RESEARCH CENTRE

**Inria Center
at Université Côte d'Azur**

2022
ACTIVITY REPORT

Team

AYANA

## AI and Remote Sensing on board for the New Space

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions)

**DOMAIN**

**Perception, Cognition and Interaction**

**THEME**

**Vision, perception and multimedia interpretation**

*Inria*

# Contents

# Team AYANA

*Creation of the Team: 2020 January 01*

# Keywords

## Computer sciences and digital sciences

A2.3.1. – Embedded systems

A3.4.1. – Supervised learning

A3.4.2. – Unsupervised learning

A3.4.5. – Bayesian methods

A3.4.6. – Neural networks

A3.4.8. – Deep learning

A5.4.1. – Object recognition

A5.4.5. – Object tracking and motion analysis

A5.4.6. – Object localization

A5.9.2. – Estimation, modeling

A5.9.4. – Signal processing over graphs

A5.9.6. – Optimization tools

A9.3. – Signal analysis

## Other research topics and application domains

B3.3.1. – Earth and subsoil

B3.3.2. – Water: sea & ocean, lake & river

B3.3.3. – Nearshore

B3.4. – Risks

B3.5. – Agronomy

B3.6. – Ecology

B8.3. – Urbanism and urban planning

B8.4. – Security and personal assistance

# 1 Team members, visitors, external collaborators

**Research Scientist**

- Josiane Zerubia [Team leader, INRIA, Senior Researcher, HDR]

**Post-Doctoral Fellows**

- Camilo Aguilar Herrera [INRIA, until Sep 2022]

- Bilel Kanoun [INRIA, Geoazur , until Mar 2022]

**PhD Students**

- Jules Mabon [INRIA]

- Martina Pastorino [UNIV GENES, INRIA]

**Administrative Assistant**

- Nathalie Bellesso [INRIA]

# 2 Overall objectives

The AYANA AEx is an interdisciplinary project using knowledge in stochastic modeling, image processing, artificial intelligence, remote sensing and embedded electronics/computing. The aerospace sector is expanding and changing ("New Space"). It is currently undergoing a great many changes both from the point of view of the sensors at the spectral level (uncooled IRT, far ultraviolet, etc.) and at the material level (the arrival of nano-technologies or the new generation of "Systems on chips" (SoCs) for example), that from the point of view of the carriers of these sensors: high resolution geostationary satellites; Leo-type low-orbiting satellites; or mini-satellites and industrial cube-sats in constellation. AYANA will work on a large number of data, consisting of very large images, having very varied resolutions and spectral components, and forming time series at frequencies of 1 to 60 Hz. For the embedded electronics/computing part, AYANA will work in close collaboration with specialists in the field located in Europe, working at space agencies and/or for industrial contractors.

# 3 Research program

## 3.1 FAULTS R GEMS: Properties of faults, a key to realistic generic earthquake modeling and hazard simulation

Decades of research on earthquakes have yielded meager prospects for earthquake predictability: we cannot predict the time, location and magnitude of a forthcoming earthquake with sufficient accuracy for immediate societal value. Therefore, the best we can do is to mitigate their impact by anticipating the most "destructive properties" of the largest earthquakes to come: longest extent of rupture zones, largest magnitudes, amplitudes of displacements, accelerations of the ground. This topic has motivated many studies in last decades. Yet, despite these efforts, major discrepancies still remain between available model outputs and natural earthquake behaviors. An important source of discrepancy is related to the incomplete integration of actual geometrical and mechanical properties of earthquake causative faults in existing rupture models. We first aim to document the compliance of rocks in natural permanent damage zones. These data –key to earthquake modeling– are presently lacking. A second objective is to introduce the observed macroscopic fault properties –compliant permanent damage, segmentation, maturity– into 3D dynamic earthquake models. A third objective is to conduct a pilot study aiming at examining the gain of prior fault property and rupture scenario knowledge for Earthquake Early Warning (EEW). This research project is partially funded by the ANR Fault R Gems, whose PI is Prof. I. Manighetti from

Geoazur. Two sucessive postdocs (Barham Jafrasteh and Bilel Kanoun) have worked on this research topic funded by UCA-Jedi.

### 3.2 Probabilistic models on graphs and machine learning in remote sensing applied to natural disaster response

We currently develop novel probabilistic graphical models combined with machine learning in order to manage natural disasters such as earthquakes and flooding. The first model will introduce a semantic component to the graph at the current scale of the hierarchical graph, and will necessitate a new graph probabilistic model. The quad-tree proposed by Ihsen Hedhli in AYIN team in 2016 is no longer fit to resolve this issue. Applications from urban reconstruction or reforestation after natural disasters will be achieved on images from Pleiades optical satellites (provided by the French Space Agency, CNES) and CosmoSKyMed radar satellites (provided by the Italian Space Agency, ASI). This project is conducted in partnership with the University of Genoa (Prof. G.Moser and Prof. S. Serpico) via the co-supervision of a PhD student, financed by the Italian government. The PhD student, Martina Pastorino, has worked with Josiane Zerubia and Gabriele Moser in 2020 during her double Master degree at both University Genoa and IMT Atlantique. She started her PhD in co-supervision between University of Genoa DITEN (Prof. Moser) and Inria (Prof. Zerubia) in November 2020.

### 3.3 Marked point process models for object detection and tracking in temporal series of high resolution images

The model proposed by Paula Craciun's PhD thesis in 2015 in AYIN team, supposed the speed of tracked objects of interest in a series of satellite images to be quasi-constant between two frames. However this hypothesis is very limiting, particularly when objects of interest are subject to strong and sudden acceleration or deceleration. The model we proposed within AYIN team is then no longer viable. Two solutions will be considered within AYANA team : either a generalist model of marked point processes (MPP), supposing an unknown and variable velocity, or a multi-model of MPPs, simpler with regard to the velocity that can have a very limited amount of values (ie, quasi-constant velocity for each MPP). The whole model will have to be redesigned, and is to be tested with data from a constellation of small satellites, where the objects of interest can be for instance "speed-boats" or "go-fast cars". Some comparaisons with deep learning based methods belonging to Airbus Defense and Space (Airbus DS) will be done. Then this new model should be brought to diverses platforms (ground based or on-board). The modeling and ground-based application part will be studied within a PhD thesis (Jules Mabon), and the on-board version will be developed as part of a postdoctoral project (Camilo Aguilar-Hererra), both being conducted in partnership with Airbus DS, and the company Erems for the onboard hardware. This project is financed by BPI France within the LiChIE contract. This research project has started in 2020 for a duration of 4 years.

## 4 Application domains

Our research is applied within all Earth observation domains such as: urban planning, precision farming, natural disaster management, geological features detection, geospatial mapping, and security management.

