Weakly supervised detection of marine mammals

Pierre Allain & Minh-Tan Pham

WIPSEA, F-35000 Rennes, France
Univ. Bretagne Sud - IRISA, F-56000 Vannes, France

MACLEAN – CAP&RFIAP

05/07/2022

Introduction

Context

- Data annotation for fully-supervised approach → time-consuming and expensive
- A majority (>95%) of acquired images are empty (a single session of an aerial survey could provide thousands of images)

  → visual analysis: laborious and time-consuming

> 95% < 5%

Introduction

Objective: weakly-supervised anomaly detection

- Reduce the gap between fully supervised and unsupervised object detection
- Benefit available data with low annotation effort (empty vs non-empty)

Outputs will be served for:

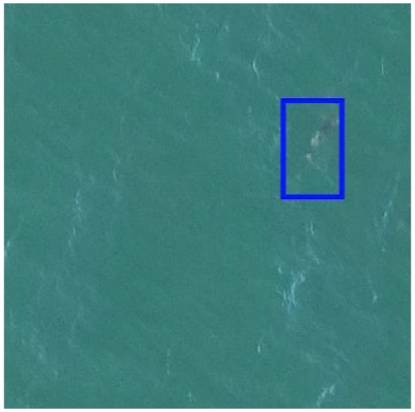
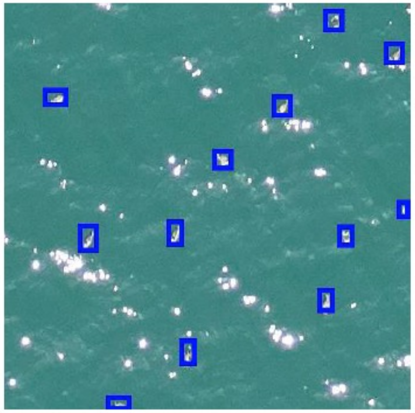
- Animal discovery in new flights
- Quick density estimation
- Animal proposals in annotation task

Introduction

Challenges

- Animals at different depth levels from the sea surface
- Sun glitters and wave crests
- Various background *w.r.t.* weather, season, geography location, etc.



State-of-the-art

Weakly supervised: train only with empty samples → learn data normality

Deep embedding-based approach

- Extract feature vectors using pre-trained network
- Use a **distance metric** in the feature space
- Use **distributions or statistical tests** in the feature space

Reconstruction-based approach

- Use generative models (VAE, GAN, etc.) to reconstruct normal images
- Discover anomalies thanks to their **poor reconstruction**

In practice: they are mixed and/or combined

Our work (in the context of SEMMACAPE project)

Deep embedding-based approach

Article

Weakly Supervised Detection of Marine Animals in High Resolution Aerial Images

Paul Berg [1], Deise Santana Maia [2], Minh-Tan Pham [1,*] and Sébastien Lefèvre [1]

Remote Sensing, 2022

Deep embedding + Reconstruction-based approach

Leveraging Vector-Quantized Variational Autoencoder Inner Metrics for Anomaly Detection

