



HAL
open science

Évaluation de méthodes de détection d'anomalies non supervisée pour l'auscultation des ouvrages d'art

Brice Marc, Philippe Foucher, Florence Forbes, Pierre Charbonnier

► To cite this version:

Brice Marc, Philippe Foucher, Florence Forbes, Pierre Charbonnier. Évaluation de méthodes de détection d'anomalies non supervisée pour l'auscultation des ouvrages d'art. RFIAP 2024 - Congrès Reconnaissance des Formes, Image, Apprentissage et Perception, AFRIF (Association Française pour la Reconnaissance et l'Interprétation des Formes), Jul 2024, Lille, France. pp.1-3. hal-04620019

HAL Id: hal-04620019

<https://hal.science/hal-04620019>

Submitted on 21 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Évaluation de méthodes de détection d'anomalies non supervisée pour l'auscultation des ouvrages d'art

B. Marc^{1,2}

P. Foucher¹

F. Forbes²

P. Charbonnier¹

¹ Équipe de recherche ENDSUM, Cerema, Strasbourg

² Équipe-projet STATIFY, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble

brice.marc@cerema.fr

Résumé

Les méthodes automatiques utilisées pour l'auscultation des ouvrages d'art sont, pour la plupart, supervisées et donc, coûteuses en termes d'acquisition et d'annotation d'anomalies. Les méthodes de détection d'anomalies non supervisée offrent une alternative à ces méthodes grâce à des modèles n'apprenant qu'avec des images saines. Afin d'étudier les performances de ces méthodes appliquées au cas des ouvrages d'art, nous avons acquis des images de parois d'un tunnel routier pour constituer un jeu de données. Nous avons ensuite déployé et testé sur ces données cinq modèles de l'état de l'art, qui mettent en œuvre chacun des principaux paradigmes de méthodes non supervisées. Sur ces expériences, le modèle fondé sur les *normalizing flow* multi-échelles (MSFlow) s'avère le plus performant pour relever les anomalies sur les parois du tunnel.

1 Introduction

L'inspection des ouvrages d'art vise à garantir la sécurité et la durabilité des infrastructures. Indispensable à l'auscultation, le relevé des anomalies est une tâche fastidieuse, qu'il est souhaitable d'automatiser. De nombreux travaux de recherche se fondent sur les méthodes d'apprentissage profond pour effectuer ces relevés, mais la très grande majorité des approches, supervisées, nécessitent la constitution de bases d'images de défauts. Cela demeure un défi important, aussi bien en termes de coût de production des annotations qu'en termes de représentativité des désordres, qui peuvent prendre des apparences très variables. Face à ces difficultés, les méthodes de détection non supervisée constituent une alternative intéressante.

L'objectif de cette étude exploratoire est de mettre en œuvre différentes méthodes non supervisées pour la détection d'anomalies sur ouvrages d'art. Pour cela, nous nous appuyons sur les familles d'architectures décrites dans [1], à savoir les méthodes par : Représentation ; *Normalizing Flows* (NF) ; Reconstruction dont distillation de connaissance (KD) ; Augmentation de données (AD) et nous les appliquons à des images acquises dans un tunnel en service. Les méthodes par représentation visent à extraire les caractéristiques (*features*) des images saines afin de définir leur distribution. Lors des tests, sont catégorisées comme anomalies les *features* s'éloignant de ces distributions. L'approche NF est également une méthode de représentation apprenant à transformer la distribution des images saines vers une distribution normale. Lors de l'inférence, les anomalies se situant hors de la distribution peuvent ainsi être identifiées. Les méthodes par reconstruction apprennent à reconstruire des images saines, éventuellement bruitées ou en partie masquées, après une projection dans un espace latent. Avec ces modèles, il est attendu une mauvaise reconstruction des défauts, permettant de les détecter en comparant l'image d'entrée et sa reconstruction. Les méthodes par KD se basent sur un duo de réseaux dits « enseignant » et « élève ». Le premier extrait des *features* que le deuxième doit apprendre à reconstruire. Les différences entre les *features* des deux réseaux permettent de détecter les anomalies. Enfin, les méthodes avec AD sont des méthodes auto-supervisées qui ajoutent des anomalies synthétiques aux images saines afin de constituer les vérités terrains pour l'apprentissage. Les modèles testés ont obtenu d'excellentes performances sur la base d'anomalies MVtec [2]. Les mêmes procédés d'apprentissage sont appliqués en repartant de zéro avec les images d'ouvrages. Notons que les images de cette base sont acquises en milieu contrôlé, alors que nous les appliquons ici à des images acquises *in situ*.

