



الجمهورية الجزائرية الديمقراطية الشعبية

People's Democratic Republic of Algeria

وزارة التعليم العالي والبحث العلمي

Ministry of Higher Education and Scientific Research

جامعة محمد المهدى ابن باديس - مستغانم

Abdelhamid Ibn Badis University of Mostaganem



THESE

Présentée à la

FACULTÉ DES SCIENCES ET DE LA TECHNOLOGIE

DÉPARTEMENT GÉNIE ÉLECTRIQUE

Laboratoire Signaux et Systèmes

Pour l'obtention du diplôme

DOCTORAT

Option : Signaux et Systèmes

Par

CHOUAI Mohamed

Systeme semi-automatique de détection des éléments d'intérêt dans un scanne de bagage

Soutenue le 21/06/2020 devant la commission d'examen :

Président du jury	OULD MAMMAR Madani	MCA	Université Mostaganem
Examineurs	AHMED Fouatih Zoubir	Pr	Université USTO
	LOUKIL Abdelhamid	Pr	Université USTO
	BENOUDNINE Hadjira	Pr	Université Mostaganem
Encadrant	MERAH Mostefa	MCA	Université Mostaganem
Co-encadrant	MIMI Malika	Pr	Université Mostaganem

Résumé

Depuis plusieurs décennies, les avions sont des cibles intéressantes pour les terroristes. Si l'on se penche sur l'histoire des attaques contre des avions, les bombes sont l'une des principales préoccupations, c'est-à-dire les engins explosifs improvisés. La base de données mondiale sur le terrorisme en 2017 recense 893 attaques d'aéroports ou d'avions avec des explosifs, dont 247 ont eu lieu après 2001. Face aux risques accrus, en particulier depuis le 11 septembre 2001, les aéroports et les gouvernements ont accru leurs investissements dans la sûreté de l'aviation. Ces dernières années, l'apprentissage en profondeur a été utilisé avec succès dans la reconnaissance d'images et de vidéos. L'objectif de notre équipe est de développer un système de détection utilisant l'apprentissage automatique et l'apprentissage approfondi pour détecter efficacement les objets menaçants (armes blanches, armes à feu, explosifs). Notre contribution décrit des algorithmes très précis pour la détection d'objets menace dans l'imagerie complexe de bagage radiologique. Deux problèmes principaux sont abordés : i) la détection des objets menaces masquées ou chevauchés par d'autres objets ; et ii) minimiser le taux des fausses alarmes sur les images ne contenant pas des objets menaces qui contiennent souvent des motifs similaires. Cette thèse tente également d'apporter une contribution au domaine de la reconnaissance d'objets dans les tests par rayons X dans le domaine de l'inspection des bagages en évaluant différentes stratégies de vision par ordinateur proposées ces dernières années.

mots clés : avions, terroriste, bombe, aéroports, rayons X , objets menaçant, l'apprentissage automatique, l'apprentissage en profondeur, détection, vision par ordinateur

Abstract

For several decades, airplanes have been attractive targets for terrorists. If you look at the history of attacks on airplanes, explosives are one of the main concerns, that is, improvised explosive devices. The global database on terrorism in 2017 lists 893 attacks on airports or planes with explosives, of which 247 occurred after 2001. Faced with increased risks, in particular since September 11, 2001, airports and governments have increased their investments in aviation security. In recent years, deep learning has been used successfully in image and video recognition. The goal of our team is to develop a detection system using

machine learning and deep learning to effectively detect threatening objects (knives, firearms, explosives). Our contribution describes very precise algorithms for the detection of threat objects in complex imaging of radiological baggage. Two main problems are addressed: i) detection of threatened objects masked or overlapped by other objects; and ii) minimize the false alarms rate on images not containing threatened objects which often contain similar objects. This thesis also attempts to make a contribution to the field of object recognition in X-ray tests in the field of baggage inspection by evaluating different computer vision strategies proposed in recent years.

keywords: planes, terrorist, bomb, airports, x-rays, threatening objects, machine learning, deep learning, detection, computer vision

ملخص

منذ عدة عقود ، كانت الطائرات أهدافًا جذابة للإرهابيين. إذا نظرت إلى تاريخ الهجمات على الطائرات ، فإن المتفجرات هي أحد الشواغل الرئيسية ، وهي الأجهزة المتفجرة المرتجلة. تدرج قاعدة البيانات العالمية عن الإرهاب في عام 2017 893 هجومًا على المطارات أو الطائرات بالمتفجرات ، منها 247 حدثت بعد عام 2001. وفي مواجهة المخاطر المتزايدة ، خاصة منذ 11 سبتمبر 2001 ، زادت المطارات والحكومات من استثماراتها في أمن الطيران. في السنوات الأخيرة ، تم استخدام التعلم العميق بنجاح في التعرف على الصور والفيديو. هدف فريقنا هو تطوير نظام الكشف باستخدام التعلم الآلي والتعلم العميق لاكتشاف الأشياء المهددة (السكاكين والأسلحة النارية والمتفجرات) بفعالية. توضح مساهمتنا خوارزميات دقيقة للغاية للكشف عن الأجسام المهددة في التصوير المعقد للأمتعة الإشعاعية. تتم معالجة مشكلتين رئيسيتين: (1) الكشف عن الأشياء المهددة التي تحجبها أو تتداخل معها الأشياء الأخرى (2) تقليل معدل الإنذارات الخاطئة على الصور التي لا تحتوي على أشياء مهددة والتي تحتوي غالبًا على أشياء مماثلة. تحاول هذه الأطروحة أيضًا المساهمة في مجال التعرف على الأشياء في اختبارات الأشعة السينية في مجال فحص الأمتعة من خلال تقييم استراتيجيات رؤية الكمبيوتر المختلفة المقترحة في السنوات الأخيرة.

الكلمات الرئيسية: طائرات ، إرهابي ، قنبلة ، مطارات ، أشعة سينية ، أشياء مهددة ، التعلم الآلي ، التعلم العميق ، الكشف ، رؤية الكمبيوتر

Remerciements et dédicaces

Nous devons remercier Allah Tout-Puissant pour toute la volonté et le courage qu'il nous a donné pour mener à bien ce travail.

L'achèvement de ce travail mené sur plusieurs années procure une grande satisfaction. Il est l'occasion de se remémorer les différentes embûches qu'il a fallu surmonter, mais surtout les personnes qui m'ont permis d'en arriver là.

En premier lieu, je remercie mes parents (YASMINA BENATCHI et HOUARI) qui ont su croire en moi et qui m'ont apporté toute leur aide quand j'en ai eu besoin. Cette thèse leur est dédiée à 200%. Je remercie également toute ma famille (qui ne cesse de s'agrandir) qui a contribué de près ou de loin à ce que je suis devenu. Je souhaite également exprimer ma sincère gratitude à mes quatre frangins Miloud, Hamida, Anouar et Nourine qui ont eu le droit et le privilège de me supporter pendant plus de vingt ans. Je veux ajouter à ce passage les fils de mon frère Miloud : Houari et Farah, également la fille de mon frère Hamida: Yasmina (quand vous serez grand, vous lirez ce passage, alors je vous aime).

Nous tenons aussi à remercier le Président et les membres du jury qui m'ont fait l'honneur de bien vouloir étudier avec attention notre travail, qui ont bien daigné siéger la soutenance de notre thèse.

Je voudrais remercier mon directeur de thèse, Dr MERAH Mostefa, pour ses conseils incommensurables, sa patience et son soutien tout au long de mon parcours de doctorat. J'ai beaucoup appris en travaillant directement avec lui, pour lequel il reçoit mes sincères remerciements. De même, à madame la professeur MIMI Malika, sa confiance en mes capacités a permis une grande partie de mon travail de création, je la remercie pour toute la patience et la disponibilité dont elle a fait preuve à mon égard. Je tiens également à remercier Dr ABED Mansour, Dr OULD MAMMAR Madani et Dr BENTOUMI Mohamed de leurs conseils utiles et de leur participation à mon comité.

Cette recherche n'aurait pu être réalisée sans le soutien de monsieur le professeur José-Luis SANCHO-GOMEZ de l'université polytechnique de Carthagène en Espagne. Nous remercions chaleureusement tous les membres de l'équipe de recherche TDAM (traitement des données et apprentissage automatique) et toutes les personnes qui ont travaillé avec moi sur le projet à l'université polytechnique de Carthagène. Je voudrais remercier tous mes amis à Cartagena, en Espagne, pour leur amitié et leur soutien au cours de mon parcours de doctorat sur tous mes amis Marouane SALHAOUI, Mohamed el Amine OUELD MAMMAR et Ahmed ELLOUMI.

Je ne souhaite pas oublier les administrateurs de l'université polytechnique de Carthagène (David MORAL, Ana Isabel MOSTAZA FERNÁNDEZ et Marisa RUBIO SOLE) pour leur attention et leur efficacité qui ont permis que je passe deux années à l'UPCT dans de bonnes conditions grâce à leur professionnalisme.

Je tiens également à remercier tous mes amis et collègues de l'équipe de recherche au laboratoire LSS à Mostaganem pour leur aide, leur soutien et leur gentillesse. Je souhaite également les remercier pour la confiance dont ils ont fait preuve à mon égard en me laissant m'impliquer au sein de la vie collective de l'équipe. Les rapports humains dont j'ai profité à leur côté ont fait naître de réels liens d'amitié qui à mes yeux n'ont pas de prix. Qu'ils soient tous assurés de ma plus profonde gratitude et estime. J'exprime en particulier ma gratitude à Sara SENINETE, Mohammed GUEZGOUZ et Arezki LARBI pour toute l'aide qu'ils ont su m'apporter.

Une dédicace spéciale à Sara SENINETE, pour lui dire que je suis profondément reconnaissant de ce que vous avez fait pour moi. Une chose est sûre : je n'oublierai jamais. Merci ... c'est un mot trop simple. Ce que je voudrais exprimer est au-dessus de cela. Je suis à la fois touché et reconnaissant de l'aide que vous m'avez apportée ... et je ne vous remercierai jamais assez.

Je souhaite également remercier tous mes amis de grande date qui ont su m'apporter confiance et écoute à tous les moments. Je les remercie pour avoir supporté mes blagues vaseuses et mes jeux stupides, mais bon on ne change pas du jour au lendemain. Ainsi, je remercie particulièrement Sadek, Mohamed, Maamoune, Salah-eddine et tous les autres pour avoir partagé un grand nombre d'années à mes côtés (il y en aura d'autres c'est sûr).

Je tiens également à remercier tous mes amis formidables qui j'ai l'honneur de les rencontrer à l'Université de Skikda, qui ont su me donner la confiance et à l'écoute de tous les instants. Un grand merci à Cherief Abdelwahid, Hakkar Wissam, Lakhdar Sidi Salah, Khalfi Mohamed El Amine, je veux vous dire que vous faites partie intégrante de ma vie (et même plus), sans oublier les mères de Wissam et Lakhder " Laatra et Houria" (sinon je serai puni), qui sont des femmes très sensibles et merveilleuses. Assurez-vous de ma sincère et profonde amitié et gratitude. Je sais, cela ne doit pas toujours être facile de me côtoyer à la fois sur un terrain et dans la vie de tous les jours...

Bien entendu, cette liste n'est pas exhaustive et je remercie tous ceux et celles qui me connaissent et qui me permettent de me sentir exister...Merci à toutes et à tous.

Table des matières

Résumé	ii
Remerciements et dédicaces	iv
Table des matières	vi
Liste des abréviations.....	ix
Liste des tableaux	xi
Liste des figures.....	xii
Introduction générale	1
Chapitre 1 Détection par rayons X	5
1.1 Introduction	5
1.2 Historique	5
1.3 Génération de rayons X et spectres	9
1.3.1 Génération de rayons X	10
1.3.2 Les spectres caractéristiques.....	11
1.4 Principes de la détection par rayons X.....	12
1.5 Le problème de la détection	13
1.5.1 Défi actuel : les explosifs artisanaux et leurs précurseurs	16
1.6 Système de technologies de détection à rayons X actuel	18
1.7 Systèmes d'imagerie par rayons X	21
1.7.1 Systèmes de dépistage par rayons X conventionnels.....	22
1.7.2 Systèmes d'imagerie biénergie à rayons X.....	23
1.7.3 Système d'imagerie par rayons X avec dispersion.....	25
1.7.4 Système de tomographie par ordinateur (CT).....	26
1.7.5 Techniques de rétrodiffusion	28

1.8	Les problèmes opérationnels	29
1.8.1	Le processus d'inspection	29
1.8.2	Objets interdits dans les bagages de cabine ou les bagages enregistrés dans le cadre juridique	31
1.8.3	Considérations opérationnelles pour l'inspection des bagages	32
1.9	Conclusion.....	35
Chapitre 2 État de l'art		36
2.1	Introduction	36
2.2	Amélioration de la qualité de l'image	36
2.3	Segmentation.....	47
2.4	Compréhension de l'image	53
2.4.1	Classification.....	53
2.4.2	Détection.....	54
2.5	Conclusion.....	56
Chapitre 3 Notion opératoire et méthodes utilisées		57
3.1	Introduction	57
3.2	Amélioration de la qualité d'image.....	58
3.2.1	Amélioration du contraste.....	59
3.2.2	Autres techniques utilisées.....	65
3.2.3	Notre technique proposée de dé-bruitage d'image	65
3.2.4	Evaluation par signal-to-noise ratio (PSNR)	71
3.3	Segmentation.....	72
3.3.1	Segmentation basée sur le clustering en utilisant l'apprentissage automatique .	72
3.3.2	Segmentation sémantique profonde.....	82
3.4	Détection d'objets	88
3.4.1	Après notre modèle de segmentation proposé	89
3.4.2	Utilisant le deep learning	89
3.5	Extraction de caractéristiques.....	90
3.5.1	Caractéristiques de formes.....	91

3.5.2	Caractéristiques de couleur	91
3.5.3	Caractéristiques de texture	91
3.5.4	Notre modèle proposé pour l'extraction des caractéristiques	92
3.6	Sélection des caractéristiques	95
3.6.1	Brève description des méthodes de sélection des caractéristiques utilisées	95
3.7	Classification	96
3.8	Conclusion	97
Chapitre 4 Résultats et discussions		98
4.1	Introduction	98
4.2	Base de données	98
4.3	Amélioration de la qualité d'image	98
4.4	Segmentation	99
4.4.1	Segmentation basée sur le clustering en utilisant l'apprentissage automatique	99
4.4.2	Segmentation sémantique profonde	104
4.5	Détection d'objets	111
4.5.1	Par les méthodes de machine learning	111
4.5.2	Par la méthode YOLOv3 modifiée	113
4.5.3	Après la Segmentation sémantique profonde	114
4.6	Extraction de caractéristiques	116
4.6.1	Par notre modèle proposé	116
4.7	Sélection des caractéristiques et classification	118
4.7.1	Classification par notre modèle proposé	119
4.8	Conclusion	123
Conclusions générale		124
Production scientifique		127
Bibliographie		128

Liste des abréviations

μ	<i>coefficient d'atténuation linéaire</i>
A	<i>poids atomique</i>
AAE	<i>Adversarial AutoEncoder</i>
AE	<i>Autoencoder</i>
AHE	<i>Adaptive histogram equalization</i>
BMU	<i>Best Matching Unit</i>
BoVW	<i>Bag-of-Visual-Words</i>
CEAC	<i>communauté internationale de contrôle des bagages</i>
CH	<i>Calinski-Harabasz index</i>
CLAHE	<i>Contrast-limited adaptive histogram equalization</i>
CNN	<i>Réseau neuronal convolutif</i>
CT	<i>tomographie par ordinateur</i>
DB	<i>Davies-Bouldin index</i>
DECT	<i>tomographie par rayons X à double énergie</i>
DI	<i>Dunn index</i>
DWT	<i>transformée en ondelettes discrète</i>
EDS	<i>systèmes de détection d'explosifs</i>
EDSCB	<i>Explosive detection systems for cabin baggage</i>
EM	<i>Expectation-Maximization</i>
FAA	<i>Federal Aviation Administration</i>
FCN	<i>Fully Convolutional Network</i>
GLCM	<i>Gray-Level Co-Occurrence Matrix</i>
H.E	<i>high energy</i>
HI	<i>Hartigan index</i>
HME	<i>Homemade Explosives</i>
HOG	<i>histogramme de gradients orientés</i>
HSV	<i>hue, saturation, lightness</i>
HTDS	<i>Hi-TECH DETECTION SYSTEMS</i>
IDWT	<i>transformée en ondelettes discrète inverse</i>
IED	<i>sengins explosifs improvisés</i>
IEF	<i>fonctions d'amélioration de l'image</i>
KNN	<i>K-Nearest Neighbors</i>
KR	<i>Kernel</i>
L.E	<i>low energy</i>
LDA	<i>Linear discriminant analysis</i>
LReLU	<i>Leaky Rectified Linear Unit</i>
LSVM	<i>SVM avec un noyaux linéaire</i>

ML.....	<i>apprentissage machine</i>
MLP	<i>Multilayer perceptron</i>
MSE.....	<i>Mean squared error</i>
N_A.....	<i>nombre d'Avogadro</i>
NB	<i>Naive Bayes</i>
NMR	<i>Nuclear Magnetic Resonance</i>
PDF	<i>densité de probabilité</i>
PIT.....	<i>test d'élément interdit</i>
polynomial	<i>SVM avec un noyaux polynomial</i>
PRMF.....	<i>filtre médian de restauration de pixels</i>
PSNR.....	<i>Evaluation par signal-to-noise ratio</i>
RANSAC	<i>Random sample consensus</i>
RBFSVM	<i>SVM avec un noyaux radial basique</i>
RF.....	<i>Random Forest</i>
RGB	<i>red, green, blue</i>
SDT	<i>spike detection technique</i>
SOM.....	<i>cartes auto-organisatrices</i>
SVM.....	<i>machine à vecteurs de support</i>
SymRG.....	<i>Symmetric Region-Growing</i>
TIP.....	<i>Threat Image Projection</i>
TNA	<i>Thermal Neutron Analysis</i>
TSA.....	<i>Transport Security Administration</i>
UV	<i>Ultraviolets</i>
VD.....	<i>Vapeur Detection</i>
W_φ.....	<i>coefficient d'approximation</i>
XRT.....	<i>X-Ray Tuto</i>
Zeff.....	<i>numéro atomique effectif</i>
ρ	<i>masse volumique</i>

Liste des tableaux

Tableau 1 : Valeurs typiques de μ et $H1/2$ pour le rayonnement monoénergétique.....	15
Tableau 2 : Sélection des explosifs et des milieux explosifs précurseurs,.....	17
Tableau 3 : Les règles de sélection de la CEAC (communauté internationale de contrôle des bagages)	30
Tableau 4: La différence entre l'amélioration du contraste local et l'amélioration du contraste global.....	60
Tableau 5: Structures des réseaux de codeurs et de décodeurs. Dans lequel les blocs de décodage reproduisent le processus de codage dans l'ordre inverse.....	94
Tableau 6: Matrices de confusion. Ces matrices nous permettent de visualiser les performances des algorithmes utilisés après la tâche de fusion (deux classes).....	101
Tableau 7: Le résultat principal des meilleures (Naive Bayes) et des pires méthodes (Random Forest).....	101
Tableau 8: Résultat de la segmentation montrant la performance comparative entre le système proposé et les autres méthodes testées.....	105
Tableau 9: Le résultat moyen de la précision, les limites F1 et IoU entre le système proposé et les autres méthodes testées.....	107
Tableau 10: La performance du YOLOv3 modifiée par rapport à l'original pour notre proposition	113
Tableau 11: Résultat de précision / écart type moyen de classification avec la meilleure méthode de sélection des caractéristiques.....	118
Tableau 12: Performances de classification pour détecter les armes à feu, les objets explosifs et les objets non menaçants.....	122

Liste des figures

Figure 1: Description schématique du tube à rayons X.	10
Figure 2 : Schéma d'une installation de radiographie par projection standard permettant d'imager des objets volumineux.....	14
Figure 3 : Exemple 1 d'une image fournie par un système dépistage par rayons X conventionnels monoénergétiques.....	23
Figure 4 : Exemple 2 d'une image fournie par un système dépistage par rayons X conventionnels monoénergétiques.....	23
Figure 5 : Exemple d'image fournie par un système de dépistage par rayons X à double énergie.	24
Figure 6 : Schéma du tomographe CT	27
Figure 7 : Schéma d'une situation de géométrie en rétrodiffusion.....	29
Figure 8 : une image fournie par un système de radiographie X-ray à double énergie se joint au bagage d'origine.	37
Figure 9 : amélioration de la qualité de l'image par le filtre niveaux de gris.....	40
Figure 10 : amélioration de la qualité de l'image par amélioration locale.	41
Figure 11 : amélioration de la qualité de l'image par amélioration de la puissance de pénétration élevée.	41
Figure 12 : amélioration de la qualité de l'image par amélioration de la faible puissance de pénétration.....	42
Figure 13 : amélioration de la qualité de l'image par super amélioration.	42
Figure 14 : amélioration de la qualité de l'image par exclusion des produits organiques.	43
Figure 15 : amélioration de la qualité de l'image par exclusion inorganique.	43
Figure 16 : amélioration de la qualité de l'image par le filtre négatif.	44
Figure 17 : amélioration de la qualité de l'image par éclaircissement.	44
Figure 18 : amélioration de la qualité de l'image par assombrissement.	45
Figure 19 : amélioration de la qualité de l'image par balayage des niveaux de gris.....	45
Figure 20 : amélioration de la qualité de l'image par traitement de pseudo-couleur.	46

Figure 21 : amélioration de la qualité de l'image par amélioration des substances organiques suspectes Z7 / Z8 / Z9.	46
Figure 22 : amélioration de la qualité de l'image par amélioration des bords.	47
Figure 23: Schéma fonctionnel d'un système de détection permettant de détecter des objets dans des images numérisées de balayage de bagages à rayons X.	57
Figure 24: Un organigramme de l'approche d'amélioration d'image proposée.	66
Figure 25: Les images radiologiques originales, ils représentent respectivement de haut au bas: image à basse énergie, image haute énergie, l'image en niveaux de gris, l'image fusionnée.	69
Figure 26: Un diagramme sur la procédure de la segmentation par machine learning.	73
Figure 27: Principe de fonctionnement de la méthode regroupement hiérarchique.	75
Figure 28: Dendrogramme présente la sortie principale du clustering hiérarchique.	75
Figure 29: L'architecture du FCN.	83
Figure 30: L'architecture du SegNet.	84
Figure 31: L'architecture de l'U-Net.	84
Figure 32: L'architecture du DeconvNet.	85
Figure 33: L'architecture du RedNet.	85
Figure 34: L'architecture du système de segmentation sémantique proposée (SemSegAAE).	86
Figure 35: Un schéma synoptique illustre la préparation du ground truth et de la procédure de détection d'objet.	88
Figure 36: La procédure d'apprentissage proposée en utilisant l'apprentissage des caractéristiques avec l'adversarial autoencoder étiqueté.	92
Figure 37: Schéma de l'adversarial Autoencoder étiqueté proposé.	93
Figure 38: Variation des valeurs de PSNR pour 100 images applique pour les différentes méthodes testées.	99
Figure 39: Processus d'optimisation des hyper-paramètres via la méthode CV.	102
Figure 40: Pourcentage moyen de la meilleure segmentation obtenue par l'indice de validation de la classification.	103

Figure 41: Temps de calcul moyen obtenu pour chaque méthode après avoir testé 1000 images différentes.	103
Figure 42: Quelques résultats expérimentaux de segmentation sur des images de test.	109
Figure 43: Résultats comparatifs entre la segmentation obtenue par les méthodes considérées et la vérité de terrain.	110
Figure 44 : La figure montre quelques résultats de détection d'objet extraits de notre ensemble de données.....	112
Figure 45: Exemples de détection d'objet.	115
Figure 46: La représentation latente des objets d'apprentissage (première colonne) et la représentation des codes latents pour certains objets de test (deuxième colonne).	117
Figure 47 : Les matrices de confusion montrent clairement que notre système de proposition basé sur AAE à 5 dimensions et espace permet d'obtenir les meilleurs résultats.	121

Introduction générale

L'industrie du transport aérien et en particulier le transport commercial est un élément important de l'économie mondiale actuelle. Chaque année, des centaines de millions de personnes utilisent des avions pour voyager dans le monde entier. Les flux nationaux et internationaux de passagers, de courrier et de cargaison dépendent de la sûreté du système de l'aviation civile. Les avions en vol, de par leur nature et leur conception, sont extrêmement sensibles aux explosives et constituent une cible attrayante pour les terroristes. La base de données mondiale sur le terrorisme en 2017 [1] recense 893 attaques d'aéroports ou d'avions avec des explosifs, dont 247 ont eu lieu après 2001. Récemment, le 29 juillet 2017, un complot terroriste a été déjoué à l'aéroport de Sydney lorsqu'un engin piégé a été dissimulé dans un sac [2].

Depuis le 11 septembre 2001, le contrôle aérien par les scanners à rayons X est devenu plus qu'important. Le processus de contrôle est complexe, car les objets menaçants sont très difficiles à détecter lorsqu'ils sont placés dans des sacs fermés, superposés par d'autres objets et/ou orientés dans différentes positions rendant ainsi difficile leur détection. Dans le contrôle des bagages, où la complexité de l'inspection est très élevée, les opérateurs humains sont encore utilisés. Néanmoins, ils ont seulement quelques secondes pour décider si un sac contient un objet interdit ou pas, ainsi, les performances de détection ne dépassent guère les 90%. Bien que, l'opérateur humain reste incontournable dans ce type de technologie, il présente un certain nombre de facteurs de risque tels que la distraction ou les erreurs d'inattention, ce qui constitue un risque majeur, d'où une aide à la décision par ordinateur est nécessaire.

Face aux risques accrus, les aéroports et les gouvernements ont accru leurs investissements dans la sûreté de l'aviation [3]. Au cours des cinq dernières années, des systèmes de détection d'explosifs au contrôle des bagages (EDSCB) sont disponibles [4]. Il y a peu de pays tels que les États-Unis qui utilisent ces systèmes [5], [6].

Les chercheurs dans le domaine d'inspection des bagages peuvent être grossièrement divisés en quatre groupes :

- 1) Le groupe « système » chargé de la configuration, des tests et du développement du système, ainsi que des logiciels associés, il fournit des images de rayons X exploitable pour le groupe traitement d'images ;
- 2) Le groupe « physique » qui travaille sur la physique de la détection par rayons X et ses simulations, c'est ce groupe qui coordonne les efforts de recherche des autres groupes;
- 3) Le groupe « traitement d'image » chargé de développer des algorithmes avancés de traitement des images afin d'extraire des informations pertinentes d'atténuation des rayons X concernant les objets complexes. Le groupe envoie ensuite, les informations relatives à chaque objet au 2^{ème} groupe qui assure la caractérisation des matériaux ;
- 4) Le groupe « développement » dont le rôle est de développer de nouvelles méthodes pour l'amélioration de la caractérisation des matériaux. Ce groupe renvoie ainsi, les nouvelles propositions au système en vue d'une modification ultérieure.

Les intensités des rayons X de chaque objet ne sont pas uniquement liées à la composition de son matériau, qui reflète la propriété intrinsèque, mais également à certains paramètres extrinsèques, tels que la position, l'orientation et l'épaisseur. Pour une application pratique, tous les paramètres extrinsèques ci-dessus sont indéterminés et ils affecteront négativement les performances de la caractérisation des matériaux. C'est l'un des problèmes majeurs des systèmes existants de détection à rayons X à double énergie.

Le nombre d'objets figurant dans un bagage de passager est souvent très important. Différents objets peuvent apparaître dans un sac dans des orientations arbitraires de forme et de géométrie différentes ; et la plupart d'entre eux chevauchent d'autres objets. Particulièrement, les explosifs sont généralement mélangés avec des matières inoffensives, ce qui rend leur détection très difficile. Ainsi, le but de notre équipe de recherche est de détecter et d'identifier tous les objets à menaces présents dans le bagage et cela d'une manière semi-automatique en prenant en considération le chevauchement des différents objets.

Ces dernières années, l'apprentissage en profondeur a été utilisé avec succès dans la reconnaissance des images et des vidéos [7], [8]. L'objectif affilié à notre équipe est de

développer un système utilisant l'apprentissage automatique et l'apprentissage approfondi pour détecter efficacement les objets à menace (armes blanches, armes à feu, explosifs).

L'application des méthodes de vision artificielle aux images radiographiques des bagages présente de nombreux avantages par rapport à l'inspection manuelle, notamment un débit d'inspection élevé dû à l'automatisation. Toutefois, comme indiqué dans les travaux connexes (Chapitre 2), peu de travaux ont été publiés sur le traitement automatisé des images par rayons X des bagages, probablement en raison de la difficulté d'obtenir des ensembles de données étiquetés suffisamment volumineux et aussi dû au caractère confidentiel de ses données.

Une détection automatique précise et fiable des objets à risques est plus que nécessaire, cependant, ce type de détection souffre de deux problèmes majeurs : i) la détection des objets à risque, masqués ou chevauchés par d'autres objets ; et ii) le taux de fausse alarme sur les images ne contenant pas des objets à risque, ce qui à force, peut nuire à l'utilisation de ce type de système.

Dans cette thèse nous présentons des systèmes performants de détection d'objets à risque (armes blanches, armes à feu et les explosives) dans les images radiographiques des bagages souvent complexe, après avoir testé différentes stratégies de vision par ordinateur proposées ces dernières années.

Ainsi, différentes étapes de système liées aux traitements d'images sont abordées dans ce travail, comme présentées ci-dessous :

- Amélioration de la qualité des images : généralement pour fournir une image plus claire à l'œil humain, et aussi pour l'utiliser comme étape de prétraitement dans la détection automatique des objets.
- Séparation des images radiographique présentant des chevauchements, en particulier une séparation en images organiques et images inorganiques,
- Détection individuelle de tous les objets présents dans les bagages,
- Réduction du nombre de ressources, par des techniques de sélection de caractéristiques,
- Classification des objets, par l'intelligence artificielle, permettant de formuler des jugements compliqués avec précision, rapidité et cohérence,

- Détection des objets à risques sans passer par la segmentation, à l'aide d'approches d'apprentissage en profondeur.

Cette thèse est organisée de la manière suivante : le Chapitre 1 présente une vue globale sur l'état actuel de l'inspection des bagages dans les aéroports et leurs problèmes opérationnels, ainsi, que les technologies de détection par rayons X. Le Chapitre 2 décrit un bref aperçu sur l'état de l'art des systèmes d'inspection semi-automatique. Le Chapitre 3 détaille l'étude expérimentale et les méthodes utilisées. Les résultats expérimentaux sont présentés au Chapitre 4. Une conclusion générale clôture cette thèse.

Chapitre 1

Détection par rayons X

1.1 Introduction

Le rayon X est une forme de rayonnement électromagnétique qui peut traverser différents matériaux. Les rayons X peuvent être utilisés dans les systèmes d'inspections des bagages dans les aéroports utilisant ainsi, des détecteurs capables de représenter simultanément, des matériaux de poids atomique élevé (des armes à feu) et des matériaux de poids atomique faible (des explosifs plastiques). Pour faire face aux défis de la détection des objets à risques, des efforts énormes sont déployés pour augmenter les performances des systèmes d'inspection. Dans ce chapitre, nous verrons l'évolution des techniques matérielles d'imagerie à rayons X au cours des dernières décennies, pour l'inspection des bagages ainsi que les problèmes opérationnels rencontrés dans ce domaine.

1.2 Historique

L'industrie du transport aérien commercial est un élément important de l'économie mondiale actuelle. Chaque année, des centaines de millions de personnes utilisent des avions pour voyager dans le monde entier. Les flux nationaux et internationaux de passagers, de courrier et de cargaison dépendent de la sûreté du système de l'aviation civile. Les avions en vol, par leur nature et leur conception, sont extrêmement sensibles aux bombes et constituent une cible attrayante pour les terroristes, dont l'objectif est d'infliger de nombreuses victimes. De 1985 à 1997, huit avions commerciaux ont été perdus ou endommagés en raison d'attaques terroristes présumées [9], [10], et environ 1100 personnes sont mortes dans ces tragédies :

- Le 23 juin 1985, un Boeing 747 d'Air India s'est écrasé dans la mer à la suite de l'explosion de la soute.

- Le 29 novembre 1987, le vol 858 de Korean Air a été détruit par un engin explosif dans la cabine.
- Le 21 décembre 1988, le vol 103 de Pan Am a été détruit par une bombe de 12 onces dissimulée dans une radio portable au-dessus de Lockerbie, en Écosse.
- Le 19 septembre 1989, un vol UTA a été détruit au-dessus du Sahara à la suite d'une explosion cachée dans le cargo.
- Le 27 novembre 1989, un Boeing 727 Avianca a été détruit par un engin explosif dans la cabine.
- Le 17 juillet 1994, un avion EMB-110 de la compagnie Alas Chiricanas s'est écrasé après l'explosion d'une bombe dans la cabine pendant le vol de Colon City à Panama City.
- Le 11 décembre 1994, un Boeing 727 de Philippine Airlines a été attaqué en vol par l'explosion d'une bombe dans la cabine.
- Le 9 juillet 1997, un engin explosif dans la cabine des passagers a explosé l'avion de la compagnie aérienne TAM Mercosur.

En réponse à une série d'attaques terroristes contre l'aviation, le congrès américain a ordonné, par des amendements à la loi fédérale sur l'aviation de 1958, que la Federal Aviation Administration (FAA), en sa qualité de membre du département des transports (Department Of Transportation DOT), assume la sûreté de l'aviation civile. La nécessité de mesures spéciales de sécurité de l'aviation a été reconnue pour la première fois en 1969. Depuis lors, la FAA a mis au point et administré un certain nombre de programmes destinés à prévenir les actes criminels contre le secteur de l'aviation. À ses débuts, le programme était axé sur la menace de détournement d'avions et comprenait le développement de technologies et d'équipements de détection des armes, tels que les détecteurs des métaux et les appareils à rayons X. En 1974, le congrès américain a adopté la loi anti-détournement qui établit essentiellement la responsabilité de la FAA de mener un programme de recherche, d'ingénierie et de développement dans le domaine de la sécurité de l'aviation afin de contrer la menace terroriste. En 1976, précipité par un attentat à la bombe à l'aéroport LaGuardia de New York, la FAA a lancé un programme de recherche et développement sur la détection des explosifs. Les experts se sont réunis pour

discuter de la menace et ont recommandé un programme de sécurité à long terme et les technologies associées. Outre les techniques de rayons X, plusieurs technologies ont été suggérées et étudiées, notamment l'analyse par neutrons thermiques (TNA), la résonance magnétique nucléaire (NMR) et la détection de vapeur (VD).

L'attentat à la bombe contre le vol 103 de la Pan Am, qui a coûté la vie à 270 personnes, a contribué à l'application de la loi sur la sécurité de l'aviation (Aviation Security Improvement Act) de 1990. En août 1989, le décret 12686 a créé la commission présidentielle sur la sécurité de l'aviation et le terrorisme. Un rapport publié en mai 1990 recommandait à la FAA de poursuivre un programme plus rigoureux de recherche et développement en matière de sûreté de l'aviation afin de contrer la menace terroriste. En novembre 1990, le congrès a adopté la loi sur l'amélioration de la sécurité de l'aviation afin d'accélérer la recherche et le développement des technologies et des procédures de lutte contre le terrorisme. La loi demandait à la FAA de financer la mise au point des systèmes de détection d'explosifs (EDS), d'établir des normes de certification EDS et de tester et certifier les EDS potentiels en vue d'un déploiement éventuel. La FAA a lancé le mouvement pour développer et évaluer la performance de la nouvelle technologie.

Entre 1991 et 1996, la FAA a investi 153,3 millions de dollars dans le développement de technologies spécialement conçues pour détecter les explosifs dissimulés. Ces dépenses ont permis de financer environ 85 projets de développement d'une nouvelle technologie de détection des explosifs [11]. La FAA s'appuie principalement sur des contrats et des subventions avec des entreprises privées et des instituts de recherche pour développer ces technologies et effectue quelques recherches internes limitées.

Organisé et parrainé par la FAA, le premier symposium international sur les technologies de détection des explosifs s'est tenu les 13 - 15 novembre 1991 à Atlantic City, dans le New Jersey, USA. Son objectif était double: 1) fournir un forum pour l'échange d'idées, d'informations et de technologies entre les responsables de programmes, les scientifiques et les ingénieurs travaillant dans le domaine de la sûreté de l'aviation civile; et 2) déterminer l'orientation de la recherche et du développement dans le domaine critique de la sûreté de l'aviation civile mondiale [12]. Au total, 430 experts de la sûreté de l'aviation civile des États-Unis et de

l'étranger ont participé à la conférence. Plus de 100 communications sur la détection en masse, la détection de vapeur, le traitement du signal et la simulation, et l'intégration des systèmes ont été présentés au symposium par des scientifiques et ingénieurs de dix-neuf universités, de 14 laboratoires nationaux et de défense et des organisations du Canada , France, Allemagne et Royaume-Uni. Plus tard, la FAA a continué de parrainer ou de co-parrainer plusieurs autres symposiums, notamment le deuxième symposium sur les technologies de détection des explosifs et la conférence sur les technologies de sécurité de l'aviation, en novembre 1996.

La FAA estime que les explosifs placés dans les bagages enregistrés constituent la plus grande menace pour l'aviation [12]. En septembre 1993, la FAA a publié une norme de certification selon laquelle les systèmes de détection d'explosifs pour les bagages enregistrés doivent être conformes avant de pouvoir être déployés. La norme est classifiée et définit certains critères de performance minimaux pour 1) les substances explosives à détecter ; 2) la probabilité de détection, par explosif ; 3) la quantité d'explosif et 4) le nombre de bagages traités par heure. En outre, la norme spécifie le taux maximal admissible de fausse alarme, par explosif. Pour minimiser les erreurs humaines, la norme exige également que les dispositifs déclenchent automatiquement une alarme lorsque des explosifs sont suspectés. Cette fonctionnalité contraste avec les appareils à rayons X classiques, dans lesquels l'opérateur doit regarder l'écran de rayons X de chaque sac pour déterminer s'il contient une menace.

Le CTX5000, fabriqué par InVision Technologies, a réussi le test des exigences minimales au centre technique de la FAA en décembre 1994, devenant ainsi le premier appareil à rayons X certifié par la FAA pour la détection des explosifs. Ce système de détection d'explosif est basé sur le principe de la tomographie assistée par ordinateur, qui permet de déterminer les formes et les propriétés des objets, ce qui permet de les identifier comme des matières potentiellement explosives avec plus de certitude que dans les systèmes de radiographie en vue unique. À l'heure actuelle, aucun autre système n'a été certifié par la FAA pour la détection des explosifs [13].

L'accident du vol TWA 800 au large de Long Island en juillet 1996, bien qu'il ait été établi par la suite qu'aucune activité criminelle n'était causée, a conduit à la création de la commission de la maison blanche sur « la sécurité et la sécurité de l'aviation ». Le président a appelé la

commission à élaborer rapidement un plan d'action visant à déployer des équipements de haute technologie pour la détection des explosifs. Ce plan comprend l'achat et l'installation de cinquante-quatre systèmes CTX5000 SP (version modifiée du CTX5000), de douze appareils à rayons X à double énergie non certifiés, de plusieurs appareils à résonance quadripolaires et des détecteurs de trace. Utilisé en combinaison avec le profilage des passagers (technique policière qui consiste à établir le profil psychologique d'un individu recherché à partir d'indices recueillis par les enquêteurs), le déploiement d'équipements de haute technologie peut renforcer la sécurité de l'aviation. La combinaison de ces technologies peut fournir plusieurs niveaux d'inspection pour faciliter la résolution des fausses alarmes.

Comme indiqué dans le discours liminaire du premier symposium international sur la technologie de détection des explosifs, "Dans ce secteur, il n'y a pas de solution miracle ... Il n'y aura pas un seul équipement que nous puissions utiliser sur le terrain pour résoudre tous nos problèmes." Cela devrait encourager tout le monde sur le terrain à travailler sur un équipement qui, à leur avis, fera partie du travail " [14].

