



HAL
open science

Génération photo-réaliste de défauts sur des images de surfaces extérieures d'avions acquises par un drone autonome

Bryan Deliencourt, Mathieu Giraud, Julien Miranda, Stanislas Larnier, Ariane Herbulot, Michel Devy

► To cite this version:

Bryan Deliencourt, Mathieu Giraud, Julien Miranda, Stanislas Larnier, Ariane Herbulot, et al.. Génération photo-réaliste de défauts sur des images de surfaces extérieures d'avions acquises par un drone autonome. ORASIS, journées francophones des jeunes chercheurs en vision par ordinateur, May 2019, Saint-Dié-des-Vosges, France. hal-02164953

HAL Id: hal-02164953

<https://hal.science/hal-02164953v1>

Submitted on 25 Jun 2019

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.

Génération photo-réaliste de défauts sur des images de surfaces extérieures d'avions acquises par un drone autonome

Bryan Deliencourt¹
Stanislas Larnier¹

Mathieu Giraud¹
Ariane Herbulot^{2,3}

Julien Miranda^{1,2,3}
Michel Devy^{2,3}

¹ Donecle, 201 Rue Pierre et Marie Curie, F-31670 Labège, France

² LAAS, CNRS, 7 avenue du colonel Roche, F-31400 Toulouse, France

³ Univ de Toulouse, UPS, LAAS, F-31400 Toulouse, France

(bryan.deliencourt, mathieu.giraud, julien.miranda, stanislas.larnier)@donecle.com

Résumé

Ce travail porte sur l'extension de bases de données constituées d'images de surfaces extérieures d'avion prises par des drones autonomes. Ces données sont utilisées pour entraîner et tester des algorithmes d'inspection visuelle automatique. L'efficacité de l'apprentissage dépend de la représentativité et de la quantité d'exemples. Les bases de données disponibles ne comportent qu'une faible quantité d'éléments considérés comme des défauts qui sont pourtant les plus critiques. Pour pallier ce manque, il est proposé d'utiliser également des images semi-synthétiques. Des générations de défauts sont envisagées avec de réseaux antagonistes et des méthodes classiques de traitement d'images. Ces défauts sont ensuite introduits de manière réaliste dans des images acquises par des drones.

Mots Clef

Apprentissage, images semi-synthétiques, retouche numérique, réseaux antagonistes génératifs.

Abstract

This work focuses on the extension of image databases of aircraft outer surfaces taken by autonomous drones. Those databases are used to train and test visual automatic inspection algorithms. The effectiveness of learning depends on the representativeness and the quantity of examples. Available databases have only a small amount of elements considered as defects which are however the most critical. To overcome this lack, it is proposed to use also semi-synthetic images. Generations of defects are envisaged with antagonistic networks and conventional methods of image processing. Then the defects are introduced realistically into drone-acquired images.

Keywords

Machine Learning, semi-synthetic images, image editing, generative adversarial networks.

1 Introduction

Lors des opérations de maintenance aéronautique, les inspections pratiquées pour détecter des défauts ou anomalies sur le fuselage externe sont en majorité visuelles. Ces contrôles sont sujets aux erreurs des opérateurs humains qui sont soumis à une pression croissante due à l'intensification du trafic aérien [1, 2].

Différentes solutions se sont développées afin d'automatiser les inspections dans la maintenance aéronautique et réduire les facteurs humains pouvant affecter ces processus. La recherche dans ce domaine a débuté dans les années 1990 avec des robots de surface [3, 4] puis roulants [5, 6, 7]. Récemment, des solutions commercialisables sont apparues avec des drones autonomes [8, 9]. L'objectif de ces initiatives est de rendre les opérations de contrôle plus fiables, répétables et traçables.

Plusieurs de ces plateformes robotiques emploient des techniques de traitement d'images pour réaliser des inspections à partir des acquisitions [5, 9, 10, 11, 12, 13]. Quel que soit le type d'approche employé, des données sont nécessaires pour permettre de développer, tester et valider ces techniques, en particulier pour les méthodes à base d'apprentissage. Ce besoin de données se heurte souvent au nombre restreint d'acquisitions possibles lié au coût d'immobilisation d'un avion au sol.

La section 2 présente la solution d'inspection proposée par la société Donecle et quelques exemples réels d'apprentissage issus des inspections. Plusieurs de ces drones sont en activité, permettant à l'entreprise de répondre à ce besoin avec des prises de vue répétables venant régulièrement accroître les données existantes. Un autre inconvénient apparaît alors : le manque de défauts.

En effet, la grande majorité des exemples d'apprentissage disponibles sont des éléments courants et anodins présents sur le fuselage (vis, rivet, etc.). Ces derniers sont utiles pour une bonne compréhension de la scène et du contexte, nécessaire à une classification efficace des défauts.