## 5 Highlights of the year

### 5.1 Awards

- Martina Pastorino was finalist for the Prix Laffitte, 6th edition, Ecole Mines Paris-PSL, 2022

# 6 New software and platforms

## 6.1 New software

### 6.1.1 MPP & CNN for object detection in remotely sensed images

**Name:** Marked Point Processes and Convolutional Neural Networks for object detection in remotely sensed images

**Keywords:** Detection, Satellite imagery

**Functional Description:** Implementation of the work presented in: "CNN-based energy learning for MPP object detection in satellite images" Jules Mabon, Mathias Ortner, Josiane Zerubia In Proc. 2022 IEEE International Workshop on Machine Learning for Signal Processing (MLSP) and "Point process and CNN for small objects detection in satellite images" Jules Mabon, Mathias Ortner, Josiane Zerubia In Proc. 2022 SPIE Image and Signal Processing for Remote Sensing XXVIII

**URL:** [https://github.com/Ayana-Inria/MPP_CNN_RS_object_detection](https://github.com/Ayana-Inria/MPP_CNN_RS_object_detection)

**Contact:** Jules Mabon

**Partner:** Airbus Defense and Space

### 6.1.2 FCN and Fully Connected NN for Remote Sensing Image Classification

**Name:** Fully Convolutional Network and Fully Connected Neural Network for Remote Sensing Image Classification

**Keywords:** Satellite imagery, Classification, Image segmentation

**Functional Description:** Code related to the paper:

M. Pastorino, G. Moser, S. B. Serpico, and J. Zerubia, "Fully convolutional and feedforward networks for the semantic segmentation of remotely sensed images," 2022 IEEE International Conference on Image Processing, 2022,

**URL:** [https://github.com/Ayana-Inria/FCN-FFNET_RS-semantic-segmentation](https://github.com/Ayana-Inria/FCN-FFNET_RS-semantic-segmentation)

**Contact:** Josiane Zerubia

**Partner:** University of Genoa, DITEN, Italy

### 6.1.3 Stabilizer for Satellite Videos

**Keywords:** Satellite imagery, Video sequences

**Functional Description:** A python-implemented stabilizer for satellite videos. This code was used to produce the object tracking results shown in: C. Aguilar, M. Ortner and J. Zerubia, "Adaptive Birth for the GLMB Filter for object tracking in satellite videos," 2022 IEEE 32st International Workshop on Machine Learning for Signal Processing (MLSP), 2022, pp. 1-6

**URL:** [https://github.com/Ayana-Inria/satellite-video-stabilization](https://github.com/Ayana-Inria/satellite-video-stabilization)

**Contact:** Josiane Zerubia

**Partner:** Airbus Defense and Space

### 6.1.4 GLMB filter with History-based Birth

**Name:** Python GLMB filter with History-based Birth

**Keywords:** Multi-Object Tracking, Object detection

**Functional Description:** Implementation of the work presented in: C. Aguilar, M. Ortner and J. Zerubia, "Adaptive Birth for the GLMB Filter for object tracking in satellite videos," 2022 IEEE 32st International Workshop on Machine Learning for Signal Processing (MLSP), 2022, pp. 1-6

**URL:** https://github.com/Ayana-Inria/GLMB-adaptive-birth-satellite-videos

**Contact:** Josiane Zerubia

**Partner:** Airbus Defense and Space

### 6.1.5 Automatic fault mapping using CNN

**Keyword:** Detection

**Functional Description:** Implementation of the work published in: Bilel Kanoun, Mohamed Abderrazak Cherif, Isabelle Manighetti, Yuliya Tarabalka, Josiane Zerubia. An enhanced deep learning approach for tectonic fault and fracture extraction in very high resolution optical images. ICASSP 2022 - IEEE International Conference on Acoustics, Speech, & Signal Processing, IEEE, May 2022, Singapore/Hybrid, Singapore.

**URL:** https://github.com/Ayana-Inria/Automatic_fault_mapping_using_CNN

**Contact:** Josiane Zerubia

**Partner:** Géoazur

## 7 New results

### 7.1 Semantic segmentation of remote sensing images combining hierarchical probabilistic graphical models and deep convolutional neural networks

**Participants:** Martina Pastorino, Josiane Zerubia.

**External collaborators:** Gabriele Moser *(Genoa University, DITEN, Professor)*, Sebastiano Serpico *(Genoa University, DITEN, Professor)*.

**Keywords:** image processing, stochastic models, deep learning, fully convolutional networks, multiscale analysis.

The work carried out consisted in the development of novel methods based on fully convolutional networks (FCNs) and a hierarchical probabilistic graphical model (PGM) to tackle the semantic labelling of very high resolution (VHR) remote sensing data. The logic was to take advantage of the spatial modeling capabilities of hierarchical PGMs to mitigate the impact of incomplete ground truths (GTs) and obtain accurate classification results in scenarios where spatially exhaustive GTs do not exist, such as real-world remote sensing applications. In fact, deep learning techniques have shown outstanding performances in semantic segmentation tasks, but their accuracy depends on the quantity and quality of GT used to train them. This approach makes use of a quadtree to model the interactions of the pixels, inter-layer and intra-layer and exploit the intrinsically multiscale nature of FCNs. The marginal posterior mode (MPM) criterion for inference was used in the proposed framework. The method was experimentally
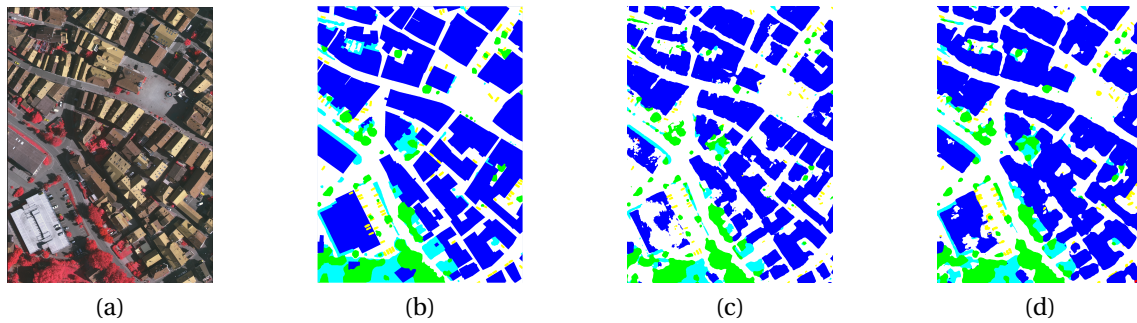
Figure 1: Ground truth and classification maps for the Vaihingen dataset: (a) IRRG image, (b) ground truth, and classification maps obtained with: (c) standard FCN (U-Net) and (d) the proposed method. Classes: buildings (blue), impervious (white), vegetation (cyan), trees (green), cars (yellow).

validated on the ISPRS 2D Semantic Labeling Challenge dataset, which includes aerial images of the cities of Vaihingen and Potsdam in Germany. As these datasets are "ideal": the class label is known for all pixels in the images, the FCN has been further trained with a dataset with several modifications, aimed at approximating the most common GTs found in real remote sensing applications (which are far from being spatially exhaustive). The results are significant, as the proposed methodology has a higher producer accuracy (the map accuracy from the point of view of the map maker) compared to the standard FCNs considered, in particular when the input training data are scarce and approach the real-world GTs available for land-cover mapping applications. The results were published on IEEE TGRS [3].

The method was further extended with the addition of fully connected neural networks (dense layers) at different convolutional blocks of the FCN, allowing to incorporate information at different spatial resolutions and define an end-to-end architecture. Spatial-contextual information is favored in this framework through a supplementary convolutional layer modeling the interactions between neighboring pixels at the same resolution. The loss function of the proposed model consists of a linear combination of the weighted cross-entropy losses of the fully connected neural networks and of the FCN. The modeling of spatial information is further addressed by the introduction of an additional loss term which allows to integrate spatial information between neighboring pixels. The goal of the proposed method is threefold: firstly, to take advantage of the intrinsic multiscale behavior of FCNs to integrate multiscale information through the addition of fully connected neural networks; secondly, to strengthen the modeling of the spatial information through an additional convolutional layer; finally, to define a loss function capable to consider both this multiscale and spatial information within an end-to-end neural model. The experimental validation was conducted again with the ISPRS 2D Semantic Labeling Challenge data set over the city of Vaihingen, Germany (see Fig.1). This approach obtains higher average classification results than the state-of-the-art techniques considered, especially in the case of scarce, suboptimal GTs. This work was submitted to ICIP'22 [10].

