

# Détection de fissure sur matériaux métalliques par apprentissage profond et thermographie laser flying-spot : approche par apprentissage progressif

Kevin HELVIG<sup>1</sup> Pauline TROUVÉ-PELOUX<sup>1</sup> Baptiste ABELOOS<sup>1</sup> Ludovic GAVÉRINA<sup>2</sup> Jean-Michel ROCHE<sup>2</sup>

<sup>1</sup>DTIS, ONERA - Université Paris-Saclay, F-91123, Palaiseau, France

<sup>2</sup>DMAS, ONERA - Université Paris-Saclay, F-92322 Châtillon, France

**Résumé** – En contrôle non destructif pour les matériaux métalliques, la thermographie "Flying-spot" permet la détection de fissures grâce au balayage d'échantillons par une source de chaleur active locale observée en spectre infrarouge. Le traitement de ces données est délicat car il faut distinguer une fissure des autres structures à la surface du matériau. Pour cela nous proposons d'utiliser le *deep learning*, capable d'exploiter les informations contextuelles, mais qui nécessite cependant une quantité importante de données annotées. Nous présentons ici une méthode d'entraînement basée sur le *curriculum learning* et les récents modèles de diffusion par débruitage pour générer des images synthétiques, augmentant la quantité de données. Nous évaluons les scores de performance de détection pour diverses architectures d'apprentissage machine et d'apprentissage profond, mettant en évidence les bénéfices de notre approche.

**Abstract** – In non-destructive testing for metallic materials, "Flying-spot" thermography allows the detection of cracks thanks to the scanning of samples by a local active heat source observed in the infrared spectrum. The data processing is difficult because it is necessary to distinguish a crack from other features on the surface of the material. For this we propose to use deep learning, capable of exploiting contextual information, but which nevertheless requires a significant amount of labeled data. We present here a training method based on curriculum learning and the recent denoising diffusion models to generate synthetic images, increasing the amount of data. We evaluate the detection scores for various machine learning and deep learning architectures, highlighting the benefits of our approach.

## 1 Introduction

La thermographie laser "flying-spot" (FST) est une technique de contrôle non-destructif (CND) utilisant une source de chaleur locale active pour balayer la surface d'un échantillon métallique. Les discontinuités dans la diffusion de chaleur sur la surface, mesurées en spectre infrarouge, révèlent la présence de défauts de type fissure. Cette méthode d'inspection a été proposée à l'origine pour la détection de fissures sur pièces métalliques aéronautiques dans les années 1960 [1]. Cette technique est efficace pour caractériser des défauts difficiles à repérer, sans instrumentation lourde, telles que les micro-fissurations de faïençage [2]. Dans [3], Krapez définit le nombre de Pecllet, permettant de paramétrer les inspections afin de maximiser le signal obtenu pour un défaut, suivant la vitesse, le rayon de la tâche laser et la diffusivité thermique du matériau. Le nombre de Pecllet maximisant la réponse de la fissure tend vers 1. Une méthode de traitement est proposée, basée sur la soustraction de deux cartographies de balayage aller et retour traversant le défaut, puis par filtrage Laplacien. Cette méthode est illustrée dans la Figure 1 présentant d'une part une différence aller-retour, de l'autre la cartographie obtenue après filtrage, faisant ressortir le signal dû à la fissure. Cette méthode pose néanmoins des problèmes dans un contexte d'automatisation : elle se révèle coûteuse en temps d'inspection et de traitement, nécessitant un recalage délicat des deux cartographies avant soustraction, et des ajustements fins pour le filtrage. Enfin, les balayages considérés sont traversants, supposant un a priori sur la localisation et l'orientation du défaut.

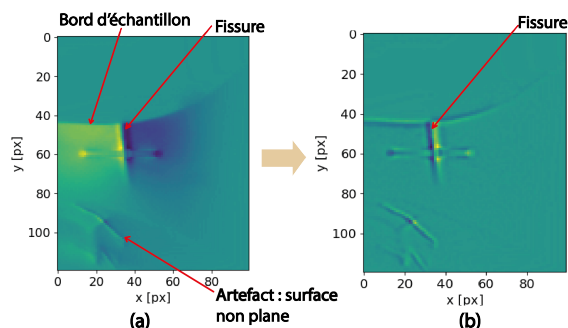


FIGURE 1 : (a) Cartographie FST aller-retour sans filtrage obtenue en suivant l'état de l'art [3] (b) Cartographie après filtrage Laplacien : si il subsiste des artefacts dus notamment à la non-planéité de l'échantillon, le signal dû à la fissure est clairement visible.