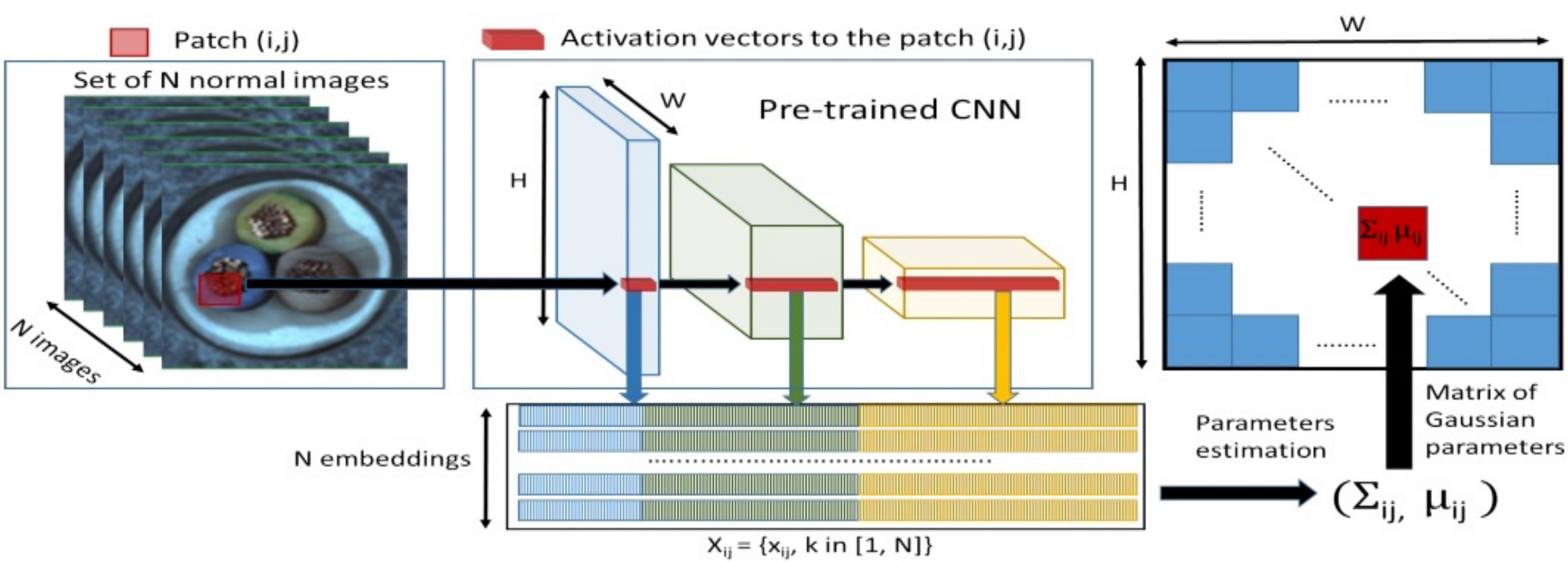
Hugo Gangloff, Minh-Tan Pham, Luc Courtrai, Sébastien Lefèvre
IRISA, Université Bretagne Sud, UMR 6074

ICPR, 2022

Patch distribution modeling & normalizing flows

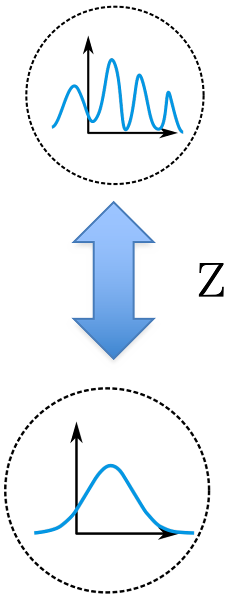
Patch distribution modeling + normalizing flows

- Based on the existing work PaDiM [1] (ICPR 2021)
  - Gaussian distribution for patches + Mahalanobis distance
  - Radom dimension down-sampling → a single Gaussian distribution