Current work involves the definition of an end-to-end neural model for semantic segmentation tasks as a fully convolutional conditional random field (CRF) network, and for the future the idea is to integrate the proposed method with transfer learning and test it with various datasets related to real world applications, such as disaster management.

## 7.2   Hierarchical probabilistic graphical models and fully convolutional networks for the semantic segmentation of multimodal SAR images

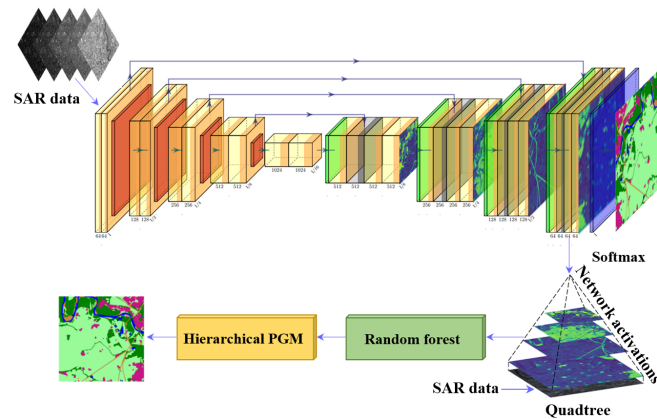**Participants:**   Martina Pastorino, Josiane Zerubia.

Figure 2: Overall architecture of the proposed method applied on SAR data.

**External collaborators:** Gabriele Moser *(Genoa University, DITEN, Professor)*, Sebastiano Serpico *(Genoa University, DITEN, Professor)*.

**Keywords:** image processing, stochastic models, deep learning, fully convolutional networks, multimodal radar images.

The method developed and tested on the optical dataset in [3] was extended to deal with synthetic aperture radar (SAR) images (see Fig. 2, 3), in the framework of a research project funded by the Italian Space Agency (ASI). This SAR images are multimission (SAOCOM and Cosmo-SkyMed) and were collected with different modalities (TOPSAR and StripMap), bands (L and X), polarizations, and several trade-offs between resolution and coverage. The possibility to deal with multimodal images offers great prospects for land cover mapping applications in the field of remote sensing. However, the development of a supervised method for the classification of SAR images – which suffer from speckle and are determined by complex scattering phenomena – presents some issues.

The results obtained by the proposed technique were also compared with those of HRNet [23], a network consisting of multiresolution subnetworks connected in parallel, and of the light-weight attention network (LWN-Attention) [19], a multiscale feature fusion method making use of multiscale information through the concatenation of feature maps associated with different scales. The proposed method, leveraging both hierarchical and long-range information attained generally higher average classification results, thus suggesting that the proposed integration of FCN and hierarchical PGM can be advantageous. The work validated on Cosmo-SkyMed data was accepted for a presentation at IGARSS'22 [9].
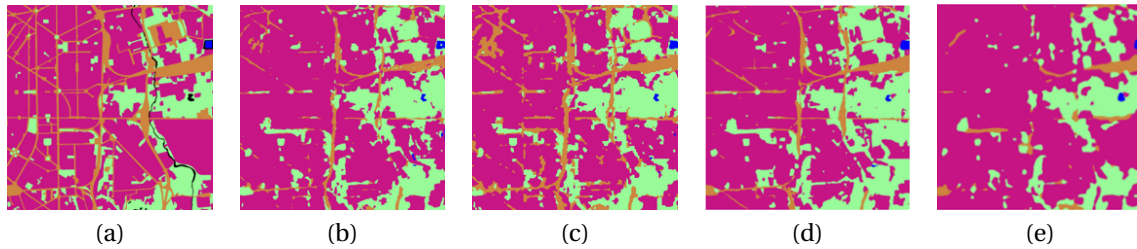
Figure 3: Ground truth and classification maps for the multimission (Cosmo-SkyMed and SAOCOM) SAR dataset, Milan, 2021: (a) ground truth, and classification maps obtained with: (b) standard FCN (U-Net), (c) the proposed method, (d) HRNet, and (e) LWN-Attention. Classes: urban (pink), vegetation (pale green), clutter (orange), water (blue). (b) to (e): Product processed under a license of the Italian Space Agency (ASI) and of the Argentinian Space Agency (CONAE); Original COSMO-SkyMed and SAOCOM Product - ©ASI, ©CONAE - (2021).

## 7.3    Extraction of curvilinear structure networks in image data using an innovative deep learning approach

**Participants:**    Bilel Kanoun, Josiane Zerubia.

**External collaborators:**    Isabelle Manighetti *(OCA, Géoazur, Senior Physicist)*.

**Keywords:** curvilinear feature extraction, machine learning, VHR optical imagery, remote sensing.

Fractures and faults are ubiquitous on Earth and contribute to a number of important processes: small scale fractures control the permeability of rock reservoirs, and larger-scale faults may produce damaging earthquakes. Therefore, identifying and mapping fractures and faults are important in geosciences, especially in earthquake hazard and geological reservoir studies. However, the delineation of fractures and faults is a challenging task due to complexity of the faults. So far, mapping fractures and faults has been done manually in optical images of the Earth surface, yet it is time consuming and it requires an expertise that may not be available. To this purpose, the idea is to explore deep learning approaches for the automatic extraction of fractures and faults. Building upon a recent prior study namely TM-MRef model [20], we developed a deep learning approach, called BK-MRef (presented at ICASSP 2022 [5]), based on a variant of a U-Net neural network, and applied it to automate fracture and fault mapping in optical images and topographic data. We performed the training of the model with a realistic knowledge of fracture and fault uneven distributions and trends (a "balanced training"), using a loss function, called the focal loss [18] , that "down-weights" the most represented features (not-a-fault) to focus on the less represented but most meaningful features (the fracture/fault pixels). In addition, we proceeded the study with another loss function that operates at both pixel and larger scales through a combined use of weighted binary Cross-entropy (wBCE) [14] and Jaccard loss [15], noted as combined loss $L''$. Such a loss function greatly improves the predictions, both quantitatively (see ROC curves in Figure 4) and qualitatively (see Figure 5). As we applied the model to a site different from those used for training, we have demonstrated it enhanced generalization capacity.
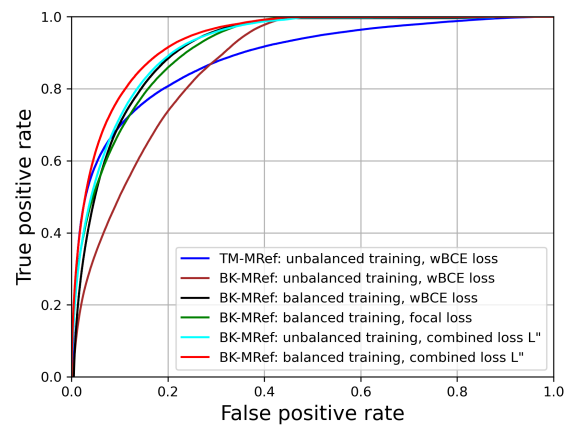
Figure 4: ROC curves for the six models with respect to refined mapping.