## 2 Positionnement et contributions

Nous proposons au contraire de travailler directement sur une seule cartographie aller, en mode dit "simple passe", sans soustraction et sans filtrage classique. Nous proposons aussi de travailler sur des balayages non-traversants, afin d'éliminer l'a priori opérateur sur le défaut, ainsi que de suivre la longueur de fissuration. Les échantillons étudiés ici sont des pièces en super-alliage revêtu : si l'amorce de fissuration est visible à l'oeil, le revêtement complique généralement la mesure de la pénétration du défaut dans le coeur du matériau. La figure 2 fournit un exemple de cartographie thermique exploitée pour cette étude, illustrant les différentes structures problématiques

pour un système de vision par ordinateur, comme les hétérogénéités matériau. Ces images sont obtenues par la moyenne et normalisation de l'ensemble des images d'un enregistrement thermique. La normalisation est effectuée directement sur les champs de température observé.

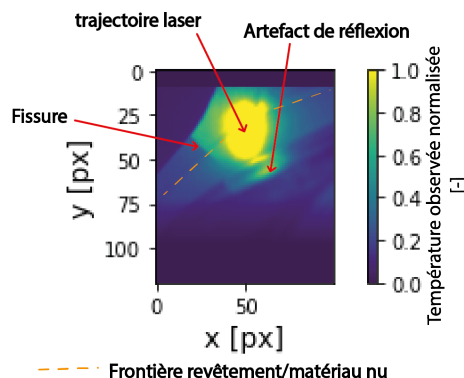


FIGURE 2 : Exemple de cartographie thermique en "simple-passe", illustrant les différentes structures interférant avec la détection du défaut.

**Revue de la littérature.** Des méthodes de traitement d'images par apprentissage profond pour des techniques d'inspections par thermographie laser sont déjà proposées dans la littérature, comme l'utilisation de réseaux récurrents pour extraire des caractéristiques temporelles des balayages thermiques [4]. Mais ces méthodes peuvent être coûteuses sur le plan calculatoire, et délicates à entraîner. De plus, la production d'un large volume de données pour la mise en application du *deep learning* associé au CND peut être difficile expérimentalement. Ainsi la **synthèse d'images** est une technique classique pour augmenter la quantité de données, par la simulation ou bien par l'utilisation de réseaux génératifs. Des architectures telles que les *generative adversarial neural networks* (GAN) ont été développés dans ce but [5]. Cependant ils se révèlent très ardues à entraîner. Les modèles de diffusion par débruitage apparaissent comme une alternative récente permettant de se passer d'un binôme discriminant-générateur pour générer des images synthétiques de qualité similaire [6]. Ces architectures présentent l'intérêt d'être faciles à entraîner, étant peu sensibles aux *mode collapses* typiques des GANs, et génériques. Ces réseaux fonctionnent en apprenant à convertir une distribution d'entrée en bruits Gaussiens. L'hypothèse de réversibilité du réseau permet ensuite de générer des données synthétiques à partir de bruits aléatoires. De son côté, le **curriculum learning** est inspiré de l'étude de l'apprentissage par la psychologie cognitive et comportementale. Il consiste en la mise en place d'un apprentissage progressif échelonné, partant de l'introduction de caractéristiques génériques simples pour aller vers des caractéristiques plus spécifiques et complexes jusqu'aux données d'application finales [7]. Cette approche est prometteuse en cas de données limitées a été utilisée en robotique et en traitement du langage. En l'état de nos connaissances celle-ci n'a pas été mise en oeuvre dans le cadre du CND par thermographie.

**Contributions.** Notre première contribution est donc la mise en place du banc d'essai de FST, puis la réalisation de données simulées et expérimentales. Nous proposons ensuite un protocole d'entraînement basé sur le *curriculum learning* faisant appel à la simulation de données et à l'utilisation des modèles de diffusion comme générateur d'images. Ce pro-

tole est mis en oeuvre pour la détection automatique de défauts de type fissures sur des échantillons métalliques par thermographie laser "simple-passe" : l'entraînement est réalisé sur nos jeux de données. Nous mettons en valeur les gains en performance de notre méthode, ainsi que les améliorations obtenues en termes de généralisation face à des échantillons inconnus.

### 3 Données

Cette section décrit la production de données pour la mise en place de notre protocole d'entraînement.