Pour les programmes de recherche et développement actuels et futurs, les projets de la FAA visant à développer des technologies de détection des bagages enregistrés comprennent des efforts pour améliorer le système certifié, développer de nouvelles technologies et évaluer diverses technologies. La FAA estime qu'une combinaison appropriée des systèmes qui, individuellement, ne répondent pas aux exigences de certification pourrait éventuellement fonctionner ensemble pour détecter les quantités, les configurations et les types de matières explosives requises par l'Aviation Security Improvement Act de 1990.

1.3 Génération de rayons X et spectres

Les rayons X sont un rayonnement électromagnétique dont les longueurs d'onde se situent dans la plage de $\lambda = 0,001$ à 10 nm (nanomètre), c'est-à-dire les énergies des photons $E = 120$ eV à 1,2 MeV [15]. Les photons d'énergie plus élevée (longueur d'onde plus courte) sont appelés rayons γ et les photons de rayonnement d'énergie plus faible (longueur d'onde plus longue) sont appelés ultraviolets (UV). Ces définitions peuvent être trouvées dans [16], [17], [18], [19], [20].

1.3.1 Génération de rayons X

C'est Roentgen qui a annoncé pour la première fois la découverte de ce rayonnement en 1895 au cours de quelques expériences avec des rayons cathodiques, qui sont des flux d'électrons dans des tubes à vide. Dans ses recherches sur les propriétés des rayons X, Roentgen découvrit que les rayons X sont générés à chaque fois que les rayons cathodiques frappent un corps solide et qu'il les stoppe. Ainsi, afin de générer des rayons X, trois composants principaux sont nécessaires, à savoir une source de particules chargées, un moyen de les accélérer et une cible solide pour les arrêter [21]. Les électrons sont généralement utilisés comme particules chargées et sont accélérés par une tension allant de quelques centaines de volts à des centaines de kilovolts. Lorsque des électrons accélérés frappent une cible, l'énergie cinétique de ces électrons est convertie en chaleur et en rayons X.

Le tube à rayons X est un type de générateur, il est largement utilisé dans les systèmes de balayage à rayons X. Les tubes à rayons X consistent principalement en une cible métallique solide utilisée comme anode, un filament utilisé en tant que cathode et deux circuits permettant de chauffer le filament et de diriger les électrons vers l'anode. Une source de rayons X est représentée schématiquement à la Figure 1. Le tungstène est couramment utilisé comme matériau cible en raison de sa bonne conductivité thermique et de son point de fusion élevé [22].

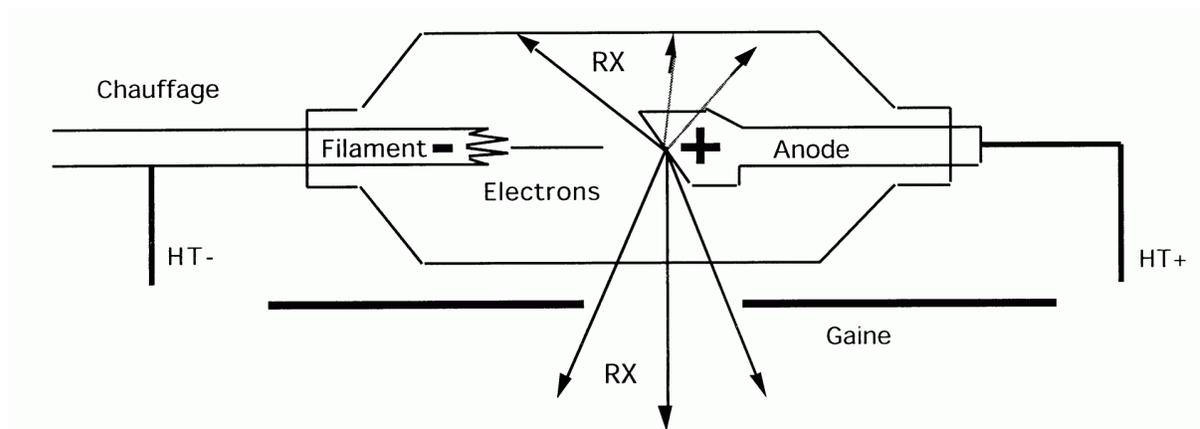


Figure 1: Description schématique du tube à rayons X.

1.3.2 Les spectres caractéristiques

Certains électrons de la cathode, lorsqu'ils heurtent la cible (connue sous le nom d'anode sur la Figure 1), peuvent avoir suffisamment d'énergie pour produire des ions en éliminant les électrons internes de l'atome, même le plus à l'intérieur, ou le K-shell. Un tel ion a un anneau de basse énergie dans sa structure électronique, et ce vide est rapidement rempli lorsqu'un de ses électrons dans un état d'énergie plus élevé tombe à ce niveau de basse énergie. Lorsqu'un électron tombe dans un niveau d'énergie faible, l'énergie est libérée sous la forme de rayons X ayant des longueurs d'onde particulières caractéristiques du matériau cible. Par conséquent, les rayons X générés par ce mécanisme sont appelés rayons X caractéristiques.

Les lignes d'émission caractéristiques ne peuvent résulter que de transitions d'électrons d'atomes dans le matériau cible entre des orbites autorisées, comme le prévoit une théorie quantique selon laquelle chaque électron d'un atome donné se déplace sur une orbitale caractérisée par quatre critères [23]:

- n , le nombre quantique principal, ou nombre de Bohr, est associé à des orbitales successives. n prend un entier positif 1, 2, 3, 4, ..., qui désigne respectivement les couches électroniques K, L, M, N, ...
- l , le nombre quantique angulaire, est une mesure du moment angulaire orbital qui explique l'existence d'orbitales elliptiques et circulaires à électrons. Il peut prendre toutes les valeurs entières comprises entre 0 et $(n-1)$ inclus. $l = 0$ correspond à une orbite circulaire.
- s , le nombre quantique de spin, ne peut prendre que deux valeurs possibles : $+1/2$ et $-1/2$. Il est souvent représenté par $+$ et $-$.
- j , le nombre quantique angulaire total. Sa valeur est calculée en tant que $l \pm s$ et doit être supérieure à 0. Chaque valeur possible de j pour une orbitale particulière détermine un niveau d'énergie unique.

1.4 Principes de la détection par rayons X

Les méthodes de détection par rayons X sont les plus largement utilisées pour inspecter les bagages à l'aéroport. Cette section présente quelques résultats en physique des rayons X, essentielle aux systèmes d'imagerie par rayons X.

Lorsqu'un faisceau de rayons X d'énergie E et d'intensité initiale I_0 traverse un matériau d'épaisseur t , son intensité I est réduite de manière exponentielle [24]

$$I = I_0 e^{-n\sigma t} \quad (1.1)$$

Où t est communément donné en cm, I et I_0 sont donnés en keV et σ est la section efficace totale en $\text{cm}^2 / \text{atome}$. La quantité σ est la somme de deux sections efficaces indépendantes dans la plage d'énergie utilisée généralement pour filtrer les bagages dans les aéroports : section transversale photoélectrique σ_{pe} et section transversale de diffusion des photons σ_{sc} . σ_{pe} et σ_{sc} sont les fonctions du nombre atomique et de l'énergie des rayons X.

$$\sigma(Z, t) = \sigma_{pe}(Z^{4.5}, E^{-3}) + Z\sigma_{sc}(E) \quad (1.2)$$

Pour une substance élémentaire donnée, le nombre d'atomes par unité de volume, n , est calculé à partir de

$$n = \frac{N_A \rho}{A} \quad (1.3)$$

Où N_A (6.02252×10^{23} atomes / mole) est le nombre d'Avogadro, A est le poids atomique en g / mole, ρ est la masse volumique de la substance en g / cm^3 et n est en atomes / cm^3 .

La section efficace totale par atome σ , multipliée par le nombre d'atomes par unité de volume de matière, est le coefficient d'atténuation linéaire μ , qui est par centimètre de parcours dans la matière. La fonction utilisée pour calculer le coefficient d'atténuation linéaire μ est la suivante :

$$\mu = \sigma \rho \frac{N_A}{A} \quad (1.4)$$

Si le coefficient d'atténuation linéaire, μ , est utilisé, l'atténuation des photons dans une matière homogène peut être définie par la loi exponentielle :

$$I = I_0 e^{-\mu t} \quad (1.5)$$

En prenant le logarithme des deux côtés et en l'écrivant explicitement, on obtient,

$$\ln\left(\frac{I_0}{I}\right) = \mu t = \sigma(Z, E) \frac{N_A}{A} \rho t \quad (1.6)$$

D'après l'équation (1.6), l'atténuation des rayons X à travers un objet est fonction de la composition, de la densité, de l'épaisseur et de la distribution spectrale des matériaux. Il existe un certain nombre de méthodes dans lesquelles les interactions fondamentales décrites ci-dessus peuvent être utilisées pour identifier des matériaux.

1.5 Le problème de la détection

L'utilisation de la technologie des rayons X pour l'inspection d'objets volumineux dépend entièrement de la nature pénétrante et de la capacité des rayons X énergétiques à distinguer les matériaux, les bagages enregistrés étant un excellent exemple de ce besoin. Bien entendu, avant de comprendre pourquoi ce comportement caractéristique devrait être, il faut d'abord reconnaître que la mise en œuvre pratique d'un système de contrôle par rayons X n'est une réalité qu'en raison d'un certain nombre de développements technologiques remarquables qui permettent un débit élevé, une sensibilité élevée d'arrangement de dépistage. Au cœur de ces développements, on trouve la source de rayons X compacte à haut flux, le plus souvent, mais non exclusivement, un verre sous vide (Evacuated glass) ou, plus récemment, un système de rayons X à tube céramique métal (parfois appelé enveloppe) comprenant une cathode émettrice d'électrons à chaud et une anode cible refroidie à haut point de fusion de grand numéro atomique (Z). Ceci assure un flux abondant de rayons X tout en limitant les dommages causés à l'enveloppe et à l'anode par la charge, résultant du désir d'avoir un point focal fin et donc une capacité de résolution spatiale élevée. L'alimentation en énergie est à son tour fournie par un générateur hautement stabilisé qui génère un potentiel d'accélération et des mA, dont les choix sont adaptés à la qualité des images et le débit, limités par la puissance thermique du tube.

Enfin, et dans les termes les plus larges, le système est complété par la fourniture d'un système de détection 2D à haute efficacité dont la conception est adaptée à une résolution spatiale donnée, cette dernière étant également influencée par le point focal défini par la disposition des tubes focalisés et la charge de puissance associée. La Figure 2 montre un schéma de configuration standard de radiographie par projection, dont l'évolution a été influencée par les divers facteurs décrits ci-dessus, ainsi que par les filtres et les grilles de durcissement, ces derniers inhibant la réception des images diffuses obstruant le rayonnement diffus.

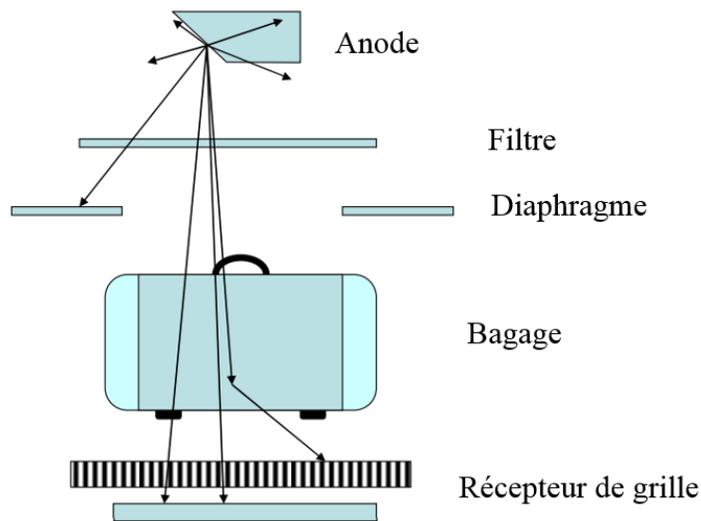


Figure 2 : Schéma d'une installation de radiographie par projection standard permettant d'imager des objets volumineux.

Des efforts énormes ont été déployés pour tenter de faire face aux très nombreux facteurs qui réduisent les performances, notamment les conceptions exploratoires d'anodes, par exemple, en métal liquide. Ce dernier est destiné à éviter les dommages de surface des systèmes d'anodes solides, dont l'ampleur devient plus prononcée avec l'âge, ce qui élargit le point focal, diminuant ainsi la résolution spatiale. Pour une excellente discussion sur ce qui précède, voir [25].

La situation dans la Figure 2 étant maintenant établie, on peut revenir à la nature pénétrante et à la capacité de discrimination matérielle des rayons X énergétiques. Les dépendances de l'imagerie par rayons X sont intimement liées au fait que les rayons x (et gamma) interagissent faiblement avec la matière via la force électromagnétique, la perte d'énergie limitée associée

par unité de longueur de trajet, également appelée transfert d'énergie linéaire, LET, étant typiquement exprimé en $\text{keV } \mu\text{m}^{-1}$. Ceci fournit à la capacité de pénétration quatre principaux mécanismes d'atténuation réduisant le flux de rayons X d'un faisceau confiné de photons qui se transportent à travers un milieu donné pour former l'image de transmission conventionnelle. Pour des énergies de photons jusqu'à 200 keV (au-dessus desquelles le contraste de l'image diminue considérablement, et aussi parce que la protection des installations de contrôle contre le rayonnement devient plus lourde), le Tableau 1 indique les capacités de pénétration indicatives, aussi bien en termes de coefficient d'atténuation linéaire combiné en tenant compte des fractions pondérales des composés.

Tableau 1 : Valeurs typiques de μ et H1/2 pour le rayonnement monoénergétique [26].

Énergie (keV)	Matériel	Numéro atomique	Densité (kg m ⁻³)	μ (mm ⁻¹)	H1 / 2 (mm)
30	Eau	7.5	10^3	0.036	19
60				0.020	35
200				0.014	50
30	Aluminium	13	1.27×10^3	0.160	4.3
60				0.050	13.9
200				0.020	35
30	Plomb	82	11.4×10^3	33	2×10^{-2}
60				5.5	0.13
200				1.1	0.6

Les systèmes de contrôle aux rayons X affichent des images en niveaux de gris, représentant les variations de forme, d'épaisseur et de composition d'un objet, reflétées dans le coefficient d'atténuation. Par exemple, pour un détecteur numérique à 5 bits, 32 niveaux de gris sont disponibles, tandis que pour un détecteur à 10 bits, les niveaux de gris sont étirés à 1024 gradations distinctes. Les variations de contraste contenues dans une image de transmission représentent le changement relatif d'atténuation entre une partie d'un support et une autre. Les caractéristiques épaisses, à haute densité et à Z élevé produisent une réduction du flux de

photons plus importante que les objets minces, de faible densité et à faible nombre atomique. Ainsi, dans certaines limites de détection, il est possible de mesurer quantitativement ces caractéristiques et de développer des corrélations reflétant un seuil de décision. Une limitation particulière de cette absorptiométrie lorsqu'on examine la présence de constituants à faible Z dans des explosifs, par exemple, est la variation limitée de la densité physique, ρ , de ces substances.

1.5.1 Défi actuel : les explosifs artisanaux et leurs précurseurs

La variation de la densité des explosifs artisanaux (HomeMade Explosives HME) dépend souvent de manière sensible du procédé de fabrication. Les fabricants de bombes illégales semblent bien comprendre le fonctionnement des équipements de contrôle aux rayons X et dans la mesure du possible, ils chercheraient probablement à adapter leur technologie afin que le dispositif devienne indétectable, en particulier en ce qui concerne les systèmes d'imagerie de transmission de rayons X familiaux.

Le peroxyde d'hydrogène (densité pour une solution à 30%, 1.110 g cm^{-3} à 20°C ; densité pour H_2O_2 pur, 1.450 g cm^{-3} à 20°C) semble être le précurseur de choix à l'heure actuelle, ayant également été l'ingrédient principal des attentats à la bombe perpétrés à Londres le 7 juillet 2005, qui ont coûté la vie à 52 passagers d'un métro et de bus. Il a été rapporté que les ingrédients de fabrication de bombe étaient plus faciles à acheter que les pilules en grand nombre [27]. À la suite du complot transatlantique de Londres de 2006, dans lequel le peroxyde d'hydrogène aurait été mélangé pendant le vol avec un carburant tel que l'acétone afin de créer un mélange explosif, les passagers ont dû retirer les liquides de leurs bagages à main avant de monter à bord.

Compte tenu des observations qui précèdent, plusieurs solutions ont été proposées pour relever le défi de l'identification des liquides aux fins d'inspection de sécurité, notamment la diffusion Raman, la résonance quadripolaire nucléaire et la spectrométrie de mobilité ionique. Malheureusement, ces approches ne sont généralement pas applicables et peuvent échouer lorsqu'elles sont utilisées, par exemple, sur des flacons sous vide hermétiquement fermés. Il est clairement souhaitable que l'introduction d'une solution technologique permette à l'utilisateur final de conserver les tactiques, techniques et procédures actuelles sans modifier

de manière importante ses pratiques actuelles. Avant d'examiner certaines solutions possibles pour les liquides, nous examinons ci-dessous les pratiques prédominantes en matière d'imagerie par rayons X.

Tableau 2 : Sélection des explosifs et des milieux explosifs précurseurs, classés par numéro atomique effectif, Z_{eff} , pour la plage d'énergie de pénétration des rayons X présentant un intérêt, ainsi que par le nombre d'électrons contributeurs par atome de C, H, N, O, nombre total d'électrons par molécule, densité moléculaire et masse, ρ_{rel} , par rapport à l'eau. Sont également inclus un certain nombre de supports inoffensifs pour lesquels la discrimination par l'utilisation de l'absorptiométrie par rayons X peut être contestée [28].

	La nature	C	H	N	O	n_e	Z_{eff}	ρ_{rel}
Polyethylene	Polymère	2	4			16	55	-0.9
PMMA	Polymère	5	8		2	54	65	118
Acetone peroxyde: TATP (dimer)	Explosive	6	12		4	80	67	07
Acetone peroxide: TATP (trimer)	Explosive	9	18		6	120	67	
HMTD	Explosive	6	12	2	4	94	68	09
Sorbitol	Sucre	6	14		6	98	69	15
Uric acid (C ₅ H ₄ N ₄ O ₃)	Précurseur	5	4	4	3	86	69	19
Sugar(Saccharose)	Sucre	12	22		11	182	69	16
Fructose	Sucre	6	12		6	96	70	15
Oetrin	Polymère	1	2		1	16	70	14
TNT	Explosive	7	5	3	6	116	71	17
Ammonium picrate (Ounnfte)	Explosive	6	6	4	7	126	71	17
Nitromethane	Explosive	1	3	1	2	32	71	11
Cydonfte (ROX)	Explosive	3	6	6	6	114	72	18
HMX (Octogen)	Explosive	4	8	8	8	152	72	19
Nitrocellulose	Explosive	6	7	3	11	152	73	13
PETN (Penthrite)	Explosive	5	8	4	12	162	74	18

Ammonium nitrate	Explosive	4	2	3	42	74	18
H ₂ O	Eau	2		1	10	74	10
H ₂ O ₂ ,	Précurseur	2		2	18	77	15
H ₂ O ₂ (1/3) + H ₂ O (2/3)	Précurseur						
Black Powder	Carbonifère						09
Chlorate powder (ClCO ₃ ,)	Sel inorganique						145
C4 (ROX + Liant)	Explosive						14
Semtex	Explosive						148
Dynamite	Explosive						14

1.6 Système de technologies de détection à rayons X actuel

Les méthodes de détection par rayons X sont le moyen le plus courant d'inspecter les bagages dans les aéroports. Il y a plusieurs raisons techniques :

- 1) La technologie des rayons X permet de déterminer certaines caractéristiques importantes des matériaux d'intérêt. Les informations les plus utiles que la technologie des rayons X peuvent fournir, sont liées à la densité d'un objet et à son numéro atomique effectif (Zeff). Le terme "numéro atomique effectif" est le numéro atomique de l'élément unique hypothétique qui donne la même atténuation des rayons X qu'un composé ou un mélange en cours de mesure [29]. Une radiographie peut fournir d'autres valeurs caractéristiques sur un objet, mais aucune de ces valeurs n'est aussi efficace que ces deux informations en termes de caractérisation des matériaux. Théoriquement, le type de matériau d'un objet peut être déterminé de manière unique en utilisant la densité et Zeff [30].
- 2) La technologie à rayons X est développée depuis près d'un siècle.
- 3) La technologie à rayons X offre davantage de sécurité aux êtres humains, car elle représente le contenu des sacs de bagages que certaines autres technologies ne peuvent pas les représenter, comme la résonance magnétique nucléaire.

4) La physique des rayons X est bien comprise ; La technologie à rayons X est également très sophistiquée en termes de source de rayons X, de dispositif électronique et de détecteur.

Le système EDS à rayons X peut considérablement améliorer la capacité des compagnies aériennes à détecter les explosifs dissimulés avant leur embarquement dans un avion. Alors que la plupart de ces technologies sont encore en développement, un certain nombre de périphériques sont maintenant disponibles dans le commerce. Cependant, aucun des appareils disponibles dans le commerce n'est illimité. Le système EDS basé sur des techniques de scanographie, d'une valeur d'environ un million de dollars chacun, présente une probabilité élevée de détection de différents matériaux explosifs avec une gamme de configurations et un taux relativement faible des fausses alarmes, mais son débit n'est pas suffisant et c'est très coûteux. Les systèmes de balayage à double énergie, décrit plus en détail dans ce qui suit, ne détectent pas les quantités et les configurations de la gamme complète d'explosifs spécifiée dans les normes, mais ils ont un débit de balayage élevé et sont relativement bon marché. Les systèmes d'imagerie par rayons X avec détection de dispersion présentent généralement des défauts similaires à ceux des systèmes à double énergie. Cependant, un tel système donne des informations relatives à la densité et à un débit plus élevé que les systèmes de scanographie.

Voici les observations générales sur les technologies de détection de rayons X qui ont des implications importantes pour leur utilisation dans les aéroports [31]:

- Premièrement, ces dispositifs varient dans leur capacité à détecter les types, les quantités et les formes d'explosifs.
- Deuxièmement, les dispositifs de détection d'explosifs produisent généralement un certain nombre des fausses alarmes qui doivent être résolues par une intervention humaine ou par des moyens techniques. Ces fausses alarmes se produisent parce que les dispositifs utilisent diverses technologies pour identifier des caractéristiques, telles que les formes, les densités et autres propriétés, afin d'indiquer un potentiel explosif. Compte tenu du volume énorme de passagers, de bagages et de cargaison traités par l'aéroport, même des taux de fausse alarme relativement modestes pourraient faire en sorte que plusieurs centaines, voire des milliers d'objets par jour nécessitent un examen plus approfondi.

- Troisièmement, et le plus important, ces dispositifs dépendent en fin de compte de l'être humain pour résoudre les alarmes. Cette activité peut aller d'une inspection plus minutieuse d'une image radiographique et d'un appel du jugement à une recherche manuelle de l'objet en question. La détection ultime des explosifs dépend des mesures supplémentaires prises par le personnel de sécurité et de son jugement correct pour déterminer la présence d'un explosif. Étant donné que bon nombre des alarmes des dispositifs indiquent uniquement la présence potentielle d'explosifs, la détection des explosifs nécessite une intervention humaine. Plus le taux de fausse alarme est élevé, plus le système doit être basé sur le jugement humain. Cette confiance pourrait constituer un maillon faible du processus de détection des explosifs. De plus, l'appui sur des jugements humains a des implications pour la sélection et la formation des opérateurs pour les nouveaux équipements.
- Quatrièmement, bien que ces dispositifs puissent augmenter considérablement la probabilité de découvrir un explosif, leurs performances sur le terrain risquent de ne pas être aussi performantes que dans les tests de laboratoire. Par exemple, le système CTX5000 certifié par la FAA n'a pas obtenu d'aussi bons résultats lors des tests opérationnels effectués dans deux aéroports que lors des tests de certification de la FAA [31]. La nécessité de s'appuyer sur les opérateurs pour résoudre les fausses alarmes existe toujours.

Un problème majeur dans ce type de système est le taux de fausse alarme. Il est causé par trois facteurs :

- Premièrement, certains matériaux ne sont pas séparables des systèmes basés sur des méthodes d'atténuation des rayons X, conçues pour distinguer les matériaux en fonction de leur densité physique et de leur numéro atomique effectif. Dans la reconnaissance statistique des modèles, le terme «non séparable» signifie qu'il est impossible de trouver une fonction de décision qui distingue exactement une classe des autres. Sinon, on dit qu'ils sont séparables. En effet, de nombreux matériaux possèdent des profils d'atténuation similaires à ceux des explosifs. Par exemple, dans le domaine (Zeff, densité), certains plastiques et caoutchoucs tombent dans la zone des poudres noires et

les autres dans la zone des poudres sans fumée [32]. Si toutes les poudres noires et toutes les poudres sans fumée sont menacées par notre système d'alarme automatisé, nous déclencherions une fausse alarme pour certains plastiques et caoutchoucs. Sinon, nous manquerions la détection de certains explosifs.

- Deuxièmement, la caractérisation d'un matériau donné peut s'étendre sur un plus grand nombre de fonctions en raison de la méthode réellement utilisée. Par exemple, la plupart des explosifs peuvent être séparés des matières inoffensives en utilisant des informations bidimensionnelles (Zeff, densité). Mais pour obtenir Zeff et la densité, il est nécessaire de connaître exactement la composition du matériau, les fractions de poids et ses informations géométriques pour chaque objet dans le sac de bagage. La vision technologique n'a pas été prouvée qu'elle pourrait effectivement possible d'obtenir une telle information en utilisant le système de radiographie à double énergie moderne. Dans de telles circonstances, certains paramètres extrinsèques, tels que l'épaisseur et l'orientation de ce matériau, auraient une incidence négative sur la caractérisation du matériau.
- Troisièmement, il existe des erreurs dans tout système de détection. Les erreurs dans un système modifient la distribution statistique de toute classe. Il ne fait aucun doute que cela diminue les performances de classification du système en augmentant le taux de fausses alertes ou en diminuant le taux de détection.

1.7 Systèmes d'imagerie par rayons X

Il existe des avantages ou des inconvénients particuliers des deux approches prédominantes de l'imagerie par rayons X menant à la formation d'images 2D ou 3D. L'approche de transmission 2D classique fournit clairement de nombreuses informations sur les caractéristiques particulières d'un objet donné, notamment la position, l'orientation et le volume, mais souffre indéniablement du fait que des caractéristiques sous-jacentes peuvent masquer la présence de menaces potentielles. En d'autres termes, une caractéristique faible d'un plan peut être masquée par une caractéristique forte d'un autre. De plus, une radiographie à transmission ne donne aucune information sur la profondeur d'une caractéristique. Par

conséquent, à première vue, il semblerait qu'il soit absolument nécessaire d'obtenir l'image d'une coupe isolée à l'aide de techniques permettant de la disséquer efficacement des couches potentiellement obscurcissantes, au-dessus ou au-dessous. Cela amène à l'idée d'exploiter ce que l'on appelle maintenant la tomosynthèse ou simplement la tomодensitométrie. Cependant, la tomодensitométrie conventionnelle à rayons X présente l'inconvénient de ne pas être spécifique à un matériau et ne permet pas non plus de distinguer les liquides des solides. La réalité est que les informations supplémentaires obtenues peuvent ne pas correspondre aux dépenses et aux efforts nécessaires pour capturer l'image.

Pour un examen détaillé de l'évolution des techniques d'imagerie par rayons X au cours des dernières décennies, le lecteur intéressé est invité à consulter les travaux de [33], [34] se concentrant sur les applications industrielles et [35] sur la détection des explosifs.

1.7.1 Systèmes de dépistage par rayons X conventionnels

Les systèmes classiques de criblage aux rayons X mesurent l'énergie atténuée des rayons X après qu'elle a traversé un objet numérisé. Ils ont été introduits pour lutter contre le détournement en 1970. Les appareils ont été conçus pour fournir des images hautes résolution avec une bonne plage dynamique de niveaux de gris afin que l'opérateur de sécurité puisse identifier des armes, telles que des fusils et des couteaux, fabriqués dans des métaux présentant des formes caractéristiques.

Certains des premiers appareils à rayons X proposés en tant que systèmes de détection d'explosifs consistaient pour la plupart en une adaptation des technologies existantes aux problèmes rencontrés dans l'EDS, de sorte que ces appareils pouvaient créer des explosifs évidents fabriqués à partir de matériaux légers sans forme caractéristique. Deux exemples d'images sont donnés aux Figure 3 et Figure 4.

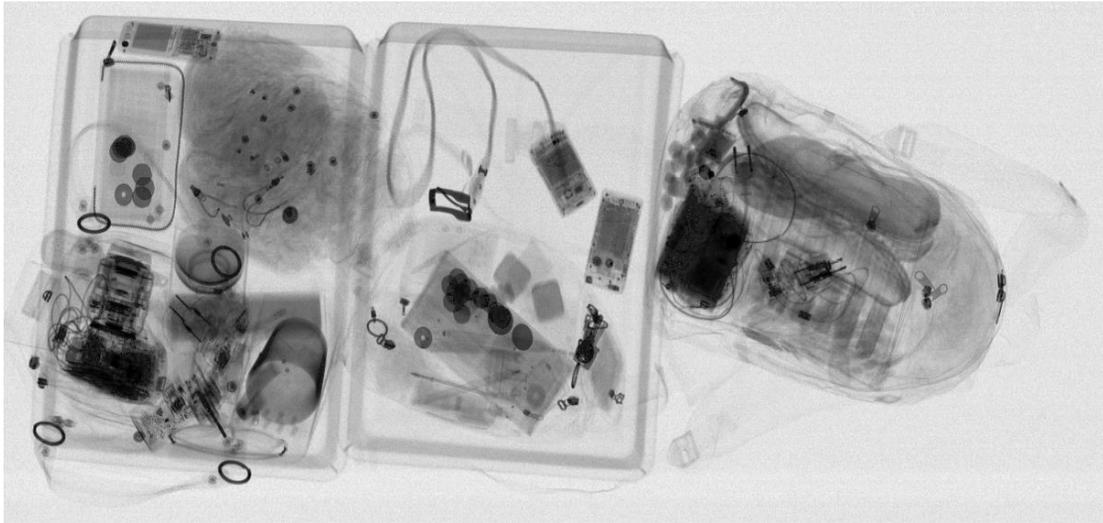


Figure 3 : Exemple 1 d'une image fournie par un système dépitage par rayons X conventionnels monoénergétiques.

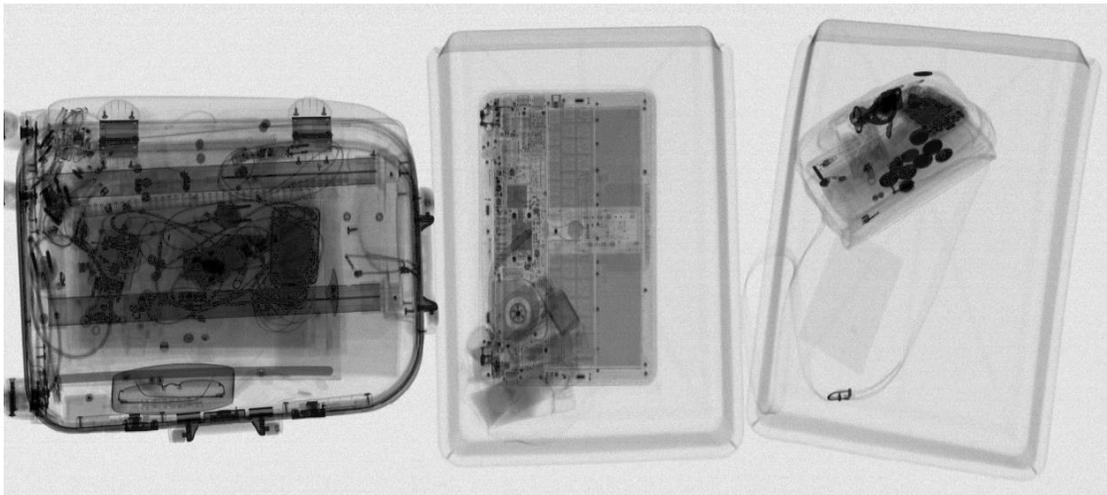


Figure 4 : Exemple 2 d'une image fournie par un système dépitage par rayons X conventionnels monoénergétiques.

1.7.2 Systèmes d'imagerie biénergie à rayons X

Certains systèmes d'imagerie à rayons X commerciaux dans les aéroports intègrent désormais une analyse à double énergie permettant d'estimer les numéros atomiques des matériaux contenus dans les bagages, une information essentielle pour l'opérateur de sécurité [36], un exemple d'image est donné dans la Figure 5.

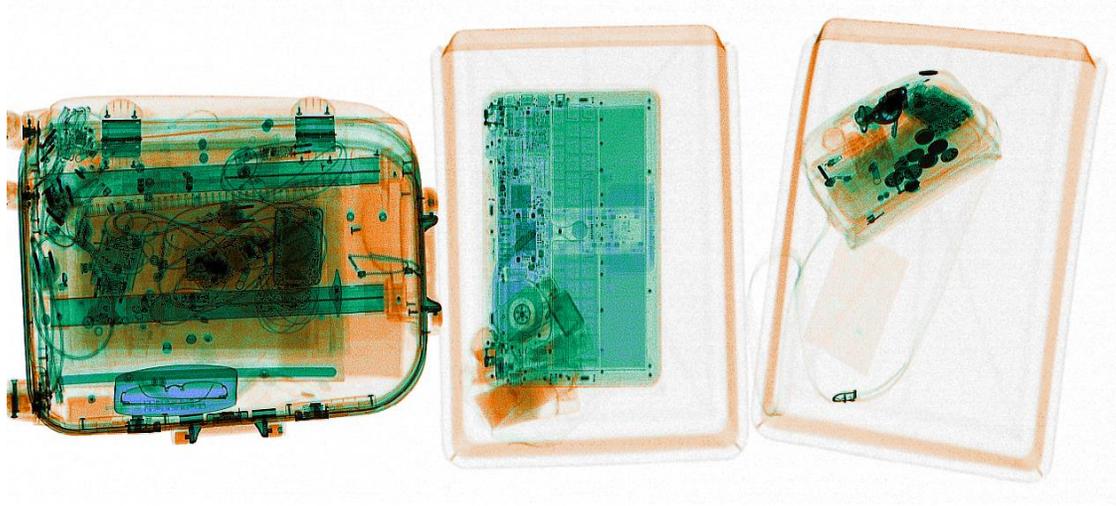


Figure 5 : Exemple d'image fournie par un système de dépistage par rayons X à double énergie.

Pour voir la puissance et les limites de la méthode, le signal d'atténuation logarithmique obtenu à l'énergie des rayons X E_1 est donné dans l'équation suivante [37],

$$\ln\left(\frac{I_{01}}{I_1}\right) = \sigma(Z, E_1) \frac{N_A}{A} \rho t \quad (1.7)$$

Il existe trop d'inconnues, telles que la section efficace σ , la densité ρ et l'épaisseur t , dans l'équation précédente pour obtenir beaucoup d'informations à partir d'une seule mesure. En particulier, un matériau mince à haut Z aura le même signal d'atténuation qu'un matériau épais à faible Z . Cependant, deux mesures obtenues avec deux énergies différentes peuvent révéler une partie du mystère d'un objet unique.

À la deuxième énergie de rayons X E_2 , nous avons également

$$\ln\left(\frac{I_{02}}{I_2}\right) = \sigma(Z, E_2) \frac{N_A}{A} \rho t \quad (1.8)$$

La division des deux dernières équations élimine la densité et l'épaisseur, et s'appelle double énergie qui peut être utilisée pour calculer le numéro atomique,

$$\frac{\ln\left(\frac{I_{01}}{I_1}\right)}{\ln\left(\frac{I_{02}}{I_2}\right)} = \frac{\sigma(Z, E_1)}{\sigma(Z, E_2)} \quad (1.9)$$

La méthode de double énergie appliquée à un objet simple donne également la densité surfacique ρt (g/cm^2) en utilisant l'équation (1.7) ou (1.8), et donne à son tour une mesure de la densité et de l'épaisseur en utilisant d'abord une connaissance a priori entre numéro atomique et densité. Mais les objets simples auxquels s'applique l'équation (1.9) sont rarement vus aux portes des aéroports. Le contenu des bagages varie énormément et n'est emballé que de temps en temps. Par conséquent, la méthode à double énergie appliquée au bagage réel ne donne plus le numéro atomique de l'objet, mais un numéro atomique effectif Z_{eff} , et la densité réelle est mal connue. Le terme «numéro atomique effectif» est le numéro atomique de cet élément unique hypothétique qui donne la même atténuation des rayons X qu'un composé ou un mélange en cours de mesure.

Dans le monde réel, l'analyse simple à double énergie, bien qu'elle soit avantageuse par rapport à l'analyse à énergie unique, est facilement confondue et peut difficilement être utilisée pour construire l'EDS automatisé. Pour améliorer la méthode, il faut séparer les objets dans une image complexe. Plusieurs approches ont été vigoureusement explorées [37], telles que la transmission avec diffusion et la tomographie par ordinateur.

1.7.3 Système d'imagerie par rayons X avec dispersion

Les deux méthodes ci-dessus utilisent uniquement des techniques d'imagerie par transmission. Une autre technique utilisée par les systèmes avancés d'inspection par rayons X des bagages est la collecte de données d'images à partir d'énergie de rayons X dispersée [37], [38], [39]. Les données d'image peuvent être collectées à partir de l'énergie diffusée des rayons X qui est réfléchi vers la source de rayons X ou de l'énergie diffusée des rayons X qui est passée à travers l'objet balayé, appelées respectivement images à rétrodiffusion et à diffusion vers l'avant. Les systèmes qui collectent des données d'images de rayons X dispersées collectent également généralement des images de transmission conventionnelles.

Deux caractéristiques des explosifs plastiques les rendent uniques dans leurs interactions avec les rayons X. Premièrement, ils sont constitués d'éléments ayant de faibles nombres atomiques. Deuxièmement, leur densité physique est environ 30 à 50% plus élevée que celle des plastiques ordinaires et cinq à huit fois plus élevée que celle des vêtements [39]. Alors que les méthodes

à double énergie tentent d'exploiter la première de ces caractéristiques, une image en dispersion exploite à la fois les caractéristiques de faible Z et de haute densité des explosifs plastiques.

L'imagerie en transmission standard mesure le nombre de photons de rayons X retirés d'un faisceau à chaque emplacement du sac ; peu importe comment la radiographie est retirée du faisceau de crayon (pencilbeam). Par conséquent, la section transversale atomique qui détermine l'atténuation en un emplacement quelconque des images de transmission est la section transversale interactive totale des rayons X est la somme des sections efficaces photoélectriques et de diffusion.

Cependant, une image de dispersion est créée d'une manière différente et reflète les sections transversales de l'interaction différemment. Chaque emplacement le long du faisceau de crayon dans les objets ajoute à la quantité de dispersion vue dans l'image de dispersion. La quantité dépend du nombre de photons de rayons X atteignant l'emplacement, de la densité électronique à cet emplacement, de la section transversale de diffusion et de l'atténuation subie par les photons diffusés à la sortie du sac. Par conséquent, si une image de transmission d'objets dans un sac montre que deux matériaux ont la même atténuation totale, ils pourraient tout de même avoir une quantité de dispersion très différente.

Une technique d'utilisation des images en dispersion a été présentée dans [39], qui utilise uniquement l'histogramme en dispersion pour définir la menace. La raison en est que si un sac contient une bombe compacte, la bombe apparaîtrait dans l'image comme une région de dispersion intense. Dans l'histogramme, il apparaîtrait une augmentation significative du nombre de pixels au niveau de dispersion élevée.

