



IN PARTNERSHIP WITH:  
**Ecole Centrale Paris**

Activity Report 2015

## **Project-Team GALEN**

Organ Modeling through Extraction,  
Representation and Understanding of Medical  
Image Content

RESEARCH CENTER  
**Saclay - Île-de-France**

THEME  
**Computational Neuroscience and  
Medicine**



## Table of contents

<b>1. Members</b>	<b>1</b>
<b>2. Overall Objectives</b>	<b>2</b>
<b>3. Research Program</b>	<b>2</b>
3.1. Shape, Grouping and Recognition	2
3.2. Machine Learning & Structured Prediction	4
3.3. Self-Paced Learning with Missing Information	4
3.4. Discrete Biomedical Image Perception	5
<b>4. Application Domains</b>	<b>6</b>
4.1. Lung Tumor Detection and Characterization	6
4.2. Co-segmentation and Co-registration of Subcortical Brain Structures	6
<b>5. Highlights of the Year</b>	<b>7</b>
<b>6. New Software and Platforms</b>	<b>7</b>
6.1. DISD	7
6.2. DPMS	7
6.3. DROP	7
6.4. FastPD	8
6.5. GraPeS	8
6.6. HOAP-SVM	8
6.7. LBSD	9
6.8. TeXMeG	9
6.9. mrf-registration	9
<b>7. New Results</b>	<b>9</b>
7.1. Optimizing Average Precision	9
7.2. Region-based Semantic Segmentation	10
7.3. Parsimonious Labeling	10
7.4. Structured Learning and Inference	10
7.5. Novel graph kernels	10
7.6. Structured Sparsity Regularization & Statistical Hypothesis Testing	10
7.7. High-Order MRF models	11
7.8. Graph-based registration and segmentation	11
7.9. Object Detection from RGB-Depth images	11
7.10. Deep CNN for Modelling Deformations and Semantic Segmentation	12
7.11. Learning Low-level Image Representations with Deep CNN	12
7.12. Human-Limb Segmentation for Intelligent Mobility Assistance Robots	12
<b>8. Partnerships and Cooperations</b>	<b>13</b>
8.1. Regional Initiatives	13
8.2. National Initiatives	13
8.3. European Initiatives	14
8.3.1.1. DIOCLES	14
8.3.1.2. I-SUPPORT	14
8.3.1.3. MOBOT	15
8.3.1.4. RECONFIG	16
8.3.1.5. Strategie	16
8.4. International Initiatives	17
8.5. International Research Visitors	17
<b>9. Dissemination</b>	<b>17</b>
9.1. Promoting Scientific Activities	17
9.1.1. Scientific events organisation	17
9.1.2. Scientific events selection	18

9.1.2.1.	Chair of conference program committees	18
9.1.2.2.	Member of the conference program committees	18
9.1.2.3.	Reviewer	18
9.1.3.	Journal	18
9.1.3.1.	Member of the editorial boards	18
9.1.3.2.	Reviewer - Reviewing activities	18
9.1.4.	Invited talks	18
9.1.5.	Scientific expertise	18
9.2.	Teaching - Supervision - Juries	19
9.2.1.	Teaching	19
9.2.2.	Supervision	19
9.2.3.	Juries	20
9.3.	Popularization	20
<b>10.</b>	<b>Bibliography</b> .....	<b>20</b>

# Project-Team GALEN

*Creation of the Team: 2008 February 15, updated into Project-Team: 2014 January 01*

## Keywords:

### Computer Science and Digital Science:

3. - Data and knowledge
5. - Interaction, multimedia and robotics
8. - Artificial intelligence

### Other Research Topics and Application Domains:

1. - Life sciences
2. - Health

## 1. Members

### Research Scientist

Matthew Blaschko [Inria, Researcher, HdR]

### Faculty Members

Nikos Paragios [Team leader, CentraleSupélec, Professor, HdR]

Iasonas Kokkinos [CentraleSupélec, Associate Professor, HdR]

Pawan Kumar Mudigonda [CentraleSupélec, Associate Professor, until 31/08/2015, HdR]

Dimitrios Samaras [CentraleSupélec, Visiting Professor]

### Engineer

Rafael Marini Silva [CentraleSupélec]

### PhD Students

Puneet Dokania [Inria, from Oct 2015 until Dec 2015]

Stavros Alchatzidis [CentraleSupélec]

Eugène Belilovsky [CentraleSupélec]

Maxim Berman [CentraleSupélec]

Diane Bouchacourt [CentraleSupélec]

Wacha Bounliphone [CentraleSupélec]

Siddhartha Chandra [Inria]

Vivien Fecamp [Ecole CentraleSupélec]

Enzo Ferrante [CentraleSupélec]

Riza Guler [Inria, from Sep 2015]

Hariprasad Kannan [CentraleSupélec]

Stefan Kinauer [CentraleSupélec]

Evgenios Kornaropoulos [CentraleSupélec]

Huu Dien Khue Le [CentraleSupélec]

Yohann Salaun [CentraleSupélec]

Abhishek Sharma [Inria, until Jul 2015]

Stavros Tsogkas [CentraleSupélec]

Jiaqian Yu [CentraleSupélec]

### Visiting Scientist

Evangelia Zacharaki [CentraleSupélec, from Sep 2015]

### Administrative Assistant

Alexandra Merlin [Inria]

**Others**

Adam Harley [Inria, INTERN, from Jun 2015]  
Mihir Sahasrabudhe [CentraleSupélec, INTERN, from Mar 2015 until Dec 2015]  
Mahsa Shakeri [CentraleSupélec, INTERN, from May 2015 until Sep 2015]  
Sharma Abhishek [Inria, INTERN, until Jul 2015]  
Natalia Leclercq [CentraleSupélec]

## 2. Overall Objectives

### 2.1. GALEN@Centrale-Paris

Computational vision is one of the most challenging research domains in engineering sciences. The aim is to reproduce human visual perception through intelligent processing of visual data. The application domains span from computer aided diagnosis to industrial automation & robotics. The most common mathematical formulation to address such a challenge is through mathematical modeling. In such a context, first the solution of the desired vision task is expressed in the form of a parameterized mathematical model. Given such a model, the next task consists of associating the model parameters with the available observations, which is often called the model-to-data association. The aim of this task is to determine the impact of a parameter choice to the observations and eventually maximize/minimize the adequacy of these parameters with the visual observations. In simple words, the better the solution is, the better it will be able to express and fit the data. This is often achieved through the definition of an objective function on the parametric space of the model. Last, but not least given the definition of the objective function, visual perception is addressed through its optimization with respect to the model parameters. To summarize, computation visual perception involves three aspects, a task-specific definition of a parametric model, a data-specific association of this model with the available observations and last the optimization of the model parameters given the objective and the observations.

Such a chain processing inherits important shortcomings. The curse of dimensionality is often used to express the importance of the model complexity. In simple words, the higher the complexity of the model is, the better its expressive power will be with counter effect the increase of the difficulty of the inference process. Non-linearity is another issue to be addressed which simply states that the association between the model and the data is a (highly) non-linear function and therefore direct inference is almost infeasible. The impact of this aspect is enforced from the curse of non-convexity that characterizes the objective function. Often it lives in high-dimensional spaces and is ill posed making exact inference problematic (in many cases not possible) and computationally expensive. Last, but not least modularity and scalability is another important concern to be addressed in the context of computational vision. The use of task-specific modeling and algorithmic solutions make their portability infeasible and therefore transfer of knowledge from one task to another is not straightforward while the methods do not always scale well with respect either to the dimensionality of the representation or the data.

GALEN aims at proposing innovative techniques towards automatic structuring, interpretation and longitudinal modeling of visual data. In order to address these fundamental problems of computational perception, GALEN investigates the use of discrete models of varying complexity. These methods exhibit an important number of strengths such as their ability to be modular with respect to the input measurements (clinical data), the nature of the model (certain constraints are imposed from computational perspective in terms of the level of interactions), and the model-to-data association while being computationally efficient.

## 3. Research Program

### 3.1. Shape, Grouping and Recognition

A general framework for the fundamental problems of image segmentation, object recognition and scene analysis is the interpretation of an image in terms of a set of symbols and relations among them. Abstractly

stated, image interpretation amounts to mapping an observed image,  $X$  to a set of symbols  $Y$ . Of particular interest are the symbols  $Y^*$  that *optimally explain the underlying image*, as measured by a scoring function  $s$  that aims at distinguishing correct (consistent with human labellings) from incorrect interpretations:

$$Y^* = \operatorname{argmax}_Y s(X, Y) \quad (1)$$

Applying this framework requires (a) identifying which symbols and relations to use (b) learning a scoring function  $s$  from training data and (c) optimizing over  $Y$  in Eq. 1.

