

# **Application des réseaux neuronaux artificiels à l'identification des principaux éléments des trains pour un système de détection de pantographes défectueux automatique.**

ALAIN RIVERO<sup>1,2,3</sup>, DIEGO ARENAS<sup>3</sup>, FLORIAN NAPIERALA<sup>4</sup>, PHILIPPE VANHEEGHE<sup>2,3</sup>

1. SNCF-Réseau, Direction générale industrielle et ingénierie, Département Transformation Technologie et Digital, Paris, France
2. Univ. Lille, CNRS, Centrale Lille, UMR 9189 – CRISTAL, Lille, France.
3. IRT Railenium, Valenciennes, France.
4. Talan, Paris, France

[alain.rivero@reseau.sncf.fr](mailto:alain.rivero@reseau.sncf.fr), [Diego.Arenas@railenium.eu](mailto:Diego.Arenas@railenium.eu), [Florian.napierala@talan.com](mailto:Florian.napierala@talan.com), [Philippe.vanheeghe@centralelille.fr](mailto:Philippe.vanheeghe@centralelille.fr)

**Résumé** - Les arrachements de caténaires ont de très lourdes répercussions sur la régularité des circulations et sur les coûts de maintenance d'un réseau ferré. Les contraintes importantes imposées aux pantographes entraînent des casses et des taux d'usures significatives qui sont à l'origine de ce type d'incident. Afin de prévenir ces arrachages, SNCF Réseau développe un système de détection des pantographes défectueux. Les premiers essais se sont focalisés sur la reconnaissance des éléments globaux (face avant, toiture et pantographe) en se basant, dans un premier temps, sur l'utilisation de classifieurs de type Radial Basis Function (RBF). Les mauvais résultats obtenus par le premier prototype nous a conduits à mettre en œuvre dans un second système des réseaux convolutifs profonds (CNN).

**Abstract** – Catenary pullout has a very serious impact on the regularity of traffic and on the maintenance costs of a rail network. The important constraints imposed on the pantographs cause breakages and significant wear rates which are at the origin of this type of incident. In order to prevent these pull-outs, SNCF Réseau develop a new system to detect defective pantographs. The first tests focused on the recognition of global elements (front, roof and pantograph) based initially on RBF classifiers. The poor results obtained by this first system led us to implement into a second system a deep convolution networks.

## **1 Introduction**

Les réseaux neuronaux artificiels (RNA) constituent une méthode efficace dans la détection de défauts et plus généralement pour résoudre des problèmes nécessitant d'analyser le contenu d'images. En effet, ce type de problème est difficilement modélisable par des approches statistiques ou algorithmiques dites classiques. Le système que nous avons testé fin 2018 pour la détection des défauts de pantographes en est la parfaite illustration. Ces architectures permettent de développer des systèmes compacts, légers et d'un coût modique (trois fois moins cher que les systèmes commercialisés actuellement). Du fait de sa compacité, les futurs capteurs peuvent s'installer sur les portiques de signalisation ferroviaires, les ouvrages d'art ou tous autres supports disponibles. L'utilisation des RNA dans le domaine ferroviaire n'est pas nouveau, de nombreux travaux, notamment ceux de Shen et al. (2018), ont démontré la validité de cette voie d'étude.

Cet article est organisé de la manière suivante : nous traiterons dans un premier temps de l'état de l'art dans le domaine de la surveillance des éléments d'infrastructure. Dans le paragraphe 2 nous décrirons les méthodes de détection disponibles sur le marché. Ensuite, paragraphe 3, nous traiterons des principes de mise en œuvre du dispositif et nous décrirons le principe de fonctionnement. Le paragraphe 4 présente les résultats expérimentaux obtenus en détaillant tout d'abord les résultats obtenus par les méthodes de classification (RBF), puis les modifications apportées au système (utilisation de réseaux convolutifs profonds et l'amélioration de la qualité des images) ainsi que leur

évaluation. Nous terminerons par une conclusion et les perspectives du dispositif.