1.7.4 Système de tomographie par ordinateur (CT)

Dans les années 1970, l'utilisation des ordinateurs en association avec des compteurs à ionisation ou à scintillation a marqué le début d'une nouvelle ère pour la radiographie à rayons X, particulièrement adaptée à l'imagerie de supports à faible nombre atomique (en particulier pour l'imagerie des tissus mous en clinique, puis dans du matériel industriel). Au niveau essentiel, un faisceau étroit traverse l'objet étudié et traverse un collimateur de sortie soigneusement aligné pour lui permettre d'être reçu par un détecteur. Le faisceau du crayon est conçu pour être mis en rotation autour de l'objet, le faisceau transmis tombant sur un réseau de

détecteurs, surmontant ainsi le problème de superposition et permettant de détecter les variations du coefficient d'atténuation linéaire dans les milieux à faible numéro atomique de 0,05% (Figure 6).

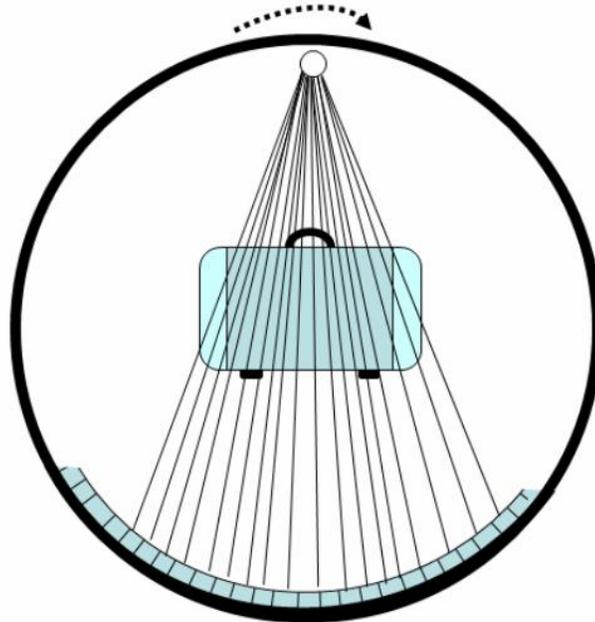


Figure 6 : Schéma du tomographe CT. De nombreuses sources de rayons X autour de la circonférence du dispositif sont activées séquentiellement pour produire un faisceau de rayonnement en éventail qui tourne efficacement autour d'un axe perpendiculaire sur le plan de la thèse entourée d'un anneau de détecteurs afin de permettre des vitesses de balayage rapides.

La tomographie par ordinateur, initialement développée pour obtenir des images tridimensionnelles d'êtres humains, est maintenant appliquée à la surveillance des bagages [40], [41], [42]. Comme indiqué précédemment, la transmission des rayons X à travers un objet est fonction de la composition du matériau (numéro atomique effectif), de la densité et de l'épaisseur. Pour chaque trajet de rayons X, le coefficient d'atténuation linéaire total est la somme des coefficients d'atténuation de chaque élément du trajet. CT produit des images en coupe transversale d'un objet en reconstruisant une matrice de coefficients d'atténuation des rayons X. Ces images peuvent être rendues en volume pour fournir des informations en trois dimensions. Dans les images rendues en coupe ou en volume, l'algorithme de reconstruction détermine le coefficient d'atténuation linéaire pour chaque élément de volume de l'image.

Comme la dimension de l'élément de volume est connue par la géométrie du système de balayage, le coefficient d'atténuation devient fonction de deux variables seulement, le nombre atomique effectif et la densité. La séparation de ces variables est mise en œuvre en prenant des données de tomодensitométrie à deux énergies de rayons X et en résolvant essentiellement deux équations à deux inconnues. Ainsi, les éléments de volume peuvent être mappés en fonction de Z_{eff} et de la densité, et la corrélation de ces variables peut fournir à la fois une identification des caractéristiques et des matériaux [43].

Crawford et Martz [34] ont observé que les tomодensitomètres utilisés pour détecter les explosifs dans les bagages enregistrés et les bagages à main sont généralement très similaires aux tomодensitomètres multicoupes à une et deux énergies d'imagerie médicale. Parmi les dispositifs récemment développés pour la détection des explosifs, certains utilisent de multiples sources et réseaux de détecteurs pour éliminer la rotation mécanique d'un portique, ainsi que des détecteurs à comptage de photons pour une imagerie spectrale. Crawford et Martz ajoutent que la nécessité d'un nombre plus limité de vues réduit les coûts.

CT fournit la distribution de densité résolue spatialement comme information principale. Ainsi, à la fois la densité et la forme/volume sont des caractéristiques utiles pour discriminer les matières inoffensives des menaces.

1.7.5 Techniques de rétrodiffusion

Dans cette technique, les photons rétrodiffusés Compton sont observés à l'aide d'un réseau de détecteurs placé du même côté de l'objet étudié que la source [44] (voir Figure 7). Les événements rétrodiffusés enregistrés révèlent effectivement la différence entre les sections transversales de diffusion Compton du support dominant et les objets numérisés qu'il contient. Ceci peut être utilisé pour reproduire la forme de tels détails. Parmi les avantages de la technique, il y a la réduction de la superposition et la possibilité d'imager des objets dans des situations où le côté éloigné d'un objet est inaccessible.

Pour un atténuateur donné, l'intensité des photons rétrodiffusés Compton diminue avec l'augmentation de l'énergie des photons et, pour cette raison, les applications pratiques des mesures de rétrodiffusion de Compton ont été limitées. La géométrie dicte en outre que la plus

grande sensibilité de détection est obtenue au niveau de la surface ou de la surface par rapport à un matériau plus profondément incrusté dans un emballage.

La technique de rétrodiffusion est une technique offrant la forme et la distribution spatiale des caractéristiques, étant uniquement dépendante de la densité électronique.

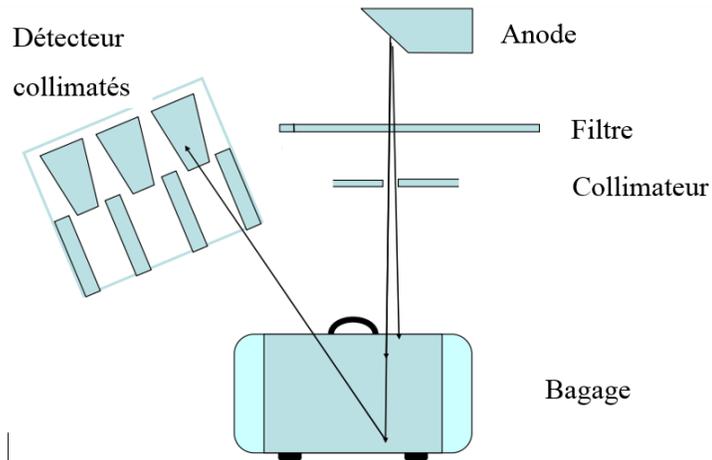


Figure 7 : Schéma d'une situation de géométrie en rétrodiffusion.

1.8 Les problèmes opérationnels

1.8.1 Le processus d'inspection

Il existe des limites bien connues concernant la taille des bagages de cabine, souvent exprimées en pouces linéaires (somme de la hauteur, de la largeur et de la longueur), généralement définies à environ 45 pouces linéaires. Cependant, il est moins connu que des limites physiques sont imposées aux bagages enregistrés.

Il existe certes de légères variations entre les vols domestiques et les vols internationaux, mais la limite standard est de 62 pouces linéaires et une limite de poids supérieure à 50 lbs (environ 20 kg), en fonction du transporteur et du type de billet émis. Les bagages en dehors de ces limites peuvent être interdits de passage ou être passibles des pénalités pour excédent de bagages [45].

Les matériaux de construction utilisés dans un sac de voyage typique peuvent être dérivés de plastique, de tissu, de cuir, de tissus synthétiques ou naturels, d'armatures métalliques ou d'un

mélange de ceux-ci. Toutefois, pour le personnel chargé de l'inspection, le contenu du sac revêt une importance primordiale, car, associé au matériel et à la construction des sacs et des bagages, crée une scène de fond dans laquelle des cibles potentielles peuvent être présentes. À cet égard, il est pratiquement impossible de faire des hypothèses sur le contenu des bagages de cabine. Hormis les divers objets interdits tels que des armes (armes à feu), des munitions, des denrées alimentaires diverses, du matériel sensible (par exemple de la pornographie) que différentes agences frontalières appliquent, on ne peut faire aucune hypothèse sur le contenu des bagages de cabine. Cela a été exploité par des groupes terroristes, de sorte que les explosifs peuvent souvent sembler, à la suite d'une inspection par rayons X, ressembler à des objets bénins tous les jours. Par exemple, la dynamite peut avoir l'apparence d'une barre de massépain. La communauté internationale de contrôle des bagages CEAC a adopté un processus de contrôle à cinq niveaux, résumé dans le Tableau 3.

Tableau 3 : Les règles de sélection de la CEAC (communauté internationale de contrôle des bagages) [46].

Niveau	Action
1	Évaluation de l'image radiographique par la machine à rayons X.
2	Analyse de l'opérateur de l'image de niveau 1 sur un ou plusieurs postes de travail, réalisée pendant le transport du sac.
3	Analyse plus approfondie de l'image de niveau 1 d'origine sur un ou plusieurs postes de travail ou soumission du sac à un processus de radiographie distinct utilisant une technologie de radiographie différente, par exemple, la tomographie par ordinateur (CT).
4	Rassembler le passager et le sac et effectuer une recherche manuelle.
5	Si le passager est introuvable, le sac est considéré comme une alerte à la bombe et traité en conséquence.

1.8.2 Objets interdits dans les bagages de cabine ou les bagages enregistrés dans le cadre juridique

Bien que sur le plan international, on puisse s'attendre à des variations mineures dans ces éléments, les listes les mieux connues sont probablement celles de la US Transport Security Administration (TSA), détaillées et régulièrement mises à jour en ce qui concerne les articles autorisés et interdits [47]. Les niveaux actuels de contrôle résultent de l'adoption de H.R. 3150 (Helicopter Fuel System Safety Act), loi de 2001 sur la sécurisation des transports pour l'Amérique. En vertu de cette loi, le contrôle de sécurité des compagnies aériennes est fédéralisé dans le cadre d'une nouvelle «Transportation Security Administration» (TSA) au sein du département américain des transports. La TSA doit : 1) adopter des normes plus strictes pour le contrôle des passagers et des bagages des compagnies aériennes ; 2) procéder à une vérification approfondie des antécédents de tout le personnel de sécurité de l'aéroport ; 3) assurer la formation et la mise à l'essai de tout le personnel de sécurité de l'aéroport ; et 4) nommer des fonctionnaires fédéraux en uniforme pour superviser le contrôle de tous les passagers et de tous les bagages dans les aéroports.

Dans les aéroports de l'UE, la sécurité de l'aviation civile est régie par les règlements (CE) n°300/2008 [48] et n°185/2010 [49], entrés en vigueur le 29 avril 2010. L'annexe du règlement (CE) n°185/2010 contient une liste des articles dont le transport est interdit dans la soute d'un avion. Cette liste d'objets interdits est accessible au public et comprend les substances et engins explosifs et incendiaires susceptibles d'être utilisés pour causer des blessures graves ou constituer une menace pour la sécurité des avions. Parmi les articles énumérés comme interdits, être transporté dans la soute d'un avion est les munitions. Toutefois, conformément au point 5.4.2 de l'annexe du règlement (CE) n°185/2010, des dérogations peuvent être accordées à condition que l'autorité compétente dispose de règles nationales permettant le transport de l'objet et que les règles de sécurité applicables soient respectées.

Singh *et al.* [50] ont mis en lumière le fait que le contenu des bagages se divise en deux grandes catégories : les objets solides (ou liquides) et les objets textiles (vêtements, objets à base de tissu). En termes d'identification aisée des risques potentiels liés aux objets bénins, que ce soit

par la machine ou par l'opérateur, il en découle un certain nombre de problèmes : le textile peut apparaître (au moins en projection 2D) autour d'un objet solide.

1.8.3 Considérations opérationnelles pour l'inspection des bagages

1.8.3.1 Projection avant / après le 11 septembre

Jusqu'en 2001, les conceptions relatives à l'inspection des bagages aux États-Unis et en Europe occidentale divergeaient radicalement - les politiques s'étaient développées à la suite des types de menaces très différents apparus : seuls 5% des bagages enregistrés aux États-Unis étaient effectivement contrôlés. Avant le 11 septembre, les entreprises de sécurité privées employaient 162000 contrôleurs de bagages effectuaient les bagages enregistrés, ce qui ne prenait que très peu de temps (12 heures) de formation avant de commencer les travaux de contrôle officiels. De plus, ces entreprises versaient généralement des salaires médiocres (cité dans un rapport comme étant moins que de travailler dans un établissement de restauration rapide [51]), ce qui avait pour effet de faire baisser le moral des employés et de favoriser le roulement de personnel. Le processus d'inspection signifiait que le nombre de contrôleurs était passé à 56000 (mais ensuite réduit à 46000), et que le personnel chargé d'inspection disposait généralement de plus de 100 heures de formation [52]. Il suffit généralement de 3 à 6 secondes à l'utilisateur pour prendre une décision sur une seule image de bagage de rayons X [53], la nécessité d'une formation poussée va de soi.

Après le 11 septembre, l'un des principaux changements introduits par la nouvelle TSA était que 100% de tous les bagages de soute seraient soumis à une analyse des éléments de menace, en particulier les engins explosifs improvisés (IEDs) [54]. À l'époque, la technologie de contrôle semi-automatique disponible n'était capable de contrôler que 150 à 200 bagages à l'heure avec une FPR (Taux de fausses alarmes) relativement élevée de 30% [52], ce qui entraînait des retards importants pour les passagers [54]. Ces investissements en matière de contrôle ont également inévitablement entraîné des retards dans les transports aériens nationaux et internationaux. Blalock *et al.* [52] ont constaté que cela réduirait le volume de passagers de 5% (et par la suite un taux de mortalité imputable sur la route civile plus élevé en raison de

choix de transport), fournissant une forte motivation pour des technologies de débit améliorées ainsi que des méthodes de dépistage efficaces et efficientes.

Plusieurs tests différents peuvent être utilisés pour les agents de contrôle, notamment le test de reconnaissance d'objet et le test d'élément interdit (PIT). Le PIT émis pour établir si un examinateur comprend quels éléments sont interdits et quels éléments sont acceptés. Il consiste en un ensemble d'images de couteaux, d'objets tranchants et d'explosifs, et est conçu pour évaluer les connaissances de l'ensemble des compétences d'un examinateur. En revanche, l'ORT (Object Recognition Test ou le test de reconnaissance d'objet) est conçu pour évaluer les capacités de détection d'un cribleur à la complexité du sac. Ces facteurs ont été largement associés à trois facteurs principaux basés sur l'image : la complexité, la superposition et le point de vue de la cible.

1.8.3.2 Formation des agents d'inspection

Dans le cadre du développement en cours des processus de formation au contrôle des bagages, plusieurs types de tests et d'objets de test ont été développés aux États-Unis et en Europe et ont suscité une activité importante dans le développement et la compréhension de la psychophysique du processus de détection des bagages.

Aux États-Unis, le système SPEARS (système d'évaluation des compétences en matière d'évaluation et de compte rendu) avait déjà été mis à l'essai dans 19 aéroports américains avant les attentats du 11 septembre [55]. SPEARS comportait un certain nombre de tests de perception pour identifier les personnes ayant une aptitude à la tâche de filtrage et pour aider à former ceux qui allaient devenir agents d'inspection des bagages. Le système SPEARS fonctionnait visuellement de la même manière que l'actuel appareil de radiographie, ou XRT (X-Ray Tuto), en ce sens qu'il superpose des objets de menace dans des images de bagage bénignes, avec des niveaux de difficulté croissants. En 1998, la FAA a annoncé qu'elle utiliserait SPEARS pour la formation et qu'elle pourrait être utilisée par les compagnies aériennes pour faciliter le recrutement de personnel de contrôle qualifié. Le système Threat Image Projection (TIP) [50] a permis de projeter de manière aléatoire des objets de menace à partir d'une base de données volumineuse de plusieurs milliers. Il était prévu que TIP soit une fonctionnalité standard intégrée à tous les systèmes d'inspection de bagages X-ray utilisés dans

les aéroports américains. L'une des méthodes standard d'évaluation de la compétence d'inspection de bagages aux États-Unis consistait à soumettre les contrôleurs au test de compétence, dans lequel une relation linéaire a été trouvée entre le taux de réussite et le temps de décision [56].

L'étude de l'efficacité de la formation des experts en dépistage a suscité un intérêt scientifique considérable. Plusieurs études ont examiné les différences entre les performances des observateurs naïfs et non entraînés par rapport à celles d'observateurs qualifiés utilisant l'eye tracking pour analyser le comportement de l'observateur [57]. Dans les tests d'inspection simulés, il a été constaté que les agents de contrôle fonctionnaient plus efficacement que les observateurs naïfs, qu'ils analysaient et déterminaient une décision, en utilisant des stratégies de recherche plus élaborées que les individus naïfs. Liu & Gale [57] ont également constaté que les observateurs formés détectaient plus rapidement les menaces et étaient plus précis. Cependant, il a été noté que les deux groupes avaient mal détecté les engins piégés, même si l'observateur entraîné s'est concentré sur la menace, mais l'a ignoré invariablement, soulignant la nécessité d'une formation appropriée sur cette menace, la plus difficile à détecter. Ceci était également corroboré par d'autres travaux antérieurs [56], qui montraient des réductions similaires lorsque la tâche de détection devenait plus difficile (c'est-à-dire que la complexité de l'image des bagages augmentait).

Il existe trois facteurs de base basés sur l'image qui régiront les performances d'un individu : le point de vue ou l'orientation de la menace par rapport à l'observateur, la superposition (d'autres objets sur / sous la menace) et la complexité du dépistage des rayons X [58]. Les tests effectués avec le personnel des aéroports européens suggèrent, sans surprise, de fortes corrélations entre les performances de la XRT (X-Ray Tuto) et les scores obtenus lors du test d'évaluation des compétences par rayons X utilisés pour permettre aux agents d'accréditer dans plusieurs aéroports européens.

D'autres travaux [53] ont également révélé que les performances des observateurs sont considérablement améliorées lorsque la couleur appropriée est utilisée pour mettre en évidence les menaces potentielles. Cependant, les mêmes auteurs soulignent également certaines des limites du cadre théorique très idéalisé de détection des signaux utilisé pour dériver les

caractéristiques ROC : la performance du contrôle des bagages peut se dégrader en raison de nombreux facteurs de la vie réelle tels que l'état de vigilance, les problèmes de personnel ou les conditions environnementales (par exemple le bruit), qui peuvent affecter négativement les performances obtenues dans une situation de test.

1.9 Conclusion

Les machines d'inspection utilisées dans les aéroports sont généralement basées sur un système à rayons X à double énergie. Étant donné que des différents matériaux absorbent les rayons X à différents niveaux, l'image sur le moniteur permet à l'opérateur de la machine de voir des articles distincts à l'intérieur d'un bagage. Les éléments sont généralement colorés sur le moniteur d'affichage. En fonction de la plage d'énergie qui traverse l'objet, trois principales catégories sont représentées : organique, inorganique, métal. Différents modèles de scanner de bagages à rayons X ont été présentés dans ce chapitre. Nous avons également abordé le défi des explosifs artisanaux et de leurs précurseurs, ainsi que les problèmes opérationnels du processus d'inspection. Dans le chapitre suivant, nous allons nous intéresser à une revue littéraire des méthodes logicielles des processus de traitement d'image, utilisées dans le domaine de l'inspection des bagages.

Chapitre 2

État de l'art

2.1 Introduction

Le traitement des images est défini, comme étant un processus dont l'exécution a pour but de faciliter la compréhension humaine des images ainsi que pour des cas d'utilisation automatisée des images. Dans la littérature, nous avons identifié trois axes dans le traitement des images : l'amélioration de la qualité d'image, la segmentation et la compréhension d'image. Dans ce chapitre d'état de l'art, nombreux travaux de recherche de ces différents axes, sont présentés, des résumés des approches de chaque application sont présentés.

2.2 Amélioration de la qualité de l'image

Les systèmes à rayons X classiques détectent efficacement les objets de composition métallique (couteaux et armes à feu), tandis que les armes à faible densité pourraient facilement ne pas être contrôlées. Les objets tels que les pistolets et les couteaux métalliques, qui se caractérisent par des réponses à haute densité sur les images à rayons X, sont également facilement repérables. Cependant, des objets comme le plastique, le verre et les objets coupants en bois se caractérisent par des réponses très faibles et de faibles densités dans les projections de rayons X et sont très difficiles à distinguer par les agents de contrôle. La principale limite de ces dispositifs est qu'ils ne peuvent pas faire la distinction entre une mince feuille de l'absorbeur fort et une plaque épaisse de l'absorbeur faible.



Figure 8 : une image fournie par un système de radiographie X-ray à double énergie se joint au bagage d'origine.

Les systèmes à rayons X à double énergie permettent d'identifier des matériaux dans les bagages en comparant deux images des bagages obtenues à deux niveaux différents d'énergie aux rayons X (haute et basse énergie) (voir Figure 8). Lorsque les rayons X de haute énergie pénètrent dans des objets, l'absorption d'énergie dépend principalement de la densité du matériau. Plus la densité est élevée, plus l'absorption d'énergie par l'objet est élevée, et donc plus l'image est sombre. Cependant, pour les rayons X à basse énergie, l'absorption d'énergie dépend principalement du numéro atomique effectif du matériau ainsi que de l'épaisseur de l'objet. Par conséquent, les zones de matériaux à haute densité, tels que le métal, sont sombres sur les images à rayons X à basse et haute énergie, mais les zones à éléments plus clairs sont représentées comme des régions plus sombres dans les images à basse énergie par rapport aux images à haute énergie. Les systèmes commerciaux à double énergie utilisent les informations des deux images pour estimer le numéro atomique des matériaux. Théoriquement, le type de matériau d'un objet peut être déterminé de manière unique à l'aide de sa densité et de son numéro atomique. Cependant, ces dispositifs fournissent des informations médiocres sur la densité des objets et seule une estimation du nombre atomique peut être générée, le numéro atomique effectif qui représente le numéro atomique de tout objet traversé par le photon aux rayons x. En raison de ces limitations, le taux de fausses alarmes des systèmes de détection de bagages à rayons X à double énergie atteint environ 30% [59]. Selon la Federal Aviation Administration (FAA), un taux de fausse alarme acceptable est inférieur à 5%.

La tâche d'inspection dans la plupart des aéroports du monde repose toujours sur la détection manuelle d'objets potentiellement menaçants par des experts humains; pour cette raison, la formation en matière de sécurité s'appuie largement sur le test de reconnaissance d'objet (ORT) pour qualifier les contrôleurs de bagages humains dans les aéroports [60], et diverses fonctions dites «d'amélioration de l'image» facilitant la prise de décision humaine qui est un domaine reconnu de besoin critique. Une étude de l'effet des fonctions d'amélioration de l'image sur les performances de détection des rayons X est disponible dans [61]. L'objectif principal de ces fonctions d'amélioration d'image est de traiter une image de sorte que le résultat soit plus approprié que l'image originale pour une application spécifique. Dans les images à rayons X, les améliorations apportées à l'image peuvent augmenter la visibilité des objets dans le bagage et supprimer les bruits de fond.

Dans la littérature, plusieurs méthodes ont permis de progresser dans le traitement des images de bagages radiographiques. Les méthodes peuvent être regroupées en deux groupes principaux : celles conçues pour une image à rayons X conventionnelle unique et celles utilisant les informations de deux images obtenues à deux niveaux d'énergie de rayons X différents (systèmes à rayons X à double énergie) pour obtenir une image fusionnée.

Les méthodes du premier groupe améliorent la qualité d'image à l'aide de techniques traditionnelles telles que de désencombrement et de pseudo-coloration. Nercessian *et al.* [62] ont proposé une amélioration de l'image via le transfert d'histogramme par transformation logarithmique (LTHS). Le procédé améliore une image en prenant d'abord sa transformée en cosinus discrète (DCT) et en transformant les coefficients d'amplitude DCT dans le domaine logarithmique. L'image améliorée est obtenue via la transformée en cosinus discrète inverse (IDCT). Des méthodes telles que l'étirement de contraste (contrast stretching), la correction d'intensité gamma, l'ajustement logarithmique de l'intensité et l'égalisation des histogrammes sont devenues des méthodes populaires d'amélioration de l'image et ont été appliquées dans [63], [64] par Abidi *et al.* dans des images radiographiques de bagages contenant des objets de menace de faible densité. Ils ont combiné ces méthodes à des techniques de déclassement et ont obtenu une augmentation significative du taux de détection des menaces pouvant atteindre 62%, par rapport à l'image d'origine.

La deuxième catégorie basée sur la double énergie est axée sur la combinaison des informations provenant de l'image aux rayons X à basse énergie et de l'image aux rayons X à haute énergie, de sorte que l'image combinée produite soit plus appropriée pour une interprétation réussie des agents d'inspection. Les systèmes commerciaux d'inspection des bagages à rayons X à double énergie comportent une analyse à double énergie pour estimer le nombre atomique de matériaux dans les bagages fusionnant l'information d'images à basse énergie et à haute énergie. Généralement, l'image fusionnée est une image colorée où les nuances rouges représentent un matériau organique, les nuances bleues représentent un matériau inorganique et les nuances vertes représentent un mélange entre inorganique et organique. Zheng [65] a décrit une méthode de fusion d'images à rayons X à double énergie utilisant des informations spatiales. Le procédé classe d'abord chaque pixel d'une scène de bagages en pixels d'arrière-plan ou en pixels de détail, puis met en surbrillance les pixels de détail en partant du principe que les pixels de détail portent les caractéristiques d'intérêt de la scène de bagages. Zhiyu Chen *et al.* [66] ont développé une approche combinatoire basée sur la transformation en ondelettes. Le procédé transforme d'abord chaque image en images du coefficient d'ondelettes correspondantes à l'aide de la transformée en ondelettes discrète (DWT). Ensuite, suivant une règle de fusion, les coefficients d'ondelettes fusionnés sont calculés à partir des coefficients d'ondelettes des images sources. La DWT inverse (IDWT) est ensuite appliquée aux coefficients d'ondelettes pour obtenir l'image fusionnée. Étant donné que le bruit de fond est souvent amplifié pendant le processus de fusion, les auteurs appliquent une technique de réduction du bruit de fond. L'image atténuée est ensuite traitée à l'aide d'une technique d'amélioration basée sur l'histogramme pour reconstruire l'image finale.

Les systèmes d'inspection à rayons X à la pointe de la technologie offrent une variété de fonctions dites «d'amélioration de l'image» (IEF).

Ces dernières années, l'importance du contrôle par rayons X des bagages dans les aéroports a considérablement augmenté. La qualité d'image des anciens appareils de dépistage par rayons X avait parfois besoin d'être améliorée. Par exemple, une première version d'un algorithme de coloration en tant que fonction d'amélioration n'a pas pour but d'accroître les performances de détection des objets menaçants ; elle l'a en réalité altérée. Cela était dû à l'occlusion de parties

d'objet par l'algorithme de coloration opaque [67], [68]. Mais il y a eu un progrès technologique important ces dernières années, en particulier pour les appareils de dépistage aux rayons X, qui offrent aujourd'hui une qualité d'image élevée et diverses fonctions d'amélioration de l'image (IEF). L'objectif principal de ces fonctions est de traiter une image de sorte que le résultat soit plus approprié que l'image d'origine pour une application spécifique, par exemple le contrôle par rayons X dans les aéroports [69]. Dans les images à rayons X, les améliorations apportées à l'image peuvent augmenter la visibilité des objets dans le sac et supprimer les bruits de fond. Les quatorze techniques d'amélioration d'image suivantes sont utilisées dans les aéroports actuellement pour les images radiographiques sur l'écran lorsqu'elles fonctionnent sur un appareil à rayons X. Chaque pixel du format d'image utilisé dans ces machines à rayons X a un matériau et une valeur de luminance. Pour afficher les images sur un écran, les valeurs des pixels sont codées par couleur : le logiciel affiche l'image avec quatre types de couleurs selon une catégorie : orange pour la matière organique, bleu pour la matière inorganique, vert pour mixte et noir (ou rouge) pour des objets aux propriétés incertaines (ou difficiles à pénétrer). La valeur de luminance définit la luminance du pixel.

1. Image couleur / Image noir et blanc : le filtre niveaux de gris supprime les informations sur le matériau de l'image et affiche uniquement la valeur de luminance.



Figure 9 : amélioration de la qualité de l'image par le filtre niveaux de gris.

2. Amélioration locale : lorsqu'un objet est caché derrière un autre ou qu'il a des propriétés matérielles incertaines.



Figure 10 : amélioration de la qualité de l'image par amélioration locale.

3. Amélioration de la puissance de pénétration élevée : afin de modifier le contraste de l'image.



Figure 11 : amélioration de la qualité de l'image par amélioration de la puissance de pénétration élevée.

4. Amélioration de la faible puissance de pénétration : cette sélection améliore le contraste des zones les plus sombres (faible pénétration).



Figure 12 : amélioration de la qualité de l'image par amélioration de la faible puissance de pénétration.

5. Super amélioration : Cette sélection affiche les objets cachés derrière des objets épais.



Figure 13 : amélioration de la qualité de l'image par super amélioration.

6. Exclusion des produits organiques : Cette sélection active toutes les matières organiques en gris, ce qui permet à l'opérateur de voir les matières inorganiques (pistolets, couteaux, boîtes de conserve, etc.).



Figure 14 : amélioration de la qualité de l'image par exclusion des produits organiques.

7. Exclusion inorganique : Cette sélection active tous les matériaux inorganiques en gris, ce qui permet à l'opérateur de voir les matériaux organiques (drogues, explosifs, etc.).



Figure 15 : amélioration de la qualité de l'image par exclusion inorganique.

8. Négatif : cette sélection permet à l'opérateur de mettre en évidence moins d'objets absorbants rayons X, comme un fil métallique.



Figure 16 : amélioration de la qualité de l'image par le filtre négatif.

9. Éclaircissement : Cette sélection permet à l'opérateur d'éclaircir l'image.



Figure 17 : amélioration de la qualité de l'image par éclaircissement.

10. Assombrissement : cette sélection permet à l'opérateur d'assombrir l'image.



Figure 18 : amélioration de la qualité de l'image par assombrissement.

11. Balayage des niveaux de gris : Cette sélection permet à l'opérateur d'obscurcir les zones claires et d'éclaircir les zones sombres.

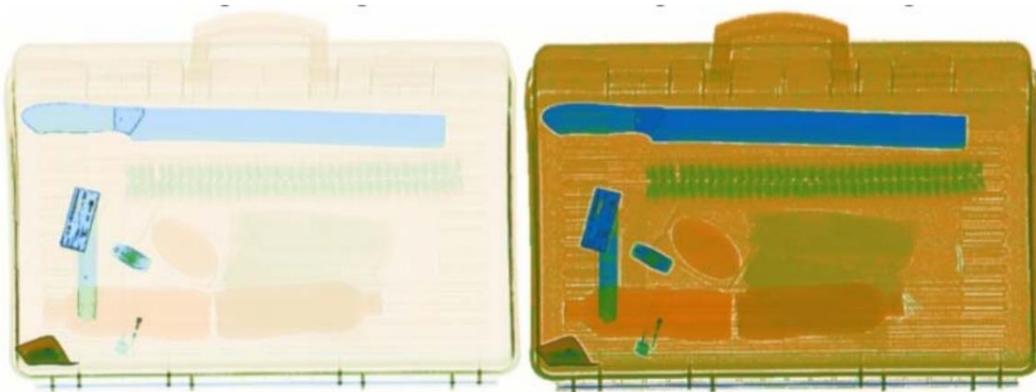


Figure 19 : amélioration de la qualité de l'image par balayage des niveaux de gris.

12. Traitement de pseudo-couleur : Cette sélection permet à l'opérateur de mieux identifier les objets. La combinaison des fonctions "Traitement de pseudo-couleur" et "Amélioration de la puissance de pénétration élevée" permet de visualiser des objets hauts densité.



Figure 20 : amélioration de la qualité de l'image par traitement de pseudo-couleur.

13. Amélioration des substances organiques suspectes Z7 / Z8 / Z9: Ce bouton permet à l'opérateur de mettre en surbrillance certains objets en fonction de leur numéro atomique. 7 - eau, explosifs plastiques, 8 - drogues impures ou explosives, 9 - drogues pures.



Figure 21 : amélioration de la qualité de l'image par amélioration des substances organiques suspectes Z7 / Z8 / Z9.

14. Amélioration des bords : permet à l'opérateur de rendre l'image originale plus nette.



Figure 22 : amélioration de la qualité de l'image par amélioration des bords.

2.3 Segmentation

La segmentation des images est un problème fondamental de la vision par ordinateur et a été à l'origine d'une vaste ressource documentaire. Dans le contexte du contrôle des bagages, l'objectif de la segmentation est généralement pragmatique, c'est-à-dire que sa mise en œuvre est destinée à faciliter la détection ultérieure des menaces. Malgré son importance, les travaux antérieurs sur la segmentation automatisée d'objets 2D/3D inconnus à partir d'images à rayon X de bagage sont limités. Les différences de qualité et de nature des images de contrôle de sécurité par rapport aux images médicales ont limité le succès des techniques de segmentation médicale dans ce domaine [70]. En outre, la dépendance de la majorité des techniques de segmentation médicale de pointe vis-à-vis d'informations a priori en particulier nuit à leur pertinence pour les images d'inspection de sécurité, qui nécessitent la segmentation de plusieurs objets inconnus. Cette hypothèse a été vérifiée par Megherbi *et al.* [70], qui ont étudié l'efficacité des techniques classiques de segmentation médicale lorsqu'elles étaient appliquées directement sur des tomodensitométries de bagage de faible qualité. Ils ont été démontrés que la réussite de la segmentation dépend d'un réglage minutieux des paramètres par cas. De plus, même après l'optimisation des paramètres, les effets du bruit d'image (malgré le pré filtrage), l'encombrement et le manque de connaissances préalables se traduisent par une segmentation nettement inférieure à celle observée dans le domaine médical.

Fu *et al.* [71] tentent de segmenter des objets suspects. Ils introduisent une approche de regroupement hybride qui ne requiert pas de priorité sur le nombre de grappes ou la taille des grappes (clusters), mais une priorité sur le niveau de pas, qui détermine le nombre de niveaux de quantification dans l'image groupée en fonction de la valeur maximale de l'image. La mise en cluster hybride effectue le clustering suivie par la croissance de la région (region growing). Pour le regroupement, chaque pixel est d'abord comparé à la moyenne de son voisinage. Si le pixel est proche de la moyenne, sa valeur est attribuée à la quantification de cette moyenne. Si ce n'est pas proche, ils divisent le voisinage en quadrants, calculent les moyennes et fixent la valeur du pixel à la moyenne quantifiée du quadrant le plus proche. Ils affirment que cela est plus rapide que le regroupement récursif de k-means et le regroupement de Leader utilisé par Ogorodnikov et Petrounine [72], [73]. Il est bénéfique, car il crée des grappes continues et non disjointes. Après le regroupement, ils fusionnent les régions.

Dans les bagages CT, plusieurs propositions de segmentation mono-énergie et double-énergie ont été proposées, certaines basées sur l'apprentissage automatique que nous examinons ici. Dans la segmentation des bagages CT, les algorithmes doivent prendre en charge un nombre variable et inconnu de bagages, avec une grande variabilité dans leurs formes, types et tailles [74]. Cela contraste avec le domaine médical, où les tâches de segmentation sont pré-spécifiées, par exemple une segmentation d'un organe particulier [74]. Par conséquent, les chercheurs en bagages ont cherché à concevoir des algorithmes non supervisés qui ne supposent aucune hypothèse sur le nombre d'objets ou sur leur composition.

Mery *et al.* [75] ont étudié la reconnaissance d'objets en appliquant différents techniques modernes de vision par ordinateur. Leurs images ont été segmentées à l'aide d'une méthode de classification adaptative k-means. Abidi *et al.* [76] ont amélioré un système de détection d'armes dans des images radiologiques à énergie unique par pseudo-coloration, utilisant une segmentation à seuils multiples.

Matthew *et al.* [77] ont présenté un système automatisé d'inspection des bagages avec des robots. Ils ont utilisé la bibliothèque de nuages de points (PCL) [78] pour mettre en œuvre la segmentation du plan, leur permettant de supprimer des surfaces supplémentaires de l'image. Ainsi que Hassanpour *et al.* [79], ils se sont intéressés au développement de la détection de

matériau dans les bagages où l'image est segmentée en sous-régions en fonction du niveau d'intensité des pixels, suivie d'un algorithme de détection de recouvrement afin d'estimer les régions appartenant à chaque couche d'intensité. De plus, Domingo *et al.* [80] ont travaillé sur la technologie de vision par ordinateur pour les tests par rayons X. Ils ont utilisé les histogrammes, la détection de contour, des opérations morphologiques et un filtrage pour la phase de segmentation. De plus, Megherbi *et al.* [81] ont proposé un système de détection des menaces potentielles lors du filtrage des bagages de CT. Ils ont utilisé la segmentation des approches médicales, qui consistent en des méthodes basées sur les régions (effectuer la segmentation en trouvant des régions cohérentes en fonction de certains critères) et des méthodes basées sur les limites (trouver les limites des objets d'intérêt). De plus, en 2013 [70], ils ont étudié l'application des techniques de segmentation CT au contrôle automatisé des bagages et des colis; en particulier, ils ont exploré et mis l'accent sur la croissance de la région connectée (confidence connected region growing), la connectivité floue (fuzzy connectedness), Ligne de partage des eaux (watershed) et la mise en correspondance rapide (fast matching). Dans [82], Liang *et al.* ont proposé une méthode de segmentation automatique d'images par rayons X pour la détection de menaces de faible intensité. Ils ont utilisé des techniques de multi-seuillage et de regroupement de données pour le processus de segmentation. Turcsany *et al.* [83] ont amélioré le système de reconnaissance d'objets en utilisant une segmentation approximative au premier plan en tronquant les intensités de pixels en niveaux de gris (a coarse foreground segmentation by truncating grayscale pixel intensities). Martin *et al.* [84] proposent un framework basé sur l'apprentissage pour la segmentation conjointe et l'identification d'objets dans des images de tomographie par rayons X à double énergie (DECT). Ils se sont concentrés sur la segmentation et l'identification d'un petit ensemble d'objets présentant un intérêt, en considérant tout le reste comme arrière-plan. De même, Babaheidarian *et al.* [85] se sont concentrés sur le problème de l'identification des matériaux dans les images DECT, ils ont donc proposé un nouvel algorithme pour la segmentation conjointe et la classification des régions matérielles. De plus, Muthukkumarasamy *et al.* [86] visaient à développer un système intelligent de détection d'objets interdits pour renforcer la sécurité de l'aviation en convertissant chaque image en niveaux de gris après les avoir segmentés à l'aide d'un

algorithme à seuils multiples. Heitz *et al.* [87] ont présenté une méthode permettant de séparer des objets dans un ensemble d'images aux rayons X en utilisant la propriété d'atténuation d'additivité au niveau des pixels (somme des atténuations de tous les objets traversés par les rayons X). Cette méthode exploite plusieurs vues de projection de la même scène sous des angles légèrement différents pour produire une estimation précise des propriétés d'atténuation des objets de la scène.

Wiley *et al.* [88] présentent une méthode de croissance de régions en 3D basée sur la technologie de segmentation médicale Stratovan Tumbler [89]. La technique est composée de cinq étapes: 1) la définition d'un noyau 3D; 2) détermination des critères de mouvement du noyau; 3) initialisation de la graine (seed initialization); 4) remplissage et 5) scission et fusion. Les résultats optimaux sont obtenus en utilisant un noyau sphérique, à condition que la taille du noyau (déterminée automatiquement, en fonction de la quantité d'encombrement local) soit inférieure à l'objet segmenté et supérieure à tous les trous prévus dans sa limite. Les critères de mouvement pour un noyau donné sont déterminés automatiquement à l'aide d'une procédure d'apprentissage.