#### 3.1 Données expérimentales

Le banc de FST de l'ONERA est illustré Figure 3. Le laser travaille à une puissance fixée entre 0.5 et 3 W. La longueur d'onde est de 532 nm. Une lentille dichroïque permet de renvoyer la source de chaleur vers la surface de l'échantillon, qui est fixé sur un support motorisé permettant sa translation dans les 3 directions. La lentille laisse passer le rayonnement infrarouge vers la caméra infrarouge MWIR refroidie (3 à 5 microns), qui permet de suivre l'évolution de l'échauffement sur la surface observée. Ce choix de bande de travail permet d'étudier des échauffements relativement faibles. La théorie permet d'isoler des paramètres de travail optimisés (Peclet tendant vers 1), correspondant à un couple vitesse-rayon de tache laser donné [3]. Nous avons donc utilisé le banc pour générer deux bases de données à partir de multiples balayages dans la région autour du défaut : une base de données avec des réglages expérimentaux variés (base B), qui servira à l'entraînement d'un réseau pour la génération d'images synthétiques, et une base de test pour laquelle les paramètres expérimentaux sont optimaux, c'est-à-dire optimisés pour avoir un Peclet proche de 1 (base C), conformément à la théorie [3]. Des variations d'orientation du balayage sont aussi incorporées pour les deux bases. Les échantillons utilisés dans cette étude sont numérotés de 1 à 6, suivant une longueur de fissure croissante. Les échantillons 1,3,5 et 6 sont utilisés pour les entraînements (modèle de diffusion et réseaux de classification). Les échantillons 2 et 4 servent à vérifier les capacités de généralisation des réseaux à des échantillons inconnus, dont des examens n'ont pas été intégrés lors de l'entraînement initial. Le numéro 2 présente en particulier un marquage atypique du revêtement de surface.

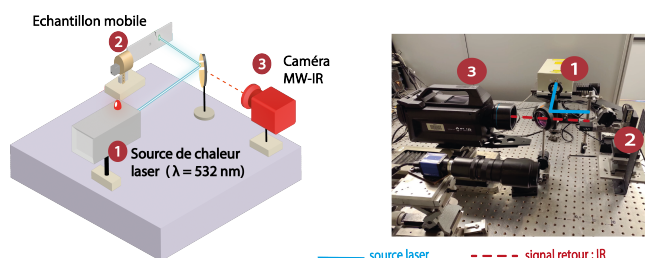


FIGURE 3 : Illustration du banc d'essai FST de l'Onera.

#### 3.2 Données simulées

Les données simulées ont été produites à l'aide du logiciel multi-physique éléments-finis Comsol. La simulation présente

la trajectoire de la source de chaleur sur la surface, sur laquelle se trouve ou non une fissure. Celle-ci est simulée par une résistance thermique linéique. Les simulations ne présentent pas de revêtement ou d'hétérogénéités en surface. Les paramètres d'inspection tels que la vitesse, le rayon de tâche laser et la puissance sont élargis autour du Peclet optimal afin d'augmenter la quantité d'images produites. Nous avons pu simuler une base de données (base A) contenant 13,860 images thermiques sans fissures et 15,115 avec fissures. La figure 4 (a) présente un exemple d'image simulée générée en suivant ce protocole.

### 3.3 Récapitulatif des bases de données

Le Tableau 1 fournit un résumé des différentes bases de données, avec leur origine et leur principale utilisation. Le tableau rappelle les échantillons employés, la quantité d'images thermiques et les conditions expérimentales utilisées. Pour chaque classe, nous avons entraîné un modèle de diffusion par débruitage simple [6], pour obtenir un générateur de données. Ils ont été entraînés grâce à la base B, et ce nouveau dataset est appelé Bb. Ils nous permettent de générer des données synthétiques à paramètres de travail élargis. La base C est scindée entre données d'entraînement et de test, avec un rapport 2/3.

TABLE 1 : Tableau résumant les différentes bases de données employées pour cette étude.