2 Méthodes et Expérimentations

Afin de mesurer l'applicabilité des différents modèles au cas concret de l'auscultation des ouvrages d'art, une campagne d'acquisition de données a été menée. Nous avons pris, au moyen d'un appareil photographique *Canon EOS 5D Mark III* des

images, de 5760×3840 pixels, des deux parois d'un tunnel routier. Après rognage pour ne conserver que la paroi du tunnel, les images ont été divisées verticalement en quatre sous-images, puis redimensionnées en 512×512 pixels et annotées manuellement. Au total, le lot d'entraînement compte 107 images saines, et le lot de test contient 118 images dont 4 saines, ce qui reste du même ordre de grandeur que les jeux de la base MVTEC de chaque catégorie, qui comptent généralement environ 250 images pour l'entraînement et une centaine d'images pour les tests.

L'approche PNI [3] est une méthode de représentation où la probabilité, à partir de laquelle la vraisemblance est calculée, est conditionnelle aux informations de voisinage et de position dans l'espace des *features*. MSFlow [4] est un modèle NF visant à bénéficier d'une information multi-échelle. Pour cela, l'architecture applique des modules NF parallèles asymétriques à des *features* issues des différentes couches d'un extracteur pré-entraîné. La méthode DDAD [5] reconstruit, par un procédé de diffusion, une image préalablement bruitée. Dans cette approche, les images originale et reconstruite sont comparées à la fois au niveau des pixels et des *features*. Dans les méthodes KD standards, le réseau « élève » a tendance à copier correctement le réseau « enseignant », y compris en présence d'anomalies. Pour éviter cet écueil, le modèle RD [6] propose une structure inversée avec un encodage des *features* par l'« enseignant » et un décodage par l'« élève ». MemSeg [7] est un modèle auto-supervisé utilisant des anomalies synthétiques générées par bruit de Perlin. Il synthétise des anomalies texturées en collant des parties d'images suivant du bruit de Perlin avec d'autres textures ou des anomalies structurelles en déplaçant des portions de l'image.

Les différents modèles ont été entraînés en utilisant les paramètres proposés par leurs auteurs respectifs et les répertoires Github indiqués dans les références. Les procédés d'entraînement ont été appliqués sur les données du tunnel en repartant de zéro.

Les scores AUROC et AUPRO sont utilisés pour évaluer les modèles à l'échelle du pixel. Le premier est l'aire sous la courbe du taux de vrais positifs en fonction du taux de faux négatifs obtenus avec différents seuils de binarisation des scores de probabilité. Le second est similaire à AUROC mais prend en compte l'information de connexité des anomalies. Cela permet de considérer de la même manière les petites et grandes anomalies.

Les modèles MSFlow et RD sont les plus performants, avec un léger avantage pour le premier. Les scores sont ainsi égaux à 92,22 % (MSFlow) et 91,57 % (RD) pour AUROC, et à 88,52 % (MSFlow) et 88,38 % (RD) pour AUPRO. La méthode PNI est moins performante sur les petites anomalies, comme nous le montre la différence de score entre AUROC (91,51 %) et AUPRO (80,20 %). Ce dernier est 8 % inférieur aux scores AUPRO des modèles MSFlow et RD. Enfin, les performances des modèles DDAD et MemSeg sont significativement inférieures aux autres modèles avec des scores AUROC de 71,89 % pour DDAD et 79,26 % pour MemSeg. L'écart de score AUPRO est plus net : 71,07 % pour DDAD et 68,81 % pour MemSeg.

La figure 1, présente sur la page 3, illustre les performances des modèles. On observe que tous les modèles arrivent à détecter la majorité de l'épaufrure présente sur l'image de la 1^{re} ligne. MemSeg et DDAD n'en segmentent toutefois qu'une partie. Les petites fissures à droite de l'image sont grossièrement repérées par les modèles. Sur l'image de la 2^e ligne, la fissure verticale est plus ou moins nettement détectée par tous les modèles, à l'exception de MemSeg, qui n'en relève qu'une faible partie. MSFlow est le seul modèle à localiser tous les impacts en haut de l'image, mais considère aussi à tort certaines zones comme des anomalies. RD et PNI relèvent seulement légèrement deux impacts tandis que DDAD et MemSeg n'en décèlent aucun. Enfin, sur l'image de la 3^e ligne, aucun modèle n'arrive à distinguer la fissure fine en bas à gauche. Malgré des exemples similaires présents dans le lot d'apprentissage, tous les modèles sauf DDAD relèvent le graffiti en haut de l'image ou les deux brides de serrage de cauffrage en bas comme des anomalies.