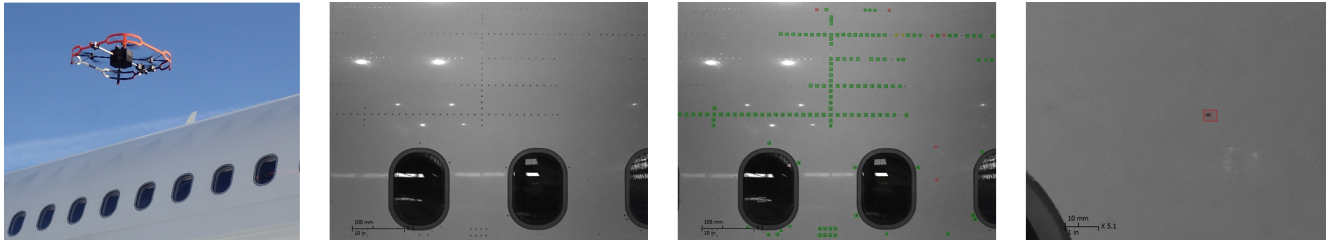


FIGURE 1 – De gauche à droite, vol du drone, image acquise, extraction de régions d’intérêt, détection d’un défaut potentiel.

Pour pallier ce manque de certains éléments, ainsi que ce déséquilibre de classes entre les objets courants et rares, il est envisagé d’utiliser des données synthétiques ou semi-synthétiques. La génération d’images pour compléter les ensembles d’apprentissage ou de validation se retrouve dans bien d’autres domaines appliqués comme par exemple la reconnaissance de texte dans des scènes naturelles [14], la navigation spatiale [15] ou encore la biométrie [16, 17]. La section 3 présente différentes approches de génération de défauts synthétiques de petite taille (quelques centimètres à quelques millimètres). Ces derniers sont ensuite incorporés dans des images réelles afin d’obtenir des données semi-synthétiques. La section 4 détaille ce processus de génération de scènes contenant les différents types de défauts rencontrés à la surface d’un avion. La section 5 a pour but d’évaluer la qualité des images ainsi générées et l’intérêt pour l’apprentissage sur ces bases de données étendues artificiellement.

2 Système actuel d’inspection

Une inspection visuelle par le système Donecle a pour étapes : l’acquisition d’images, leurs analyses automatiques puis la validation des résultats par un expert. Cette dernière opération constitue un mécanisme de contrôle autant qu’un moyen d’étiqueter les données nécessaire aux algorithmes d’apprentissage supervisé.

2.1 Acquisition automatique d’images

L’inspection peut avoir lieu à l’intérieur d’un hangar ou à l’extérieur sur le tarmac. Un technicien de maintenance se rend avec le drone près de l’appareil à inspecter, choisit une mission et le place dans sa zone de décollage.

Lors du vol du drone le système de localisation emploie une technologie de positionnement laser. Un algorithme de navigation permet de planifier puis de contrôler les mouvements du drone lors de l’acquisition des données en exploitant en temps réel sa position par rapport à l’avion [18, 19]. Les capteurs de navigation garantissent également la sécurité des opérations en prévenant les collisions avec la main-d’œuvre et les équipements. La figure 1 (à gauche) montre le drone en vol et une acquisition.

2.2 Analyse automatique des images

Une fois le vol terminé, l’analyse automatique des images peut être effectuée sur la tablette de l’opérateur. L’objectif

est de décrire la scène capturée en reconnaissant les objets qui s’y trouvent pour déceler les défauts sur le fuselage. Ce problème de reconnaissance d’objets nécessite de détecter les régions d’intérêt et d’en déterminer la nature.

La figure 1 (à droite) montre le résultat de l’analyse automatique d’une image acquise et un défaut potentiel signalé. Les systèmes les plus performants tels que [20, 21] sont constitués d’un réseau de neurones entraîné de bout en bout (*end-to-end*) pour détecter et identifier les objets de la scène. Ces modèles sont efficaces au sens du score de *Mean Average Precision* mais nécessitent de nombreuses données du fait de leur complexité (grand nombre de paramètres à apprendre).

Un apprentissage par transfert permet de limiter le nombre de données nécessaire en utilisant des représentations apprises sur des jeux de référence, néanmoins la spécificité des acquisitions nécessite un ré-apprentissage conséquent et des bases de données de plusieurs centaines d’éléments pour chaque classe. De plus ces réseaux nécessitent une grande quantité de mémoire GPU, ce qui limite la taille des images pouvant être traitées sur des machines portables en un temps d’inspection raisonnable. Les images doivent donc être traitées dans une résolution plus basse dans un premier temps.

Pour effectuer les tâches de détection et de classification il est possible d’utiliser un réseau de classification en complément de la reconnaissance d’objets dans la scène. Ce réseau travaille sur des images en résolution complète.

Une combinaison de modèles coopératifs de ce type est actuellement utilisée par Donecle. Un protocole de création de bases de données par des humains a été mis en place.

2.3 Bases annotées de données

L’opérateur en charge de l’inspection doit valider ou réfuter chaque zone d’image contenant des défauts détectés et peut également en ajouter si besoin. Ces zones permettent d’alimenter la base de données existante composée d’objets extraits dans les scènes. Des campagnes d’annotation poussées sont également réalisées sur d’autres éléments (e.g. vis, rivet, signalétique) de l’avion.

La figure 2 présente quelques zones annotées. En pratique, une inspection totale d’avion fournit bien plus d’éléments (e.g. vis, rivet, signalétique) intacts que d’éléments défectueux. Les défauts graves (e.g. fuite, impact) sont encore moins courants.

