

Automatisation de l'inspection de conteneurs maritimes : une preuve de concept

S. Tabti¹, F. Vallée¹, D. Poon Shu Han¹, M. Elion^{1*}, Y. Alouini¹

¹ Fieldbox, Quai Armand Lalande, Bordeaux, 33300

stabti@fieldbox.ai

Résumé

Cet article présente une preuve de concept d'automatisation de l'inspection de conteneurs maritimes développée pour un client en Asie. Un modèle de détection de défauts (Yolov5) est entraîné sur des vidéos d'intérieur de conteneurs collectées grâce à un rover. Ce projet couvre toutes les étapes de développement d'un modèle d'IA, de l'exploration d'un gros volume de données historiques, la définition d'un système d'acquisition de données adapté à une automatisation de l'inspection, à l'annotation des images et l'entraînement du modèle sur un jeu de données frugal.

Mots-clés

Automatisation, inspection de conteneurs, détection de défauts, Yolov5, transfer learning.

Abstract

This article presents a proof of concept developed for a client in Asia. It deals with automated visual inspection of shipping containers thanks to a deep learning model (Yolov5) trained to detect defects inside containers using videos collected by a rover. This project covered all the steps to build an AI model, from historic data exploration on a big set of pictures taken by container surveyors, defining a new data collection process adapted to future inspection automation, to image labeling and model training on a small dataset.

Keywords

Automation, container inspection, defect detection, Yolov5, transfer learning.

1 Introduction

L'inspection de conteneurs maritimes est un défi logistique majeur [5]. En particulier, inspecter d'éventuels défauts sur les conteneurs est important pour leur maintenance et la prévention de dommages plus graves. A titre d'exemple, à Singapour, 500 conteneurs sont inspectés chaque jour, dont 30% comportent des défauts. L'inspection d'un conteneur peut prendre entre 5min (défauts mineurs) à 40min (défauts majeurs). Automatiser ce procédé grâce à des modèles de vision par ordinateur serait bénéfique, néanmoins cela est

*Ce travail a été réalisé sur une période où M. Elion faisait partie de Fieldbox avant de rejoindre une autre entreprise.

complexe à plusieurs niveaux. Nous abordons dans cet article deux principaux verrous.

Le premier verrou consiste en l'absence d'un système de collecte de données adapté à cette automatisation. En effet, chez notre client, le procédé actuel implique qu'un camion porte-conteneur se gare dans une zone dédiée, qu'un inspecteur analyse l'extérieur, l'intérieur vide et les informations inscrites sur le conteneur. Puis, des photos des dommages observés sont prises à l'aide d'une tablette avec un point de vue très proche du défaut. Ce procédé est inadapté si on vise l'automatisation de la prise de vue et de la détection des défauts car les points de vues des photos sont trop hétérogènes et dépendent de l'inspecteur. Par la suite, l'inspecteur renseigne la localisation des défauts, leurs catégories, et émet si nécessaire des recommandations de nettoyage et de réparation sur une application.

Un second verrou concerne la variabilité des types de défauts à identifier qui sont plus d'une centaine et soumis à la norme [ISO 9897](#) General Container Standard. Ceci implique plusieurs points. Premièrement, la constitution d'une base de données intégrant cette variété va demander un temps considérable. Ensuite, le niveau d'expertise requis pour l'annotation des images est tel qu'il sera nécessaire que des inspecteurs la prennent en charge, ce qui entraîne un investissement humain, en temps et financier. Enfin, plusieurs modèles d'Intelligence Artificielle (IA) seront nécessaires pour localiser et classer les différents types de défauts. En effet, pour certains défauts un modèle de détection d'objets est pertinent mais pour d'autres, comme l'évaluation de la lisibilité d'informations inscrites sur les conteneurs, il faudrait faire de la reconnaissance de caractères puis des post-traitements.