One of the main themes of our work is the development of methods that jointly address (a,b,c) in a shape-grouping framework in order to reliably extract, describe, model and detect shape information from natural and medical images. A principal motivation for using a shape-based framework is the understanding that shape- and more generally, grouping- based representations can go all the way from image features to objects. Regarding aspect (a), image representation, we cater for the extraction of image features that respect the shape properties of image structures. Such features are typically constructed to be purely geometric (e.g. boundaries, symmetry axes, image segments), or appearance-based, such as image descriptors. The use of machine learning has been shown to facilitate the robust and efficient extraction of such features, while the grouping of local evidence is known to be necessary to disambiguate the potentially noisy local measurements. In our research we have worked on improving feature extraction, proposing novel blends of invariant geometric- and appearance- based features, as well as grouping algorithms that allow for the efficient construction of optimal assemblies of local features.

Regarding aspect (b) we have worked on learning scoring functions for detection with deformable models that can exploit the developed low-level representations, while also being amenable to efficient optimization. Our works in this direction build on the graph-based framework to construct models that reflect the shape properties of the structure being modeled. We have used discriminative learning to exploit boundary- and symmetry-based representations for the construction of hierarchical models for shape detection, while for medical images we have developed methods for the end-to-end discriminative training of deformable contour models that combine low-level descriptors with contour-based organ boundary representations.

Regarding aspect (c) we have developed algorithms which implement top-down/bottom-up computation both in deterministic and stochastic optimization. The main idea is that ‘bottom-up’, image-based guidance is necessary for efficient detection, while ‘top-down’, object-based knowledge can disambiguate and help reliably interpret a given image; a combination of both modes of operation is necessary to combine accuracy with efficiency. In particular we have developed novel techniques for object detection that employ combinatorial optimization tools (A\* and Branch-and-Bound) to tame the combinatorial complexity, achieving a best-case performance that is logarithmic in the number of pixels.

In the long run we aim at scaling up shape-based methods to 3D detection and pose estimation and large-scale object detection. One aspect which seems central to this is the development of appropriate mid-level representations. This is a problem that has received increased interest lately in the 2D case and is relatively mature, but in 3D it has been pursued primarily through ad-hoc schemes. We anticipate that questions pertaining to part sharing in 3D will be addressed most successfully by relying on explicit 3D representations. On the one hand depth sensors, such as Microsoft’s Kinect, are now cheap enough to bring surface modeling and matching into the mainstream of computer vision - so these advances may be directly exploitable at test time for detection. On the other hand, even if we do not use depth information at test time, having 3D information can simplify the modeling task during training. In on-going work with collaborators we have started exploring combinations of such aspects, namely (i) the use of surface analysis tools to match surfaces from depth sensors (ii) using branch-and-bound for efficient inference in 3D space and (iii) groupwise-registration to build statistical 3D surface models. In the coming years we intend to pursue a tighter integration of these different directions for scalable 3D object recognition.

## 3.2. Machine Learning & Structured Prediction

The foundation of statistical inference is to learn a function that minimizes the expected loss of a prediction with respect to some unknown distribution

$$\mathcal{R}(f) = \int \ell(f, x, y) dP(x, y), \quad (2)$$

where  $\ell(f, x, y)$  is a problem specific loss function that encodes a penalty for predicting  $f(x)$  when the correct prediction is  $y$ . In our case, we consider  $x$  to be a medical image, and  $y$  to be some prediction, e.g. the segmentation of a tumor, or a kinematic model of the skeleton. The loss function,  $\ell$ , is informed by the costs associated with making a specific misprediction. As a concrete example, if the true spatial extent of a tumor is encoded in  $y$ ,  $f(x)$  may make mistakes in classifying healthy tissue as a tumor, and mistakes in classifying diseased tissue as healthy. The loss function should encode the potential physiological damage resulting from erroneously targeting healthy tissue for irradiation, as well as the risk from missing a portion of the tumor.

A key problem is that the distribution  $P$  is unknown, and any algorithm that is to estimate  $f$  from labeled training examples must additionally make an implicit estimate of  $P$ . A central technology of empirical inference is to approximate  $\mathcal{R}(f)$  with the empirical risk,

$$\mathcal{R}(f) \approx \widehat{\mathcal{R}}(f) = \frac{1}{n} \sum_{i=1}^n \ell(f, x_i, y_i), \quad (3)$$

which makes an implicit assumption that the training samples  $(x_i, y_i)$  are drawn i.i.d. from  $P$ . Direct minimization of  $\widehat{\mathcal{R}}(f)$  leads to overfitting when the function class  $f \in \mathcal{F}$  is too rich, and regularization is required:

$$\min_{f \in \mathcal{F}} \lambda \Omega(\|f\|) + \widehat{\mathcal{R}}(f), \quad (4)$$

where  $\Omega$  is a monotonically increasing function that penalizes complex functions.

Equation Eq. 4 is very well studied in classical statistics for the case that the output,  $y \in \mathcal{Y}$ , is a binary or scalar prediction, but this is not the case in most medical imaging prediction tasks of interest. Instead, complex interdependencies in the output space leads to difficulties in modeling inference as a binary prediction problem. One may attempt to model e.g. tumor segmentation as a series of binary predictions at each voxel in a medical image, but this violates the i.i.d. sampling assumption implicit in Equation Eq. 3. Furthermore, we typically gain performance by appropriately modeling the inter-relationships between voxel predictions, e.g. by incorporating pairwise and higher order potentials that encode prior knowledge about the problem domain. It is in this context that we develop statistical methods appropriate to structured prediction in the medical imaging setting.

## 3.3. Self-Paced Learning with Missing Information

Many tasks in artificial intelligence are solved by building a model whose parameters encode the prior domain knowledge and the likelihood of the observed data. In order to use such models in practice, we need to estimate its parameters automatically using training data. The most prevalent paradigm of parameter estimation is supervised learning, which requires the collection of the inputs  $x_i$  and the desired outputs  $y_i$ . However, such an approach has two main disadvantages. First, obtaining the ground-truth annotation of high-level applications, such as a tight bounding box around all the objects present in an image, is often expensive. This prohibits the use of a large training dataset, which is essential for learning the existing complex models. Second, in many applications, particularly in the field of medical image analysis, obtaining the ground-truth annotation may not be feasible. For example, even the experts may disagree on the correct segmentation of a microscopical image due to the similarities between the appearance of the foreground and background.



In order to address the deficiencies of supervised learning, researchers have started to focus on the problem of parameter estimation with data that contains hidden variables. The hidden variables model the missing information in the annotations. Obtaining such data is practically more feasible: image-level labels (‘contains car’, ‘does not contain person’) instead of tight bounding boxes; partial segmentation of medical images. Formally, the parameters  $\mathbf{w}$  of the model are learned by minimizing the following objective:

$$\min_{\mathbf{w} \in \mathcal{W}} R(\mathbf{w}) + \sum_{i=1}^n \Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w})). \quad (5)$$

Here,  $\mathcal{W}$  represents the space of all parameters,  $n$  is the number of training samples,  $R(\cdot)$  is a regularization function, and  $\Delta(\cdot)$  is a measure of the difference between the ground-truth output  $y_i$  and the predicted output and hidden variable pair  $(y_i(\mathbf{w}), h_i(\mathbf{w}))$ .

Previous attempts at minimizing the above objective function treat all the training samples equally. This is in stark contrast to how a child learns: first focus on easy samples (‘learn to add two natural numbers’) before moving on to more complex samples (‘learn to add two complex numbers’). In our work, we capture this intuition using a novel, iterative algorithm called self-paced learning (SPL). At an iteration  $t$ , SPL minimizes the following objective function:

$$\min_{\mathbf{w} \in \mathcal{W}, \mathbf{v} \in \{0,1\}^n} R(\mathbf{w}) + \sum_{i=1}^n v_i \Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w})) - \mu_t \sum_{i=1}^n v_i. \quad (6)$$

Here, samples with  $v_i = 0$  are discarded during the iteration  $t$ , since the corresponding loss is multiplied by 0. The term  $\mu_t$  is a threshold that governs how many samples are discarded. It is annealed at each iteration, allowing the learner to estimate the parameters using more and more samples, until all samples are used. Our results already demonstrate that SPL estimates accurate parameters for various applications such as image classification, discriminative motif finding, handwritten digit recognition and semantic segmentation. We will investigate the use of SPL to estimate the parameters of the models of medical imaging applications, such as segmentation and registration, that are being developed in the GALEN team. The ability to handle missing information is extremely important in this domain due to the similarities between foreground and background appearances (which results in ambiguities in annotations). We will also develop methods that are capable of minimizing more general loss functions that depend on the (unknown) value of the hidden variables, that is,

$$\min_{\mathbf{w} \in \mathcal{W}, \theta \in \Theta} R(\mathbf{w}) + \sum_{i=1}^n \sum_{h_i \in \mathcal{H}} \Pr(h_i | x_i, y_i; \theta) \Delta(y_i, h_i, y_i(\mathbf{w}), h_i(\mathbf{w})). \quad (7)$$

