Multi–modal unsupervised fault detection system via deep AutoEncoder neural network

Francesco Cordoni[1], Luca Di Persio[2], Riccardo Muradore[2]

1: University of Trento - Department of Civil, Environmental and Mechanical Engineering
2: University of Verona - Departiment of computer Science

francesco.cordoni@unitn.it, luca.dipersio@univr.it, riccardo.muradore@univr.it

Abstract

Multi-modal unsupervised Deep Learning based algorithms have been shown to be very effective for fault detection industrial purposes. In this paper, thermal images and electric measurements from the world of the refrigeration industry, are analysed to train and validate a system able to detect faults at run time.

1 Introduction

The last decade has seen a constant increasing attention to the development of Deep Learning (DL) algorithms. Given their incredible impact in many research areas, during recent years a major effort has been devoted to transfer theoretical methods to industrial scenarios, [Muhuri *et al.*, 2019]. In particular, the manufacture industry represents one of the fields which can benefit more from advantages provided by a DL extensive use. Indeed, manufacturing industry is characterized by highly complex processes characterized by many variables each of which concurs realizing significant difference on the final product. To guarantee high-quality production standards, every step of modern industrial processes need to be efficiency optimized. Therefore, all steps involved in the final outcome must be constantly monitored to detect possible malfunctioning as early as possible.

Predictive maintenance aims at real-time monitoring industrial systems trying to realize early interventions to prevent failures. Hence, it is clear that an effective predictive maintenance system allows to obtain relevant economic savings avoiding harmful failures while optimizing manufacturing processes quality. Achieving previous goals implies treating enormous quantity of data [Cordoni *et al.*, 2020], a task in which DL proved to be particularly efficient.

The present paper shows a fully unsupervised DL based fault detection algorithm feeded by data provided by several sensors mounted on refrigerators. In particular, such sensors communicate: temperature, current and power measurements. Data have been collected for any machine at the end of the manufacturing process to check quality via a predetermined testing procedure realized by thermal cameras, placed along the test line, snapshotting refrigerators' backs to characterize temperatures for different regions of interest.

The main ingredient of our work is the Auto-Encoder (AE) Neural Network (NN) architecture. An AE, [Vincent *et al.*, 2010], is an unsupervised type of NN typically used to encode inputs into a lower dimensional space to then decode the new learnt representation aiming at reconstructing the original dataset as close as possible. The *decoded representation* is used to isolate anomalous data which should in principle correspond to faulty devices.

In our case, the considered dataset is extremely unbalanced, the vast majority of data being corresponding to healthy devices. The latter represents a major problem in any fault detection algorithm, an issue efficiently addressed by AE models. In fact, since AE is trained to encode and decode input data, anomalous input should be more difficult to be properly reconstructed, hence *reconstruction error*, namely the square error between the input and the AE output, is expected to be higher for anomalous input. It is worth stressing that in doing so, no label on the status of the system is used so that the AE implying that AE is a fully unsupervised technique.

2 Method

Three different pre-processing techniques have been applied to thermal images before passing the input to the AE. They are labelled *manual*, *semi-automatic* and *automatic*. In the *manual* procedure, a technician manually selects specific regions of interest belonging to processed images to identify the system's state chacterized by several features such as maximum, mean and minimum temperatures.

The *semi-automatic* and *automatic* techniques rely on a deep convolution NN to extract relevant features and their spatial locations from thermal images. A pre-trained deep convolutional NN model, namely the *Visual Geometry Group* (VGG) NN, [Simonyan e Zisserman, 2014], has been used to extract such features from thermal images. Such features are then passed, alongside to structured temperatures and power measurements, to the AE. The major difference between the *semi-automatic* and the *automatic* technique is that the latter considers the whole image, while the former analyses only the relevant image portions, according to experts selection realized within the *manual* approach, given to the VGG network.

As regard the chosen method, some remarks are in order. AE is a natural and powerful tool to treat highly imbalanced dataset. In this direction, several other methods can be used

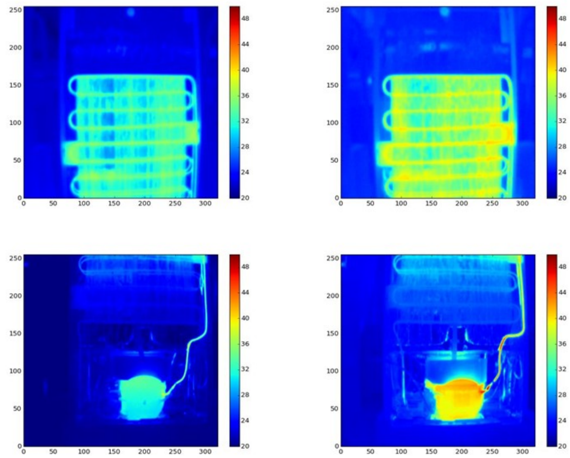

Figura 1: Typical thermal camera images. Left Column: Healthy device; Right column: Faulty device. Top: upper part of the refrigerator; Bottom: Lower part of the refrigerator.