Song *et al.* (TeleSecurity Sciences, Inc.) [90] présentent une approche séquentielle composée de trois étapes : 1) prétraitement (par filtrage bilatéral 2D [91]), 2) segmentation d'objet et 3) post-traitement. La segmentation des objets est réalisée selon une approche séquentielle «Segment et carve» (SC), reposant sur le principe que les objets simples doivent être segmentés en premier. Les objets segmentés à chaque étape sont taillés dans l'image avant de passer à l'étape suivante. La segmentation est réalisée à l'aide de la technique Symmetric Region-Growing (SymRG) [92] - une technique de croissance de région sans pépins (c'est-à-dire non supervisée) basée sur une fonction symétrique et qui est invariante des conditions de départ. Au total, cinq étapes sont proposées, chacune ciblant des objets ayant des caractéristiques différentes : 1) objets homogènes en vrac ; 2) objets homogènes d'épaisseur moyenne ; 3) objets homogènes en forme de feuille ; 4) objets métalliques homogènes et 5) objets hétérogènes. Chaque étape est composée de cinq sous-étapes : 1) génération de masque binaire par seuillage ; 2) prétraitement du masque ; 3) segmentation par SymRG; 4) correction des limites et 5) sculpture d'objets. Les cinq étapes nécessitent chacune un réglage des paramètres. Une fois la

procédure SC séquentielle en cinq étapes terminées, les objets segmentés de chaque étape sont soumis à de nombreuses opérations de post-traitement visant à corriger la sur et la sous-segmentation. En particulier, le fractionnement d'objet s'effectue en quatre étapes : 1) le fractionnement par analyse d'histogramme ; 2) scission par RANSAC ; 3) scission par regroupement k-moyen récursif et 4) scission par ouverture morphologique. La fusion d'objets est effectuée en fonction de trois seuils : 1) la proximité spatiale ; 2) intensité moyenne et 3) type d'objet. Bien que l'étude montre des segmentations de hautes qualités pour des objets sélectionnés, les résultats des analyses complètes ne sont pas présentés. L'approche est extrêmement compliquée (avec un grand jeu de paramètres) et des performances optimales nécessitent un réglage minutieux des paramètres.

Harvey *et al.* [90], présentent une technique basée sur la classe d'algorithmes multi-échelles des tamis (multi-scale sieves) [93], [94]. Les tamis fonctionnent en filtrant les signaux d'entrée pour éliminer les extremas d'intensité à des échelles spécifiques. Dans le contexte de la segmentation d'images, les objets sémantiquement significatifs sont supprimés à des échelles spécifiques (généralement plus élevées). L'approche proposée est composée de quatre étapes: 1) tamiser le volume d'entrée à quatre échelles logarithmiquement espacées; 2) calculer quatre volumes de canal; 3) étiquetez les volumes de canal à l'aide de l'analyse par composant connecté de Sedgewick [95] et 4) fusionnez les volumes de canal étiquetés en un seul volume étiqueté. La fusion est effectuée en déterminant les similitudes entre les histogrammes de densité pour chaque objet marqué dans chaque volume de canal en utilisant le test de Kolmogorov-Smirnov (K-S) [96], qui calcule la probabilité que les histogrammes aient été tirés de la même distribution. Les forces et les faiblesses spécifiques de l'approche ne sont pas suffisamment détaillées [90]. Une observation intéressante est qu'en raison de la segmentation de tous les objets à toutes les échelles, au moins un canal contient toujours une segmentation d'un objet. Harvey *et al.* [90] proposent donc qu'une approche plus appropriée (par rapport à la fusion de canaux) serait de passer les volumes de canal directement dans un système d'intelligence artificielle (par exemple, un classificateur, un détecteur d'objet ou un détecteur de région saillante).

Feng *et al.* [90] présentent une technique 3D. L'approche est composée de trois étapes : 1) génération de cartes de semences (seed map) ; 2) croissance régionale adaptative et 3) fusion. Les cartes de semences sont générées en localisant des régions homogènes suffisamment grandes dans le volume d'entrée. Les régions homogènes sont déterminées par seuillage de la carte de gradient de Sobel du volume [97], tandis que la taille de la région est déterminée par une analyse en composantes connexes. Les régions de semences sont cultivées par culture dynamique [98], où le seuil de croissance de la région n'est pas constant. Pour compenser la variation d'intensité au sein des objets (due aux artefacts de tomographie), le seuil de croissance de la région est modélisé comme une fonction non linéaire de l'intensité moyenne de la région. À la fin de la croissance de la région, des paires d'objets en contact (c'est-à-dire ceux partageant une arête commune) sont fusionnées en fonction de leur similarité dans un espace de fonctions 2D (caractérisant la texture et l'intensité moyennes). Cette heuristique de fusion est appliquée de manière récursive. Il est démontré que la technique est sensible au réglage des paramètres et susceptible de sous-segmentation (survenant dans environ 15% des cas).

Mouton et Breckon [99] ont récemment présenté une technique basée sur les matériaux pour la segmentation d'objets 3D inconnus à partir d'images CT-baggage encombrées et à faible résolution. Les segmentations grossières initiales sont générées à partir d'images [100] à réduction d'artefact métallique utilisant l'indice de double énergie (DEI) [101], le seuillage et l'analyse des composants connectés. La qualité des composants individuels de la segmentation grossière est déterminée à l'aide d'une nouvelle mesure d'évaluation basée sur une forêt aléatoire (Random Forest Score (RFS)), conçue pour reconnaître les segments de haute qualité contenant un seul objet. Les segments d'objets individuels de qualité médiocre sont soumis à une procédure de partitionnement d'objet qui divise les objets fusionnés en régions détectées automatiquement à l'aide d'opérations morphologiques et d'une analyse des composants connectés. Dans une évaluation comparative, il est démontré que la technique produit des segmentations comparables à l'état de la technique [102] avec une réduction significative du temps de traitement. Un manque d'informations sur la vérité sur le terrain est souligné comme une limite dans l'évaluation de la performance de l'étude.

2.4 Compréhension de l'image

2.4.1 Classification

La classification des objets est une tâche importante pour l'identification (étiquetage sémantique) d'objets particuliers par rapport à d'autres, c'est-à-dire qu'il s'agit d'une menace ou non. Pour la classification des objets à rayons X, la plupart des travaux antérieurs proposent des approches traditionnelles d'apprentissage automatique basées sur un schéma de représentation des caractéristiques du sac de mots visuels (Bag-of-Visual-Words BoVW), utilisant des classificateurs tels que la machine à vecteurs de support (SVM) [103], [104], [105], [106], [107].

Le travail de [103] examine le concept de BoVW dans l'imagerie des bagages à rayons X utilisant la classification SVM avec plusieurs représentations d'entités (DoG, DoG + EIPR, DoG + Harris) atteignant des performances de 0,7 de rappel (recall), 0,29 précision et 0,57 précision moyenne. Turcsany *et al.* [83] ont suivi une approche similaire et ont prolongé le travail présenté dans [103] utilisant la BoVW avec des descripteurs SURF et d'un classificateur SVM, associée à une version modifiée de la génération du livre de codes (codebook generation), donne 0,99 taux de vrais positifs et 0,04 taux de faux positifs [83]. Les approches BoVW avec descripteur de caractéristique et classification SVM sont également utilisées dans [108] pour la classification des images radiologiques à simple et double vues, avec des précisions moyennes optimales obtenues pour les armes à feu (0,95) et les ordinateurs portables (0,98). Mery *et al.* [107] proposent une approche de reconnaissance qui applique la détection aux images à vue unique pour trouver des objets d'intérêt, puis leur fait correspondre des images aux rayons X à vues multiples pour une précision de 0,96 et un rappel de 0,93 pour 120 objets. Une approche BoVW est en outre utilisée dans [106] où un dictionnaire est formé pour chaque classe qui consiste en des descripteurs de caractéristiques de patches d'image recadrés de façon aléatoire. Les performances du modèle sont évaluées en ajustant une classification de représentation fragmentée aux descripteurs de caractéristiques extraites de patches de test coupé au hasard, et des dictionnaires adaptatifs sont obtenus à partir de la phase de formation. La procédure expérimentale montre des résultats prometteurs pour la classification des patches.

Kundegorski *et al.* [105] explorent de manière exhaustive l'utilisation de divers descripteurs de points de caractéristique dans un modèle BoVW. Ceci concerne la détection des menaces basée sur la classification des images dans l'imagerie par rayons X de sécurité des bagages, en utilisant un détecteur de caractéristiques FAST-SURF et une combinaison de descripteurs donnant une performance maximale avec une classification SVM (détection d'arme à feu à 2 classes : précision 94.0%).

L'étude de [109] compare une approche BoVW et une approche CNN, explorant l'utilisation de l'apprentissage par transfert. Des expériences ont montré que CNN surpasse la méthode BoVW, même lorsque les entités sont transférées de manière abstraite à partir d'un autre problème de classification. À la suite des travaux antérieurs de [109], [75] explore de manière exhaustive l'utilisation de différentes approches de classification dans le domaine des bagages X-ray en utilisant dix techniques différentes, y compris BoVW, les représentations clairsemées et CNN. Les expériences montrent des résultats parallèles à [109], soutenant la supériorité généralisée des caractéristiques de CNN, mais sans autre considération du problème de détection d'objet initial (localisation), ni d'exploration exhaustive de la performance de CNN au sens large.

Samet Akcay *et al.* [110] ont examiné l'utilisation de réseaux de neurones à convolution profonde (CNN) avec apprentissage par transfert pour la classification des images et les problèmes de détection. Dans leur contexte de disponibilité limitée d'exemples de données sur les objets d'intérêt, ils ont utilisé un paradigme d'apprentissage par transfert tel qu'un CNN préformée. Pour permettre une comparaison cohérente d'espace-fonction entre cette approche et les représentations d'espace-espace traditionnelles, ils ont également formé le classifieur SVM (support vector machine) sur les entités CNN. Ils ont montré que les fonctionnalités CNN ajustées offrent des performances supérieures à celles des fonctionnalités classiques conçues à la main pour les tâches de classification des objets dans ce contexte. Au total, ils ont atteint une précision de 0,994 sur la base des fonctionnalités AlexNet formées avec le classificateur SVM.

2.4.2 Détection

Une tâche essentielle restante dans ce domaine de problème est celle de la détection dans laquelle les objets d'intérêt sont localisés dans l'image globale de rayons X, généralement

désignés par un cadre de sélection ou un contour de forme. La détection étant un problème complexe, les modèles fondés sur la détection dans l'imagerie par rayons X des bagages sont nettement plus limités dans la littérature.

Dans [111], la détection des régions d'intérêt (ROI) dans les images à rayons X est effectuée via un modèle géométrique de l'objet, en estimant la structure à partir du mouvement. Les régions potentielles obtenues à partir de l'étape de segmentation sont ensuite suivies en fonction de leur similarité, ce qui permet d'obtenir 0,943 vrais positifs et 0,056 faux positifs sur un petit jeu de données non encombré.

Franzel *et al.* [112] proposent une approche de détection par fenêtre glissante utilisant un classifieur SVM linéaire et un histogramme de gradients orientés (HOG) [113]. Comme HOG n'est pas totalement invariant par rotation, ils complètent leur approche en détectant des orientations variables. L'étape suivante, appelée intégration multivues, consiste à fusionner les détections d'images à vue unique à rayons X, obtenues à partir de plusieurs points de vue dans un scanner à rayons X moderne, pour éviter les fausses détections et rechercher l'intersection des détections vraies. Il a été démontré que la détection à vues multiples offrait des performances de détection supérieures à la détection à vue unique pour les armes de poing (mAP: 0.645). De même [114] explore la détection d'objets dans l'imagerie de bagages par rayons X en évaluant diverses combinaisons de détecteurs de caractéristiques et de descripteurs à la main, en utilisant un algorithme de branchement et lié et un classificateur SVM structurel (mAP: 0.881 pour 6400 images d'armes de poing, d'ordinateurs portables et de bouteilles en verre).

Un ensemble de travaux liés cible également l'utilisation de techniques BoVW dans le cadre de la détection d'objet de tâche hautement associée dans l'imagerie de sécurité des bagages de tomographie 3D [115], [116], [99]. Une revue détaillée est présentée dans [117], [118].

Très récemment, Samet Akcay *et al.* [110] ont également exploré l'applicabilité de plusieurs paradigmes de détection pilotés par CNN, tels que les CNN à fenêtre glissante (SW-CNN), les CNN à région plus rapide (F-RCNN), les réseaux entièrement à convolution à base de région (R-FCN) et YOLOv2. Ils ont formé de nombreux réseaux traitants des détections simples et multiples sur des variantes SW-CNN / F-RCNN / R-FCN / YOLOv2. YOLOv2, Faster-RCNN

et R-FCN ont fourni des résultats supérieurs aux approches plus traditionnelles SW-CNN. Avec l'utilisation de YOLOv2 et d'images d'entrée de taille 544×544 , ils ont atteint une précision moyenne de 0,885 (mAP) pour un problème de détection d'objet de six classes. La même approche, avec une entrée de taille 416×416 , donne 0,974 mAP pour le problème de détection d'arme à feu à deux classes et nécessite environ 100 ms par image. Dans l'ensemble, ils ont illustré les performances comparées de ces techniques et montré que les stratégies de localisation d'objets s'adaptent bien aux images de sécurité à rayons X encombrées, où les techniques de classification échouent.

2.5 Conclusion

Dans ce chapitre, des travaux de recherche dans le domaine traitement des images radiographique des bagages sont présentés comme étant une solution aux différentes limitations actuelles. L'état de l'art montre que les techniques d'apprentissage automatique n'ont pas été presque utilisées dans les différentes étapes du traitement des images radiographiques de bagages. Cela peut s'expliquer par la difficulté d'obtenir des ensembles de données étiquetés suffisamment volumineux et de leurs confidentialités. Le chapitre suivant décrit la mise en œuvre des méthodes utilisées dans nos travaux de recherche. Nous décrivons les différentes étapes suivies pour le développement des systèmes de détection qui permettent d'identifier les objets (à menace ou non) dans les bagages.

Chapitre 3

Notion opératoire et méthodes utilisées

3.1 Introduction

La Figure 23 montre un schéma fonctionnel résumant les étapes suivies dans le système de détection des objets à partir des images numérisées de balayage à rayons X de bagages. Au départ, nous avons commencé par la présentation d'une étude des différentes techniques d'amélioration de la qualité des images, y compris la technique de débruitage d'image proposée. Suivie par une brève discussion sur les méthodes de segmentation des images, en nous concentrant sur la segmentation sémantique qui fait partie d'une de nos contributions dans ce travail. Avant de finir, nous entamons la tâche de détection d'objet, et l'extraction des caractéristiques on utilisant le deep learning une sélection des caractéristiques est aussi mise en œuvre. Finalement, la tâche de classification est réalisée.

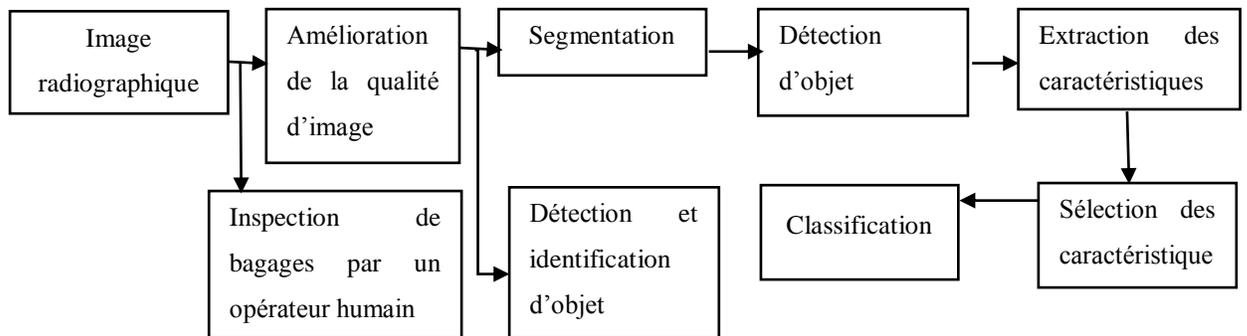


Figure 23: Schéma fonctionnel d'un système de détection permettant de détecter des objets dans des images numérisées de balayage de bagages à rayons X.

3.2 Amélioration de la qualité d'image

La segmentation des images en couleurs est un sujet de recherche très émergent dans le domaine de l'analyse d'images en couleurs et de la reconnaissance de formes. De nombreux algorithmes à la pointe de la technologie ont été développés à cette fin. Mais, souvent, les résultats de la segmentation de ces algorithmes semblent souffrir d'erreurs de classification et de sur-segmentation. Cela s'explique par la dégradation de la qualité de l'image lors de l'acquisition, de la transmission et de la conversion de l'espace colorimétrique. Il devient donc nécessaire de disposer d'une technique d'amélioration d'image efficace capable de supprimer les pixels ou bruits redondants de l'image couleur avant de procéder à la segmentation finale. Dans ce chapitre, un effort a été fait pour étudier et analyser différentes techniques d'amélioration d'image afin de déterminer celle qui convient le mieux à la segmentation des images en couleur. En outre, cette étude comparative est réalisée séparément sur trois espaces colorimétriques bien connus, RGB, HSV et LAB, afin de déterminer quel espace colorimétrique prend en charge la tâche de segmentation plus efficacement par rapport à ces techniques d'amélioration.

Les images en couleur contiennent une grande quantité d'informations. Mais cette information est quelque peu cachée, de sorte que les yeux de l'homme ont tendance à ne pas les analyser. Plus important encore, il est très difficile d'obtenir de réels changements dans les caractéristiques des informations telles que l'intensité, la couleur, la texture, etc. Nous avons donc besoin d'une technique efficace de segmentation d'images couleur pour les analyser. Mais le résultat de toute technique de segmentation d'images couleur dépend totalement de la qualité de l'image concernée. En particulier, dans le cas de l'image radiographique des bagages, la qualité de l'image se dégrade en raison des bruits généralement impliqués lors des processus de capture, de transmission et d'acquisition de l'image. Ainsi, la segmentation de telles images bruitées ne produit pas un résultat d'analyse efficace. Par conséquent, nous avons besoin de techniques de prétraitement pour supprimer les artefacts, les valeurs aberrantes ou nous pouvons dire des bruits dans les images avant de passer à l'étape d'analyse suivante. L'amélioration de l'image est une technique de prétraitement dans laquelle notre objectif est de supprimer le bruit tout en préservant l'intégrité des contours et les autres informations détaillées

[119], [120]. En réalité, les bruits ne peuvent être complètement réduits que lorsque les causes réelles de leur formation sont étudiées, nous ne pouvons pas enquêter complètement sur eux. La seule chose que nous puissions faire est donc d'introduire des techniques basées sur des équations mathématiques pour éliminer le plus possible les bruits [121].

Les techniques d'amélioration des images couleur impliquent plus d'efforts que les techniques d'amélioration des images grises pour les deux raisons suivantes [121]:

- Dans le cas des images couleur, nous devons considérer les vecteurs plutôt que les scalaires.
- En outre, pour les images en couleur, la complexité de la perception de l'image est encore un fait considérable.

Les techniques vectorielles sont très difficiles à mettre en œuvre sur le plan informatique. Ainsi, les techniques à base monochromatique ont toujours la préférence lorsque des canaux séparés d'une image couleur sont améliorés. L'espace colorimétrique compte toujours beaucoup pour le traitement des images couleur. Le RGB est le plus commun quand on parle d'image couleur. Cet espace colorimétrique comprend trois composants : rouge, vert et bleu. La segmentation monochromatique implique donc l'analyse de ces trois canaux et nous ne pouvons garantir un très bon résultat pour cet espace colorimétrique. Pour cette raison, l'option suivante à choisir pour les espaces colorimétriques pour lesquels nous disposons d'un canal séparé pour la mesure de la luminosité. HSV et LAB sont les deux espaces colorimétriques les plus populaires répondant à ce critère. Dans ce chapitre, nous avons adopté ces deux espaces colorimétriques pour notre étude expérimentale. Une illustration détaillée des espaces colorimétriques LAB et HSV sont disponibles dans [122].

3.2.1 Amélioration du contraste

L'amélioration du contraste est un processus par lequel l'intensité de pixel de l'image est modifiée [123]. Généralement, le terme «contraste» fait référence à la séparation des zones sombres et des zones claires présentes dans une image. L'amélioration du contraste a pour avantage de lever l'ambiguïté qui pourrait normalement se produire entre différentes régions d'une image. L'amélioration du contraste peut être classée en deux catégories : (1)

Amélioration du contraste local ; et (2) Amélioration du contraste global. Le Tableau 4 les illustre plus précisément :

Tableau 4: La différence entre l'amélioration du contraste local et l'amélioration du contraste global.

	Amélioration du contraste local	Amélioration du contraste global
1	Dans ce type de techniques d'amélioration du contraste, une petite fenêtre est une diapositive séquentielle de chaque pixel de l'image d'entrée et seuls les blocs de pixels de cette fenêtre sont améliorés. Cela signifie que les informations locales sont utilisées intelligemment dans ce cas.	Dans ce type de techniques d'amélioration du contraste, les informations globales de l'histogramme sont prises en compte pour l'amélioration. Étant donné que l'ensemble de l'image est considérée à la fois, les informations locales sont alors ignorées.
2	La luminosité locale future est considérablement améliorée dans ce cas. Le rapport de contraste peut donc être amélioré dans toutes les régions de l'image.	Comme la luminosité globale est envisagée dans ce cas, cela limite donc le rapport de contraste dans certaines parties de l'image. Cela entraîne des pertes de contraste significatives dans de petites régions de l'image, en particulier à l'arrière-plan.
3	La technique d'amélioration du contraste local est complexe en termes de calcul. Parfois, cela implique un coût élevé en raison de la prise en compte de sous-blocs chevauchants.	Le principal avantage de l'amélioration globale du contraste est sa simplicité de calcul et son aptitude à l'amélioration globale de l'image.
4	AHE et CLAHE sont des exemples de deux techniques d'amélioration du	Des exemples de deux techniques d'amélioration du contraste global

	contraste local fréquemment adoptées.	largement adoptées sont l'égalisation d'histogramme et la spécification d'histogramme.
--	---------------------------------------	--

Dans les sous-sections suivantes, on discute d'abord de l'amélioration du contraste global, puis des techniques d'amélioration du contraste local. Des expériences appropriées montrant leur effet sur les images radiographiques couleur des bagages sont présentées dans le chapitre suivant.

3.2.1.1 Amélioration du contraste global

3.2.1.1.1 Égalisation d'histogramme

L'histogramme est un graphique qui montre la fréquence d'apparition de données dans l'ensemble des données. Il trace le nombre de pixels pour chaque valeur tonale [124]. Considérons une image avec G niveaux d'intensité possibles totaux. Ensuite, l'histogramme de l'image dans [0, G-1] est défini comme une fonction discrète :

$$p(r_k) = \frac{n_k}{n} \quad (3.1)$$

Où,

r_k est le $k^{\text{ième}}$ niveau d'intensité dans l'intervalle.

n_k est le nombre de pixels de l'image dont le niveau d'intensité est r_k .

n est le nombre total de pixels de l'image.

L'égalisation d'histogramme est une technique d'amélioration de l'image qui améliore le contraste d'une image en répartissant les valeurs d'intensité sur toute la plage dynamique disponible [125], [124]. Ceci est réalisé grâce à une fonction de transformation $T(r)$, qui peut être définie par la fonction de distribution cumulative (CDF) d'une fonction de densité de probabilité (PDF) donnée de niveaux de gris dans une image.

3.2.1.1.1.1 Cas continu

Ceci concerne les niveaux d'intensité correspondant à des quantités continues normalisées dans la plage [0, 1].

Soit, $P_r(r)$ est le PDF des niveaux d'intensité. Ensuite, la transformation requise sur les niveaux d'entrée pour obtenir le niveau de sortie S est la suivante :

$$S = T(r) = \int_0^r P_r(w)dw \quad (3.2)$$

Où w est une variable fictive d'intégration. Ensuite, on peut montrer que le PDF des niveaux de sortie est uniforme [125], c.-à-d.

$$P = \begin{cases} 1, & \text{pour } 0 \leq S \leq 1 \\ 0, & \text{autrement} \end{cases} \quad (3.3)$$

La transformation ci-dessus génère une image dont les niveaux d'intensité sont tout aussi probables et couvre également toute la plage [0, 1]. Ce processus d'égalisation du niveau d'intensité produit une image avec une plage dynamique accrue avec une tendance à un contraste plus élevé.

3.2.1.1.1.2 Cas discret

Dans le cas de quantités discrètes, nous traitons des sommations et, par conséquent, la transformation d'égalisation donne :

$$S_k = T(r_k) = \sum_{j=0}^k P_r(r_j) = \sum_{j=1}^k \frac{n_j}{n} \quad \text{pour } k = 1,2,3 \dots L \quad (3.4)$$

où S_k est la valeur d'intensité de l'image de sortie correspondant à la valeur r_k dans l'image d'entrée. Pour notre étude, nous avons d'abord converti l'image RVB en image HSV, car l'espace colorimétrique HSV convient mieux à la segmentation d'images couleur. Cela est dû à la propriété que l'espace colorimétrique HSV représente la couleur de la même manière que l'œil humain peut percevoir. De plus, comme nous avons besoin du canal de luminance, nous avons donc extrait le canal V de l'image convertie en HSV et avons effectué l'opération

d'égalisation de l'histogramme sur celle-ci. Enfin, remplacer le canal V d'origine par l'histogramme égalisé ; nous obtenons la version améliorée de l'image dans l'espace colorimétrique HSV. Donc, avec une étape supplémentaire (c'est-à-dire la conversion de l'image HSV en RGB), nous obtenons la version égalisée de l'histogramme requise de l'image radiographique couleur originale de bagage. Les résultats expérimentaux sont présentés dans le chapitre suivant.

3.2.1.1.2 Spécification de l'histogramme

Il s'agit d'une technique par laquelle la forme de l'histogramme d'une image est changée en une forme spécifique que nous souhaitons [125], [126]. La spécification d'histogramme est également appelée correspondance d'histogramme. Contrairement à l'égalisation idéale de l'histogramme à espacement égal, l'histogramme est ici explicitement spécifié. Si nous considérons que les niveaux de gris continus sont normalisés aux intervalles [0,1] et que r et z sont les niveaux d'intensité de l'image d'entrée et de l'image de sortie. Disons que $p_r(r)$ est la fonction de densité de probabilité de l'image en entrée et que $p_z(z)$ est celle qui est en sortie. Maintenant, appliquons d'abord la transformation :

$$S = T(r) = \int_0^r P_r(w)dw \quad (3.5)$$

Cependant, il existe quelques inconvénients à la spécification d'histogramme :

- (1) Nous n'avons pas de règle définie pour spécifier un histogramme optimal.
- (2) Chaque tâche d'amélioration donnée doit être analysée au cas par cas.
- (3) Dans l'ensemble, il s'agit d'un processus d'essai et d'erreur.

3.2.1.2 Amélioration du contraste local

3.2.1.2.1 Equalisation Adaptive Histogram (AHE)

AHE diffère de l'égalisation ordinaire des histogrammes en ce sens qu'il n'est pas global et qu'il calcule de nombreux histogrammes correspondant à différentes sections d'une image [127]. Ainsi, il est possible d'améliorer le contraste local d'une image via AHE. Avec AHE, les

informations de toutes les plages d'intensité d'une image peuvent être visualisées simultanément, ce qui résout le problème de nombreux périphériques ordinaires incapables de représenter la plage d'intensité dynamique complète. Ici, tout d'abord, une région contextuelle est définie pour chaque pixel de l'image. La région contextuelle est la région centrée autour de ce pixel particulier. Ensuite, les valeurs d'intensité de cette région sont utilisées pour trouver la fonction de mappage d'égalisation d'histogramme. La fonction de mappage ainsi obtenue est appliquée au pixel en cours de traitement dans la région et, par conséquent, l'image résultante produite après chaque pixel de l'image est mappée différemment. Il en résulte que la distribution locale des intensités et l'amélioration finale sont basées sur la zone locale plutôt que sur la totalité de la zone globale de l'image. C'est le principal avantage d'AHE. Mais, parfois, l'AHE a tendance à trop améliorer le contenu en bruit qui peut exister dans un bloc local homogène de l'image en mappant une plage courte de pixels sur un large.

3.2.1.2.2 Égalisation d'histogramme adaptatif limité au contraste (CLAHE)

CLAHE est synonyme d'égalisation d'histogramme adaptatif limité au contraste. Il s'agit d'une technique d'amélioration du contraste local, une version améliorée de l'AHE [123] [128]. En réalité, AHE souffre d'une amplification des bruits pouvant exister dans certaines régions homogènes, il est donc nécessaire de limiter le rehaussement du contraste. Grâce à CLAHE, l'image d'entrée est partitionnée en régions non superposées de taille égale (approximative) et une égalisation d'histogramme est appliquée à chacune d'entre elles [123]. Ensuite, chaque histogramme est découpé selon une limite de découpage basée sur le contraste souhaité et sur la taille de la région voisine. Ensuite, une interpolation bilinéaire est appliquée pour éliminer les limites des régions et leur donner un aspect plus lisse, comme s'il n'y avait pas de limites. CLAHE est basé sur le concept que si nous mettons une limite ou coupons la hauteur de l'histogramme de la corbeille concernée à un certain niveau, la pente de la fonction CDF peut être limitée automatiquement. Cela limite l'amplification souhaitée du contraste local. Donc, un mécanisme efficace est nécessaire pour déterminer les points auxquels le découpage doit être effectué et pour redistribuer les pixels découpés.

3.2.2 Autres techniques utilisées

Les principaux problèmes des images radiologiques à double énergie sont le bruit (bruit de fond) et le faible contraste. Différents chercheurs ont utilisé différents algorithmes / techniques pour résoudre ces problèmes, mais comme indiqué précédemment, il existe toujours un taux de fausses alertes supérieur à 20% et le vérificateur a du mal à détecter les objets de contrebande / illicites [129]. Les principales techniques de filtrage utilisées pour l'élimination du bruit sont les filtres moyen et médian.

3.2.2.1 Technique de filtrage moyenne

Elle est utilisée pour éliminer le bruit. Le filtre moyen est utile pour éliminer le bruit de grain d'une image. Ce filtre est le plus simple et le plus fondamental des algorithmes. Ce filtre est également appelé filtre moyen. Le niveau de gris du pixel central est remplacé par la moyenne des niveaux de gris des pixels environnants. La taille du kernel dépend de l'intensité du bruit et de la taille des détails significatifs de l'image traitée. L'inconvénient du filtre moyen est qu'il est pauvre en préservation de bord [130].

3.2.2.2 Technique de filtrage médiane

Le filtre médian est une technique de filtrage numérique non linéaire souvent utilisée pour supprimer le bruit. Il fournit de meilleurs résultats que les techniques de filtrage moyennes, car il préserve les contours. Mais l'inconvénient est que si nous prenons une grande fenêtre, c'est-à-dire 5×5 ou 7×7 , cela conduit à un flou [131]. Le niveau de gris du pixel central est remplacé par la valeur médiane de tous les pixels de la fenêtre d'analyse centrée sur le pixel. La taille du noyau dépend de la fréquence du bruit et de la taille des détails significatifs de l'image traitée.

3.2.3 Notre technique proposée de dé-bruitage d'image

Une nouvelle approche consistant à combiner des images à rayons X hautes énergie et basse énergie, à la supprimer du bruit et à améliorer la qualité de l'image fusionnée à l'aide de la spécification de l'histogramme en améliorant le contraste. La Figure 25 présente un

organigramme de l'approche proposée. L'approche proposée comprend trois étapes principales, à savoir la fusion d'images, le débruitage et l'amélioration d'image.

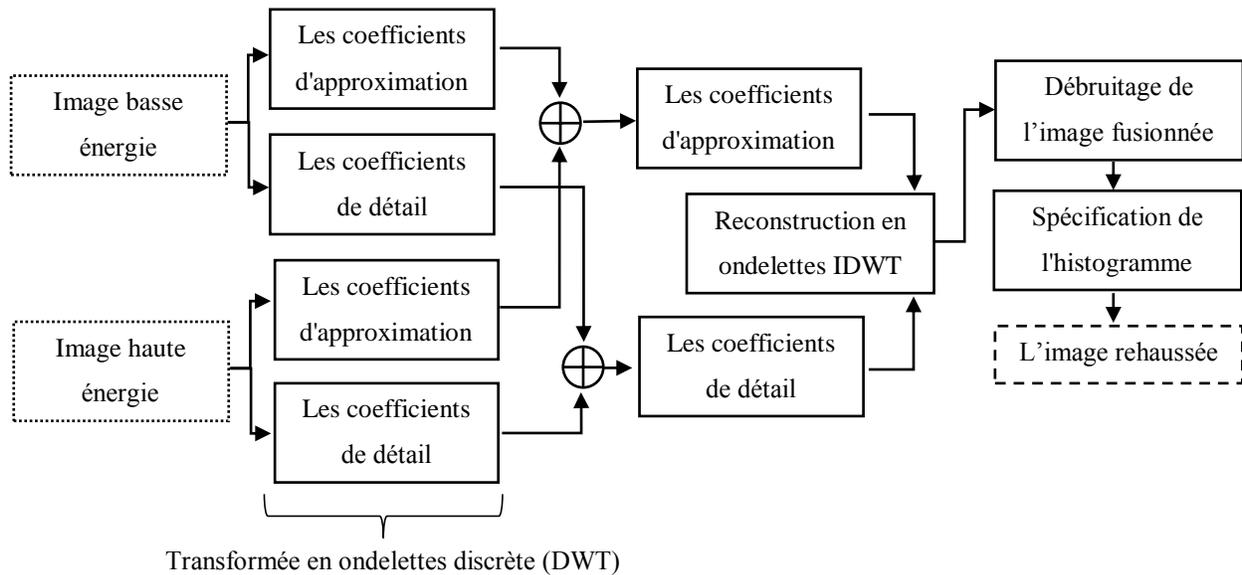


Figure 24: Un organigramme de l'approche d'amélioration d'image proposée.

3.2.3.1 La fusion d'images

Dans la fusion d'images, image haute énergie et image basse énergie sont combinées en une seule image afin d'intégrer des informations complémentaires provenant des deux images. Il y a trois étapes dans ce processus.

3.2.3.1.1 Transformée en ondelettes discrètes (DWT)

Dans cette étape, l'image haute énergie et l'image basse énergie sont combinées en une seule image pour intégrer des informations complémentaires provenant des deux images. Ici, nous effectuons le traitement DWT sur les images à énergies basse et haute pour obtenir des coefficients d'approximation et des coefficients de détail. Une famille d'ondelettes et une base d'ondelettes capable de représenter les détails de l'image doivent être sélectionnés. Une règle de sélection pratique dans les applications de traitement d'image consiste à utiliser une base d'ondelettes pouvant présenter suffisamment de variations de détail, quelle que soit sa famille. Une autre question à déterminer est le nombre d'échelles nécessaires à la décomposition. Trop

peu d'échelles entraîneront la perte d'un trop grand nombre de détails dans l'image fusionnée, et trop d'échelles donneraient une image brouillée grossière difficile à interpréter par les agents de contrôle. Il ressort de la revue de la littérature que 4 échelles donnent généralement de bons résultats [132].

3.2.3.1.2 Traitement des coefficients

Appliquez un filtre passe-bas aux coefficients d'approximation de L et H, respectivement, pour générer les coefficients d'approximation de l'image fusionnée. L'idée derrière cette étape est qu'une approximation régulière d'une scène donnée peut permettre de mieux distinguer les caractéristiques importantes de la scène. Nous générons chacun des coefficients d'approximation de l'image fusionnée, F, en faisant la moyenne des coefficients d'approximation correspondants de L et H, comme indiqué dans l'équation suivante :

$$W_{\varphi F} = \frac{1}{2} (W_{\varphi L} + W_{\varphi H}) \quad (3.6)$$

Où $W_{\varphi F}$, $W_{\varphi L}$ et $W_{\varphi H}$ sont les coefficients d'approximation de F, L et H, respectivement.

La combinaison des coefficients de détail correspondants de L et H permet d'obtenir les coefficients de détail de l'image fusionnée F. L'objectif de cette étape est d'incorporer des détails uniques provenant de L ou de H dans l'image fusionnée et de rendre plus visibles les détails existant dans les deux images. Nous calculons les coefficients de détail à toutes les échelles de décomposition de l'image fusionnée en faisant la somme des coefficients de détail correspondants de L et H, comme indiqué dans l'équation suivante :

$$W_{\psi F} = W_{\psi L} + W_{\psi H} \quad (3.7)$$

Où $W_{\psi F}$, $W_{\psi L}$ et $W_{\psi H}$ sont les coefficients de détail de F, L et H, respectivement.

3.2.3.1.3 Reconstruction en ondelettes IDWT

L'image fusionnée peut être obtenue en implémentant IDWT en utilisant les coefficients d'approximation et les coefficients de détail de l'étape 2.

Après une transformation discrète en ondelettes, nous obtenons une bonne image de détail qui contient les informations complémentaires des images H.E et L.E, mais l'image fusionnée est toujours bruyante en raison de la chute aléatoire de photons, de la discontinuité et de la non-uniformité des rayons X. les résultats dans le chapitre suivant montrent que l'image fusionnée est toujours bruyante, nous allons donc réduire le bruit de l'image fusionnée.

3.2.3.2 Débruitage de l'image fusionnée

Les images multi-capteurs ont généralement des arrière-plans bruités, comme dans les images radiologiques originales de la Figure 25. Bien que les images fusionnées révèlent généralement des informations plus détaillées, le bruit de fond existe toujours et est même amplifié. Par conséquent, une opération de suppression du bruit est nécessaire pour obtenir un bon résultat d'amélioration et de fusion. La méthode d'élimination du bruit est décrite ci-dessous.



Figure 25: Les images radiologiques originales, ils représentent respectivement de haut au bas: image à basse énergie, image haute énergie, l'image en niveaux de gris, l'image fusionnée.

Les premiers pixels bruyants de l'image corrompue sont identifiés à l'aide d'une technique de détection de pointes (spike detection technique SDT). Il est suivi par un filtre médian de restauration de pixels (PRMF) permettant de récupérer les pixels corrompus identifiés à l'aide du SDT. La technique PRMF est très efficace pour éliminer le bruit. Il utilise une fenêtre de filtrage 3×3 pour le filtrage du bruit. Notre technique de SDT suivie par PRMF est capable de produire des images de haute qualité et évite les images floues par rapport aux autres techniques de débruitage. Il convient également aux images en couleur.

3.2.3.2.1 Technique de détection de pointes

La technique de détection de pointes est appliquée à l'image corrompue pour déterminer les niveaux de bruit d'impulsion supérieur et inférieur (NL I et NL II) dans l'image corrompue. La carte binaire de l'image corrompue est construite en utilisant ces niveaux de bruit supérieur et inférieur. Si un pixel de l'image se situe à l'intérieur de ces deux limites, il est alors considéré comme «non corrompu» et «0» est attribué à la carte binaire à la position correspondante du pixel de l'image considéré. Si le pixel considéré se situe dans la plage de bruit, il est considéré comme « corrompu » et « 1 » est attribué à la carte binaire à la position correspondante.

3.2.3.2.2 Filtre médian de restauration de pixels (PRMF)

Le filtre médian de restauration de pixels est une technique robuste. Il est très efficace pour éliminer le bruit et est supérieur aux techniques de médiane adaptative. La technique PRMF proposée utilise une fenêtre de filtrage 3×3 pour éliminer le bruit impulsif. Cela simplifie l'algorithme de PRMF et peut être utilisé pour obtenir des images de meilleure qualité.

En fin de compte, une technique d'atténuation de bruit est appliquée pour récupérer l'image.

L'ensemble des étapes impliquées dans le filtrage d'une image bruyante est donné ci-dessous.