Le développement d'une solution d'automatisation d'une telle envergure pour un client commence traditionnellement par un premier projet avec un périmètre réduit dit "preuve de concept" (ou POC pour Proof Of Concept). L'objet de cet article est de présenter ce POC mené entre fin 2021 et mi-2022. Nous commençons par un bref état de l'art en inspection de conteneurs en section 2. Puis, grâce à une exploration de données historiques, nous définissons une liste de défauts prioritaires pour ce POC en section 3. La section

4 décrit le système d’acquisition de données qui permettra à l’avenir l’automatisation de l’inspection de l’intérieur et de l’extérieur des conteneurs. En particulier, nous évoquons le choix des équipements, la campagne de collecte de données d’entraînement et leur annotation. Les images de défauts collectées étant trop peu nombreuses, le périmètre du projet a été réduit. En conséquence, le modèle de détection de défauts (localisation et classification) entraîné est dédié uniquement à l’intérieur des conteneurs et décrit en section 5. Pour conclure, nos retours d’expériences et perspectives sont partagés en section 6.

2 Bref état de l’art en inspection de conteneurs

Plusieurs solutions commerciales et travaux de recherche existent afin d’automatiser une partie du procédé d’inspection de conteneurs et ceci sur différents types de données. Par exemple, l’entreprise [MEP](#) propose un système où un camion chargé d’un conteneur passe dans un scanner à rayons X. Aucun modèle d’IA n’apparaît dans leur documentation. En revanche, plusieurs articles de recherche développent des modèles de classification pour déterminer si les conteneurs contiennent des armes ou des objets de contre-bande, ex : [1]. L’entreprise [LASE](#), quant à elle, a proposé un système basé sur 3 capteurs lasers permettant d’obtenir un profil 3D d’un conteneur qui est analysé afin de détecter des défauts de surface.

En ce qui concerne l’analyse d’images, plusieurs solutions commerciales existent, ex : [VISY](#) et [ProjAIX](#). À notre connaissance, ces dernières se focalisent sur la reconnaissance des caractères pour récupérer les informations inscrites sur les conteneurs et sur la détection de quelques défauts extérieurs. Cependant, elles ne permettent pas de classifier les défauts détectés. Plusieurs travaux de recherche, ex : [2], sont développés pour la détection de corrosion sur les conteneurs. D’autres proposent des modèles de classification d’un nombre limité de défauts. Par exemple, le modèle proposé dans [6] est entraîné sur une base de données de photographies de conteneurs avec des prises de vues très hétérogènes.

Le jeu de données mentionné dans [6] risque d’être peu adapté pour l’entraînement d’un modèle d’IA qui, une fois déployé, serait appliqué à des flux vidéos avec des points de vues très différents et potentiellement loin des défauts. Or c’est l’objectif que nous nous sommes fixé pour automatiser la prise de vue des images aussi bien que la détection de défauts. De plus, après la campagne de collecte de données, le périmètre de ce POC va se concentrer sur l’intérieur des conteneurs, peu traité dans la littérature. La localisation des défauts dans l’image sera garantie par un modèle de détection d’objets traçant des rectangles (bounding boxes) autour des défauts plutôt qu’un modèle de classification qui ne prédit qu’un label sans localisation.

Catégorie	Porte	Extérieur	Intérieur	Inscriptions	Nettoyage
#types de défauts	49	27	20	18	2
#images	52164	49045	66680	7176	11412

TABLE 1 – Résumé des données historiques transmises par le client. Plus de cent types de défauts ont été recensés et plus de 180K images partagées. On rappelle que ces données sont inadaptées à l’automatisation proposée et donc inutilisées pour l’entraînement des futurs modèles.