to mitigate the effect of imbalanced classes. Among the most used it is worth mentioning the re-population of classes to obtain a more balance proportions of samples, and the choice of a suitable weighted error to account for a different number of samples belonging to a certain class. Therefore, classical supervised methods, either ML or DL based, can be trained on imbalanced dataset. In this sense, the choice of the proposed model is mainly focused on avoiding using labels as much as possible. Unfortunately, both repopulating the dataset and weighting the training error strongly rely on labels. On the contrary, the proposed method does not use labelling except to optimize the *application-dependent threshold* above which the reconstruction error highlights a faulty device. Let us underline that previous optimization methodology implies the developed method is not fully unsupervised. Indeed, experts' knowledges about the process itself is exploited to set the hyper-parameters. Nonetheless, all DL models are trained in a fully unsupervised manner. Finally, we remark that also supervised *Convolutional Neural Network* (CNN) classification model with F_1-optimization has been tried, obtaining not satisfactory results when compared to those provided by the model presented in this work, particularly because much more information is required to realize a negligible gain.

3 Experimental results

Figure 1 shows examples of thermal images used to train and test the proposed methods. In the figure both healthy and faulty devices are reported. The initial dataset consists of ∼12.000 samples and it has been then split into a train set (∼ 8.500) and a test set (∼ 3.500). Given the fact that the dataset is highly unbalanced (only 30 faults), with the vast majority of samples corresponding to healthy devices, the best threshold above which the AE reconstructed error results in a fault signal, has been chosen optimizing the F_1 score in the train set, hence considering both precision and recall.

Tables 1 and 2 report: confusion matrices, Accuracy, PPV, TPR and F1 for the three methods, respectively.

		Predicted class	
		Fault	**No Fault**
Actual class	**Fault**	3 | 4 | 5	5 | 4 | 3
	No Fault	4 | 6 | 8	3579 | 3577 | 3575

Tabella 1: Confusion matrix for test set for the *manual method* | *semi-automatic method* | *automatic method*

Method	Accuracy	PPV	TPR	F1
Manual	**99.7**%	**42**%	37.5%	40%
Semi-automatic	**99.7**%	40%	50%	44%
Automatic	99.6%	38%	**62**%	**47**%

Tabella 2: Results for three methods; in bold it is highlighted the best performing method according to the specific metric.

Experimental results show that the VGG is able to extract in a fully automatic way relevant features providing results which are very close to those obtained by the manual case. Although the three methods show similar results, VGG based methods have the clear advantage that little or no at all human intervention is required. In particular, since the DL-based method is fully automatic, it is extremely scalable and it can be generalized to consider different devices with little effort.

4 Projects and Funding

The activity of this research was funded by the POR / FESR of the Veneto Region on the "PreMANI – Manifattura Predittiva: progettazione, sviluppo e implementazione di soluzioni di Digital Manufacturing per la previsione del la Qualità e la Manutenzione Intelligente", 2014-2020.

This work was also partly supported by the project "Informatica per Industria 4.0" funded by MIUR "Dipartimenti di Eccellenza" 2018-2022.

5 Conclusions and Challenges

The main contributions of the paper are to:

1. develop a highly sophisticated quality test that allows to consider measurements of different nature, efficiently including both structured data and images;
2. develop an unsupervised deep learning-based algorithm capable of efficiently identify faulty devices in a highly unbalanced dataset using both structured and unstructured data simultaneously.

Becasue of the high scalability of the proposed approach, we aim at applying it to different industrial sectors, the automotive being an elected one, particularly with respect to electric engines, where real-time thermal images analysis could prevent production failures potentially responsible of dangerous issues during on road car use.

A further and more ambitious goal is to exploit the multi–modal unsupervised AE-based technique to analyze medical images in order to early detect abnormal growth of tissues in specific body regions.

Riferimenti bibliografici

[Cordoni *et al.*, 2020] Francesco Giuseppe Cordoni, Gianluca Bacchiega, Giulio Bondani, Robert Radu, e Riccardo Muradore. A deep learning unsupervised approach for fault diagnosis of household appliances. In *IFAC online, 21st IFAC world conference*, 2020.

[Muhuri *et al.*, 2019] Pranab K Muhuri, Amit K Shukla, e Ajith Abraham. Industry 4.0: A bibliometric analysis and detailed overview. *Engineering applications of artificial intelligence*, 78:218–235, 2019.

[Simonyan e Zisserman, 2014] Karen Simonyan e Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[Vincent *et al.*, 2010] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, e Pierre-Antoine Manzagol. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research*, 11(Dec):3371–3408, 2010.