3 Conclusion

Dans cette contribution, nous avons entraîné et évalué 5 modèles d'apprentissage non supervisé pour détecter des anomalies sur un jeu restreint d'images de parois d'un tunnel routier, acquises *in-situ*. MSFlow est la méthode enregistrant les meilleurs résultats, suivie de près par les modèles RD et PNI, qui ne localisent toutefois pas certains défauts. Les méthodes DDAD et MemSeg sont bien moins performantes. Les méthodes par reconstruction ont du mal à reconstruire parfaitement les parties saines des images complexes sans reconstruire les anomalies, ce qui explique les faibles résultats du modèle DDAD. La méthode MemSeg est, ici, pénalisée par le manque de réalisme des anomalies synthétiques qu'elle génère pour l'apprentissage.

Même si le modèle MSFlow permet de relever la plupart des défauts, les détections comportent de nombreux faux positifs, liés à des éléments rares présents dans le tunnel. Réduire le nombre de faux positifs semble donc être une piste intéressante pour améliorer les performances des méthodes de détection d'anomalies appliquée à l'auscultation des ouvrages d'art.

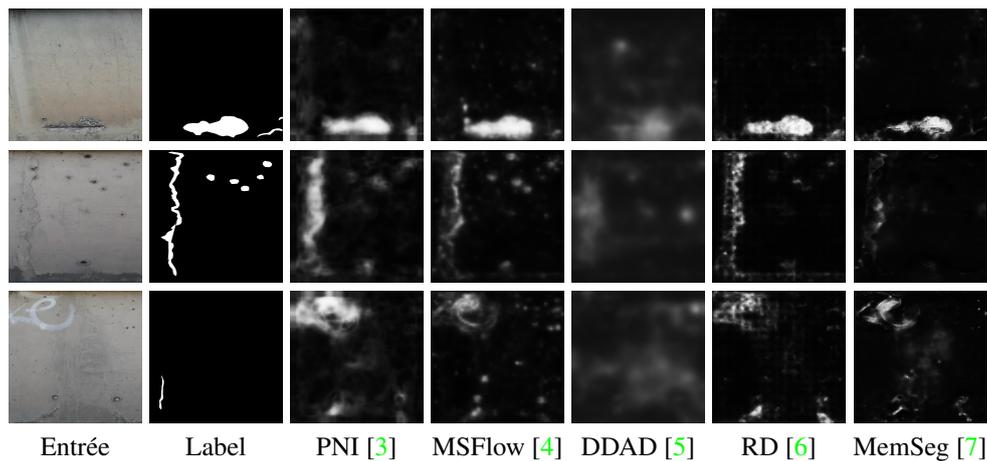


FIGURE 1 – Exemples de résultats obtenus pour chaque modèle

Références

- [1] Y. Cui, Z. Liu, and S. Lian. A Survey on Unsupervised Anomaly Detection Algorithms for Industrial Images. *IEEE Access*, 2023.
- [2] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger. MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. In *IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, 2019.
- [3] J. Bae, J. Lee, and S. Kim. PNI : Industrial anomaly detection using position and neighborhood information. In *Proc. of the IEEE/CVF International Conf. on Computer Vision*, 2023. https://github.com/wogur110/PNI_anomaly_detection.
- [4] Y. Zhou, X. Xu, J. Song, F. Shen, and HT. Shen. MSFlow : Multiscale Flow-Based Framework for Unsupervised Anomaly Detection. *IEEE Transactions on Neural Networks and Learning Systems*, 2024. <https://github.com/cool-xuan/msflow>.
- [5] A Mousakhan, T Brox, and J Tayyub. Anomaly Detection With Conditioned Denoising Diffusion Models. *arXiv preprint arXiv :2305.15956*, 2023. <https://github.com/arimousa/DDAD>.
- [6] H. Deng and X. Li. Anomaly detection via reverse distillation from one-class embedding. In *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, 2022. <https://github.com/hq-deng/RD4AD>.
- [7] M. Yang, P. Wu, and H. Feng. MemSeg : A semi-supervised method for image surface defect detection using differences and commonalities. *Engineering Applications of Artificial Intelligence*, 2023. <https://github.com/TooTouch/MemSeg>.