DEEL France-Quebec

COOPÉRATION FRANCE-QUÉBEC

NEWSLETTER #11

June 2021

TOULOUSE & QUEBEC JOIN FORCES TO DEVELOP AI FOR CRITICAL SYSTEMS

CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2021 : ACHIEVING ROBUSTNESS IN CLASSIFICATION USING OPTIMAL TRANSPORT WITH HINGE REGULARIZATION

Do you think Spock could have a moustache?

Robustness of Neural Network to adversarial samples [1] is an important challenge, and L-lipschitz property for Neural Network was pointed out as an intrinsic positive factor for robustness. A function f is said L-Lipschitz when its first derivatives are bounded by L, or equivalently when function outputs distance ||f(x+ε)-f(x)|| is lower than L times the distance of its inputs L.||ε|| , certifying that variations on outputs are bounded within a neighbourhood of x.

But learning 1-lipschitz Neural Networks is hard and often lead to poor accuracy. **DEEL core team has proposed and presented, at CVPR'21** (http://cvpr2021.thecvf.com/), a new loss for classification, called hKR, using Optimal Transport with Hinge Regularization, that both optimize robustness and accuracy. Mathematical proofs are also provided on existence and uniqueness of the solution, link with optimal transport, and provable robustness (the output f(x) directly encodes the robustness). And experimentally, adversarial attacks becomes counterfactuals (Adding a moustache to Spock)

The DEEL-LIP library has also been developed to easily construct and learn L-lipschitz convolutional neural networks with guaranties for each layer, and to export weights into conventional layers for inference after learning. This library is already available on github (https://github.com/deel-ai/deel-lip)

Fig: Adversarial samples found on Neural Networks learnt with hKR loss are no more adversarial but counterfactual (left: MNIST 0 vs 8 classification; right: Moustache/No-Moustache Celeb-A classification; Original image/Adversarial found/Difference)

Adversarial samples found on Neural Networks learnt with hKR loss are no more adversarial but counterfactual (left: MNIST 0 vs 8 classification; right: Moustache/No-Moustache Celeb-A classification; Original image/Adversarial found/Difference)

 # Thibaut BOISSIN, Eustasio DEL BARRIO, Alberto GONZALEZ-SANZ, Jean-Michel LOUBES, Franck MAMALET, Mathieu SERRURIER

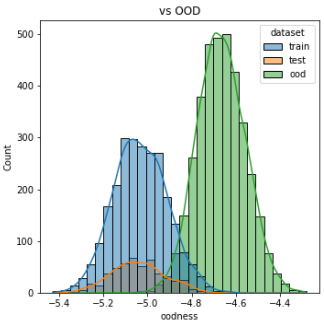
The paper about #Achieving Robustness in Classification using optimal transport with hinge regularization accepted to CVPR 2021.

Congrats to Thibaut BOISSIN, Eustasio DEL BARRIO, Alberto GONZALEZ-SANZ, Jean-Michel LOUBES, Franck MAMALET, Mathieu SERRURIER !

https://arxiv.org/abs/2006.06520

CAN WE BE SURE IN A DOMAIN WE DO NOT KNOW ?

"*The question is quickly answered*" (sic). With neural networks we have the same problem of overconfidence using unknown input and this is a big issue with critical systems. Those unknown inputs are called **Out-of-distribution (OOD)** as they are unlikely to be sampled from the training distribution contrarily to the In Distribution (ID) data.



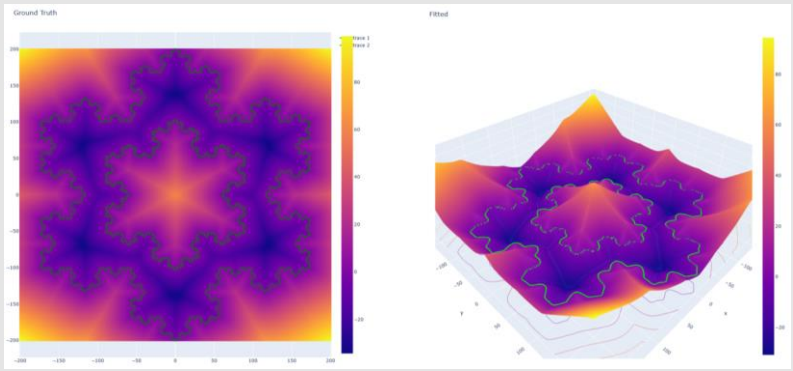
Detecting OOD is a difficult challenge. After reviewing the current SoTA , our current approach is to create a representation of the ID data in a space with a coherent density function so that data which representation lies into low density space can be considered as OOD. For that, we are working on Variational Autoencoders (VAE), Normalizing Flows (NF) and 1-Lipschitz networks and tweaking data distributions and distances.

Louis BETHUNE, Camille CHAPDELAINE, Adrien ELFASSI, Raphaël PUGET, David VIGOUROUX

FOCUS ON A PhD STUDENT, Louis BETHUNE

The emergence of deep neural networks have caused a major revolution in the field of artificial intelligence, allowing to tackle difficult problems such as image annotation, speech recognition, natural language processing, sometimes outperforming humans. However, there is still major flaws preventing their deployment at large scale and within critical systems. They require a huge amount of data, which may be prohibetly expensive in some situations. Moreover, they are vulnerable to adversarial attacks: an imperceptible change in the input (such as removing the right pixel) can cause the network to change its decision. Finally, the decision rules inferred during training are not easily interpretable.