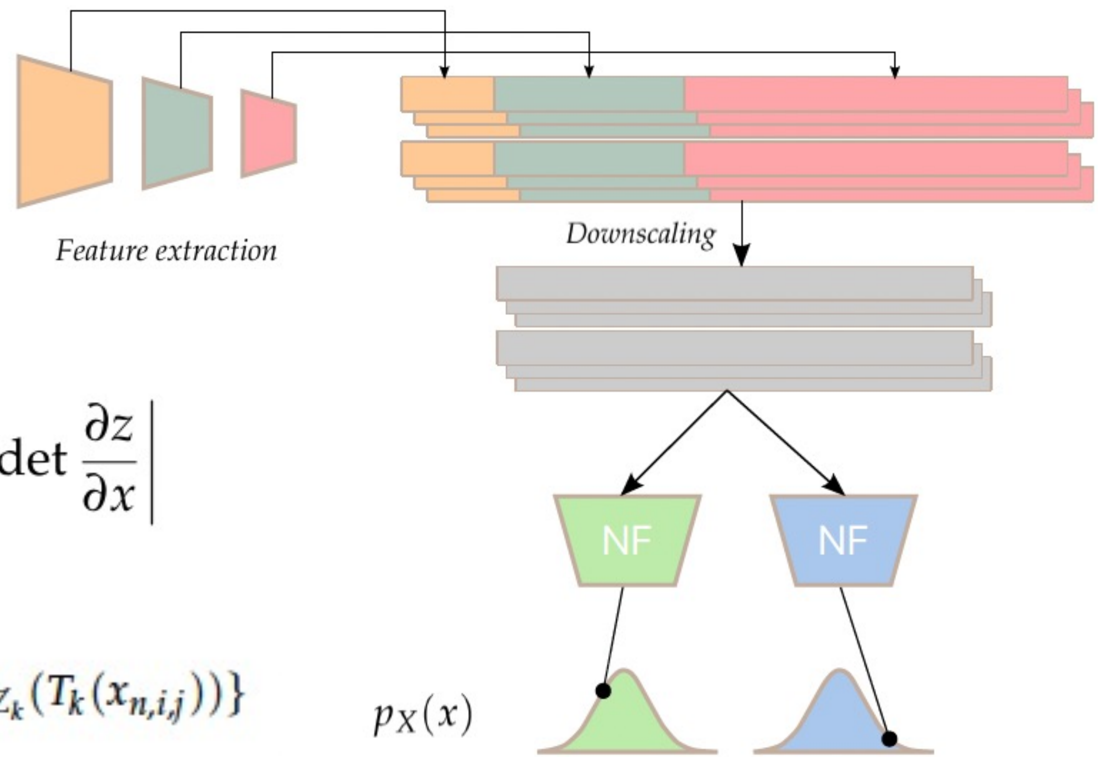
Patch distribution modeling + normalizing flows

- Our approach: use normalizing flow (NF) to transform patch distribution into Gaussian distribution (multi-head)



$$Z = T(X)$$

$$p_X(x) = p_Z(z)\left|\det \frac{\partial z}{\partial x}\right|$$



$$\mathcal{L} = \frac{1}{N \times W \times H} \sum_{n,i,j} \min_{k}\{-\log p_{Z_k}(T_k(x_{n,i,j}))\}$$

T(.): invertible transformation adopted by using Masked Autoregressive Flow (MAF)