Here,  $\theta$  is the parameter vector of the distribution of the hidden variables  $h_i$  given the input  $x_i$  and output  $y_i$ , and needs to be estimated together with the model parameters  $\mathbf{w}$ . The use of a more general loss function will allow us to better exploit the freely available data with missing information. For example, consider the case where  $y_i$  is a binary indicator for the presence of a type of cell in a microscopical image, and  $h_i$  is a tight bounding box around the cell. While the loss function  $\Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w}))$  can be used to learn to classify an image as containing a particular cell or not, the more general loss function  $\Delta(y_i, h_i, y_i(\mathbf{w}), h_i(\mathbf{w}))$  can be used to learn to detect the cell as well (since  $h_i$  models its location)

### 3.4. Discrete Biomedical Image Perception

A wide variety of tasks in medical image analysis can be formulated as discrete labeling problems. In very simple terms, a discrete optimization problem can be stated as follows: we are given a discrete set of variables  $\mathcal{V}$ , all of which are vertices in a graph  $\mathcal{G}$ . The edges of this graph (denoted by  $\mathcal{E}$ ) encode the variables’ relationships. We are also given as input a discrete set of labels  $\mathcal{L}$ . We must then assign one label from  $\mathcal{L}$  to each variable in  $\mathcal{V}$ . However, each time we choose to assign a label, say,  $x_{p_1}$  to a variable  $p_1$ , we are forced to pay a price according to the so-called *singleton* potential function  $g_p(x_p)$ , while each time we choose to

assign a pair of labels, say,  $x_{p_1}$  and  $x_{p_2}$  to two interrelated variables  $p_1$  and  $p_2$  (two nodes that are connected by an edge in the graph  $\mathcal{G}$ ), we are also forced to pay another price, which is now determined by the so called *pairwise* potential function  $f_{p_1 p_2}(x_{p_1}, x_{p_2})$ . Both the singleton and pairwise potential functions are problem specific and are thus assumed to be provided as input.

Our goal is then to choose a labeling which will allow us to pay the smallest total price. In other words, based on what we have mentioned above, we want to choose a labeling that minimizes the sum of all the MRF potentials, or equivalently the MRF energy. This amounts to solving the following optimization problem:

$$\arg \min_{\{x_p\}} \mathcal{P}(g, f) = \sum_{p \in \mathcal{V}} g_p(x_p) + \sum_{(p_1, p_2) \in \mathcal{E}} f_{p_1 p_2}(x_{p_1}, x_{p_2}). \quad (8)$$

The use of such a model can describe a number of challenging problems in medical image analysis. However these simplistic models can only account for simple interactions between variables, a rather constrained scenario for high-level medical imaging perception tasks. One can augment the expression power of this model through higher order interactions between variables, or a number of cliques  $\{C_i, i \in [1, n] = \{p_{i^1}, \dots, p_{i^{|C_i|}}\}\}$  of order  $|C_i|$  that will augment the definition of  $\mathcal{V}$  and will introduce hyper-vertices:

$$\arg \min_{\{x_p\}} \mathcal{P}(g, f) = \sum_{p \in \mathcal{V}} g_p(x_p) + \sum_{(p_1, p_2) \in \mathcal{E}} f_{p_1 p_2}(x_{p_1}, x_{p_2}) + \sum_{C_i \in \mathcal{E}} f_{p_1 \dots p_n}(x_{p_{i^1}}, \dots, x_{p_{i^{|C_i|}}}). \quad (9)$$

where  $f_{p_1 \dots p_n}$  is the price to pay for associating the labels  $(x_{p_{i^1}}, \dots, x_{p_{i^{|C_i|}}})$  to the nodes  $(p_1 \dots p_{i^{|C_i|}})$ . Parameter inference, addressed by minimizing the problem above, is the most critical aspect in computational medicine and efficient optimization algorithms are to be evaluated both in terms of computational complexity as well as of inference performance. State of the art methods include deterministic and non-deterministic annealing, genetic algorithms, max-flow/min-cut techniques and relaxation. These methods offer certain strengths while exhibiting certain limitations, mostly related to the amount of interactions which can be tolerated among neighborhood nodes. In the area of medical imaging where domain knowledge is quite strong, one would expect that such interactions should be enforced at the largest scale possible.

## 4. Application Domains

### 4.1. Lung Tumor Detection and Characterization

The use of Diffusion Weighted MR Imaging (DWI) is investigated as an alternative tool to radiologists for tumor detection, tumor characterization, distinguishing tumor tissue from non-tumor tissue, and monitoring and predicting treatment response. In collaboration with Hôpitaux Universitaires Henri-Mondor in Paris, France and Chang Gung Memorial Hospital – Linkou in Taipei, Taiwan we investigate the use of model-based methods of 3D image registration, clustering and segmentation towards the development of a framework for automatic interpretation of images, and in particular extraction of meaningful biomarkers in aggressive lymphomas.

### 4.2. Co-segmentation and Co-registration of Subcortical Brain Structures

New algorithms to perform co-segmentation and co-registration of subcortical brain structures on MRI images are investigated in collaboration with Ecole Polytechnique de Montreal and the Sainte-Justine Hospital Research Center from Montreal. Brain subcortical structures are involved in different neurodegenerative and neuropsychiatric disorders, including schizophrenia, Alzheimers disease, attention deficit, and subtypes of epilepsy. Segmenting these parts of the brain enables a physician to extract indicators, facilitating their quantitative analysis and characterization. We are investigating how estimated maps of semantic labels (obtained using machine learning techniques) can be used as a surrogate for unlabelled data. We are exploring how to combine them with multi-population deformable registration to improve both alignment and segmentation of these challenging brain structures.

## 5. Highlights of the Year

### 5.1. Highlights of the Year

#### 5.1.1. Awards

- Pr. Iasonas Kokkinos was appointed associate editor for the Computer Vision and Image Understanding Journal.
- Pr. Pawan Kumar was appointed associate editor for the Computer Vision and Image Understanding Journal.
- Pr. Nikos Paragios was admitted as a senior fellow at the Institut Universitaire de France in the section of Mathematics.

## 6. New Software and Platforms

### 6.1. DISD

Dense Image and Surface Descriptors

FUNCTIONAL DESCRIPTION

Scale-Invariant Descriptor, Scale-Invariant Heat Kernel Signatures DISD implements the SID, SI-HKS and ISC descriptors. SID (Scale-Invariant Descriptor) is a densely computable, scale- and rotation- invariant descriptor. We use a log-polar grid around every point to turn rotation/scalings into translation, and then use the Fourier Transform Modulus (FTM) to achieve invariance. SI-HKS (Scale-Invariant Heat Kernel Signatures) extract scale-invariant shape signatures by exploiting the fact that surface scaling amounts to multiplication and scaling of a properly sampled HKS descriptor. We apply the FTM trick on HKS to achieve invariance to scale changes. ISC (Intrinsic Shape Context) constructs a net-like grid around every surface point by shooting outwards and tracking geodesics. This allows us to build a meta-descriptor on top of HKS/SI-HKS that takes neighborhood into account, while being invariant to surface isometries.

- Participants: Iasonas Kokkinos and Eduard Trulls
- Contact: Iasonas Kokkinos
- URL: <http://vision.mas.ecp.fr/Personnel/iasonas/descriptors.html>

### 6.2. DPMS

FUNCTIONAL DESCRIPTION

Dpms implements branch-and-bound object detection, cutting down the complexity of detection from linear in the number of pixels to logarithmic.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos
- URL: <http://cvn.ecp.fr/personnel/iasonas/dpms.html>

### 6.3. DROP

KEYWORDS: Health - Merging - Registration of 2D and 3D multimodal images - Medical imaging

FUNCTIONAL DESCRIPTION

Drop is a software programme that registers images originating from one or more modes by quickly and efficiently calculating a non-rigid / deformable field of deformation. Drop is a new, quick and effective registration tool based on new algorithms that do not require a cost function derivative.

- Partner: Centrale Paris
- Contact: Nikolaos Paragios
- URL: <http://campar.in.tum.de/Main/Drop>

## 6.4. FastPD

KEYWORD: Medical imaging

FUNCTIONAL DESCRIPTION

FastPD is an optimization platform in C++ for the computer vision and medical imaging community.

- Contact: Nikolaos Paragios
- URL: <http://www.csd.uoc.gr/~komod/FastPD/>

## 6.5. GraPeS

Grammar Parser for Shapes

FUNCTIONAL DESCRIPTION It is a software for parsing facade images using shape grammars. Grapes implement a parsing methods based on Reinforcement Learning principles. It optimizes simultaneously the topology of the parse tree as well as the associated parameters. GraPeS comes along with predefined shape grammars as XML files and defines three kinds of rewards. However, it also offers the possibility to create new grammars and to provide custom rewards in text files, widening the scope of potential applications. The name of the software comes from the aspect of the parse tree of the binary split grammars involved in the process.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos
- URL: <http://vision.mas.ecp.fr/Personnel/teboul/grapesPage/index.php>

## 6.6. HOAP-SVM

High-Order Average Precision SVM

SCIENTIFIC DESCRIPTION

We consider the problem of using high-order information (for example, persons in the same image tend to perform the same action) to improve the accuracy of ranking (specifically, average precision). We develop two learning frameworks. The high-order binary SVM (HOB-SVM) optimizes a convex upper bound of the surrogate 0-1 loss function. The high-order average precision SVM (HOAP-SVM) optimizes a difference-of-convex upper bound on the average precision loss function.

