

DEEL MAG #02

TOULOUSE & QUEBEC JOIN FORCES TO DEVELOP AI FOR CRITICAL SYSTEMS



DEEL at ERTS 2020 in Toulouse on January

1

On 29-31 January took place [ERTS 2020](#) “Embedded real time systems” conference. The DEEL Team was present in this conference through:

- ▶ The talk of Eric JENN presenting **the certification mission paper on *Current Challenges in the Certification of Machine Learning for Safety Critical Systems*** # Eric Jenn, Alexandre Albore, Franck Mamalet, Grégory Flandin, Christophe Gabreau, Hervé Delseny, Adrien Gauffriau, Hugues Bonnin, Lucian Alecu, Jérémy Pirard, Baptiste Lefevre, Jean-Marc Gabriel, Cyril Cappi, Laurent Gardès, Sylvaine Picard, Gilles Dulon, Brice Beltran, Jean-Christophe Bianic, Mathieu Damour, Kevin Delmas, Claire Pagetti)
- ▶ The chair of Adrien Gauffriau on a **formal verification session**
- ▶ The panel of Claire Pagetti on "**Certification of machine learning for safety critical applications**" # with the participation of Hugues Bonnin, Christophe Gabreau, Franck Mamalet with the presence of Guillaume Soudain – EASA & Loic Correnson – CEA

AI CERTIFICATION MISSION

The successes of Machine Learning in solving difficult problems or improving significantly the performance of existing systems, make their dissemination ineluctable in safety critical systems. But existing development assurance standards used for certification (e.g. DO-178C, ISO26262) of classical systems, are incompatible with data-driven development process of Machine Learning based ones. The certification frameworks should therefore evolve to ensure that errors stemming from these activities are mitigated.



In DEEL project, the “Certification Mission” gathers experts on AI and safety critical embedded systems and certification, from key players in the automotive, railways, and aeronautical domains, with a strong connection with the core Team, to identify the main challenges to the Certification of Machine Learning for Safety Critical Systems.

First year of this workgroup was devoted to both AI and multi-sector certification acculturations, and to a global analysis of off-line supervised ML to produce a White Paper (due to early Q2.20) including identified challenges. After a consultation period of two months, the working group decided to continue this mission for 2021, focusing for next months on two use case challenges between AI and Certification experts (ACAS-Xu and image processing), and a special theme on probabilistic assessment.

Franck MAMALET

LANDCOVER DATA SET



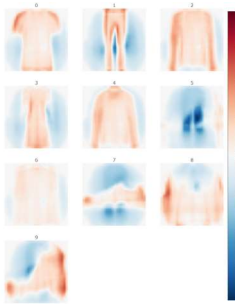
3 remote sensing datasets have been made available to the DEEL team. It is composed of Sentinel2 images: Landcover (100k 32x32 images with ground truth for 5 classes, ~300MB), Landcover Resolution (100k 32x32 images and ground truth for 5 classes, multiple resolutions: 10m and 20m, ~300MB), EuroSAT (27k 64x64 images and ground truth for 10 classes, no atmospheric correction, ~300MB). A specific problematic focuses on the bias due to

atmospheric phenomenon that can influence decisions made by a Deep Learning model. This bias can be measured in EuroSAT with an indicator called the Disparate Impact that evaluates whether a specific bias introduces a difference of treatment between sensible and non-sensible data. A regularization method, based on Optimal Transport theory and the Wasserstein-2 distance, has been developed and shows that we can correct any model during its training phase and make fair decisions with respect to the bias, with a low cost of performance.

Laurent RISSER, Quentin VINCENOT, Nicolas COUELLAN, Jean-Michel LOUBES

EXPLAINABILITY CHALLENGE

To get confidence in the results computed by the artificial intelligence models, the different actors involved in the life of an industrial product should understand the underlying logic of these models: how these models respond to a specific input and which main logics the model follows. In ML community, there are no consensus on a definition of the explicability. Several approaches are explored in the literature.



In DEEL, we first did a review of the literature and we wrote a **state of the art document**. We did a **traceability between technics, audience (Explainability for who), problematics (for what) and industrial datasets**. It allows identifying if a technology is suitable for a specific audience or a kind of problem. We also identify advantages or disadvantages and the degree of trust on the understanding of the model that can provided these techniques.

We selected the **development axes for the Explainability challenge based on the joint vision of industrials and researchers**. The first axis is oriented towards “research” which will explore the formal methods and the metrics for Explainability. For the second axis, we will **experiment existing state of the art technics** (internal model analysis and build interpretable features) on industrial usescases. In parallel of this top down approach, we experiment on industrial usescase the **Ethik library** designed by DEEL researchers, which can predict the importance of a variable in the model decision process.

LOUBES Jean-Michel, Edouard PAUWELS, Mario MARCHAND, Raphaël PUGET, Adrien GAUFFRIAUX, Mélanie DUCOFFE, Frederic BOISNARD, Pierre BLANC-PAQUE, Bertrand CAYSSIOLS, Agustin Martin PICARD, Jean-Marc GABRIEL, Baptiste LEFEVRE, Florence DE GRANCEY, Franck MAMALET, David VIGOUROUX, Mikael CAPELLE, Mathieu DAMOUR, Camille CHAPDELAINE, Thomas SOUMARMON



KEY DATES

26 & 27 mar	Certification mission Workshop
23 & 24 apr	Certification mission Workshop
26 & 27 may	Mobilit.AI
28 may	DEEL - International Operational Committee (COMOPS#4 with Quebec partners)



PROJECT MANAGEMENT

Project Contract R2V1 has been issued and signed by 10/11 partners. [ANR report available](#) on DEEL SharePoint.

✚ 1 secondment for the certification mission (20%), ✚ 1 trainee