Patch distribution modeling + normalizing flows

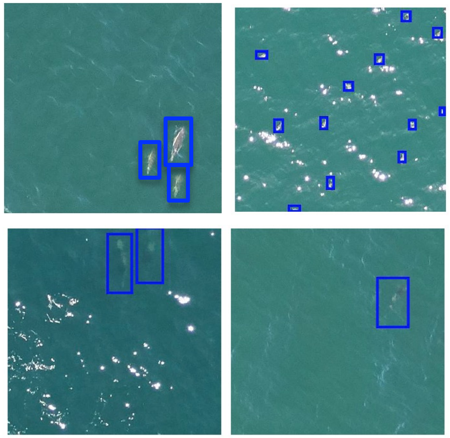


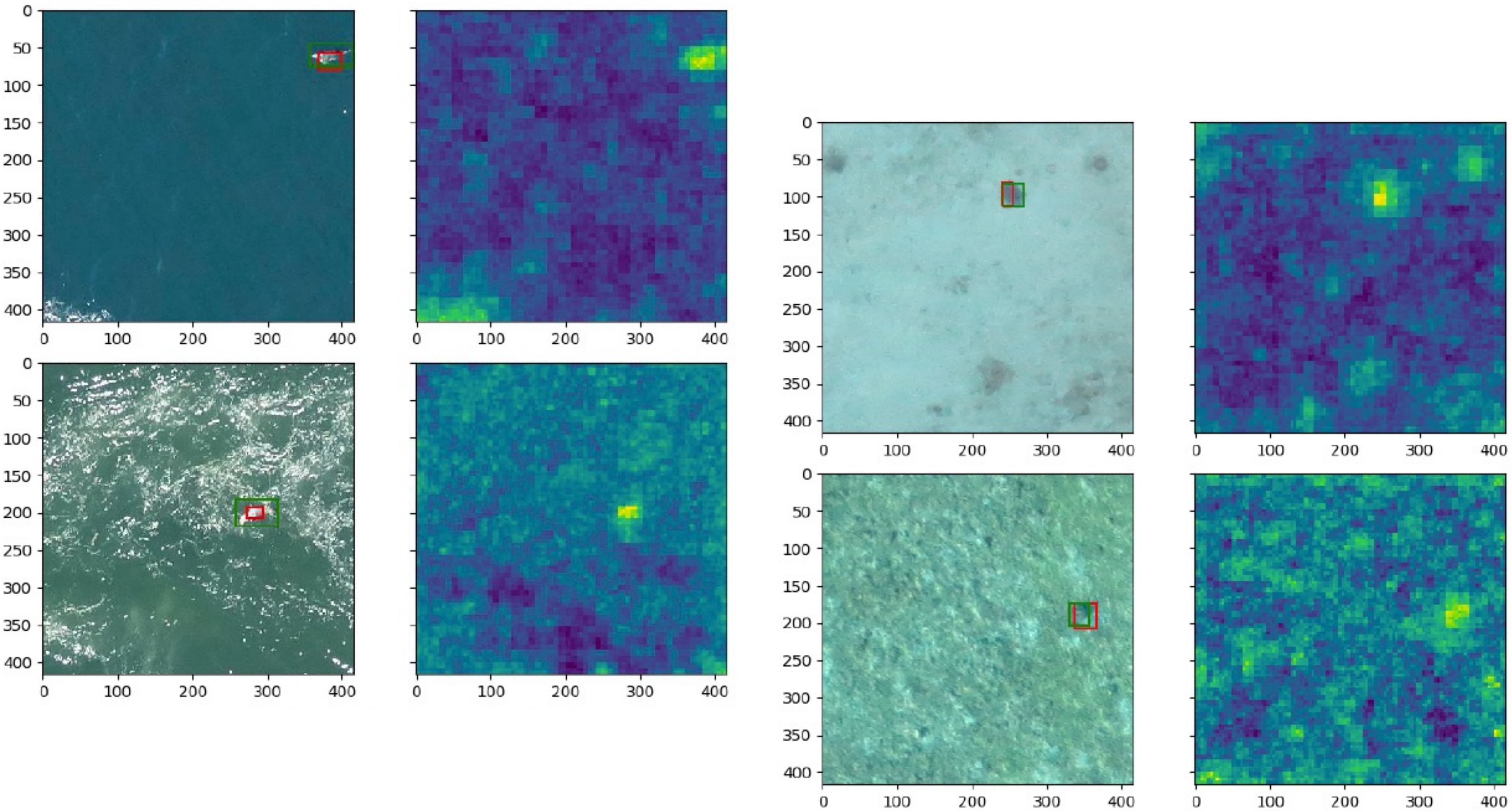
Table 1. Results on the Semmacape dataset. All the models have been trained using the same set of 1000 training images.

Method	F1 Score	Recall	Precision	AUROC
PaDiM [11]	0.383	0.434	0.343	0.606
OrthoAD [12]	0.458	0.373	0.594	0.795
AnoVAEGAN [13]	0.469	0.531	0.420	0.697
Ours, 1× MAF [33]	0.530	0.757	0.408	0.919
Ours, 2× MAF [33]	0.486	0.523	0.455	0.869

Table 2. Results on the Kelonia dataset. All the models have been trained using the same set of 1000 training images.

Method	F1 Score	Recall	Precision	AUROC
PaDiM [11]	0.504	0.443	0.586	0.431
OrthoAD [12]	0.571	0.514	0.643	0.431
AnoVAEGAN [13]	0.051	0.033	0.107	0.469
Ours, 1× MAF [33]	0.568	0.559	0.578	0.410
Ours, 2× MAF [33]	0.584	0.566	0.604	0.391