3 Exploration et analyse des données historiques

Plus d’une centaine de types de défauts de conteneurs existent et sont répartis entre les catégories suivantes : extérieur, intérieur, inscriptions et nettoyage. Afin de définir une liste de défauts sur lesquels se concentrer pour construire un premier modèle d’IA pour ce POC, une phase d’exploration et d’analyse de données historiques est menée. L’extrait d’archives de données de 2019 et 2020 transmis par le client comporte : 1/ des fichiers JSON avec pour chaque type de défaut les noms de fichiers des photos prises par les inspecteurs, 2/ les photos associées (plus de 90 Gigaoctets d’images). Analyser les fichiers JSON a permis d’avoir une vision globale de ces données résumée en tableau 1 et de calculer pour chaque catégorie le nombre de photos collectées en 2019 et en 2020 pour chaque type de défaut.

Il est important de séparer l’exploration des données datant de 2019 et de 2020 pour éviter un biais d’analyse dû au COVID qui a eu un fort impact sur les activités commerciales en Asie, et ceci même jusqu’à la fin du POC en 2022. L’analyse de ces graphes montre également que d’une année à une autre, ce ne sont pas nécessairement les mêmes types de défauts qui sont les plus récurrents. De plus, il y a un fort déséquilibre dans la répartition des défauts, ce qui doit être un point d’attention pour la modélisation afin d’éviter de favoriser uniquement les types de défauts majoritaires. La figure 1 illustre nos observations dans le cas de la distribution des principaux types de défauts d’intérieur de conteneurs et leur nettoyage.

L’analyse de ces graphes pour chaque catégorie (extérieur, intérieur, ... etc.) permet de se concentrer sur l’exploration d’images comportant les types de défauts au nombre d’occurrences les plus élevés en 2019, pour éviter le biais du COVID, et de ne conserver qu’une vingtaine de types de défauts importants aux yeux du client et dont on attend que la visibilité soit suffisante sur les données qui seront collectées avec le système d’acquisition décrit en section 4.

4 Système d’acquisition des données, collecte et labellisation

Pour la campagne de collecte de données menée entre Février et Mars 2022 et un potentiel déploiement futur de nos modèles d’IA sur des flux de données acquis automatique-

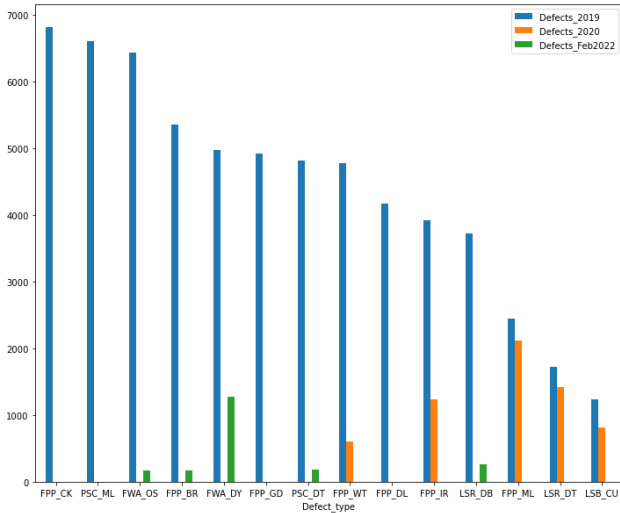


FIGURE 1 – Distribution des types de défauts intérieurs de conteneurs. Les années 2019 et 2020 représentent les données historiques avec des photos prises par des inspecteurs. L’année 2022 représente la campagne de collecte de données sur le mois de Février servant à l’entraînement du modèle. On note un déséquilibre dans la répartition des types de défauts, des défauts majoritaires différents d’une année à l’autre, un fort impact du COVID en 2020 et un faible volume de données collecté en 2022.

ment, nous avons mis en place un système d’acquisition de vidéos grâce à un ensemble de caméras de surveillance pour l’inspection de l’extérieur des conteneurs. Le modèle de caméra a été sélectionné avec soin en fonction de ses propriétés (résolution de 2K, Field Of View, auto-focus, ...etc.). Nous avons opté pour la marque Hikvision. La marque Axis est une très bonne option également. La configuration d’installation des caméras proposée en figure 2 garantit la visibilité des défauts avec peu de distorsions, tout en s’adaptant à la structure du site industriel. Les vidéos sont stockées sur un équipement de type NAS (Network Attached Storage), puis synchronisées à intervalles réguliers vers un espace de stockage cloud (type bucket) accessible à distance pour effectuer des tâches d’analyse et de modélisation.