Étape I : Appliquer la technique de détection de pointes (SDT) et déterminer le niveau de bruit I et le niveau de bruit II (NLI et NLII).

Étape II : En utilisant NLI et NLII, construisez une carte binaire (BM) de l'image corrompue.

"0" signifie "pixels non endommagés"

«1» signifie «pixels corrompus»

Étape III : Appliquer le filtre médian de restauration des pixels (PRMF) à l'image bruyante X

$R = \text{PRMF}(X)$

Étape IV : Une nouvelle carte binaire (NBM) est obtenue simultanément en mettant à jour l'ancienne valeur de carte par «0» pour les pixels modifiés.

Étape V : Si NBM a des pixels "corrompus", passez à l'étape III.

Les étapes (III-V) sont répétées jusqu'à ce que le nombre de pixels corrompus dans NBM soit réduit à zéro.

3.2.3.3 Spécification de l'histogramme

L'égalisation par histogramme améliore l'ensemble de l'image. Cela égalisait l'image, mais réduisait le niveau de gris. Nous allons donc utiliser la technique de spécification d'histogramme pour améliorer l'image en particulier les régions intéressantes. Ici, tout d'abord égalisé l'image d'origine, puis recherchez la fonction de transformation $G(z)$ à partir de l'histogramme cible spécifié. Finalement, appliquez la transformation inverse de l'image égalisée en utilisant G^{-1} . Après application de la spécification de l'histogramme, l'image résultante contient des informations plus complémentaires, sans bruit et présentant davantage de contraste. Les images résultantes sont présentées dans le chapitre suivant.

3.2.4 Evaluation par signal-to-noise ratio (PSNR)

Le rapport signal sur bruit PSNR, est un terme technique qui désigne le rapport entre la puissance maximale possible d'un signal et la puissance de corruption du bruit qui affecte la fidélité de sa représentation. Du fait que de nombreux signaux ont une plage dynamique très large, le PSNR est généralement exprimé en échelle logarithmique de décibels. Ce rapport est souvent utilisé comme mesure de qualité entre l'image originale et une image compressée. Plus le PSNR est élevé, meilleure est la qualité de l'image compressée ou reconstruite [133].

Pour calculer le PSNR, nous devons d'abord calculer l'erreur quadratique moyenne en utilisant l'équation suivante :

Étant donné une image monochrome $m \times n$ sans bruit I_1 et son approximation bruyante I_2 , MSE est définie comme suit :

$$\text{MSE} = \frac{\sum_{m,n} [I_1(m,n) - I_2(m,n)]^2}{m * n} \quad (3.8)$$

Où m et n sont le nombre de lignes et de colonnes dans les images d'entrées, respectivement. Ensuite, le PSNR est calculé en utilisant l'équation suivante :

$$\text{PSNR} = 10 \log_{10} \left(\frac{R^2}{\text{MSE}} \right) \quad (3.9)$$

R est la fluctuation maximale du type de données d'image en entrée. Par exemple, si l'image d'entrée est codée sur 8 bits, R est égal à 255.

Il est très important de mentionner que nous avons injecté un bruit similaire au bruit produit par le scanner à rayons X dans l'image d'entrée, afin que nous puissions calculer le PSNR entre l'image améliorée et l'image bruitée.

3.3 Segmentation

3.3.1 Segmentation basée sur le clustering en utilisant l'apprentissage automatique

Le clustering (regroupement) consiste à diviser un ensemble de données en un certain nombre de groupes, de sorte que les points de données du même groupe se ressemblent davantage aux autres points de données du même groupe que ceux des autres groupes. Ces groupes sont appelés clusters. La mesure de la similarité sur laquelle les clusters sont modélisés peut être définie par la distance euclidienne, la distance probabiliste ou une autre métrique. Un diagramme sur la procédure de la segmentation proposée est illustré dans la Figure 26.

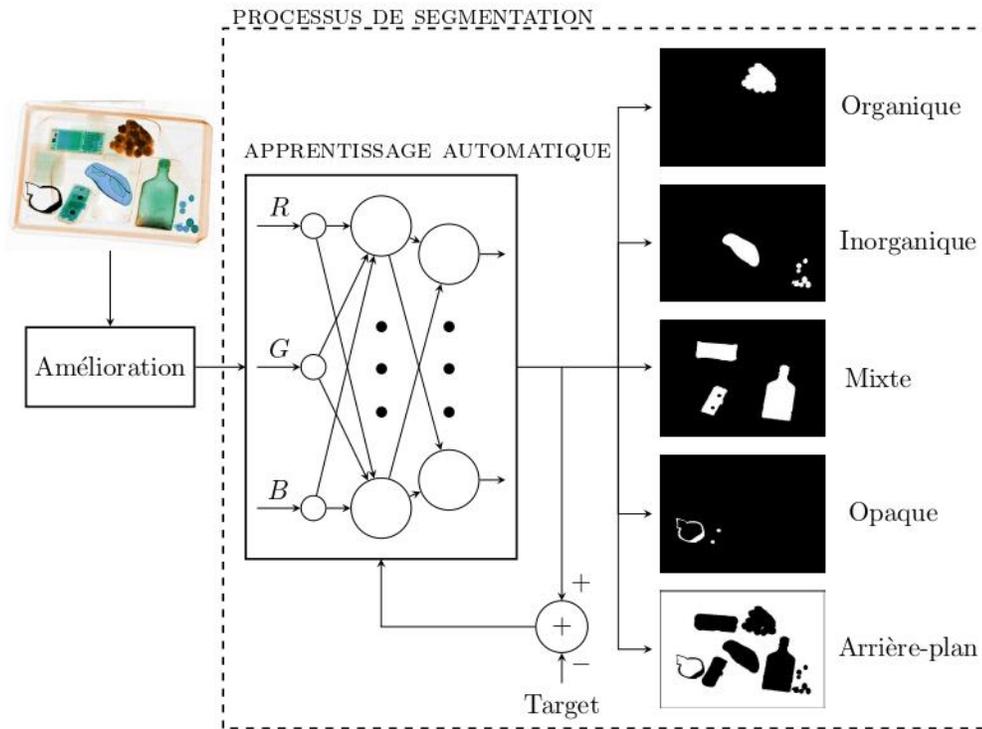


Figure 26: Un diagramme sur la procédure de la segmentation par machine learning.

Pour notre application, nous nous sommes concentrés sur le clustering avec deux manières : supervisées et non supervisées, nous détaillerons ensuite toutes les méthodes utilisées

3.3.1.1 Méthode d'apprentissage non supervisée

Contrairement aux méthodes supervisées, les méthodes non supervisées ne nécessitent aucune donnée étiquetée. Au lieu de cela, ils essaient de trouver automatiquement des modèles intéressants dans des données non étiquetées. L'apprentissage non supervisé est aujourd'hui l'un des problèmes les plus difficiles et les plus intéressants de la vision par ordinateur et de l'apprentissage automatique. De nombreux chercheurs pensent qu'apprendre de vastes collections de vidéos non étiquetées pourrait aider à décoder des questions difficiles concernant la nature de l'intelligence et de l'apprentissage. De plus, comme les images et les vidéos non étiquetées sont faciles à collecter à un coût relativement bas, un apprentissage non supervisé

pourrait avoir une réelle valeur pratique dans de nombreuses applications de vision par ordinateur et de robotique.

Nous proposons une application d'apprentissage machine (ML) non supervisée pour la segmentation d'images. Pour cela, nous avons utilisé la segmentation d'image de pixel basée sur la couleur pour séparer les objets organiques, inorganiques, mélangés et opaques de l'arrière-plan. Deuxièmement, ces cinq images résultantes sont réduites lors de la phase dite de « fusion » et sont classées qu'en deux : organique et inorganique. Nous avons utilisé les méthodes suivant :

3.3.1.1.1 Regroupement hiérarchique

La classification hiérarchique commence par traiter chaque observation comme un cluster distinct. Ensuite, il exécute à plusieurs reprises les deux étapes suivantes : (1) identifier les deux clusters les plus proches l'un de l'autre et (2) fusionner les deux clusters les plus similaires. Cela continue jusqu'à ce que tous les clusters soient fusionnés [134]. Ceci est illustré dans la Figure 27 ci-dessous.

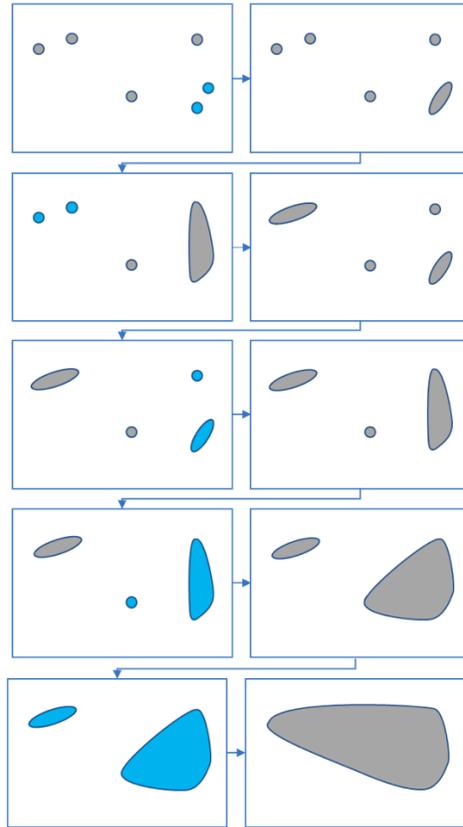


Figure 27: Principe de fonctionnement de la méthode regroupement hiérarchique.

La sortie principale du clustering hiérarchique est un dendrogramme, qui montre la relation hiérarchique entre les clusters. Une illustration est présentée dans la Figure 28.

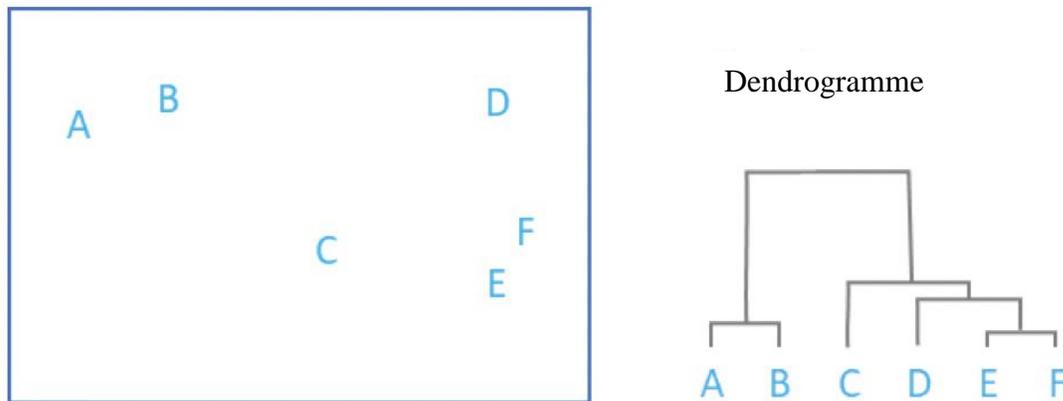


Figure 28: Dendrogramme présente la sortie principale du clustering hiérarchique.

3.3.1.1.2 Le partitionnement en k-moyennes

Le clustering K-means est l'un des algorithmes d'apprentissage automatique non supervisé le plus simple et le plus répandu. L'algorithme K-means dans l'exploration de données commence par un premier groupe de centroïdes sélectionnés au hasard, qui servent de points de départ pour chaque cluster, puis effectue des calculs itératifs (répétitifs) pour optimiser la position des centroïdes. Il arrête de créer et d'optimiser des clusters lorsque les centroïdes sont stabilisés (leurs valeurs ne changent pas, car la mise en cluster a réussi) ou que le nombre défini d'itérations a été atteint [135].

3.3.1.1.3 Modèles de mélange gaussien

Un modèle de mélange gaussien est un modèle probabiliste qui suppose que tous les points de données sont générés à partir d'un mélange d'un nombre fini de distributions gaussiennes avec des paramètres inconnus. Un modèle de mélange gaussien est paramétré par deux types de valeurs, les poids des composants du mélange et la moyenne des composants et les variances / covariances. Pour un modèle de mélange gaussien à composantes k , la k -ième composante a une moyenne de μ_k et une variance de σ_k pour le cas univarié et une moyenne de $\bar{\mu}_k$ et une matrice de covariance de Σ_k pour le cas multivarié [136].

Expectation maximization (EM) est la technique la plus couramment utilisée pour estimer les paramètres du modèle de mélange. L'algorithme Expectation-Maximization (EM) est un moyen itératif de rechercher des estimations du maximum de vraisemblance pour les paramètres de modèle lorsque les données sont incomplètes ou comportent des points de données manquants ou des variables cachées. EM choisit des valeurs aléatoires pour les points de données manquants et estime un nouvel ensemble de données. Ces nouvelles valeurs sont ensuite utilisées de manière récursive pour estimer de meilleures premières données, en remplissant les points manquants, jusqu'à ce que les valeurs soient corrigées [137].

3.3.1.1.4 Carte auto adaptative

Les cartes auto adaptatives, cartes auto-organisatrices ou cartes topologiques forment une classe de réseau de neurones artificiels fondée sur des méthodes d'apprentissage non supervisées. SOM fournit une technique de visualisation des données qui aide à comprendre

les données de grande dimension en réduisant les dimensions des données à une carte. SOM représente également le concept de regroupement en regroupant des données similaires [138]. Par conséquent, on peut dire que SOM réduit les dimensions des données et affiche des similitudes entre les données.

Avec SOM, le regroupement est effectué en mettant en concurrence plusieurs unités pour l'objet en cours. Une fois les données entrées dans le système, le réseau de neurones artificiels est formé en fournissant des informations sur les entrées. Le vecteur de poids de l'unité le plus proche de l'objet en cours devient l'unité gagnante ou active. Au cours de la phase d'apprentissage, les valeurs des variables d'entrée sont progressivement ajustées afin de préserver les relations de voisinage existant dans l'ensemble de données d'entrée. À mesure qu'il se rapproche de l'objet en entrée, les poids de l'unité gagnante sont ajustés, ainsi que de ses voisins.

En résumé, l'apprentissage se fait en plusieurs étapes et au cours de nombreuses itérations. :

1. Les poids de chaque nœud sont initialisés.
2. Un vecteur est choisi au hasard dans l'ensemble des données d'apprentissage.
3. Chaque nœud est examiné pour calculer les poids qui ressemblent le plus au vecteur d'entrée. Le nœud gagnant est généralement appelé unité de meilleure correspondance (Best Matching Unit BMU).
4. Ensuite, le voisinage du BMU est calculé. La quantité de voisins diminue avec le temps.
5. Le poids gagnant est récompensé par le fait de ressembler davantage au vecteur échantillon. Les voisins deviennent également davantage comme le vecteur échantillon. Plus un nœud est proche du BMU, plus ses poids sont modifiés et plus le voisin est éloigné du BMU, moins il en apprend.
6. Répétez l'étape 3 pour N itérations.

3.3.1.2 Mesures d'évaluation de la segmentation

Le meilleur algorithme de segmentation est sélectionné par quatre mesures d'évaluation de la classification, décrites ci-après.

- Le critère de Davies-Bouldin (DB) repose sur la relation entre les mesures inter-clusters et intra-clusters. En particulier, l'index DB est donné par :

$$DB = \frac{1}{k} \sum_{i=1}^k R_i \quad (3.10)$$

Où

$$R_i = \max_{i=1, \dots, k; i \neq j} D_{ij} \quad (3.11)$$

Et

$$D_{ij} = \frac{\bar{d}_i + \bar{d}_j}{d_{ij}} \quad (3.12)$$

Où k est le nombre de partitionnements, \bar{d}_i et \bar{d}_j sont les distances moyennes entre chaque échantillon dans les $i^{\text{ème}}$ et $j^{\text{ème}}$ cluster jusqu'à leurs centroïdes respectifs \bar{C}_i et \bar{C}_j , et d_{ij} est la distance euclidienne entre ces centroïdes. La plus petite valeur de DB correspond à la meilleure méthode de segmentation [139].

- Le critère de Calinski-Harabasz (CH) est décrit par le ratio de variance suivant :

$$CH = \frac{A N - k}{B k - 1} \quad (3.13)$$

Où A est la variance globale entre les clusters, B est la variance globale au sein des clusters et N le nombre d'observations. La variance globale entre les clusters A est définie comme suit:

$$A = \sum_{i=1}^k n_i \|C_i - \mathbf{m}\|^2 \quad (3.14)$$

Où \mathbf{m} est la moyenne globale des données de l'échantillon et n_i sont les échantillons de la $i^{\text{ème}}$ cluster, et C_i est le centroïde de la $i^{\text{ème}}$.

La variance globale intra-cluster B est défini comme suit:

$$B = \sum_{i=1}^k \sum_{x \in c_i} \|x - m_i\|^2 \quad (3.15)$$

Où x est un échantillon et c_i est le $i^{\text{ème}}$ cluster. La meilleure segmentation est la solution offrant la plus haute valeur Calinski-Harabasz [140].

- L'indice de Dunn (DI) est le rapport entre la distance de séparation minimale entre deux éléments de classes différentes et la distance maximale entre deux éléments de la même classe. C'est un indice qui ne repose pas sur une distance particulière et qui peut donc être utilisé dans une grande variété de situations. Il est exprimé comme suit

$$DI = \frac{d_{\min}}{d_{\max}} \quad (3.16)$$

Où d_{\min} est la distance minimale entre les échantillons de différents clusters et d_{\max} est la plus grande distance intra-cluster. Ces deux distances sont définies par

$$d_{\min} = \min_{i \neq j} d_{ij} \quad (3.17)$$

$$d_{\max} = \max_{1 \leq i \leq k} D_i \quad (3.18)$$

Où d_{ij} est la distance entre les groupes c_i et c_j définie par la distance entre leurs échantillons les plus proches et D_i est la plus grande distance séparant deux échantillons différents dans le même groupe. La meilleure méthode de segmentation est celle qui donne la valeur la plus élevée de DI [141].

La dernière mesure utilisée pour l'évaluation est l'indice de Hartigan (HI), qui est calculé comme suit

$$HI(k) = (n - k - 1) \frac{E(k) - E(k + 1)}{E(k + 1)} \quad (3.19)$$

Où n est le nombre d'entités (dans notre cas, les pixels à classer), k est le nombre de clusters, $k = 1, 2, \dots, n - 2$ et $E(k)$ est donné par

$$E(k) = \sum_{j=1}^k \sum_{i \in c_j} (\mathbf{x}_i - C_j)(\mathbf{x}_i - C_j)^T \quad (3.20)$$

\mathbf{x}_i est le $i^{\text{ème}}$ vecteur p-dimensionnel des échantillons du cluster c_i et C_j est le centroïde de la $j^{\text{ème}}$ cluster. La méthode de segmentation optimale est celle qui minimise le HI [142].

3.3.1.3 Méthode d'apprentissage supervisée

3.3.1.3.1 Le perceptron multicouche (multilayer perceptron)

Le MLP est un réseau de neurones artificiels organisés en couches et où l'information circule dans une direction, de la couche d'entrée à la couche de sortie.

Les neurones sont connectés les uns aux autres par des connexions pondérées. Ce sont les poids de ces connexions qui régissent le fonctionnement du réseau et "programment" une application de l'espace des entrées à l'espace de sortie au moyen d'une transformation non linéaire [143], [144].

La valeur des sorties dépend des paramètres liés à la structure du réseau de neurones : connectivité entre les neurones, fonctions d'agrégation et d'activation, ainsi que pondérations. Les différences entre ces sorties et les sorties souhaitées forment des erreurs qui sont corrigées via l'algorithme de back propagation (BP). La manière de quantifier cette erreur peut varier en fonction du type d'apprentissage à effectuer. BP est une procédure à gradient itératif qui tend à réduire l'erreur globale pas à pas. Il est toutefois possible que les poids soient empilés dans un minimum local. C'est pourquoi un terme de momentum est généralement ajouté à la formule de propagation pour aider l'algorithme de gradient à sortir de ce minimum local [145].

3.3.1.3.2 K-voisins les plus proches (K-Nearest Neighbors KNN)

La méthode KNN est une méthode de voisinage non paramétrique, cela signifie que l'algorithme permet de classifier sans faire d'hypothèses sur la fonction qui relie la variable dépendante aux variables indépendantes [146]. Dans le contexte de la classification d'une nouvelle observation x , l'idée simple est de faire voter cette observation par les voisins les plus

proches. La classe de x est déterminée en fonction de la classe de majorité parmi les K voisins les plus proches de l'observation x .

3.3.1.3.3 Naive Bayes (NB)

Classificateur NB est un algorithme populaire dans l'apprentissage automatique. C'est un algorithme d'apprentissage supervisé utilisé pour la classification. Le classifieur NB est basé sur le théorème de Bayes et les probabilités conditionnelles. Un avantage de cette méthode est la simplicité de la programmation, la facilité d'estimation des paramètres et sa rapidité (même sur de très grandes bases de données) [147], [148].

3.3.1.3.4 La forêt aléatoire (Random Forest RF)

RF est un ensemble d'arbres décisionnels binaires permettant de voter pour la classe la plus populaire. Par conséquent, plutôt que d'utiliser toutes les données pour construire les arbres, nous choisissons au hasard pour chaque arbre un sous-ensemble de données de sorte les arbres se distinguent les uns des autres par le sous-échantillon de données sur lequel ils ont été formés. En utilisant des ensembles d'arbres, nous pouvons améliorer considérablement la précision de prévision des nouvelles données et obtenir une reconnaissance très rapide, ce dont nous avons besoin pour notre application. En outre, il est efficace sur les gros intrants sans oublier la robustesse des valeurs aberrantes [149].

3.3.1.3.5 Une machine à vecteurs de support (Support Vector Machine SVM)

SVM est un réseau de classification binaire formé par un algorithme supervisé. Le but de l'algorithme SVM est de trouver la séparation entre deux classes en partant du principe que plus la séparation est grande, plus la classification est robuste. L'algorithme sélectionne l'hyperplan qui sépare les données en deux classes distinctes afin de maximiser la distance entre l'hyperplan et les observations les plus proches de l'échantillon d'apprentissage. Les points les plus proches utilisés pour la détermination de l'hyperplan sont appelés vecteurs de support [150].

Tant que nous voulons faire une classification de quatre classes et que le SVM est un classificateur binaire, il est nécessaire d'introduire la multiclass SVM, ceci peut être appliqué de deux manières One-vs-All et One-vs-One :

- Dans le premier cas, le problème original de la classe multiple K est traité avec K classificateurs et le classement est donné par le classificateur qui répond le mieux.
- Dans le second cas, le problème est transformé en $k \times \frac{k-1}{2}$ classificateurs binaires : toute classe est en effet comparée à chaque autre. Le classement est établi par un vote à la majorité ou par un graphe de décision acyclique [151].

3.3.1.3.6 L'analyse discriminante linéaire (Linear discriminant analysis LDA)

LDA est une méthode de classification qui suppose que différentes classes génèrent des données basées sur différentes distributions gaussiennes. Il est basé sur le concept de recherche d'une combinaison linéaire de variables (prédicteurs) qui séparent au mieux deux classes. Chaque classe (Y) génère des données (X) en utilisant une distribution normale multivariée. C'est-à-dire que le modèle suppose que X a une distribution de mélange gaussienne. Pour l'analyse discriminante linéaire, le modèle a la même matrice de covariance pour chaque classe, seules les moyennes varient [152].

3.3.2 Segmentation sémantique profonde

La segmentation sémantique est un élément de base de la compréhension de l'image. En classant tous les pixels d'une image de manière dense, il est alors possible de construire des représentations abstraites intéressées par les objets et leurs formes. Ses applications incluent la compréhension d'images, la conduite autonome, la reconnaissance d'objets, la traduction automatique et la vision industrielle. La segmentation sémantique est un domaine de recherche très actif en raison de son importance dans les applications du monde réel [153], [154], [155], [156]. L'objectif est de classer chaque pixel de l'image dans une catégorie prédéfinie spécifique, i.e., Destinée à étiqueter chaque pixel de l'image comme appartenant à une classe. Cette affectation est faite en fonction du contexte des différents objets de l'image.

3.3.2.1 Réseaux entièrement convolutionnels (Fully Convolutional Networks FCN)

Le réseau FCN est une extension du CNN classique. L'idée principale est de faire en sorte que le CNN classique prenne en entrée des images de taille arbitraire [157]. Le FCN est construit uniquement à partir de couches connectées localement, telles que la convolution, le pooling et l'upsampling. Cela réduit le nombre de paramètres et le temps de calcul. Les principales contributions du réseau FCN sont la vulgarisation de l'utilisation de réseaux convolutionnels de bout en bout pour la segmentation sémantique et l'introduction de connexions sautées pour améliorer la grosseur de l'upsampling. Il existe des variantes de l'architecture FCN, qui diffèrent principalement par la précision spatiale de leur sortie. Pour notre application, nous utiliserons les variantes FCN-32, FCN-16 et FCN-8. La Figure 29 montre l'architecture du modèle.

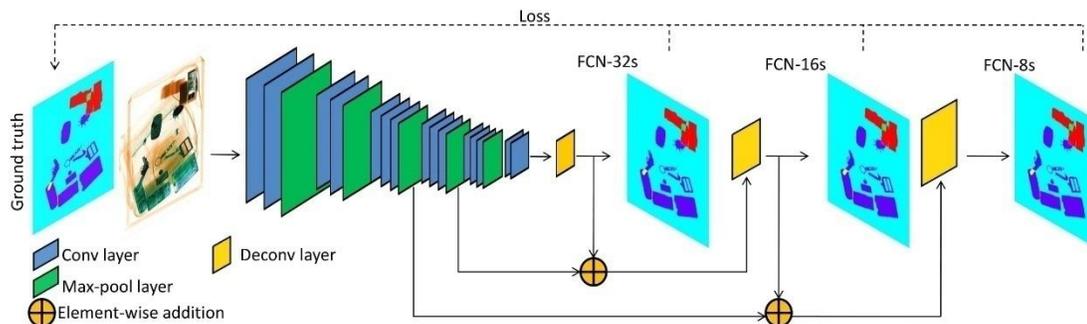


Figure 29: L'architecture du FCN.

3.3.2.2 SegNet

SegNet est une architecture de codeur / décodeur convolutionnelle profonde. Le codeur applique la convolution, la normalisation par lots (batch normalization) et une opération de non-linéarité à regroupement maximal, tout en stockant l'index de la valeur extraite de chaque fenêtre. Le décodeur est similaire au codeur, à la différence qu'il n'a pas de non-linéarité et qu'il sur-échantillonne son entrée en utilisant les index précédemment stockés à l'étape de codage. L'architecture du modèle est illustrée à la Figure 30. Après la phase de décodage, la sortie est acheminée vers une couche de classification finale au niveau du pixel [158].

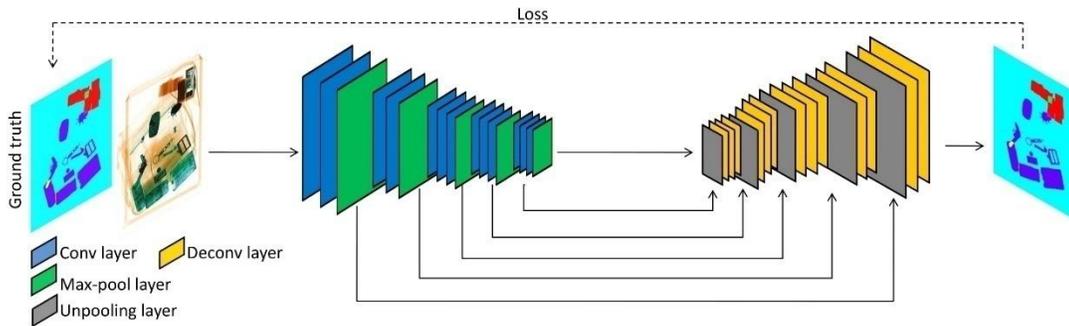


Figure 30: L'architecture du SegNet.

3.3.2.3 U-Net

U-Net est un CNN profond basé sur une architecture codeur-décodeur spécialement conçu pour la segmentation d'images biomédicales. L'architecture illustrée dans la Figure 31 contient deux chemins. Le premier chemin est le codeur qui est une suite traditionnelle de couches de convolution et de max-pooling. Le second chemin est le décodeur (chemin à développement symétrique) qui permet une localisation précise à l'aide de convolutions transposées [159].

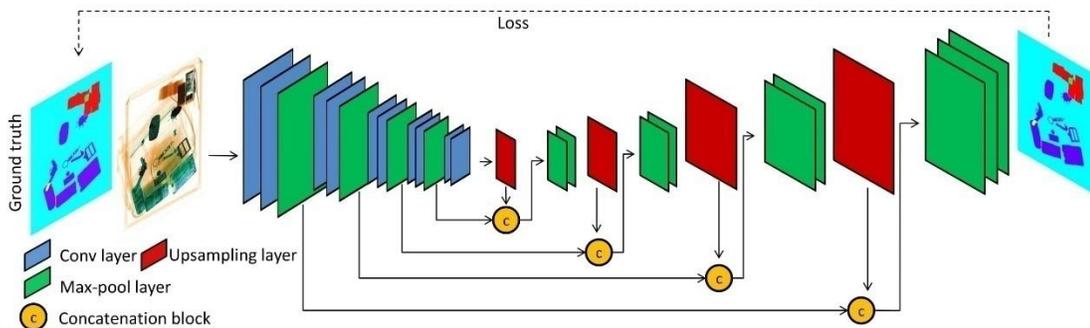


Figure 31: L'architecture de l'U-Net.

3.3.2.4 DeconvNet

DeconvNet proposé dans [160] est un CNN profond basé sur une architecture codeur-décodeur combinant des propositions de région ascendante avec un réseau de déconvolution à plusieurs couches. La première partie est un réseau de convolution semblable à celui du réseau FCN, avec des couches de convolution et de pooling. La deuxième partie est le réseau de déconvolution qui est une nouvelle partie. Le réseau de déconvolution est composé de couches

de déconvolution et de couches de unpooling qui identifient les étiquettes de classe en pixels et prédisent les masques de segmentation. L'architecture globale est montrée dans la Figure 32.

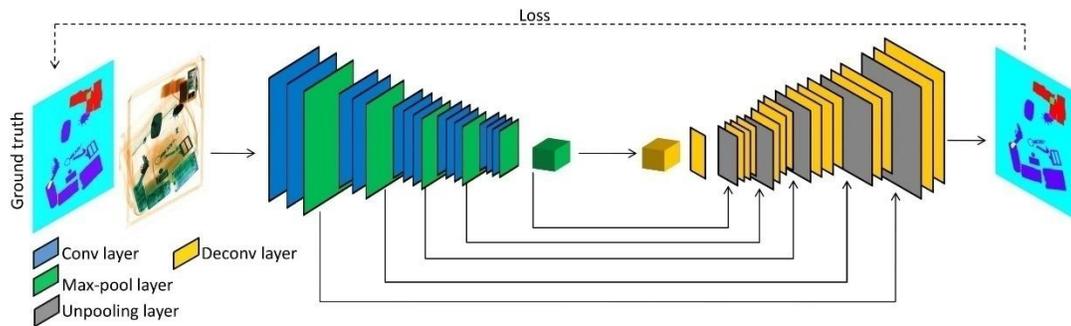


Figure 32: L'architecture du DeconvNet.

3.3.2.5 RedNet

RedNet est un réseau résiduel codeur-décodeur qui combine la connexion saut courte dans l'unité résiduelle et la connexion saut longue entre le codeur et le décodeur pour une inférence sémantique précise [161]. Les couches sont composées par couche de convolution, couche de déconvolution et addition. Afin de ne pas perdre les informations de localisation et de dépenser de la mémoire supplémentaire, les auteurs débarrassent le pooling maximal. L'architecture de ce réseau est illustrée à la Figure 33.

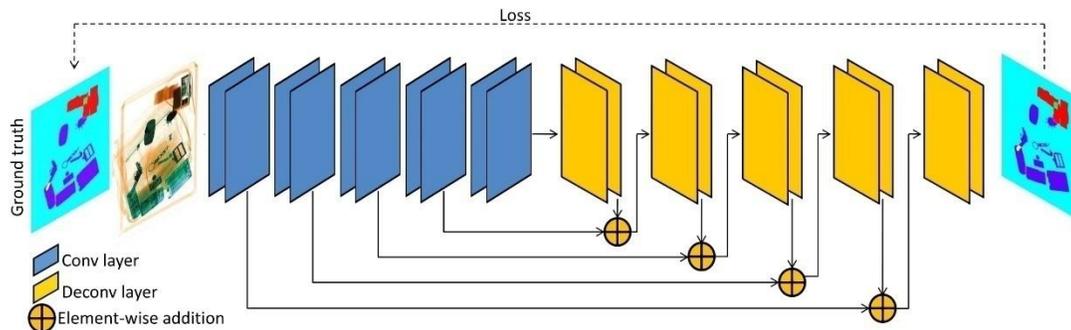


Figure 33: L'architecture du RedNet.

3.3.2.6 Notre modèle proposé (SemSegAAE)

Dans ce travail, une méthode de segmentation sémantique basée Adversarial AutoEncoder (AAE) est présentée. Cette méthode est très efficace en termes de mémoire, d'exploitation, de temps de calcul et qu'elle est compétitive par rapport aux l'état de l'art de la segmentation

sémantique. De cette manière, l'architecture proposée est conçue pour former un classifieur de label de pixel de bout en bout au moyen d'une descente de gradient stochastique et de la fonction de perte d'entropie croisée. Un schéma de SemSegAAE est présenté dans la Figure 34.

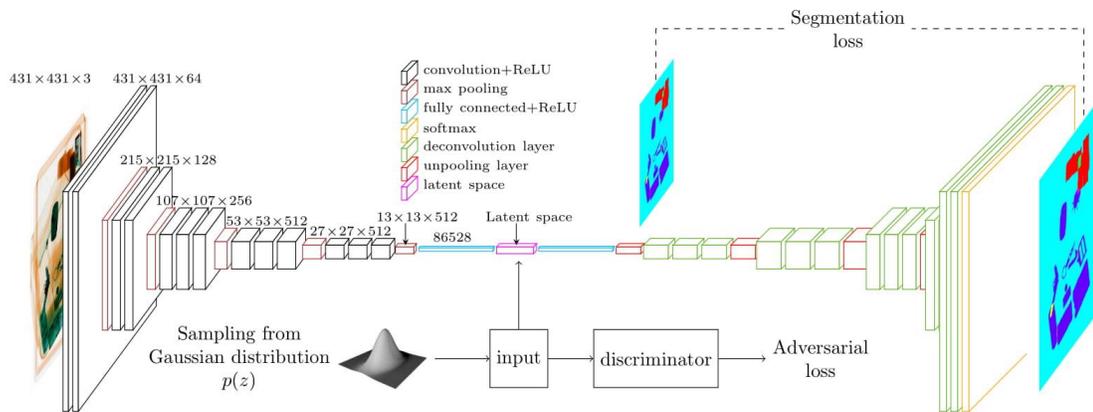


Figure 34: L'architecture du système de segmentation sémantique proposée (SemSegAAE).

AAE est similaire à un autoencodeur, mais le codeur est formé en manière « défavorable » (adversely trained) pour le forcer à produire une distribution d'espace latent spécifique. AAE entretient une relation directe avec les réseaux génératifs d'adversaire (GAN), organisés en trois blocs principaux : le codeur, le décodeur et le discriminateur [162]. Le codeur prend une entrée x et produit une sortie h (espace latent) en compressant l'image de sorte qu'elle occupe un espace dimensionnel inférieur. Le décodeur prend le h et tente de reconstruire l'entrée à sa sortie en minimisant la perte de reconstruction entre l'image d'origine et l'image de sortie du décodeur. Le discriminateur permet de discriminer entre des échantillons générés à l'aide d'une distribution préalable et des échantillons à partir de code caché.

La formation d'un AAE comporte deux phases : la phase de reconstruction et la phase de régularisation. Le premier vise à former à la fois le codeur et le décodeur pour minimiser la perte de reconstruction, le second dans le but de former le discriminateur (pour classifier la sortie du codeur et l'entrée aléatoire) et le générateur (forçant le codeur à sortir du code latent avec distribution souhaitée) [162].

Nous avons choisi AAE pour apprendre l'espace latent des fonctionnalités, car il peut produire des vecteurs de caractéristiques avec une distribution souhaitée. Dans ce travail, l'AAE est

formé afin de produire une distribution gaussienne symétrique de l'espace des fonctions. De cette manière, toutes les entités de l'espace latent ne sont pas corrélées. i.e. chaque entité, capture une caractéristique unique et spécifique de l'image d'entrée. En d'autres termes, AAE agit comme un puissant extracteur de fonctionnalités.

3.3.2.6.1 Architecture proposée

Notre approche de segmentation suit une architecture CNN codeur-décodeur symétrique utilisant un réseau AAE. Le codeur de l'AAE est identique aux couches de convolution de VGG16 [163]. Il est constitué de couches convolutives de filtres 3×3 , répétés plusieurs fois, modifiant le nombre de noyaux. Pour les besoins d'extraction robuste, les opérations de normalisation par lots (batch normalization) [164] et ReLu [165] sont effectuées après l'opération de convolution. Le réseau de décodage reproduit le processus de codage dans l'ordre inverse. Les réseaux de décodeurs se terminent par un classifieur de pixels avec une couche de sortie softmax. Cette architecture est montrée dans la Figure 34.

Le réseau est formé sur un GPU NVIDIA GeForce GTX 1050ti. Nous arrêtons l'entraînement lorsque la perte de segmentation ne diminue plus.

3.3.2.6.2 Préparation des données

Les images utilisées dans cette étude provenaient de la base de données privée HDTS (Hi-Tech Detection Systems Society) contenant 2000 images de bagages à rayons X numérisées avec une taille moyenne de 800 pixels par 800 pixels. Pour résoudre le problème des objets qui se chevauchent, une segmentation manuelle a été préparée en étiquetant chaque pixel comme appartenant à l'une des six classes suivantes : objets organiques, objets inorganiques, chevauchement entre organique / organique, chevauchement entre inorganique / inorganique, chevauchement entre organique / inorganique et le fond.

La préparation du ground truth et la procédure de détection d'objet sont illustrées dans un diagramme synoptique à la Figure 35 (a été effectuée de manière manuelle avec l'aide d'un expert). Dans l'image d'origine, deux cercles orange et deux carrés bleus représentent respectivement deux objets organiques et inorganiques. L'image de ground truth montre les six catégories souhaitées dans lesquelles le pixel de l'image d'origine doit être classé. On peut

observer qu'il existe trois types de chevauchement : organique / organique, inorganique / inorganique et organique / inorganique. Après application de la méthode de segmentation sémantique proposée, six masques correspondant aux six classes sont obtenus. Enfin, la phase de détection des objets est terminée en utilisant les trois règles qu'on va les mentionner dans la Section 3.4.1.

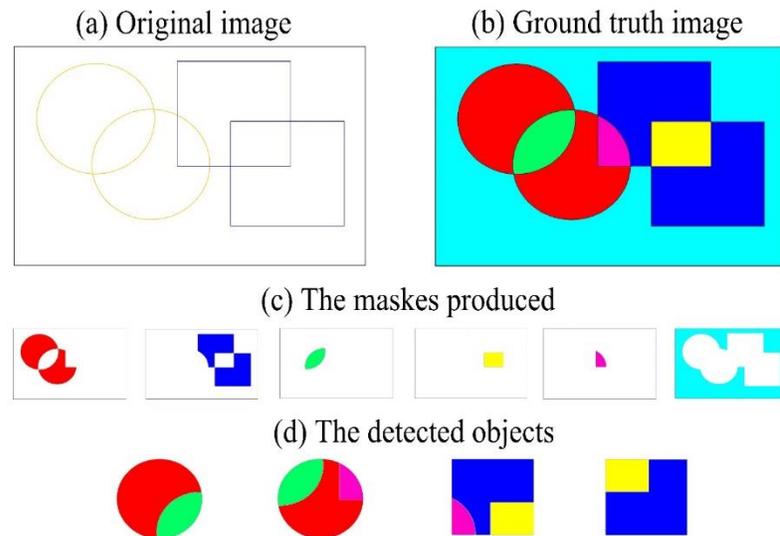


Figure 35: Un schéma synoptique illustre la préparation du ground truth et de la procédure de détection d'objet. Dans l'image d'origine (a), deux cercles orange et deux carrés bleus représentent respectivement deux objets organiques et inorganiques. L'image de ground truth (b) montre les six catégories souhaitées dans lesquelles le pixel de l'image d'origine doit être classé. On peut observer qu'il existe trois types de chevauchement : organique / organique, inorganique / inorganique et organique / inorganique. Après application de la méthode de segmentation sémantique proposée, six masques correspondant aux six classes sont obtenus. Après la phase de détection d'objet, les objets en (d) seront détectés.