Patch distribution modeling + normalizing flows

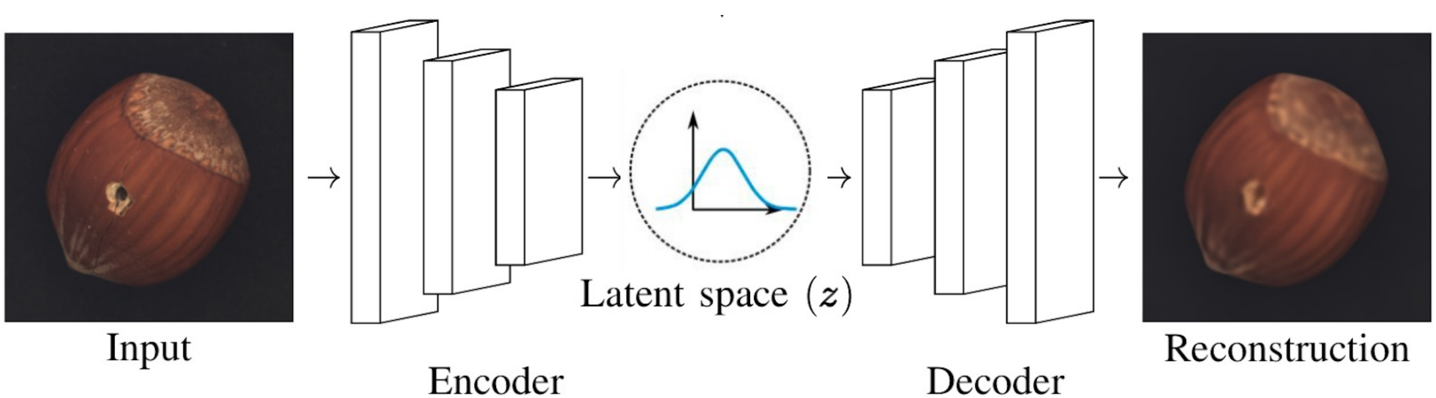


Remark: the anomaly threshold selection is significant (and data-dependent) !!!

Vector-quantized Variational autoencoder
& Inner metrics

Vector-quantized Variational autoencoder + Inner metrics

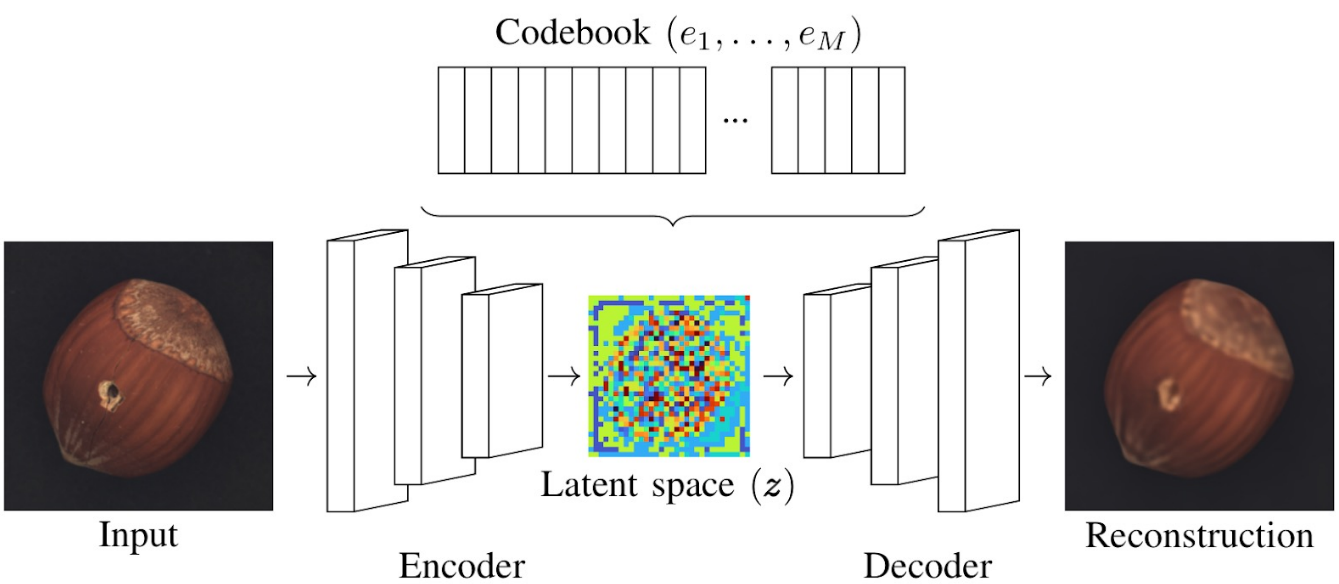
- **VAE-like generative model for anomaly detection**
  - Popular and widely-used in vision
  - Learning of data normality within the latent space with reconstruction + regularization losses