My work focus on **learning representations**: rather than training a neural network specialized in a task, we look for a general architecture able to infer invariances from data, to compress information by separating signal from noise, and to encode semantic similarities in geometric proximity. This is quite a challenge! More recently I explored Lipschitz constrained neural networks.



[1]: by constraining their smoothness, they are naturally immune to adversarial attacks and they yield robustness certificates. We shown that they are powerful enough to solve all classification tasks, and they generalize well on unseen examples [2]. These first results give us full confidence that Lipschitz Neural Networks are a promising tool for representation learning, and ultimately overcoming deep neural networks weaknesses

Figure 1. Syntetic classification task with two classes. The frontier between classes is defined with Von Koch Snowflake, the support of class -1 is the interior ring while the suport of class +1 are the center and the exterior. Despite the fractale structure of the frontier, the Lipschitz Neural network fit perfectly the ground truth smoothly.
The benchmark of the algorithms will be performed on satellite images and time series anomaly detection tasks, with data of Thales Alenia Space, in collaboration with team leaded by Marc Spigai.

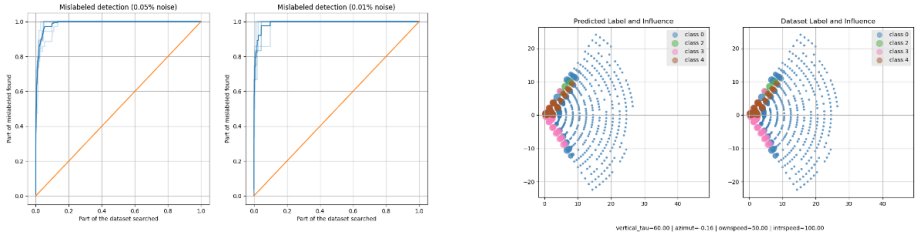
[1] Serrurier, Mathieu, Franck Mamalet, Alberto González-Sanz, Thibaut Boissin, Jean-Michel Loubes, and Eustasio del Barrio. "Achieving robustness in classification using optimal transport with hinge regularization." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 505-514. 2021. https://arxiv.org/abs/2006.06520
[2] Béthune, Louis, Alberto González-Sanz, Franck Mamalet, and Mathieu Serrurier. "The Many Faces of 1-Lipschitz Neural Networks." (2021). https://arxiv.org/abs/2104.05097

 # Louis BETHUNE, Mathieu SERRURIER, Marc SPIGAI

DETECTING DECISION BOUNDARIES AND MISLABELED DATA THROUGH INFLUENCE FUNCTIONS

Back in the 1980's, some robust statistics researchers created **the concept of ``Influence Function''** to measure the impact a change in the training distribution would have on a given statistical estimator. Traditionally, such a mathematical device required a considerable amount of computational power to compute, but with the advent of auto-differentiation frameworks, some recent work has found a way to approximate it using the neural network's gradients. Right from the start, we surmised these influence functions could be useful for the detection of samples or groups of samples of interest, such as mislabeled examples, data-points located near discontinuities of the approximated function or of those near the decision boundary.

As such, we have analyzed **its suitability for the detection of interesting regions on the ACAS Xu use-case, as well as the detection of (synthetically generated) mislabeled examples** on the CIFAR-10 image classification dataset, with considerable success. Then, **we gathered these results and wrote a comprehensive report, and developed a python library implementing this functionality, both of which will be available soon.**



 # Thomas FEL, Jean-Michel LOUBES, Edouard PAUWELS, Agustin MARTIN PICARD, David VIGOUROUX, Quentin VINCENOT, Petr ZAMOLODTCHIKOV

MobiliT.AI 2021 FORUM : A SUCCESS FOR THIS 2nd EVENT !

This 2sde edition of the forum - in virtual mode - will have gathered about **30 international specialists who debated on certification issues & guarantees of artificial intelligence, etc., in the mobility and transportation sectors (aeronautics, automotive, rail, space, drones...).**

Some figures: 600 registrations, 14 conferences, 3 tutorials, 7 scientific posters, 3 round tables on strategic topics related to the development of AI: certification and theoretical guarantees, certification and hybrid approaches, and "embarcability".

In the background, the advances in AI over the last few years have highlighted the need to cope with the legislative framework, to understand algorithms, to define the potential ethical impact of an application, or to ensure the reliability of an airplane or a car.

Soon on the new DEEL Project website: the whole MobiliT.AI 2021 Program, videos and posters of the PhD students → **www.deel.ai**

Thank you for your participation, and see you in 2022 for the 3rd MobiliT.AI forum !

#Organizing Committee 2021

KEY DATES & INFORMATIONS

Certification Mission	Next workshops : August, 25th & 26th – September, 29th & 30th - (Dates to be confirmed)
« Les Carrefours DEEL »	Next Carrefours DEEL: September, 2nd
Annual CONFIANCE.AI days	October, 5th & 6th