(a) Optical image



(b) Ground truth (refined manual mapping)



(c) TM-MRef: unbalanced training, wBCE loss



(d) BK-MRef: unbalanced training, wBCE loss



(e) BK-MRef: balanced training, wBCE loss



(f) BK-MRef: balanced training, focal loss



(g) BK-MRef: unbalanced training, combined loss $L''$



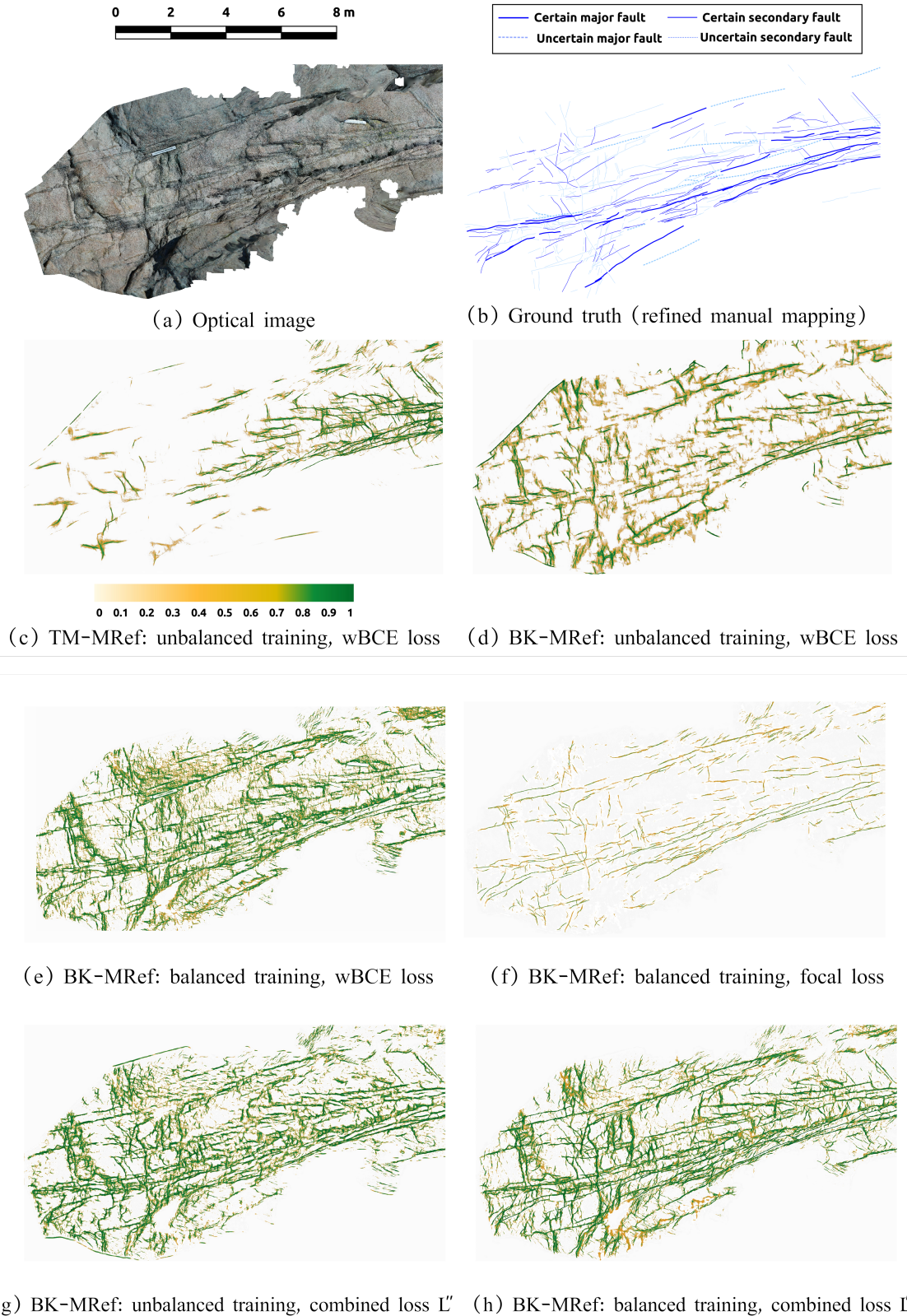(h) BK-MRef: balanced training, combined loss $L''$

Figure 5: Predictions in the test zone with the different approaches. (a) Optical ortho-image of the test zone, (b) Manual expert mapping of fracture and fault traces (refined mapping), (c) Predictions using TM-MRef model [20] and (d, e, f, g, h) variants of BK-MRef model.

### 7.4   Combining Stochastic geometry and deep learning for small objects detection in remote sensing datasets

**Participants:**   Jules Mabon, Josiane Zerubia.

**External collaborators:**   Mathias Ortner *(Airbus Defense and Space, Senior Data Scientist).*

**Keywords:** object detection, deep learning, fully convolutional networks, stochastic models, energy based models.

Unmanned aerial vehicles and low-orbit satellites, including cubesats, are increasingly used for wide area surveillance, which results in large amounts of multi-source data that have to be processed and analyzed. These sensor platforms capture vast ground areas at frequencies varying from 1 to 60 Hz depending on the sensor. The number of moving objects in such data is typically very high, accounting for up to thousands of objects. On one hand, stochastic geometry has proven to be extremely powerful for capturing object positions within images using a prior model. For instance, it is possible to introduce a regularizing term that accounts for relative positions of objects (see previous research results in both AYIN and ARIANA teams), in order to represent specific patterns. One major drawback of this approach is that the amount of marks that can be handled has to be limited in order to avoid an explosion of the problem dimensionality. Moreover, parameter optimization is more complex when objects have complex shapes. On the other hand, deep learning approaches have largely proven over the recent years to be extremely efficient in building representations that match object signatures. We first tackled the problem of static detection. We propose to merge both approaches and build a stochastic point process model using a data term learned through a convolutional neural network (CNN); more specifically a Unet architecture, allowing for dense predictions on any image size. This approach allows us to leverage the generalization capabilities of neural networks while regularizing the results through prior energies in the point process model. The availability of detection map sourced from the CNN -used for the data term of the point process- allows us to leverage a birth map within the sampling procedure thus reduce the inference time. Our overall goal is to intertwine the sampling and learning procedure as much as feasible, in order to avoid the ad-hoc tweaking of parameters that is often necessary with point process modeling. We perform tests of our CNN and point process joint model on the DOTA dataset and on aerial data provided by Airbus DS (as part of our collaboration within the LiChIE contract funded by BPI France), subsampled to a spatial resolution of 50 cm per pixel. In our approach, we build a model to extract oriented rectangles (position, length, with and angle) that match the vehicles in the images. At first, we set the relative weights of the energies of our model (weighting the relative importance of the data term and the various prior terms), these results are published in [8, 11]. We elaborate on that model by introducing a procedure to learn the weights; this method is presented in [7].