Authors of the research paper: Puneet K. Dokania, A. Behl, C. V. Jawahar and M. Pawan Kumar

FUNCTIONAL DESCRIPTION

The software provides a convenient API for learning to rank with high-order information. The samples are ranked according to a score that is proportional to the difference of max-marginals of the positive and the negative class. The parameters of the score function are computed by minimizing an upper bound on the average precision loss. The software also provides an instantiation of the API for ranking samples according to their relevance to an action, using the poselet features. The following learning algorithms are included in the API:

(1) Multiclass-SVM (2) AP-SVM (3) High Order Binary SVM (HOB-SVM) (4) High Order AP-SVM (HOAP-SVM) (5) M4 Learning (unpublished work)

The API is developed in C/C++ by Puneet K. Dokania.

- Participants: Puneet Dokania and Pawan Kumar
- Contact: Puneet Dokania
- URL: <http://puneetkdokania.github.io/projects/ranking-highorder/ranking-highorder.html>

## 6.7. LBSD

Learning-Based Symmetry Detection

FUNCTIONAL DESCRIPTION

LBSD implements the learning-based approach to symmetry detection. It includes the code for running a detector, alongside with the ground-truth symmetry annotations that we have introduced for the Berkeley Segmentation Dataset (BSD) benchmark.

- Participant: Stavros Tsogkas
- Contact: Stavros Tsogkas
- URL: [https://github.com/tsogkas/oid\\_1.0](https://github.com/tsogkas/oid_1.0)

## 6.8. TeXMeG

FUNCTIONAL DESCRIPTION

Texture, modulation, generative models, segmentation, TeXMeG is a front-end for texture analysis and edge detection platform in Matlab that relies on Gabor filtering and image demodulation. Includes frequency- and time- based definition of Gabor- and other Quadrature-pair filterbanks, demodulation with the Regularized Energy Separation Algorithm and Texture/Edge/Smooth classification based on MDL criterion.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos
- URL: <http://cvsp.cs.ntua.gr/software/texture/>

## 6.9. mrf-registration

KEYWORDS: Health - Medical imaging

FUNCTIONAL DESCRIPTION

Deformable image and volume registration, is a deformable registration platform in C++ for the medical imaging community. This is the first publicly available platform which contains most of the existing metrics to perform registration under the same concept. The platform is used for clinical research from approximately 3,000 users worldwide.

- Participant: Nikolaos Paragyios
- Contact: Nikolaos Paragyios
- URL: <http://www.mrf-registration.net/>

# 7. New Results

## 7.1. Optimizing Average Precision

**Participants:** Pawan Kumar

Average precision (AP) is one of the most commonly used measures for ranking. However, due to the inefficiency of optimizing it during learning, a common approach is to use surrogate loss functions such as 0-1 loss. We have developed a novel latent AP-SVM classifier [1], that minimizes a carefully designed upper bound on the AP-based loss function over weakly supervised samples. Using publicly available datasets, we demonstrate the advantage of our approach over standard loss-based learning frameworks on three challenging problems: action classification, character recognition and object detection.

## 7.2. Region-based Semantic Segmentation

**Participants:** Pawan Kumar

In [9] we consider the problem of parameter estimation and energy minimization for a region-based semantic segmentation model. The main problem we face in the context of energy minimization, is the large number of putative pixel-to-region assignments. We address this problem by designing an accurate linear programming based approach for selecting the best set of regions from a large dictionary, which is constructed by merging and intersecting segments obtained from multiple bottom-up over-segmentations. The lack of fully supervised data is tackled by using a latent structural SVM formulation, where the latent variables model any missing information in the human annotation. Using large, publicly available datasets we show that our methods are able to significantly improve the accuracy of the region-based model.

## 7.3. Parsimonious Labeling

**Participants:** Puneet Dokania, Pawan Kumar

In [22], we propose a new family of discrete energy minimization problems, which we call parsimonious labeling, that is to use as few labels as possible. This allows us to capture useful cues for important computer vision applications such as stereo correspondence and image denoising. Furthermore, we propose an efficient graph-cuts based algorithm for the parsimonious labeling problem that provides strong theoretical guarantees on the quality of the solution. Using both synthetic and standard real datasets, we show that our algorithm significantly outperforms other graph-cuts based approaches.

## 7.4. Structured Learning and Inference

**Participants:** Jiaqian Yu, Matthew Blaschko

We have developed computationally efficient structured output prediction methods for learning with non-modular losses [19], [29], [40]. We both demonstrate the feasibility of learning with submodular losses, as well as show that learning with multiple correct outputs can lead to NP-hard optimization problems even when learning with a single correct output is feasible.

## 7.5. Novel graph kernels

**Participants:** Katerina Gkirtzou, Matthew Blaschko

We have developed a novel family of graph kernels that are capable of incorporating local curvature properties of 3D meshes [6]. We have additionally demonstrated their application to the modelling of interdependencies between different brain regions in an fMRI based classification task for predicting cocaine addiction.

## 7.6. Structured Sparsity Regularization & Statistical Hypothesis Testing

**Participants:** Eugene Belilovsky, Wacha Bounliphone, Katerina Gkirtzou, Andreas Argyriou, Matthew Blaschko

We have developed novel methods for structured sparsity regularization & hypothesis testing. We have applied these methods to fMRI [3], [2], [36] and the analysis of large medical databases [10]. We have also developed novel statistical hypothesis tests for relative dependency [21], [37] and similarity [14]. We have applied these methods to the problem of identifying relative dependencies between languages using a multi-lingual corpus, and for discovering the relative relationships between gliomas and genetic information. Additionally, we have shown the application of relative tests to the problem of model selection in deep generative models, and currently an important question in machine learning.

## 7.7. High-Order MRF models

**Participants:** Nikos Komodakis, Nikos Paragios

We developed a very general algorithm for structured prediction learning [7] that is able to efficiently handle discrete MRFs/CRFs (including both pairwise and higher-order models) so long as they can admit a decomposition into tractable subproblems. By properly combining dual decomposition with a max-margin learning method, the framework manages to reduce the training of a complex high-order MRF to the parallel training of a series of simple slave MRFs that are much easier to handle.

## 7.8. Graph-based registration and segmentation

**Participants:** Enzo Ferrante, Vivien Fecamp, Aimilia Gastounioli, Bharat Singh, Stavros Alchatzidis, Nikos Paragios

Deformable image registration plays a fundamental role in many clinical applications. We investigated the use of graphical models in the context of a particular type of image registration problem, known as slice-to-volume registration. We introduced a scalable, modular and flexible formulation that can accommodate low-rank [5] and high order [16] terms, that simultaneously selects the plane and estimates the in-plane deformation through a single shot optimization approach. We applied our models on simulated and real-data in the context of ultrasound and magnetic resonance registration, demonstrating the potential of our methods.

We also developed a novel methodology for graph-based motion-driven segmentation [24] and applied it for carotid plaque segmentation in ultrasound images. We identified the plaque region by exploiting kinematic dependencies between the atherosclerotic and the normal arterial wall. The methodology exploits group-wise image registration towards recovering the deformation field, on which information theory criteria are used to determine dominant motion classes and a map reflecting kinematic dependencies, which is then segmented using Markov random fields.

Moreover, in order to address the problem of general purpose multi-modal deformable registration/fusion we developed a novel and robust method using a metric defined in an appropriate sub-space which is adaptive to the image-content/image-modality [18]. We adopted a graph-based formulation that assumes that intensities of corresponding pixels in the two image domains are related through an unknown piece-wise constant linear function. This relation is propagated to an appropriate sub-space (wavelets coefficients) where a criterion that couples the estimation of the deformation field with optimal transport function on the subspace and the smoothness of the deformation is considered.

## 7.9. Object Detection from RGB-Depth images

**Participant:** Siddhartha Chandra, Iasonas Kokkinos

In [11] we explore RGB-Depth representations for the training of Deformable Models. and describe strategies to improve an object detection pipeline by introducing viewpoint based mixture components. Our contributions are threefold. First, we use surface-based object representations (3D mesh models) from available 3D object model repositories to learn strongly supervised viewpoint classifiers. Second, we develop a geometric dataset augmentation scheme that uses scene geometry to ‘take another look’ at the training data, simulating the effect of camera viewpoint changes. Third, to better exploit depth information, we develop a novel depth-based dense feature extraction method that provides a robust statistical description of scene geometry. We evaluate our learned detectors on the common NYU dataset, and demonstrate that each of our advances results in systematic performance improvements over the traditional detection pipeline.