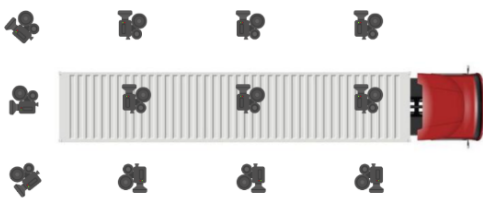


FIGURE 2 – Schéma d’installation des caméras pour l’inspection de l’extérieur des conteneurs (vue du dessus d’un camion remorque porte-conteneur).

Concernant l’équipement permettant d’acquérir semi-automatiquement des vidéos à l’intérieur des conteneurs, nous avons sélectionné pour ce POC le rover (drone rou-

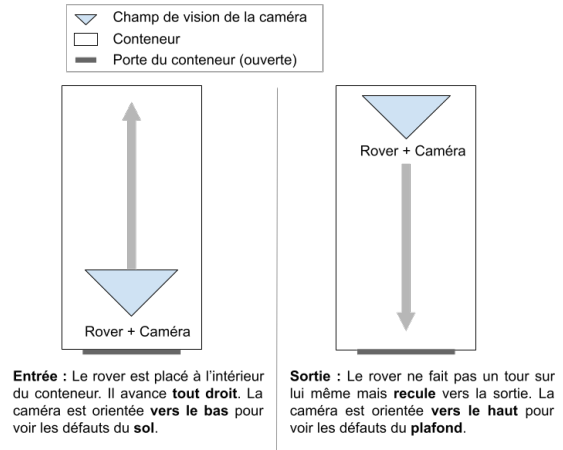


FIGURE 3 – Recommandations de pilotage du rover pour maximiser la qualité et la visibilité des défauts collectés lors d’inspections de l’intérieur des conteneurs. À gauche : préconisations lorsque le rover est placé dans le conteneur et qu’il se déplace vers le fond de celui-ci. À droite : préconisations pour que le rover sorte sans être ébloui par la lumière extérieure.

lant) DJI RoboMaster S1 d’une résolution de 2K. Le mode opératoire que nous avons mis en place est le suivant : un inspecteur ouvre la porte du conteneur, installe le rover à l’intérieur, puis le pilote en suivant le chemin que nous avons préconisé afin de maximiser la visibilité de tous les types de défauts. Pour cela, le rover avance tout droit en orientant la caméra vers le bas pour observer les défauts sur le plancher et en bas du conteneur. Arrivé au bout, il ne fait pas un tour sur lui même afin d’éviter un contre-jour provoqué par la lumière de l’extérieur. Il fait plutôt une marche arrière en pointant la caméra vers le haut pour observer les défauts sur le plafond et en haut du conteneur. Ce procédé est schématisé en figure 3. La carte SD du rover est régulièrement extraite par l’inspecteur pour télécharger les données sur un espace de stockage dans le cloud.

Au fur et à mesure de la phase de collecte de vidéos, les images les plus pertinentes pour la modélisation sont extraites et labellisées par des inspecteurs de conteneurs, vu le niveau d’expertise élevé requis pour cette tâche. Pour cela, l’outil open-source [Label Studio](#) leur a été mis à disposition et nous les avons formés à son utilisation. L’histogramme vert en fig.1 montre la distribution des défauts les plus récurrents collectés. Les activités commerciales en Asie étant toujours impactées par le COVID en 2022, un nombre insuffisant d’images pour la majorité des types de défauts furent effectivement collectées. Les défauts majoritaires concernant l’intérieur des conteneurs et leur nettoyage, nous avons réduit le périmètre du POC à l’entraînement d’un modèle de détection de 5 types de défauts sur les images du rover décrits en tableau 2.