## **2 Méthodes de détection**

Un pantographe en mauvais état a des effets catastrophiques sur les caténaires (arc électrique, arrachage de caténaire, etc.). En 2012, sur l'ensemble du réseau, 1301 trains ont été impactés produisant 448.266 minutes perdues. Afin de repérer et de prévenir l'apparition de ces phénomènes, des systèmes de surveillance par caméra intelligente peuvent être mis en place. L'analyse des pantographes sur des trains en circulation n'est pas nouvelle, plusieurs sociétés ont développé des systèmes permettant d'identifier les pantographes défectueux (Camlin, PantoInspect, Riccardo).

Dans le domaine de l'identification des défauts, l'utilisation d'un programme qui « apprend » est nettement plus efficace qu'un programme qui se contente d'appliquer la même stratégie quel que soit la situation rencontrée. Les défauts sont par nature très différents des uns des autres et ils sont positionnés aléatoirement dans l'image, il est donc difficile de trouver une forme générique analysable par les méthodes « classiques ». De plus les Réseaux neuronaux artificiels ne nécessitent pas d'extraire préalablement les caractéristiques des défauts. Ils ont une excellente capacité de généralisation lorsque le défaut subit un décalage, un changement de taille ou de point de vue. La représentation de l'objet est suffisamment précise pour différencier les défauts d'une situation normale rendant leurs emplois parfaitement adaptés à notre problématique.

Les difficultés dans l'utilisation des algorithmes de traitement d'image pour la reconnaissance des défauts peuvent être résumées de la manière suivante :

- Performance de l'UC afin d'obtenir des temps de traitement raisonnable qui explique la taille des dispositifs commercialisés (environ 100kg) ;
- La grande diversité des formes des défauts à reconnaître, l'apparition d'une nouvelle forme génère le développement de nouveaux algorithmes ;
- La sensibilité de ces traitements à la variation de la taille des défauts. Les systèmes actuels sont sensibles aux variations de température qui peuvent déformer la bande de carbone ;
- La complexité de la mise à jour afin d'intégrer de nouvelles formes (modification du code, série de test et monté de version) ;
- L'obligation d'intégrer plusieurs méthodes afin de cerner le défaut (descripteurs HOG (Histogram of Gradient) et BoW (Bag of Word) ou la transformée de Hough utilisée pour la détection des pantographes).

Outre les prix prohibitifs des systèmes actuels (environ 150 keuros), ces difficultés expliquent les principales critiques exprimées par les clients à l'encontre des systèmes existants.

### 3 Description du dispositif de mesure

Le premier prototype se compose de deux caméras permettant d'analyser les trains sous deux angles, appliquant les travaux de Yang et Painsavoine (2003). Chaque caméra fonctionne selon deux modes : le mode « veille » (attente d'un train) et le mode d'analyse. Dès qu'un train est détecté par la reconnaissance de la face avant de l'engin de traction ou de la rame, le système ausculte la toiture et recherche les pantographes jusqu'à la disparition complète du convoi du champ des caméras (figure 2). Le système peut fonctionner indistinctement dans les deux sens de circulation des trains. Les trains circulent à des vitesses comprises entre 10 et 160 km/h. Le dispositif s'installe sur les portiques de signalisation ferroviaire sans nécessiter de travaux d'adaptation préalable.

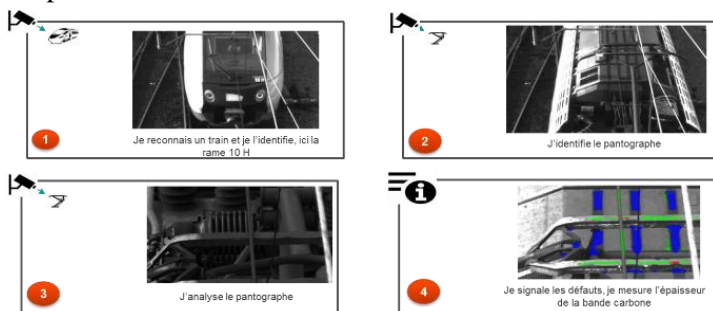


Figure 1 - Principe de fonctionnement du système

Le système est alimenté en 220v, il consomme moins de 30w. Chaque caméra est équipée d'un processeur spécifique neuroFPGA ayant une fréquence d'échantillonnage de 100 MHz hébergeant les algorithmes de traitement des images et d'un Raspberry PI hébergeant les classifieurs SVM avec un noyau