Base	Source	Nimage	Echantillons	Réglages
A	Simulation	28,975	-	Paramètres élargis
B	Experimental	600	1,3,5+6	Paramètres élargis
Bb	Modèle de diffusion	20.000	-	-
C	Experimental	330	1,3,5+6	Paramètres optimaux
D	Experimental	21	2,4	Paramètres optimaux

## 4 Méthode

### 4.1 Entraînement par curriculum learning

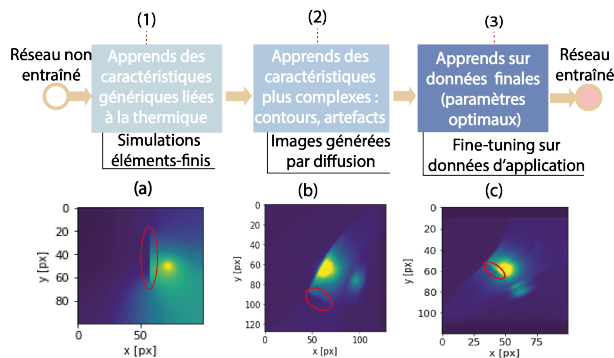


FIGURE 4 : Protocole d'entraînement proposé. Les images thermiques (a), (b) et (c) correspondent respectivement à une image simulée par éléments finis, à une image générée par le modèle de diffusion par débruitage, et à une image réelle finale. Le défaut est entouré en rouge dans chaque cas.

La figure 4 décrit le protocole d'entraînement proposé basé sur le *curriculum learning*, fournissant un exemple d'image d'entraînement pour chaque étape. Le protocole d'entraînement commence par un apprentissage de *features* génériques sur données simulées par éléments finis, liées aux aspects thermiques tels que la diffusion de chaleur sur la surface (étape 1). Ensuite le réseau est ré-entraîné sur des images synthétiques produites par nos modèles de diffusion (étape 2). Cette étape

nous permet de finaliser l'apprentissage en passant des caractéristiques génériques à leurs pendant plus spécifiques, liées à la diffusion de chaleur sur un matériau non-homogène, ou aux structures. Un *fine-tuning* est ensuite réalisé sur un échantillon limité de données réelles (étape 3). Des augmentations typiques telles que les inversions horizontales et verticales, ainsi que des rotations aléatoires sont appliquées pour l'entraînement de toutes les architectures. Pour toutes les architectures les poids sont initialisés aléatoirement.

### 4.2 Architectures et métriques

Nous testons diverses architectures d'apprentissage profond pour la classification binaire des patches avec ou sans défaut extraits directement des cartographie thermiques. D'une part nous évaluons un réseau convolutif de classification, VGG 13 [8]. Puis nous comparons avec des architectures basées sur le mécanisme d'attention, les *transformers*. La *self-attention* permet un apprentissage hiérarchisé associant les différentes régions de l'image d'entrée. Cette approche semble pertinente pour notre application présentant à la fois un éclairage local et des phénomènes multi-échelles, ici thermiques, comme indiqué dans différents travaux [9]. Nous retenons donc deux réseaux basés sur le mécanisme d'attention, les réseaux Swin et CaiT [10, 11]. Notre *baseline* issue de l'apprentissage machine traditionnel est constituée d'une architecture d'apprentissage machine traditionnelle, utilisant un filtre *Histogram of oriented gradient* (HOG) comme extracteur de caractéristiques, et un séparateur à vastes marges (SVM) pour réaliser la classification [12]. Elle fournit un repère de performances pour les méthodes classiques.

Pour l'ensemble des réseaux retenus, les métriques ont été évaluées sur le sous-échantillon de test de la base C. Nous avons évalué l'*accuracy*, le F1-score, la précision et enfin le rappel [13].

## 5 Résultats

### 5.1 Résultats généraux

Le Tableau 2 résume les résultats obtenus en test sur la base C, par les différentes architectures sélectionnées, pour un entraînement direct (base C, jeu d'entraînement) ou après entraînement par notre méthode de *curriculum learning*. Notre méthode permet des gains significatifs de performances pour toutes les architectures retenues, comparée à la méthode de *machine learning* ainsi que vis-à-vis de l'entraînement direct à partir de la base C seulement. D'un autre côté, les différences de performances entre VGG13 et les architectures basées sur l'attention illustrent l'intérêt de ce mécanisme pour notre application.

### 5.2 Capacités de généralisation

Les capacités de généralisation a des échantillons non-examinés pour la création des bases de données C ont été testées. Le tableau 3 illustre les performances obtenues pour les deux échantillons, en test seulement, pour deux architectures, en entraînement direct et par entraînement progressif. Nous constatons que l'approche proposée permet des gains de performance pour les deux architectures. La baisse significative de performance pour l'échantillon numéro 2 s'explique

Méthode	Architecture	Test-accuracy [%]	F1	Précision	Rappel
<i>Baseline</i>	HOG+SVM	83	0.83	0.82	0.83
Entraînement direct	VGG13	92	0.92	0.95	0.90
	Swin	87	0.86	0.92	0.86
	CaiT	90	0.91	0.93	0.89
<i>Curriculum learning</i>	VGG13	97	0.97	0.96	0.99
	Swin	96	0.96	0.94	0.98
	CaiT	99	0.98	0.98	0.99

TABLE 2 : Scores de classification obtenus par les architectures retenues en entraînement direct et grâce au protocole d’entraînement proposé par *curriculum learning*. (Base C : sous-échantillon de test).