## 7.10. Deep CNN for Modelling Deformations and Semantic Segmentation

**Participant:** Iasonas Kokkinos

Invariance to deformations in Deep Convolutional Neural Networks (DCNN) is commonly achieved by using multiple ‘max-pooling’ (MP) layers. In [26] we show that alternative methods of modeling deformations can improve the accuracy and efficiency of DCNNs. For this, (i) we introduce epitomic convolution as an alternative to the common convolution-MP cascade of DCNNs, (ii) we introduce a Multiple Instance Learning algorithm to accommodate global translation and scaling in image classification and (iii) we develop a DCNN sliding window detector that explicitly, but efficiently, searches over the object’s position, scale, and aspect ratio. We provide competitive image classification and localization results on the ImageNet dataset and object detection results on Pascal VOC2007.

In [25] we bring together methods from DCNNs and probabilistic graphical models for addressing the task of pixel-level classification (“semantic image segmentation”). We overcome the poor localization property of deep networks by combining the responses at the final DCNN layer with a fully connected Conditional Random Field (CRF). Qualitatively, our “DeepLab” system is able to localize segment boundaries at a level of accuracy which is beyond previous methods.

## 7.11. Learning Low-level Image Representations with Deep CNN

**Participant:** Iasonas Kokkinos

In [27] we propose a novel framework for learning local image descriptors in a discriminative manner. For this purpose we explore a siamese architecture of DCNNs, with a Hinge embedding loss on the L2 distance between descriptors. Since a siamese architecture uses pairs rather than single image patches to train, there exist a large number of positive samples and an exponential number of negative samples. We propose to explore this space with a stochastic sampling of the training set, in combination with an aggressive mining strategy over both the positive and negative samples which we denote as “fracking”. We perform a thorough evaluation of the architecture hyper-parameters, and demonstrate large performance gains compared to both standard CNN learning strategies, hand-crafted image descriptors like SIFT, and the state-of-the-art on learned descriptors: up to 2.5x vs SIFT and 1.5x vs the state-of-the-art in terms of the area under the curve (AUC) of the Precision-Recall curve.

In [4] we explore connections between DCNNs and texture understanding. First, instead of focusing on texture instance and material category recognition, we propose a human-interpretable vocabulary of texture attributes to describe common texture patterns, complemented by a new describable texture dataset for benchmarking. Second, we look at the problem of recognizing materials and texture attributes in realistic imaging conditions, including when textures appear in clutter, developing corresponding benchmarks on top of the recently proposed OpenSurfaces dataset. Third, we revisit classic texture representations, including bag-of-visual-words and the Fisher vectors, in the context of deep learning and show that these have excellent efficiency and generalization properties if the convolutional layers of a deep model are used as filter banks. We obtain in this manner state-of-the-art performance in numerous datasets well beyond textures, an efficient method to apply deep features to image regions, as well as benefit in transferring features from one domain to another.

In [35] we propose a new DCNN architecture that learns pixel embeddings, such that pairwise distances between the embeddings can be used to infer whether or not the pixels lie on the same region. That is, for any two pixels on the same object, the embeddings are trained to be similar; for any pair that straddles an object boundary, the embeddings are trained to be dissimilar. Experimental results show that when this embedding network is used in conjunction with a DCNN trained on semantic segmentation, there is a systematic improvement in per-pixel classification accuracy. Our contributions are integrated in the popular Caffe deep learning framework, and consist in straightforward modifications to convolution routines. As such, they can be exploited for any task involving convolution layers.

## 7.12. Human-Limb Segmentation for Intelligent Mobility Assistance Robots

**Participants:** Siddhartha Chandra, Stavros Tsogkas, Iasonas Kokkinos



We developed a computer vision component [12] to be used as part of an intelligent robotic assistant. This component exploits RGB and depth information extracted from Kinect sensors mounted on the robot, to accurately segment human limbs, using fully convolutional neural networks (FCNNs). We trained our network using an in-house Human-Limb dataset composed of video frames, and described a scheme for dynamically exploiting RGB and depth data in a single framework for training and testing. Our method demonstrated promising performance, being very efficient at the same time, with a run-time of 8 frames per second on a single GPU.

## 8. Partnerships and Cooperations

### 8.1. Regional Initiatives

#### 8.1.1. Excellence Clusters

- Program: DIGITEO (Chair)
  - Project acronym: SubSample
  - Project title: Identification and prediction of Salient brain States through probabilistic structure learning towards fusion of imaging and genomic data
  - Duration: 01/2012-12/2015
  - Coordinator: ECP - FR
- Program: DIGITEO
  - Project acronym: SOPRANO
  - Project title: Structured Output Prediction on Large Scale Neuroscience Data
  - Duration: 3/2013-3/2016
  - Coordinator: Ecole Centrale Paris - FR
- Program: MEDICEN
  - Project acronym: ADOC
  - Project title: ADOC – Diagnostic peropératoire numérique en chirurgie du cancer
  - Duration: 11/2011-09/2015
  - Coordinator: LLTECH - FR

### 8.2. National Initiatives

#### 8.2.1. ANR

- Program: ANR Blanc International
  - Project acronym: ADAMANTIUS
  - Project title: Automatic Detection And characterization of residual Masses in pAtients with lymphomas through fusioN of whole-body diffusion-weighted mrI on 3T and 18F-flUorodeoxyglucoSe pet/ct
  - Duration: 9/2012-8/2015
  - Coordinator: CHU Henri Mondor - FR
- Program: ANR JCJC
  - Project acronym: HICORE
  - Project title: Hierarchical COMpositional REpresentations for Computer Vision
  - Duration: 10/2010-9/2014
  - Coordinator: ECP - FR

- Program: ANR JCJC
  - Project acronym: LearnCost
  - Project title: Learning Model Constraints for Structured Prediction
  - Duration: 2014-2018
  - Coordinator: Inria Saclay - FR
- Program: ITMOs Cancer & Technologies pour la santé d'Aviesan / INCa
  - Project acronym: CURATOR
  - Project title: Slice-to-Image Deformable Registration towards Image-based Surgery Navigation & Guidance
  - Duration: 12/2013-11/2015
  - Coordinator: ECP - FR

## 8.3. European Initiatives

### 8.3.1. FP7 & H2020 Projects

#### 8.3.1.1. DIOCLEs

Title: Discrete bIOimaging perCeption for Longitudinal Organ modElling and computEr-aided diagnosiS

Type: FP7

Instrument: European Research Council

Duration: September 2011 - August 2016

Coordinator: Nikos Paragios

Partner: Ecole Centrale de Paris (FR)

Inria contact: Nikos Paragios

Recent hardware developments from the medical device manufacturers have made possible non-invasive/in-vivo acquisition of anatomical and physiological measurements. Despite enormous progress made on the field of biomedical image analysis still a huge gap exists between clinical research and clinical use. The aim of this proposal is three-fold. First we would like to introduce a novel biomedical image perception framework for clinical use towards disease screening and drug evaluation. Such a framework is expected to be modular (can be used in various clinical settings), computationally efficient (would not require specialized hardware), and can provide a quantitative and qualitative anatomo-pathological indices. Second, leverage progress made on the field of machine learning along with novel, efficient, compact representation of clinical bio-markers toward computer aided diagnosis. Last, using these emerging multi-dimensional signals, we would like to perform longitudinal modelling and understanding the effects of aging to a number of organs and diseases that do not present pre-disease indicators such as brain neurological diseases, muscular diseases, certain forms of cancer, etc.

#### 8.3.1.2. I-SUPPORT

Title: ICT-Supported Bath Robots

Programm: FP7

Duration: March 2015 - March 2018

Coordinator: Robotnik Automation S.L.L.

Partners:

Bethanien Krankenhaus - Geriatisches Zentrum - Gemeinnutzige GMBH (Germany)

Fondazione Santa Lucia (Italy)

Institute of Communication and Computer Systems (Greece)  
Karlsruher Institut für Technologie (Germany)  
Theofanis Alexandridis Kai Sia Ee (OMEGATECH) (Greece)  
Robotnik Automation Sll (Spain)  
Scuola Superiore di Studi Universitari E di Perfezionamento Sant'Anna (Italy)  
Frankfurt University of Applied Sciences (Germany)

Inria contact: Iasonas Kokkinos

The I-SUPPORT project envisions the development and integration of an innovative, modular, ICT-supported service robotics system that supports and enhances older adults' motion and force abilities and assists them in successfully, safely and independently completing the entire sequence of bathing tasks, such as properly washing their back, their upper parts, their lower limbs, their buttocks and groin, and to effectively use the towel for drying purposes. Advanced modules of cognition, sensing, context awareness and actuation will be developed and seamlessly integrated into the service robotics system to enable the robotic bathing system to adapt to the frail elderly population' capabilities and the frail elderly to interact in a master-slave mode, thus, performing bathing activities in an intuitive and safe way. Adaptation and integration of state-of-the-art, cost-effective, soft-robotic manipulators will provide the hardware constituents, which, together with advanced human-robot force/compliance control that will be developed within the proposed project, will form the basis for a safe physical human-robot interaction that complies with the most up-to-date safety standards. Human behavioural, sociological, safety, ethical and acceptability aspects, as well as financial factors related to the proposed service robotic infrastructure will be thoroughly investigated and evaluated so that the I-SUPPORT end result is a close-to-market prototype, applicable to realistic living settings.