3.4 Détection d'objets

Une fois que l'image segmentée est obtenue, elle est binarisée par seuillage à l'aide de la méthode d'Otsu et en appliquant la technique de marquage des composants associés (labeling technique of related components), regroupés dans un ensemble appelé composant connecté. Chaque ensemble est séparé des autres et peut ensuite être facilement isolé [166], [167], [168]. L'identification de ces objets permet de prédire la catégorie de l'objet dans les étapes suivantes.

Nous voudrions mentionner que pour notre processus de segmentation, la détection des objets sera divisée en deux vues, organique et inorganique.

3.4.1 Après notre modèle de segmentation proposé

Le processus de détection d'objet est mis en œuvre après la tâche de segmentation, en appliquant ces trois règles de fusion :

- Un chevauchement entre deux objets du même matériau est considéré comme appartenant aux deux.
- Un chevauchement entre deux matériaux différents est considéré comme appartenant aux deux.
- L'arrière-plan sera ignoré.

Après cette opération, nous pouvons obtenir deux masques, l'un pour les objets organiques et l'autre pour les objets inorganiques, en appliquant la technique de marquage des composants associés [168] effectuée un étiquetage bidimensionnel des composants (les relations de contiguïté des rangées de pixels, quant à elles, analysent les relations de contiguïté de pixels individuels), nous pouvons détecter tous les objets présentés dans le bagage.

3.4.2 Utilisant le deep learning

Pour notre travail, nous avons proposé une détection d'objet utilisant une version modifiée de la méthode YOLOv3. En particulier, cette méthode est utilisée, car elle permet d'obtenir de meilleurs résultats. Quelques informations sur YOLOv3 et la modification appliquée sont décrites ci-après.

YOLOv3 a été choisi, car elle offre actuellement les meilleures performances de détection dans le benchmark PASCAL VOC, tout en conservant le calcul en temps réel, critère très important pour les systèmes de sécurité. YOLO est un réseau de neurones de convolution de bout en bout qui détecte des objets en fonction d'une prédiction de boîtes de sélection et de probabilités de classe. La première version de YOLO [169] a été publiée en 2016, suivie de la version 2 en 2017 [170] et de la dernière version en 2018 [171], qui inclut de nouvelles fonctionnalités ainsi que des modifications du réseau existant. . La dernière version est celle utilisée dans cette étude.

Le réseau divise l'image d'entrée en une grille $S \times S$. Si le centre d'un objet est situé dans une grille, il appartiendra à cette grille en tant qu'objet détecté. Chaque grille prédit les cadres de sélection et la confiance correspondante (dans la manière dont une boîte contient un objet), ainsi que la précision de la boîte. La grille prédit également la probabilité qu'un objet appartienne à une classe. Le point fort de YOLOv3 est qu'il examine l'ensemble de l'image au cours de la formation (diminue le temps de calcul) en aidant à acquérir des informations contextuelles sur une classe donnée et son environnement.

Au lieu de prédire 5 boîtes de limites arbitraires comme le fait YOLOv2, YOLOv3 prédit 9 boîtes d'ancrage (anchors) avec 3 échelles différentes (les 9 boîtes de limites les plus répétitives calculées par la mise en cluster à k-means), des détections à différentes échelles aident à résoudre le problème de la détection de petits objets, lesquels est un problème avec YOLOv2. Avec cette technique, la diversité des prédictions peut être maintenue et chaque prédiction se concentre sur une forme spécifique. Par conséquent, la formation initiale sera plus stable.

Dans notre étude, nous utilisons une version modifiée de YOLOv3. Notre modification est proposée dans la méthode de calcul d'ancrage, dans laquelle les algorithmes de regroupement des k-moyennes sont remplacés par la carte auto-organisée (SOM), car il a été prouvé qu'elle produisait de meilleurs résultats.

Une autre question importante qui sera examinée est l'effet de la variation de la taille de l'image d'entrée sur les performances de la détection. Ainsi, nous appliquerons 416×416 , 512×512 et 608×608 comme taille d'entrée pour former notre modèle.

3.5 Extraction de caractéristiques

Lorsque les objets de l'image ont été détectés, l'extraction et la sélection des caractéristiques sont effectuées afin de faciliter la classification finale. Plusieurs caractéristiques pertinentes utiles pour distinguer les objets dangereux des objets inoffensifs ont été sélectionnées. Ceux-ci sont :

3.5.1 Caractéristiques de formes

- Surface : correspond au nombre de pixels de la région.
- Périmètre : est la distance autour de la limite de l'objet.
- Convexité : mesure de la courbure, il s'agit du rapport entre le périmètre convexe de l'objet et le périmètre arrière de l'objet.
- Excentricité : Après avoir limité notre objet à l'ellipse minimale, l'excentricité est le rapport entre la distance entre les foyers de cette ellipse et la longueur de son axe principal.
- Rectangularité : nous indique combien notre objet est similaire au rectangle. Après avoir limité notre objet dans le rectangle minimal, nous calculons le rapport entre l'aire de ce rectangle sur l'aire de la surface de l'objet.
- Solidité : nous donne l'information sur la proportion de pixels dans la coque convexe, c'est le rapport entre la surface et la surface convexe.
- Circularité : indique combien notre objet est similaire au cercle. C'est le rapport entre l'aire et le carré du périmètre.

3.5.2 Caractéristiques de couleur

Moyenne, variance, écart type, skewness, kurtosis (pour les 3 canaux RGB).

3.5.3 Caractéristiques de texture

La texture est utilisée pour caractériser un aspect homogène de la surface d'un objet. Pour cela, nous avons utilisé la matrice de cooccurrence de niveaux de gris (Gray-Level Co-Occurrence Matrix GLCM) pour extraire les caractéristiques suivantes :

Contraste, dissimilarité, énergie, entropie, homogénéité, différence inverse normalisée, moment de différence inverse normalisée, permission de cluster, nuance de cluster, autocorrélation et corrélation.

3.5.4 Notre modèle proposé pour l'extraction des caractéristiques

Une fois que les objets de l'image radiographique ont été détectés, leurs caractéristiques sont apprises par un réseau autoencodeur (AE). Un AE utilise une structure symétrique composée de deux blocs principaux : une partie codeur qui compresse l'entrée dans une représentation de petite taille (qui contient les données des caractéristiques) et une partie décodeur qui est conçue pour reconstruire l'entrée des caractéristiques extraites par le codeur. La Figure 36 décrit la procédure d'apprentissage proposée.

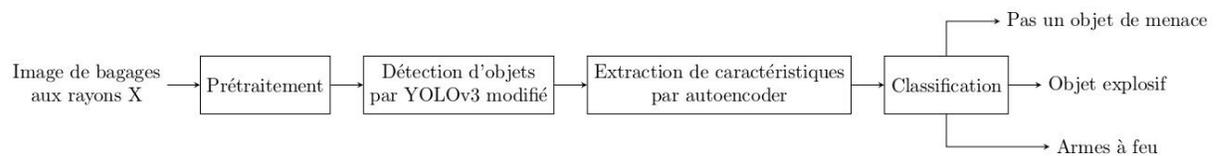


Figure 36: La procédure d'apprentissage proposée en utilisant l'apprentissage des caractéristiques avec l'adversarial autoencodeur étiqueté.

Dans l'approche proposée, nous utiliserons le terme "Adversarial AE étiqueté" (AAE) [162]. Cette méthode apprend les paramètres de caractéristiques dans une distribution de probabilité (pour notre cas, la distribution gaussienne) nous donnant une distribution spécifique de chaque catégorie de classe. Au moyen d'un schéma d'apprentissage supervisé, un AAE représente l'entrée dans un espace latent en rendant la distribution latente aussi proche que possible d'une distribution "préalable" souhaitée. Le principal avantage du choix de l'AAE par rapport à l'AE standard est que les informations préalables permettent l'injection de connaissances de domaine, permettant d'estimer l'incertitude de la prévision et de la rendre plus adaptée à notre système. Par conséquent, la sortie du système de détection est l'entrée de l'AAE, pour obtenir les caractéristiques de l'objet comme sortie AAE.

Nous examinerons l'effet de la modification de la dimension de l'espace latent (nous l'augmentons jusqu'à ce que la performance de classification se stabilise ou diminue). Notre système ne se concentre pas sur l'évaluation des images générées. L'objectif est plutôt de déterminer la représentation latente à utiliser comme espace de caractéristiques.

3.5.4.1 Adversarial Autoencoder étiqueté (labeled Adversarial Autoencoder)

Afin de mieux structurer la distribution du code caché et de mapper les codes latents de chaque classe sur une distribution gaussienne spécifique, nous allons utiliser un AAE régularisé. Pour cela, un code « one-hot » de l'étiquette est transmis au discriminateur. Dans notre implémentation, nous avons utilisé un mélange de 3 distributions gaussiennes (explosif, arme à feu et objets non menaçants). Nous avons formé ce système de manière supervisée. À cette fin, aucune dimension supplémentaire n'est supposée pour le vecteur one-hot. Le schéma proposé est montré dans la Figure 37.

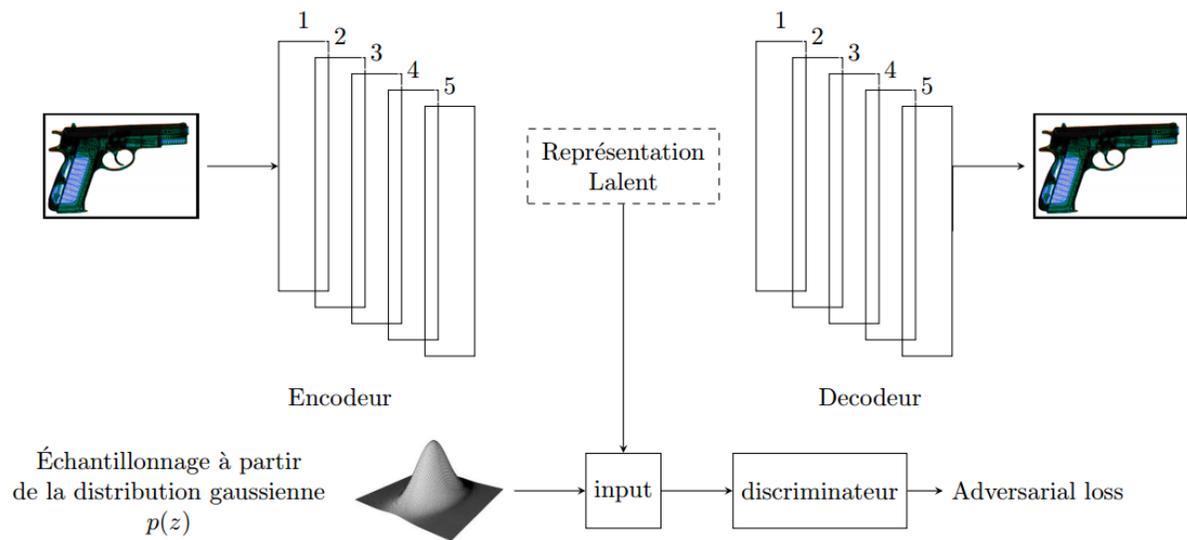


Figure 37: Schéma de l'adversarial Autoencoder étiqueté proposé.

Le codeur et le décodeur sont construits selon le Tableau 5. Les deux architectures sont composées d'une séquence de blocs de couches convolutives, chacune étant suivie par une procédure de normalisation et des activations de Leaky Rectified Linear Unit (LReLU). Toutes les couches convolutives reposent sur un noyau de taille 3×3 . Le réseau de codage se termine par une couche entièrement connectée, où d est égal à la dimension de la représentation latente \mathbf{z} . Le réseau de décodage réplique le processus de codage dans l'ordre inverse. Il commence par mapper les entités extraites dans un espace de dimensions $7 \times 7 \times 128$ en utilisant un réseau entièrement connecté. Chaque bloc de couches est composé d'une couche convolutionnelle,

suivie d'une convolution à pas fractionné -fractionally strided convolution- (convolution transposée de la foulée 2). La dernière convolution mappe ces activations au domaine de valeur de pixel RGB.

Tableau 5: Structures des réseaux de codeurs et de décodeurs. Dans lequel les blocs de décodage reproduisent le processus de codage dans l'ordre inverse.

Couche de convolution	Noyau	Stride	Padding	Sortie
conv(nb=32, LReLU)	3x3	2	0	25x25x32
conv(nb=64, LReLU)	3x3	1	1	25x25x64
conv(nb=64, LReLU)	3x3	2	1	13x13x64
conv(nb=128, LReLU)	3x3	1	1	13x13x128
conv(nb=128, LReLU)	3x3	2	1	7x7x128
Flatten	N/A	N/A	N/A	6272
Réseau entièrement connecté				
(nb=d)	N/A	N/A	N/A	d

Le discriminateur est composé de 2 réseaux entièrement connectés de 64 neurones chacun, suivis d'une normalisation et de LReLU. La troisième couche se termine par un seul neurone de sortie.

La formation de l'AAE comporte deux phases : la phase de reconstruction et la phase de régularisation. Le premier vise à former à la fois le codeur et le décodeur afin de minimiser les pertes de reconstruction. La deuxième phase est composée de deux tâches. Dans le premier cas, le discriminateur est formé pour faire la distinction entre la sortie du codeur et les échantillons générés à partir des distributions connues ; dans la deuxième tâche, les poids du discriminateur sont fixes et le codeur est entraîné à tromper le discriminateur. De cette manière, la distribution spatiale latente tend vers les distributions connues [162]. Tous les paramètres des trois composants sont mis à jour simultanément à chaque itération.

Après avoir eu les caractéristiques, nous avons choisi de les classer avec Support Vector Machine (SVM). SVM est un classificateur binaire. Pour notre application, nous avons 3 classes (objets explosifs, armes à feu, objets non menaçants). Donc, la classification multi-

classes est nécessaire. SVM peut être étendu pour des problèmes multi-classes en utilisant l'approche dite one-vs-all [172]. Pour donner à notre modèle plus de flexibilité pour rechercher un espace entre les vecteurs de support, nous allons former notre modèle avec le noyau polynomial. Les algorithmes d'optimisation et leurs vitesses d'apprentissage sont choisis de telle sorte que le réseau converge correctement.

3.6 Sélection des caractéristiques

À partir de l'ensemble des caractéristiques extraites, une procédure permettant de les classer est appliquée. Nous utilisons le modèle "filtre" (indépendant de tout algorithme d'apprentissage) qui effectue l'étape de classement. Il a pour objectif de sélectionner le sous-ensemble de caractéristiques le plus pertinent en fonction de certains critères d'évaluation afin de classer les caractéristiques du meilleur au pire. Cette évaluation est faite indépendamment de l'étape de classification.

La sélection des fonctionnalités nous permet de simplifier les modèles, d'améliorer la généralisation en réduisant le surapprentissage, de réduire les effets de la malédiction de la dimensionnalité et de réduire le temps d'apprentissage.

3.6.1 Brève description des méthodes de sélection des caractéristiques utilisées

3.6.1.1 Sélection de caractéristiques latentes infinies (Infinite Latent Feature Selection)

Dans [173], les auteurs présentent un algorithme de sélection de caractéristiques basé sur un graphe latent probabiliste. Son objectif est de classer les caractéristiques du meilleur au pire. Pour cela, il suit trois étapes. La première consiste à mapper les valeurs de distribution des entités \bar{x} à un petit ensemble nominal (quantification). Cette étape permet d'affecter un descripteur f_i à chaque fonctionnalité brut x_i . Deuxièmement, créez un graphique entièrement connecté non dirigé pour modéliser la probabilité que les entités soient pertinentes les unes par

rapport aux autres, en utilisant l'analyse sémantique probabiliste latente (PLSA) pour apprendre automatiquement les poids. La dernière étape est le classement par validation croisée.

3.6.1.2 Score laplacien

Le score laplacien est intrinsèquement basé sur les techniques spectrales de réduction de la dimensionnalité (cartes propres laplaciennes) et sur les projections de préservation de la localisation (LPP). Cette méthode évalue les caractéristiques et les classe en fonction de leur pouvoir de préservation de localité [174].

3.6.1.3 Sélection de caractéristiques à partir de corrélations d'ordre supérieur

Cette méthode est basée sur l'idée que les bonnes caractéristiques sont celles qui sont fortement corrélées à la variable prédite, mais non corrélées les unes aux autres [175].

3.6.1.4 Algorithme Relief

L'algorithme Relief est limité à des problèmes à deux classes [176]. C'est l'une des principales méthodes rapides de sélection des fonctionnalités de filtrage. L'idée est d'estimer la qualité des entités en fonction de la capacité de leurs valeurs à distinguer des instances proches les unes des autres. Le score d'une entité est d'autant plus grand que les données d'une même classe ont des valeurs similaires et que les données de différentes classes sont bien séparées.

3.7 Classification

Pour cette tâche, nous avons utilisé les mêmes méthodes que celles utilisées dans la tâche de segmentation basée sur le clustering en utilisant l'apprentissage automatique dans la Section 3.3.1.3. Pour cette étape, nous avons effectué une optimisation d'hyper-paramètres afin de choisir les optimaux pour les algorithmes d'apprentissage en modifiant plusieurs fois ces hyper-paramètres de chaque méthode et en traçant l'erreur quadratique moyenne. Ces paramètres sont le nombre de nœuds pour le MLP, le nombre d'arbres pour le RF et le nombre de K pour le KNN. Pour le cas de SVM, trois noyaux différents sont évalués : linéaire, RBF et polynomial. Après cela, nous répétons la formation et l'évaluation pour choisir la méthode de

sélection des caractéristiques (avec un numéro de caractéristiques spécifique) et donner les meilleures performances avec chaque méthode (en comparant la précision du jeu de tests).

Au final, nous avons obtenu la meilleure méthode de classification avec un hyper-paramètre spécifique, le nombre optimal de caractéristiques classées par méthode de sélection de caractéristiques spécifiques, qui nous donne la meilleure précision possible, la précision de la variance et les temps de calcul.

3.8 Conclusion

Différentes étapes liées aux traitements d'images sont abordées dans ce travail. Ainsi, le choix de la méthode d'amélioration de l'image la plus performante est une phase plus qu'importante pour une analyse et un traitement efficaces. De ce fait, nous avons développé une approche de fusion et amélioration des images radiographiques de bagages à double énergie, et pour montrer les performances du système proposé, une analyse comparative a été effectuée sur de nombreux modèles et d'algorithmes d'amélioration d'image bien connus. Le deuxième point de notre travail concerne l'étape de segmentation, dans laquelle nous avons choisi d'utiliser la segmentation basée sur le clustering par l'apprentissage automatique et la segmentation sémantique profonde s'appuyant pour cette dernière, sur un modèle développé en ce sens de ce travail, basé sur l'Adversarial AutoEncoder. Une fois l'image segmentée obtenue, une détection d'objets est effectuée afin d'identifier tous les objets présentés dans les bagages suivie d'une extraction et d'une sélection des caractéristiques afin de faciliter la classification.

Chapitre 4

Résultats et discussions

4.1 Introduction

La reconnaissance automatique (ou semi-automatique) des éléments d'intérêts dans un scanner de bagage aura une incidence directe sur les performances des opérateurs humains dans l'environnement opérationnel. Dans ce chapitre, nous montrerons et discuterons les résultats des méthodes et des modèles proposés dans le chapitre précédent, concernant, l'amélioration de la qualité de l'image, le développement d'un système permettant de séparer l'image radiographique en zones pouvant se chevaucher, la création de détecteurs d'objets capables de distinguer tous les objets présents, la réduction du nombre de ressources nécessaires après la phase de détection d'objets, la classification des objets et finalement, la création d'un détecteur des objets à risques sans passer par la segmentation.

4.2 Base de données

Les images utilisées dans notre étude proviennent de la base de données privée de la Hi-Tech Detection Systems Society (HDTS) avec 1000 images de rayons X à double énergie, qui contiennent environ 9000 objets (explosifs, armes à feu et objets non menaçants).

4.3 Amélioration de la qualité d'image

Pour montrer la supériorité de notre système proposé, une analyse comparative a été réalisée incluant cinq méthodes : amélioration du contraste global par égalisation d'histogramme, amélioration du contraste local par égalisation adaptative de l'histogramme (AHE) et par égalisation d'histogramme adaptatif limité au contraste (CLAHE), technique de filtrage moyenne et technique de filtrage médiane.

Les images utilisées dans cette étude provenaient de la base de données privée HDTS (Hi-Tech Detection Systems Society) contenant 2000 images des bagages à rayons X numérisées avec une taille moyenne de 800 pixels par 800 pixels. Nous avons utilisé 100 images pour tester notre système proposé.

La Figure 38 montre différentes valeurs de PSNR avec les différentes méthodes testées. Le graphique montre clairement que notre méthode proposée est la meilleure technique d'amélioration d'image pour l'analyse d'images HDTS comparée aux autres cinq méthodes testées.

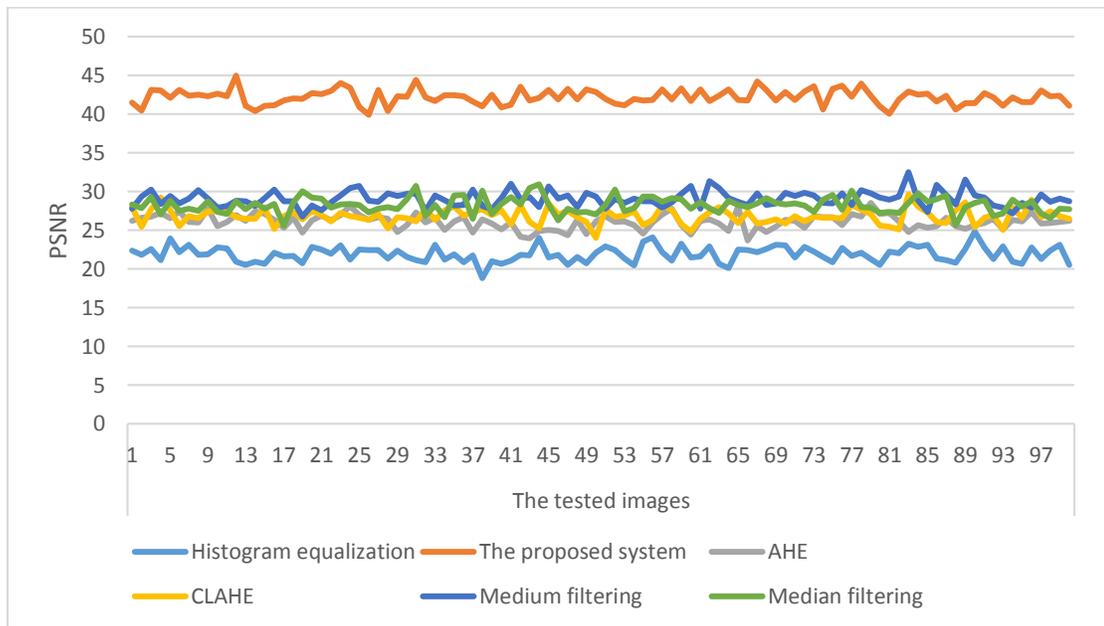


Figure 38: Variation des valeurs de PSNR pour 100 images applique pour les différentes méthodes testées.

4.4 Segmentation

4.4.1 Segmentation basée sur le clustering en utilisant l'apprentissage automatique

Les méthodes décrites dans la Section 3.3.1.1 ont été testées sur la base de données HDTS. Une optimisation des hyper-paramètres à l'aide de la validation croisée (CV) de toutes les méthodes est nécessaire pour assurer la meilleure segmentation. L'ensemble de données est divisé en trois ensembles (formation, validation et test) organisés en paires entrée cible (input-target). Les

vecteurs d'entrée sont composés des pixels d'image contenant les valeurs colorimétriques (Rouge (R), Vert (G) et Bleu (B)) et les vecteurs cibles indiquent la valeur correspondante du matériau (organique, inorganique, mixte, opaque et l'arrière-plan). Une fois les jeux de données sont disponibles, la formation et la recherche des valeurs optimales d'hyper-paramètres sont effectuées. Ces paramètres sont le nombre de neurones pour le MLP, le nombre d'arbres pour le RF, le numéro de voisinage K pour le KNN et ceux concernant les différents noyaux (KR) utilisés dans le réseau SVM. Les noyaux linéaire (LSVM), radial basique (RBF SVM) et polynomial (PSVM) (KR) ont été explorés :

$$KR(x, x') = \langle x, x' \rangle \quad (4.1)$$

$$KR(x, x') = \exp(-\|x - x'\|^2) \quad (4.2)$$

$$KR(x, x') = (1 + \langle x, x' \rangle)^\delta \quad (4.3)$$

Où x et x' sont des vecteurs dans l'espace d'entrée, δ est le paramètre à optimiser pour le SVM avec le noyau polynomial. Notez qu'il n'y a pas d'hyper-paramètres à optimiser pour les méthodes NB et LDA.

L'utilisation arbitraire d'hyper-paramètres dans différents algorithmes d'apprentissage machine conduit à des problèmes de sur-adaptation et de sous-adaptation, la principale cause de la piètre performance des modèles prédictifs. L'hyper-paramètre de toutes les méthodes a été obtenu par une stratégie de CV. Ces résultats sont présentés à la Figure 39.

Une fois que toutes les images sont segmentées, le meilleur pourcentage de segmentation est obtenu (pour chaque index d'évaluation), comme illustré dans la Figure 40. Le temps de calcul moyen est également montré dans la Figure 41.

Dans le Tableau 6, les matrices de confusion obtenues de toutes les méthodes testées sont affichées. La terminologie et les dérivées montrant le taux de faux positifs, le taux de faux négatifs, la précision et le score F1 [177] des meilleures et des pires méthodes sont présentées dans le Tableau 7. Il est important de mentionner que dans un but de simplification, les matrices de confusion sont obtenues une fois la tâche de fusion effectuée.

Tableau 6: Matrices de confusion. Ces matrices nous permettent de visualiser les performances des algorithmes utilisés après la tâche de fusion (deux classes).

	MLP			RF	
	Organique	Inorganique		Organique	Inorganique
Organique	94.2%	4.1%	Organique	86.3%	22.9%
Inorganique	5.8%	95.9%	Inorganique	13.7%	77.1%
	LDA			LSVM	
	Organique	Inorganique		Organique	Inorganique
Organique	91.1%	10.2%	Organique	89.4%	8.1%
Inorganique	8.9%	89.8%	Inorganique	10.6%	91.9%
	KNN			NB	
	Organique	Inorganique		Organique	Inorganique
Organique	90.5%	8.6%	Organique	97.5%	0.8%
Inorganique	9.5%	91.4%	Inorganique	2.5%	99.2%
	RBF SVM			PolySVM	
	Organique	Inorganique		Organique	Inorganique
Organique	90.1%	12.9%	Organique	91.5%	14.3%
Inorganique	9.9%	87.1%	Inorganique	8.5%	85.7%

Tableau 7: Le résultat principal des meilleures (Naive Bayes) et des pires méthodes (Random Forest).

	Naive Bayes	Random Forest
False positive rate	0.80 %	22.90 %
False negative rate	2.60 %	13.70 %
Précision	98.40 %	81.70 %
F1 score	98.30 %	82.50 %

Dans le Tableau 7, la terminologie et les dérivations des matrices de confusion montrent clairement que Naive Bayes présente la meilleure performance et RF le pire avec un taux de faux positifs égal à 0,8 %, un taux de faux négatifs de 2,6 % et avec une précision égale à 98,4

%. Sinon, RF a du mal à distinguer les pixels organiques des pixels inorganiques, avec un total de taux de faux positifs et de taux de faux négatifs de 22,9 % et 13,7 %, respectivement, et un score F1 de 82,5 %.

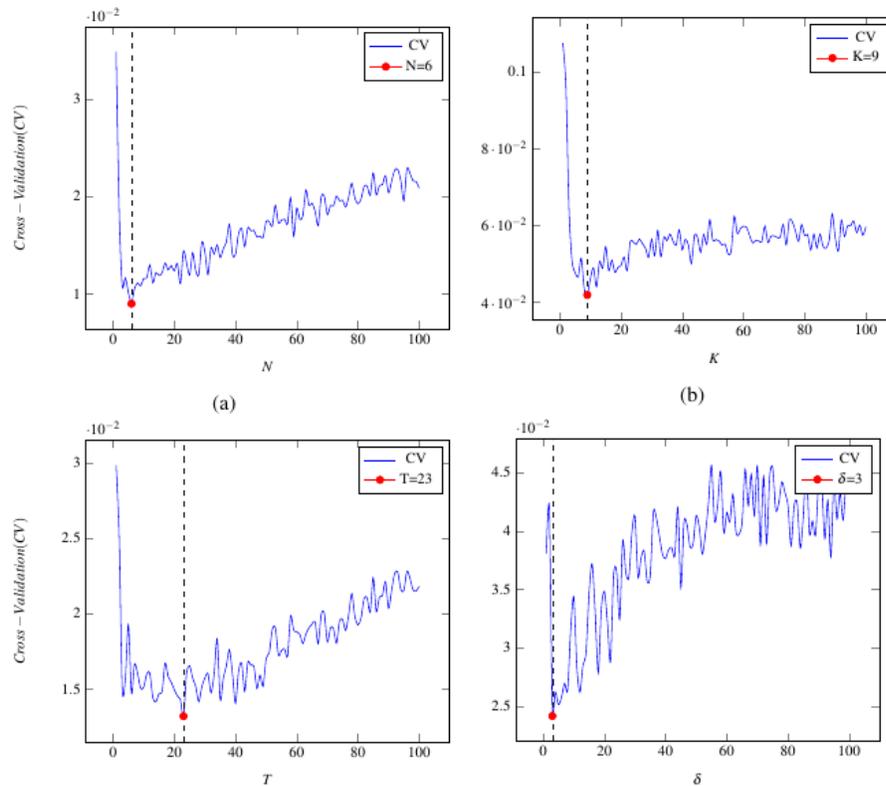


Figure 39: Processus d'optimisation des hyper-paramètres via la méthode CV. Les figures (a), (b), (c) et (d) indiquent les valeurs optimales (points rouges) des paramètres N , K , T et δ , respectivement, étant N le nombre de neurones du MLP, K le nombre de voisinages du KNN, T le nombre d'arbres pour le RF et δ le numéro d'ordre du SVM à noyau polynomial.

Figure 39 (a) montre que la sur-adaptation survient lorsque le nombre de nœuds du réseau MLP dépasse 6 et que la sous-adaptation se produit pour un nombre inférieur à cette valeur, ce qui implique que ce nombre est optimal. De même, les figures (b), (c) et (d) montrent que 9, 23 et 3 sont les valeurs optimales pour les hyper-paramètres de KNN, RF et SVM polynomial, respectivement.

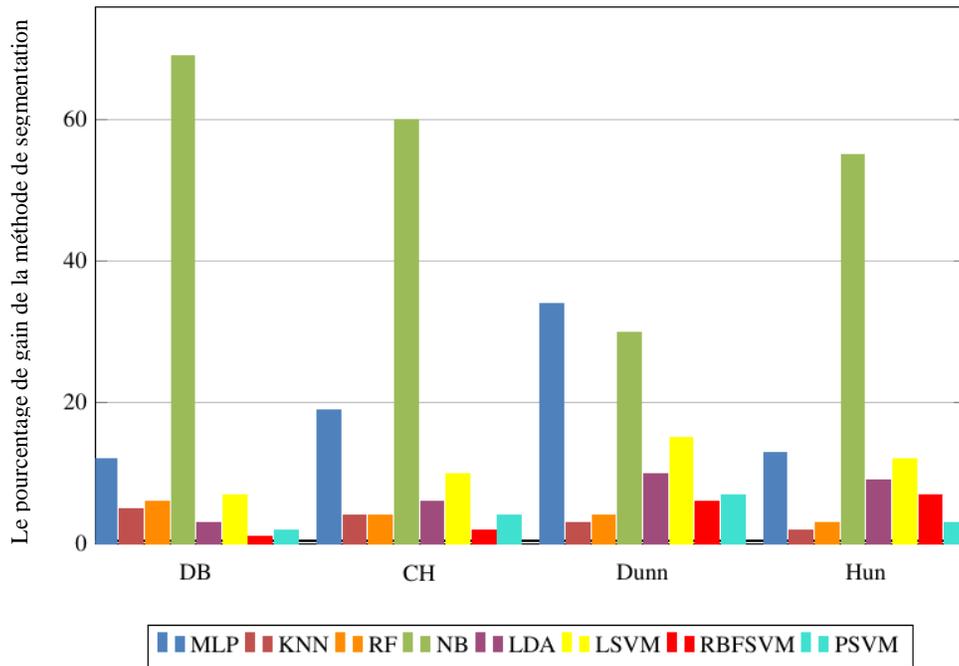


Figure 40: Pourcentage moyen de la meilleure segmentation obtenue par l'indice de validation de la classification : indice de Davies-Bouldin (DB), indice de Calinski-Harabasz (CH), indice de Dunn (DI) et indice de Hartigan (HI).

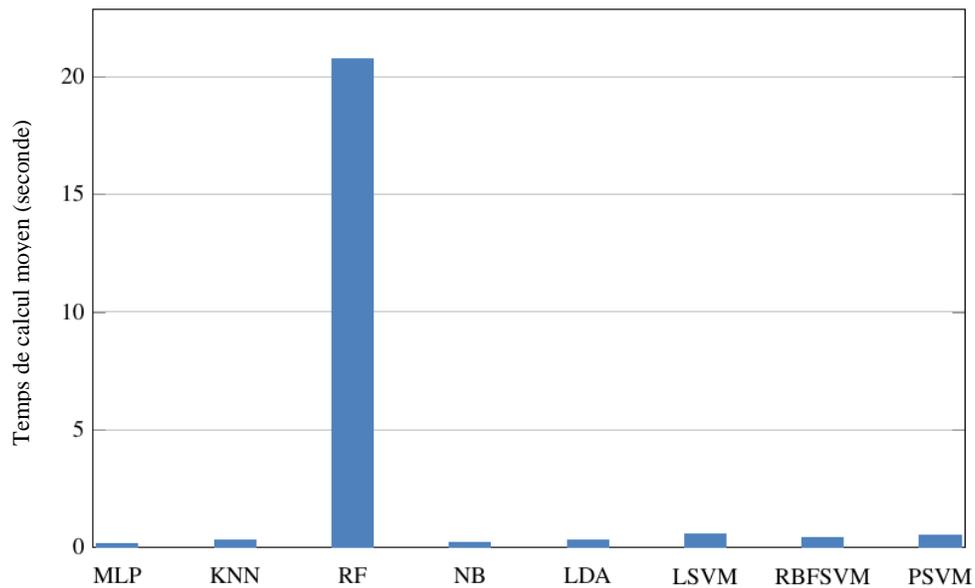


Figure 41: Temps de calcul moyen obtenu pour chaque méthode après avoir testé 1000 images différentes.

Dans la Figure 40, l'index de Davies-Bouldin indique que la méthode Naive Bayes est la meilleure, évaluée à 69,66 % avec un temps de calcul moyen de 0,22 seconde par image en segmentation (montré dans la Figure 41). En outre, les indices de Calinski-harabasz et Hartigan ont également confirmé que Naive Bayes est le meilleur algorithme, avec des taux de 60,23 % et 55,97 % respectivement. L'indice Dunn indique que MLP est la meilleure méthode suivie par Naive Bayes. RF est la pire des méthodes, à part le temps de calcul qui dépasse la valeur moyenne, qui est de 20 secondes pour chaque image, comme illustrée dans la Figure 41. Malgré son modèle de conception "naïf" et ses hypothèses simples, la conclusion est que Naive Bayes est la meilleure méthode de segmentation utilisée dans cette application. Ceci est corroboré par d'autres études, telles que [178], qui ont montré qu'il existait des raisons théoriques à l'efficacité inattendue des classificateurs Naive Bayes, notamment en raison de leur indépendance conditionnelle.

4.4.2 Segmentation sémantique profonde

La base de données HDTS introduite dans le chapitre précédent est partitionnée en un jeu d'apprentissage contenant 80% des données (1600 images) et un jeu d'essai contenant 20 % du jeu de données (400 images). Les résultats comparatifs sont calculés à partir de la classification finale au niveau des pixels : une étiquette de pixel est considérée correctement détectée si elle est identique à l'étiquette de vérité terrain (ground truth). Les performances sont mesurées en termes de précision, de limite F1 et d'intersection de pixels (Intersection Over-Union (IoU)) entre les six classes.

Pour la formation du SemSegAAE, nous avons utilisé le réseau VGG-16 qui a été pré-formé sur ImageNet. Nous avons adapté le réseau VGG-16 à notre jeu de données en minimisant la fonction de perte d'entropie croisée à l'aide d'un algorithme de descente de gradient stochastique. Nous utilisons un mini-lot de six images et un taux d'apprentissage initial égal à 0,001, qui est multiplié par 0,1 toutes les 800 itérations (un taux d'apprentissage de 0,01 pour la couche de classificateur final). Les paramètres de perte de mouvement (momentum) et de poids (weight decay) sont établis à 0,9 et à 0,005, respectivement.

Pour l'entraînement et les étapes de test, nous redimensionnons toutes les images d'entrée et les cartes sémantiques de vérité terrain en une résolution de 431×431×3. Pour la phase de test, l'image sera rétablie à sa taille d'origine après l'opération du décodeur.

La méthode proposée a montré des résultats très prometteurs. Le Tableau 8 répertorie la comparaison des performances, en termes de précision, de scores limites F1 et IoU, entre le système proposé et les méthodes testées. Les résultats moyens sur les six classes sont affichés dans le Tableau 9. Ces résultats démontrent que le SemSegAAE fournit de meilleurs résultats que les autres algorithmes testés montrant l'efficacité du système dans la segmentation sémantique. En particulier, la méthode obtient 80,17 %, 76,28 % et 76,84 % correspondant à la précision globale, à la limite moyenne F1 et à la moyenne d'IoU. D'autre part, Tableau 8 montre comment le SegNet donne de bons résultats, en particulier dans la segmentation des objets inorganiques et dans la segmentation des objets qui se chevauchent avec des objets inorganiques.

Tableau 8: Résultat de la segmentation montrant la performance comparative entre le système proposé et les autres méthodes testées.