In figure 6 we compare our results on DOTA with the CNN method from [27], along with MPP constrast based methods from [17] and [16]. The model MPP+CNN* corresponds to manual energy weights, while MPP+CNN[†] uses learned energy weights. In Figure 7 we compare results of our method against [27] on aerial data provided by Airbus DS.
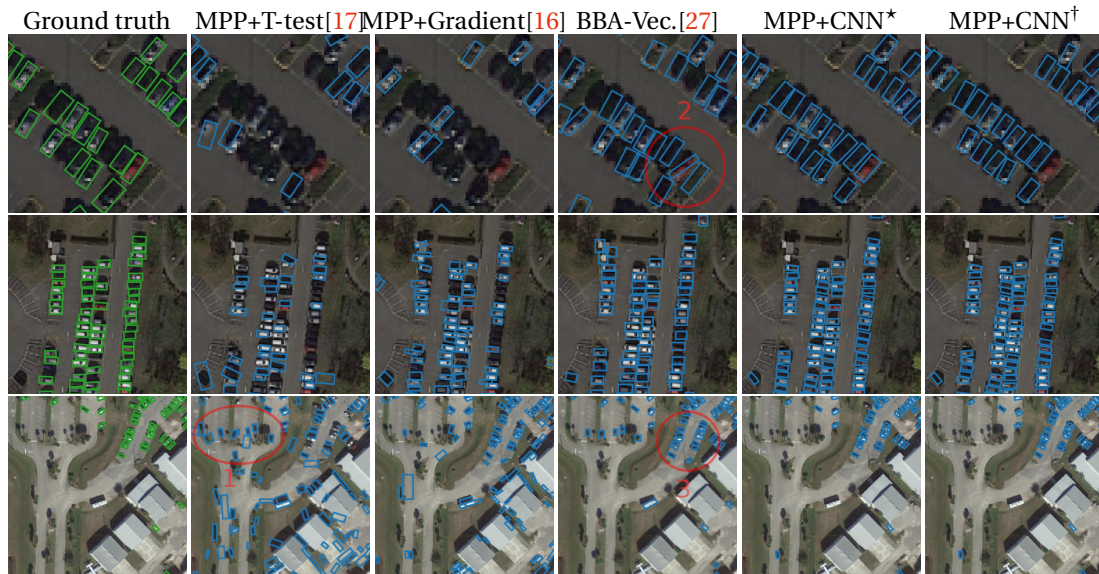
Figure 6: Inference results for various models. For example 1: false positives, 2-3: misaligned objects
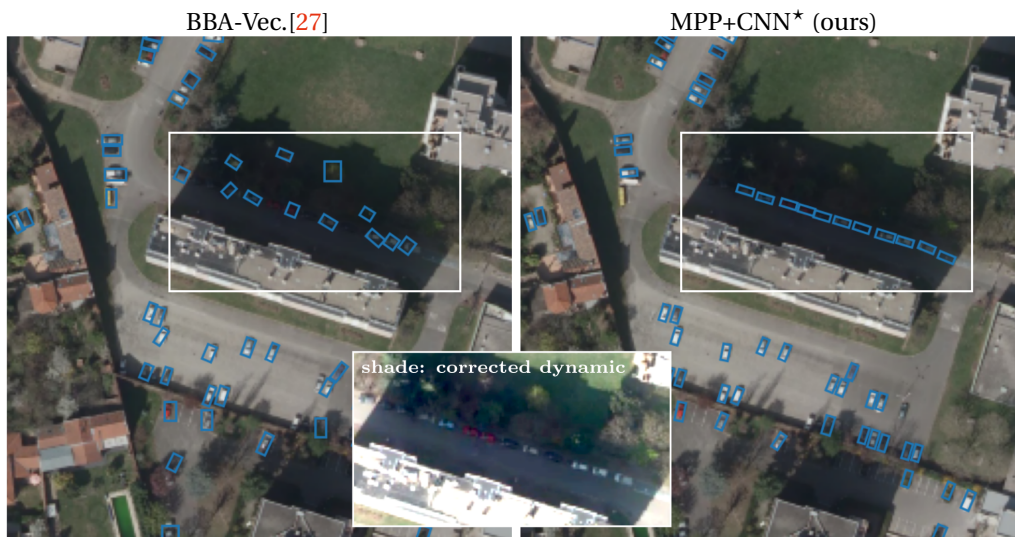


Figure 7: Detection results on samples of the data provided by Airbus, aerial photos subsampled to 50 cm/px, no ground truth available. [copyright Airbus Defense and Space]

## 7.5  Small Object Detection and Tracking in Satellite Videos with Neural Networks and Random Finite Sets

**Participants:**    Camilo Aguilar, Josiane Zerubia.

**External collaborators:**    Mathias Ortner *(Airbus Defense and Space, Senior Data Scientist)*.

**Keywords:** object detection, object tracking, convolutional neural networks, Random Finite Sets.

This work presents a novel object detection and tracking approach for satellite videos. Specifically, the project focuses on tracking moving ground vehicles in low-resolution images (1 m/pixel). The main challenge arises from detecting and tracking dense populations of small objects in cluttered scenes. Vehicles span few pixels of width and height, lack distinctive features, and are often surrounded by texture rich backgrounds. We propose a track-by-detection approach to detect and track small moving targets by using a convolutional neural network and a Bayesian tracker. Our object detection consists of a two-step process based on motion and a patch-based convolutional neural network (CNN). The first stage performs a lightweight motion detection operator to obtain rough target locations. The second stage uses this information combined with a CNN to refine the detection results. In addition, we adopt an online track-by-detection approach by using the Probability Hypothesis Density (PHD) [25] filter to convert detections into tracks. The PHD filter offers a robust multi-object Bayesian data-association framework that performs well in cluttered environments, keeps track of missed detections, and presents remarkable computational advantages over different Bayesian filters. We test our method across various cases of a challenging dataset: a low-resolution satellite video comprising numerous small moving objects. We demonstrate the proposed method outperforms competing approaches across different scenarios with both object detection and object tracking metrics (Figure 8). These have been published in MLSP 2022 [4] and Frontiers in Signal Processing[2]. Currently, our work involves developing better and computationally feasible approximations for the Random Finite Set (RFS) posterior using state-of-the-art filters [26] and novel CNN architectures [24].
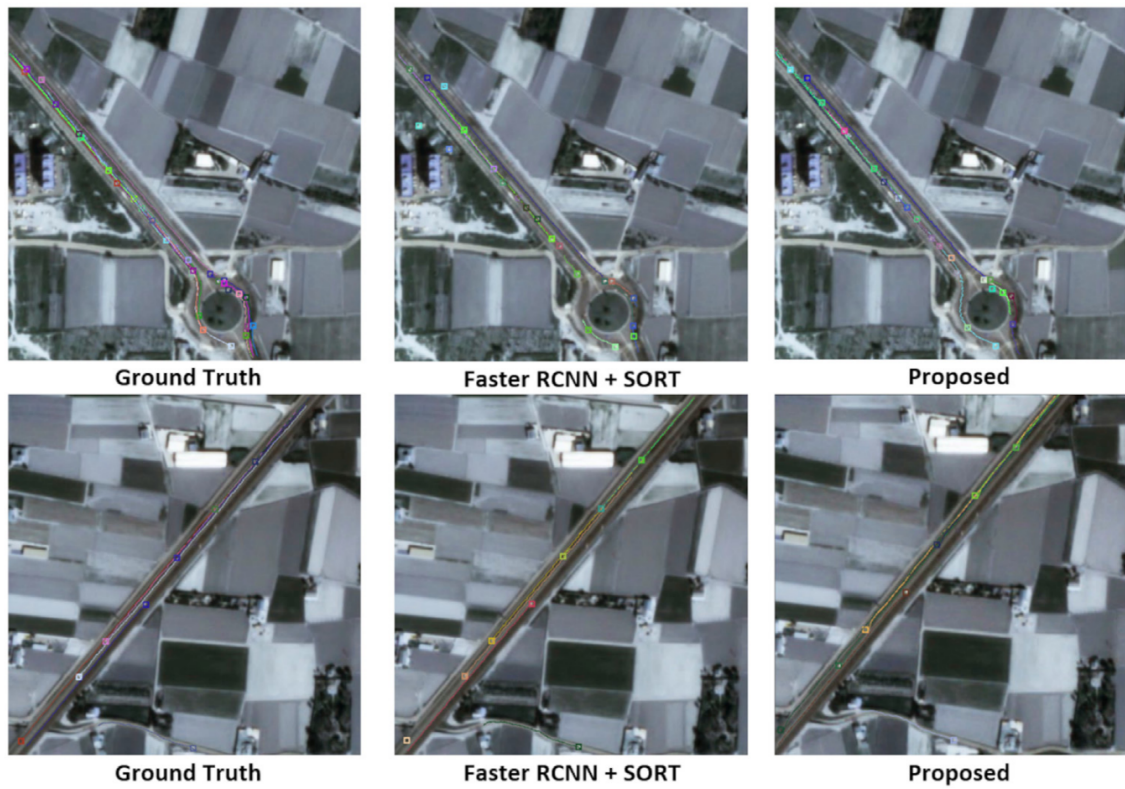
Figure 8: Sample Object Tracking. The square denotes the object current location and the line the object past locations. First column: ground truth marks. Second Column: Faster RCNN [22] and SORT [12]. Third column: proposed tracking algorithm