$$\mathcal{L}_{\boldsymbol{\theta},\boldsymbol{\varphi}}(\boldsymbol{x}) = \underbrace{\mathbb{E}_{q_{\boldsymbol{\varphi}}(\boldsymbol{z}|\boldsymbol{x})}[\log p_{\boldsymbol{\theta}}(\boldsymbol{x}|\boldsymbol{z})]}_{\text{reconstruction term}} - \underbrace{\mathbb{KL}(q_{\boldsymbol{\varphi}}(\boldsymbol{z}|\boldsymbol{x})||p_{\boldsymbol{\theta}}(\boldsymbol{z}))}_{\text{regularization term}}$$

Vector-quantized Variational autoencoder + Inner metrics

- **Our approach**
  - Adopt VQ-VAE (NeurIPS’17): popular approach with high reconstruction quality
  - Propose a novel inner metric called Alignment Map to improve the detection



$$\mathcal{L}^{VQ-VAE}_{\boldsymbol{\theta},\boldsymbol{\varphi},\boldsymbol{e}}(\boldsymbol{x}) = \log p_{\boldsymbol{\theta}}(\boldsymbol{x}|z_{\mathrm{Dec}_{\boldsymbol{\theta}}(\boldsymbol{x})}) + \underbrace{\|\mathrm{sg}[z_{\mathrm{Enc}_{\boldsymbol{\varphi}}(\boldsymbol{x})}] - \boldsymbol{e}\|_2^2}_{\text{alignment term}} + \beta\|z_{\mathrm{Enc}_{\boldsymbol{\varphi}}(\boldsymbol{x})} - \mathrm{sg}[\boldsymbol{e}]\|_2^2,$$

Vector-quantized Variational autoencoder + Inner metrics

- **Our approach**
  - Adopt VQ-VAE (NeurIPS'17): popular approach with high reconstruction quality
  - Propose a novel inner metric called Alignment Map to improve the detection

4:45-5:00	**Estimation de flot optique basé évènements en temps réel**	Vincent Brebion (UTC); Julien Moreau (UTC); Franck Davoine (UTC)
5:00-5:15	**Analyse de situations conflictuelles dans l'habitacle d'un véhicule par apprentissage profond**	Quentin Portes (Renault Software Labs)
5:15-5:30	**Autoencodeurs variationnels à registre de vecteurs pour la détection d'anomalies**	Hugo Gangloff (Telecom Sudparis); Minh-Tan Pham (IRISA); Luc Courtrai (IRISA); Sébastien Lefèvre (IRISA)
5:30-6:30	**AG AFRIF**	
7:00-11:00	**SOCIAL EVENT**	