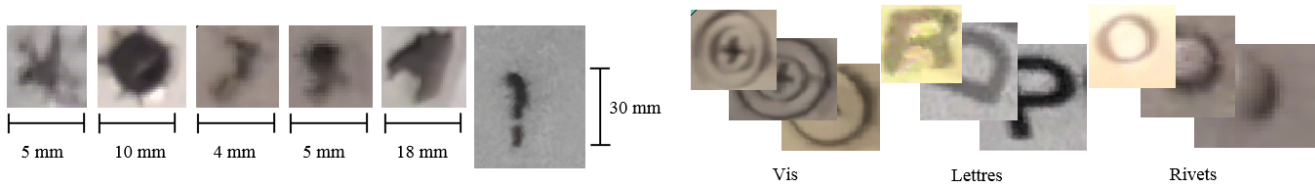


FIGURE 2 – Exemples d’éléments détectés. À droite, défauts et échelle, à gauche, autres objets annotés.

Dans le cadre de la présente étude, une base de données annotées d’environ 50 000 éléments est considérée. Conformément à la distribution observée lors d’inspections visuelles, certaines classes sont très représentées (vis et rivets consistent près de 60% de la base de données) alors que d’autres sont bien plus rares (certains types de défauts représentent moins de 1% de cette base).

L’étiquetage des données nécessaire aux réseaux de reconnaissance d’objets dans une scène présente deux inconvénients majeurs comparé à l’annotation pour la classification (où chaque image est associée à une étiquette).

Tout d’abord le format d’entrée nécessaire à ce type d’annotation rend la tâche laborieuse : chaque objet présent dans la scène doit être caractérisé par un rectangle englobant et une classe. Si des outils spécifiques pour l’annotation ont été développés, la présence de plusieurs dizaines, voire centaines d’objets de petite taille dans une seule acquisition ne permet pas de généraliser un procédé rapide et ergonomique.

De plus l’imprécision sur le rectangle englobant l’objet, et les erreurs dues à l’omission sont inévitables et induisent un biais qui est difficilement quantifiable. En effet la fonction de coût utilisée pendant l’entraînement est sensible au recouvrement des rectangles englobants, et un objet non annoté dans la scène sera considéré à tort comme un exemple négatif.

Ces limitations incitent à trouver de nouveaux moyens d’obtention des données nécessaires à l’apprentissage.

3 Augmentation des défauts

Pour répondre au manque de données d’apprentissage, plusieurs méthodes permettant de produire artificiellement de nouveaux exemples de défauts ont été utilisées. On peut ensuite effectuer un apprentissage sur des ensembles contenant des données réelles et synthétiques dans des proportions maîtrisées. Les scènes ainsi créées doivent être considérées photo réalistes afin de pouvoir être comparées et mélangées aux données réelles.

On évalue cette exigence selon trois critères : les objets placés dans la scène doivent être réalistes, l’intégration d’un objet (fondu avec le fond de la scène) doit être naturelle, et la position de l’objet dans la scène doit être cohérente.

L’objectif est d’étendre les bases de données d’entraînement sans perte de représentativité tout en respectant cette contrainte. La notion de contrôle (maîtrise et connaissance) des transformations appliquées est également importante

pour une meilleure compréhension des effets produits du système final après apprentissage.

La manière la plus courante d’étendre les bases de données consiste à appliquer des combinaisons de transformations élémentaires aux images existantes.

3.1 Transformations géométriques

Ainsi ces modifications peuvent prendre la forme de rotation, translation, symétrie, agrandissement. ou addition de bruit, et sont toujours combinées dans un souci de réalisme. Par exemple, les rivets étant tous de même taille et photographiés à la même distance, un changement d’échelle ne leur est pas appliqué. En outre, les homothéties conservent les propriétés géométrie des objets, ce qui constitue un gage de cohérence.

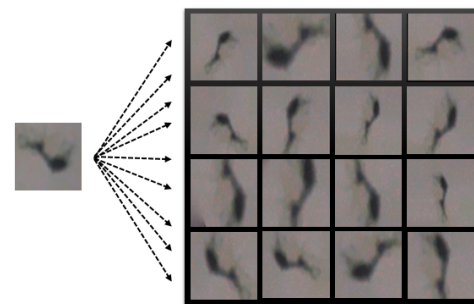


FIGURE 3 – Transformations appliquées à un défaut. A gauche la zone d’intérêt extraite d’une acquisition, à droite de multiples transformations de cette zone.

Ces techniques ne permettent pas de produire un grand nombre d’éléments et aboutissent à des bases de données dont la représentativité est faible en comparaison des jeux exclusivement constitués d’acquisitions. Cependant elles permettent à la fois un excellent contrôle des données synthétiques qui sont de surcroît presque indiscernables des acquisitions.

Pour aller plus loin, on distingue alors deux cas : les défauts qui surviennent sur une partie lisse du fuselage et ceux qui correspondent à la détérioration d’un élément déjà présent (vis, rivet, marquage, etc.). Ils peuvent être traités par des techniques différentes.

On propose ensuite de décrire plusieurs techniques s’appuyant sur la déformation d’objets existants puis une méthode de génération de défauts par un réseau de neurone qui peut être utilisé sur une portion lisse du fuselage.

3.2 Emploi de modèles spécifiques

Pour certains objets l'utilisation d'un modèle de vieillissement (évolution dans le temps) est adapté au phénomène de dégradation qu'il peut subir. Par exemple pour décrire l'évolution d'un rivet correctement peint vers un rivet dépoli. Ainsi il est possible dans ces cas-là d'utiliser des opérations de traitement d'image pour dégrader un objet en un défaut de manière graduelle, quantitative et réversible.

