!