	Organique	Inorganique	Chevauchement Organique / Organique	Chevauchement inorganique / inorganique	Chevauchement Organique / inorganique	Background
SemSegAAE						
Précision %	86.83	85.32	71.76	69.80	79.43	87.89
The boundary F1 %	79.42	82.76	64.08	69.62	72.11	89.67
IoU %	83.65	85.92	67.21	64.75	73.63	85.86
FCN-32s						
Précision %	53.15	57.23	31.79	43.71	47.46	58.92
The boundary F1 %	48.74	57.67	36.11	46.53	47.14	56.70
IoU %	59.97	62.83	42.24	37.66	48.66	60.89
FCN-16s						
Précision %	54.52	59.89	34.45	46.37	50.12	61.58

The boundary F1 %	50.11	60.33	38.77	49.19	49.80	59.36
IoU %	61.34	65.49	44.90	40.32	51.32	63.55
	FCN-8s					
Précision %	65.72	65.06	39.62	51.54	55.29	66.75
The boundary F1 %	61.31	65.50	43.94	54.36	54.97	64.53
IoU %	72.54	70.66	50.07	45.49	56.49	68.72
	SegNet					
Précision %	79.95	88.20	59.88	78.68	70.55	82.01
The boundary F1 %	73.54	90.64	61.20	71.50	70.23	79.79
IoU %	81.77	87.80	65.33	62.63	71.75	83.98
	U-Net					
Précision %	59.69	71.09	45.65	57.57	61.32	72.78
The boundary F1 %	55.28	71.53	49.97	60.39	61.00	70.56
IoU %	66.51	76.69	56.10	51.52	62.52	74.75
	Deconvet					
Précision %	51.86	58.52	33.08	45.00	48.75	60.21
The boundary F1 %	47.45	58.96	37.40	47.82	48.43	57.99
IoU %	58.68	64.12	43.53	38.95	49.95	62.18
	RedNet					
Précision %	55.87	61.24	35.80	47.72	51.47	62.93
The boundary F1 %	51.46	61.68	40.12	50.54	51.15	60.71
IoU %	62.69	66.84	46.25	41.67	52.67	64.90

Dans l'ensemble, les résultats de la segmentation sémantique révèlent la difficulté d'identifier la classe de chevauchement organique / organique. Parmi les réseaux de convolution, les FCN-

32 ont la capacité inférieure de distinguer cette classe. Les résultats confirment également l'importance des connexions sautées pour la segmentation sémantique, car tirer parti des connexions sautées dans les FCN-16 et les FCN-8 a considérablement contribué à l'amélioration des résultats de la segmentation par rapport aux FCN-32 en récupérant davantage de détails spatiaux. Cependant, beaucoup de détails ont été obtenus avec les FCN-8 avec une précision améliorée de 13,89 % (pour la classe organique / organique), en raison de l'utilisation de deux connexions sautées par rapport aux FCN-16. De plus, toutes les méthodes réussissent à discriminer la classe d'arrière-plan. Ceci est en partie attribué à la disponibilité d'un plus grand nombre d'échantillons d'apprentissage pour la classe d'arrière-plan que pour les autres. Théoriquement, il faudrait améliorer l'exactitude de toutes les classes en incluant un plus grand nombre d'échantillons d'apprentissage.

Tableau 9: Le résultat moyen de la précision, les limites F1 et IoU entre le système proposé et les autres méthodes testées.

	Précision %	The boundary F1 %	IoU %
SemSegAAE	80.17	76.28	76.84
FCN-32s	48.71	48.81	52.04
FCN-16s	51.16	51.26	54.49
FCN-8s	57.33	57.44	60.67
SEGNET	76.54	74.48	75.54
U-NET	61.35	61.46	64.69
Deconvnet	49.57	49.67	52.90
REDNET	52.50	52.61	55.84

La supériorité du réseau SemSegAAE par rapport aux méthodes de segmentation sémantique de pointe est attribuée à la manière entraînée défavorablement (adversely trained manner) du codeur qui oblige à produire une distribution spatiale latente spécifique, dans notre cas une distribution gaussienne symétrique produisant des caractéristiques non corrélées. Cela a enrichi les informations contextuelles, ce qui est très important pour distinguer différentes classes. Par exemple, la classe de chevauchement entre organique / organique est la plus difficile à

discriminer en raison de la variance élevée intra-classe et faible entre-classes; cependant, le système proposé a obtenu une précision de 71,76 %, car son étiquetage basé sur les pixels est clair, précis et comparable à l'image de vérité terrain.

Quelques résultats de segmentation d'images sont montrés à la Figure 42, des résultats comparatifs entre la segmentation obtenue par les méthodes considérées et la vérité de terrain désirée sont montrés à la Figure 43. Sur cette figure, les régions vertes et magenta mettent en évidence les zones où les résultats de la segmentation diffèrent de la vérité de terrain attendue (la couleur verte indique la segmentation faussement positive et la couleur magenta indique la segmentation vraie négative). Les deux figures montrent des résultats satisfaisants par rapport aux autres. Néanmoins, le résultat global des résultats de la FCN-8 est le plus affecté par les points de dispersion bruyants, en particulier autour des bords. Ceci peut être observé par la segmentation granulaire à la limite entre les objets organiques et l'arrière-plan. Cela est dû au fait que le réseau propage l'image à travers plusieurs couches de convolution et de pooling alternées. La résolution des cartes d'entités en sortie est ensuite échantillonnée. Par conséquent, les prédictions directes de FCN sont généralement en basse résolution, ce qui donne des limites relativement floues aux objets.

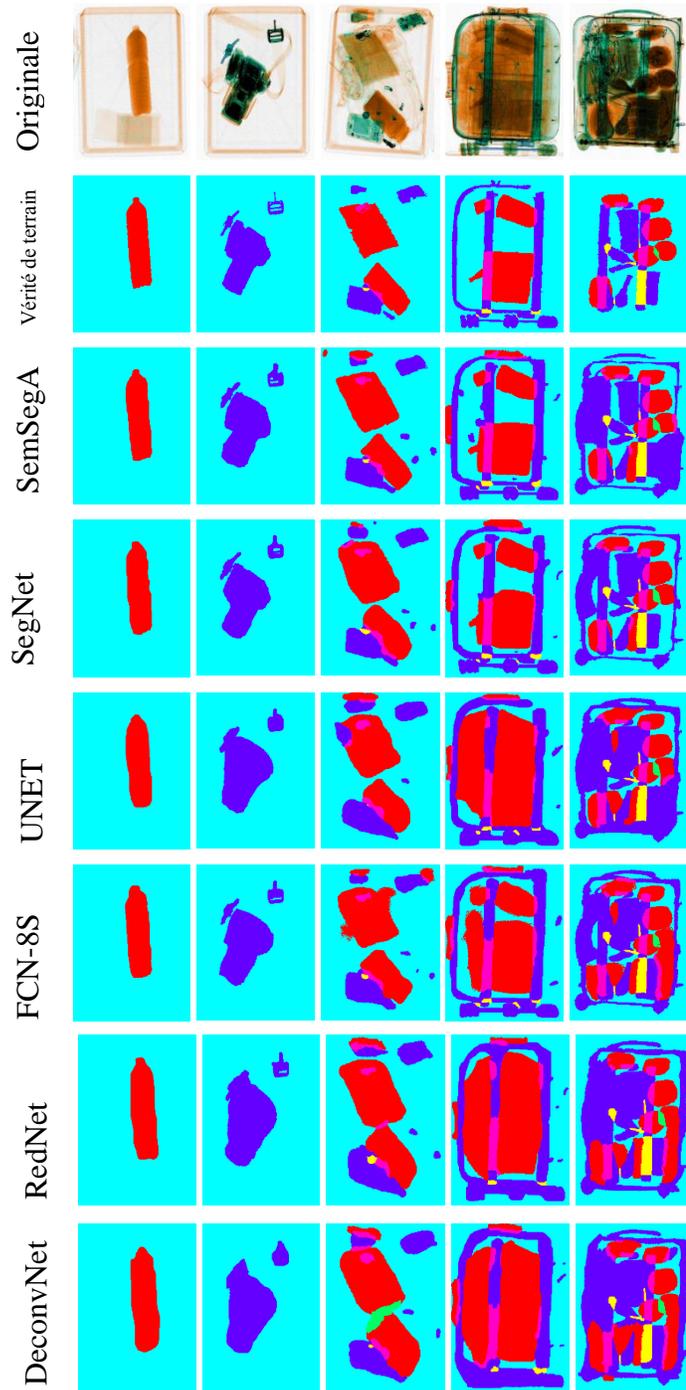


Figure 42: Quelques résultats expérimentaux de segmentation sur des images de test. Les colonnes représentent les images testées avec une complexité faible à élevée et les lignes représentent les différentes méthodes considérées. On peut observer comment la méthode proposée (troisième ligne) produit la segmentation la plus similaire à la solution souhaitée (deuxième ligne).

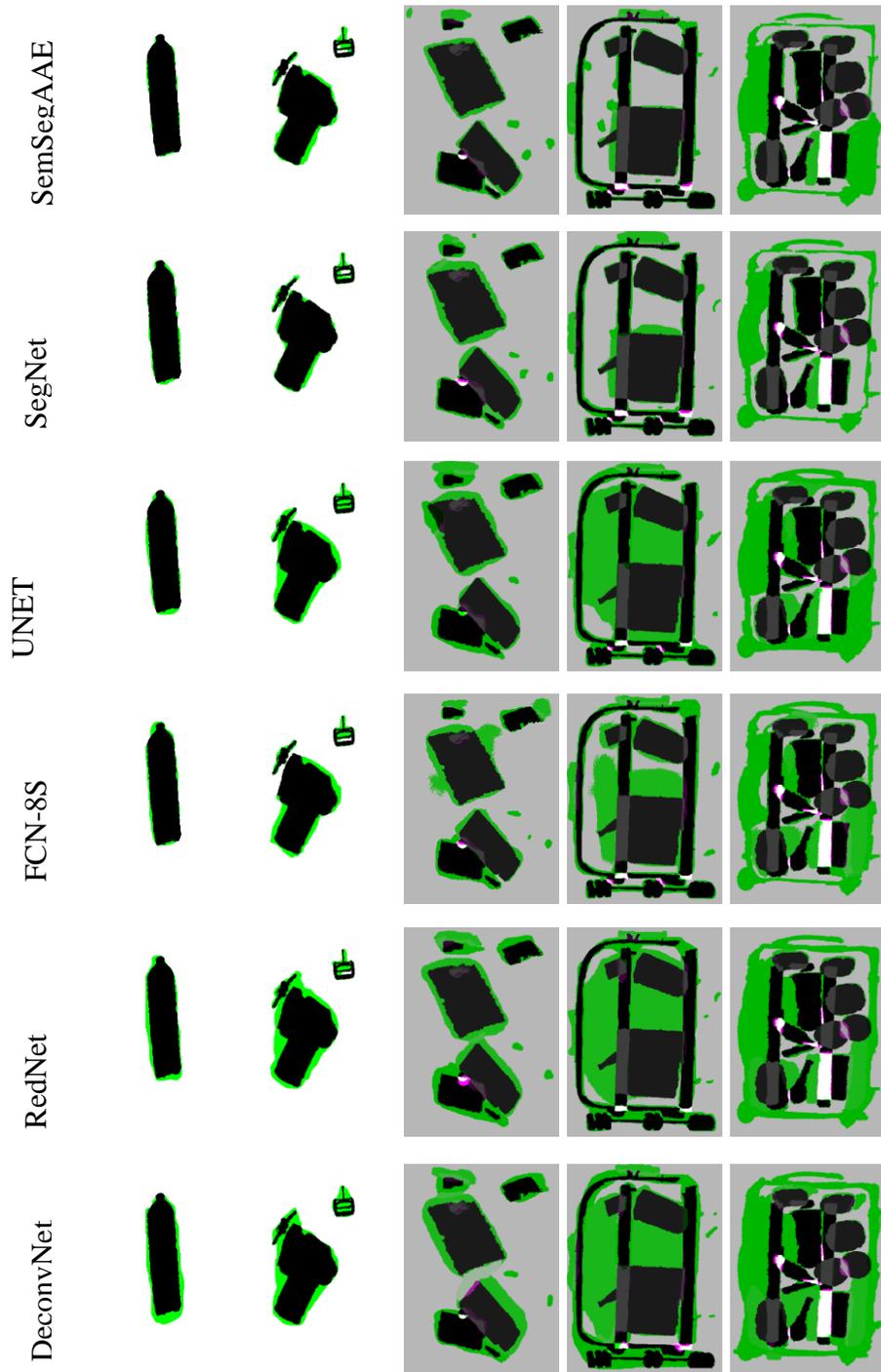


Figure 43: Résultats comparatifs entre la segmentation obtenue par les méthodes considérées et la vérité de terrain. Les régions vertes et magenta mettent en évidence les zones où les résultats de la segmentation diffèrent de la vérité au sol attendue.

4.5 Détection d'objets

4.5.1 Par les méthodes de machine learning

Après le processus de détection d'objet, composé d'une binarisation et d'un seuillage utilisant la méthode d'Otsu, une évaluation de nos méthodes a été réalisée au moyen d'un ensemble d'expériences. La Figure 44 illustre des exemples qualitatifs tirés de notre ensemble de données. Sur cette figure, les images de sortie apparaissent avec des rectangles entourant les objets détectés, dans lesquels les zones rouge, bleues, vertes et noires représentent respectivement des objets organiques, inorganiques, mixtes et opaques. Notez que les résultats obtenus sont très satisfaisants avec un très petit nombre de fausses détections, ce qui est très encourageant pour les travaux futurs (classification des objets).

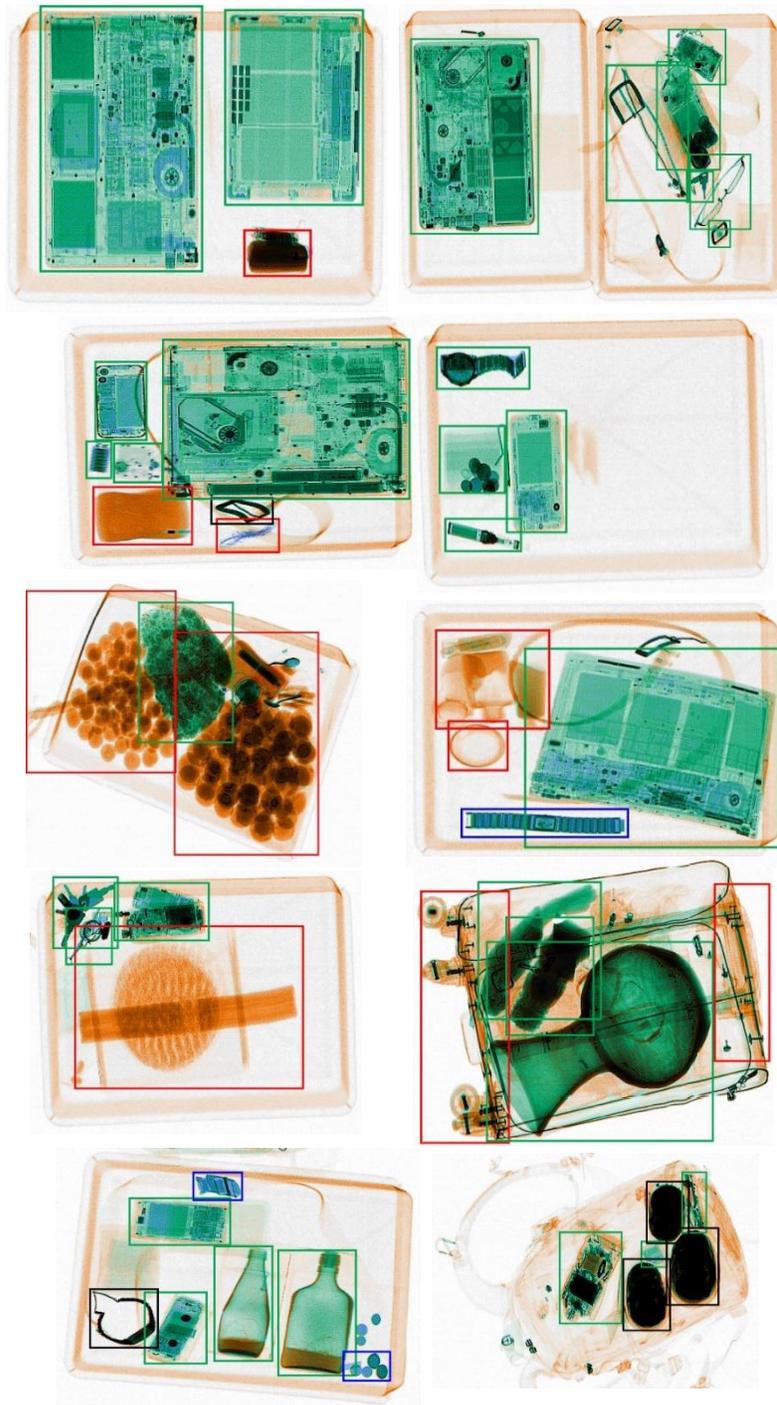


Figure 44 : La figure montre quelques résultats de détection d'objet extraits de notre ensemble de données. Les zones rouges, bleues, vertes et noires représentent respectivement les objets organiques, inorganiques, mixtes et opaques. Les résultats montrent la détection d'objet parfaite avec pratiquement aucune détection fausse.

4.5.2 Par la méthode YOLOv3 modifiée

Dans le Tableau 10, une comparaison de la méthode proposée (YOLOv3 modifiée) et de la méthode traditionnelle est présentée, avec un examen de l'effet de la variation de la taille de l'image en entrée sur les performances de la détection. L'objectif est de mesurer la précision des prévisions et l'efficacité de la recherche de tous les objets (mean Average Precision (mAP) et F1-Score), ainsi que le chevauchement entre l'objet détecté et la vérité de terrain (précision, rappel et intersection sur l'Union (IoU)), ainsi que le temps de calcul.

Tableau 10: La performance du YOLOv3 modifiée par rapport à l'original pour notre proposition

Taille de l'image d'entrée	Performances	YOLOv3 Original	YOLOv3 modifiée
416×416	mAP	72.26%	79.28%
	F1-score	78.45%	78.45%
	IoU	64.69%	67.69%
	Recall	81.92%	86.71%
	Precision	80.94%	80.94%
	time (ms)	334	350
512×512	mAP	75.78%	80.72%
	F1-score	81.98%	81.98%
	IoU	63.39%	64.39%
	Recall	79.34%	83.67%
	Precision	83.22%	83.22%
	time (ms)	389	403
608×608	mAP	76.36%	92.36%
	F1-score	81.12%	81.12%
	IoU	64.71%	81.71%
	Recall	82.19%	92.32%
	Precision	86.54%	86.54%
	time (ms)	408	425

Les résultats prouvent que le changement de méthode d'ancrage (de la k-moyennes à la SOM) améliore la coïncidence entre la vérité de terrain donnée et les boîtes englobantes détectées pour chaque détection avec une précision constante des prédictions, mais avec un temps de prédiction plus long.

Le résultat illustre également l'impact de la modification de la taille de l'image d'entrée sur les performances de la détection. La taille 608×608 donne de meilleurs résultats. On peut observer que les performances et la taille de l'image sont proportionnelles, car l'augmentation de la taille de l'image permet de conserver les détails de celle-ci.

4.5.3 Après la Segmentation sémantique profonde

Pour montrer les performances du système proposé pour distinguer les objets qui se chevauchent, nous utilisons des images segmentées pour détecter les objets dans le bagage. Pour cela, nous proposons une méthode de détection basée sur une technique de marquage des composants associés [166], [168], après avoir fusionné les six masques résultant de la phase de segmentation en deux, l'un pour l'objet organique et l'autre pour les objets inorganiques (selon les règles mentionnées dans la Section 3.4.1). Cette technique est appliquée pour obtenir des ensembles disjoints faciles à isoler. Le modèle renvoie les étiquettes des objets détectés, ainsi que leurs coordonnées. À l'aide de ces valeurs, l'application génère une nouvelle image remplie de rectangles entourant les objets détectés, dans laquelle les zones rouges et bleues représentent respectivement des objets organiques et inorganiques. Des images des d'objets séparés peuvent être générées pour les catégories d'objets organiques et inorganiques, permettant à l'opérateur humain de discriminer les objets sélectionnés afin de faciliter le diagnostic. De plus, les objets séparés constitueront l'entrée de la détection automatique d'objet de menace. On peut remarquer que les résultats obtenus sont très satisfaisants avec un petit nombre de fausses détections, comme le montre la Figure 45, ce qui est très encourageant pour les travaux futurs, par exemple la classification des objets. Il est très important de mentionner que notre étude est basée sur le chevauchement entre deux objets, que le chevauchement se produise entre le même matériau ou des matériaux différents. Par conséquent, il existe une limite dans le cas d'un chevauchement entre trois objets (ou plus).

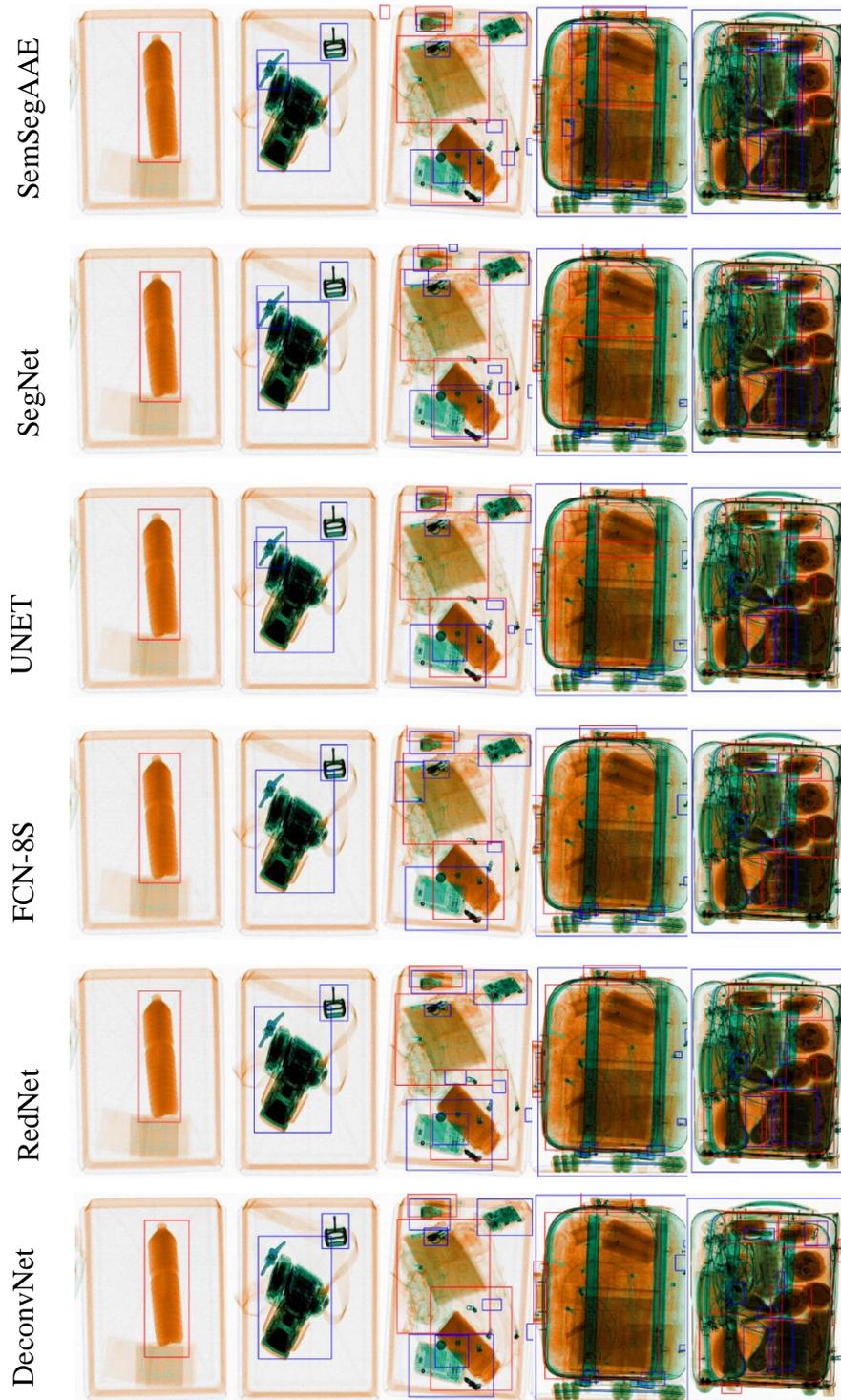


Figure 45: Exemples de détection d'objet. Les cases rouges et bleues représentent respectivement les objets organiques et inorganiques. On peut remarquer que les résultats obtenus avec le système proposé (première rangée) sont très satisfaisants avec un petit nombre de fausses détections.

4.6 Extraction de caractéristiques

Après la détection des objets, nous passons à l'extraction des caractéristiques. Les caractéristiques extraites sont les suivantes : caractéristiques de forme, de couleur et de texture. Chaque objet a une somme de 41 caractéristiques. Un problème majeur peut être posé lorsque nous nous retrouvons avec un grand nombre de fonctionnalités, est la malédiction de la dimensionnalité et de nombreux autres problèmes [179]. Pour cela, nous avons réduit le nombre de dimensions des données d'entrée afin d'obtenir la meilleure précision de classification.

4.6.1 Par notre modèle proposé

En entraînant l'AAE à utiliser nos objets (y compris 2809 objets explosifs, 2898 armes à feu et 3256 objets non menaçants) pour obtenir les caractéristiques latentes des distributions gaussiennes de différentes dimensions. Les AEs ont été formés par descente de gradient stochastique à l'aide de l'optimiseur Adam [180]. Tous les algorithmes d'apprentissage et d'extraction de code latent ont été réalisés à l'aide de TensorFlow [181] sur un GPU d'une carte graphique NVIDIA Geforce GTX 1050ti s'exécutant sur le système d'exploitation Windows.

La Figure 46 montre la représentation latente des objets d'entraînement, ainsi que les codes latents de certains objets de test. On peut constater que l'augmentation de la dimension des entités augmente la séparation du fait que les entités ont plus de flexibilité.

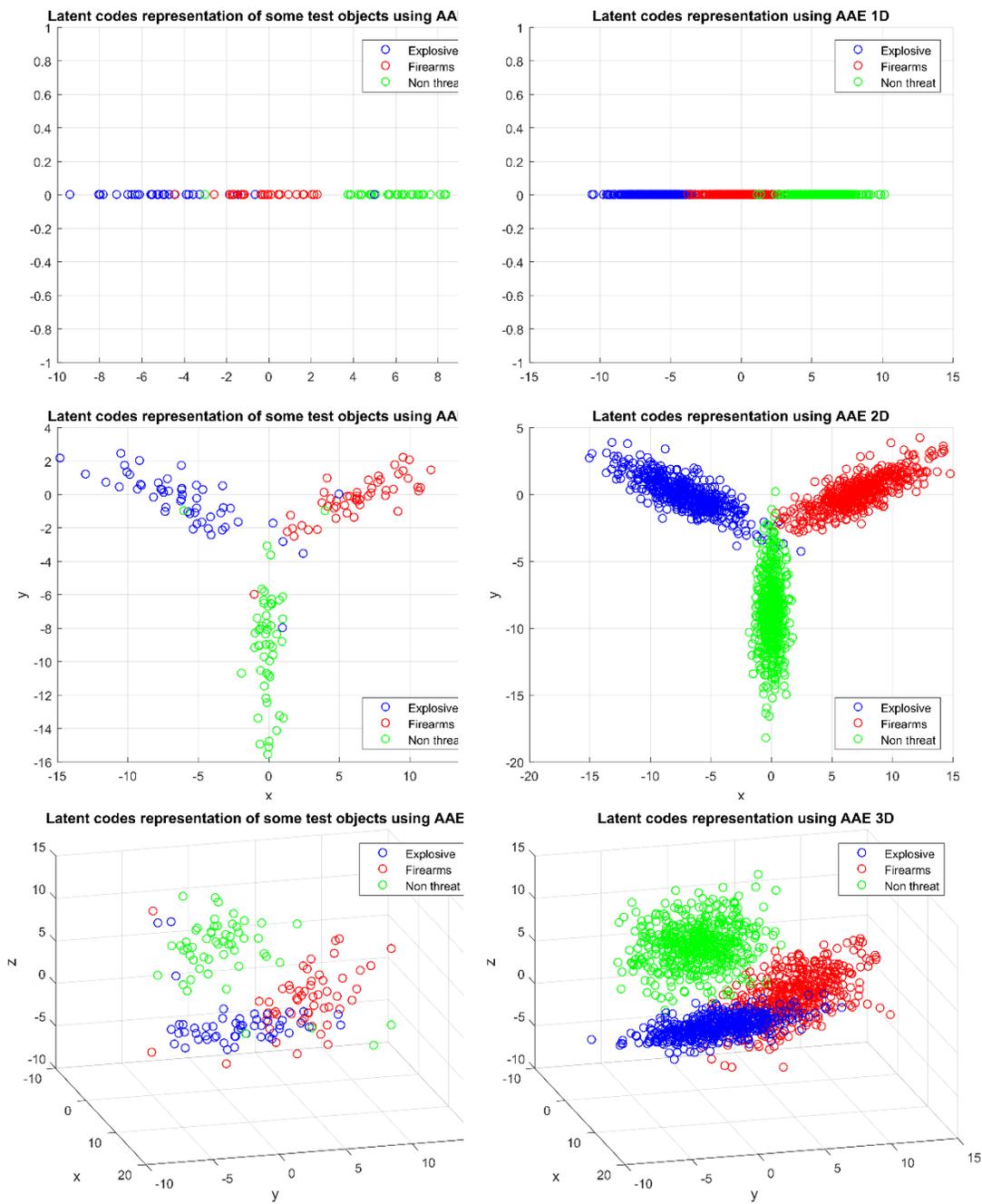


Figure 46: La représentation latente des objets d'apprentissage (première colonne) et la représentation des codes latents pour certains objets de test (deuxième colonne).

4.7 Sélection des caractéristiques et classification

Il est très important de mentionner que nous avons optimisé le nombre de caractéristiques après les avoir classées avec chaque algorithme de sélection de caractéristiques.

Le tableau suivant indique la précision et l'écart type pour chaque méthode (avec la méthode de sélection des meilleures caractéristiques).

Tableau 11: Résultat de précision / écart type moyen de classification avec la meilleure méthode de sélection des caractéristiques.

Inorganique				
	Menace / non menace		Arme blanche/ arme à feu	
MLP	98.99 ± 0.69 %	Relief	94.82 ± 2.04 %	Relief
KNN	92.52 ± 1.47 %	CFS	98.79 ± 2.07 %	Relief
RF	95.77 ± 1.04 %	Laplacian	95.14 ± 1.36 %	Relief
Linear SVM	98.79 ± 0.56 %	Relief	92.52 ± 1.53 %	ILFS
SVM RBF	96.39 ± 1.13 %	Relief	92.02 ± 1.63 %	Relief
Polynomial SVM	98.64 ± 0.62 %	Laplacian	93.39 ± 1.98 %	Relief
Organique				
	Menace / non menace		Arme blanche/ arme à feu	
MLP	98.76 ± 0.93 %	Laplacian	96.86 ± 2.01 %	Laplacian
KNN	96.14 ± 0.89 %	CFS	94.15 ± 1.61 %	Relief
RF	98.48 ± 0.56 %	Relief	94.89 ± 1.99 %	Laplacian
Linear SVM	98.99 ± 0.67 %	Laplacian	97.24 ± 1.19 %	CFS
SVM RBF	96.74 ± 1.12 %	Relief	94.53 ± 1.76 %	Relief
Polynomial SVM	98.79 ± 0.71 %	Laplacian	95.69 ± 1.55 %	Relief

Comme tendance générale, nous observons que la méthode de sélection des caractéristiques Relief est prédominante, bien que [182] ait constaté que la méthode ILFS est la meilleure

méthode appliquée à plusieurs bases de données. Après avoir utilisé le test z pour comparer les différents résultats obtenus et pour la classification menace / non-menace, nous avons constaté que le MLP est le meilleur pour les objets inorganiques avec une précision de $98,99 \pm 0,69\%$. Le SVM linéaire pour les objets organiques avec une précision de $98,99 \pm 0,67\%$, si nous passons à la classification binaire arme blanche/ armes à feu, le KNN fournit la meilleure performance $98.79 \pm 2.07 \%$ pour l'objet inorganique et la SVM linéaire présente une précision supérieure avec $97,24 \pm 1,19\%$ pour l'objet organique.

4.7.1 Classification par notre modèle proposé

Pour prouver les performances de classification en utilisant les fonctionnalités extraites des AE, nous avons utilisé 1531 objets à tester, dont 512 explosifs, 498 armes à feu et 521 objets non menaçants. Nous avons comparé notre stratégie de proposition à différentes méthodes d'apprentissage en profondeur pour notre tâche de détection / classification d'objets. Ces méthodes sont :

- Faster Region-based Convolutional Neural Network (FRCNN) [183] avec different architectures: Inception V2, ResNet-50 and ResNet-101.
- Single Shot multibox Detector (SSD) [184] avec different architectures: Inception V2, Mobilenet V1 and Mobilenet V2.
- Region-based Fully Convolutional Networks (R-FCN) [185] avec ResNet-101 architecture.
- YOLOv3 Real-Time Object Detection [171] avec Darknet-53 architecture.

Afin de réduire le temps de formation et d'améliorer les performances, l'approche par transfert d'apprentissage et le réseau pré-formé sur le jeu de données, ImageNet a été considéré [186]. Nous avons utilisé 800 images (contenant environ 7000 objets) pour la formation et 200 (environ 2 000 objets) pour les tests.

Les matrices de la Figure 47 montrent clairement que tous les systèmes testés ont du mal à faire la distinction entre les objets explosifs et les objets non menaçants en comparant l'explosif aux objets de l'arme à feu. Cela s'explique par le fait que les objets explosifs que nous avons utilisés dans cette étude sont de type plastique, donc de matière organique, identique à de nombreuses matières organiques courantes du type CHON (carbone, hydrogène, oxygène et azote), tels que le sucre, le chocolat et le miel [187].

Le Tableau 12 indique les résultats de la classification entre les armes à feu, les objets explosifs et les objets non menaçants pour les systèmes proposés, ainsi que pour FRCNN, SSD, R-FCN et YOLOv3, avec différentes architectures de réseau. La classification est évaluée par la précision (ACC), le score F1 et le coefficient de corrélation de Matthews (MCC). D'une manière générale, nous observons que les performances augmentent avec la complexité globale du réseau, telle que les performances supérieures obtenues avec FRCNN avec l'architecture ResNet-101. Dans l'ensemble, le système proposé à 5 dimensions AAE (AAE 5D) offre les meilleures performances avec une précision de 97,92% et 96,42% respectivement dans la détection des armes à feu et des objets explosifs. Cela est dû à la capacité de l'AAE à générer de nouveaux objets menacés, ce qui offre l'avantage de les détecter facilement. L'extraction des caractéristiques AAE à une dimension est peu performante par rapport à d'autres approches de classification. Cette faible performance est principalement due au chevauchement des caractéristiques extraites.

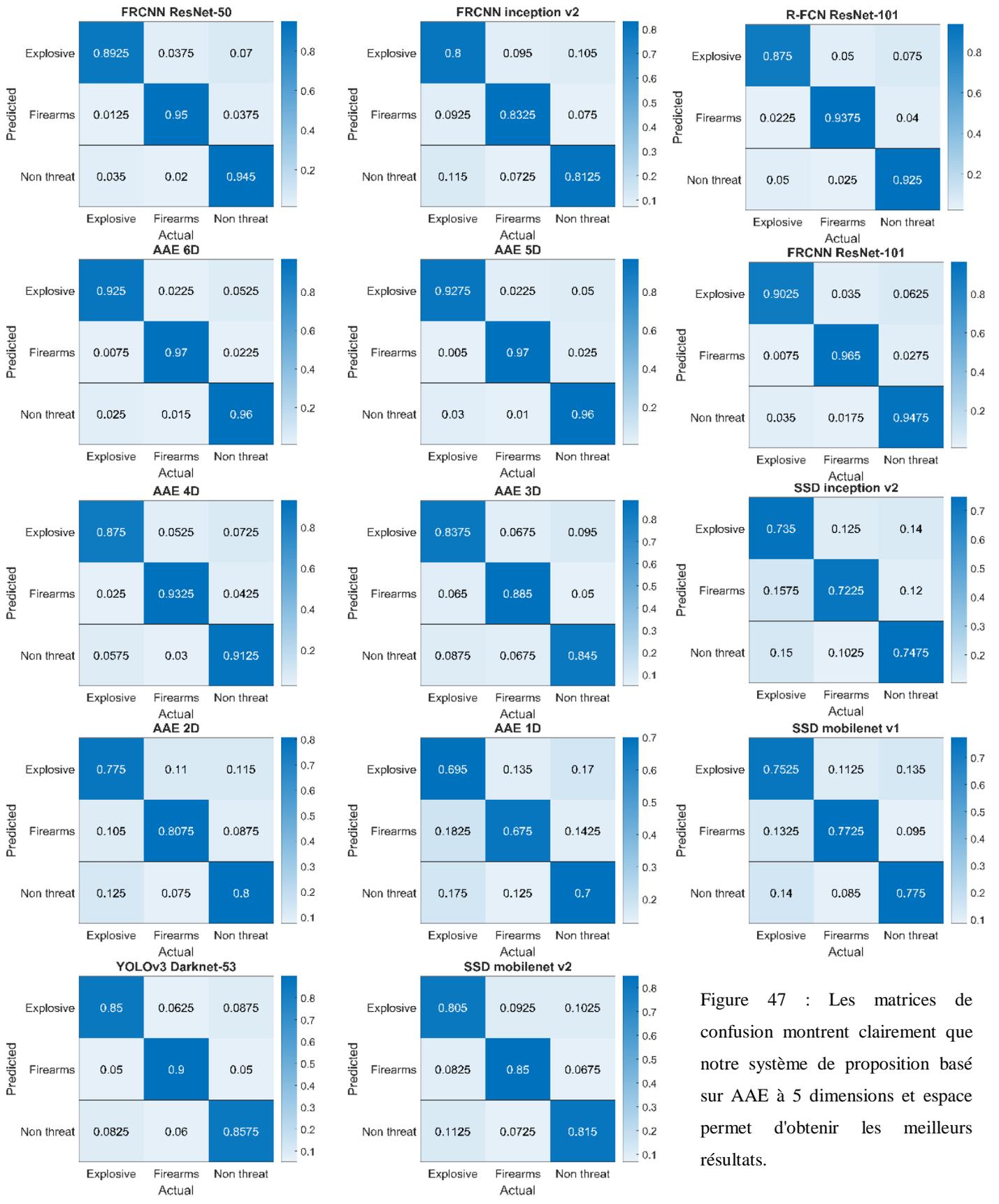


Figure 47 : Les matrices de confusion montrent clairement que notre système de proposition basé sur AAE à 5 dimensions et espace permet d'obtenir les meilleurs résultats.

Tableau 12: Performances de classification pour détecter les armes à feu, les objets explosifs et les objets non menaçants.

	Objets non menaçants			Objets explosifs			Armes à feu		
	ACC	F1-Score	MCC	ACC	F1-Score	MCC	ACC	F1-Score	MCC
AAE 1D	79.58%	84.64%	54.21%	77.92%	67.72%	50.99%	80.50%	69.77%	55.47%
AAE 2D	86.58%	89.93%	69.83%	84.83%	77.31%	65.92%	87.42%	81.05%	71.64%
AAE 3D	90.00%	92.52%	77.45%	89.50%	84.17%	76.32%	91.67%	87.62%	81.35%
AAE 4D	93.25%	94.90%	84.93%	93.08%	89.40%	84.31%	95.00%	92.56%	88.80%
AAE 5D	96.17%	97.10%	91.48%	96.42%	94.52%	91.90%	97.92%	96.88%	95.32%
AAE 6D	96.17%	97.10%	91.48%	96.42%	94.51%	91.90%	97.75%	96.64%	94.95%
FRCNN inception v2	87.75%	90.83%	72.39%	86.42%	79.70%	69.50%	88.83%	83.25%	74.88%
FRCNN ResNet-50	94.58%	95.88%	88.04%	94.83%	92.01%	88.29%	96.42%	94.65%	91.95%
FRCNN ResNet-101	95.25%	96.40%	89.45%	95.33%	92.80%	89.43%	97.08%	95.66%	93.47%
SSD inception v2	82.92%	87.16%	61.64%	80.92%	71.97%	57.54%	83.17%	74.10%	61.69%
SSD mobilenet v1	84.83%	88.61%	65.92%	82.67%	74.32%	61.25%	85.83%	78.43%	67.90%
SSD mobilenet v2	88.17%	91.16%	73.28%	87.00%	80.50%	70.75%	89.50%	84.37%	76.47%
R-FCN ResNet-101	93.67%	95.20%	85.93%	93.42%	89.86%	85.06%	95.42%	93.17%	89.72%
YOLOv3 Darknet-53	90.67%	93.01%	78.97%	90.58%	85.75%	78.73%	92.58%	89.00%	83.42%

4.8 Conclusion

Les résultats de la phase d'amélioration d'image a montré que la méthode proposée est la technique d'amélioration d'image la plus performante pour l'analyse des images de la base de données HDTS comparé aux autres méthodes testées. Une autre contribution au domaine de la reconnaissance d'objets dans les tests aux rayons X pour le contrôle des bagages est présentée dans ce travail, par la proposition d'une nouvelle architecture d'apprentissage profond pour la segmentation sémantique et la détection d'objets. L'architecture proposée combine un réseau profond de convolution avec un AutoEncoder Adversarial (AAE) qui agit comme un outil performant d'extraction de caractéristiques. D'un autre côté, pour la phase de détection, l'expérimentation démontre que la modification apportée à la méthode YOLOv3 originale a amélioré les performances de la détection.