### 7.6 A Two-Stage Road Segmentation Approach for Remote Sensing Images

**Participants:** Josiane Zerubia.

**External collaborators:** Tianyu Li *(Purdue University, US, PhD student)*, Mary Comer *(Purdue University, US, Associate professor)*.

**Keywords:** remote sensing, road segmentation, convolutional neural network (CNN), two stage learning.

This work consisted of the development of a two-stage learning strategy for road segmentation in remote sensing images. With the development of new technology in satellites, cameras and communications, massive amounts of high-resolution remotely-sensed images of geographical surfaces have become accessible to researchers. To help in analyzing these massive image datasets, automatic road segmentation has become more and more important. Many road segmentation methods based on CNNs have been proposed for remote sensing images in recent years. Although these techniques show great performance in various applications, there are still problems in road segmentation, due to the existence of complex backgrounds, illumination changes, and occlusions due to trees and cars. To alleviate such problems, we propose a two-stage strategy for road segmentation. The diagram of the two-stage segmentation approach is given in Figure 9. A selected network (such as ResUnet) is trained in stage one with the original RGB training samples. Then the network trained in stage one is used to generate the training samples for stage two. To be specific, when an RGB training sample is fed to the trained network in stage one, a probability map and a weight map are generated. The probability map is attached to the RGB training sample as a 4 dimension input to the U-Net-like network in stage two. The weight map is used for calculating the loss function in stage two. Our experiments on the Massachusetts road dataset [21] and the DeepGlobe dataset [13] show the average IoU can increase up to 3% from stage one to stage two, which achieves state-of-the-art results on these datasets. Moreover, from the qualitative results, such as the examples in Figure 10, obvious improvements from stage one to stage two can be seen: fewer false positives and better connection of road lines.

This work was presented at ICPRw 2022 - 26th International Conference on Pattern Recognition workshops (PRRS 2022) Montréal, Canada, August 21, 2022 [6].
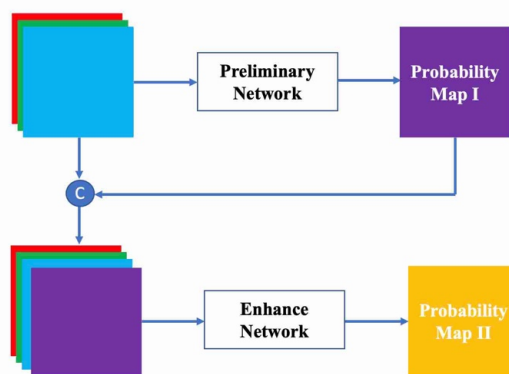


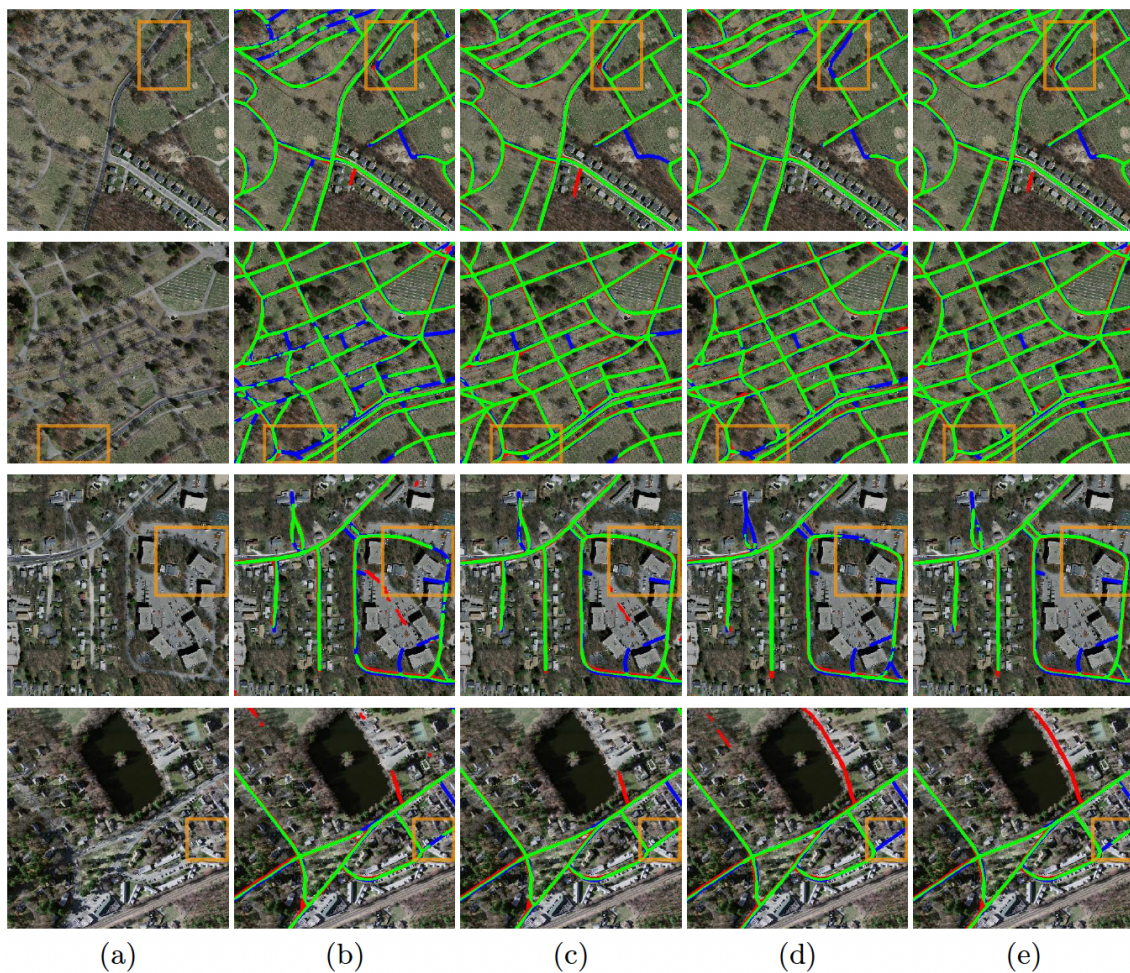Figure 9: The diagram of the two-stage segmentation approach.