### 8.3.1.3. *MOBOT*

Title: Intelligent Active MObility Aid RoBOT integrating Multimodal Communication

Programm: FP7

Duration: February 2013 - January 2016

Coordinator: Technische Universität München

Partners:

Bartlomiej Marcin Stanczyk (Poland)  
Athena Research and Innovation Center in Information Communication & Knowledge Technologies (Greece)  
Bethanien Krankenhaus - Geriatisches Zentrum - Gemeinnutzige (Germany)  
Diaplasia Rehabilitation Center (Greece)  
Ecole Centrale des Arts et Manufactures (France)  
Technische Universitaet Muenchen (Germany)  
Ruprecht-Karls-Universitaet Heidelberg (Germany)

Inria contact: Iasonas Kokkinos

Mobility disabilities are prevalent in our ageing society and impede activities important for the independent living of elderly people and their quality of life. The MOBOT project aims at supporting mobility and thus enforcing fitness and vitality by developing intelligent active mobility assistance robots for indoor environments that provide user-centred, context-adaptive and natural support. Our driving concept envisions cognitive robotic assistants that act (a) proactively by realizing an autonomous and context-specific monitoring of human activities and by subsequently reasoning on meaningful user behavioural patterns, as well as (b) adaptively and interactively, by analysing multi-sensory and physiological signals related to gait and postural stability, and by performing adaptive compliance control for optimal physical support and active fall prevention. Towards these

targets, a multimodal action recognition system will be developed to monitor, analyse and predict user actions with a high level of accuracy and detail. The main thrust of our approach will be the enhancement of computer vision techniques with modalities such as range sensor images, haptic information as well as command-level speech and gesture recognition. Data-driven multimodal human behaviour analysis will be conducted and behavioural patterns will be extracted. Findings will be imported into a multimodal human-robot communication system, involving both verbal and nonverbal communication and will be conceptually and systemically synthesised into mobility assistance models taking into consideration safety critical requirements. All these modules will be incorporated in a behaviour-based and context-aware robot control framework. Direct involvement of end-user groups will ensure that actual user needs are addressed. Finally, user trials will be conducted to evaluate and benchmark the overall system and to demonstrate the vital role of MOBOT technologies for Europe's service robotics.

#### 8.3.1.4. RECONFIG

Type: FP7

Defi: Cognitive Systems and Robotics

Instrument: Specific Targeted Research Project

Objectif: Cognitive Systems and Robotics

Duration: February 2013 - January 2016

Coordinator: Dimos Dimarogonas

Partner: KTH (SE)

Inria contact: Iasonas Kokkinos

The RECONFIG project aims at exploiting recent developments in vision, robotics, and control to tackle coordination in heterogeneous multi-robot systems. Such systems hold promise for achieving robustness by leveraging upon the complementary capabilities of different agents and efficiency by allowing sub-tasks to be completed by the most suitable agent. A key challenge is that agent composition in current multi-robot systems needs to be constant and pre-defined. Moreover, the coordination of heterogeneous multi-agent systems has not been considered in manipulative scenarios. We propose a reconfigurable and adaptive decentralized coordination framework for heterogeneous multiple & multi-DOF robot systems. Agent coordination is held via two types of information exchange: (i) at an implicit level, e.g., when robots are in contact with each other and can sense the contact, and (ii) at an explicit level, using symbols grounded to each embodiment, e.g, when one robot notifies one other about the existence of an object of interest in its vicinity.

#### 8.3.1.5. Strategie

Title: Statistically Efficient Structured Prediction for Computer Vision and Medical Imaging

Programm: FP7

Duration: January 2014 - December 2017

Coordinator: Inria

Inria contact: Matthew Blaschko

'Inference in medical imaging is an important step for disease diagnosis, tissue segmentation, alignment with an anatomical atlas, and a wide range of other applications. However, imperfections in imaging sensors, physical limitations of imaging technologies, and variation in the human population mean that statistical methods are essential for high performance. Statistical learning makes use of human provided ground truth to enable computers to automatically make predictions on future examples without human intervention. At the heart of statistical learning methods is risk minimization - the minimization of the expected loss on a previously unseen image. Textbook methods in statistical learning are not generally designed to minimize the expected loss for loss functions appropriate to medical imaging, which may be asymmetric and non-modular. Furthermore, these methods often do not have the capacity to model interdependencies in the prediction space,

such as those arising from spatial priors, and constraints arising from the volumetric layout of human anatomy. We aim to develop new statistical learning methods that have these capabilities, to develop efficient learning algorithms, to apply them to a key task in medical imaging (tumor segmentation), and to prove their convergence to optimal predictors. To achieve this, we will leverage the structured prediction framework, which has shown impressive empirical results on a wide range of learning tasks. While theoretical results giving learning rates are available for some algorithms, necessary and sufficient conditions for consistency are not known for structured prediction. We will consequently address this issue, which is of key importance for algorithms that will be applied to life critical applications, e.g. segmentation of brain tumors that will subsequently be targeted by radiation therapy or removed by surgery. Project components will address both theoretical and practical issues.'

## 8.4. International Initiatives

### 8.4.1. Inria International Partners

#### 8.4.1.1. Informal International Partners

- University of Oxford – Collaborative research with Andrea Vedaldi. Collaboration Topic: Deep Learning for Texture Recognition.
- Google Research – Collaborative research with George Papandreou. Collaboration Topic: Deep Learning for Semantic Segmentation.
- Universitat Polytechnical de Catalonia – Collaborative research with Francesc Moreno. Collaboration Topic: Deep Learning for Image Descriptors.
- Ecole Polytechnique Federale de Lausanne (EPFL) – Collaborative research with Eduard Trulls. Collaboration Topic: Deep Learning for Image Descriptors.
- University of California at Los Angeles – Collaborative research with Alan Yuille. Collaboration Topic: Deep Learning for Semantic Segmentation.
- University of Massachusetts at Amherst – Collaborative research with Subhransu Maji. Collaboration Topic: Deep Learning for Texture Recognition.
- Ryerson University – Collaborative research with Kostas Derpanis. Collaboration Topic: Deep Learning for Learning Segmentation.
- Ecole Polytechnique de Montreal - Collaborative research with Samuel Kadoury. Collaboration Topic: Multi-atlas co-segmentation, co-registration.
- University of Pennsylvania - Collaborative research with Aristeidis Sotiras. Collaboration Topic: Higher Order Graphs in biomedical image analysis.

## 8.5. International Research Visitors

### 8.5.1. Visits of International Scientists

- Angst, Roland. Max Planck Center for Visual Computing and Communication, GE (April 2015)
- Professor Maragos, Petros: Technical University of Athens, GR (13-20 November 2015)

## 9. Dissemination

### 9.1. Promoting Scientific Activities

#### 9.1.1. Scientific events organisation

##### 9.1.1.1. General chair, scientific chair

- Kokkinos, Iasonas: (i) Co-Organizer of From Image Statistics to Probabilistic Deep Learning in conjunction with ICCV 2015, (ii) Co-Organizer of Search and Planning for Inference and Learning (SPIL), in conjunction with CVPR 2015
- Paragios, Nikos: (i) Organizer of 3rd Biomedical Image Analysis Summer School : Modalities, Methodologies & Clinical Research

### **9.1.2. Scientific events selection**

#### *9.1.2.1. Chair of conference program committees*

- Blaschko, Matthew: British Machine Vision Conference (BMVC)
- Kokkinos, Iasonas: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
- Paragios, Nikos: Medical Image Computing and Computer Assisted Intervention (MICCAI)

#### *9.1.2.2. Member of the conference program committees*

- Blaschko, Matthew: International Conference on Computer Vision (ICCV), Annual Conference on Neural Information Processing Systems (NIPS)
- Kokkinos, Iasonas: International Conference on Computer Vision (ICCV), Neural Information Processing Systems (NIPS), International Conference on Learning Representations (ICLR)

#### *9.1.2.3. Reviewer*

The members of the team reviewed numerous papers for several international conferences, such as for the annual conferences on Computer Vision and Pattern Recognition (CVPR), and Medical Image Computing and Computer Assisted Intervention (MICCAI)

### **9.1.3. Journal**

#### *9.1.3.1. Member of the editorial boards*

- Paragios, Nikos: Medical Image Analysis Journal (MedIA), SIAM Journal on Imaging Sciences
- Kokkinos, Iasonas: Image and Vision Computing Journal (IVC), Computer Vision and Image Understanding Journal (CVIU).
- Kumar, Pawan: Computer Vision and Image Understanding Journal (CVIU).