Conclusions générale

Dans cette thèse, nous avons développé la reconnaissance automatique des éléments d'intérêt dans un scanner de bagages, qui dans l'environnement opérationnel auront un impact direct sur les performances des opérateurs humains soumis à des erreurs d'interprétation ou d'inattention.

Le but ainsi, de cette recherche est de proposer un système d'aide à la décision par l'application des méthodes de vision artificielle et la reconnaissance d'objets aux images radiographiques des bagages dans le domaine de l'inspection des bagages en comparant notre travail avec différentes stratégies de vision par ordinateur proposées ces dernières années.

Une liste des tâches que nous avons exécutées est présentée ci-dessous :

- Amélioration de qualité des images : Une nouvelle approche est développée qui constitue à combiner des images à rayons X hautes énergie et basse énergie, en supprimant le bruit et en améliorant la qualité de l'image fusionnée.
- Mise au point d'un système permettant de séparer l'image radiographique en zones et notamment les objets présentant des chevauchements : particulièrement en séparant l'image aux rayons X en images organiques et inorganiques, cela en tenant compte du chevauchement entre les objets du même matériau ainsi que ceux des différents matériaux. Notre contribution dans cette étape est le développement d'une méthode de segmentation sémantique basée sur Adversarial AutoEncoder (AAE). Nous avons montré l'efficacité de cette méthode en termes d'exploitation, de temps de calcul et de compétitivité par rapport aux méthodes présentés dans l'état de l'art de la segmentation sémantique.
- Proposition d'un détecteur d'objet capable de distinguer tous les objets présents dans les bagages (organiques et inorganiques).
- Réduction du nombre de ressources, nécessaires pour décrire un grand ensemble de données à l'aide des techniques d'extraction des caractéristiques, suivies des techniques de sélection des caractéristiques

- Classification des objets, en imitant le processus de la pensée humaine, qui permet de formuler des jugements compliqués avec précision, rapidité et cohérence.
- Création d'un détecteur d'objet de menace sans passer par la segmentation à l'aide des approches d'apprentissage en profondeur. Nous avons proposé une détection d'objet utilisant une version modifiée de la méthode YOLOv3.

En premier lieu, pour l'amélioration de la qualité d'image, les résultats montrent que notre méthode proposée est la meilleure technique pour l'analyse des images HTDS comparée aux autres cinq méthodes testées. En second lieu, et pour l'étape qui suit, particulièrement pour la segmentation par les méthodes de machine learning, le Naive Bayes présente la meilleure performance avec un taux de faux positif égal à 0,8 %, un taux de faux négatif de 2,6 % et avec une précision égale à 98,4 %. En outre, notre modèle proposé pour la segmentation sémantique a fourni des résultats meilleurs que les autres algorithmes testés montrant l'efficacité du système. En particulier, le système a atteint 80,17 %, 76,28 % et 76,84 % correspondant à la précision globale, à la limite moyenne F1 et à la moyenne d'IoU. D'autre part, pour la détection d'objet par machine learning, nous avons noté que les résultats obtenus sont très satisfaisants avec un très petit nombre de fausse détection. Nous avons proposé aussi une modification de la fameuse méthode YOLOv3, les résultats prouvent que le changement de méthode d'ancrage (de la k-moyennes à la SOM) améliore la coïncidence entre la vérité de terrain donnée et les boîtes englobantes détectées pour chaque détection avec une précision constante des prédictions, mais avec un temps de prédiction plus long. En dernier lieu, et pour l'étape de classification, le système proposé offre les meilleures performances avec une précision de 97,92% et 96,42% respectivement dans la détection des armes à feu et des objets explosifs.

Bien que les résultats obtenus dans cette thèse soient plus que satisfaisantes et prometteuses, néanmoins, il existe certaines limites à cette recherche. Premièrement, il est difficile d'obtenir des informations complètes sur les systèmes commerciaux existants, voire sur les données obtenues à partir de ces systèmes, pour des raisons de propriété (que les systèmes soient certifiés ou non). Aussi, les données nécessaires à nos recherches ont été collectées de HTDS; une société spécialisée dans la distribution et la maintenance des systèmes de détection haute

technologie en France et à l'international, la limitation est que l'ensemble des données collectées était limité au petit ensemble de bagages.

En perspective, les futurs travaux seront consacrés à :

- La détection des objets menaçants au moyen d'images à rayons X à double et à vues multiples, ainsi que de la détection d'objets explosifs au moyen de l'apprentissage en profondeur.
- Exploiter des images à rayons X à double vue (2 vues sous différents angles, généralement du haut et du côté) afin de détecter les objets menaçants et les batteries lithium-ion dans les bagages a soute dans les aéroports.

Production scientifique

Articles

1. Chouai, M., Merah, M., Sancho-Gómez, J. *et al.* Supervised feature learning by adversarial autoencoder approach for object classification in dual X-ray image of luggage. *J Intell Manuf* 31, 1101–1112 (2020). <https://doi.org/10.1007/s10845-019-01498-5>
2. Chouai, M., Mustefa, M. & Mimi, M. CH-Net: Deep adversarial autoencoders for semantic segmentation in X-ray images of cabin baggage screening at airports. *J Transp Secur* (2020). <https://doi.org/10.1007/s12198-020-00211-5>

Conférences internationales

1. Chouai, M., Merah, M., Sancho-Gómez, J. *et al.* (2019, november). Dual-energy x-ray images enhancement based on a discrete wavelet transform fusion technique for luggage inspection at airport. In 2019 6th international conference on image and signal processing and their applications (ISPA) (pp. 1-6). IEEE.
2. Chouai, M., Merah, M., Sancho-Gómez, J. L., & Mimi, M. (2020, March). A machine learning color-based segmentation for object detection within dual X-ray baggage images. In Proceedings of the 3rd International Conference on Networking, Information Systems & Security (pp. 1-11).
3. Chouai, M., Merah, M., Mimi, M (2019, December).color-based segmentation using Kohonen Self-organisation map (SOM) for x-ray airline images. In 20175th international conference on image and signal processing and their applications (ISPA).

Bibliographie

- [1] G. Team, "Global Terrorism Database," 15 November 2017. [Online]. Available: <https://www.start.umd.edu/gtd>. [Accessed 15 November 2017].
- [2] W. Tom and B. Jonathan, "Islamic State behind Australians' foiled Etihad meat-mincer bomb plot—Police," 4 August 2017. [Online]. Available: <https://www.reuters.com/article/us-australia-security-raids/islamic-state-behind-australians-foiled-etihad-meat-mincer-bomb-plot-police-idUSKBN1AJ367>.
- [3] D. Gillen and W. G. Morrison, "Aviation security: Costing, pricing, finance and performance," *Journal of Air Transport Management*, vol. 48, pp. 1-12, 2015.
- [4] Y. Sterchi and A. Schwaninger, "A first simulation on optimizing EDS for cabin baggage screening regarding throughput," in *International Carnahan Conference on Security Technology (ICCST)*, Taipei, Taiwan, 2015.
- [5] P. Neffenger, "Advanced Integrated Passenger and Baggage Screening," United States. Department of Homeland Security, USA, 2015.
- [6] G. Pochet, "Smart Security: Alternative Detection Methods and Unpredictability," 2016.
- [7] L. Y. B. Yann and H. Geoffrey, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436–444, 2015.
- [8] Y. Bengio, A. Courville and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798-1828, 2013.
- [9] A. K. Novakoff, "FAA bulk technology overview for explosives detection," *International Society for Optics and Photonics*, vol. 1824, pp. 2-13, April 1993.
- [10] O. o. C. A. Security, "Criminal acts against civil aviation," *Federal Aviation Administration*, 1997.
- [11] *Aviation security: technology's role in addressing vulnerabilities*, 1996.

- [12] S. M. Khan, P. A. Polski, C. E. Robbins, B. B. Boubli, R. H. Doney, R. Redman, P. Morvan, T. Gozani, J. Bartko and D. Syme, "Proceedings of the First International Symposium on Explosive Detection Technology," Federal aviation administration technical center, Atlantic city, 1992.
- [13] N. R. C. a. other, Assessment of Technologies Deployed to Improve Aviation Security: First Report, Washington, USA: National Academies Press, 1999.
- [14] C. E. Robbins, "Keynote address," in *Proceedings of the First International Symposium on Explosives Detection Technology*, Atlantica , November 1991.
- [15] N. P. Chen, "Simulation of x-ray tungsten source characteristic spectra," Spatial Data Analysis Lab, Virginia Tech, 1996.
- [16] G. L. Clark, Applied X-rays, McGraw-Hill, 1927.
- [17] S. E. Liverhant, Outline of atomic physics, New York: Regents Publishing Co., 1966.
- [18] E. F. Kaelble, Handbook of X-rays: for diffraction, emission, absorption, and microscopy, New York: McGraw-Hill, 1967.
- [19] N. A. Dyson, X-rays in Atomic and Nuclear Physics, Cambridge University Press, 2005.
- [20] H. Haken and H. C. Wolf, The physics of atoms and quanta: introduction to experiments and theory, Springer Science & Business Media, 2005.
- [21] J. G. Brown, X-rays and Their Applications, London: Springer Science & Business Media, 1966.
- [22] A. G. .: B. C. J. Michette, X-ray Science and Technology, CRC Press, 1993.
- [23] R. Van Grieken and A. Markowicz, Handbook of X-ray Spectrometry, CRC press, 1993.
- [24] W. H. McMaster, N. Kerr Del Grande, J. H. Mallett and H. J. H., "Compilation of X-Ray Cross Sections," Springfield, 1992.
- [25] M. Médicos, "The Function of X-Ray Tubes," Online marketplace for trading medical equipment, 14 August 2011. [Online]. Available: <http://used-medicequipmentblog.blogspot.com/2011/08/function-of-x-ray-tubes.html>.
- [26] P. P. Dendy and B. Heaton, Physics for diagnostic radiology, CRC press, 2011.

- [27] BBC, "7/7 inquests: Coroner warns over bomb ingredient," BBC News website, 1 February 2011. [Online]. Available: <https://www.bbc.co.uk/news/uk-12337575>.
- [28] R. P. Technologies, "Simulant explosives," Renful Premier Technologies, [Online]. Available: <https://simulants.co.uk/explosive-simulants/>.
- [29] J. Vanier and C. Audoin, *The quantum physics of atomic frequency standards*, CRC Press, 1989.
- [30] R. F. Eilbert and K. D. Krug, "Aspects of image reconstruction in Vivid Technology's dual-energy x-ray system for explosives detection," *Applications of Signal and Image Processing in Explosives Detection Systems*, vol. 1824, pp. 127-144, 1993.
- [31] K. O. Fultz, "Aviation security: technology's role in addressing vulnerabilities," 1996.
- [32] R. Krauss and W. Hughes, "Signatures of explosives by elemental composition analysis," in *Proceedings of the Second Explosives Detection Technology Symposium & Aviation Security Technology Conference*, Atlantic City, NJ, 1996.
- [33] G. Joe and T. Gerd-Rüdiger, "X-ray imaging methods over the last 25 years—new advances and capabilities," in *AIP Conference Proceedings*, Ames, Iowa, USA, 2001.
- [34] H. Martz and C. Crawford, "Explosive detection in aviation applications using CT," Lawrence Livermore National Lab. (LLNL), Livermore, CA, USA, 2011.
- [35] G. Harding, "X-ray scatter tomography for explosives detection," *Radiation Physics and Chemistry*, vol. 71, no. 3-4, pp. 869-881, 2004.
- [36] K. KRUG and J. STEIN, "Advanced dual energy x-ray for explosives," in *FAA, Proceedings of the First International Symposium on Explosive Detection Technology*, 1992.
- [37] L. GRODZINS, "Photons in- photons out: Non-destructive inspection of containers using x ray and gamma ray techniques," in *FAA, Proceedings of the First International Symposium on Explosive Detection Technology*, 1992.
- [38] D. SCHAFER, M. Annis and M. Hacker, "New x-ray technology for the detection of explosives," in *FAA, Proceedings of the First International Symposium on Explosive Detection Technology*, 1992.

- [39] M. Annis, P. Bjorkholm and D. Schafer, "Automatic detection of explosives using x-ray imaging explosives using x-ray imaging," in *Access Security Screening: Challenges and Solutions*, West Conshohocken, 1992.
- [40] F. Roder, "The evolution of Computed Tomography(CT) as an explosives detection modality," in *FAA, Proceedings of the First International Symposium on Explosive Detection Technology*, 1992.
- [41] F. L. Roder and B. G. Stebler, "Image processing requirements for computed-tomographic explosives detection," in *Applications of Signal and Image Processing in Explosives Detection Systems*, 1993.
- [42] R. Smith and P. Krall, "Full volume dual-energy high speed computed tomography (CT) explosives detection system," in *Proceedings of the Second Explosives Detection Symposium and Aviation Security Conference*, 1996.
- [43] R. Ryon, K. Dolan, D. Schneberk, H. Martz and R. Rikard, "Explosives detection limitations using dual-energy radiography and computed tomography Ryon, RW and Dolan, KW and Schneberk, DJ and Martz, HE and Rikard, RD," Lawrence Livermore National Lab., CA (United States), 1991.
- [44] G. Harding and J. Kosanetzky, "Scattered X-ray beam nondestructive testing," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 280, no. 2-3, pp. 517-528, 1989.
- [45] esky, "Bagages: limites de taille et de poids," 10 juin 2019. [Online]. Available: <https://www.esky.fr/guide-de-voyage/volsbillets-davion/bagages/bagages-limites-de-taille-et-de-poids>.
- [46] R. H. T. Ltd, "Hold Baggage Screening Systems," [Online]. Available: <https://www.copybook.com/companies/robson-handling-technology-ltd/articles/hold-baggage-screening-systems>.
- [47] T. S. Administration, "Security Screening: What Can I Bring?," [Online]. Available: <https://www.tsa.gov/travel/security-screening/whatcanibring/all>.
- [48] P. européen, "Règlement (CE) n° 300/2008," Conseil de l'Union européenne, 2008.
- [49] P. Européen, "Règlement (UE) n° 185/2010," Conseil de l'Union européenne, 2010.

- [50] S. Singh and M. Singh, "Explosives detection systems (EDS) for aviation security," *Signal processing*, vol. 83, no. 1, pp. 31-55, 2003.
- [51] S. Sascha, "Flying 10 Years After 9-11: The Good, the Bad, and the Ugly," frommers, 2011. [Online]. Available: <https://www.frommers.com/tips/airfare/flying-10-years-after-9-11-the-good-the-bad-and-the-ugly#ixzz1b8s8UkJH>.
- [52] G. Blalock, V. Kadiyali and D. H. Simon, "The impact of post-9/11 airport security measures on the demand for air travel," *The Journal of Law and Economics*, vol. 50, no. 4, pp. 731-755, 2007.
- [53] L. Al-Fandi, N. Mabe and M. T. Khasawneh, "The impact of threat type and image color on baggage screening performance: A preliminary investigation," in *Proceedings of the 10th Asia Pacific Industrial Engineering and Management Systems (APIEMS) Conference*, Kitakyushu, Japan, 2009.
- [54] E. A. Bretz, "Delayed arrival for US baggage screening?," *IEEE Spectrum*, vol. 39, no. 5, pp. 16-19, 2002.
- [55] J. Janowitz, R. Malone, J. Brasher, D. Fischer and B. Schultz, "Test and Evaluation Report for the Explosive Device Detection Baseline Study," Galaxy Scientific Corp Egg Harbor Township, 1995.
- [56] A. Wales, T. Halbherr and A. Schwaninger, "Using speed measures to predict performance in X-ray luggage screening tasks," in *43rd Annual International Carnahan Conference on Security Technology*, Zürich, Switzerland, 2009.
- [57] X. Liu and A. Gale, "Searching for possible threat items to safe air travel: human error and training," in *International Conference on Engineering Psychology and Cognitive Ergonomics*, Beijing, China, 2007.
- [58] S. Michel, M. Mendes and A. Schwaninger, "Can the difficulty level reached in computer-based training predict results in x-ray image interpretation tests?," in *44th Annual 2010 IEEE International Carnahan Conference on Security Technology*, San Jose, California, USA, 2010.
- [59] S. Singh and M. Singh, "Explosives detection systems (EDS) for aviation security," *Signal processing*, vol. 83, no. 1, pp. 31-55, 2003.
- [60] D. Hardmeier, F. Hofer and A. Schwaninger, "The X-ray object recognition test (X-ray ORT)-a reliable and valid instrument for measuring visual abilities needed in X-ray screening," in

Proceedings 39th Annual 2005 International Carnahan Conference on Security Technology, Albuquerque, NM, USA, 2005.

- [61] M. Stefan, K. Saskia, R. Markus and A. Schwaninger, "The effect of image enhancement functions on x-ray," in *Proceedings of the 4th International Aviation Security Technology Symposium*, Washington, D.C., USA, 2006.
- [62] S. Nercessian, K. Panetta and S. Agaian, "Automatic detection of potential threat objects in X-ray luggage scan images," in *IEEE Conference on Technologies for Homeland Security*, IEEE, 2008.
- [63] B. R. Abidi, J. Liang, M. Mitckes and M. A. Abidi, "Improving the detection of low-density weapons in x-ray luggage scans using image enhancement and novel scene-decluttering techniques," *Journal of Electronic Imaging*, vol. 13, no. 3, pp. 523-539, 2004.
- [64] B. Abidi, M. Mitckes, M. A. Abidi and J. Liang, "Grayscale enhancement techniques of x-ray images of carry-on luggage," *Sixth International Conference on Quality Control by Artificial Vision*, vol. 5132, pp. 579-592, 2003.
- [65] Y. Zheng, "X-Ray Image Processing and Visualization for Remote Assistance of Airport Luggage Screeners," 2004.
- [66] Z. Chen, Y. Zheng, B. R. Abidi, D. L. Page and M. A. Abidi, "A combinational approach to the fusion, de-noising and enhancement of dual-energy x-ray luggage images," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, San Diego, CA, USA, 2005.
- [67] A. Schwaninger, B. K. Kersten and M. Groner, "Objekterkennung und Signaldetektion: Anwendungen in der Praxis," in *Praxisfelder der Wahrnehmungspsychologie*, Verlag Hans Huber, 2005, pp. 106-132.
- [68] A. Schwaninger, "X-ray Imagery: enhancing the value of the pixels," *Aviation Security International*, pp. 16-21, 2005.
- [69] R. Gonzalez and R. Woods, *Digital Image Processing (2nd ed.)*, Prentice-Hall International Edition, 2002.
- [70] N. Megherbi, T. P. Breckon and G. T. Flitton, "Investigating existing medical CT segmentation techniques within automated baggage and package inspection," in *Optics and Photonics for*

Counterterrorism, Crime Fighting and Defence IX; and Optical Materials and Biomaterials in Security and Defence Systems Technology X, Dresden, Germany, 2013.

- [71] K. Fu, C. Guest and P. Das, "Segmentation of suspicious objects in an x-ray image using automated region filling approach," in *Signal and Data Processing of Small Targets*, San Diego, California, USA, 2009.
- [72] S. Ogorodnikov and V. Petrunin, "Processing of interlaced images in 4-10 MeV dual energy customs system for material recognition," *Physical Review Special Topics-Accelerators and Beams*, vol. 5, no. 10, pp. 1-11, 2002.
- [73] S. Ogorodnikov, V. Petrunin and M. Vorogushin, "Radioscopic discrimination of materials in 1/10 MeV range for customs applications," in *Proceeding of the 8th European Particle Accelerator Conference*, Paris, France, 2002.
- [74] L. a. S. V. a. K. T. Grady, C. Alvino and C. Bahlmann, "Automatic segmentation of unknown objects, with application to baggage security," in *European Conference on Computer Vision*, Berlin, Germany, 2012.
- [75] D. Mery, E. Svec, M. Arias, V. Rizzo, J. M. Saavedra and S. Banerjee, "Modern computer vision techniques for x-ray testing in baggage inspection," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 4, pp. 682-692, 2016.
- [76] B. R. Abidi, Y. Zheng, A. V. Gribok and M. A. Abidi, "Improving weapon detection in single energy X-ray images through pseudocoloring," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 36, no. 6, pp. 784--796, 2006.
- [77] M. P. DeDonato, V. Dimitrov and T. Padir, "Towards an automated checked baggage inspection system augmented with robots," in *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense XIII*, 2014.
- [78] B. R. Radu and C. Steve, "3D is here: Point Cloud Library (PCL)," in *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, 2011.
- [79] R. Hassanpour, "Illicit Material Detection using Dual-Energy X-Ray Images," *International Arab Journal of Information Technology (IAJIT)*, vol. 13, no. 14, 2016.
- [80] D. Mery, "Computer vision technology for X-ray testing," *Insight-non-destructive testing and condition monitoring*, vol. 56, no. 3, pp. 147-155, 2014.

- [81] N. Megherbi, G. T. Flitton and T. P. Breckon, "A classifier based approach for the detection of potential threats in CT based baggage screening," in *17th IEEE International Conference on Image Processing (ICIP)*, Hong Kong, Hong Kong, 2010.
- [82] J. Liang, B. R. Abidi and M. A. Abidi, "Automatic X-ray image segmentation for threat detection," in *Proceedings Fifth International Conference on Computational Intelligence and Multimedia Applications. ICCIMA 2003*, Xi'an, China, 2003.
- [83] D. Turcsany, A. Mouton and T. P. Breckon, "Improving feature-based object recognition for X-ray baggage security screening using primed visualwords," in *IEEE International Conference on Industrial Technology (ICIT)*, 2013, 2013.
- [84] L. Martin, A. Tuysuzoglu, W. C. Karl and P. Ishwar, "Learning-based object identification and segmentation using dual-energy ct images for security," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4069-4081, 2015.
- [85] P. Babaheidarian and D. Castañón, "Joint Segmentation and Material Recognition in Dual-Energy CT Images," *International Symposium on Electronic Imaging, Computational Imaging XV*, vol. 6, no. 17, pp. 30-35, 2017.
- [86] V. Muthukkumarasamy, M. Blumenstein, J. Jo and S. Green, "Intelligent illicit object detection system for enhanced aviation security," in *International Conference on Simulated Evolution and Learning (SEAL '04)*, Busan, Korea, 2004.
- [87] G. Heitz and G. Chechik, "Object separation in X-ray image sets," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, 2010.
- [88] D. F. Wiley, D. Ghosh and C. Woodhouse, "Automatic segmentation of CT scans of checked baggage," in *Proceedings of the 2nd International Meeting on Image Formation in X-ray CT*, Salt Lake City, Utah, USA, 2012.
- [89] D. Wiley, "Analysis of anatomic regions delineated from image data," Google Patents, 2012.
- [90] D. o. h. s. center, "Segmentation of objects from volumetric CT data," Northeastern University, Boston, 2011.
- [91] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Bombay, India, 1998.

- [92] S.-Y. Wan and W. E. Higgins, "Symmetric region growing," in *International Conference on Image Processing*, Vancouver, BC, Canada, 2000.
- [93] J. A. Bangham, P. Chardaire, C. J. Pye and P. D. Ling, "Multiscale nonlinear decomposition: the sieve decomposition theorem," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 5, pp. 529-539, 1996.
- [94] J. A. Bangham, J. R. Hidalgo, R. W. Harvey and G. C. Cawley, "The Segmentation of Images via Scale-Space Trees," in *British Machine Vision Conference*, Southampton, UK, 1998.
- [95] R. Sedgewick, *Algorithms in C*. Addison-Wesley Professional, 3rd edition, 1998.
- [96] Y. Lan, R. Harvey and R. Perez-Torres, "Finding stable salient contours," in *Proceedings of the British Machine Vision Conference*, Oxford, UK, 2010.
- [97] C. Solomon and T. Breckon, *Fundamentals of Digital Image Processing: A practical approach with examples in Matlab*, John Wiley & Sons, 2011.
- [98] U. Luis Garcia, S. Eli, V. Sreenath Rao, A. Vincent, S. Mark and B. Ranjit, "Automatic image segmentation by dynamic region growth and multiresolution merging," *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2275-2288, 2009.
- [99] A. Mouton and T. P. Breckon, "Materials-based 3D segmentation of unknown objects from dual-energy computed tomography imagery in baggage security screening," *Pattern Recognition*, vol. 48, no. 6, pp. 1961-1978, 2015.
- [100] A. Mouton, N. Megherbi, K. Van Slambrouck, J. Nuyts and T. P. Breckon, "A distance driven method for metal artefact reduction in computed tomography".
- [101] T. Johnson, *Medical radiology/diagnostic imaging: dual energy CT in clinical practice*, Springer, 2011.
- [102] L. Grady, "Fast, quality, segmentation of large volumes-isoperimetric distance trees," in *European Conference on Computer Vision*, Graz, Austria, 2006.
- [103] M. Baştan, M. R. Yousefi and T. M. Breuel, "Visual words on baggage X-ray images," in *International Conference on Computer Analysis of Images and Patterns*, Seville, Spain, 2011.

- [104] D. Turcsany, A. Mouton and T. P. Breckon, "Improving feature-based object recognition for X-ray baggage security screening using primed visualwords," in *IEEE International Conference on Industrial Technology (ICIT)*, Cape Town, South Africa, 2013.
- [105] M. E. a. A. S. a. D. M. a. M. A. a. B. T. P. Kundegorski, "On using feature descriptors as visual words for object detection within x-ray baggage security screening," in *International Conference on Imaging for Crime Detection and Prevention*, Madrid, Spain, 2016.
- [106] D. Mery, E. Svec and M. Arias, "Object recognition in baggage inspection using adaptive sparse representations of X-ray images," in *7th Pacific-Rim Symposium: Image and Video Technology*, Auckland, New Zealand, 2015.
- [107] D. Mery, V. Riffo, I. Zuccar and C. Pieringer, "Object recognition in X-ray testing using an efficient search algorithm in multiple views," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 59, no. 2, pp. 85-92, 2017.
- [108] M. Bastan, W. Byeon and T. M. Breuel, "Object recognition in multiview dual energy x-ray images," in *British Machine Vision Conference*, Bristol-UK, 2013.
- [109] S. Akçay, M. E. Kundegorski, M. Devereux and T. P. Breckon, "Transfer learning using convolutional neural networks for object classification within x-ray baggage security imagery," in *IEEE International Conference on Image Processing (ICIP)*, Phoenix, Arizona, USA, 2016.
- [110] S. Akçay, M. E. Kundegorski, C. G. Willcocks and T. P. Breckon, "Using deep convolutional neural network architectures for object classification and detection within x-ray baggage security imagery," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2203-2215, 2018.
- [111] D. Mery, "Automated detection in complex objects using a tracking algorithm in multiple x-ray views," in *Computer Vision and Pattern Recognition Workshops (CVPRW)*, Colorado Springs, CO, USA, 2011.
- [112] T. Franzel, U. Schmidt and S. Roth, "Object detection in multi-view X-ray images," in *Joint DAGM (German Association for Pattern Recognition) and OAGM Symposium*, Graz, Austria, 2012.
- [113] N. Dalal and B. Triggs, "Histograms of oriented gradients for human," in *International Conference on computer vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005.

- [114] M. Baştan, "Multi-view object detection in dual-energy x-ray images," *Machine Vision and Applications*, vol. 26, no. 7-8, pp. 1045--1060, 2015.
- [115] G. Flitton, T. P. Breckon and N. Megherbi, "A comparison of 3D interest point descriptors with application to airport baggage object detection in complex CT imagery," *Pattern Recognition*, vol. 46, no. 9, pp. 2420-2436, 2013.
- [116] G. Flitton, A. Mouton and T. P. Breckon, "Object classification in 3D baggage security computed tomography imagery using visual codebooks," *Pattern Recognition*, vol. 48, no. 8, pp. 2489-2499, 2015.
- [117] A. Mouton and T. P. Breckon, "A review of automated image understanding within 3D baggage computed tomography security screening," *Journal of X-ray science and technology*, vol. 23, no. 5, pp. 531-555, 2015.
- [118] T. W. Rogers, N. Jaccard, E. J. Morton and L. D. Griffin, "Automated x-ray image analysis for cargo security: Critical review and future promise," *Journal of X-ray science and technology*, vol. 25, no. 1, pp. 33-56, 2017.
- [119] J. M. Blackledge, *Digital Image Processing (1st Edition)*, Woodhead Publishing, 2005.
- [120] S. H. Malik and T. A. Lone, "Comparative study of digital image enhancement approaches," in *International Conference on Computer Communication and Informatics*, Coimbatore, Tamilnadu, 2014.
- [121] A. Koschan and M. Abidi, *Digital color image processing*, John Wiley & Sons, 2008.
- [122] R. Sève, *Science de la couleur: Aspects physiques et perceptifs*, Chalagam Edition, 2009.
- [123] S. Gupta and Y. Kaur, "Review of different local and global contrast enhancement techniques for a digital image," *International Journal of Computer Applications*, vol. 100, no. 18, pp. 18-23, 2014.
- [124] D. J. Bora and A. K. Gupta, "A new efficient color image segmentation approach based on combination of histogram equalization with watershed algorithm," *International Journal of Computer Science Engineering*, vol. 4, no. 6, pp. 156-167, 2016.
- [125] C. G. Rafael and E. W. Richard, *Digital Image Processing*, New Delhi: Dorling Kindersley, 2014.

- [126] V. Rajamani and P. a. J. S. Babu, "A Review of various global contrast enhancement techniques for still images using histogram Modification Framework," *International Journal of Engineering Trends and Technology*, vol. 4, no. 4, pp. 1045-1048, 2013.
- [127] I. J. o. C. Applications, "Review of different local and global contrast enhancement techniques for a digital image," *International Journal of Computer Applications*, vol. 100, no. 18, pp. 18-23, 2014.
- [128] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355-368, 1987.
- [129] S. Singh and M. Singh, "Explosives detection systems (EDS) for aviation security," *Signal processing*, vol. 83, no. 1, pp. 31-55, 2003.
- [130] P. Patidar, M. Gupta, S. Srivastava and A. K. Nagawat, "Image de-noising by various filters for different noise," *International journal of computer applications*, vol. 9, no. 4, pp. 45-50, 2010.
- [131] G. R. Arce, *Nonlinear signal processing: a statistical approach*, John Wiley & Sons, 2005.
- [132] C. E. Heil and D. F. Walnut, "Continuous and discrete wavelet transforms," *SIAM review*, vol. 31, no. 4, pp. 628--666, 1989.
- [133] D. H. Johnson, "Signal-to-noise ratio," *Scholarpedia*, vol. 1, no. 12, 2006.
- [134] O. Maimon and L. Rokach, *Data mining and knowledge discovery handbook*, Springer, 2005.
- [135] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Oakland, CA, USA, 1967.
- [136] B. S. Everitt, *Finite mixture distributions*, Wiley Online Library, 1981, pp. 25-57.
- [137] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1-22, 1977.
- [138] T. Kohonen, "The self-organizing map," *Invited paper in proceedings of the IEEE*, vol. 78, no. 9, pp. 1464-1480, 1990.

- [139] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE transactions on pattern analysis and machine intelligence*, no. 2, pp. 224-227, 1979.
- [140] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1-27, 1974.
- [141] J. C. Dunn, "Well-separated clusters and optimal fuzzy partitions," *Journal of cybernetics*, vol. 4, no. 1, pp. 95-104, 1974.
- [142] J. A. Hartigan, *Clustering algorithms*, Wiley, 1975.
- [143] C. Ronan and B. Samy, "Links between perceptrons, MLPs and SVMs," in *The twenty-first International Conference on Machine Learning (ICML)*, Banff, Alberta, Canada, 2004.
- [144] H. Simon, *Neural Networks: A Comprehensive Foundation (2nd Edition)*, Prentice Hall, 1998.
- [145] A. A. James, *An Introduction to Neural Networks*, A Bradford Book, 1995.
- [146] S. A. Naomi, "An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression," *The American Statistician*, vol. 46, no. 3, pp. 175-185, 1990.
- [147] K. Mehmed, *Data Mining: Concepts, Models, Methods, and Algorithms*, Wiley-IEEE Press, 2011, pp. 146-149.
- [148] H. Jiawei and K. Micheline, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2006.
- [149] B. Leo, *Machine Learning*, Springer International Publishing, 2001, pp. 5-32.
- [150] S. Suthaharan, *Machine Learning Models and Algorithms for Big Data Classification*, Springer US, 2016.
- [151] W. Zhe and X. Xiangyang, *Support Vector Machines Applications, Chapter 2: Multi-Class Support Vector Machine*, Springer International Publishing, 2014.
- [152] F. Ronald Aylmer, "The use of multiple measurements in taxonomic problems," *Annals of eugenics*, vol. 7, no. 2, pp. 179-188, 1936.
- [153] E. Romera, J. M. Alvarez, L. M. Bergasa and R. Arroyo, "Erfnet: Efficient residual factorized convnet for real-time semantic segmentation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 1, pp. 263-272, 2017.

- [154] H. Zhao, X. Qi, X. Shen, J. Shi and J. Jia, "Icnet for real-time semantic segmentation on high-resolution images," in *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, 2018.
- [155] S. Mehta, M. Rastegari, A. Caspi, L. Shapiro and H. Hajishirzi, "Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, 2018.
- [156] M. Siam, M. Gamal, M. Abdel-Razek, S. Yogamani, M. Jagersand and H. Zhang, "A comparative study of real-time semantic segmentation for autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Salt Lake City, UT, USA, 2018.
- [157] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, Boston, MA, USA, 2015.
- [158] V. Badrinarayanan and A. .: C. R. Kendall, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481-2495, 2017.
- [159] O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *18th International Conference on Medical image computing and computer-assisted intervention*, Munich, Germany, 2015.
- [160] H. Noh, S. Hong and B. Han, "Learning deconvolution network for semantic segmentation," in *Proceedings of the IEEE international conference on computer vision (ICCV)*, Santiago, Chile, 2015.
- [161] J. Jiang, L. Zheng, F. Luo and Z. Zhang, "Rednet: Residual encoder-decoder network for indoor rgb-d semantic segmentation," *arXiv preprint arXiv:1806.01054*, 2018.
- [162] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow and B. Frey, "Adversarial autoencoders," *arXiv preprint arXiv:1511.05644*, 2015.
- [163] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [164] J. L. Ba, J. R. Kiros and G. E. Hinton, "Layer normalization," *arXiv preprint arXiv:1607.06450*, 2016.

- [165] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv preprint arXiv:1803.08375*, 2018.
- [166] S. Mehmet and S. Bülent, "Survey over image thresholding techniques and quantitative performance evaluation," *Electronic Imaging*, vol. 13, no. 1, pp. 159-179, 2004.
- [167] O. Nobuyuki, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [168] d. S. Luigi and B. Andrea, "A simple and efficient connected components labeling algorithm," in *International Conference on Image Analysis and Processing*, Venice, Italy, 1999.
- [169] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, Nevada, USA, 2016.
- [170] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Hawaii, USA, 2017.
- [171] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
- [172] Y. Ahuja and S. K. Yadav, "Multiclass classification and support vector machine," *Global Journal of Computer Science and Technology Interdisciplinary*, vol. 12, no. 11, pp. 14-20, 2012.
- [173] G. Roffo, S. Melzi, U. Castellani and A. Vinciarelli, "Infinite latent feature selection: A probabilistic latent graph-based ranking approach," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017.
- [174] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Advances in neural information processing systems*, Vancouver, British Columbia, Canada, 2002.
- [175] M. A. Hall, "Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning," in *Proceedings of the 17th international conference on machine learning (ICML-2000)*, Stanford, CA, USA, 2000.
- [176] K. Kira and L. A. Rendell, "A practical approach to feature selection," in *Machine Learning Proceedings*, 1992.

- [177] S. Yutaka, "The truth of the F-measure," *Teaching, Tutorial materials*, vol. 1, no. 5, pp. 1-5, 2007.
- [178] Z. Harry, "The Optimality of Naïve Bayes," in *In FLAIRS2004 conference*, Florida, USA, 2004.
- [179] C.-h. Chen, *Handbook of pattern recognition and computer vision*, World Scientific, 2015.
- [180] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [181] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving and M. a. o. Isard, "Tensorflow: A system for large-scale machine learning," in *12th Symposium on Operating Systems Design and Implementation*, 2016.
- [182] R. Giorgio, M. Simone, C. Umberto and V. Alessandro, "Infinite Latent Feature Selection: A Probabilistic Latent Graph-Based Ranking Approach," in *IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017.
- [183] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015.
- [184] 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*, 2016.
- [185] J. Dai, Y. Li, K. He and J. Sun, "R-fcn: Object detection via region-based fully convolutional networks," in *Advances in neural information processing systems*, 2016.
- [186] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla and M. a. o. Bernstein, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211-252, 2015.
- [187] S. Singh and M. Singh, "Explosives detection systems (EDS) for aviation security," *Signal processing*, vol. 83, no. 1, pp. 31-55, 2003.
- [188] A. a. G. J. Abadie, "Terrorism and the world economy," *European Economic Review*, vol. 52, no. 1, pp. 1--27, 2008.
- [189] P. Baum, *Violence in the skies: a history of aircraft hijacking and bombing*, Summersdale Publishers LTD, 2016.

- [190] T. Fawcett, "An introduction to ROC analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861-874, 2006.
- [191] F. H. Attix and W. C. Roesch, *Radiation Dosimetry*, Academic Press Inc, 1968.
- [192] J. G. Brown, *X-Rays and Their Applications*, London Iliffe Books, 1966.
- [193] N. A. Dyson, *X-rays in Atomic and Nuclear Physics*, Second Edition, Cambridge University Press, 1990.
- [194] R. D. Knight, *Physics for scientists and engineers*, Pearson Higher Ed., 2017.
- [195] E. M. Hussein, "Detection of explosive materials using nuclear radiation: a critical review," *X-Ray Detector Physics and Applications*, vol. 1736, pp. 130--138, 1993.
- [196] D. J. Mitchell, K. Thompson, L. T. Harding, G. G. Thoreson, L. Theisen and J. E. Parmeter, "Characterization of X-Ray Generator Beam Profiles," Sandia National Lab.(SNL-NM), Albuquerque, NM, USA, 2013.
- [197] T. Takahashi, H. Itoh, T. Shimada and H. Takeuchi, "Design of integrated radiation detectors with a-Si photodiodes on ceramic scintillators for use in x-ray computed tomography," *IEEE Transactions on Nuclear Science*, vol. 37, no. 3, pp. 1478--1482, 1990.
- [198] L. a. S. V. Grady, T. Kohlberger, C. Alvino and C. Bahlmann, "Automatic segmentation of unknown objects, with application to baggage security," in *12th European Conference on Computer Vision (ECCV)*, Florence, Italy, 2012.
- [199] L. Grady and C. V. Alvino, "The piecewise smooth Mumford--Shah functional on an arbitrary graph," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2547-2561, 2009.
- [200] H. K. Tuy, "A post-processing algorithm to reduce metallic clip artifacts in CT images," *European Radiology*, vol. 3, no. 2, pp. 129-134, 1993.