Défaut	Description	TP
FWA-DY	Plancher sale	0.72
LSR-DB	Débris sur anneau d'arrimage	0.18
PSC-DT	Tôle froissée	0.11
FPP-BR	Plancher abîmé	-
FWA-OS	Tâche d'huile sur le plancher	0.25

TABLE 2 – Description des défauts d'intérieur de conteneurs sur lequel Yolov5 a été entraîné et performances sur des images de test (TP : True Positives).

5 Détection de défauts avec Yolov5

Pour la détection de défauts à l'intérieur de conteneurs, le modèle **Yolov5 (medium)** développé par Ultralytics et préentraîné sur le dataset COCO a été utilisé. Alors que peu de travaux dédiés à la détection de défauts à l'intérieur de conteneurs sont proposés dans la littérature et que la majorité proposent des modèles de classification, notre choix s'est porté sur ce modèle de détection d'objets car la localisation des défauts dans les images et leur catégorisation est importante pour nos clients. De plus, l'inférence des modèles de type Yolo (ex : Yolov4 [3]) est très rapide, en particulier si on dispose d'une puissance de calcul adaptée. Un transfer learning a été appliqué pour spécialiser le modèle préentraîné aux données du client. Pour cela, les images acquises en Février 2022 ont été utilisées. On rappelle que le nombre total d'images pour chaque type de défauts est représenté en fig.1, en vert. Des images sans défauts (ou : background images) sont ajoutées à ce dataset à hauteur de 10% du nombre total d'images tel qu'il est **recommandé** par Ultralytics afin de réduire les faux positifs.

Le data splitting entre entraînement/validation ($\sim 80\%$ des images) et test ($\sim 20\%$ des images) a été réalisé de façon à éviter toute fuite de données menant à du sur-apprentissage (ex : ne pas intégrer des images extraites de la même vidéo dans le même split) et à maintenir la même distribution des types de défauts dans chaque split. De plus, pour pallier la frugalité et le déséquilibre du jeu de données (cf. fig. 1), nous avons généré 500 images par des opérations de base d'augmentation de données d'entraînement (flip horizontal, flip vertical, translations, changements de contraste). Au final, après nettoyage et augmentation, le jeu d'entraînement et de validation est composé d'un peu plus de 1600 images, ce qui est très faible. Le jeu de test est composé d'approximativement 300 images. L'optimiseur sélectionné est une descente de gradient stochastique, la taille des batchs est de 8 et le paramètre de label smoothing est fixé à 0.15 afin de limiter l'impact d'erreurs d'annotation par les inspecteurs. Le nombre d'epochs maximal est limité à 100.

Des métriques de performances du modèle sur des images de test sont présentées en dernière colonne du tableau 2. Il s'agit de TP (True Positives) par type de défaut normalisé par le nombre d'instances (nombre de boxes) correspondantes à ce type de défaut. Ces résultats sont pour



FIGURE 4 – Détections réussies sur des vidéos de test.

la plupart faibles et paraissent corrélés au nombre d'images collectées par type de défaut. En effet, le nombre d'images est trop limité pour espérer des performances élevées (Ultralytics **recommande** au moins 1500 images par classe, ce dont nous sommes loin). Il faudrait que la campagne de collecte de données soit reconduite pour plus longtemps. Cependant, on rappelle que ce score dépend du seuil sur le degré de confiance pour la sélection de boxes candidates à la Non-Max Suppression et du seuil sur l'IoU (Intersection over Union). Le seuil par défaut de l'IoU à 0.45 n'est pas assez permissif, ce qui a pour conséquence d'observer aucun TP de FPP-BR. Un travail futur pour pousser plus loin l'analyse des résultats consisterait à développer un post-traitement permettant d'optimiser ces seuils. Ceci est encouragé par des détections prometteuses obtenues sur des vidéos de test, cf. quelques exemples en figure 4. On précise que lorsque le modèle doute, la classe background est prédite plutôt que d'autres classes.