Figure 10: Segmentation results from the Massachusetts dataset: green represents true positive (TP), red represents false positive (FP), blue represents false negative (FN). (a): Original image; (b): ResUnet; (c): ResUnet+CUnet; (d) D-LinkNet; (e) DLinkNet+CUnet.

# 8    Bilateral contracts and grants with industry

## 8.1    Bilateral contracts with industry

### 8.1.1    LiChIE contract with Airbus Defense and Space funded by BPI France

**Participants:**    Camilo Aguilar, Jules Mabon, Josiane Zerubia.

**External collaborators:**    Mathias Ortner *(Airbus Defense and Space, Senior Data Scientist).*

Automatic object detection and tracking on sequences of images taken from various constellations of satellites.

# 9    Partnerships and cooperations

## 9.1    National initiatives

**Participants:**    Bilel Kanoun, Josiane Zerubia.

**External collaborators:**    Isabelle Manighetti *(OCA, Géoazur, Senior Physicist).*

AYANA team is part of the ANR project FAULTS R GEMS (2017-2023, PI Geoazur) dedicated to realistic generic earthquake modeling and hazard simulation.

# 10    Dissemination

## 10.1    Promoting scientific activities

### 10.1.1    Scientific events: organisation

**Member of the organizing committees**

- Josiane Zerubia was part of the organizing committee and member of the National Scientific Committee of the ISPRS congress 2020 (digital), 2021 (virtual) and 2022 (in person), in Nice, France

### 10.1.2    Scientific events: selection

**Member of the conference program committees**

- Josiane Zerubia was a member of the conference program committee of SPIE Remote Sensing'22 (Berlin, Germany)

**Reviewer**

- Martina Pastorino did reviewing for the conference IEEE IGARSS'22

- Josiane Zerubia did reviewing for the conferences GRETSI'22, IEEE ICASSP'22, IEEE EMBC'22, IEEE ICIP'22, IEEE IGARSS'22, IAPR ICPR'22, IEEE/EURASIP EUSIPCO'22, SPIE Remote Sensing'22 and ISPRS Congress 2022

### 10.1.3 Journal

**Member of the editorial boards**

- Josiane Zerubia was a member of the editorial boards of IEEE SPMag (2018-2022), and Fondation and Trend in Signal Processing (2007- )

**Reviewer - reviewing activities**

- Martina Pastorino did reviews for the journals MDPI Remote Sensing, IEEE JSTARS, and IEEE TGRS

- Josiane Zerubia did reviews for IEEE SPMag and as lead guest editor for a special issue of IEEE SPMag

- Camilo Aguilar did reviews for IEEE Transactions on Geoscience and Remote Sensing

### 10.1.4 Invited talks

**Conferences and workshops**

- Martina Pastorino presented at CIRM (Centre International de Rencontres Mathématiques) Workshop on Apprentissage Automatique et Traitement du Signal sur Graphes / Machine Learning and Signal Processing on Graphs, Marseille, France, 7-11 Nov. 2022

**Talks**

- Martina Pastorino presented at the PhD Seminar, Inria d'Université Côte d'Azur, Mar. 2022

- Josiane Zerubia presented Ayana research work at CNES, Airbus DS and IRT St Exupéry, Toulouse, May 17-19 2022

- Jules Mabon, Camilo Aguilar, and Martina Pastorino presented to CNES Data Campus during their visit at Inria center in Sophia Antipolis, Sep. 9, 2022

- Jules Mabon presented ar Journées communes GéoSto MIA, Rouen, France, 22-23 Sep. 2022

- Martina Pastorino presented a seminar at Obelix team invited by Professor Sébastien Lefèvre, Université Bretagne Sud, Vannes, Nov. 2022

### 10.1.5 Leadership within the scientific community

- Josiane Zerubia is IEEE (2002- ), EURASIP (2019- ) and IAPR Fellow (2020- )

- Josiane Zerubia is a member of the Teaching Board of the Doctoral School STIET at University of Genoa, Italy (2018- )

- Josiane Zerubia is a member of the Scientific Council of GdR Geosto (Stochastic Geometry) from CNRS (2021- )

- Josiane Zerubia is a member of the Scientific Council of the IMPT-TDMA summer school (2022-2023)

### 10.1.6 Scientific expertise

- Josiane Zerubia was member of the IEEE SPS Awards Board (2020-2022)

- Member of IAPR Fellow Committee (2021-2022)– Josiane Zerubia

- Josiane Zerubia is a member of the STEREO IV evaluation panel (Belgian Research Programme for Earth Observation) for the Belgian Science Policy (BELSPO) in Brussels on Sep. 20-22, 2022

- Josiane Zerubia did consulting for Pikaïros, Aug. 2022

### 10.1.7    Research administration

- Josiane Zerubia is a member of various committees at Inria-SAM (CEP, CC, CB)

## 10.2    Teaching - Supervision - Juries

### 10.2.1    Teaching

- Josiane Zerubia taught masters course on Remote Sensing: 13.5h eq. TD (9h of lectures), Master RISKS, Université Côte d'Azur, France (2019-2021). This course was given to Masters students

### 10.2.2    Supervision

- Co-supervision of a Master student in 2022 for the Master in Internet and Multimedia Engineering, University of Genoa – Martina Pastorino

- Supervising two postdocs and two PhD students within the AYANA team – Josiane Zerubia

### 10.2.3    Juries

- Member of the recruiting committee (reviewer) for an Inria Chair Junior Professor position at University of Montpellier, Mar. 2022 – Josiane Zerubia

## 10.3    Popularization

- Participation to the day "Girls, Mathematics and Computer Science" at Lycée du Val d'Argens au Muy dans le Var (83), Apr. 2022 – Martina Pastorino, Josiane Zerubia

### 10.3.1    Internal or external Inria responsibilities

- Secretary of the ADSTIC (Doctoral Association of the SophiaTech campus), which organizes scientific, social and sportive events for the PhD students, postdocs and non-permanent researchers working both at Inria and in the SophiaTech campus – Jules Mabon

- Regular contacts with Federique Segond, responsible for the Defense and Security mission at Inria, regarding the aerospace research field, since Apr. 2022 – Josiane Zerubia

- Organization of CNES Data Campus visit at Inria center in Sophia Antipolis and University of Montpellier (remotely), Sep. 9, 2022. Scientific presentations from Stars, Titane, Ayana and future Evergreen teams and technical presentation by Inria STIP – Josiane Zerubia

# 11    Scientific production

## 11.1    Major publications

[1]    I. Hedhli, G. Moser, S. B. Serpico and J. Zerubia. 'Hierarchical Markov random fields for high resolution land cover classification of multisensor and multiresolution image time series'. In: *Change Detection and Image Time Series Analysis*. Vol. 2. Supervised Methods. ISTE-Wiley, Dec. 2021, pp. 1–27. URL: https://hal.inria.fr/hal-03466612.

## 11.2    Publications of the year

**International journals**

[2]    C. Aguilar, M. Ortner and J. Zerubia. 'Small Object Detection and Tracking in Satellite Videos With Motion Informed-CNN and GM-PHD Filter'. In: *Frontiers in Signal Processing* 2 (29th Apr. 2022). DOI: 10.3389/frsip.2022.827160. URL: https://hal.inria.fr/hal-03655022.

[3]     M. Pastorino, G. Moser, S. B. Serpico and J. Zerubia. 'Semantic Segmentation of Remote Sensing
        Images through Fully Convolutional Neural Networks and Hierarchical Probabilistic Graphical
        Models'. In: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), pp. 1–16. DOI: 10.110
        9/TGRS.2022.3141996. URL: https://hal.inria.fr/hal-03534026.