Ces méthodes sont appliquées aux vis, aux rivets et aux marquages signalétiques présents sur le fuselage. Ces approches présentent aussi l'intérêt de valider des algorithmes de détection ou d'évaluation sur des défauts qui seront probablement vu un jour mais qui ne peuvent pas être créés en réel sur un avion en service.

Simulation de dévissage par rotation de vis. La vérification des vis qui fixent certaines parties du fuselage sont des sujets majeurs dans la maintenance aéronautique. Des recherches en vision par ordinateur [22, 23] ont été menées pour détecter le dévissage voire la perte de tels éléments qui peuvent compromettre l'avion entier. Le modèle évolutif proposé simule une rotation de la tête de vis. Les étapes d'un dévissage sont présentées en figure 4. Un trait rouge est placé dans l'axe du sillon de la tête de vis par les opérateurs sur chaque vis dans les endroits à risque.

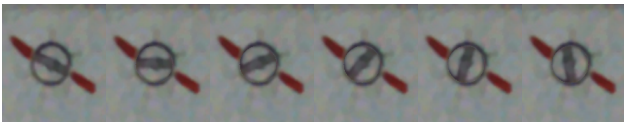


FIGURE 4 – Simulation d'un dévissage.

L'utilisation de ce modèle présente les propriétés d'une transformation géométrique simple : elle garantit la cohérence et une connaissance quantitative de la déformation mais ne permet qu'une extension limitée de jeu de défaut et produit des exemples peu représentatifs de la diversité rencontrée en conditions réelles.

Perte de peinture sur une tête de fixation. Les fixations, rivets ou vis, perdent souvent leur peinture avant le reste du fuselage. Des images de rivets à différents stades sont présentées dans la figure 5.



FIGURE 5 – Exemple de rivets avec à gauche, des rivets acceptables, et à droite, des rivets avec pertes de peinture.

Un modèle d'évolution permet de générer de manière contrôlée la dégradation progressive d'un rivet. Le défaut de peinture simulé débute aléatoirement par une attaque qui peut être périphérique ou centrale. Les pixels endommagés sont obtenus à chaque itération par une dilatation progressive de la zone impactée et d'autres zones similaires

apparaissent ou non sur le rivet suivant un processus aléatoire. Le taux de détérioration détermine l'ampleur de l'assombrissement de la zone. La figure 6 illustre une perte de peinture progressive sur des têtes de rivet.

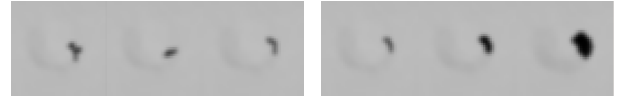


FIGURE 6 – Exemple de rivets synthétiques avec à gauche, des débuts aléatoires de perte de peinture, et à droite, une perte qui s'accroît progressivement.

L'utilisation de ce modèle permet un grand contrôle quantitatif de la déformation et une extension modérée du jeu de défauts. Néanmoins ces éléments dégradés simulés sont souvent présents en grand nombre dans les acquisitions. Il peut donc être utile de faire appel à ces modèles pour l'évaluation quantitative de certains défauts, qui correspond alors à une tâche de régression en apprentissage.

Génération d'écritures synthétiques defectueuses. Les marquages textuels ou symboliques sont régulièrement inspectés et peuvent subir des dégradations. Le système d'inspection doit identifier les symboles ou caractères, détecter de potentiels défauts, et évaluer leur lisibilité.

La génération de texte synthétique permet d'utiliser l'apprentissage profond pour des situations présentant peu de lots d'apprentissage. Jaderberg et al. [14] ont illustré la pertinence des résultats des systèmes appris sur des éléments synthétiques sur les jeux de données classiques (ICDAR, Street view text, etc.). Pour cela, un texte issu d'un lot de mots envisagés est généré. Les polices d'écriture employées par les algorithmes de génération sont celles utilisées par les compagnies aériennes. Ce texte subit ensuite des déformations aléatoires (rotation, perspective, floutage) pour prendre en compte les différentes situations de prise de photo. La dernière étape est de fonder ce nouveau texte dans une image pré-sélectionnée comme vide d'éléments (pour éviter les risques de superposition objet réel/texte généré) dans la section 4.

Certains marquages sont peints avec des pochoirs et peuvent présenter des défauts tels que des coulures. Une étape de détérioration de l'image est ajoutée dans le processus précédent afin d'obtenir tous les cas envisageables dans le lot d'apprentissage. Un intérêt particulier est apporté aux défauts de peinture, qui subissent des aléas tels que la coulure ou l'effacement des peintures. Ces défauts sont générés en propageant la couleur du texte dans une direction normale au bord de la lettre, en un point choisi au hasard. La longueur du défaut est contrôlable, ainsi que la stratégie de propagation (changement de direction de la trace, épaisseur, ainsi que son effacement continu selon la distance à la lettre).

Les figures 7 et 8 illustrent de telles générations.

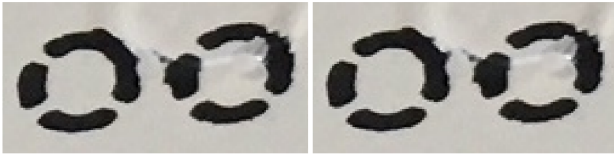


FIGURE 7 – À gauche, l'image originale contenant une coulure réelle sur le 2^{ème} O, à droite la même image en simulant une coulure sur le premier.