#### *9.1.3.2. Reviewer - Reviewing activities*

- Blaschko, Matthew: IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI), Journal of Machine Learning Research (JMLR)
- Kokkinos, Iasonas: International Journal of Computer Vision, IEEE Transactions on Pattern Analysis and Machine Intelligence, Computer Vision and Image Understanding (CVIU)
- Ferrante, Enzo: IEEE Transactions on Medical Imaging (T-MI), Medical Image Analysis (MedIA), Computerized Medical Imaging and Graphics (CMIG)
- Zacharaki, Evangelia: IEEE Transactions on Medical Imaging (T-MI), Transactions on Biomedical Engineering
- Dokania, Puneet: Computer Vision and Image Understanding (CVIU)

### **9.1.4. Invited talks**

- Paragios, Nikos: Hong Kong University of State and Technology, Chinese University of Hong Kong, Swiss Federal Institute of Technology in Zurich (ETHZ), Queen Mary University of London, Bayesian and Graphical Models in Biomedical Imaging (in conjunction with MICCAI), National Technical University of Athens

### **9.1.5. Scientific expertise**

- Paragios, Nikos: (i) member of the Scientific Council of the SAFRAN Conglomerate, (ii) Member of the advisory board of the Data HealthCare Institute

## 9.2. Teaching - Supervision - Juries

### 9.2.1. Teaching

#### Masters

Blaschko, Matthew:

- Master: Neural Information Processing, Spring, Ecole CentraleSupélec Paris, FR

Kokkinos, Iasonas

- Master: Machine Learning for Computer Vision, 24, M2, Ecole Normale Supérieure-Cachan, FR
- Master: Introduction to Deep Learning, 24, M2, CentraleSupélec, FR
- Master: Introduction to Signal Processing, 36, M1, CentraleSupélec, FR

### 9.2.2. Supervision

PhD in progress : Eugene Belilovsky, Structured Output Prediction on Large Scale Neuroscience Data, Université Paris-Saclay & KU Leuven, 2014-2017, M. Blaschko

PhD in progress : Jiaqian Yu, Structured Prediction Methods for Computer Vision and Medical Imaging, Université Paris-Saclay, 2014-2017, M. Blaschko

PhD in progress : Wacha Bounliphone, Statistical tools for Imaging-Genetics data integration, 2013-2016, Université Paris-Saclay & KU Leuven, M. Blaschko & A. Tenenhaus

PhD in progress : Amal Rannen, Deep Neural Networks for OCT image analysis, Yonsei University & KU Leuven, M. Blaschko & Y.M. Jung

PhD in progress : Jose Ignacio Orlando, Computer assisted fundus image analysis for eye disease detection, Universidad Nacional del Centro de la Provincia de Buenos Aires, M. Blaschko & M. del Fresno

PhD in progress : Puneet Kumar Dokania, Learning to Rank with Missing and High-Order Information, 2012-2015, M. Pawan Kumar

PhD in progress : Diane Bouchacourt, Large Scale Diverse Learning for Structured Output Prediction, 2014-2017, M. Pawan Kumar

PhD in progress: Haithem Boussaid, Efficient Inference and Learning in Graphical Models for Multi-organ Shape Segmentation, 2011-2015, I. Kokkinos

PhD in progress: Stavros Tsogkas, Learning structured mid-level representations for object recognition, 2011-2015, I. Kokkinos

PhD in progress: Siddhartha Chandra, Efficient Learning and Optimization for 3D Visual Data, 2013-2016, I. Kokkinos, Pawan Kumar

PhD in progress: Stefan Kinauer, Surface-based representations for high-level vision tasks, 2013-2016, I. Kokkinos.

PhD in progress: Alp Guler, Learning 3D representations for high-level vision, 2016-2018, I. Kokkinos.

PhD in progress : Stavros Alchatzidis, Message Passing Methods, Parallel Architectures & Visual Processing, 2011-2014, Nikos Paragios

PhD in progress : Enzo Ferrante, 2D-to-3D Multi-Modal Deformable Image Fusion, 2012-2015, N. Paragios

PhD in progress : Vivien Fecamp, Linear-Deformable Multi-Modal Deformable Image Fusion, 2012-2015, N. Paragios

PhD in progress : Evgenios Kornaropoulos, Diffusion Coefficient: a novel computer aided biomarker, 2013-2016, N. Paragios

PhD in progress : Maxim Berman, Learning Higher Order Graphical Models, 2014-2017, N. Paragios, I. Kokkinos

PhD in progress : Hariprasad Kannan, Efficient Inference on Higher Order Graphs, 2014-2017, N. Paragios

PhD in progress : Huu Dien Khue Le, Graph-based Visual Perception : Theories and Applications, 2014-2017, N. Paragios

### 9.2.3. Juries

- **Matthew Blaschko**
  - **Grant Reviewing Services:** European Research Council (ERC), The French National Research Agency (ANR).
- **Iasonas Kokkinos**
  - **PhD Thesis Participation:** Mateusz Kozinski, Ecole des Ponts ParisTech (PhD).
- **Paragios, Nikos**
  - **PhD Thesis Participation**
  - **Grant Reviewing Services**

## 9.3. Popularization

- **Blaschko, Matthew**
  - **Presentations:** V Argentina Applied Maths Conference (MACI 2015)
- **Kokkinos, Iasonas**
  - **Presentations:** University of California at Los Angeles, Imperial College, Boston University, Xerox Research Center Europe

# 10. Bibliography

## Publications of the year

### Articles in International Peer-Reviewed Journals

- [1] A. BEHL, P. MOHAPATRA, C. JAWAHAR, M. PAWAN KUMAR. *Optimizing Average Precision using Weakly Supervised Data*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2015, vol. 37, n<sup>o</sup> 12, pp. 2545-2557 [DOI : 10.1109/TPAMI.2015.2414435], <https://hal.inria.fr/hal-01223977>
- [2] E. BELILOVSKY, A. ARGYRIOU, G. VAROQUAUX, M. B. BLASCHKO. *Convex relaxations of penalties for sparse correlated variables with bounded total variation*, in "Machine Learning", June 2015, pp. 1-21 [DOI : 10.1007/s10994-015-5511-2], <https://hal.inria.fr/hal-01167861>
- [3] E. BELILOVSKY, K. GKIRTZOU, M. MISYRLIS, A. KONOVA, J. HONORIO, N. ALIA-KLEIN, R. GOLDSTEIN, D. SAMARAS, M. BLASCHKO. *Predictive sparse modeling of fMRI data for improved classification, regression, and visualization using the k-support norm*, in "Computerized Medical Imaging and Graphics", 2015, 1 p. [DOI : 10.1016/j.COMPMEDIMAG.2015.03.007], <https://hal.inria.fr/hal-01141082>
- [4] M. CIMPOI, S. MAJI, I. KOKKINOS, A. VEDALDI. *Deep filter banks for texture recognition, description, and segmentation*, in "International Journal of Computer Vision", January 2016, <https://hal.inria.fr/hal-01263622>



- [5] E. FERRANTE, V. FECAMP, N. PARAGIOS. *Slice-to-Volume Deformable Registration: Efficient one shot consensus between plane selection and in-plane deformation*, in "International Journal for Computer Assisted Radiology and Surgery (IJCARS)", 2015, vol. 10, n<sup>o</sup> 6, pp. 791-800 [DOI : 10.1007/s11548-015-1205-2], <https://hal.inria.fr/hal-01118871>
- [6] K. GKIRTZOU, M. BLASCHKO. *The Pyramid Quantized Weisfeiler-Lehman Graph Representation*, in "Neuro-computing", September 2015 [DOI : 10.1016/J.NEUCOM.2015.09.023], <https://hal.inria.fr/hal-01205717>
- [7] N. KOMODAKIS, B. XIANG, N. PARAGIOS. *A Framework for Efficient Structured Max-Margin Learning of High-Order MRF models*, in "PAMI", 2015, <https://hal.archives-ouvertes.fr/hal-01245673>
- [8] H. LANGET, C. RIDDELL, A. RESHEF, Y. TROUSSET, A. TENENHAUS, E. LAHALLE, G. FLEURY, P. NIKOS. *Compressed-sensing-based content-driven hierarchical reconstruction: Theory and application to C-arm cone-beam tomography*, in "Medical Physics", August 2015, vol. 42, n<sup>o</sup> 9 [DOI : 10.1118/1.4928144], <https://hal.archives-ouvertes.fr/hal-01230277>
- [9] M. PAWAN KUMAR, H. TURKI, D. PRESTON, D. KOLLER. *Parameter Estimation and Energy Minimization for Region-based Semantic Segmentation*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2015, vol. 37, n<sup>o</sup> 7, pp. 1373-1386 [DOI : 10.1109/TPAMI.2014.2372766], <https://hal.inria.fr/hal-01223979>
- [10] H. SIDAHMED, E. PROKOFYEVA, M. BLASCHKO. *Discovering Predictors of Mental Health Service Utilization with k-support Regularized Logistic Regression*, in "Information Sciences", 2015, pp. 1-15 [DOI : 10.1016/J.INS.2015.03.069], <https://hal.inria.fr/hal-01139786>