Ech.	Nimage	Acc. (CaiT, direct)	Acc.(CaiT, curriculum)	Acc.(VGG13, direct)	Acc.(VGG13, curriculum)
2	11	58.8	82.3	41.2	76.5
4	10	90.0	95.8	77.8	90.0

TABLE 3 : Évaluation de l’*accuracy* sur les échantillons 2 et 4, non vus lors de l’entraînement.

par une surface atypique (marbrure due à un poinçonnage en surface), qui pourrait altérer l’inférence du réseau.

## 6 Conclusions et perspectives

Dans cet article, nous avons présenté une méthode de détection automatique de défaut dans les matériaux métalliques par FST, reposant sur des réseaux de détection entraînés par une approche de *curriculum learning*. Cette approche est mise en oeuvre à l’aide de données simulées, de données synthétiques par diffusion et de données expérimentales. Nous avons montré que notre méthode d’entraînement permettait des gains en termes de performances par rapport à un apprentissage direct. Il a aussi été relevé des gains de performance plus importants pour les architectures basées sur l’attention. Nous avons aussi montré que ce protocole d’entraînement augmentait les performances de généralisation à des échantillons inconnus.

Nous avons relevé plusieurs possibilités pour élargir ce travail. La comparaison entre entraînement direct et *curriculum learning* peut être évidemment étendue à d’autres architectures. Ensuite une comparaison avec d’autres modèles génératifs paraît pertinente, tel qu’un *variational auto-encoder* ou un GAN [14, 5]. D’autre part nous pourrions nous tourner vers des modèles de diffusion plus avancés [15]. L’approche d’entraînement proposée n’est pas limitée à la technique d’inspection par FST. Son application à d’autres techniques de CND, comme la thermographie flash, pourrait être mise en oeuvre à l’avenir.

## Références

- [1] E. J. KUBIAK : Infrared Detection of Fatigue Cracks and Other Near-Surface Defects. *Applied Optics*, 7(9), 1968. Publisher : Optical Society of America.
- [2] T. ARCHER *et al.* : Use of laser spot thermography for the non-destructive imaging of thermal fatigue microcracking of a coated ceramic matrix composite. *Quantitative InfraRed Thermography Journal*, 18, 2019.
- [3] J-C. KRAPEZ : Résolution spatiale de la camera thermique source volante. *International Journal of Thermal Sciences*, 1999.
- [4] W. SHI *et al.* : A technique combining laser spot thermography and neural network for surface crack detection in laser engineered net shaping. *Optics and Lasers in Engineering*, 138, 2021.
- [5] I. GOODFELLOW *et al.* : Generative Adversarial Nets. *In Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.
- [6] J. HO *et al.* : Denoising Diffusion Probabilistic Models, 2020. Number : arXiv :2006.11239 [cs, stat].
- [7] Y. BENGIO *et al.* : Curriculum learning. *In Proceedings of the 26th Annual International Conference on Machine Learning, ICML ’09*, New York, NY, USA, 2009. Association for Computing Machinery.
- [8] K. SIMONYAN *et al.* : Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv :1409.1556 [cs]*, 2015.
- [9] M. RAGHU *et al.* : Do vision transformers see like convolutional neural networks?, 2022. arXiv :2108.08810 [cs].
- [10] Z. LIU *et al.* : Swin Transformer : Hierarchical Vision Transformer using Shifted Windows. *In 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 2021. IEEE.
- [11] H. TOUVRON *et al.* : Going deeper with Image Transformers, 2021. Number : arXiv :2103.17239 [cs].
- [12] N. DALAL *et al.* : Histograms of Oriented Gradients for Human Detection. *In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, volume 1, San Diego, CA, USA, 2005. IEEE.
- [13] D. M. W. POWERS : Evaluation : from precision, recall and f-measure to roc, informedness, markedness and correlation. *ArXiv*, abs/2010.16061, 2011.
- [14] I. GULRAJANI *et al.* : PixelVAE : A Latent Variable Model for Natural Images, 2016. arXiv :1611.05013 [cs].
- [15] J. HO *et al.* : Cascaded Diffusion Models for High Fidelity Image Generation, 2021. arXiv :2106.15282 [cs].