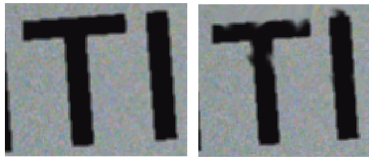


FIGURE 8 – À gauche, un texte intégré dans une scène et le même texte altéré à droite.

3.3 Génération via des réseaux antagonistes

Une autre approche pour produire de nouveaux exemples d'apprentissage est basée sur l'utilisation de réseaux antagonistes génératifs (*Generative Adversarial Network* ou GAN en anglais) [24]. Celle-ci permet la génération d'images à partir d'un jeu de données non étiquetées (apprentissage non supervisé). Ce réseau fait intervenir deux réseaux concurrents : un génératif et un discriminant.

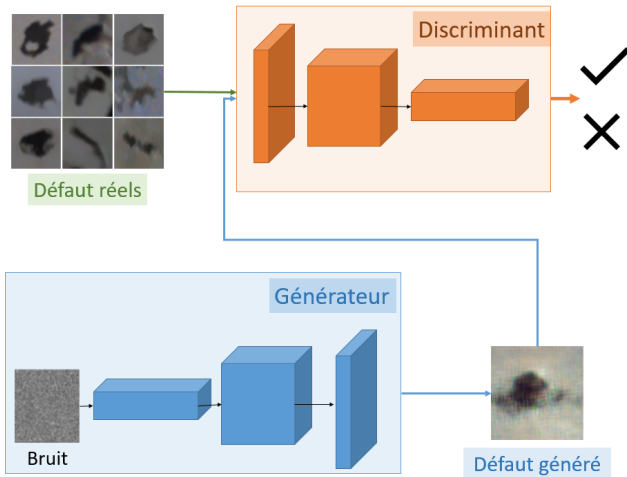


FIGURE 9 – Fonctionnement des réseaux antagonistes.

À partir de valeurs pseudo-aléatoires (bruit), le réseau génératif génère l'image d'un défaut pour la soumettre au réseau discriminant. Entraîné avec une base de données de défauts réels, le réseau discriminant doit déterminer si le défaut soumis par le générateur est réel (issu d'une acquisition) ou non (issu du générateur). Ces derniers entrent en compétition dans une situation de jeu à somme non nulle,

dans laquelle le succès de l'un des joueurs implique l'échec de l'autre. Le réseau génératif apprend à produire des défauts de plus en plus réalistes pour tromper le réseau discriminant qui cherche à différencier les défauts réels des défauts générés. La figure 9 illustre leur fonctionnement.

Les réseaux antagonistes se déclinent en différentes versions, utilisées pour de nombreuses applications.

La variante des réseaux antagonistes utilisée ici est un *Deep Convolutional GAN* (DCGAN) [25]. L'utilisation de couches de convolution revient à ajouter une contrainte spatiale aux représentations apprises, ce qui explique le succès de tels réseaux pour la génération d'images.

Les figures 10 et 11 présentent respectivement des défauts issus d'acquisitions et d'autres générés via un GAN.

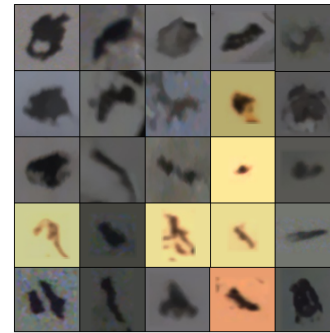


FIGURE 10 – Défauts extraits des acquisitions.

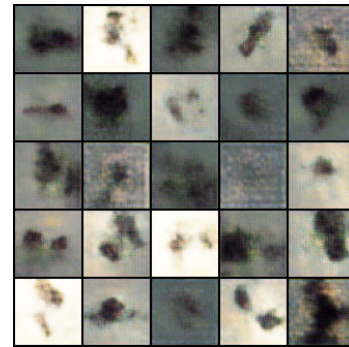


FIGURE 11 – Défauts générés avec un GAN.

Les réseaux antagonistes permettent une grande extension des bases de données sans perte de représentativité, et la création de défauts sans objet réel a priori. Néanmoins ils ne permettent pas un contrôle quantitatif de données ainsi produites : la taille, la forme et la géométrie des défauts ainsi générés sont le reflet des caractéristiques apprises mais ne sont pas directement maîtrisées.

4 Intégration dans une scène

Le réseau utilisé ici permet la localisation et la caractérisation des différents éléments présents sur l'avion à partir de données dans le format détaillé précédemment.

Cette partie traite de la création automatisée de scènes semi-synthétique respectant ce formalisme. Les éléments (réels ou synthétiques) sont insérés aléatoirement dans une image prise par un drone.

4.1 Intégration d'un texte

La fusion d'un texte est plus immédiate qu'un défaut. Étant généré, on connaît le masque du texte, qui est ensuite nécessaire à la fusion. Cette fusion se fait en plusieurs étapes : le masque H_{texte} du texte est obtenu en isolant les lettres écrites (étape réalisée lors de la conception du texte afin de faciliter la segmentation : un simple seuillage absolu est appliqué). Il est ensuite flouté avec un flou gaussien afin d'atténuer les bords contrastés. L'image résultante est obtenue selon l'opération suivante :

$M_{fusion} = H_{texte} \odot M_{orig} + (1 - H_{texte}) \odot M_{fond}$
avec M_{orig} image du texte, M_{fond} image vide d'éléments, M_{fusion} image résultante, \odot opérateur du produit matriciel de Hadamard.