### Invited Conferences

- [11] S. CHANDRA, G. CHRYSOS, I. KOKKINOS. *Surface Based Object Detection in RGBD Images*, in "British Machine Vision Conference", Swansea, Wales, United Kingdom, September 2016 [DOI : 10.5244/C.29.187], <https://hal.inria.fr/hal-01263930>
- [12] S. CHANDRA, S. TSOGLAS, I. KOKKINOS. *Accurate Human-Limb Segmentation in RGB-D images for Intelligent Mobility Assistance Robots*, in "ACVR, ICCV 2015", Santiago, Chile, December 2015, <https://hal.inria.fr/hal-01263621>

### International Conferences with Proceedings

- [13] S. ALCHATZIDIS, A. SOTIRAS, N. PARAGIOS. *Local atlas selection for discrete multi-atlas segmentation*, in "International Symposium on Biomedical Imaging", New York, United States, April 2015, <https://hal.archives-ouvertes.fr/hal-01145510>
- [14] W. BOUNLIPHONE, E. BELILOVSKY, M. BLASCHKO, I. ANTONOGLIOU, A. GRETTON. *A Test of Relative Similarity For Model Selection in Generative Models*, in "International Conference on Learning Representations", San Juan, Puerto Rico, May 2016, <https://hal.inria.fr/hal-01232290>
- [15] V. FECAMP, A. SOTIRAS, N. PARAGIOS. *Modular linear iconic matching using higher order graphs*, in "ISBI 2015 - IEEE 12th International Symposium on Biomedical Imaging", New York, United States, April 2015, pp. 1097 - 1101 [DOI : 10.1109/ISBI.2015.7164063], <https://hal.inria.fr/hal-01223970>

- [16] E. FERRANTE, V. FECAMP, N. PARAGIOS. *Implicit Planar and In-Plane Deformable Mapping in Medical Images Through High Order Graphs*, in "IEEE International Symposium on BIOMEDICAL IMAGING: From Nano to Macro (ISBI)", Ney York, United States, April 2015, <https://hal.inria.fr/hal-01130724>
- [17] M. SHAKERI, S. TSOVKAS, E. FERRANTE, S. LIPPE, S. KADOURY, N. PARAGIOS, I. KOKKINOS. *Sub-cortical brain structure segmentation using F-CNN's*, in "ISBI 2016: International Symposium on Biomedical Imaging", Prague, Czech Republic, 2016, <https://hal.inria.fr/hal-01265500>
- [18] B. SINGH, S. ALCHATZIDIS, N. PARAGIOS. *Quasi real-time sub-space 3D deformable fusion*, in "Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on", New York, NY , United States, April 2015 [DOI : 10.1109/ISBI.2015.7164000], <https://hal.inria.fr/hal-01270453>

### Conferences without Proceedings

- [19] M. BLASCHKO, J. YU. *Hardness Results for Structured Learning and Inference with Multiple Correct Outputs*, in "Constructive Machine Learning Workshop at ICML", Lille, France, July 2015, <https://hal.inria.fr/hal-01165337>
- [20] D. BOUCHACOURT, S. NOWOZIN, M. P. KUMAR. *Entropy-based Latent Structured Output Prediction*, in "International Conference on Computer Vision (ICCV)", Santiago, Chile, December 2015, <https://hal.inria.fr/hal-01223968>
- [21] W. BOUNLIPHONE, A. GRETTON, A. TENENHAUS, M. BLASCHKO. *A low variance consistent test of relative dependency*, in "International Conference on Machine Learning", Lille, France, July 2015, <https://hal.inria.fr/hal-01005828>
- [22] P. K. DOKANIA, M. P. KUMAR. *Parsimonious Labeling*, in "ICCV 2015 - International Conference on Computer Vision 2015", Santiago, Chile, December 2015, <https://hal.inria.fr/hal-01223973>
- [23] V. FÉCAMP, A. SOTIRAS, N. PARAGIOS. *Simultaneous Linear and Deformable Registration*, in "Medical Imaging Computing and Computer Assisted Interventions", Munich, Germany, October 2015, <https://hal.inria.fr/hal-01223991>
- [24] A. GASTOUNIOTI, A. SOTIRAS, K. S. NIKITA, N. PARAGIOS. *Graph-Based Motion-Driven Segmentation of the Carotid Atherosclerotic Plaque in 2D Ultrasound Sequences*, in "Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015", Munich, Germany, October 2015 [DOI : 10.1007/978-3-319-24574-4\_66], <https://hal.inria.fr/hal-01270448>
- [25] C. LIANG-CHIEH, G. PAPANDREOU, I. KOKKINOS, K. MURPHY, A. YUILLE. *Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs*, in "International Conference on Learning Representations", San Diego, United States, May 2015, <https://hal.inria.fr/hal-01263610>
- [26] G. PAPANDREOU, I. KOKKINOS, P.-A. SAVALLE. *Modeling Local and Global Deformations in Deep Learning: Epitomic Convolution, Multiple Instance Learning, and Sliding Window Detection*, in "IEEE Conference on Computer Vision and Pattern Recognition", Boston, United States, June 2015, <https://hal.inria.fr/hal-01263611>

- [27] E. SIMO-SERRA, E. TRULLS, L. FERRAZ, I. KOKKINOS, P. FUA, F. MORENO-NOGUER. *Deep Convolutional Feature Point Descriptors*, in "International Conference on Computer Vision", Santiago, Chile, December 2015, <https://hal.inria.fr/hal-01263627>
- [28] M. VAKALOPOULOU, K. KARANTZALOS, N. KOMODAKIS, N. PARAGIOS. *Building Detection in Very high Resolution Multispectral Data with Deep Learning Features*, in "IGARSS 2015", Champs-sur-Marne, France, July 2015, <https://hal.archives-ouvertes.fr/hal-01245705>
- [29] J. YU, M. BLASCHKO. *Learning Submodular Losses with the Lovász Hinge*, in "International Conference on Machine Learning", Lille, France, July 2015, <https://hal.inria.fr/hal-01151823>

### Books or Proceedings Editing

- [30] N. PARAGIOS, J. DUNCAN, N. AYACHE (editors). *Handbook of Biomedical Imaging: Methodologies and Clinical Research*, Springer, 2015, 590 p. [DOI : 10.1007/978-0-387-09749-7], <https://hal.inria.fr/inria-00616178>

### Research Reports

- [31] E. FERRANTE, N. PARAGIOS. *Graph Based Slice-to-Volume Deformable Registration*, Inria Saclay ; Center for Visual Computing, CentraleSupélec, Inria, Université Paris-Saclay, October 2015, n° RR-8803, 30 p. , <https://hal.inria.fr/hal-01220005>
- [32] R. GADDE, R. MARLET, N. PARAGIOS. *Learning grammars for architecture-specific facade parsing*, Ecole des Ponts Paris-Tech ; Ecole Centrale Paris, October 2015, n° RR-8600, <https://hal.inria.fr/hal-01069379>
- [33] P. KUMAR, N. KOMODAKIS, N. PARAGIOS. *(Hyper)-Graphs Inference via Convex Relaxations and Move Making Algorithms: Contributions and Applications in artificial vision*, Inria, October 2015, n° RR-8798, 65 p. , <https://hal.inria.fr/hal-01223027>
- [34] Y. ZENG, C. WANG, X. GU, D. SAMARAS, N. PARAGIOS. *Higher-order Graph Principles towards Non-rigid Surface Registration*, Inria Saclay - Ile-de-France ; Inria, September 2015, n° RR-8607, 31 p. , <https://hal.inria.fr/hal-01086052>

### Other Publications

- [35] H. ADAM, I. KOKKINOS, K. DERPANIS. *Learning Dense Convolutional Embeddings for Semantic Segmentation*, 2015, working paper or preprint, <https://hal.inria.fr/hal-01263620>
- [36] E. BELILOVSKY, G. VAROQUAUX, M. B. BLASCHKO. *Hypothesis Testing for Differences in Gaussian Graphical Models: Applications to Brain Connectivity*, December 2015, working paper or preprint, <https://hal.inria.fr/hal-01248844>
- [37] W. BOUNLIPHONE, A. GRETTON, A. TENENHAUS, M. B. BLASCHKO. *Kernel non-parametric tests of relative dependency*, Elsevier, December 2015, vol. ISBN 978-9963-2227-0-4, pp. 165-166, International conference of the ERCIM WG on computational and methodological statistics - CMStatistics 2015, Poster, <https://hal.inria.fr/hal-01263194>
- [38] I. KOKKINOS. *Surpassing Humans in Boundary Detection using Deep Learning*, 2015, working paper or preprint, <https://hal.inria.fr/hal-01263617>

- [39] S. TSOBKAS, I. KOKKINOS, G. PAPANDREOU, A. VEDALDI. *Deep Learning for Semantic Part Segmentation with High-Level Guidance*, 2015, working paper or preprint, <https://hal.inria.fr/hal-01264942>
- [40] J. YU, M. BLASCHKO. *The Lovász Hinge: A Convex Surrogate for Submodular Losses*, December 2015, working paper or preprint, <https://hal.inria.fr/hal-01241626>