gaussien (RBF). Le système de prise de vue utilise un bloc optique COMPUTAR dont les caractéristiques sont décrites dans le tableau 1 et un capteur APTINA produisant des images au format jpeg avec une résolution de 752x480 en niveaux de gris et un taux de compression de 80%.

Tableau 1 – Caractéristique du bloc optique COMPUTAR utilisé pour le premier prototype

Temps de réponse	4.5sec
Ration ouverture max	1:1.2
Focale	25 mm
Temps d'exposition	1/900s
Angle de prise de vue	58° +/- 5°

## 4 Résultats expérimentaux

Dans un premier temps, nous avons testé un classifieur RBF pour la détection des faces avant, toitures et des pantographes. Les performances médiocres obtenues par ce premier dispositif expérimental nous ont conduits à réorienter nos recherches en utilisant les réseaux convolutifs profonds (He et al., 2016).

### 4.1 Evaluation des réseaux de type RBF

Les résultats obtenus montrent que la première version du dispositif est fortement dépendante de l'orientation des caméras. En effet, chaque installation a nécessité un apprentissage spécifique. Les mesures effectuées sur le site de Savigny (ligne du RER C dans la banlieue Est de Paris) entre juillet 2017 et octobre 2017 montrent une grande instabilité des taux de détection des trains. Sur les 77 trains quotidiens planifiés entre 6h30 du matin et 20h30, seuls 15 trains en moyenne ont été détectés par le système.

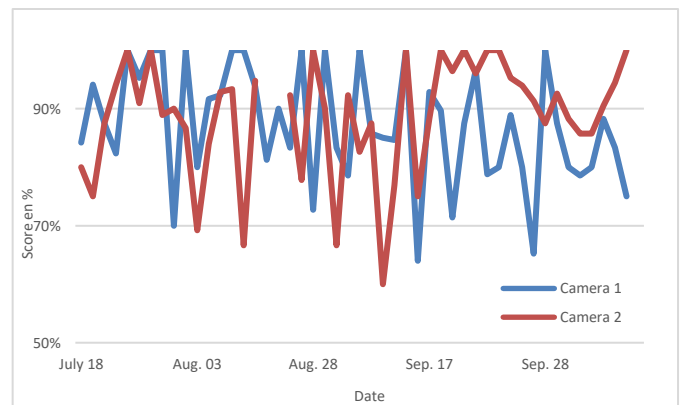


Figure 2 – fluctuation du taux de reconnaissance des trains des deux caméras du premier prototype.

La première version de l'algorithme de détection des numéros de rame situés sur la face avant des engins moteurs, implémenté dans le premier dispositif expérimental est très en deçà des performances obtenues par les travaux de Wang et al. (2015). Nous avons également mesuré la performance globale du système (PGS). (Figure 3) en utilisant un indice qui combine le taux de détection des trains, le taux de reconnaissance des numéros des trains, le taux de reconnaissance des pantographes et le taux de détection de pantographe par train.

$$Score = \frac{N_{TV}}{N_{TC}} \cdot \frac{N_{NT}}{N_{NI}} \cdot \frac{N_{PI}}{2} \cdot \frac{N_{FD}}{2} \quad (1)$$

Où :

$N_{TV}$  : Nombre de trains vu par les caméras ;

$N_{TC}$  : Nombre de trains ayant circulés ;

$N_{NT}$  : Nombre de numéro de trains correctement identifié ;

$N_{NI}$  : Nombre de numéros identifiables ;

$N_{PI}$  : Nombre de pantographes identifiés par les caméras ;

$N_{FD}$  : Nombre de fausses détections.