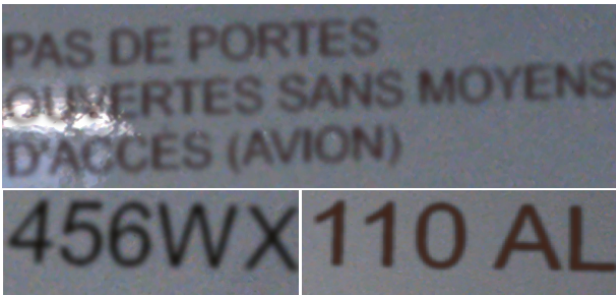


FIGURE 12 – Fusion de différents textes dans l'image

L'image ainsi fusionnée donne l'impression d'un texte fondu dans l'image comme on peut le voir dans figure 12. Une fois le texte fusionné avec le fond de l'image, les positions et tailles de chacune des lettres (après les multiples transformations) sont enregistrées : ce nouvel échantillon est utilisable par un algorithme d'apprentissage supervisé.

4.2 Insertion de plusieurs défauts

Le défaut à importer dans l'image, qu'il soit créé (c.f. section 3.3) ou sélectionné dans un lot de défauts réels, doit être isolé de son fond pour l'insertion dans l'image courante. L'optimisation par *GrabCut* [26, 27] est employée pour répondre à ce problème. Considérant que les défauts sont pleins et homogènes en leur centre, on choisit de placer une ancre d'initialisation du premier plan en ce centre. Le masque ainsi obtenu est utilisé de la même manière qu'en section 4.1. Un exemple est donné sur la figure 13.

4.3 Cas particulier d'insertion

Un cas particulier concerne des autocollants positionnés sur le fuselage, qui doivent être repérés pour faire un suivi particulier en un point de la structure. Ces autocollants n'étant pas homogènes, la méthode *GrabCut* ne donne plus de résultats suffisants.

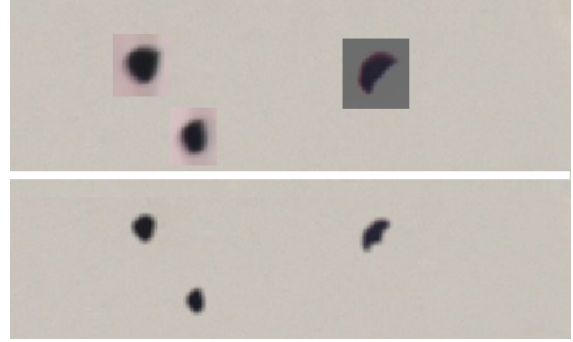


FIGURE 13 – Insertion de plusieurs défauts : en haut trois éléments à insérer sur le fuselage, en bas le résultat de leur intégration photo-réaliste.

N'ayant que peu d'échantillons de ces autocollants, leurs masques sont créés à la main, et le nombre d'échantillon est augmenté artificiellement grâce à la méthode vue en section 3.1 (les transformations sont alors aussi appliquées au masque en question). Un tel exemple d'intégration est visible dans la figure 14.

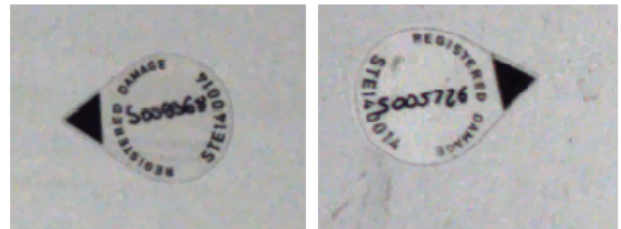


FIGURE 14 – À gauche, un autocollant sur une image réelle, à droite, un autocollant intégré d'une autre image.

5 Évaluation des performances liées à l'augmentation des données

5.1 Influence sur la classification

On évalue la pertinence des techniques de création de d'objets détaillées plus haut en procédant à l'entraînement d'un système sur des ensembles d'apprentissage constitués d'une part croissante d'images de synthèses issues des différentes techniques dans proportions variables. Afin de déterminer la stratégie d'augmentation des bases de données le déséquilibre naturel des classes est à prendre en compte. En effet, dans [28] il est établi que le déséquilibre des classes induit une baisse des performances des algorithmes d'apprentissage par réseau de neurones, et que la solution la plus efficace consiste à augmenter artificiellement les exemples de classes sous-représentées. Ainsi on procède à la balance des classes par augmentation sélective, et on compare les résultats en classification des systèmes entraînés. Ce faisant on cherche à déterminer le taux d'augmentation maximal acceptable.

Pour une répartition très déséquilibrée des éléments dans les classes on observe le taux de succès ainsi que les scores de précisions et rappel pour les défauts sur l'ensemble de test pour différents jeux d'entraînement :

- données réelles uniquement.
- égalisation artificielle : les classes moins représentées sont artificiellement étendues pour que leur cardinal égale celui de la classe la plus représentée.
- augmentation d'un facteur α : après égalisation, toutes les classes sont artificiellement étendues.

	réelles	égalisation	$\alpha = 2$	$\alpha = 4$
succès	0.921	0.945	0.961	0.960
précision	0.73	0.79	0.84	0.83
rappel	0.74	0.85	0.89	0.89

TABLE 1 – Performances en classification avec entraînement sur des bases réelles, égalisées et étendues : taux de succès global, précision et rappel pour les défauts.