**International peer-reviewed conferences**

[4]     C. Aguilar, M. Ortner and J. Zerubia. 'Adaptive birth for the GLMB filter for object tracking in
        satellite videos'. In: MLSP 2022 - IEEE International workshop on machine learning for signal
        processing. Xi'an, China, 22nd Aug. 2022. DOI: 10.1109/MLSP55214.2022.9943411. URL: https:
        //hal.inria.fr/hal-03715704.

[5]     B. Kanoun, M. A. Cherif, I. Manighetti, Y. Tarabalka and J. Zerubia. 'An enhanced deep learning
        approach for tectonic fault and fracture extraction in very high resolution optical images'. In:
        ICASSP 2022 - IEEE International Conference on Acoustics, Speech, & Signal Processing. Singa-
        pore/Hybrid, Singapore, 22nd May 2022. DOI: 10.1109/ICASSP43922.2022.9747007. URL:
        https://hal.inria.fr/hal-03577214.

[6]     T. Li, M. Comer and J. Zerubia. 'A Two-Stage Road Segmentation Approach for Remote Sensing
        Images'. In: ICPRw 2022 - 26th International Conference on Pattern Recognition workshops (PRRS
        2022). Montréal, Canada, 21st Aug. 2022. URL: https://hal.inria.fr/hal-03810488.

[7]     J. Mabon, M. Ortner and J. Zerubia. 'CNN-based energy learning for MPP object detection in
        satellite images'. In: MLSP 2022 - IEEE International workshop on machine learning for signal
        processing. Xi'an, China, 22nd Aug. 2022. DOI: 10.1109/MLSP55214.2022.9943312. URL: https:
        //hal.inria.fr/hal-03715331.

[8]     J. Mabon, J. Zerubia and M. Ortner. 'Point process and CNN for small object detection in satellite
        images'. In: SPIE, Image and Signal Processing for Remote Sensing XXVIII. Berlin, Germany, 5th Sept.
        2022. DOI: 10.1117/12.2635848. URL: https://inria.hal.science/hal-03738027.

[9]     M. Pastorino, G. Moser, S. Serpico and J. Zerubia. 'Semantic Segmentation of Sar Images Through
        Fully Convolutional Networks and Hierarchical Probabilistic Graphical Models'. In: IEEE IGARSS
        2022- International Geoscience and Remote Sensing Symposium. Kuala Lumpur / Virtual, Malaysia,
        17th July 2022. DOI: 10.1109/IGARSS46834.2022.9883111. URL: https://hal.inria.fr/hal
        -03655029.

[10]    M. Pastorino, G. Moser, S. B. Serpico and J. Zerubia. 'Fully convolutional and feedforward networks
        for the semantic segmentation of remotely sensed images'. In: ICIP 2022- IEEE International
        Conference on Image Processing. Bordeaux, France, 16th Oct. 2022. DOI: 10.1109/ICIP46576.20
        22.9897336. URL: https://hal.inria.fr/hal-03720693.

**National peer-reviewed Conferences**

[11]    J. Mabon, M. Ortner and J. Zerubia. 'Processus ponctuels marqués et réseaux de neurones convo-
        lutifs pour la détection d'objets dans des images de télédétection'. In: GRETSI 2022 - XXVIIIème
        Colloque Francophone de Traitement du Signal et des Images. Nancy, France, 6th Sept. 2022. URL:
        https://hal.inria.fr/hal-03715337.

## 11.3   Cited publications

[12]    A. Bewley, Z. Ge, L. Ott, F. Ramos and B. Upcroft. 'Simple online and realtime tracking'. In: *2016
        IEEE international conference on image processing (ICIP)*. IEEE. 2016, pp. 3464–3468.

[13]    I. Demir, K. Koperski, D. Lindenbaum, G. Pang, J. Huang, S. Basu, F. Hughes, D. Tuia and R. Raskar.
        'DeepGlobe 2018: A challenge to parse the earth through satellite images'. In: *IEEE/CVF Conference
        on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2018).

[14]    Y. Ho and S. Wookey. 'The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of
        Mislabeling'. In: *IEEE Access* 8 (2020), pp. 4806–4813. DOI: 10.1109/ACCESS.2019.2962617.

[15]   P. Jaccard. 'The distribution of the flora in the Alpine zone.1'. In: *New Phytologist* 11.2 (1912), pp. 37–50. DOI: https://doi.org/10.1111/j.1469-8137.1912.tb05611.x.

[16]   M. S. Kulikova, I. H. Jermyn, X. Descombes, E. Zhizhina and J. Zerubia. 'Extraction of Arbitrarily-Shaped Objects Using Stochastic Multiple Birth-and-Death Dynamics and Active Contours'. In: *Proc. Computational Imaging VIII*. SPIE. 2010.

[17]   C. Lacoste, X. Descombes and J. Zerubia. 'Point processes for unsupervised line network extraction in remote sensing'. In: *IEEE TPAMI* 27.10 (2005), pp. 1568–1579.

[18]   T.-Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollar. 'Focal Loss for Dense Object Detection'. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2017, pp. 2980–2988.

[19]   S. Liu, C. He, H. Bai, Y. Zhang and J. Cheng. 'Light-Weight Attention Semantic Segmentation Network for High-Resolution Remote Sensing Images'. In: *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. Waikoloa, HI, USA, Sept. 2020, pp. 2595–2598.

[20]   L. Mattéo, I. Manighetti, Y. Tarabalka, J.-M. Gaucel, M. van den Ende, A. Mercier, O. Tasar, N. Girard, F. Leclerc, T. Giampetro, S. Dominguez and J. Malavieille. 'Automatic Fault Mapping in Remote Optical Images and Topographic Data With Deep Learning'. In: *Journal of Geophysical Research: Solid Earth* 126.4 (2021). e2020JB021269 2020JB021269, e2020JB021269. DOI: https://doi.org/10.1029/2020JB021269.

[21]   V. Mnih. 'Machine Learning for Aerial Image Labeling'. PhD thesis. University of Toronto, 2013.

[22]   S. Ren, K. He, R. B. Girshick and J. Sun. 'Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.' In: *NIPS*. Ed. by C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama and R. Garnett. MIT Press, 2015, pp. 91–99.

[23]   K. Sun, B. Xiao, D. Liu and J. Wang. 'Deep High-Resolution Representation Learning for Human Pose Estimation'. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5686–5696, 2019.

[24]   A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin. 'Attention is All you Need'. In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett. Vol. 30. Curran Associates, Inc., 2017, pp. 5998–6008.

[25]   B.-N. Vo and W.-K. Ma. 'The Gaussian Mixture Probability Hypothesis Density Filter'. In: *IEEE Transactions on Signal Processing* 54 (Dec. 2006), pp. 4091–4104. DOI: 10.1109/TSP.2006.881190.

[26]   B.-N. Vo, B.-T. Vo and H. G. Hoang. 'An Efficient Implementation of the Generalized Labeled Multi-Bernoulli Filter'. In: *IEEE Transactions on Signal Processing* 65.8 (2017), pp. 1975–1987. DOI: 10.1109/TSP.2016.2641392.

[27]   J. Yi, P. Wu, B. Liu, Q. Huang, H. Qu and D. Metaxas. 'Oriented Object Detection in Aerial Images with Box Boundary-Aware Vectors'. In: *Proc. IEEE Winter Conf. on Applications of Computer Vision*. 2021.