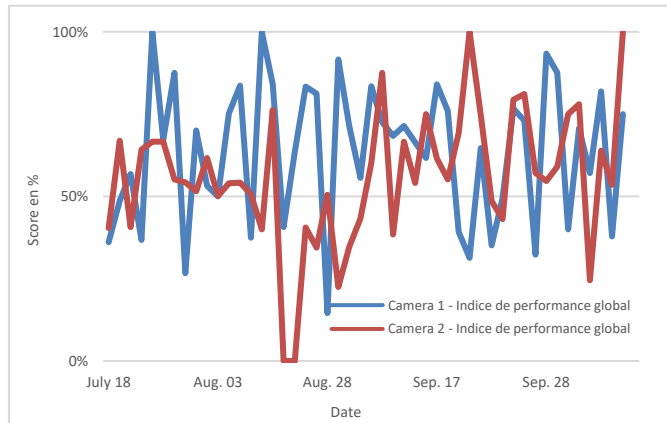


Figure 3 – Taux de performance global du système pour chaque caméra composant le premier prototype.

Ces premiers résultats expérimentaux montrent que cette première version du système n'est pas adaptée à l'identification de formes complexes dans un environnement fortement bruité, malgré les avantages intrinsèques des réseaux RBF, notamment : la simplicité architecturale ou la tolérance aux distorsions. Les raisons de l'échec de cette version sont à rechercher dans la mauvaise qualité des images, le bloc optique n'étant pas adapté au problème posé, et les limitations du processeur FPGA qui nous a contraints à réduire la taille des images avant leurs traitements par la Raspberry Pi. Lors des essais suivants, nous avons changé de technologie en testant les réseaux de convolution profonds et une optique plus performante.

## 4.2 Evaluation des réseaux convolutifs

Les réseaux de convolutions sont particulièrement bien adaptés pour le traitement des images (Krizhevsky et al, 2012). Créer un nouveau réseau de neurones convolutif est coûteux en termes d'expertise, de matériel et de quantité de données annotées nécessaires. Cependant, Il est possible d'obtenir un modèle performant avec peu d'images, une capacité de calcul limitée en se servant d'un modèle pré-entraîné. L'intérêt de cette méthode est double : on utilise une architecture optimisée et l'on profite des capacités d'extraction de caractéristiques apprises sur un jeu de données de qualité. Cette technique de « transfer learning », théorisée par Thrun et Pratt (1998) consiste à extraire une partie, ou la totalité, de la connaissance déjà apprise par un réseau profond sur un problème spécifique puis à utiliser cette connaissance pour obtenir un autre réseau profond

similaire mais résolvant une autre problématique. Nous avons testé plusieurs configurations :

- Plusieurs modèles de réseaux : ResNet 50, VGG16, VGG19, MobileNet, Xception, InceptionV3, InceptionResNet, NASNetLarge, NASNetMobile et DenseNet (121, 201),
- Deux qualités d'images : celles issues de la première version du système installée sur le réseau et celles prises à l'aide d'un appareil photographique installé sur un pont surplombant une voie ferrée.

## 4.3 L'évaluation du modèle ResNet 50

Le premier modèle testé sur les images produites par les caméras installées sur le réseau est le ResNet 50 (He et al., 2016). Plusieurs raisons ont orienté ce choix :

- Cette architecture est très efficace dans le processus d'apprentissage et ses performances dans la reconnaissance des objets sont très intéressantes.
- Sa taille compacte permet son emploi sur un système embarqué.

Pour les phases d'entraînement et de test nous disposons de 12 251 images réparties entre les cinq classes. Les images ont été divisées en deux groupes : la base d'entraînement comportant 11 107 images (90% de la totalité des images) et la base de test avec 1 144 images (10% de la totalité des images). À l'issue de la phase d'apprentissage, nous avons obtenu un taux de précision global moyenne de 92%.