Les résultats, rassemblés dans la table 1, permettent d'établir l'intérêt de l'extension artificielle des bases de données pour la tâche de classification, jusqu'à un certain point. En outre, on observe que l'extension ne présente un intérêt que dans une certaine limite (coefficient α).

Nous n'avons pas pu établir la pertinence de l'utilisation de réseaux génératifs par rapport aux matrices de transformation géométriques (approches présentées dans la section 3). L'utilisation de technique de réseaux génératifs plus récentes, telles que les GANs progressifs [29] ou l'ajout de conditions sur l'entraînement de ces réseaux devront être étudiés afin de parfaire cette approche.

5.2 Influence sur la détection

Afin de valider les méthodes de création de scènes pour étendre artificiellement les jeux d'apprentissage pour la détection, nous avons considéré un système entraîné uniquement sur des scènes réelles annotées. Nous avons comparé les performances de ce système de détection confronté à des scènes semi-synthétique créées en utilisant les méthodes détaillées précédemment et des scènes réelles jamais observées. Les performances sur ces deux catégories de scènes ne sont pas significativement différentes, nous considérons donc qu'un tel système se comporte de manière similaire dans les deux situations.

Un apprentissage sur des scènes semi-synthétiques doit désormais être effectué pour observer si la suppression des incertitudes (bruit d'annotation) induit un meilleur apprentissage pour la détection d'objets.

6 Conclusions et perspectives

Pour l'application décrite, nous avons présenté plusieurs méthodes d'extension de bases de données en vue de l'utilisation d'algorithmes d'apprentissage. Nous avons montré qu'elles permettent d'améliorer significativement les performances en classification dans notre cas, en particulier pour les défauts présents sur le fuselage pour lesquels il est

crucial d'obtenir un score de rappel le plus élevé possible. Un des avantages de nos acquisitions qui facilite l'insertion est la régularité de la distance d'acquisition.

Les différentes approches abordées ici ne permettent cependant pas une extension des jeux de données d'apprentissage sans perte de leur représentativité, ce qui limite la capacité de généralisation des modèles entraînés sur des bases de données artificiellement étendues.

Différentes méthodes de génération de défauts proposées dans cet article peuvent être améliorées. De nouveaux modèles de dégradation doivent être proposés pour les renforcements, fuites, fissures dans la peinture du fuselage, salissures. Pour ces dernières, une séparation fond-texture pourrait permettre d'apprendre les textures qui donnent un aspect sale. Des modèles d'illuminations permettraient de diversifier encore les bases de données.

L'utilisation de modèles génératifs plus sophistiqués pourrait permettre une extension plus importante des bases de données, permettant de diviser les classes de défauts en sous classes plus précises.

Pour la détection d'objets, les résultats préliminaires indiquent que la méthode d'extension des bases de données via la création de scènes semi-synthétiques est potentiellement un moyen de rendre la tâche d'annotation moins laborieuse tout en s'affranchissant des imprécisions humaines. Un apprentissage sur ces scènes doit permettre de valider cette hypothèse. Ces travaux ouvrent la voie à de nouveaux processus d'entraînements de systèmes performant nécessitant de moins en moins de données annotées par les humains et permettant la reconnaissance d'éléments rares. Ils sont ici présentés dans le cadre d'une application à l'industrie aéronautique, mais de nombreuses applications, par exemple médicales [30] font face au problème du déséquilibre des classes alors même que les classes les moins représentées sont les plus importantes.

Si l'utilisation de transformations spatiales semble pertinent, le choix empirique ou aléatoire des combinaisons et paramètres ne permet pas de déterminer avec une grande confiance l'efficacité a priori de ces méthodes. Une alternative pour venir à bout de l'imprécision en annotation pourrait venir de l'emploi d'un *Spatial Transformer Network* comme algorithme de transformation géométrique. Ce réseau est entraîné selon un mode non supervisée avec pour critère la capacité d'apprentissage supplémentaire permise par les transformations apprises [31]. Ce bloc agit comme un pré-traitement appliqué à toutes les images avant leur entrée dans le reste du réseau, et s'entraîne conjointement : il faut apprendre ce pré-traitement dynamique.

Remerciements

Les auteurs remercient les étudiants de l'UPSSITECH Ander Jimenez-Garcia, Anwar Alfatayri, Julien Sanchez et Marvin Jaubert pour leur travail préliminaire sur le sujet qu'ils ont effectué au cours d'un projet école avec l'entreprise Donecle. Les auteurs remercient également les membres actuels et passés de Donecle.