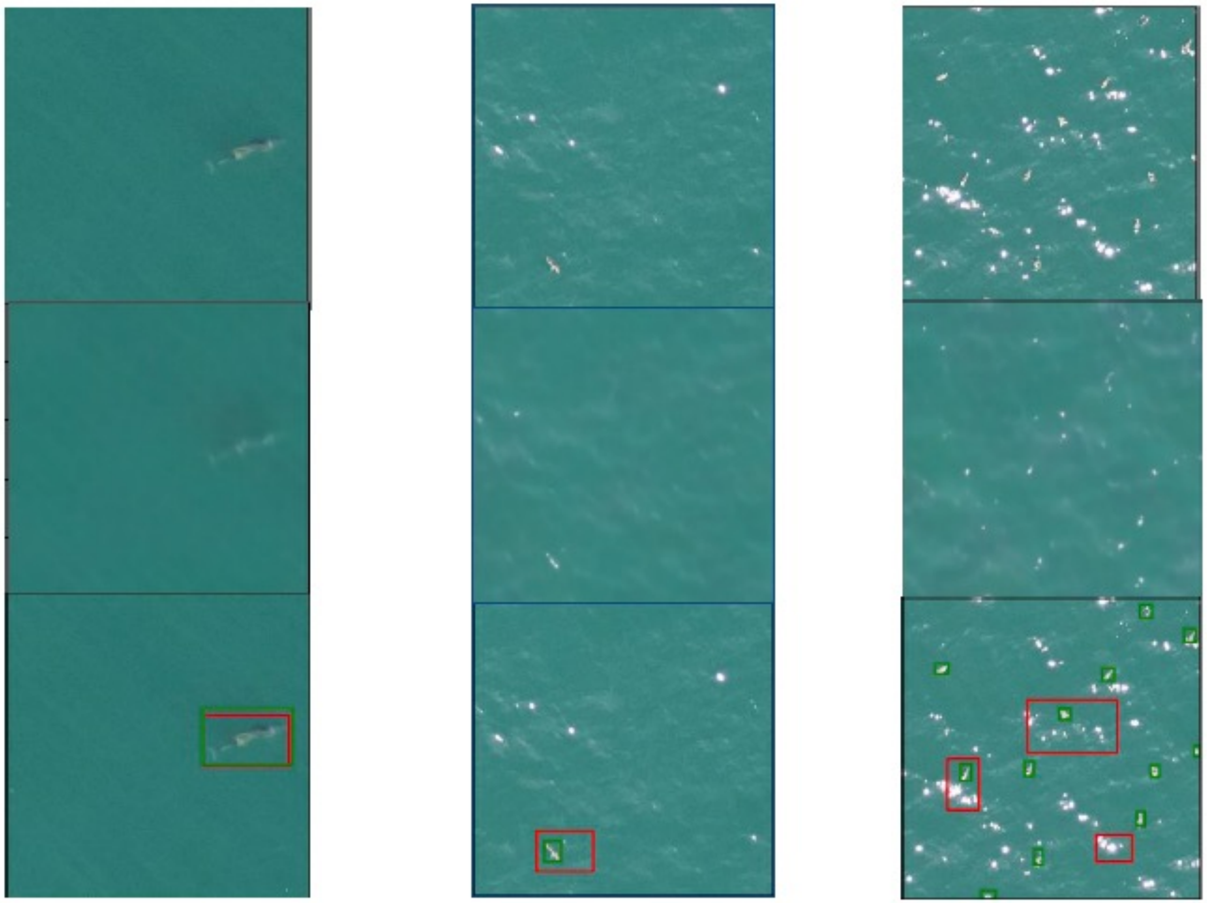
For more details about our methodology

- ICPR'22 paper (available online)
- ORAL presentation of Hugo Gangloff on **Thursday 5:15-5:30pm**

Vector-quantized Variational autoencoder + Inner metrics

Preliminary results show VQ-VAE+AM provides better performance than PaDiM+NF

Illustration of reconstructed images: successful and failed cases



REFERENCES:

Defard, T., Setkov, A., Loesch, A., & Audigier, R. (2021, January). Padim: a patch distribution modeling framework for anomaly detection and localization. In International Conference on Pattern Recognition (pp. 475-489). Springer, Cham.

Berg, P., Santana Maia, D., Pham, M. T., & Lefèvre, S. (2022). Weakly Supervised Detection of Marine Animals in High Resolution Aerial Images. *Remote Sensing*, *14*(2), 339.

Gangloff, H., Pham, M. T., Courtrai, L., & Lefèvre, S. (2022). Leveraging Vector-Quantized Variational Autoencoder Inner Metrics for Anomaly Detection.

Salehi, M., Mirzaei, H., Hendrycks, D., Li, Y., Rohban, M. H., & Sabokrou, M. (2021). A unified survey on anomaly, novelty, open-set, and out-of-distribution detection: Solutions and future challenges. *arXiv preprint arXiv:2110.14051*.

Van Den Oord, A., & Vinyals, O. (2017). Neural discrete representation learning. *Advances in neural information processing systems*, *30*.

Vo, H. V., Bach, F., Cho, M., Han, K., LeCun, Y., Pérez, P., & Ponce, J. (2019). Unsupervised image matching and object discovery as optimization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8287-8296).

Weakly supervised detection of marine mammals

Pierre Allain & Minh-Tan Pham

WIPSEA, F-35000 Rennes, France
Univ. Bretagne Sud - IRISA, F-56000 Vannes, France

MACLEAN – CAP&RFIAP

05/07/2022

Institut de Recherche en Informatique et Systèmes Aléatoires