Les premiers résultats nous ont permis d'établir les constats suivants : le réseau ResNet 50 est efficace pour la reconnaissance des faces avant, pantographes et rail. Un nombre non négligeable d'images correspondant à des toitures (10%) ont été confondues avec des faces avant dû au fait qu'il est parfois difficile de différencier une face avant d'une toiture notamment quand la photo montre le haut de la face avant. Ces premiers résultats sont donc très satisfaisants. En effet, nous avons obtenu des très bonnes performances dans la reconnaissance des éléments distinctifs.

## 4.4 L'évaluation des autres modèles de réseau

Afin d'améliorer les résultats obtenus avec le ResNet nous avons réduit le nombre de classes en fusionnant les classes faces avant et numéros. Ce nouveau dataset comporte 11 841 images, réparties en 3 groupes : apprentissage, validation et test. La répartition des images a été également changée, les images sont réparties aléatoirement dans trois sous-groupes : 70% pour la phase d'apprentissage, 15% pour la validation et 15% pour le test. L'apprentissage des réseaux étant sensible aux classes déséquilibrées nous les avons rééquilibrés en prenant aléatoirement pour chaque classe  $n$  images, où  $n$  est la cardinalité minimale de l'ensemble des classes. Les premiers résultats montrent que le cœur du problème se situe au niveau de la différenciation entre les pantographes et les toitures.

Les indicateurs utilisés pour cette étude sont les suivants :

- La précision (P) : rapport entre les éléments correctement prédits et le total des éléments prédits ;
- Le rappel (R) : rapport entre les éléments correctement prédits et tous les éléments d'une classe ;
- Le F1-score (F1) : moyenne pondérée de la précision et du rappel.

Tableau 2 - Comparaison des performances des réseaux artificiels en utilisant le premier dataset

6,	Faces avant			Pantographes			Toitures			
	P	R	F1	P	R	F1	P	R	F1	
VGG16	0.95	0.95	0.95	0.95	0.95	0.93	0.94	0.89	0.91	0.93
VGG19	0.93	0.97	0.95	0.92	0.95	0.94	0.92	0.85	0.92	0.92
Xception	0.74	0.76	0.75	0.71	0.63	0.67	0.63	0.68	0.69	0.69
ResNet 50	0.92	0.96	0.94	0.89	0.95	0.92	0.91	0.81	0.91	0.91
MobileNet	0.95	0.86	0.90	0.97	0.64	0.77	0.66	0.95	0.82	0.82
Inception V3	0.73	0.70	0.72	0.64	0.75	0.69	0.64	0.57	0.67	0.67
InceptionResnet	0.55	0.63	0.59	0.49	0.58	0.53	0.60	0.40	0.53	0.53
DenseNet 121	0.80	0.76	0.78	0.69	0.76	0.73	0.67	0.64	0.72	0.72
DenseNet 201	0.62	0.91	0.74	0.75	0.51	0.60	0.63	0.53	0.64	0.64
NASNetLarge	0.78	0.73	0.76	0.64	0.80	0.71	0.73	0.59	0.71	0.71
NASNetMobile	0.74	0.72	0.73	0.62	0.68	0.65	0.64	0.61	0.67	0.67

Un second essai, en supprimant la classe rail a permis d'améliorer la performance de classification de l'ensemble des modèles. Ce résultat milite pour scinder le processus de détection des éléments en deux traitements :

- Un premier système qui détecte l'arrivée d'un train (identification de la face avant), par l'utilisation d'un laser, d'un radar, ou la détection d'un bruit.
- Un second système qui analyse le train (détection des numéros et des pantographes), cette analyse s'effectuant à l'aide de réseaux neuronaux.

#### 4.5 Changement de la qualité des images

Après avoir considérablement amélioré la définition des images, nous avons testé l'ensemble de nos réseaux sur le nouveau dataset. Le nouveau dataset est constitué de 6297 images. En analysant les résultats obtenus (tableau 3) on constate une amélioration notable des résultats obtenus. Nous avons toujours une légère confusion entre les pantographes et les toitures. Afin de pallier ce problème, il semble judicieux d'utiliser des architectures de réseaux adaptées non plus à la classification mais à la détection d'objets.