Références

- [1] D. A. Marx and R. C. Graeber, *Human error in aircraft maintenance*. Johnstone, McDonald & Fuller, Avebury Technical, Aldershot, 1994.
- [2] C. G. Drury, “Human reliability in civil aircraft inspection,” in *Human Factors and Medicine Panel (HFM) Workshop*, 1999.
- [3] M. Siegel, “Remote and automated inspection : Status and prospects,” in *1st Joint DoD/FAA/NASA Conference on Aging Aircraft*, juillet 1997.
- [4] M. Siegel, P. Gunatilake, and G. Podnar, “Robotic assistants for aircraft inspectors,” *IEEE Instrumentation Measurement Magazine*, vol. 1, pp. 16–30, mars 1998.
- [5] I. Jovančević, S. Larnier, J.-J. Orteu, and T. Sentenac, “Automated exterior inspection of an aircraft with a pan-tilt-zoom camera mounted on a mobile robot,” *Journal of Electronic Imaging*, vol. 24, novembre 2015.
- [6] F. Donadio, J. Frejaille, S. Larnier, and S. Vetault, “Human-robot collaboration to perform aircraft inspection in working environment,” in *5th International conference on Machine Control and Guidance*, 2016.
- [7] F. Donadio, J. Frejaille, S. Larnier, and S. Vetault, “Artificial Intelligence and Collaborative Robot to Improve Airport Operations,” in *Online Engineering & Internet of Things*, pp. 973–986, Springer International Publishing, 2018.
- [8] L. Bjerregaard, “Aircraft drone inspection technology,” *MRO Network*, janvier 2018.
- [9] J. Miranda, S. Larnier, and M. Claybrough, “Caractérisation d’objets sur des images acquises par drone,” in *Conférence Reconnaissance des Formes, Image, Apprentissage et Perception*, 2018.
- [10] I. Jovančević, I. Viana, J.-J. Orteu, T. Sentenac, and S. Larnier, “Matching CAD model and image features for robot navigation and inspection of an aircraft,” in *International Conference on Pattern Recognition Applications and Methods (ICPRAM 2016)*, février 2016.
- [11] I. Jovančević, A. Arafat, J.-J. Orteu, and T. Sentenac, “Airplane tire inspection by image processing techniques,” in *5th Mediterranean Conference on Embedded Computing (MECO’2016)*, juin 2016.
- [12] T. Villemot, S. Larnier, and S. Vetault, “Détection d’amers visuels pour la navigation d’un robot autonome autour d’un avion et son inspection,” in *Reconnaissance des Formes et l’Intelligence Artificielle*, juin 2016.
- [13] J. R. Leiva, T. Villemot, G. Dangoumeau, M. A. Bauda, and S. Larnier, “Automatic visual detection and verification of exterior aircraft elements,” in *2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM)*, mai 2017.
- [14] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, “Synthetic data and artificial neural networks for natural scene text recognition,” *NIPS Deep Learning Workshop*, 2014.
- [15] S. Li, R. Lu, L. Zhang, and Y. Peng, “Image processing algorithms for deep-space autonomous optical navigation,” *Journal of Navigation*, vol. 66, no. 4, pp. 605–623, 2013.
- [16] R. Cappelli, D. Maio, and D. Maltoni, “Synthetic fingerprint-database generation,” in *Object recognition supported by user interaction for service robots*, vol. 3, pp. 744–747, août 2002.
- [17] S. Shah and A. Ross, “Generating synthetic irises by feature agglomeration,” *2006 International Conference on Image Processing*, pp. 317–320, 2006.
- [18] M. Claybrough, “System and method for automatically inspecting surfaces. Patent WO 2016203151,” 2016.
- [19] A. Deruaz-Pepin, “Method and system for determining the position of a moving craft. Patent WO 2017121936,” 2017.
- [20] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “Ssd : Single shot multibox detector,” in *European conference on computer vision*, pp. 21–37, Springer, 2016.
- [21] J. Redmon and A. Farhadi, “Yolov3 : An incremental improvement,” *arXiv preprint arXiv :1804.02767*, 2018.
- [22] M. Rice, L. Li, G. Ying, M. Wan, E. T. Lim, G. Feng, J. Ng, M. Nicole, T. Jin-Li, and V. S. Babu, “Automating the visual inspection of aircraft,” in *Aerospace Technology and Engineering Conference*, 2018.
- [23] J. Miranda, S. Larnier, A. Herbulot, and M. Devy, “Uav-based inspection of airplane exterior screws with computer vision,” in *14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications.*, 2019.
- [24] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” *Advances in Neural Information Processing Systems*, vol. 27, pp. 2672–2680, 2014.
- [25] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *CoRR*, 2015.
- [26] C. Rother, V. Kolmogorov, and A. Blake, “Grabcut - interactive foreground extraction using iterated graph cuts,” *ACM Transactions on Graphics (SIGGRAPH)*, août 2004.
- [27] C. Rother, V. Kolmogorov, Y. Boykov, and A. Blake, “Interactive foreground extraction using graph cut,” Tech. Rep. MSR-TR-2011-46, Microsoft, 2011.
- [28] M. Buda, A. Maki, and M. A. Mazurowski, “A systematic study of the class imbalance problem in convolutional neural networks,” *Neural Networks*, vol. 106, pp. 249–259, 2018.
- [29] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive growing of gans for improved quality, stability, and variation,” *arXiv*, 2017.
- [30] B. Krawczyk, M. Galar, Ł. Jeleń, and F. Herrera, “Evolutionary undersampling boosting for imbalanced classification of breast cancer malignancy,” *Applied Soft Computing*, vol. 38, pp. 714–726, 2016.
- [31] M. Jaderberg, K. Simonyan, A. Zisserman, et al., “Spatial transformer networks,” in *Advances in neural information processing systems*, pp. 2017–2025, 2015.