Detection of anomalies using Deep Learning Methods

3-year PhD position

The detection of anomalies is a frequent topic in a number of fields (surveilance of networks for attacks, healthcare, text data [1], detection of material defects [2], etc).

In materials we want to detect harmful defects, regardless their shape, and ignore machininng traces or acquisition artifacts that are visible but not real defects. In healthcare, during a screening, we want to raise alert when some other, potentially lethal, disease is present, instead letting the patient go home.

The difficulty resides in incomplete or unavailable databases for such tasks. In control, some defects or faults are rare, or can have undetermined form, in healthcare, any lethal pathology should be detected and only those that a classifier has been trained for.

Using classical statistical methods the anomalies are regarded as outliers to some distribution and detected as such [3]. In machine learning this paradigm is known as unsupervised learning. More complex setups, such as images, have been studied using deep learning (DL) techniques. In the context of DL, Uzunova et al. use autoencoders to learn the internal structure in data [4]. The outliers are detected using the recontruction error commited by the second part of the autoenconder. Another method, based on learning the normality in patches has been proposed Napoletano at al. [2] to detect defects in materials. More generally, Hendrycks and Gimpel [5] show that a soft-max classifier can be trained to show a lower max-probability on out-of-distribution examples. When a sample of out-of-distribution data is available, a more performant classifier can be trained with the technique called Outlier Exposure [6].

The lack of robustness of machine learning in detecting unseen patterns slows or even blockades its application in industrial or medical context. In this PhD we are interested indeveloping a robust methodology in detection of abnormalities in images from various domains, such as material control or healthcare, to detect unseen defects, or pathologies, unseen during the training phase, able of reliable deployment in industrial or medical context.

Working plan: Working towards the PhD withing this subject will involve

- studying the literature, learning coding tools and existing algorithms
- Experimentation with datasets and development of new algorithms
- Evaluation and comparison with the state of the art
- Publishing results in conferences and journals

Location : 3 yrs PhD at the Centre for Mathematical Morphology (CMM), MINES ParisTech. The campus is located in the centre of the historical, royal town of Fontainebleau, one hour from Paris, in the heart of one the most beautiful forest of France. The team of CMM is ten permanent researchers and 15 PhD candidates, post-doctoral fellows and trainees.

Starting date : September 1 2019

Required competences : Image processing ; probability and statistics ; python or matlab coding skills

Contact : An excellent school note record, CV and a cover letter is required to apply. Prospective candidates should contact the : petr.dokladal@mines-paristech.fr (PhD director)

Literature :

[1] Gorokhov, Oleg, & Petrovskiy, Mikhail, & Mashechkin, Igor. (2017). Convolutional Neural Networks for Unsupervised Anomaly Detection in Text Data. 500-507. 10.1007/978-3-319-68935-7_54.

[2] Napoletano, P. & Piccoli, F. & Schettini, R. (2018). Anomaly detection in nanofibrous materials by CNN-based self-similarity. *Sensors*, *18*(1), 209.

[3] Reed, I. S., & Yu, X. (1990). Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *38*(10), 1760-1770.

[4] Uzunova, Hristina & Handels, Heinz & Ehrhardt, Jan. (2018). Unsupervised Pathology Detection in Medical Images using Learning-based Methods. 10.1007/978-3-662-56537-7_30.

[5] Hendrycks, D., & Gimpel, K. (2016). A baseline for detecting misclassified and out-of-distribution examples in neural networks. *arXiv preprint arXiv:1610.02136*.

[6] Hendrycks, D., Mazeika, M., & Dietterich, T. G. (2018). Deep anomaly detection with outlier exposure. *arXiv preprint arXiv:1812.04606*.