Tableau 3 - Comparaison des performances des réseaux artificiels en utilisant le second dataset.

6,	Faces avant			Pantographes			Toitures		
	P	R	F1	P	R	F1	P	R	F1
VGG16	0.98	0.96	0.97	0.93	0.91	0.92	0.89	0.92	0.90
VGG19	0.99	0.98	0.98	0.89	0.98	0.93	0.95	0.87	0.91
Xception	0.90	0.82	0.86	0.79	0.93	0.86	0.89	0.82	0.85
ResNet 50	0.99	1.00	0.99	0.94	0.99	0.96	0.99	0.92	0.96
MobileNet	0.91	0.95	0.93	0.85	0.98	0.91	0.96	0.77	0.86
Inception V3	0.79	0.72	0.75	0.76	0.86	0.81	0.82	0.78	0.80
InceptionResnet	0.49	0.72	0.58	0.66	0.53	0.59	0.78	0.58	0.66
DenseNet 121	0.85	0.88	0.87	0.81	0.87	0.84	0.87	0.77	0.82
DenseNet 201	0.89	0.90	0.90	0.80	0.96	0.87	0.93	0.74	0.82
NASNetLarge	0.81	0.70	0.75	0.75	0.85	0.80	0.82	0.83	0.82
NASNetMobile	0.60	0.70	0.64	0.65	0.64	0.64	0.85	0.72	0.78

## 5 Conclusion et perspectives

Les réseaux convolutifs profonds montrent un fort potentiel pour la mise au point d'un système de détection des défauts des pantographes. Associés à des cartes GPU (Nvidia, Qualcomm, Intel) ils permettent de développer des systèmes compacts capables d'analyser des images

en temps réel. Les tests réalisés sur un panel représentatif de modèles montrent que le modèle ResNet50 est le plus performant dans le cadre de nos expérimentations et dans les conditions spécifiques retenues lors des tests, notamment les contraintes liées à la puissance de calcul disponible. Enfin nous avons démontré qu'il est indispensable de disposer d'images de qualité si l'on veut obtenir un système performant. Nous poursuivons nos travaux par la recherche de défauts sur les pantographes.

Ce dispositif ouvre plusieurs perspectives d'applications dans le domaine ferroviaire, outre la détection des pantographes défectueux, ce type d'architectures pourrait également servir à :

- Détecter certains défauts sur le matériel roulant,
- Réaliser l'inventaire automatisé des équipements d'infrastructures
- Surveiller l'infrastructure (Contrôle des rails, des attaches, de la géométrie de la voie),
- Détecter les personnes divagantes sur les voies ou mesurer l'activité de personnes (analyse de posture) dans les gares ou dans les emprises.

## 6 Bibliographie

He, K., Zhang, X., Ren, S. and Sun, J., (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778.

Krizhevsky A., Sutskever I., Hinton G., (2012). ImageNet classification with deep convolutional neural networks. In Proc. 26th Annual Conf. on Neural Information Processing Systems, Lake Tahoe, NV, 3-6 December 2012, pp. 1090-1098.

Shen Y., Liu Z., Chang L. (2018). A Pantograph Horn Detection Method Based on Deep Learning Network, IEEE 3rd Optoelectronics Global Conference (OGC), Shenzhen, 2018, pp. 85-89.

Thrun S., Pratt L. (1998). Learning to learn: Introduction and Overview. In: Thrun S., Pratt L. (eds) learning to learning, Springer, Boston MA.

Yang F., Paindavoine M., (2003). Implementation of an RBF neural network on embedded systems: Real-time face tracking and identity verification. IEEE Transactions on Neural Networks, 14(5), pp. 1162-1175.

Wang Z., Zhou B., Jegelka J. (2016). Optimization as estimation with gaussian processes in bandit settings. AISTATS, 2016.

Zhang C., Li P., Sun G., Guan Y., Xiao B., & Cong J., (2015). Optimizing fpga-based accelerator design for deep convolutional neural networks. In Proceedings of the 2015 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (pp. 161-170). ACM. February 2015.