En fin de projet, des tests en conditions réelles réalisés sur site sur des vidéos inédites ont satisfait le client. En revanche, ces tests nous ont permis d'observer un sur-apprentissage sur la classe la plus récurrente : FWA-DY. En effet, le modèle prédit trop souvent cette classe. Après investigation auprès des inspecteurs, nous avons réalisé que les critères de décision sur ce défaut n'étaient pas assez bien formalisés et variaient d'un inspecteur à un autre, ce qui a biaisé le modèle pour prédire plus souvent qu'il ne faut que le plancher est sale. Encourager les annotateurs à avoir des critères de décisions clairs et avoir un processus de révision des labels par un deuxième annotateur est une bonne pratique pour limiter ce type de biais.

6 Conclusion et perspectives

Cet article décrit un projet concret d'intelligence artificielle mené pour un client en Asie dans un temps limité. L'objectif est de développer une preuve de concept d'automatisation d'inspection de conteneurs maritimes. Pour cela, un système de prise de vue quasi-automatique de conteneurs chargés sur des camions a été mis en place. Un réseau de caméra permet l'inspection de l'extérieur, tandis qu'un rover permet l'inspection de l'intérieur. La

campagne de collecte et d'annotation d'images nous a permis de constituer un jeu d'entraînement. Cependant, ce dernier étant d'un volume trop faible, le périmètre du projet a été réduit à la détection de 5 types de défauts à l'intérieur des conteneurs. Les performances du modèle Yolov5 sont améliorables mais encourageantes vu les résultats obtenus sur des vidéos de test.

Pour une prochaine phase commerciale, une campagne de collecte de données plus longue et une labellisation de meilleure qualité auront un impact positif sur les résultats. Cependant, le chemin est long pour l'automatisation de l'inspection de conteneurs. Plusieurs travaux sont envisagés, ex : d'autres types de défauts (intérieurs et extérieurs) doivent être intégrés, des post-traitements doivent être mis en place, les modèles développés doivent être déployés sur site et une application métier doit être développée pour faciliter le travail quotidien des inspecteurs. De plus, en fonction des temps d'inférence requis pour la détection de défauts, d'autres modèles peuvent être expérimentés. L'implémentation de Yolov5 par Ultralytics est satisfaisante pour ce POC mais d'autres implémentations avec des licences permettant une meilleure protection du code produit existent, ex : des transformers pour une détection d'objets interprétable [4] sont une bonne option.

Remerciements

C. Collignon, R. Yang, K. Vaibhav, J. Budynek, A. Coutant, A. Cherifi, P. Oreistein, P. Chin, J. Sauer.

Références

- [1] M. Abdolshah, M. Teimouri, and R. Rahmani. Classification of X-ray images of shipping containers. *Expert Systems with Applications*, 77 :57–65, 2017.
- [2] Z. Bahrami, R. Zhang, T. Wang, and Z. Liu. An end-to-end framework for shipping container corrosion defect inspection. *IEEE Transactions on Instrumentation and Measurement*, 71 :1–14, 2022.
- [3] A. Bochkovskiy, C-Y. Wang, and H-Y. M. Liao. Yolov4 : Optimal speed and accuracy of object detection. *arXiv preprint arXiv :2004.10934*, 2020.
- [4] H. Chefer, S. Gur, and L. Wolf. Generic attention-model explainability for interpreting bi-modal and encoder-decoder transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 397–406, October 2021.
- [5] N. Hoffmann, R. Stahlbock, and S. Voß. A decision model on the repair and maintenance of shipping containers. *Journal of Shipping and trade*, 5 :1–21, 2020.
- [6] Z. Wang, J. Gao, Q. Zeng, and Y. Sun. Multitype damage detection of container using CNN based on transfer learning. *Mathematical Problems in Engineering*, 2021 :1–12, 2021.